5-1-2013

A Holarctic perspective on mammalian evolution: The evolutionary and biogeographic history of red-backed voles and their close relatives (Rodentia: Arvicolinae)

Brooks Kohli

Brooks A. Kohli

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

Steven Poe

Christopher C. Witt

A Holarctic perspective on mammalian evolution: The evolutionary and biogeographic history of red-backed voles and their close relatives (Rodentia: Arvicolinae)

by

BROOKS A. KOHLI

B.S., Biological Sciences, Ohio University, 2010

THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

Biology

The University of New Mexico

Albuquerque, New Mexico

May 2013

ACKNOWLEDGEMENTS

A thesis focusing on a group of species found across the Northern Hemisphere and on three continents could not be successful without the help of many people at all stages of the research. First of all, I have to thank my advisor, Joe Cook, for his guidance and encouragement throughout the whole process and for giving me the invaluable experiences I enjoyed in New Mexico, Alaska, and beyond. I would also like to thank Chris Witt and Steve Poe for their commitment and insight that greatly improved the quality of this work. Many thanks are also due to the members and associates of the Cook lab, past and present, for their advice, encouragement, help with lab work and learning analysis programs, and especially their friendship. Special thanks to Kelly Speer for her enthusiasm and dedication in the lab that made it possible to collect such a large amount of data in a short period of time. I am indebted to all the collectors of voles, over decades of difficult field research in some of the most remote regions of the world that afforded me this wealth of samples, and the natural history museums that have preserved them and made them available to me. I am grateful to the generous funding sources that supported this research, especially the UNM Department of Biology, Graduate and Professional Student Association, and Biology Graduate Student Association. Finally, my deepest gratitude to Adrienne for her endless help, support, and understanding and for always pushing me to be my best in everything I do.

A Holarctic perspective on mammalian evolution: The evolutionary and biogeographic history of red-backed voles and their close relatives (Rodentia: Arvicolinae)

by

Brooks A. Kohli

B.S., Biology, Ohio University, 2010

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

ABSTRACT

The influence of Quaternary environmental changes on demography and geographic distribution is paramount to understanding how contemporary genetic diversity is partitioned in high-latitude species. This history of change in northern ecosystems sets the stage for forecasting how species inhabiting tundra and boreal forest will respond to the ongoing shift in climate. Utilizing broad geographic and taxonomic sampling, a multilocus species tree approach, ecological niche models, and population genetic techniques, I investigate the evolutionary and biogeographic history of a Holarctic mammal, the northern red-backed vole (*Clethrionomys rutilus)* and the systematics of the tribe it belongs to, Clethrionomyini. This tribe of forest and alpine voles is distributed throughout the Holarctic and diversified in the Northern Hemisphere over the last 5 MY. The biogeographic history of *C. rutilus* is characterized by subdivision and subsequent expansion from multiple refugia, including previously

unidentified refugia in central Asia. Several taxonomic issues are resolved by the use of multilocus data, including support of a polyphyletic *Clethrionomys*, prompting the suggestion of splitting the genus into two. Diversification in the tribe is not characterized by pulses of diversification, as previously suggested. The dynamic geographic, genealogic, and demographic history of *C. rutilus* and Clethrionomyini provides insight into the Quaternary biogeography of the northern high-latitudes in Asia and North America.

TABLE OF CONTENTS

Chapter 1

Thesis Introduction

In high-latitude ecosystems, cyclical climatic change over the past 2-3 million years played a pivotal role in creating biodiversity and shaping species distributions and genetic diversity (Webb III and Bartlein 1992; Hofreiter and Stewart 2009). During the Quaternary, cyclical build up of continental ice sheets led to significant sea level changes and shifting landscapes that structured plant and animal species. The resulting demographic and distributional changes within and among populations over time can be traced through distinct patterns of genetic variation in contemporary populations (Avise 2009). It is possible to identify barriers to dispersal, the location of presumed glacial refugia, regions recently colonized through population expansion, which enables the interpretation of genetic diversity patterns in the context of spatial and demographic processes (Avise 2000; Smith 2007). Using this information, the role that these types of events may play in incipient speciation among high-latitude species can be better understood and may help forecast how Arctic and boreal species would be affected in the future by directional climate change and expected range shifts (Cook, Hoberg et al. 2005).

Different genetic patterns result from individual species' responses to geologic and climatic change, largely influenced by unique biological constraints or innovations, dispersal abilities, and historic distributions (DeChaine 2008). Although some common phylogeographic patterns have emerged, details of the response of individual northern species to past climatic changes are often idiosyncratic (Hewitt 2004; Waltari, Hoberg et al. 2007; DeChaine 2008). For example, many species show concordance in their

persistence within major glacial refugia followed by range expansion after the recession of the ice sheets. Furthermore, species with similar ecological affinities, such as those associated with boreal forest, are not the same in every detail of their phylogenetic and distributional histories, but generally share broad phylogeographic patterns. These similarities suggest that genetic signals can be used as an independent line of evidence from which to infer the complexity of the history of the boreal forest biome (Fedorov, Goropashnaya et al. 2008). This connection between large-scale ecological processes and evolutionary change demonstrates that phylogeographic and phylogenetic analyses provide a unique and vital perspective in an era of shifting climatic conditions.

In Chapter 2, I investigate the range-wide phylogeography of *Clethrionomys rutilus*, northern red-backed voles. This is the only Holarctic species in the relatively large group of voles known as the tribe Clethrionomyini and one of only fourteen Holarctic mammals (MacDonald and Cook 2007). I sampled populations from Finland to Canada, throughout the enormous range of this species. *C. rutilus* is strongly associated with boreal forest habitat, one of the most extensive biomes in the world, but is also found in tundra habitats. These biomes were greatly impacted by climatic cycling in the Quaternary and are projected to be radically changed in response to contemporary climate change (Parmesan 2006). Due to its widespread distribution, *C. rutilus* provides nearly unparalleled insight into factors that have influenced patterns of genetic structure across spatial scales ranging from global to regional to local. My analysis detects considerable genetic structure across the species' range, revealing multiple glacial refugia, dispersal barriers, and variable demographic histories. Taken together, the results emphasize the idiosyncratic and dynamic response of northern species to environmental change.

In Chapter 3, I assess phylogenetic relationships and species limits in Clethrionomyini, a taxonomically problematic tribe, mostly represented by voles found in northern and central Asia. The historical biogeography in this dynamic, biodiversity-rich region is understudied compared to Europe and North America (Hewitt 1996). Due to much of Asia's inaccessibility, many species in this tribe have rarely or never been sequenced, limiting the ability to conclusively resolve the phylogeny. I generate DNA sequences for three previously unstudied species and add molecular data for sixteen others belonging to all five nominal genera: *Hyperacrius*, *Caryomys*, *Eothenomys*, *Alticola*, and *Clethrionomys* (Carleton and Musser 2005). Previous molecular analyses revealing cases of apparent paraphyly or polyphyly of *Clethrionomys* (Cook, Runck et al. 2004; Luo, Yang et al. 2004; Lebedev, Bannikova et al. 2007) are tested with nuclear, multilocus data for the first time. Clarifying relationships among members of Clethrionomyini at various taxonomic levels addresses questions about taxonomy and speciation within the group and more generally sheds light on the biogeographic history of northern terrestrial regions.

By studying this tribe of voles and the species with the largest distribution among them in detail, I hope to elucidate the evolutionary processes that have impacted northern species over the last several million years. These molecular data will not only inform the systematics of this group, but will provide broad comparative information for evolutionary and biogeographic research in other northern high-latitude organisms.

References

Avise, J. C. (2000). Phylogeography: The history and formation of species. Cambridge, Massachusetts, Harvard University Press.

Avise, J. C. (2009). "Phylogeography: retrospect and prospect." Journal of Biogeography **36**(1): 3-15.

Carleton, M. D. and G. G. Musser (2005). Order Rodentia. Mammal Species of the World. D. E. Wilson and D. M. Reeder. Baltimore, Maryland, The Johns Hopkins University Press.

Cook, J. A., E. P. Hoberg, et al. (2005). "Beringia: Intercontinental exchange and diversification of high latitude mammals and their parasites during the Pliocene and Quaternary." Mammal Study **30**(Suppl. S): S33-S44.

Cook, J. A., A. M. Runck, et al. (2004). "Historical biogeography at the crossroads of the northern continents: molecular phylogenetics of red-backed voles (Rodentia : Arvicolinae)." Molecular Phylogenetics and Evolution **30**(3): 767-777.

DeChaine, E. G. (2008). "A bridge or a barrier? Beringia's influence on the distribution and diversity of tundra plants." Plant Ecology & Diversity **1**(2).

Fedorov, V. B., A. V. Goropashnaya, et al. (2008). "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.

Hewitt, G. M. (1996). "Some genetic consequences of ice ages, and their role in divergence and speciation." Biological Journal of the Linnean Society **58**(3): 247-276.

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**(1442): 183-195.

Hofreiter, M. and J. Stewart (2009). "Ecological change, range fluctuations and population dynamics during the Pleistocene." Current Biology **19**(14): R584-R594.

Lebedev, V. S., A. A. Bannikova, et al. (2007). "Molecular phylogeny of the genus Alticola (Cricetidae, Rodentia) as inferred from the sequence of the cytochrome b gene." Zoologica Scripta **36**(6): 547-563.

Luo, J., D. Yang, et al. (2004). "Molecular phylogeny and biogeography of Oriental voles: genus *Eothenomys* (Muridae, Mammalia)." Molecular phylogenetics and evolution **33**(2): 349-362.

MacDonald, S. O. and J. A. Cook (2007). Checklist of the Land Mammals of Beringia. Albuquerque, NM, University of New Mexico.

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

Smith, C. I. (2007). "Historical biogeography: The new synthesis." Current Biology **17**(15): R598-R600.

Waltari, E., E. P. Hoberg, et al. (2007). "Eastward Ho: phylogeographical perspectives on colonization of hosts and parasites across the Beringian nexus." Journal of Biogeography **34**(4): 561-574.

Webb III, T. and P. Bartlein (1992). "Global changes during the last 3 million years: climatic controls and biotic responses." Annual Review of Ecology and Systematics **23**: 141-173.

Chapter 2

Phylogeography of a Holarctic rodent (*Clethrionomys rutilus*): Testing high-latitude biogeographic hypotheses and the dynamics of range shifts

Abstract

Pleistocene glacial cycles greatly influenced demography and geographic distribution of high-latitude species and played an essential role in determining how contemporary genetic diversity is partitioned. This history of change in northern ecosystems sets the stage for forecasting how species inhabiting tundra and boreal forest will respond to the ongoing alterations in climate. We use a multilocus phylogeographic approach (1 mtDNA, 3 nuclear genes) combined with species distribution models to investigate the response of a Holarctic mammal, the northern red-backed vole (*Clethrionomys rutilus)*, to environmental change. Our assessment is based on widespread sampling, a species tree approach for integration of multiple lines of genetic evidence, and population genetic techniques to elucidate the biogeographic and demographic history of northern red-backed voles. The phylogeographic history of *C. rutilus* is characterized by isolation and subsequent expansion from multiple refugia in Asia, an uncommon history among other predominantly boreal forest-associated species. Demographic and distributional changes are dynamic and emphasize the effect pre-LGM glacial-interglacial cycles had on contemporary phylogeographic structure.

Brooks A. Kohli[a], Vadim B. Fedorov[b]*, Eric Waltari[c]* and Joseph A. Cook[a]*

[a] Department of Biology and Museum of Southwestern Biology, University of New Mexico, Albuquerque NM 87131–1051, [b] Institute of Arctic Biology, University of

Alaska Fairbanks, Fairbanks, AK 99775-7000, USA, [c] City College of New York, Department of Biology, Marshak Science Building 814, 160 Convent Ave, New York, NY 10031, USA

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

1. Introduction

Evolutionary biologists have long attempted to understand the processes that create, maintain, and alter variation within species (Darwin 1859; Mayr 1970). Variation is affected by contemporary factors, but is also shaped by species' evolutionary and biogeographic histories, thus impacting ecological interactions and evolutionary trajectories (Hewitt 1996; Abbott and Brochmann 2003; Lister 2004; Thompson 2005). An essential component of biogeographic history is knowledge of the distributional changes of species, including range shifts, long-term subdivision of populations, colonization of new habitats, and changes in range size. These types of changes are typically accompanied by demographic and genealogic changes as well. The degree to which distributions are altered is largely mediated by external factors that affect connectivity, dispersal, and population dynamics, especially large-scale environmental changes (Hewitt 1996; Lister 2004).

Biomes of the northern high-latitudes were extensively impacted by climatic cycling throughout the Quaternary, about the last 2.6 million years (MY), and are projected to be strongly affected by contemporary climate change (Parmesan 2006). The Quaternary was marked by the most frequent and dramatic climatic oscillations of the entire Cenozoic (the last 65 MY) (Velichko and Wright 2005). Fluctuations between

warm periods, at times warmer than the present, and cold periods dominated by continental ice sheets had the greatest effect on environments at the highest latitudes (Miller, Brigham-Grette et al. 2010).

Because individual species differ in biological constraints, innovations, ecological affinities, and historic distributions, the effects of environmental changes influence species in different ways (Hoffmann 1981; Stewart 2009). Extensive paleoecological, paleontological and phylogeographic data suggests organisms largely responded idiosyncratically to past climatic changes, but common patterns are also emerging as phylogeographic evidence accumulates (Abbott and Brochmann 2003; Hewitt 2004; Waltari, Hoberg et al. 2007; Fedorov, Goropashnaya et al. 2008; Korsten, Ho et al. 2009). For example, many northern species show concordance in their persistence within a handful of shared glacial refugia followed by range expansion after the recession of Pleistocene ice sheets. As independently diverging lineages expanded into recently unglaciated areas, common areas of secondary contact developed (suture zones). Multiple glacial refugia, colonization routes, and contact zones, either between species or divergent populations within a species, have frequently been identified by the discovery of concordant phylogeographic patterns across taxa, primarily in Europe and North America (Hewitt 2000; Waltari, Hoberg et al. 2007; Shafer, Cullingham et al. 2010).

At the nexus of Asia and North America, Beringia is one of the best described northern refugia. Beringia spans the region currently bisected by the Bering Strait, but during Pleistocene glacial maxima and as recently as 12 KYA, the area was contiguous (i.e. the Bering Land Bridge) due to lowered sea levels. Beringia was bounded by glaciations in the Verkhoyansk Range in northeastern Siberia and the Mackenzie

Mountains in northwestern Canada during glacial periods (Hultén 1937) (Figure 1). Populations of many boreal and arctic organisms inhabited Beringia throughout the Quaternary, but their distributions within the refugium, expansion events, demographic characteristics, and dispersal barriers were variable (Fedorov 1999; Hewitt 2004; Weksler, Lanier et al. 2010; Galbreath, Cook et al. 2011). However, most amphi-Beringian species colonized the North American continent eastward from Asia, probably due to the location of massive ice sheets in North America (Waltari, Hoberg et al. 2007). This generalized pattern supports the hypothesis that Beringia was a critical source of diversity for mammals and further downstream colonization of North America (Hoberg, Kutz et al. 2003), but downplays Beringia's importance for mammalian colonization of Asia.

In contrast to the relative wealth of knowledge about mammalian biogeographic history of Europe and Beringia, the intervening region of north-central Asia, from the Ural Mountains in western Siberia to Beringia and the Pacific coast of Asia, is comparatively under-studied. Considering the size of this region and it's largely ice-free status during Quaternary glacial cycles, north-central Asia may have played a critical role in structuring high latitude biodiversity (Tarasov, Volkova et al. 2000; Ehlers and Gibbard 2007; Binney, Willis et al. 2009), perhaps in ways that differ from processes of diversification in Europe and North America. From limited phylogeographic studies of the region's mammals, two alternative biogeographic models have emerged (Hewitt 2000; Fedorov, Goropashnaya et al. 2008; Korsten, Ho et al. 2009). The first is a signal of widespread expansion throughout Asia from a single source refugium, with limited genetic differentiation across large geographic areas (Korsten, Ho et al. 2009). This

pattern is common among boreal forest species (Fedorov, Goropashnaya et al. 2008) and forms the basis for the biogeographic hypothesis that north-central Asia was largely uninhabitable by forest-associated species during extended glacial periods when cooler and drier habitats dominated the region.

In contrast, a second model is supported by other northern Eurasian mammals that exhibit deep divergence and genetic structure among geographic regions often reinforced by contemporary contact zones, a pattern indicative of historic subdivision over multiple glacial-interglacial periods (Hewitt 2000; Brunhoff, Galbreath et al. 2003; Fedorov, Goropashnaya et al. 2003; Galbreath and Cook 2004; Marmi, Lopez-Giraldez et al. 2006). This second hypothesis postulates multiple refugia, such as those identified in Europe and Beringia, and cryptic (previously poorly described) refugia in Asia, with each refugium independently contributing to the recolonization of northern Eurasia. This model has primarily, but not exclusively, been described for tundra and grassland adapted species (Brunhoff, Galbreath et al. 2003; Hope, Waltari et al. 2010; Hope, Waltari et al. 2011). A few putative Asian refugia have been inferred based on fossils, paleoecological reconstructions and comparative molecular studies (Urals, Northern Mongolia: (Markova, Smirnov et al. 1995; Tarasov, Volkova et al. 2000; Todisco, Gratton et al. 2012); SE Siberia, Amga River basin (Fedorov, Goropashnaya et al. 2008; Hope, Waltari et al. 2011), and fossils of plant material indicate the presence of many more cryptic refugia in the late Quaternary (Binney, Willis et al. 2009), but extensive phylogeographic studies of Holarctic species are needed to rigorously test hypothesized refugia.

1.1 Study species and hypotheses

The relatively few phylogeographic studies of Holarctic species have uncovered glacial refugia, dispersal barriers, and variable demographic or phylogenetic histories (Brunhoff, Galbreath et al. 2003; Fedorov, Goropashnaya et al. 2003; Hope, Waltari et al. 2010). Due to their extremely broad longitudinal distributions across multiple continents, Holarctic species provide unparalleled insight into biogeographic processes that have influenced genetic diversity and structure ranging from intercontinental to local scales.

We present a range-wide phylogeographic analysis of the northern red-backed vole (*Clethrionomys rutilus* Pallas 1779, recently *Myodes*; (Tesakov, Lebedev et al. 2010)), a Holarctic rodent whose current distribution spans the northern high-latitudes, from Scandinavia, through northern Asia to the Bering Strait, and across Alaska and northwestern Canada (Wilson and Reeder 2005) (Figure 1). Northern red-backed voles are also found on several islands with histories of Quaternary land connections, including the northern Japanese island of Hokkaido; Sakhalin Island, off the southeastern coast of Russia; and St. Lawrence Island, in the Bering Strait. Northern red-backed voles are associated with boreal forests and shrub tundra habitats throughout their range. Often the most abundant small mammal, these voles are important components of trophic webs (Guthrie 1968), and occur in most arctic or boreal habitats with sufficient cover and food (Gromov and Polyakov 1977).

C. rutilus has a range that encompasses nearly all major phylogeographic breaks recognized in northern organisms outside of Europe and is one of only fourteen mammals with a Holarctic distribution (MacDonald and Cook 2007). Fossil evidence and previous genetic studies support an Asian origin of *C. rutilus* during the Pleistocene (Gromov and Polyakov 1977; Cook, Runck et al. 2004). No fossils are known from

North America until the Holocene. Populations on either side of the Bering Strait exhibit close genetic relationships, although Asian populations are more structured (Frisman, Kartavtseva et al. 2002; Iwasa, Kartavtseva et al. 2002; Cook, Runck et al. 2004). However, these previous studies focused on relatively small geographic regions, with a limited number of genetic loci and samples. The fossil evidence, along with its distribution in North America, including an extensive parapatric contact zone with the more southerly *C. gapperi*, suggests eastward, late Pleistocene colonization through Beringia (Rausch 1963).

The primary objective of this study is to explore how the environmental history of the northern high-latitudes affected genetic diversity, demographics, and distribution of boreal organisms. To effectively test Holarctic biogeographic hypotheses, we perform a multi-gene analysis of *Clethrionomys rutilus* that extends previous work spatially and taxonomically. Using multiple lines of evidence, including phylogeographic analyses, species distribution models, and demographic reconstructions, we test alternative hypotheses to identify factors influencing temporal and spatial structure of this Holarctic mammal. Specifically, the Holarctic biogeographic history *C. rutilus* best fits which of two general models: (1) recent recolonization of a huge range from a single refugium or (2) isolation and divergence in multiple refugia followed by expansion? Secondly, we aim to describe the temporal and spatial characteristics of the biogeographic history of this widespread species to address whether a critically understudied region of the Holarctic, north-central Asia, harbored refugial populations and to identify particular episodic events that contributed most to the contemporary phylogeographic structure of *C. rutilus*.

2. Methods

2.1 Sampling and laboratory techniques

We sampled between one and five individuals from 110 localities over the Holarctic range of *C. rutilus* (Figure 1) for a total of 220 individuals, including samples representing 11 of 14 subspecies (Gromov and Polyakov 1977; Hall 1981) (Appendix A). Tissue or skin samples were loaned from the Museum of Southwestern Biology, University of Alaska Museum of the North, Siberian Zoological Museum, University of Washington Burke Museum, and Royal Ontario Museum, primarily from fieldwork over the past decade as part of the Beringian Coevolution Project. To increase geographic breadth, 92 additional sequences were retrieved from GenBank for mitochondrial DNA (mtDNA) analyses (Appendix B). All other species of *Clethrionomys* and putatively paraphyletic species of *Alticola* (Lebedev, Bannikova et al. 2007) were included in phylogenetic analyses to assess the magnitude of intraspecific variation and prevent confounding results due to hybridization with parapatric species of *Clethrionomys.* Species of other arvicolids, *Microtus* and *Dicrostonyx*, were included to root phylogenies.

Genomic DNA was extracted from frozen (-80°C) or ethanol-preserved tissue or dried skin samples using commercial kits (Qiagen Inc., Valencia, California). Ethanol-preserved tissues and skin samples were washed in STE or PBS buffer prior to extraction. We amplified four independent loci via the polymerase chain reaction (PCR), the mtDNA cytochrome-b gene (Cyt b: 799–1,143 base pairs (bp)), and 3 nuclear gene loci trimmed to balance gene length and number of individuals: 1059 bp of the first exon of the interphotoreceptor retinoid binding protein gene (IRBP) (Stanhope, Smith et al. 1996), 205 bp of mineralocorticoid receptor exon 3 (MLR) (Lyons, Laughlin et al. 1997), and

926 bp of V-ets erythroblastosis virus E26 oncogene homolog 2 (ETS2) (Lyons, Laughlin et al. 1997) genes. Primers used for double-stranded amplifications and sequencing of Cyt b were MSB05 and MSB14 (Hope, Waltari et al. 2010) or VOLE14 (Hadly, Ramakrishnan et al. 2004). For highly fragmented DNA, the internal primers MVZ04 and MVZ16 (Smith and Patton 1993) and L649 (AAAATTCCCTTCCACCCCTA; A. Hope, unpublished) were used in conjunction with primers designed for this work to amplify ~400bp fragments: L373Cleth (GCA ACA GCA TTT ATA GGC TAT GTCC), L829Cleth (GCT TAC GCC ATC CTA CG), and H885 (GCT AGG ATT AGG ATT GAT AGG). Primers used for degraded DNA were paired as follows: MSB05/MVZ04, L373Cleth/MVZ16, L649/H885Cleth, and L829Cleth/Vole14. Primer pairs IRBPA/B (Stanhope, Smith et al. 1996), MLRF/R (Lyons, Laughlin et al. 1997), and ETS2F/R (Lyons, Laughlin et al. 1997) were used to amplify and sequence IRBP, MLR and ETS2 genes, respectively. Multiple genes provide independent perspectives of the history of the species with conflicting signals potentially due to hybridization or stochastic lineage sorting among red-backed vole species. For all targeted regions, polymerase chain reaction (PCR) reagents and conditions were: 1 µl of DNA template (variable concentration); 1.5 µl each of deoxynucleoside triphosphates (10 mM), MgCl (25 mM), 10X PCR buffer, and bovine serum albumin (1%); 0.5 µl of each primer (2 mM); 0.08 µl of AmpliTaq DNA polymerase (Applied Biosystems, Foster City, California); and 6.92 µl of double-distilled H_2O to total 15-µl reactions. Polymerase chain reaction was performed in a PTC 200 thermocycler (MJ Research, Waltham, MA) with initial denaturation at 94°C for 5 min, followed by 35 cycles of denaturation at 94°C for 15 s, annealing at 50°C for 20 s, extension at 72°C for 1 min, and final extension at 72°C for 5

min, with cooling at 15°C for 10 minutes. For nuclear genes, the conditions were modified to optimize annealing temperature and maximize yield (MLR: 54°C, IRBP: 56°C, and ETS2: 63°C). Amplified PCR product was visualized using agarose gel electrophoresis and PCR cleanup was performed with ExoSAP-IT (Affymetrix Inc., Santa Clara, California). Cycle sequencing was performed using ABI BigDye version 3.1 Sequencing Kit (Applied Biosystems, Foster City, California) and the following reagents for each 10-µl cycle sequencing reaction: 2 µl of H2O, 2 µl of BigDye buffer (Applied Biosystems), 0.5 µl of primer (2 mM), 0.5 µl of ABI BigDye version 3.1 (Applied Biosystems), and 5 µl of DNA (variable concentration). Conditions for cycle sequencing followed Platt et al. (2007) and reactions were cleaned using Ethanol/EDTA/sodium acetate precipitation protocols as recommended by the supplier (Applied Biosystems). Cycle sequencing products were sequenced and scored using an ABI 3100 automated genetic analyzer (Applied Biosystems) at the University of New Mexico Molecular Biology Facility or at the University of Washington DNA Sequencing Facility.

Sequences were edited and aligned with SEQUENCHER v4.9 (GeneCodes Corporation) and checked by eye. Complementary strands were aligned to create composite sequences. To ensure genuine mtDNA was obtained, Cyt b sequences were translated to amino acids and inspected for internal stop codons and expected patterns of replacement (Irwin, Kocher et al. 1991). For nuclear genes, allelic sequences were inferred using the program PHASE v2.1 (Stephens, Smith et al. 2001), using four iterations. Results of the best goodness-of-fit to an approximate coalescent model were retained, providing two phased nuclear haplotypes (alleles). Only likelihood values over

0.9 were unambiguously assigned a base, ambiguous sites were changed to N so as not to bias phylogenetic reconstruction due to uncertainty.

2.2 Phylogeny reconstruction

Sequence alignment was completed with MEGA v5 (Tamura, Peterson et al. 2011) using the MUSCLE algorithm and validated by eye. A total of 312 (220 new + 92 GenBank) *C. rutilus* sequences were included in Cyt b analysis, ranging from 761-1143bp in length. The dataset was partitioned by codon position for all tree reconstructions. For nuclear genes, a subset of individuals was sampled from mtDNA clades (Table 1) to maximize geographic breadth. MrModeltest v2.3 (Nylander 2004) was used to determine the best model of evolution for each gene (and each codon position of Cyt b) based on the Akaike Information Criterion. Phylogenetic relationships for each gene were reconstructed using the program MrBayes 3.2.1 (Huelsenbeck and Ronquist 2001) which applies Bayesian methods and a Markov chain Monte Carlo procedure. Runs were conducted for 12-50 million generations (Table 1), sampling every 1000 generations, with 4 independent chains and a burnin of 25%. Sufficient mixing was assessed by examining the split of the standard deviations in the output, ensuring it was below 0.03 after the burnin phase.

2.3 Species tree estimation

The individual histories of genes may affect the reconstruction of phylogenies, and produce discordance between gene trees and the species tree (Rosenberg 2002). By estimating a species tree from multiple, independent loci, discord among gene trees may be reconciled. Species tree estimation from multiple loci is preferable to concatenation, which can increase the possibility of anomalous gene trees misinforming species tree

construction (Edwards, Liu et al. 2007). The software program BEAUti v7.0.1 was used to set up runs in the Bayesian reconstruction program BEAST (Drummond and Rambaut 2007), which used the *BEAST algorithm (Heled and Drummond 2010) to estimate a species tree. For species tree reconstruction, 33 individuals representing each of the mitochondrial clades and subclades identified in preliminary mtDNA analysis were selected to maximize geographic sampling (Figure 1). For all BEAST analyses, a prior was placed on the node representing the most recent common ancestor (MRCA) of the genus *Clethrionomys* based on a fossil dated to at least 2.6 MYA in the Transbaikalia region of Russia (Repenning, Fejfar et al. 1990). Other fossil evidence or biogeographic events that could be used to reliably date internal nodes is unavailable for the genus (Chaline and Graf 1988). A lognormal uncorrelated molecular clock model was initially applied to all genes. Preliminary runs indicated the clock-like evolution of some genes, so in the final runs a strict clock prior estimated the value instead of the initial relaxed clock. Models of evolution for each gene and each Cyt b codon position were input as priors as informed by ModelTest. Default settings were used for all other parameters unless otherwise suggested by the program documentation. To reduce circularity and independently test divergence times, nuclear data and mtDNA were run separately as well as together. Analyses were run for 300-500 million Markov chain Monte Carlo (MCMC) generations and convergence was assessed by examining run parameters in Tracer, making sure ESSs were >200.

2.4 Population genetic inference

Sequence divergence within and between clades was calculated by average pairwise nucleotide differences for each gene using MEGA. Standard errors were

calculated by 1000 bootstrap replicates. Population genetic parameters were calculated using Arlequin v3 (Excoffier, Laval et al. 2005) and DnaSP v5.10 (Librado and Rozas 2009) on a subset of the total Cyt b data that maximized the number of individuals and sequence length (N=281, 783bp). Individuals were assigned to groups based on inclusion in well-supported mtDNA clades. Sample size (N), number of segregating sites (S), number of haplotypes (H), haplotype diversity (Hd), average number of nucleotide differences (k), and nucleotide diversity (π), were calculated for each clade and the species to assess genetic diversity and demography. Population genetic diversity indices were also calculated for each nuclear gene for *C. rutilus* as a whole to characterize the variation at each nuclear locus. An analysis of molecular variance (AMOVA) was computed in Arlequin to identify how variation was parsed among individuals, populations, and clades for each gene.

2.5 Divergence dates and demographic history

Divergence date estimates were generated simultaneously with phylogeny reconstruction in BEAST (Drummond and Rambaut 2007). Although the species tree was estimated with and without Cyt b, only divergence dates from the four-gene species tree are reported because mtDNA is typically more informative and affords a more accurate estimate of the rate of evolution when combined with nuclear data (Sanchez-Gracia and Castresana 2012). Divergence dates were also estimated from a Cyt b gene tree produced in BEAST using the same set of priors to generate estimates not influenced by nuclear sequence data. The subsample of Cyt b used in species tree analyses was used for this gene tree (N=33).

Population size changes through time were estimated using the Extended Bayesian Skyline algorithm, as implemented in BEAST, for each Cyt b clade (Heled and Drummond 2008). To detect potential signals of demographic expansion using population genetic methods, Fu's F_S (Fu 1997), and R_2 (Ramos-Onsins and Rozas 2002) were calculated in DnaSP using 10,000 coalescent simulations to test significance. These two tests belong to separate classes of neutrality tests that can be used to detect population growth, based on the genetic expectations of rapid population growth (Ramirez-Soriano, Ramos-Onsins et al. 2008). Fu's F_S looks at the haplotype distribution of the data and the frequency of rare alleles whereas R_2 is based on the frequency of rare or unique mutations and the average difference between sequences (Fu 1997; Ramos-Onsins and Rozas 2002). Besides using different sources of information, these tests differ in their sensitivity to recombination as well as their power at different sample sizes (Ramirez-Soriano, Ramos-Onsins et al. 2008). In general, F_S and R_2 have been found to be the most powerful tests available when applied to a variety of demographic and sampling scenarios, especially identifying population expansion (Ramos-Onsins and Rozas 2002; Ramirez-Soriano, Ramos-Onsins et al. 2008). Finally, mismatch distributions of Cyt b sequences were constructed in DnaSP and a goodness-of-fit test was applied to test whether the observed data are consistent with a model of recent expansion (Slatkin and Hudson 1991; Rogers and Harpending 1992; Schneider and Excoffier 1999).

2.6 Species distribution modeling

As an independent test of biogeographic hypotheses, species distribution models (SDM) were generated for LGM and modern conditions for *C. rutilus*. For SDM

development, we used current and LGM monthly climate data at 2.5' (4 km) spatial resolution. Waltari *et al*. (2007) and Peterson & Nyári (2008) discuss the process of layer development more fully. In short, LGM climate data were based on two general circulation model (GCM) simulations: the Community Climate System Model (CCSM) (Collins, Bitz et al. 2006) and the Model for Interdisciplinary Research on Climate (MIROC, version 3.2) ((Hasumi and Emori 2004); http://www.pmip2.cnrs-gif.fr). Each model was initially generated at a spatial resolution of 2.8º, or roughly 300 x 300 km. Outputs from these models were further processed by interpolating differences between LGM and recent (pre-industrial) conditions based on the WorldClim (http://www.worldclim.org/) dataset to create monthly climate surfaces at 2.5' (4 x 4 km) spatial resolution.

SDMs for both the present and LGM were based on the 19 bioclimatic variables in the WorldClim data set (Hijmans, Cameron et al. 2005). These variables represent summaries of means and variation in temperature and precipitation, and characterize dimensions of climate considered particularly relevant in determining species distributions. Present-day SDMs were developed within a mask consisting only of land north of 30º N.

From biodiversity information systems of natural history collection data (e.g. ARCTOS (http:// arctos.database.museum/home.cfm), BOLD (Ratnasingham and Hebert 2007), and MaNIS (Stein and Wieczorek 2004)), we collated georeferenced species occurrence points of *C. rutilus*. To correct for biased sampling effort (Reddy and Davalos 2003), which can lead to model overfitting and subsequent bias, we refined these occurrences by removing localities within 100km, resulting in 126 occurrence points used

for analysis. We used the Maxent (version 3.3.3k) (Phillips, Anderson et al. 2006) program to construct present-day SDMs, and then for projecting the SDM to LGM conditions. We ran 5 replicates using the "crossvalidate" option, in which the localities are divided into five bins, with each bin used as a subset for model testing (equivalent to 20% testing).

Maxent outputs a logistic probability value, ranging from 0 to 1, an indicator of relative suitability for the species, based on the principle of maximum entropy, as constrained by the input occurrence data. We chose a low and high threshold for consideration of 'suitable' habitat based upon the Maxent outputs using contemporary climate data. The low threshold was the median value over the 5 replicates of the lowest presence threshold (LPT) (Pearson, Raxworthy et al. 2007), which is the threshold at which all occurrences in a training dataset fall into suitable habitat, or a 0% omission rate. These values are still more conservative than used in recent studies (Pearson, Raxworthy et al. 2007; Waltari, Hijmans et al. 2007). The high threshold was the median value over the 5 replicates of the Maxent-generated value "Maximum training sensitivity & specificity", which we found to have the least variance of all standard thresholds generated by Maxent (unpublished data). This threshold identified smaller areas than a lowest presence threshold that yielded zero omission error, thus resulting in a more restricted picture of potential LGM distributions. We then generated summary maps in ArcGIS 9.3 which show Maxent predictions for present day and LGM, the latter under both CCSM and MIROC climate models.

3. Results

3.1 Phylogenies and phylogeographic structure

A total of 312 individuals were included in the mtDNA gene tree (Table 1, Appendix A and B). No pseudogenes were identified. Bayesian methods revealed strong support for three geographically structured clades (Figure 2). The Western clade includes all populations from northern Europe, western Siberia and western Mongolia. The Central clade spans from central Siberia to the Bering Strait as well as eastern Mongolia and southeastern Siberia. Finally, the Eastern clade consists of all the North American samples as well as Far-East Asian localities from Hokkaido Island, Sakhalin Island, and the Kamchatka Peninsula (Figure 1). Each of the groups is reciprocally monophyletic; however, relationships among the three are uncertain.

These three widespread clades are effectively parapatric, coming into contact only in very narrow regions (Figure 1). The Western and Central clades overlap slightly along a north-south transect that roughly straddles the Yenisei River, extending from north-central Mongolia to northern Siberia. Only three populations sampled along this transect contain a mixture of individuals with Central haplotypes and Western haplotypes. The location of contact between the Eastern and Central clades is inferred in northeastern Siberia, as populations from Kamchatka are more closely related to North American samples than adjacent Asian populations only a few hundred kilometers north in Siberia.

Substructure within each group is limited except for the Eastern clade, which is composed of three subclades, two are limited to islands off the coast of far eastern Asia, Hokkaido Island and Sakhalin Island (hereafter referred to as the East Asian island subclades), and a third emanates from Beringia (North America and Kamchatka) (Figure 2). Samples from Kamchatka are nested within the North American samples of the

Beringian group rather than forming a single independent lineage, suggesting very recent contact between Kamchatka and the rest of the Beringian subclade.

In contrast to mtDNA, nuclear gene trees were geographically unstructured, reflecting limited sorting of nuclear genes in *C. rutilus* (Figure 3). Levels of variability among nuclear genes differed, with highest variability in ETS2, intermediate levels in IRBP, and very limited variation in MLR (Table 2). MLR generally displays a slow evolutionary rate in rodents (Matocq, Shurtliff et al. 2007). Nonetheless, the single informative site in MLR distinguishes eastern and western groups, with heterozygous individuals in northeast Siberian populations, east of the Verkhoyansk Range and Kolyma River region (Figure 1; locality #45-49, 54, 59, 61, 63), but only homozygotes in other populations. One allele is found in all North American populations and three Asian populations in northeastern Siberia (locality #53, 44, 62). The alternative allele is homozygous in all other populations sampled west of the Kolyma River, including the East Asian islands. The Verkhoyansk Range and Kolyma River region has repeatedly been described as a contact zone for Holarctic species (Hewitt 2000).

3.2 Species tree estimation

Species tree topologies are similar when estimated from only nuclear DNA or all genes including mtDNA. Both trees had relatively low support values for bifurcating relationships among Cyt b clades (Figure 4). However, both species trees strongly support monophyly of *C. rutilus* despite lack of support for monophyly in gene trees based on each individual nuclear gene (Figure 3). The t_{MRCA} estimate for *C. rutilus* clades is between 130-170KYA from the species trees, in contrast to 530 KYA according to the Cyt b BEAST gene tree (Figures 4 and 5). Poor resolution of the species trees and

disparity in date estimates for the origin of the species is most likely a consequence of lack of sorting among nuclear sequence data. Low variability, especially in MLR (Tables 4 and 5), and low resolution, including placement of outgroups, may be falsely compressing the t_{MRCA} estimates. Such a compression is supported by the two species trees producing similar date estimates for *C. rutilus* (170KYA from all loci, 130KYA from nuclear loci only; Figure 4).

All t_{MRCA} estimates place the ancestor of all modern mitochondrial lineages of *C. rutilus* prior to or during the Penultimate glaciations that lasted from about 300-130 KYA. Because reliable fossil estimates were used as calibrations and the nuclear dates are suspect, as mentioned above, we use dates from the Cyt b BEAST gene tree. The age of each of the three major clades can be traced to their respective common ancestors between 100-200KYA (West: 160KYA; Central: 190KYA; East, Beringian subclade: 130KYA) (Figure 5). These estimates slightly predate previous values based on a 4% per MY divergence rate for Cyt b and sampling from populations in Beringia (Hope, Takebayashi et al. 2012), but corroborate a signal of isolation through multiple glacial cycles. The more recent t_{MRCA} estimates for the East Asian island subclades dates to near the LGM (Sakhalin Island: 14KYA; Hokkaido Island: 40KYA) (Figure 5) but before the islands ceased to be a continuous landmass with the mainland about 8-10KYA (Dobson 1994).

3.3 Population differentiation and demographic history

In general, *C. rutilus* is characterized by high genetic diversity, deep sequence divergence, and signals of population expansion based on several different methods. Cyt b variation is high for *C. rutilus*, with high haplotype diversity, nucleotide diversity, and

average number of nucleotide differences between sequences (Table 2). At the clade level, haplotype diversity is high but nucleotide diversity is low, with the exception of Sakhalin Island. Summary statistics and population genetic analyses strongly corroborate the phylogenetic results showing the existence of multiple clades and subclades of *C. rutilus*. Cyt b sequence divergence between major clades and subclades ranges from 0.012- 0.033 (Table 3A). Nuclear gene divergence is much lower than Cyt b, with inter-clade values ranging from 0.002-0.014, which are comparable to many of the average intra-clade divergences at the same locus (Table 3B-D). Regarding Cyt b, the East Asian island haplotypes are most similar to one another while the Central clade shows the highest divergence (0.033), from both the Eastern clade and the Sakhalin Island subclade (Table 3A).

Monophyly of each geographic group and divergence among clades is further illustrated by the high percentage of variation (84.4%) attributed to among-group variation and low within group variation (10.2%) by the AMOVA for Cyt b (Table 4). Although the nuclear gene trees appear relatively uninformative, AMOVA analyses revealed significant underlying structure corresponding to the Cyt b clades as well. Among all genes, only ETS2 did not have the highest percentage of variation distributed among groups, but every gene had significant among group variation.

Multiple tests showed significant signals of population expansion for all clades besides Sakhalin Island. The three main clades each had strongly significant results for F_S and R_2 (Table 2). Mismatch distributions for each of the clades and subclades corroborate the hypothesis of recent expansion, as evidenced by single peaks and high frequencies of few average differences between sequences from the same clade (Figure

6). The distribution for the species as a whole contains 2-3 peaks, which is the expected distribution due to deep subdivision and divergence. Extended Bayesian Skyline Plots (EBSPs) provide a more detailed look at demographic expansion that includes estimating the timing of population growth. Because fossil dates could not be applied to this intraspecific analysis, we calibrated the runs with the Cyt b molecular clock rate estimated by the BEAST Cyt b gene tree (3.5%/MY). Previous estimates of arvicoline and Clethrionomyini Cyt b mutation rates indicate a faster rate than in most other mammals (2-7%/MY; (Conroy and Cook 1999; Brunhoff, Galbreath et al. 2003); 4.4-5.9%/MY (Lebedev, Bannikova et al. 2007)). Central and Eastern clades entered periods of population growth about 50-60KYA and continued to grow through the present (Figure 8). In contrast to other tests, the Western clade EBSP does not show a significant population expansion. This was the only case in which population expansion was not supported for any one of the major clades. Because MLR was invariable within the Western clade, it could not be used to construct the EBSP and was excluded for the Western clade's plot.

3.4 Species distribution models

Modern day predictions for suitable environmental conditions largely overlap the known distribution of *C. rutilus* (Figure 7). Notable exceptions include an extension of predicted occurrence in North America, Europe, and into the Tibetan Plateau of central Asia. The absence of *C. rutilus* from these regions in North America and Europe is often attributed to competitive exclusion by congeners *C. gapperi* and *C. glareolus*, which primarily persisted south of the ice sheets in North America and Europe respectively, as evidenced by fossils, geographically broad parapatric zones and known hybridization

zones (Deffontaine, Libois et al. 2005; Runck, Matocq et al. 2009; Melnikova, Kshnyasev et al. 2012). Similarly, the Tibetan Plateau supports many closely related genera from the Tribe Clethrionomyini, although to our knowledge, competition of arvicolines there has not been explicitly studied.

SDMs for LGM conditions were more restricted than under modern conditions, as expected given LGM climate and glacial extent (Figure 7). However, there is a high degree of overlap between the distributions at both time points. Both MIROC and CCSM models identified relatively large regions during the LGM that presumably acted as refugia in previous glacial periods in the Late Pleistocene, allowing *C. rutilus* to persist through maximal glacial extents. Although results from the two models are similar, the MIROC simulations more closely corroborate the genetic signature and biogeographic history of *C. rutilus*. The MIROC model identified an isolated region of predicted high suitability in the Altay Mountains and adjacent region in Siberia with small high probability areas extending west to the Ural Mountains. Another mostly contiguous area of high suitability stretches from eastern Mongolia to the Pacific coast and northeast into Beringia. While Beringia is commonly seen as a glacial refugium for northern mammals in North America, the other two have not been recognized as critical sources of post-glacial colonizers of boreal or tundra species. The LGM SDMs are consistent with paleoecological work that described the extent and location of forest communities in these areas during glacial periods in northern Asia (Tarasov, Volkova et al. 2000; Brubaker, Anderson et al. 2005; Binney, Willis et al. 2009) as well as fossil of *C. rutilus* in the Ural Mountains and Transbaikalia regions during glacial periods (Markova, Smirnov et al. 1995).

4. Discussion

In northern high-latitude ecosystems, cyclical environmental changes over the past 2-3 million years played a pivotal role in determining species distributions and genetic diversity (Webb III and Bartlein 1992; Hofreiter and Stewart 2009). During the Quaternary (about the past 2.5 million years), the cyclical build up of continental ice sheets led to significant sea level changes and shifting landscapes. Glacial periods occurred at regular intervals throughout the Quaternary but were marked by increased intensity and frequency toward Late Pleistocene times (Hofreiter and Stewart 2009; Miller, Brigham-Grette et al. 2010). During glacial periods, vast areas of North America and northern Europe were covered by ice sheets, sea level fluctuations repeatedly exposed low-lying land, and regional climate patterns changed, altering many organisms' distributions and creating novel interactions (Hopkins 1967; Hofreiter and Stewart 2009).

In addition to sea level changes, global climate circulation systems were altered and shifts in ice coverage directly affected the distribution of northern high-latitude ecosystems and associated organisms. Some regions were covered by ice sheets for extended periods (e.g. Canada, Scandinavia), while others experienced only localized montane glaciations or remained ice-free (e.g. north-central Asia)(Glushkova 2001; Ehlers and Gibbard 2007). For many species, isolation in multiple refugia allowed persistence through glacial periods and contributed to increased genetic diversity (Hewitt 2004; Stewart, Lister et al. 2010). Refugial populations were important sources for recolonization of large geographic areas after glaciers receded. Similarly, large unglaciated regions (e.g. Beringia) were characterized by habitat heterogeneity, which would have allowed some degree of connectivity and dispersal for a variety of taxa, but

also facilitated in situ diversification within a single large refugium (Hopkins, Matthews et al. 1982; Weksler, Lanier et al. 2010; Galbreath, Cook et al. 2011).

4.1 Phylogeographic structure and biogeographic history of C. rutilus

The biogeographic history of northern red-backed voles conforms to a model of vicariance as the main process structuring genetic diversity across its range rather than a model of recent continental-scale dispersal from a single glacial refugium. Multilocus genetic data are consistent with SDMs in defining a history of isolation in multiple glacial refugia, an unusual history for species primarily associated with boreal forests (Fedorov, Goropashnaya et al. 2008). However, *C. rutilus* also is associated with tundra and "non-analog" fossil assemblages in the western Siberia that contained a mix of forest and tundra associated species (Gromov and Polyakov 1977; Markova, Smirnov et al. 1995; Borodin 1996). Phylogeographic structure is deep, as revealed by population genetic tests, sequence divergence, and the mtDNA gene tree. Rather than events of the LGM and Holocene playing critical roles, isolation and genetic divergence were initiated during the Penultimate glaciations (300-130KYA).

Three major mitochondrial lineages, each accounting for similar proportions of the current Holarctic distribution of *C. rutilus*, most likely persisted in Beringia (Eastern clade), eastern Mongolia and/or southeastern Siberia (Central clade), and from the Altay Mountains of western Mongolia to the Ural Mountains (Western clade) (Figure 1 and 2). Separation of these lineages was nearly simultaneous prior to or during the Penultimate glaciations and they apparently remained isolated (Table 3 and 4). Lack of resolution in nuclear gene trees and species trees likely reflects insufficient time for the nuclear genetic lineages to completely sort, as expected when compared to the faster evolving mtDNA.

Alternative explanations, such as recent bouts of repeated contact and genetic exchange over time, are unlikely given such a strong opposing signal in mtDNA that indicates isolation for at least the last 150KYA (Figure 4). Large historic population sizes prior to vicariance may also have harbored high genetic variation that has contributed to long nuclear sorting times.

A weak signal of recent shared ancestry between the Eastern and Western mtDNA clades to the exclusion of the Central clade, if corroborated, would be a novel biogeographic pattern for a Holarctic species (to date, all have shared the typical correlation between genealogic affinity and geographic proximity (DeChaine 2008)). This pattern was previously suggested for *C. rutilus* based on limited sampling (Iwasa, Kartavtseva et al. 2002; Malyarchuk 2011) and is reported for the temperate Eurasian rodent, *Micromys minutes*, (Yasuda, Vogel et al. 2005), but not for any boreal or arctic species.

The three main lineages of *C. rutilus* are parapatrically distributed over the vast Holarctic, with longitudinal spans rivaling that of any other mammal (Figure 1). Northern species distributed across Asia or the Holarctic typically dispersed to their current extent from a single refugial source, or occasionally two. The three prominent sources identified for *C. rutilus* is unusual even among highly structured species (Marmi, Lopez-Giraldez et al. 2006; Korsten, Ho et al. 2009). Multiple refugia for European forest species is commonly hypothesized and complex phylogeographic structure associated with multiple Asian refugia has previously been shown in tundra (Hewitt 1999; Hewitt 2000) and boreal species(Hope, Waltari et al. 2011).

Beringia, the location of the Eastern clade refugium, also played a role in generating diversity in Asian populations of *C. rutilus*. Similar to other Beringian species, *C. rutilus* dispersed from Asia to North America through Beringia after LGM (Waltari, Hoberg et al. 2007). By persisting in western Beringia, multiple genetic lineages now exist in northeastern Siberia, as also identified for several other Holarctic species (Hewitt 2004). East Asian Islands populations, particularly from Hokkaido Island, also harbor surprising levels of genetic diversity, similar to Beringian populations. Close phylogenetic association of islands and the rest of the Eastern mtDNA clade suggests historic connections along the Pacific Rim. SDMs also indicate a corridor of suitable conditions along the Pacific coast during LGM that extended from the islands to Kamchatka and eastward across southern Beringia (Figure 7). Although the Kuril Islands appear to provide an alternative pathway for dispersal, a deep ocean trench divides the island chain and did so even during the lowest sea levels during Quaternary glacial periods (Pietsch, Bogatov et al. 2003).

MLR variation identifies a split between Beringian and Far East Island subclades that occurred before rising sea levels submerged the connection between these islands and the Asian mainland (about 8-10KYA; (Dobson 1994)). Divergence estimates corroborate this sequential pattern (Figure 5). Further sampling across the Far East Asian mainland may identify a region of contact between the Eastern and Central clades, similar to the pattern in northeast Siberia. SDMs lend support to this hypothesis because the entire region appears highly suitable for *C. rutilus* during the LGM.

4.2 Demographic history

Demographic tests revealed signals of rapid expansion (Table 2; Figure 6 and 8), began as early as 50-80KYA for Central and Eastern clades, during the glacial period prior to the LGM (Figure 8). This time frame also corresponds to the start of a well-defined interglacial period between the last two glacial advances (Svendsen, Alexanderson et al. 2004). Population growth increased for both clades during this interglacial warm period and continued through the LGM and Holocene. SDMs for the LGM show, *C. rutilus* was not highly restricted. Signatures of population growth, albeit slower during this period, reflect the idiosyncrasy of species responses to historic climate change. Because *C. rutilus* inhabits boreal forest and tundra to a greater extent than most strictly forest-associated species, overall distribution was less limited during cold periods, although barriers to dispersal isolated clades for hundreds of thousands of years. As the paleoecological and fossil record are further resolved for central Asia during the Penultimate glaciations and the Last Interglacial factors that led to phylogeographic structure in *C. rutilus* can be more explored in greater detail.

Lack of significant growth in populations of the Western clade may relate to maintenance of a relatively larger distribution than other populations. Fossils from the Ural Mountains as well as SDMs indicate possible persistence of *C. rutilus* from the Altay mountains to the Urals throughout the last 130KYA (Markova, Smirnov et al. 1995). Northern parts of this region were more steppe- or grassland-like during cold periods than at present while further east, the Altay Mountains and adjacent areas of south-central Siberia and western Mongolia were mostly ice-free and harbored vegetation during the LGM similar to contemporary landscapes (Tarasov, Volkova et al. 2000). Based on mtDNA variation, populations of *C. rutilus* near the Altay have the highest

genetic diversity of populations surveyed from there to Scandinavia, a distribution corresponding to the Western clade (Melnikova, Kshnyasev et al. 2012).

4.3 Contact zones

A zone of secondary contact along the Yenisei River between populations of the Central and Western clades without extensive penetration into the core range of either clade is consistent with relatively recent contact, perhaps resulting from expansion of Central clade population beginning about 50KYA (Figure 1 and 8). Finer scale sampling along the zone of contact may help elucidate the extent and dynamics of intermixing. Other small mammals exhibit a phylogeographic break at the Yenisei(Hope, Waltari et al. 2010), although, many other species share a common contact zone further west at the Ural Mountains (Brunhoff, Galbreath et al. 2003; Hewitt 2004). The Yenisei River defines the border between two subspecies of *C. rutilus* on the basis of dental morphology (Gromov and Polyakov 1977), the only case where subspecies assignments correspond to mtDNA lineages. Even the Far Eastern Island populations, which have historically been grouped together as a single subspecies (Gromov and Polyakov 1977), are reciprocally monophyletic and should be treated as separate based on our analyses.

A second potential contact zone, identified by MLR variation, lies to the east in the Kolyma River region in northeastern Siberia, a well-known contact zone for Holarctic taxa (Galbreath and Cook 2004; Hewitt 2004). Central and Eastern clades came into contact near the Kolyma, presumably due to expansion. The mtDNA signal of the Eastern clade was apparently then swamped by the expanding Central clade sometime after 50KYA, but the nuclear signal remains intact in the putative original region of contact near the Verkhoyansk Range and Kolyma River. Prior to this time, the

Verkhoyansk Range was heavily glaciated and prevented contact, but the glaciers were far less extensive during the LGM (Bespalyy 1984). Absence of Central clade haplotypes in Alaska and St. Lawrence Island in the Bering Strait imply the expansion of the Central Clade in extreme northeastern Siberia may have occurred more recently, even after the flooding of the Bering Land Bridge.

Because populations from Kamchatka are more closely related to individuals in North America than to adjacent populations just north in Siberia, the Central clade apparently was prevented from colonizing all of the former range of the Eastern clade in Beringia. Other mammals have maintained distinct lineages in Kamchatka (Korsten, Ho et al. 2009), confirming the inaccessibility of the peninsula for millennia. Furthermore, many trans-Beringian species have close genetic ties that cross the Bering Strait, emphasizing the role of events prior to the LGM in structuring modern high-latitude biodiversity, that contrast with contemporary biogeographic barriers (Hewitt 2004; Hope, Takebayashi et al. 2012).

4.4 Conclusion

The dynamic and idiosyncratic biogeographic history exhibited by northern red-backed voles demonstrates the importance of distinct Asian refugia for Holarctic diversification and recolonization during the Late Pleistocene. Such a complex and dynamic history of vicariance, expansion, and secondary contact within a single species emphasizes the importance of complete geographic sampling. The historical dynamics of species distributions, continually changing over time and space, shaped community dynamics, ecological complexity, and evolutionary interactions observed today. By understanding how species responded to past change, we improve our ability to forecast

the effects of future change in our rapidly changing world (Parmesan 2006). However, although *C. rutilus* shares aspects of history with other northern species, such as the location of some refugia and contact zones that reflect important biogeographic processes for arctic and boreal species, each species responds individualistically.

References

Abbott, R. J. and C. Brochmann (2003). "History and evolution of the arctic flora: in the footsteps of Eric Hulten." Molecular Ecology **12**(2): 299-313.

Abramson, N. I., V. S. Lebedev, et al. (2009). "Supraspecies Relationships in the Subfamily Arvicolinae (Rodentia, Cricetidae): An Unexpected Result of Nuclear Gene Analysis." Molecular Biology **43**(5): 834-846.

Alves, P. C., J. Melo-Ferreira, et al. (2008). "The ubiquitous mountain hare mitochondria: multiple introgressive hybridization in hares, genus *Lepus*." Philosophical Transactions of the Royal Society B: Biological Sciences **363**(1505): 2831-2839.

Avise, J. C. (2000). Phylogeography: The history and formation of species. Cambridge, Massachusetts, Harvard University Press.

Avise, J. C. (2009). "Phylogeography: retrospect and prospect." Journal of Biogeography **36**(1): 3-15.

Barraclough, T. G. and S. Nee (2001). "Phylogenetics and speciation." Trends in Ecology & Evolution **16**(7): 391-399.

Bespalyy, V. (1984). Late Pleistocene mountain glaciation in northeastern USSR. Late Quaternary Environments of the Soviet Union. A. A. Velichko. Minneapolis, University of Minnesota Press**:** 31-33.

Binney, H. A., K. J. Willis, et al. (2009). "The distribution of late-Quaternary woody taxa in northern Eurasia: evidence from a new macrofossil database." Quaternary Science Reviews **28**(23): 2445-2464.

Boratynski, Z., P. C. Alves, et al. (2011). "Introgression of mitochondrial DNA among *Myodes* voles: consequences for energetics?" BMC evolutionary biology **11**(1): 355.

Borodin, A. V. (1996). "Quaternary small mammal faunas from the west Siberian plain." Acta zoologica cracoviensia **39**: 75-81.

Brubaker, L. B., P. M. Anderson, et al. (2005). "Beringia as a glacial refugium for boreal trees and shrubs: new perspectives from mapped pollen data." Journal of Biogeography **32**(5): 833-848.

Brunhoff, C., K. E. Galbreath, et al. (2003). "Holarctic phylogeography of the root vole (*Microtus oeconomus*): implications for late Quaternary biogeography of high latitudes." Molecular Ecology **12**(4): 957-968.

Buzan, E. V., B. Krystufek, et al. (2008). "Mitochondrial phylogeny of Arvicolinae using comprehensive taxonomic sampling yields new insights." Biological Journal of the Linnean Society **94**(4): 825-835.

Carleton, M. D. and G. G. Musser (2005). Order Rodentia. Mammal Species of the World. D. E. Wilson and D. M. Reeder. Baltimore, Maryland, The Johns Hopkins University Press.

Chaline, J., P. Brunet-Lecomte, et al. (1999). Anatomy of the arvicoline radiation (Rodentia): palaeogeographical, palaeoecological history and evolutionary data. Annales Zoologici Fennici, Helsinki: Suomen Biologian Seura Vanamo, 1964-.

Chaline, J. and J. D. Graf (1988). "Phylogeny of the Arvicolidae (Rodentia): biochemical and paleontological evidence." Journal of Mammalogy: 22-33.

Collins, W. D., C. M. Bitz, et al. (2006). "The community climate system model version 3 (CCSM3)." Journal of Climate **19**(11): 2122-2143.

Conroy, C. J. and J. A. Cook (1999). "MtDNA evidence for repeated pulses of speciation within arvicoline and murid rodents." Journal of Mammalian Evolution **6**(3): 221-245.

Cook, J. A., E. P. Hoberg, et al. (2005). "Beringia: Intercontinental exchange and diversification of high latitude mammals and their parasites during the Pliocene and Quaternary." Mammal Study **30**(Suppl. S): S33-S44.

Cook, J. A., A. M. Runck, et al. (2004). "Historical biogeography at the crossroads of the northern continents: molecular phylogenetics of red-backed voles (Rodentia : Arvicolinae)." Molecular phylogenetics and evolution **30**(3): 767-777.

Cruz, R., H. Harasawa, et al. (2007). Asia. In Climate Change 2007: impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.(Eds ML Parry, OF Canziani, JP Palutikof, PJ van der Linden, CE Hanson) pp. 469-506, Cambridge University Press: Cambridge, UK.

Cusimano, N. and S. S. Renner (2010). "Slowdowns in diversification rates from real phylogenies may not be real." Systematic Biology **59**(4): 458-464.

Darwin, C. (1859). "On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life." New York: D. Appleton.

DeChaine, E. G. (2008). "A bridge or a barrier? Beringia's influence on the distribution and diversity of tundra plants." Plant Ecology & Diversity **1**(2).

Deffontaine, V., R. Libois, et al. (2005). "Beyond the Mediterranean peninsulas: evidence of central European glacial refugia for a temperate forest mammal species, the bank vole (*Clethrionomys glareolus*)." Molecular Ecology **14**(6): 1727-1739.

Dobson, M. (1994). "Patterns of distribution in Japanese land mammals." Mammal Review **24**(3): 91-111.

Doiron, S., L. Bernatchez, et al. (2002). "A comparative mitogenomic analysis of the potential adaptive value of arctic charr mtDNA introgression in brook charr populations (*Salvelinus fontinalis* Mitchill)." Molecular biology and evolution **19**(11): 1902-1909.

Drummond, A. J. and A. Rambaut (2007). "BEAST: Bayesian evolutionary analysis by sampling trees." BMC Evolutionary Biology **7**.

Edwards, S. V., L. Liu, et al. (2007). "High-resolution species trees without concatenation." Proceedings of the National Academy of Sciences **104**(14): 5936.

Ehlers, J. and P. L. Gibbard (2007). "The extent and chronology of Cenozoic global glaciation." Quaternary International **164**: 6-20.

Excoffier, L., G. Laval, et al. (2005). "Arlequin (version 3.0): an integrated software package for population genetics data analysis." Evolutionary Bioinformatics Online **1**: 47-50.

Fedorov, V. B. (1999). "Contrasting mitochondrial DNA diversity estimates in two sympatric genera of Arctic lemmings (*Dicrostonyx: Lemmus*) indicate different responses to Quaternary environmental fluctuations." Proceedings of the Royal Society of London. Series B: Biological Sciences **266**(1419): 621-626.

Fedorov, V. B., A. V. Goropashnaya, et al. (2008). "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.

Fedorov, V. B., A. V. Goropashnaya, et al. (2003). "Phylogeography of lemmings (*Lemmus*): no evidence for postglacial colonization of Arctic from the Beringian refugium." Molecular Ecology **12**(3): 725-731.

Frisman, L. V., I. V. Kartavtseva, et al. (2002). "Gene-geographical variation and genetic differentiation in red-backed voles of the genus *Clethrionomys* (Rodentia, Cricetidae) from the region of the Sea of Okhotsk." Genetika **38**(5): 655-664.

Fu, Y. X. (1997). "Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection." Genetics **147**(2): 915-925.

Galbreath, K. E. and J. A. Cook (2004). "Genetic consequences of Pleistocene glaciations for the tundra vole (*Microtus oeconomus*) in Beringia." Molecular Ecology **13**(1): 135-148.

Galbreath, K. E., J. A. Cook, et al. (2011). "Diversity and demography in Beringia: Multilocus tests of paleodistribution models reveal the complex history of arctic ground squirrels." Evolution **65**(7): 1879-1896.

Galewski, T., M.-k. Tilak, et al. (2006). "The evolutionary radiation of Arvicolinae rodents (voles and lemmings): relative contribution of nuclear and mitochondrial DNA phylogenies." Bmc Evolutionary Biology **6**(80): 1-17.

Gromov, I. M. and I. Y. Polyakov (1977). Voles (Microtinae), Fauna of the USSR, Mammals Washington, D.C., Smithsonian Institution Libraries and the National Science Foundation.

Guthrie, R. (1968). "Paleoecology of a late Pleistocene small mammal community from interior Alaska." Arctic **21**(4): 223-244.

Hadly, E. A., U. Ramakrishnan, et al. (2004). "Genetic response to climatic change: insights from ancient DNA and phylochronology." PLoS Biology **2**(10): e290.

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

Harrison, S., G. Yu, et al. (2001). "Palaeovegetation (Communications arising): diversity of temperate plants in east Asia." Nature **413**(6852): 129-130.

Hasumi, H. and S. Emori (2004). "K-1 coupled GCM (MIROC) description." Center for Climate System Research, University of Tokyo, Tokyo.

Heled, J. and A. J. Drummond (2008). "Bayesian inference of population size history from multiple loci." BMC Evolutionary Biology **8**(1): 289.

Heled, J. and A. J. Drummond (2010). "Bayesian Inference of Species Trees from Multilocus Data." Molecular Biology and Evolution **27**(3): 570-580.

Hewitt, G. (2000). "The genetic legacy of the Quaternary ice ages." Nature **405**(6789): 907-913.

Hewitt, G. M. (1996). "Some genetic consequences of ice ages, and their role in divergence and speciation." Biological Journal of the Linnean Society **58**(3): 247-276.

Hewitt, G. M. (1999). "Post-glacial re-colonization of European biota." Biological Journal of the Linnean Society **68**(1-2): 87-112.

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**(1442): 183-195.

Hewitt, G. M. (2004). "The structure of biodiversity–insights from molecular phylogeography." Frontiers in Zoology **1**(4): 1-16.

Hijmans, R. J., S. E. Cameron, et al. (2005). "Very high resolution interpolated climate surfaces for global land areas." International Journal of Climatology **25**(15): 1965-1978.

Hoberg, E. P., S. Kutz, et al. (2003). "Arctic biodiversity: From discovery to faunal baselines-Revealing the history of a dynamic ecosystem." Journal of Parasitology **89**(^ sSuppl).

Hoffmann, R. (1981). Different voles for different holes: environmental restrictions on refugial survival of mammals. Evolution today.

Hofreiter, M. and J. Stewart (2009). "Ecological change, range fluctuations and population dynamics during the Pleistocene." Current Biology **19**(14): R584-R594.

Hope, A. G., N. Takebayashi, et al. (2012). "Temporal, spatial and ecological dynamics of speciation among amphi-Beringian small mammals." Journal of Biogeography.

Hope, A. G., E. Waltari, et al. (2010). "High-latitude diversification within Eurasian least shrews and Alaska tiny shrews (Soricidae)." Journal of Mammalogy **91**(5): 1041-1057

Hope, A. G., E. Waltari, et al. (2011). "Persistence and diversification of the Holarctic shrew, *Sorex tundrensis* (Family Soricidae), in response to climate change." Molecular Ecology **20**(20): 4346-4370.

Huelsenbeck, J. P. and F. Ronquist (2001). "MRBAYES: Bayesian inference of phylogenetic trees." Bioinformatics **17**(8): 754-755.

Hultén, E. (1937). Outline of the history of arctic and boreal biota during the Quaternary period. . New York., Lehre J. Cramer.

Irwin, D. M., T. D. Kocher, et al. (1991). "Evolution of the cytochrome *b* gene of mammals." Journal of Molecular Evolution **32**(2): 128-144.

Iwasa, M. A., I. V. Kartavtseva, et al. (2002). "Local differentiation of *Clethrionomys rutilus* in northeastern Asia inferred from mitochondrial gene sequences." Mammalian Biology **67**(3): 157-166.

Iwasa, M. A. and H. Suzuki (2002). "Evolutionary significance of chromosome changes in northeastern Asiatic red-backed voles inferred with the aid of intron 1 sequences of the G6pd gene." Chromosome Research **10**(5): 419-428.

Kaneko, Y. (1992). "Identification and Morphological Characteristics of *Clethrionomys rufocanus*, *Eothenomys shanseius*, *E. inez* and *E. eva* from the USSR, Mongolia,

and Northern and Central China." Journal of the Mammalogical Society of Japan **16**(2): 71-95.

Kaneko, Y. (1996). "Morphological variation, and latitudinal and altitudinal distribution of *Eothenomys chinensis*, *E. wardi*, *E. custos*, *E. proditor*, and *E. olitor* (Rodentia, Arvicolidae) in China." Mammal Study **21**(2): 89-114.

Koh, H. S., B. K. Yang, et al. (2010). "Genetic distinctiveness of the Korean red-backed vole (*Myodes regulus*) from Korea, revealed by mitochondrial cytochrome *b* gene sequences." Biochemical Genetics: 1-8.

Korsten, M., S. Y. W. Ho, et al. (2009). "Sudden expansion of a single brown bear maternal lineage across northern continental Eurasia after the last ice age: a general demographic model for mammals?" Molecular Ecology **18**(9): 1963-1979.

Kotlik, P., V. Deffontaine, et al. (2006). "A northern glacial refugium for bank voles (*Clethrionomys glareolus*)." Proceedings of the National Academy of Sciences **103**(40): 14860-14864.

Lebedev, V. S., A. A. Bannikova, et al. (2007). "Molecular phylogeny of the genus Alticola (Cricetidae, Rodentia) as inferred from the sequence of the cytochrome b gene." Zoologica Scripta **36**(6): 547-563.

Li, J. and X. Fang (1999). "Uplift of the Tibetan Plateau and environmental changes." Chinese Science Bulletin **44**(23): 2117-2124.

Librado, P. and J. Rozas (2009). "DnaSP v5: a software for comprehensive analysis of DNA polymorphism data." Bioinformatics **25**(11): 1451-1452.

Lister, A. M. (2004). "The impact of Quaternary Ice Ages on mammalian evolution." Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **359**(1442): 221.

Liu, S., Y. Liu, et al. (2012). "Phylogeny of Oriental Voles (Rodentia: Muridae: Arvicolinae): Molecular and Morphological Evidence." Zoological Science **29**(9): 610-622.

Luo, J., D. Yang, et al. (2004). "Molecular phylogeny and biogeography of Oriental voles: genus *Eothenomys* (Muridae, Mammalia)." Molecular phylogenetics and evolution **33**(2): 349-362.

Lyons, L. A., T. F. Laughlin, et al. (1997). "Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes." Nature Genetics **15**(1): 47-56.

MacDonald, S. O. and J. A. Cook (2007). Checklist of the Land Mammals of Beringia. Albuquerque, NM, University of New Mexico.

Maddison, W. and D. Maddison (2011). Mesquite: a modular system for evolutionary analysis. Version 2.75.

Maddison, W. P. (1997). "Gene trees in species trees." Systematic biology **46**(3): 523.

Malyarchuk, B. (2011). "Adaptive intraspecific divergence: An example using the animal cytochrome b gene." Russian Journal of Genetics **47**(8): 979-986.

Markova, A. K., N. G. Smirnov, et al. (1995). "Late Pleistocene distribution and diversity of mammals in northern Eurasia." Paleontologia i Evolucio(28-29): 5-143.

Marmi, J., F. Lopez-Giraldez, et al. (2006). "Mitochondrial DNA reveals a strong phylogeographic structure in the badger across Eurasia." Molecular Ecology **15**(4): 1007-1020.

Matocq, M. D., Q. R. Shurtliff, et al. (2007). "Phylogenetics of the woodrat genus *Neotoma* (Rodentia : Muridae): Implications for the evolution of phenotypic variation in male external genitalia." Molecular Phylogenetics and Evolution **42**(3): 637-652.

Mayr, E. (1970). Populations, species, and evolution: an abridgment of animal species and evolution, Belknap Press.

Melnikova, E. N., I. A. Kshnyasev, et al. (2012). "Sympatric area of *Myodes glareolus* and *M. rutilus* (Rodentia, Cricetidae): Historic and recent hybridization." Proceedings of the Zoological Institue RAS **316**(4): 307-323.

Miller, G. H., J. Brigham-Grette, et al. (2010). "Temperature and precipitation history of the Arctic." Quaternary Science Reviews **29**(15): 1679-1715.

Nadler, C. F., N. Zhurkevich, et al. (1978). "Biochemical relationships of the Holarctic vole genera (*Clethrionomys*, *Microtus*, and *Arvicola* (Rodentia: Arvicolinae))." Canadian journal of zoology **56**(7): 1564-1575.

Ni, J., G. Yu, et al. (2010). "Palaeovegetation in China during the late Quaternary: Biome reconstructions based on a global scheme of plant functional types." Palaeogeography, Palaeoclimatology, Palaeoecology **289**(1-4): 44-61.

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

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

Pearson, R. G., C. J. Raxworthy, et al. (2007). "Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar." Journal of Biogeography **34**(1): 102-117.

Peterson, A. T. and A. S. Nyari (2008). "Ecological niche conservatism and Pleistocene refugia in the thrush-like mourner, *Schiffornis* sp., in the Neotropics." Evolution **62**(1): 173-183.

Phillips, C. J. (1969). "Review of Central Asian Voles of the Genus *Hyperacrius*, with Comments on Zoogeography, Ecology, and Ectoparasites." Journal of Mammalogy: 457-474.

Phillips, S. J., R. P. Anderson, et al. (2006). "Maximum entropy modeling of species geographic distributions." Ecological Modelling **190**(3-4): 231-259.

Pietsch, T. W., V. V. Bogatov, et al. (2003). "Biodiversity and biogeography of the islands of the Kuril Archipelago." Journal of Biogeography **30**(9): 1297-1310.

Platt, A. R., R. W. Woodhall, et al. (2007). "Improved DNA sequencing quality and efficiency using an optimized fast cycle sequencing protocol." BioTechniques **43**(1): 58.

Pybus, O. and A. Rambaut (2002). "GENIE: estimating demographic history from molecular phylogenies." Bioinformatics **18**(10): 1404-1405.

Pybus, O. G. and P. H. Harvey (2000). "Testing macro-evolutionary models using incomplete molecular phylogenies." Proceedings of the Royal Society of London. Series B: Biological Sciences **267**(1459): 2267-2272.

Ramirez-Soriano, A., S. E. Ramos-Onsins, et al. (2008). "Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination." Genetics **179**(1): 555-567.

Ramos-Onsins, S. E. and J. Rozas (2002). "Statistical properties of new neutrality tests against population growth." Molecular biology and evolution **19**(12): 2092-2100.

Ratnasingham, S. and P. D. N. Hebert (2007). "BOLD: The Barcode of Life Data System (http://www.barcodinglife.org)." Molecular Ecology Notes **7**(3): 355-364.

Rausch, R. L. (1963). "A review of the distribution of Holarctic recent mammals." Pacific Basin Biogeography: 29–44.

Reddy, S. and L. M. Davalos (2003). "Geographical sampling bias and its implications for conservation priorities in Africa." Journal of Biogeography **30**(11): 1719-1727.

Repenning, C. A. (2001). "Beringian climate during intercontinental dispersal: a mouse eye view." Quaternary Science Reviews **20**(1-3): 25-40.

Repenning, C. A., O. Fejfar, et al. (1990). Arvicolid rodent biochronology of the Northern Hemisphere. *International Symposium on the Evolution and Phylogenetic Biostratigraphy of Arvicolids (Rodentia, Mammalia)*, Prague, Pfeil-Verlag.

Robovsky, J., V. Ricankova, et al. (2008). "Phylogeny of Arvicolinae (Mammalia, Cricetidae): utility of morphological and molecular data sets in a recently radiating clade." Zoologica Scripta **37**(6): 571-590.

Rogers, A. R. and H. C. Harpending (1992). "Population Growth Makes Waves in the Distribution of Pairwise Nucleotide Differences." American Journal of Physical Anthropology(SUPPL. 14): 140.

Rosenberg, N. A. (2002). "The probability of topological concordance of gene trees and species trees." Theoretical Population Biology **61**(2): 225-247.

Runck, A. M. and J. A. Cook (2005). "Postglacial expansion of the southern red-backed vole (*Clethrionomys gapperi*) in North America." Molecular Ecology **14**(5): 1445-1456.

Runck, A. M., M. D. Matocq, et al. (2009). "Historic hybridization and persistence of a novel mito-nuclear combination in red-backed voles (genus *Myodes*)." BMC Evolutionary Biology **9**(114): 1-15.

Sanchez-Gracia, A. and J. Castresana (2012). "Impact of Deep Coalescence on the Reliability of Species Tree Inference from Different Types of DNA Markers in Mammals." PloS one **7**(1): e30239.

Schneider, S. and L. Excoffier (1999). "Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA." Genetics **152**(3): 1079-1089.

Serdyuk, N. V. and A. S. Tesakov (2006). "New form of rhizodont voles (Rodentia, Arvicolinae, Clethrionomyini) from Pleistocene of Central Altai (Russia)." Russian Journal of Theriology **5**(2): 79-83.

Shafer, A., C. Cullingham, et al. (2010). "Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America." Molecular Ecology **19**(21): 4589.

Silberfeld, T., J. W. Leigh, et al. (2010). "A multi-locus time-calibrated phylogeny of the brown algae (Heterokonta, Ochrophyta, Phaeophyceae): Investigating the evolutionary nature of the "brown algal crown radiation"." Molecular phylogenetics and evolution **56**(2): 659-674.

Slatkin, M. and R. R. Hudson (1991). "Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations." Genetics **129**(2): 555-562.

Smith, C. I. (2007). "Historical biogeography: The new synthesis." Current Biology **17**(15): R598-R600.

Smith, M. F. and J. L. Patton (1993). "The diversification of South American murid rodents: evidence from mitochondrial DNA sequence data for the akodontine tribe." Biological Journal of the Linnean Society **50**(3): 149-177.

Stanhope, M. J., M. R. Smith, et al. (1996). "Mammalian evolution and the interphotoreceptor retinoid binding protein (IRBP) gene: convincing evidence for several superordinal clades." Journal of Molecular Evolution **43**(2): 83-92.

Stein, B. R. and J. R. Wieczorek (2004). "Mammals of the world: MaNIS as an example of data integration in a distributed network environment." Biodiversity Informatics **1**(0).

Stephens, M., N. J. Smith, et al. (2001). "A new statistical method for haplotype reconstruction from population data." American Journal of Human Genetics **68**(4): 978-989.

Stewart, J. R. (2009). "The evolutionary consequence of the individualistic response to climate change." Journal of Evolutionary Biology **22**(12): 2363-2375.

Svendsen, J. I., H. Alexanderson, et al. (2004). "Late Quaternary ice sheet history of northern Eurasia." Quaternary Science Reviews **23**(11): 1229-1271.

Tamura, K., D. Peterson, et al. (2011). "MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods." Molecular biology and evolution **28**(10): 2731-2739.

Tarasov, P. E., V. S. Volkova, et al. (2000). "Last glacial maximum biomes reconstructed from pollen and plant macrofossil data from northern Eurasia." Journal of Biogeography **27**(3): 609-620.

Tegelström, H. (1987). "Transfer of mitochondrial DNA from the northern red-backed vole (*Clethrionomys rutilus*) to the bank vole (*C. glareolus*)." Journal of Molecular Evolution **24**(3): 218-227.

Tesakov, A. S., V. S. Lebedev, et al. (2010). "*Clethrionomys* Tilesius, 1850 is the valid generic name for red-backed voles and *Myodes* Pallas, 1811 is a junior synonym of *Lemmus* Link, 1795." Russian Journal of Theriology **9**(2): 83-86.

Thompson, J. N. (2005). The Geographic Mosaic of Coevolution, University of Chicago Press.

Todisco, V., P. Gratton, et al. (2012). "Mitochondrial phylogeography of the Holarctic *Parnassius phoebus* complex supports a recent refugial model for alpine butterflies." Journal of Biogeography **39**(6): 1058-1072.

Tolley, K. A., B. M. Chase, et al. (2008). "Speciation and radiations track climate transitions since the Miocene Climatic Optimum: a case study of southern African chameleons." Journal of Biogeography **35**(8): 1402-1414.

Velichko, A. A. and H. E. Wright (2005). Cenozoic climatic and environmental changes in Russia, Geological Society of Amer.

Waltari, E., R. J. Hijmans, et al. (2007). "Locating Pleistocene refugia: Comparing phylogeographic and ecological niche model predictions." PLoS One **2**(7): 1-11.

Waltari, E., E. P. Hoberg, et al. (2007). "Eastward Ho: phylogeographical perspectives on colonization of hosts and parasites across the Beringian nexus." Journal of Biogeography **34**(4): 561-574.

Webb III, T. and P. Bartlein (1992). "Global changes during the last 3 million years: climatic controls and biotic responses." Annual Review of Ecology and Systematics **23**: 141-173.

Weksler, M., H. C. Lanier, et al. (2010). "Eastern Beringian biogeography: historical and spatial genetic structure of singing voles in Alaska." Journal of Biogeography **37**(8): 1414-1431.

Wilson, D. E. and D. M. Reeder (2005). Mammal Species of the World. Baltimore, Maryland, The Johns Hopkins University Press.

Yasuda, S. P., P. Vogel, et al. (2005). "Phylogeographic patterning of mtDNA in the widely distributed harvest mouse (*Micromys minutus*) suggests dramatic cycles of range contraction and expansion during the mid-to late Pleistocene." Canadian Journal of Zoology **83**(11): 1411-1420.

Zhang, Y. and S. Ge (2007). "Molecular evolution study in China: progress and future promise." Philosophical Transactions of the Royal Society B: Biological Sciences **362**(1482): 973-986.

Zheng, S. and C. Li (1990). Comments on fossil arvicolids of China. Int. Symp. Evol. Phyl. Biostr. Arvicolids.

Zhisheng, A., J. E. Kutzbach, et al. (2001). "Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since Late Miocene times." Nature **411**(6833): 62-66.

Figures and Tables

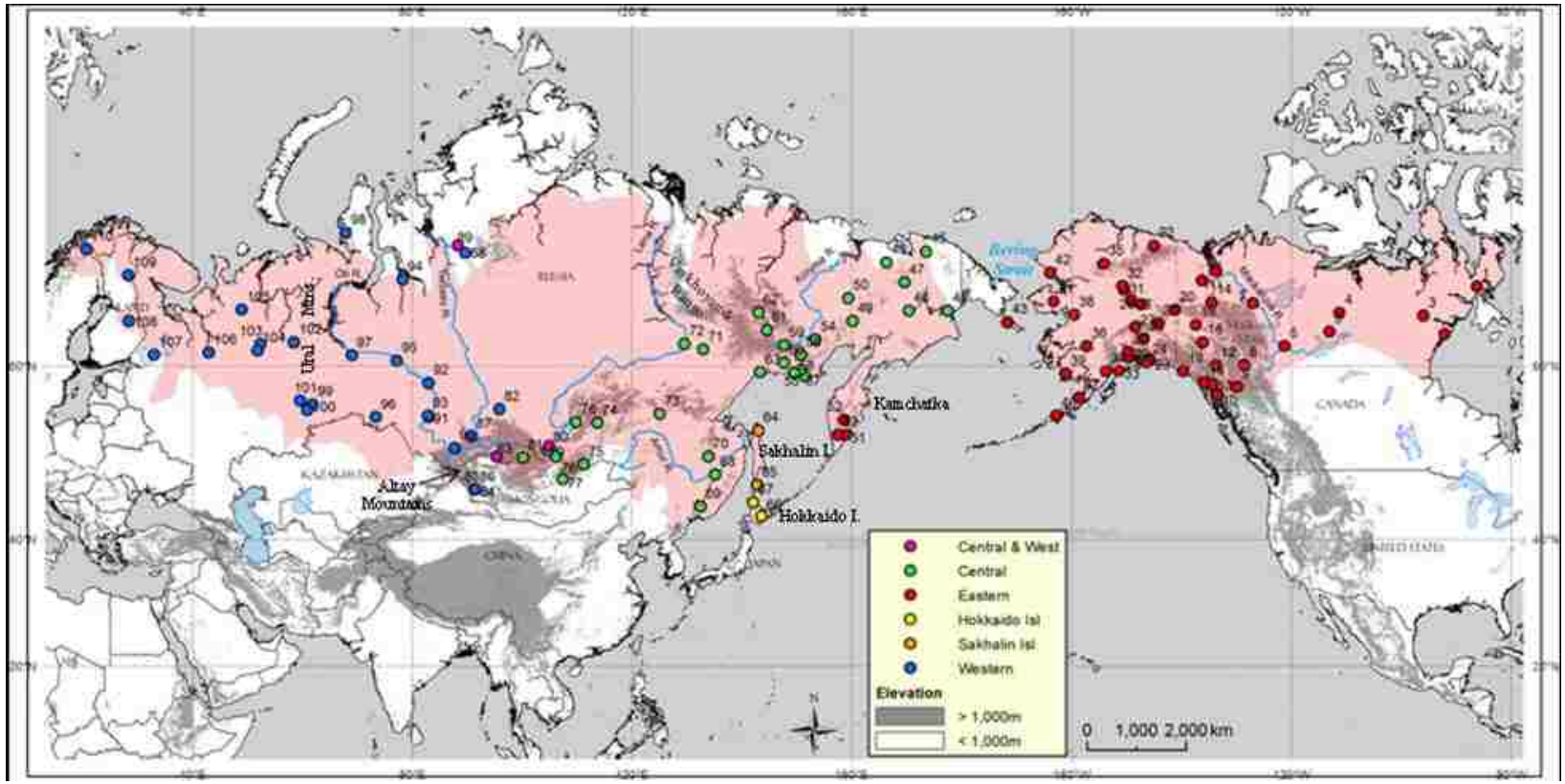


Figure 1. Distribution of *Clethrionomys rutilus* (pink, modified from IUCN and NatureServe) and sampling localities (colored dots). Colors correspond to cytochrome b clades as described in the legend. Populations containing both Central and Western haplotypes are uniquely identified in pink, these do not represent a separate clade. Localities are numbered sequentially from east to west and correspond to those in Table A of Appendix. Individuals from localities 3, 11, 29, 41, 53, 46, 48, 61, 65, 67, 68, 77, 86, 98, and 100 were used for species tree estimation.

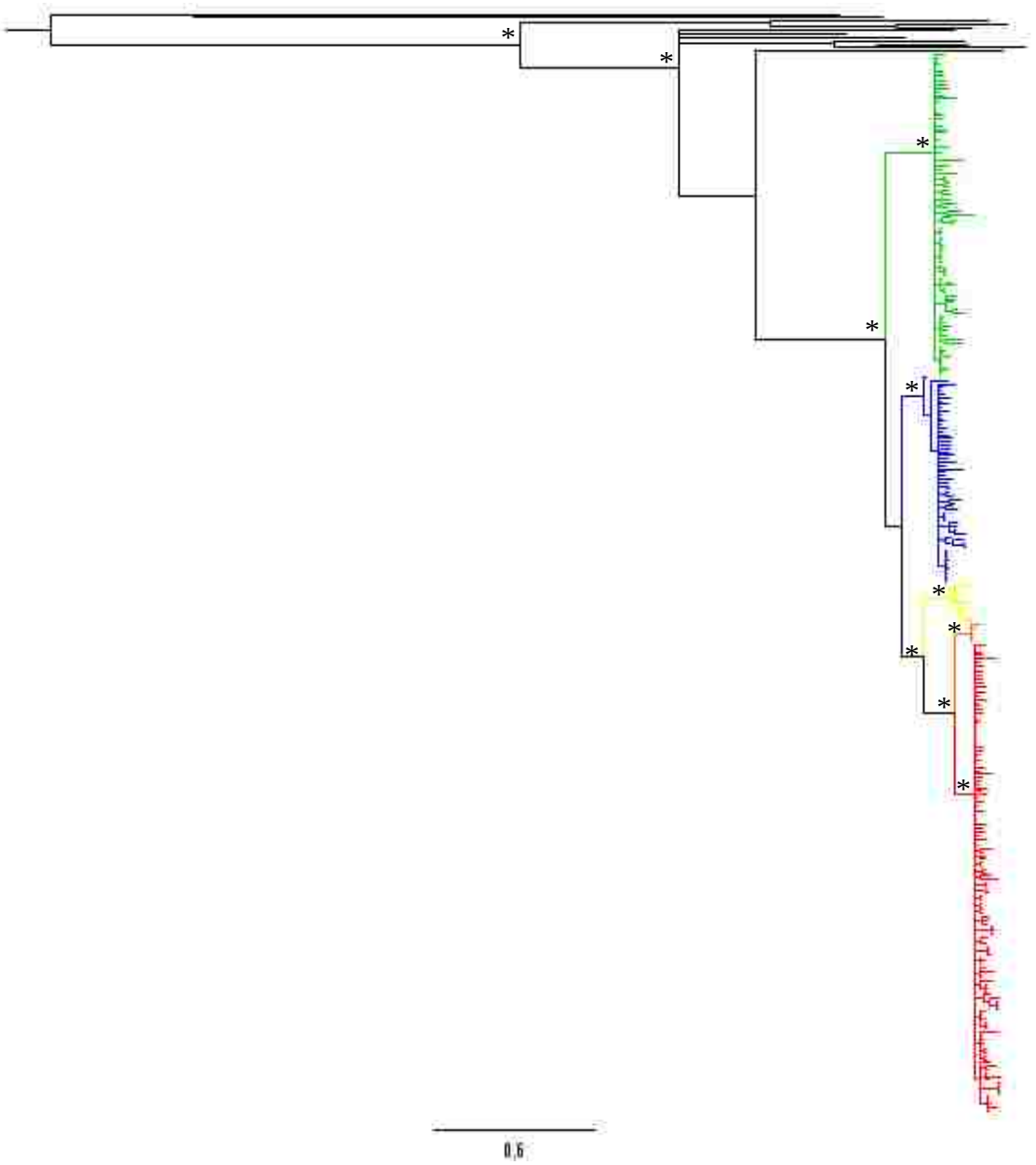


Figure 2. Bayesian Cyt b gene tree for 312 individuals of *C. rutilus* (761-1143bp). Colors correspond to Cyt b clade colors in Figure 1. One *Microtus*, one *Dicrostonyx*, and other *Clethrionomys* species are included as outgroups. An asterisk indicates posterior probability >0.95 on major nodes.

A. MLR B. ETS2 C. IRBP

Figure 3. Bayesian gene trees for three independent loci sequenced from *C. rutilus* individuals belonging to each mitochondrial clade, as indicated by branch color (see Figure 1 for legend). Black branches are outgroups. Posterior probability values >90% are reported at major nodes.

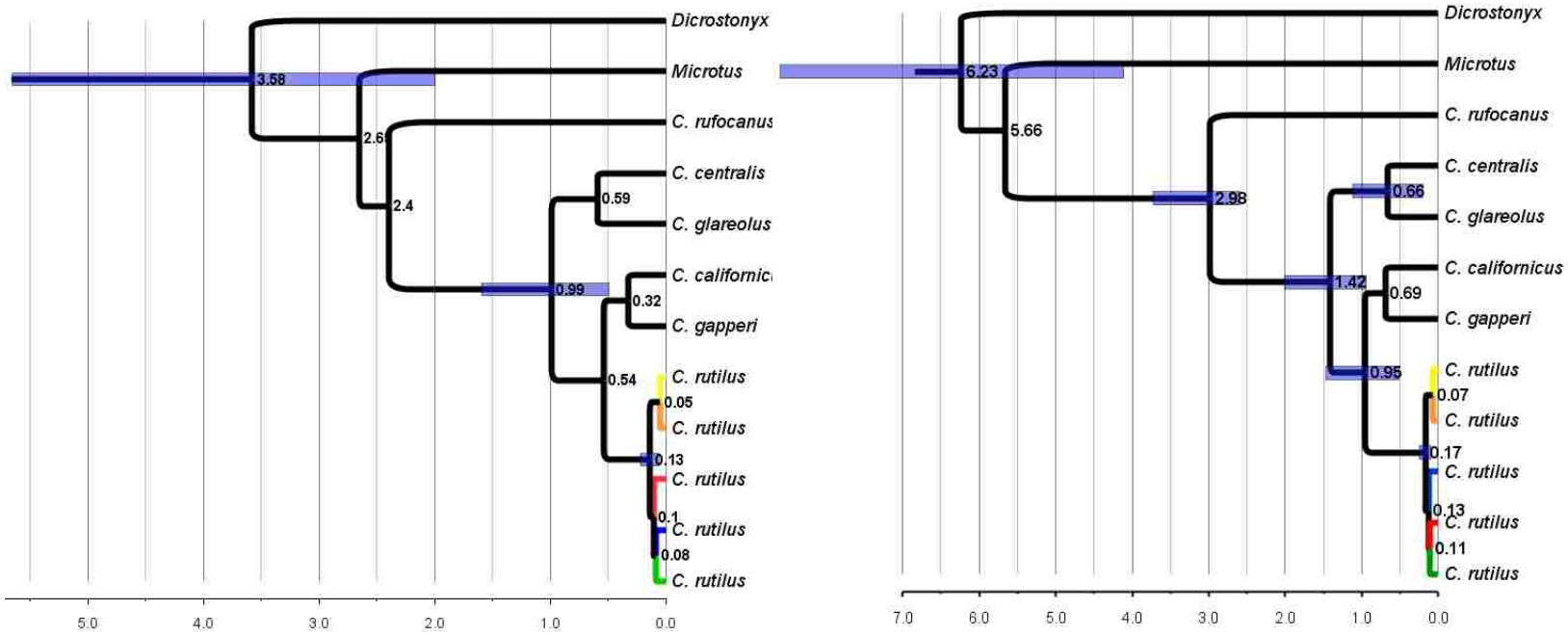


Figure 4. Species tree of *C. rutilus* clades and related outgroup species based on A) three nuclear genes and B) with nuclear and Cyt b subsets in the same analysis, constructed using *BEAST. Colors of branches correspond to *C. rutilus* clades in Figure 1. A time scale is included in B showing time from past to present (left to right) in millions of years. t_{MRCA} estimates are shown at all nodes. Blue error bars show the 95% confidence intervals around mean t_{MRCA} estimates for nodes with posterior support > 90%.

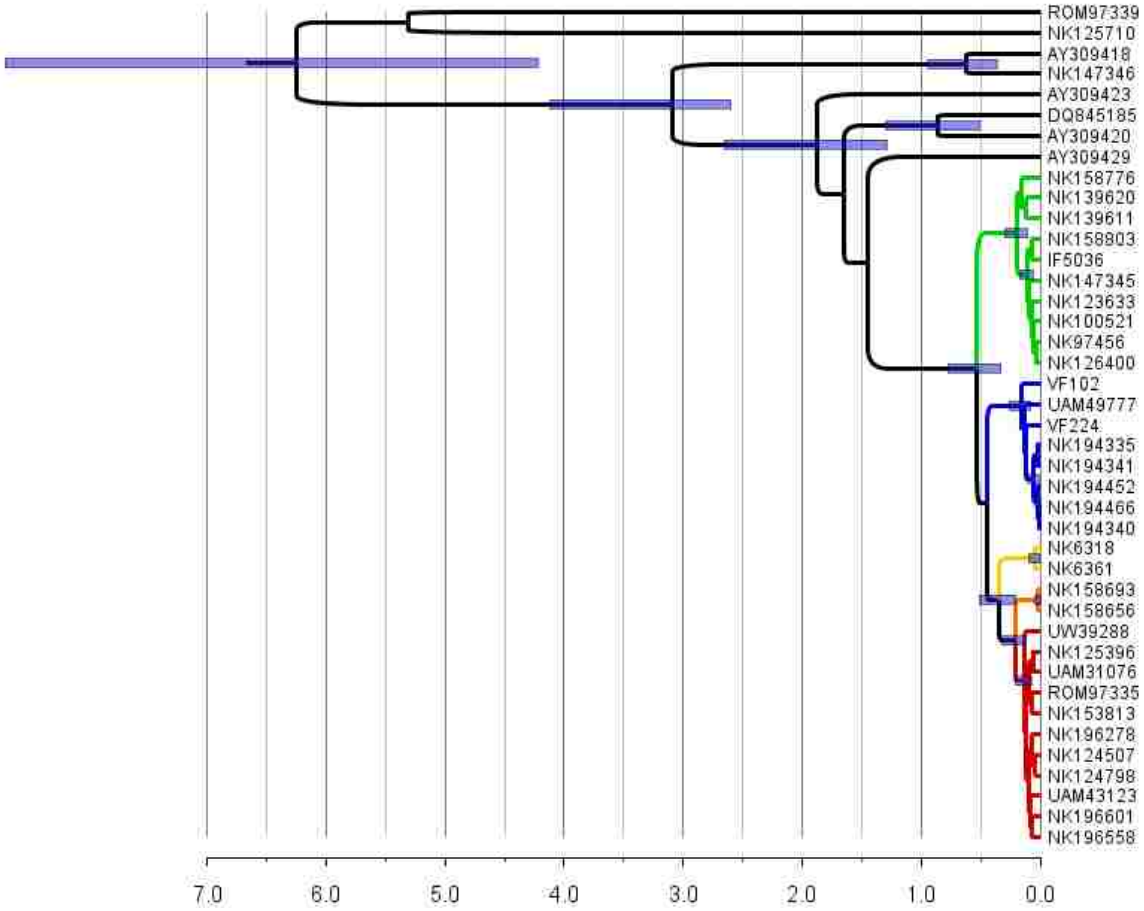


Figure 5. Cytochrome b gene tree for 33 individuals of *C. rutilus* constructed in BEAST to estimate divergence dates. Colors correspond to Cyt b clade colors in Figure 1. One *Microtus*, one *Dicrostonyx*, and other *Clethrionomys* species are included as outgroups. A time scale is included showing time from past to present (left to right) in millions of years. Blue error bars show the 95% confidence intervals around mean t_{MRCA} estimates for nodes with posterior support > 90%. t_{MRCA} estimates are shown at major nodes (in MY).

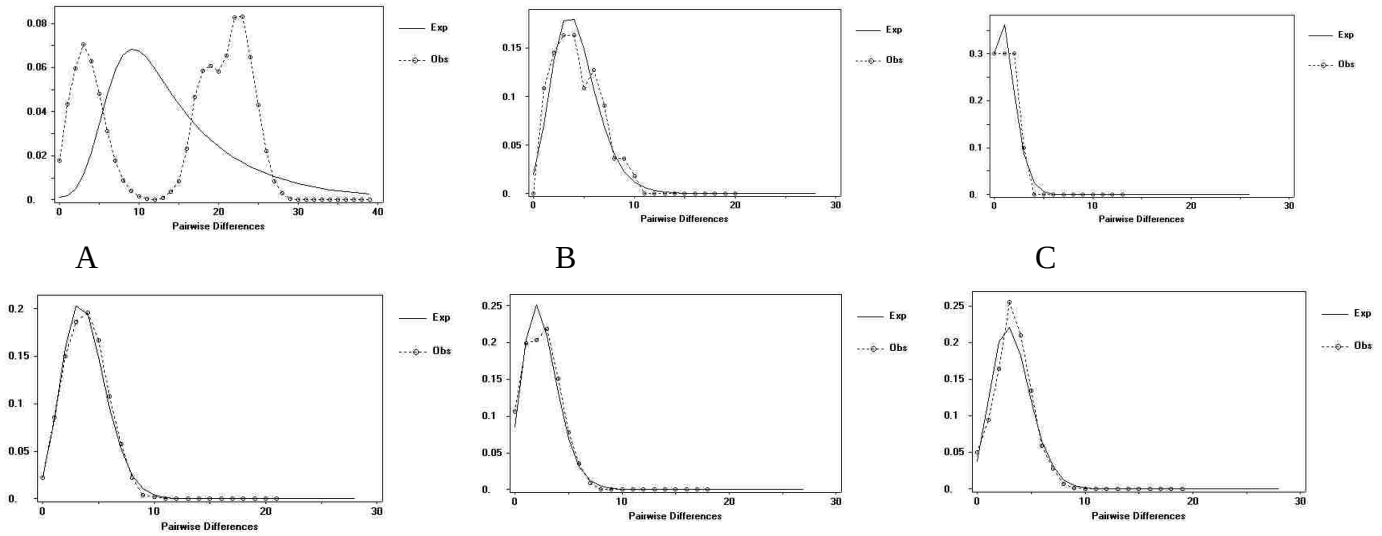


Figure 6. Mismatch distributions for Cyt b clades and the species as a whole, based on Cyt b sequence data (A) Entire species, B) Hokkaido Island, C) Sakhalin Island, D) Eastern, E) Central, F) Western). The solid lines represent the expected distribution under a model of sudden population expansion fitted to the data; the dashed lines represent frequency of pairwise nucleotide differences between sequences.

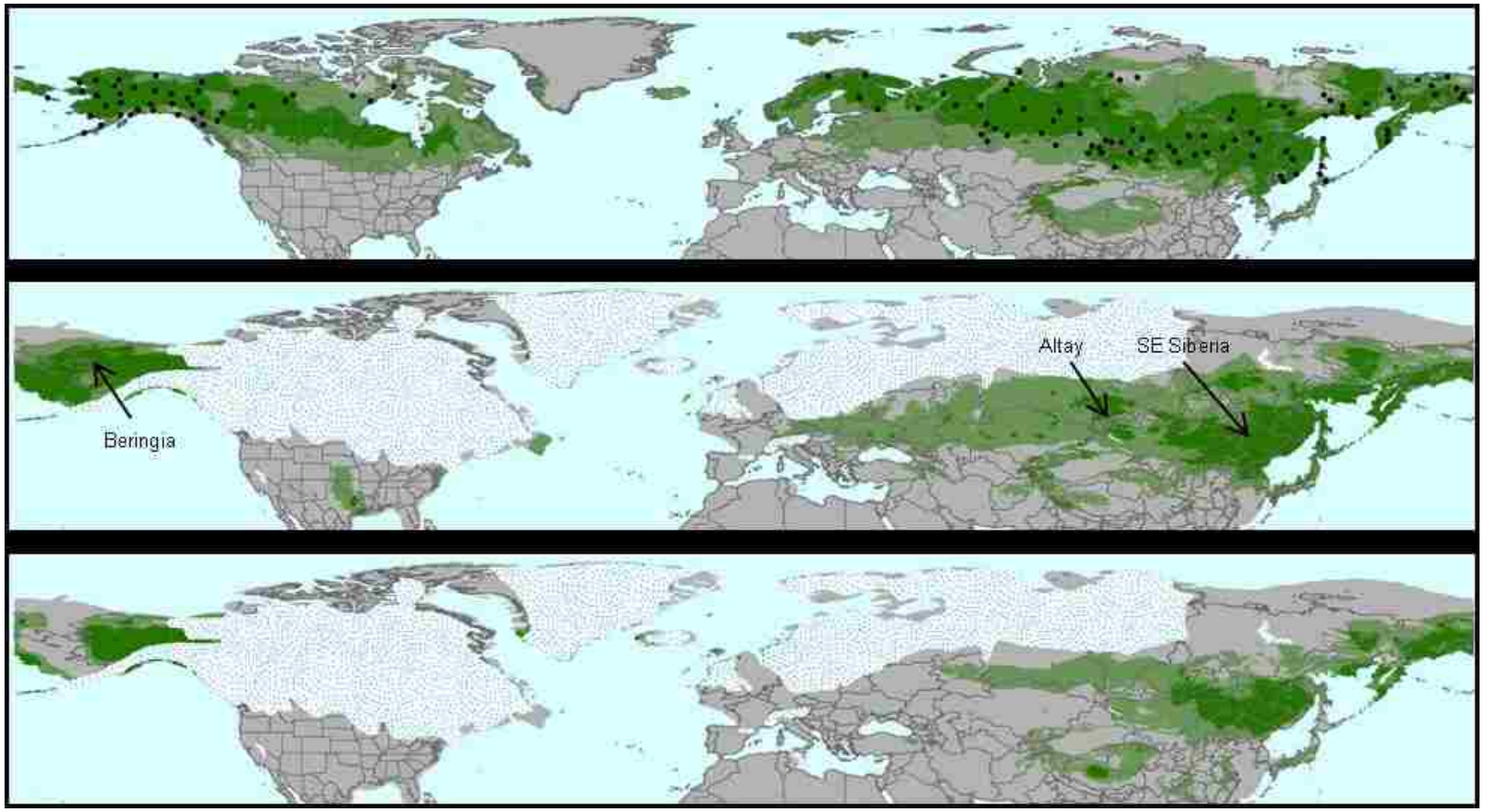


Figure 7. Species distribution models for *Clethrionomys rutilus* for present day (top panel) and for two Last Glacial Maximum scenarios based on MIROC and CCSM models (middle and bottom panels, respectively). Dots in the top panel represent sampling points used to generate the models. The shades of green represent suitable for *C. rutilus* based on a higher and a lower threshold. The stippled blue and white represents the glacial cover at LGM.

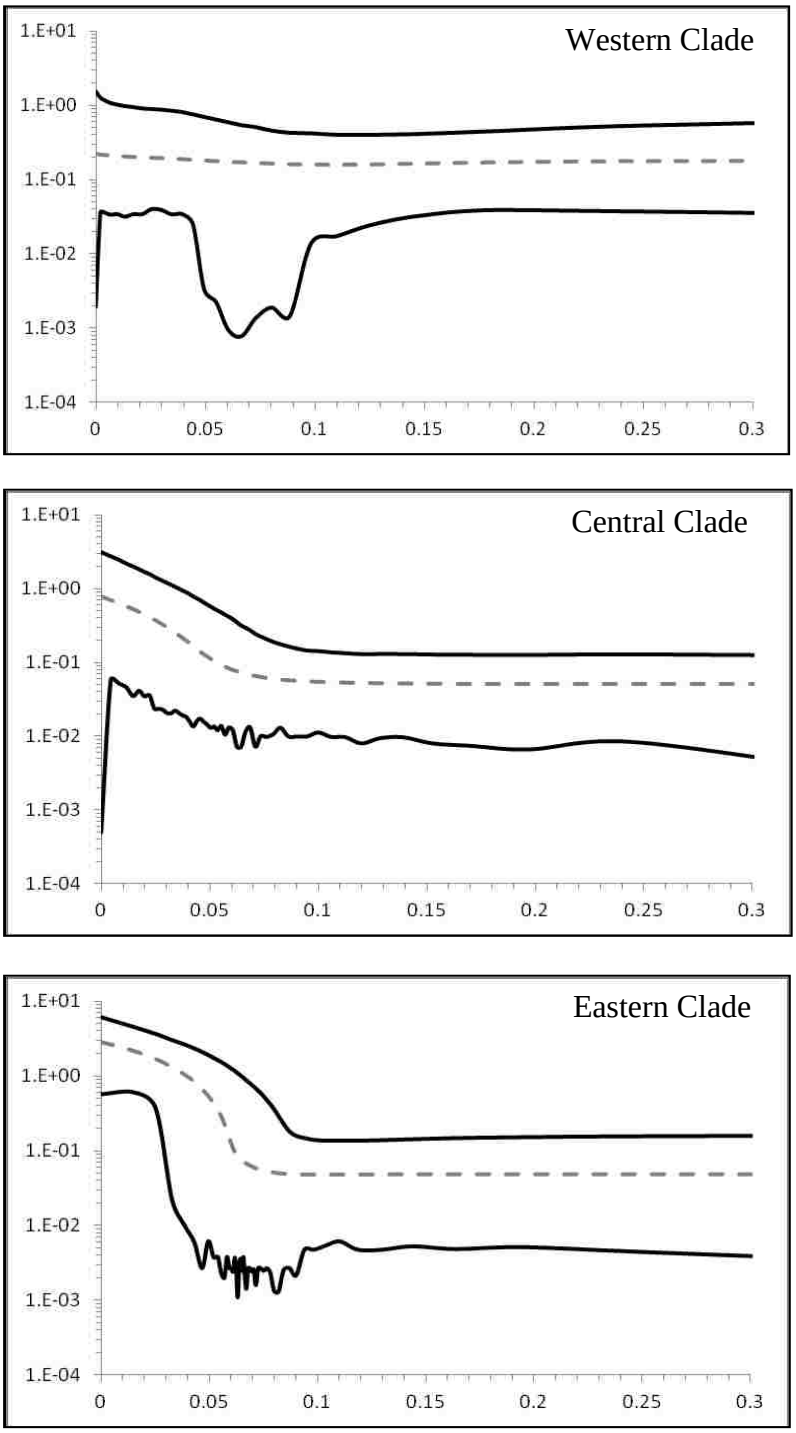


Figure 8. Extended Bayesian Skyline Plots for Western (top), Central (middle), and Eastern (bottom) Cyt b clades based on sequence data from four genes. On the x-axis is time in millions of years. On the y-axis is a log-transformed estimation of effective population size. The dashed line is the median estimated population size and solid lines are the 95% confidence interval.

Table 1. Gene sampling and information. Models of evolution were determined by ModelTest.

Gene	N (gene tree)	N (species tree)	bp	N (pop gen)	bp	Model of Evolution	# MCMC generations
ETS2	69	33	926	69	926	GTR+I+Γ	50 million
IRBP	57	33	1059	57	1059	GTR+I+Γ	50 million
MLR	76	33	205	76	205	HKY+I	12 million
Cyt b	312	33	761-1143	281	783	GTR+I+Γ	50
pos1						K80+I+ Γ	
pos2						HKY+I	
pos3						GTR+I+Γ	

Table 2. Population genetic summary statistics based on 783 bp of Cytochrome b and each nuclear gene investigated for *Clethrionomys rutilus*. Sample size (N), number of segregating sites (S), number of haplotypes (H), haplotype diversity (Hd), average number of nucleotide differences (k), nucleotide diversity (π), and three estimators of recent population expansion, Fu's F_S and Ramos-Onsins and Rozas' R_2 were analyzed. Significance of expansion statistics is indicated by asterisks: * = $p < 0.05$, **= $p < 0.0001$.

Gene/Group	**N**	**S**	**H**	**Hd**	**k**	$\boldsymbol{\pi}$	$\boldsymbol{F_S}$	$\boldsymbol{R_2}$
ETS2	69	54	50	0.870	7.59	0.0083	-26.437**	0.0655
IRBP	57	44	22	0.666	4.80	0.0047	-4.397	0.0525
MLR	76	4	5	0.536	0.59	0.0029	-1.046	0.0761
Cytochrome b	281	138	169	0.982	14.66	0.0202	-175.64**	0.0518
Cyt b clade/subclade								
Western	43	45	29	0.950	3.30	0.0043	-27.701**	0.0291**
Central	87	65	51	0.886	2.53	0.0031	-74.812**	0.0166**
Eastern (Beringian)	135	72	80	0.978	3.84	0.0050	-119.83**	0.0246**
Sakhalin	5	3	3	0.700	1.20	0.0015	-0.1858	0.2667
Hokkaido	11	18	11	1.000	4.29	0.0055	-7.781**	0.0985*

Table 3. Uncorrected average pairwise sequence divergence for each gene studied among Cyt b clades of *Clethrionomys rutilus* (bold, below diagonal). Standard error is above the diagonal. Average within clade divergence on diagonal (italicized). 1 = Sakhalin Island, 2 = Hokkaido Island, 3 = Eastern clade (Beringian subclade), 4 = Western clade, 5 = Central clade.

A. Cytochrome b

Cytb Clade	**1**	**2**	**3**	**4**	**5**
1	*0.001*	0.004	0.003	0.005	0.006
2	**0.019**	*0.005*	0.005	0.005	0.006
3	**0.012**	**0.024**	*0.005*	0.005	0.006
4	**0.024**	**0.021**	**0.026**	*0.004*	0.005
5	**0.033**	**0.031**	**0.033**	**0.025**	*0.004*

B. ETS2

Cytb Clade	1	2	3	4	5
1	*0.010*	0.002	0.002	0.002	0.002
2	**0.009**	*0.012*	0.002	0.003	0.003
3	**0.010**	**0.012**	*0.008*	0.002	0.002
4	**0.008**	**0.013**	**0.008**	*0.003*	0.001
5	**0.010**	**0.013**	**0.008**	**0.007**	*0.008*

C. IRBP

Cytb Clade	1	2	3	4	5
1	*.009*	0.002	0.002	0.002	0.002
2	**0.008**	*.010*	0.002	0.002	0.002
3	**0.007**	**0.010**	*.001*	0.003	0.001
4	**0.012**	**0.011**	**0.016**	*.007*	0.002
5	**0.008**	**0.010**	**0.005**	**0.013**	*.007*

D. MLR

Cytb Clade	1	2	3	4	5
1	*0.0*	0.002	0.005	0.0	0.002
2	**0.002**	*0.003*	0.005	0.002	0.003
3	**0.005**	**0.008**	*0.001*	0.005	0.002
4	**0.000**	**0.002**	**0.005**	*0.0*	0.002
5	**0.002**	**0.004**	**0.005**	**0.002**	*0.003*

Table 4. AMOVA results of variation partitioning in *Clethrionomys rutilus*. Populations were pooled into groups based on assignment to reciprocally monophyletic Cyt b groups. Significance denoted by asterisk ($P < 0.05$).

	Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Cyt b	Among groups	4	1806.922	10.38402*	84.37
	Among populations within groups	65	241.476	0.67246*	5.46
	Within populations	197	246.582	1.25169*	10.17
	Total	266	2294.98	12.30817	
ETS2	Among groups	4	37.334	0.33615*	8.58
	Among populations within groups	33	119.694	0.02455	0.63
	Within populations	70	249.083	3.55833	90.80
	Total	107	406.111	3.91904	
IRBP	Among groups	4	138.051	2.20437*	47.42
	Among populations within groups	32	94.901	0.51471*	11.07
	Within populations	44	84.900	1.92955*	41.51
	Total	80	317.852	4.64862	
MLR	Among groups	4	13.183	0.21427*	57.68
	Among populations within groups	34	4.672	-0.01576	-4.24
	Within populations	53	9.167	0.17296*	46.56
	Total	91	27.022	0.37146	

Appendices

Appendix A. Sampling list, localities, mtDNA clade assignment, and number of loci the individual's sequences were used for in analyses. Outgroup specimens are included at the bottom of the list. Abbreviations for collections are as follows: MSB, Museum of Southwestern Biology; UWBM, University of Washington Burke Museum; VF, Vadim Fedorov personal collection; ROM, Royal Ontario Museum; DGR, Museum of Southwestern Biology, Division of Genomic Resources.

Locality number	Sample ID	Collection	Catalog number	Cyt b Genbank Accession Number	Cyt b clade	Country	Latitude	Longitude	# of loci sequenced
1	ROM97334	ROM	97334		E	Canada	66.520	-86.230	1
2	ROM97348	ROM	97348		E	Canada	62.810	-92.100	1
3	ROM97335	ROM	97335		E	Canada	64.310	-96.010	4
3	ROM97336	ROM	97336		E	Canada	64.310	-96.010	4
3	ROM97337	ROM	97337		E	Canada	64.310	-96.010	1
3	ROM97338	ROM	97338		E	Canada	64.310	-96.010	1
4	UAM77202	UAM	77202		E	USA	64.520	-111.250	1
4	UAM77206	UAM	77206		E	USA	64.520	-111.250	1
5	UAM77250	UAM	77250		E	USA	63.000	-113.110	1
5	UAM77251	UAM	77251		E	USA	63.000	-113.110	1
5	UAM77257	UAM	77257		E	USA	63.000	-113.110	1
6	UAM85272	UAM	85272		E	USA	61.770	-121.290	1
6	UAM85286	UAM	85286		E	USA	61.770	-121.290	1
7	NK130861	DGR	10002150		E	Canada	65.184	-125.393	3
7	UAM83433	UAM	83433		E	USA	65.283	-127.833	1
7	UAM83434	UAM	83434		E	USA	65.283	-127.833	1
8	NK124785	MSB	144341		E	Canada	60.064	-128.581	1
8	NK124798	MSB	144354		E	Canada	60.107	-128.610	4
8	NK124804	MSB	144360		E	Canada	60.063	-128.620	1
8	NK181027	MSB	194002		E	Canada	60.099	-128.602	1
9	NK148457	MSB	158032		E	Canada	58.045	-129.955	4
9	NK148463	MSB	158037		E	Canada	58.045	-129.931	1
9	NK148464	MSB	158039		E	Canada	58.045	-129.931	1

9	NK148711	MSB	198839		E	Canada	58.048	-129.949	3
10	UAM44582	UAM	44582		E	USA	57.216	-133.501	1
10	UAM44591	UAM	44591		E	USA	57.216	-133.501	1
10	UAM44592	UAM	44592		E	USA	57.216	-133.501	1
11	NK153813	MSB	155943		E	Canada	67.610	-133.660	4
11	NK153814	MSB	155944		E	Canada	67.614	-133.661	1
11	NK153844	MSB	155964		E	Canada	67.614	-133.661	3
11	NK153845	MSB	155973		E	Canada	67.614	-133.661	1
12	NK181447	MSB	220896		E	Canada	60.067	-133.780	1
12	NK181454	MSB	220907		E	Canada	60.067	-133.780	1
12	NK181455	MSB	221122		E	Canada	60.067	-133.783	1
12	NK181463	MSB	221095		E	Canada	60.067	-133.780	1
14	NK181050	MSB	194075		E	Canada	65.337	-134.522	1
14	NK181054	MSB	194095		E	Canada	65.331	-134.519	1
14	NK181055	MSB	194134		E	Canada	65.331	-134.519	1
14	NK181102	MSB	194069		E	Canada	65.337	-134.522	1
15	UAM34253	UAM	34253		E	USA	58.450	-135.883	1
15	UAM43288	UAM	43288	AF272631	E	USA	58.450	-135.883	4
15	UAM43289	UAM	43289		E	USA	58.450	-135.883	1
16	NK125696	MSB	149275		E	Canada	62.116	-136.144	3
16	NK125697	MSB	149276		E	Canada	62.116	-136.144	1
16	NK125708	MSB	149289		E	Canada	62.116	-136.144	4
17	NK125152	MSB	145671	JF430949	E	Canada	66.949	-136.235	4
17	NK125231	MSB	145660	JF430950	E	Canada	66.920	-136.340	4
18	NK125002	MSB	145284		E	Canada	63.555	-137.412	1
18	NK125025	MSB	145315		E	Canada	63.555	-137.412	1
19	UAM31068	UAM	31068		E	USA	59.513	-139.679	1
19	UAM31070	UAM	31070		E	USA	59.494	-139.729	3
19	UAM31076	UAM	31068		E	USA	59.494	-139.729	4
20	NK124584	MSB	144248		E	USA	64.753	-141.230	1
20	NK124598	MSB	144001		E	USA	64.753	-141.230	1

20	NK124599	MSB	144002		E	USA	64.753	-141.230	1
22	UAM85807	UAM	85807		E	USA	69.300	-145.022	1
22	UAM85816	UAM	85816		E	USA	69.311	-145.025	1
22	UAM85820	UAM	85820		E	USA	69.311	-145.025	1
23	UAM49194	UAM	49194		E	USA	60.543	-145.758	1
24	UAM53191	UAM	53191		E	USA	60.348	-146.294	1
24	UAM53192	UAM	53192		E	USA	60.348	-146.294	1
25	NK152248	MSB	193516		E	USA	62.099	-146.362	1
25	NK152251	MSB	193515		E	USA	62.099	-146.362	1
25	NK152253	MSB	193517		E	USA	62.099	-146.362	1
25	NK152254	MSB	193518		E	USA	62.099	-146.362	1
26	NK124507	MSB	144060	JF430941	E	USA	65.149	-147.387	4
26	NK124508	MSB	144061	JF430942	E	USA	65.149	-147.387	4
29	NK125396	MSB	145458		E	USA	60.831	-149.536	4
29	NK125400	MSB	145475		E	USA	60.831	-149.536	3
29	NK125443	MSB	145425		E	USA	60.831	-149.536	1
30	NK80824	MSB	90821		E	USA	61.254	-149.688	4
30	NK80841	MSB	90822		E	USA	61.254	-149.688	3
30	NK80843	MSB	90388		E	USA	61.254	-149.688	1
30	NK80856	MSB	90825		E	USA	61.254	-149.688	1
31	NK123253	MSB	143285		E	USA	66.300	-150.430	1
33	UAM86132	UAM	86132		E	USA	59.551	-151.349	1
33	UAM86134	UAM	86134		E	USA	59.551	-151.349	1
33	UAM86135	UAM	86135		E	USA	59.551	-151.349	1
34	UAM64288	UAM	64288		E	USA	59.524	-153.773	1
34	UAM64302	UAM	64302		E	USA	59.524	-153.773	1
34	UAM64303	UAM	64303		E	USA	59.524	-153.773	1
35	NK196259	MSB	221840		E	USA	68.117	-154.124	1
35	NK196278	MSB	221742		E	USA	68.117	-154.124	4
35	NK196297	MSB	221609		E	USA	68.117	-154.124	1
35	NK196308	MSB	221583		E	USA	68.117	-154.124	4

36	UAM72707	UAM	72707	E	USA	61.773	-157.326	1
36	UAM72725	UAM	72725	E	USA	61.777	-157.333	1
36	UAM72741	UAM	72741	E	USA	61.761	-157.297	1
37	UAM64466	UAM	64466	E	USA	56.993	-158.644	1
37	UAM64467	UAM	64467	E	USA	56.993	-158.644	1
37	UAM64468	UAM	64468	E	USA	56.993	-158.644	1
38	UAM98368	UAM	98368	E	USA	64.372	-159.557	1
38	UAM98419	UAM	98419	E	USA	64.374	-159.551	1
38	UAM98471	UAM	98471	E	USA	64.374	-159.551	1
39	UAM47533	UAM	47533	E	USA	59.300	-161.115	1
39	UAM47538	UAM	47538	E	USA	59.301	-161.114	1
40	UAM36599	UAM	36599	E	USA	55.150	-162.813	4
40	UAM43122	UAM	43122	E	USA	55.150	-162.813	4
40	UAM43123	UAM	43123	E	USA	55.150	-162.813	4
41	NK196577	MSB	222045	E	USA	65.378	-163.228	1
41	NK196601	MSB	223027	E	USA	65.378	-163.228	4
41	NK196630	MSB	223126	E	USA	65.378	-163.228	3
41	NK196647	MSB	223097	E	USA	65.378	-163.228	1
42	NK196548	MSB	222012	E	USA	67.490	-163.896	1
42	NK196557	MSB	222031	E	USA	67.490	-163.896	4
42	NK196558	MSB	222068	E	USA	67.490	-163.896	4
43	MSB221177	MSB	221177	E	USA	63.775	-171.693	2
43	MSB221208	MSB	221208	E	USA	63.775	-171.693	3
43	MSB221239	MSB	221239	E	USA	63.775	-171.693	1
44	IF5122	UAM	84258	C	Russia	64.574	177.324	1
44	IF5123	UAM	84259	C	Russia	64.574	177.324	3
44	IF5175	UAM	84318	C	Russia	64.806	177.554	4
44	IF5254	UAM	84370	C	Russia	64.806	177.554	1
45	NK158474	MSB	150020	C	Russia	68.933	173.533	4
45	NK158482	MSB	150034	C	Russia	68.933	173.533	1
45	NK158484	MSB	150059	C	Russia	68.933	173.533	3

45	NK158485	MSB	150036		C	Russia	68.933	173.533	1
46	IF5035	UAM	84150		C	Russia	64.678	170.403	4
46	IF5036	UAM	84151		C	Russia	64.678	170.403	4
46	IF5037	UAM	84152		C	Russia	64.678	170.403	1
46	IF5038	UAM	84153		C	Russia	64.678	170.403	1
47	NK158316	MSB	149907		C	Russia	66.767	169.567	4
47	NK158800	MSB	193174		C	Russia	66.817	169.550	1
47	NK158801	MSB	193175		C	Russia	66.817	169.550	1
47	NK158802	MSB	193176		C	Russia	66.817	169.550	1
48	NK158803	MSB	193177		C	Russia	68.200	166.217	4
48	NK158807	MSB	193187		C	Russia	68.200	166.217	1
48	NK158808	MSB	193188		C	Russia	68.200	166.217	1
48	NK158809	MSB	193189		C	Russia	68.200	166.217	3
49	NK126469	MSB	148077		C	Russia	63.867	160.100	1
49	NK126470	MSB	148078		C	Russia	63.867	160.100	1
49	NK126471	MSB	148079	JF430961	C	Russia	63.867	160.100	3
49	NK126474	MSB	148161		C	Russia	63.867	160.100	1
49	NK126475	MSB	148165		C	Russia	63.867	160.100	4
50	UAM80568	UAM	80568		C	Russia	65.648	159.384	1
50	UAM80755	UAM	80755		C	Russia	65.648	159.384	1
50	UAM80756	UAM	80756		C	Russia	65.648	159.384	1
53	UW39277	UWBM	39277		E	Russia	53.050	157.460	1
53	UW39286	UWBM	39286		E	Russia	53.050	157.460	1
53	UW39288	UWBM	39288		E	Russia	53.050	157.460	4
53	UW39300	UWBM	39300		E	Russia	53.050	157.460	4
54	NK126068	MSB	147903		C	Russia	62.333	153.350	1
54	NK126095	MSB	147904		C	Russia	62.333	153.350	4
54	NK126211	MSB	148174	JF430960	C	Russia	62.333	153.350	3
59	NK123304	MSB	144873		C	Russia	61.847	147.662	3
59	NK123311	MSB	144862	JF430956	C	Russia	61.847	147.662	3
60	NK158274	MSB	149871		C	Russia	60.280	147.620	1

60	NK158280	MSB	149878		C	Russia	60.280	147.620	1
61	NK123599	MSB	144703		C	Russia	63.123	144.518	1
61	NK123600	MSB	144704		C	Russia	63.123	144.518	3
61	NK123633	MSB	144733		C	Russia	63.123	144.518	4
61	NK123634	MSB	144734		C	Russia	63.123	144.518	1
62	NK123420	MSB	144976		C	Russia	64.545	143.000	4
62	NK123431	MSB	144986	JF430957	C	Russia	64.545	143.000	4
62	NK123512	MSB	144618		C	Russia	64.545	143.000	1
63	NK126398	MSB	148084		C	Russia	59.433	143.300	1
63	NK126399	MSB	148085		C	Russia	59.433	143.300	1
63	NK126400	MSB	148086		C	Russia	59.433	143.300	4
63	NK126401	MSB	148087		C	Russia	59.433	143.300	4
65	NK158656	MSB	193022		SI	Russia	47.233	142.783	4
65	NK158692	MSB	193007		SI	Russia	47.233	142.783	1
65	NK158693	MSB	193008		SI	Russia	47.233	142.783	4
65	NK158694	MSB	192945		SI	Russia	47.233	142.783	1
67	NK6318	MSB	45580		HI	Japan	44.995	142.024	4
67	NK6325	MSB	45578		HI	Japan	44.995	142.024	1
67	NK6361	MSB	45628		HI	Japan	44.988	142.007	4
67	NK6380	MSB	45627		HI	Japan	44.988	142.007	1
68	NK147345	MSB	153615		C	Russia	48.300	135.050	4
68	NK158776	MSB	193200		C	Russia	48.467	134.933	4
69	UW77301	UWBM	77301		C	Russia	44.450	132.470	1
70	VF30	UAM	AF51446		C	Russia	50.631	133.824	1
72	NK139611	MSB	148506	JF430968	C	Russia	62.070	128.938	4
72	NK139620	MSB	148515	JF430969	C	Russia	62.070	128.938	4
73	VF113	UAM	AF51412		C	Russia	55.291	125.031	1
73	VF117	UAM	AF51464		C	Russia	55.291	125.031	1
74	VF17	UAM	AF51433		C	Russia	54.363	113.640	1
74	VF19	UAM	AF51434		C	Russia	54.363	113.640	1
75	VF40	UAM	AF51404		C	Russia	49.697	111.197	1

75	VF41	UAM	AF51439	C	Russia	49.697	111.197	1
75	VF42	UAM	AF52319	C	Russia	49.697	111.197	1
76	VF24	UAM	AF51432	C	Russia	54.427	109.824	1
77	NK100508	MSB	94231	C	Mongolia	47.883	107.383	1
77	NK100521	MSB	94232	C	Mongolia	47.883	107.383	4
77	NK100526	MSB	94233	C	Mongolia	47.883	107.383	1
77	NK100550	MSB	94234	C	Mongolia	47.883	107.383	4
78	VF7	UAM	AF52308	C	Russia	50.652	106.089	1
79	VF68	UAM	AF52305	W	Russia	51.852	104.905	1
79	VF69	UAM	AF51465	C	Russia	51.852	104.905	1
80	VF12	UAM	AF51405	C	Russia	51.412	104.631	1
81	NK97456	MSB	98514	C	Mongolia	50.509	100.155	4
82	VF75	UAM	AF52306	W	Russia	55.818	95.842	1
83	VF43	UAM	AF51450	W	Russia	50.652	95.346	1
83	VF46	UAM	AF51410	C	Russia	50.652	95.346	1
83	VF47	UAM	AF51435	W	Russia	50.652	95.346	1
84	NK194304	MSB	233757	W	Mongolia	46.658	91.433	2
84	NK194306	MSB	233669	W	Mongolia	46.658	91.433	1
84	NK194317	MSB	233820	W	Mongolia	46.658	91.433	1
84	NK194452	MSB	233788	W	Mongolia	46.658	91.433	4
85	NK194335	MSB	233668	W	Mongolia	46.676	91.403	4
85	NK194466	MSB	233685	W	Mongolia	46.676	91.403	4
86	NK194340	MSB	233726	W	Mongolia	46.664	91.400	4
86	NK194341	MSB	233704	W	Mongolia	46.664	91.400	4
86	NK194345	MSB	233727	W	Mongolia	46.664	91.400	1
86	NK194375	MSB	233705	W	Mongolia	46.664	91.400	1
87	VF78	UAM	AF51437	W	Russia	52.889	90.736	1
88	VF79	UAM	AF51473	W	Russia	68.827	89.624	1
88	VF80	UAM	AF51489	W	Russia	68.827	89.624	1
88	VF81	UAM	AF51482	W	Russia	68.827	89.624	1
89	VF84	UAM	AF52315	C	Russia	69.340	88.265	1

89	VF85	UAM	AF51406	C	Russia	69.340	88.265	1
89	VF87	UAM	AF52301	W	Russia	69.340	88.265	1
90	VF105	UAM	AF51486	W	Russia	51.631	87.632	1
92	VF59	UAM	AF51436	W	Russia	58.417	82.929	1
93	VF100	UAM	AF51481	W	Russia	55.114	82.843	1
93	VF98	UAM	AF51440	W	Russia	55.114	82.843	1
94	VF51	UAM	AF51449	W	Russia	67.002	78.213	1
94	VF52	UAM	AF51442	W	Russia	67.002	78.213	1
97	VF57	UAM	AF51484	W	Russia	60.954	68.997	1
98	UAM49776	UAM	49776	W	Russia	70.133	67.867	4
98	UAM49777	UAM	49777	W	Russia	70.133	67.867	3
99	VF102	Vadim Fedorov	102	W	Russia	56.349	61.657	4
99	VF103	Vadim Fedorov	103	W	Russia	56.349	61.657	1
99	VF239	Vadim Fedorov	239	W	Russia	56.269	61.966	1
99	VF265	Vadim Fedorov	265	W	Russia	56.269	61.966	1
100	VF191	Vadim Fedorov	191	W	Russia	55.740	60.870	1
100	VF192	Vadim Fedorov	192	W	Russia	55.740	60.870	1
100	VF223	Vadim Fedorov	223	W	Russia	55.740	60.870	1
100	VF224	Vadim Fedorov	224	W	Russia	55.740	60.870	4
110	UAM30003	UAM	30003	W	Finland	69.083	20.667	1
110	UAM30004	UAM	30004	W	Finland	69.083	20.667	1
110	UAM30005	UAM	30005	W	Finland	69.083	20.667	1

Outgroups

Genus	**Sample ID**	**Collection**	**Catalog Number**
C. californicus	UW78787	UWBM	78787
C. centralis	NK138064	MSB	158452
C. gapperi	NK62143	MSB	70688
C. gapperi	NK153189	MSB	193266
*C. gapperi**	NK132592	MSB	147281

*Exhibits *C. rutilus* cyt b, *C. gapperi* nucDNA

C. gapperi	UAM77807	UAM	77807
C. glareolus	NK139890	MSB	193928
C. rufocanus	NK147346	MSB	45581
C. rufocanus	NK166807	MSB	153616
Dicrostonyx groenlandicus	ROM97339	ROM	97339
Microtus sp.	MSB 149290	MSB	149290

Appendix B. GenBank sequences of Cyt b obtained for this study.

GenBank Accession Number	Locality Number	Country	Cyt b lineage
AF272632	13	United States	E
JF430951	17	Canada	E
JF430952	17	Canada	E
JF430947	18	Canada	E
JF430948	18	Canada	E
AY309427	19	United States	E
JF430946	20	United States	E
JF430945	21	Alaska	E
AY309426	23	United States	E
AF119274	25	United States	E
JF430943	26	United States	E
JF430944	26	United States	E
JF430934	27	Alaska	E
JF430935	28	Alaska	E
JF430936	28	Alaska	E
JF430954	28	Alaska	E
JF430955	28	Alaska	E
JF430953	29	United States	E
JF430938	31	United States	E
JF430939	31	United States	E
JF430940	31	United States	E
JF430937	32	Alaska	E
AB072220	51	Russia	E
AB072221	51	Russia	E
AB072222	52	Russia	E
JF430959	54	Russia	C
HQ608518	55	Russia	C
GQ301923	56	Russia	C

GQ301927	56	Russia	C
GQ301928	56	Russia	C
GQ301929	56	Russia	C
AB072215	57	Russia	C
AB072216	57	Russia	C
AB072217	57	Russia	C
GU251085	58	Russia	C
GU251087	58	Russia	C
HQ608514	58	Russia	C
AY309424	59	Russia	C
AY309425	59	Russia	C
JF430971	60	Russia	C
JF430972	60	Russia	C
JF430958	62	Russia	C
AB072224	64	Japan	SI
AB072223	65	Russia	SI
AB031581	66	Japan	HI
AB072207	66	Japan	HI
AB072208	66	Japan	HI
AB072209	66	Japan	HI
AB072210	66	Japan	HI
AB072211	66	Japan	HI
AB072212	66	Japan	HI
AB072213	66	Japan	HI
AB072214	68	Russia	C
JF430963	71	Russia	C
JF430964	71	Russia	C
JF430965	71	Russia	C
JF430962	72	Russia	C
JF430967	72	Russia	C
JF430970	72	Russia	C

AB072218	91	Russia	W
AB072219	91	Russia	W
EU035662	95	Russia	W
AF367077	96	Russia	W
EU035678	101	Russia	W
EU035679	101	Russia	W
EU232149	101	Russia	W
EU232138	102	Russia	W
EU232143	102	Russia	W
EU035676	103	Russia	W
EU035663	104	Russia	W
EU232145	104	Russia	W
EU035664	105	Russia	W
EU035677	105	Russia	W
EU035694	105	Russia	W
EU035695	105	Russia	W
EU035696	105	Russia	W
EU035661	106	Russia	W
EU232163	107	Russia	W
AF272638	108	Finland	W
JF930077	109	Finland	W
AY309428	110	Finland	W

Outgroups

Sample ID	Species	Gene
DQ845196	*Alticola macrotis*	Cyt b
AF119273	*A. lemminus*	Cyt b
AY309423	*C. californicus*	Cyt b
DQ845185	*C. centralis*	Cyt b
AY309429	*C. gapperi*	Cyt b

AY309420	*C. glareolus*	Cyt b
AB565456	*C. rex*	Cyt b
AY309418	*C. rufocanus*	Cyt b
JF906129	*Microtus fortis*	IRBP

Chapter 3

Multilocus systematics and the characteristics of a recent radiation in a Holarctic tribe of rodents (Arvicolinae: Clethrionomyini)

Abstract

The Clethrionomyini is presently composed of five genera (*Alticola*, *Caryomys*, *Clethrionomys*, *Eothenomys*, and *Hyperacrius*). This tribe of forest and alpine voles is distributed throughout the Holarctic and characterized by Late Cenozoic diversification related to repeated climatic fluctuations in the Northern Hemisphere. Phylogenetic studies using primarily mitochondrial DNA have identified distinct cases of paraphyly and polyphyly among genera that call for taxonomic revision. We apply the first multilocus approach to test phylogenetic relationships among 28 of 36 species of Clethrionomyini, including 3 (*Alticola montosa*, *A. albicaudus*, and *Hyperacrius fertilis*) not previously sequenced. Four clades are identified within the tribe and confirm that *Clethrionomys* is paraphyletic as currently delimited. Previous hypotheses regarding polyphyly of *Alticola* in regards to *A. lemminus* are not supported by multilocus nuclear analysis. *Hyperacrius* should not be considered within the clethrionomyine clade. A lineage through time plot suggests that diversification in the tribe over the last 1-5 MY did not occur in a series of pulses, as earlier proposed for Arvicolinae based on molecular and paleontological analyses.

Brooks Kohli[a], Kelly Speer[b*], C. William Kilpatrick[c*], Nyamsuren Batsaikhan[d*], Darmaa Damdinbaza[d*], Joseph A. Cook[a*]

[a] Department of Biology and Museum of Southwestern Biology, University of New Mexico, Albuquerque NM 87131–1051; [b] Florida Museum of Natural History, University of Florida, Gainesville, FL 32611; [c] Department of Biology, University of Vermont, Burlington, VT; [d] Department of Zoology, Faculty of Biology, National University of Mongolia, Ulaan Bataar, Mongolia

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

1. Introduction

The Arvicolinae (voles and lemmings) are geographically widespread rodents that inhabit grassland, forest, alpine, and tundra habitats throughout the Northern Hemisphere (Gromov and Polyakov 1977; Carleton and Musser 2005). With over 150 species, this subfamily represents one of the largest radiations of mammals (Carleton and Musser 2005) over the last 5 million years (MY) (Repenning, Fejfar et al. 1990). Their comparatively extensive fossil record has fostered detailed examinations of the subfamily's evolution and dispersal across high latitudes (Repenning, Fejfar et al. 1990; Chaline, Brunet-Lecomte et al. 1999). Paleontological (Repenning, Fejfar et al. 1990) and molecular (Conroy and Cook 1999) studies revealed punctuated bursts of dispersal and diversification correlated to environmental events such as glacial cycles, generating several clusters of closely related genera and species whose precise evolutionary relationships have proven difficult to elucidate through morphologic and molecular methods (Chaline and Graf 1988; Conroy and Cook 1999; Robovsky, Ricankova et al. 2008; Abramson, Lebedev et al. 2009).

Although the genus *Microtus* has received considerable attention, another group of arvicolines that is often cited as an example of the pulse pattern of evolution is the tribe Clethrionomyini (Conroy and Cook 1999; Buzan, Krystufek et al. 2008). Typically considered a well-supported clade within Arvicolinae, numerous unresolved relationships (i.e. polytomies) and high diversity within Clethrionomyini suggest rapid diversification (Conroy and Cook 1999). Fossil evidence indicates an Asian origin of the tribe and all its genera (Chaline and Graf 1988; Repenning, Fejfar et al. 1990; Zheng and Li 1990). The tribe's highest extant diversity is in central Asia, but members are found from southern China and the Tibetan Plateau to the Arctic Circle and across northern Europe and North America (Figure 1) (Carleton and Musser 2005). Due to this origin and distribution, diversity within Clethrionomyini often has been related to the dominant environmental change events affecting the Northern Hemisphere and Asia, such as glacial cycles (Repenning, Fejfar et al. 1990; Cook, Runck et al. 2004) and uplift of the Tibetan Plateau (Luo, Yang et al. 2004; Liu, Liu et al. 2012), over the past 3.5 MY.

Taxonomic limits of the five genera currently recognized within Clethrionomyini, *Alticola*, *Clethrionomys* (= *Myodes*), *Caryomys*, *Eothenomys*, and *Hyperacrius* (Carleton and Musser 2005), are uncertain in some cases due to sparse fossil records, convergent and parallel evolution of morphological characters, limited molecular sampling, and difficulty in obtaining samples (Gromov and Polyakov 1977; Chaline and Graf 1988; Lebedev, Bannikova et al. 2007). The vast majority of phylogenetic studies of Clethrionomyini have utilized mitochondrial DNA (mtDNA) (Cook, Runck et al. 2004; Luo, Yang et al. 2004; Lebedev, Bannikova et al. 2007; Liu, Liu et al. 2012), and when combined with other characters have revealed cases of apparent hybridization among

nominal species and paraphyly or polyphyly among genera (Cook, Runck et al. 2004; Deffontaine, Libois et al. 2005; Runck, Matocq et al. 2009). Despite efforts that greatly increased the taxonomic sampling of Clethrionomyini in molecular studies (Luo, Yang et al. 2004; Lebedev, Bannikova et al. 2007; Liu, Liu et al. 2012), no single study has included all five genera and lack of nuclear DNA sequences has precluded rigorous tests of phylogenetic relationships across the tribe.

1.1 Objectives

Evolutionary relationships among members of Clethrionomyini have been difficult to resolve by traditional methods, yet their widespread Holarctic distribution and recent diversification provides an excellent opportunity to explore the influence of Quaternary environmental changes on evolutionary and biogeographic processes across northern high latitude ecosystems. We begin to resolve taxonomic issues within Clethrionomyini and explore the mode and tempo of evolution for the tribe to assess whether it conforms to a model of punctuated diversification, as previously suggested. By including an unprecedented 28 species in a mtDNA phylogenetic analysis, and then applying a multilocus approach for 19 of tribe's 36 species, we address three primary questions about the taxonomy and evolution of the Clethrionomyini: 1) Does nominal higher-level taxonomy reflect evolutionary relationships or do multilocus analyses support previous mtDNA findings of paraphyly and polyphyly among genera? 2) Do molecular analyses delimit nominal species designations? 3) Is diversification within the tribe associated with episodic environmental events?

2. Methods

We generally follow the taxonomy of Carleton and Musser (2005), however, we retain the use of *Clethrionomys* instead of *Myodes* as the valid generic name (and extension to tribal name) for red-backed voles (Tesakov, Lebedev et al. 2010).

2.1 Brief description of Clethrionomyini

In this section, we briefly describe each genus within Clethrionomyini and the pertinent taxonomic issues. The taxonomic history of the tribe and its members has been reviewed in detail elsewhere (see Luo et al. (2004) and Liu et al. (2012) for *Caryomys* and *Eothenomys*; Lebedev et. al (2007) for *Alticola* and *Clethrionomys*; Phillips (1969) for *Hyperacrius*). *Clethrionomys* (red-backed voles) is the sole Clethrionomyini genus with a Holarctic distribution, with species found over large portions of northern Eurasia and North America (Carleton and Musser 2005). These forest-associated voles are often highly abundant and also inhabit tundra habitats in some regions. At least three species are known to hybridize and/or exhibit mitochondrial introgression. *C. gapperi* and *C. glareolus* interbreed with *C. rutilus* where their ranges meet in North America and Europe, respectively (Tegelström 1987; Deffontaine, Libois et al. 2005; Runck, Matocq et al. 2009). Further complicating the matter, *C. gapperi* may represent at least 2 species according to mtDNA relationships, with separate affiliations to *C. glareolus* and *C. rutilus* (Cook, Runck et al. 2004). Mitochondrial DNA and chromosomal evidence indicates *Clethrionomys* is paraphyletic, with *C. rufocanus* and closely related Korean (*C. regulus*) and Japanese (*C. smithii*, *C. andersoni*, *C. imaizumii*, *C. rex*) red-backed voles in the subgenus *Craseomys* (Gromov and Polyakov 1977), forming a clade basal to all other *Clethrionomys* and the genus *Alticola* (Lebedev, Bannikova et al. 2007). *C. andersoni* and *C. smithii* have previously been included in *Eothenomys*.

Species of *Alticola* (montane voles) are found mostly in central Asia, typically inhabiting steppe, rocky montane, and alpine habitats (Gromov and Polyakov 1977). The exception is *Alticola lemminus*, found in tundra habitats of northeastern Siberia. This species has the northernmost distribution of *Alticola* and is the only member of Clethrionomyini whose pelage turns white in winter (Gromov and Polyakov 1977). Based on previous mtDNA studies, *A. lemminus* and *A. macrotis* are more closely related to *Clethrionomys* than other montane voles, suggesting the need to revise higher taxonomy to eliminate polyphyly of *Alticola* (Cook, Runck et al. 2004; Lebedev, Bannikova et al. 2007).

Hyperacrius is hypothesized to be a derived descendent of *Alticola* that adapted to a semi-fossorial lifestyle (Phillips 1969), but recent morphological analysis suggests similarity to *Prometheomys*, a basal arvicoline (Robovsky, Ricankova et al. 2008). The two species of *Hyperacrius* are known only from high altitude forests and alpine meadows in northern India and northern Pakistan (Phillips 1969).

Diversity of the forest-dwelling genus *Eothenomys* is centered near the southeastern portion of the Tibetan Plateau in western China. All species have relatively limited distributions except for *E. melanogaster*, which extends across central China to Taiwan. Considerable debate has surrounded the taxonomy of *Eothenomys*, yet morphologic and molecular investigations have attempted to clarify these issues (Kaneko 1996; Luo, Yang et al. 2004; Liu, Liu et al. 2012). The most comprehensive mtDNA study concluded that *Eothenomys* should be split into several genera (Liu, Liu et al. 2012).

Caryomys is comprised of two species endemic to China and has often been included as a subgenus within *Eothenomys* (Carleton and Musser 2005). Detailed morphologic work distinguished the two extant species from each other, and from *Clethrionomys rufocanus* and *Eothenomys* (Kaneko 1992). Recent analysis combining mtDNA and morphologic data support monophyly of *Caryomys*, independent of *Eothenomys* (Liu, Liu et al. 2012).

2.2 Sampling and laboratory techniques

We sampled 19 species belonging to Clethrionomyini (Appendix A). Nine more species were represented by at least one gene sequence obtained from GenBank so that 28 of 36 species in the tribe were analyzed in all (Appendix C). The remaining species are rare in natural history collections usually because they occur in areas difficult to access. Representatives of two other arvicolid genera were used as outgroups in all analyses (*Microtus* spp. and *Dicrostonyx groenlandicus*). Geographic sampling of the tribe extends from Europe to the United States (Figure 1, Appendix A) with specimens obtained from fieldwork or from the Museum of Southwestern Biology, University of Alaska Museum of the North, University of Washington Burke Museum, and Royal Ontario Museum.

Genomic DNA was extracted from frozen (-80 °C) or ethanol-preserved tissue or dried skin samples using commercially available kits (Qiagen Inc., Valencia, California). Ethanol preserved tissues and skin samples were washed in STE or PBS buffer prior to extraction. We used the polymerase chain reaction (PCR) to amplify four independent loci and sequences were trimmed to the longest length that included the maximum number of individuals: 671-1143 bp of mtDNA cytochrome-b gene (Cyt b), and 3 nuclear

loci including 1059 bp of the first exon of interphotoreceptor retinoid binding protein gene (IRBP) (Stanhope, Smith et al. 1996), 569 bp of intron 1 of Glucose 6-phosphate dehydrogenase (G6pd) (Iwasa and Suzuki 2002), and 926 bp of V-ets erythroblastosis virus E26 oncogene homolog 2 (ETS2) (Lyons, Laughlin et al. 1997) genes. MSB05 and MSB14 (Hope, Waltari et al. 2010) or VOLE14 (Hadly, Ramakrishnan et al. 2004) primers were used for double-stranded amplifications and sequencing of Cyt b. For degraded DNA, several pairs of internal primers were used to amplify ~400bp sections (Kohli, Chapter 2). Primer pairs IRBPA/B (Stanhope, Smith et al. 1996), G6pd-int1L/H, and ETS2F/R (Lyons, Laughlin et al. 1997) were used for IRBP, G6pd and ETS2 genes, respectively. Multiple genes provide independent perspectives of evolutionary history and account for discord with existing taxonomy by identifying instances of mitochondrial capture due to hybridization between species of Clethrionomyini.

For all targeted regions, PCR reagents and conditions were: 1 µl of template DNA (variable concentration); 1.5 µl each of deoxynucleoside triphosphates (10 mM), MgCl (25 mM), 10X PCR buffer, and bovine serum albumin (1%); 0.5 µl of each primer (2 mM); 0.08 µl of AmpliTaq DNA polymerase (Applied Biosystems, Foster City, California); and 6.92 µl of double-distilled H_2O to total 15-µl reactions. Polymerase chain reaction was performed in a PTC 200 thermocycler (MJ Research, Waltham, Massachusetts) with initial denaturation at 94°C for 5 min, followed by 35 cycles of denaturation at 94°C for 15 s, annealing at 50°C for 20 s, extension at 72°C for 1 min, and final extension at 72°C for 5 min, followed by cooling at 15°C for 10 minutes. For nuclear genes, PCR conditions were modified for annealing temperature and to achieve maximum yield (G6pd and IRBP: 56°C; ETS2: 63°C). Amplified PCR product was

visualized using agarose gel electrophoresis and PCR cleanup was performed with ExoSAP-IT (Affymetrix Inc., Santa Clara, California). Cycle sequencing was carried out using the ABI BigDye version 3.1 Sequencing Kit (Applied Biosystems, Foster City, California) and the following reagents for each 10-µl cycle sequencing reaction: 2 µl of H2O, 2 µl of BigDyebuffer (Applied Biosystems), 0.5 µl of primer (2 mM), 0.5 µl of ABI BigDye version 3.1 (Applied Biosystems), and 5 µl of DNA (variable concentration). Conditions for cycle sequencing followed Platt et al. (2007), with reactions cleaned using ethanol/EDTA/sodium acetate precipitation (Applied Biosystems). Products were sequenced and scored using an ABI 3100 automated genetic analyzer (Applied Biosystems) at the University of New Mexico Molecular Biology Facility.

Sequences were edited and aligned with SEQUENCHER v4.9 (GeneCodesCorporation) and checked by eye. Complementary strands were aligned to create composite sequences for analysis. To ensure genuine mtDNA was sequenced, Cyt b sequences were translated to amino acids and inspected for internal stop codons. The program PHASE v2.1 (Stephens, Smith et al. 2001) was used to infer both allelic sequences of nuclear genes. To estimate allele frequencies, the program was run for four iterations with results of the best goodness-of-fit to an approximate coalescent model retained and outputting two nuclear haplotypes (alleles). Only likelihood values over 0.9 were unambiguously assigned a base, with ambiguous sites changed to N to avoid biasing phylogenetic construction.

2.3 Gene tree reconstruction

Sequence alignment was completed with MEGA v5 (Tamura, Peterson et al. 2011) using the MUSCLE algorithm and validated by eye. A total of 189 individuals were included in Cyt b analysis, ranging from 671-1143 bp in length. For nuclear gene trees, a subset of individuals was sampled (Table 1, Appendix B). GenBank sequences were limited for nuclear genes, but were available for IRBP (N=4) and G6pd (N=24) (Appendix C). MrModeltest v2.3 (Nylander 2004) was used to find the best model of evolution for each gene based on the Akaike Information Criterion. Cyt b was partitioned by codon position. Gene trees were reconstructed using Bayesian methods applied through the program MrBayes 3.2.1 (Huelsenbeck and Ronquist 2001). Phylogenies were reconstructed using Markov chain Monte Carlo procedures run for 10-20 million generations, sampling every 1000 generations, with 4 independent chains and a burnin of 25%. Sufficient mixing was assessed by examining the split of the standard deviations in the output, ensuring it was below 0.01 after burnin.

2.4 Species tree estimation and divergence dates

Differing gene histories can greatly affect phylogeny reconstructions, which can cause discordance between gene trees and the actual species tree (Rosenberg 2002). Species tree estimation from multiple independent loci is an effective solution and preferable to concatenation (Edwards, Liu et al. 2007). BEAUti v7.0.1 was used to set up species tree runs in the Bayesian reconstruction program BEAST (Drummond and Rambaut 2007), which applied the *BEAST algorithm (Heled and Drummond 2010). A subset of up to 4 individuals per species was used for each gene. We initially allowed all molecular clocks to vary (lognormal relaxed clock prior), but if preliminary runs supported a clock-like rate of evolution, priors were adjusted accordingly. Models of

evolution were input as priors as informed by MrModelTest. Due to a lack of convergence in long preliminary runs, mtDNA partitions were not applied in the final analysis.

Using UPGMA starting trees, 500 million MCMC generations were run for each species tree. To explore the possibility of mtDNA capture between species, nuclear data were run separately and compared to the species tree that included mtDNA. We assessed convergence by examining run parameters in TRACER, making sure ESSs were >200 and then performing at least 2 identical runs. Results were pooled using LOGCOMBINER to ensure that runs started from unique random points were converging (Drummond and Rambaut 2007).

BEAST simultaneously estimates divergence dates during phylogeny reconstruction (Drummond and Rambaut 2007) and these estimates were based on the species tree that included all genes (Sanchez-Gracia and Castresana 2012). To calibrate date estimates, information from three Clethrionomyini fossils were input as priors. The earliest known fossil assigned to each of the following genera were used to set hard minimum bounds on the age of the groups: *Alticola*, 1.5 MY (Serdyuk and Tesakov 2006); *Clethrionomys*, 2.6 MY (Repenning, Fejfar et al. 1990); *Eothenomys*, 2 MY (Liu, Liu et al. 2012). Exponential priors were then adjusted for each of the three t_{MRCA} fossil calibrations so that the 97.5% of the prior distribution fell between the hard minimum and 5.5 MY, the most recent suspected origin of Arvicolinae, according to fossil evidence (Chaline, Brunet-Lecomte et al. 1999).

2.5 Sequence divergence and genetic diversity

Sequence divergence between species was calculated by average pairwise distance for each gene in MEGA. Summary statistics were calculated using DnaSP (Librado and Rozas 2009) to assess intraspecific genetic diversity (segregating sites, haplotype diversity, and nucleotide diversity) and to compare the variability at each locus.

2.6 Characterization of evolutionary patterns

To make a first attempt at characterizing the diversification of Clethrionomyini and assess the putative link to episodic environmental changes, we investigate the rate of lineage accumulation over time, visualized as a lineage-through-time (LTT) plot (Tolley, Chase et al. 2008; Silberfeld, Leigh et al. 2010). Quantitatively, we used the gamma statistic to test the assumption of a constant diversification rate over time, applying the Monte Carlo constant rates (MCCR) test to correct for incomplete taxon sampling (Pybus and Harvey 2000). An LTT plot was created in GENIE (Pybus and Rambaut 2002), using a four-gene species tree chronogram (containing no outgroups) generated by BEAST as the input tree. Although incomplete taxonomic sampling can affect lineage accumulation curves, constructing LTT plots from unresolved or spurious relationships (e.g. due to horizontal gene transfer) can also greatly misrepresent diversification patterns (Pybus and Harvey 2000; Barraclough and Nee 2001). Given documented mitochondrial capture events among species (and potentially among genera) of Clethrionomyini and incomplete taxon sampling even for mtDNA, we chose to sacrifice partial taxonomic sampling in exchange for increased genomic sampling through the use of a multilocus species tree. Species trees should be used when complete taxon sampling cannot be

achieved because they should best reflect the true relationships of taxa and offer the most informed time estimates available for tree nodes (Barraclough and Nee 2001).

To determine whether rates of diversification varied through time, an expected or "constant-rate" lineage accumulation curve and 95% confidence intervals (CI) were generated by simulating 1000 random trees in Mesquite (Maddison and Maddison 2011). Simulations accounted for incomplete sampling by utilizing the "Uniform speciation with sampling" option to create trees with the number of taxa equal to complete taxon sampling before randomly pruning branches to match the number in the original data set (input species tree) (Maddison and Maddison 2011). This method reduces the bias that incomplete taxon sampling creates, such as artificially removing recent splitting events or rare events, by accounting for a wider array of potential branching patterns that missing taxa may represent (Pybus and Harvey 2000; Barraclough and Nee 2001). Gamma estimates for the 1000 simulated trees were then generated using GENIE to create the gamma distribution and determine the critical value for the test (one-tailed, $p<0.05$). GENIE was also used to construct LTT plots for the simulated trees and then summarized to determine the expected curve and 95% CI.

3. Results

Sequences of nuclear genes and Cyt b were obtained for 19 species in Clethrionomyini, including three species not previously sequenced (*A. montosa, A. albicaudus,* and *H. fertilis*) (Appendix B). We were unable to sequence all four genes for *A. montosa* and *H. fertilis,* but these species are included in the Cyt b and G6pd trees. No

NUMTs were identified. With the addition of GenBank sequences, 28 of the 36 species of Clethrionomyini were included in mtDNA analyses.

3.1 Gene tree reconstruction

All Bayesian gene trees were moderately- to well-resolved at all taxonomic levels (Figure 2A-D). Each gene tree supports the existence of two primary clades: one with *C. rufocanus* and associated species ("*Craseomys*"), and one with all other *Clethrionomys* and *Alticola*. A third group is not always monophyletic, but all species of *Eothenomys* and *Caryomys* are consistently basal to the two other clades.

Other than the split between the "*Craseomys*" group of red-backed voles and other *Clethrionomys,* nuclear genes support monophyly of each nominal genus. Similarly, the Cyt b tree supports current generic level designation of species, except for the two previously documented cases of apparent polyphyly (*A. macrotis* and *A. lemminus* grouping with *Clethrionomys*). This result did not change with the increased intraspecific and taxonomic sampling we added (Bayesian posterior probability (PP) of 1.0; Figure 2A). We were unable to obtain samples of *A. macrotis* to sequence nuclear genes but *A. lemminus* was successfully sequenced for all loci. The three nuclear genes disagree with the mtDNA topology, instead grouping *A. lemminus* with other *Alticola* (0.84-0.91 PP).

H. fertilis and *A. montosa* were included in Cyt b and G6pd trees. *A. montosa* consistently grouped with other *Alticola* species, as did *A. albicaudus* (Figure 2A-D). *H. fertilis* is not phylogenetically close to members of Clethrionomyini (Figure 2A, D), but instead appears more closely related to outgroups.

3.2 Species tree estimation and divergence dates

Estimating the species tree with only nuclear genes did not significantly change the topology from the species tree constructed with all loci (Figure 3). The one notable exception is the strength of support placing *A. lemminus* with other *Alticola*. Support for *Alticola* monophyly is strong (PP= 0.97) when mtDNA is excluded, compared to equivocal (PP<0.74) when both mtDNA and nuclear DNA are used to estimate the species tree. The three main clades observed in gene trees were recovered by both species tree methods as well, but with stronger support. However, deep nodes of *Clethrionomys* (excluding the "*Craseomys*"-type species) are essentially a polytomy, as also seen in individual gene trees, including Cyt b; a finding recovered in previous molecular analyses (Cook, Runck et al. 2004; Lebedev, Bannikova et al. 2007). As seen in gene trees, *Eothenomys* and *Caryomys* are basal in species tree reconstructions.

Date estimates are similar between species tree methods, with overlapping 95% confidence intervals, although the nuclear only tree has uniformly wider CIs. The estimate of time to most recent common ancestor (t_{MRCA}) of Clethrionomyini is around 4 million years ago (MYA) (Figure 3). Divergence of Clethrionomyini from *Microtus* and *Dicrostonyx* is estimated to be between 5-7 MYA. The t_{MRCA} of *Caryomys* and *Eothenomys* is roughly 3 MYA, similar in timing to the split between the *Craseomys* and *Alticola-Clethrionomys* clades. Finally, *Alticola* and *Clethrionomys* share a common ancestor from about 2 MYA. *Craseomys* has the most recent t_{MRCA} of the sampled genera (~1.4 MYA) followed by *Alticola* and *Clethrionomys* (both near 1.7 MYA), and the somewhat more ancient coalescence of *Eothenomys* (~2.5 MYA).

3.3 Sequence divergence and genetic diversity

Cyt b is the most variable gene we sampled, but each gene showed high haplotype number and haplotype diversity (Table 1). Nucleotide diversity is much higher in Cyt b than in any of the nuclear genes, as expected due to more rapid accumulation of mutations in mtDNA than in nuclear genes. However, even the least variable gene, G6pd, is phylogenetically informative for the tribe, as initially recognized in work on *Clethrionomys* (Iwasa and Suzuki 2002; Boratynski, Alves et al. 2011). Sequence divergence within and among genera is about an order of magnitude greater for Cyt b than the nuclear genes (Table 2). *Hyperacrius* consistently exhibits the highest average divergences from all other genera for Cyt b and G6pd. Comparisons among species of *Alticola* and *Clethrionomys* result in clusters of similar divergence values correspond to genera or subgenera (*Craseomys*) boundaries (Table 3).

3.4 Characterization of evolutionary patterns

Visually, there is no strong pattern of diversification rates varying over time from the LTT plot and expected curve (Figure 4). The gamma statistic for our phylogeny was -3.8, which is not statistically significant (Critical gamma value: -4.24, $p>0.05$), meaning there is no indication of changes in diversification rate through time for Clethrionomyini. However, the empirical curve does conspicuously bulge above the expected curve at two points, about 3 MYA and 1.5-1.35 MYA.

4. Discussion

Widespread and diverse groups, such as Clethrionomyini, offer broad perspectives to investigate the effects of environmental change over multiple geographic regions, biomes, and unique biogeographical histories. Understanding the phylogenetic

relationships and testing biogeographic and evolutionary hypotheses of Clethrionomyini provides insight into the critical factors that generated mammalian diversity throughout the Holarctic over the past several million years. For example, Arctic and subarctic ecosystems in the Northern Hemisphere experienced dramatic environmental change and fluctuation with periods of increased intensity and frequency during the Late Cenozoic (Webb III and Bartlein 1992; Ehlers and Gibbard 2007; Miller, Brigham-Grette et al. 2010). The impacts of glacial cycles were seen on continental scales, contributing to the restructuring of biological communities, driving diversification, and creating conditions quite different from those we observe today (Abbott and Brochmann 2003; Hewitt 2004; Lister 2004). Due to the disruption of global heat and moisture transport, the effects were strong not only in regions of vast glaciations (e.g. North America and northern Europe) but also in surrounding regions that were relatively ice free, such as the Tibetan Plateau in western China, and much of central and northern Asia (Tarasov, Volkova et al. 2000; Ni, Yu et al. 2010). Change was compounded by effects of other large scale geologic events, such as the intensification of Tibetan Plateau uplift (~1.5-3.5 MYA), that further altered regional climate patterns and created dramatic topographic variation (Li and Fang 1999; Zhisheng, Kutzbach et al. 2001).

Pulses of diversification, as recognized through fossil and genetic evidence, have been linked to these dominant factors and are thought to have been the primary contributors to diversification in northern taxa, especially arvicolines (Repenning, Fejfar et al. 1990; Conroy and Cook 1999; Buzan, Krystufek et al. 2008; Liu, Liu et al. 2012). However, to effectively address this issue and many other evolutionary questions, we need accurate taxonomy. The need for multilocus phylogenetic analysis has long been

recognized, especially in cases where hybridization, horizontal gene transfer, gene duplication, or incomplete lineage sorting are suspected (Maddison 1997; Rosenberg 2002). When these phenomena occur, there is a high probability that a gene tree will not reflect the actual history of evolution. While mtDNA has many advantages and has provided a wealth of information about phylogenetics and phylogeography (Avise 2000), this class of marker is best used to complement a suite of other independent loci to reduce the chance of inferring a species tree from a discordant gene tree (Sanchez-Gracia and Castresana 2012). Our results are the first broad application of a multilocus approach for examining Clethrionomyini, a group thought to have diversified via the pulse model identified for other arvicolines (Conroy and Cook 1999; Buzan, Krystufek et al. 2008).

4.1. *Multilocus insights and taxonomy*

Phylogenetic relationships of Clethrionomyini have proven difficult to resolve in part due to a recent radiation (potentially resulting in incomplete lineage sorting) and repeated episodes of mtDNA introgression and hybridization between species (Deffontaine, Libois et al. 2005; Runck, Matocq et al. 2009). Previous morphological, karyotypic, and genetic studies of Clethrionomyini revealed the possibility of paraphyletic and polyphyletic nominal genera (Iwasa and Suzuki 2002; Cook, Runck et al. 2004; Luo, Yang et al. 2004; Lebedev, Bannikova et al. 2007); however, until now, nuclear sequences were unavailable to test those hypotheses for the vast majority of tribal species. With the addition of three independent nuclear loci and a more complete taxon sampling of Clehtrionomyini, we can more confidently assess species limits and higher level taxonomy in the tribe, identify cases of hybridization, and explore evolutionary and biogeographical hypotheses at high latitudes.

4.1.1 Cases of mtDNA paraphyly and polyphyly

The most significant discord between nuclear DNA and mtDNA data is that our multilocus analysis does not support polyphyly of *Alticola* with respect to *A. lemminus*, whereas mtDNA alone strongly does (Cook, Runck et al. 2004; Lebedev, Bannikova et al. 2007). The arctic *A. lemminus* is atypical compared to other montane voles of the genus *Alticola* in distribution and morphological features (e.g., white winter pelage). Shortly after the divergence of *Alticola* and *Clethrionomys* about 2 MYA, *A. lemminus* appears to have captured mtDNA from *C. rutilus*, as evidenced by the low divergence of the two species' current Cyt b sequences (Table 3). Cyt b pairwise comparisons show that *A. lemminus* and *C. rutilus* are the most similar species in the *Clethrionomys-Alticola* clade, with the exception of *A. macrotis* and *C. centralis* (Table 3). The distributions of *A. lemminus* and *C. rutilus* overlap throughout much of northeastern Siberia (Figure 1). However, the magnitude of sequence divergence supports an ancient introgression event rather than recent exchange. Unfortunately, nuclear genes of *A. macrotis* could not be sequenced, but its mtDNA relationship to *C. centralis* may represent a second case of mitochondrial introgression rather than indicating polyphyly of *Alticola*.

At least two other species of *Clethrionomys* have acquired the mtDNA of *C. rutilus* through historic or modern introgression (Deffontaine, Libois et al. 2005; Runck, Matocq et al. 2009). In *C. gapperi* and *C. glareolus*, individuals harboring foreign mtDNA are found predominantly at the northern edge of the respective species range, where they are parapatric or sympatric with *C. rutilus*. We identified one hybrid individual in our analysis that possessed the mtDNA signature of *C. rutilus* but *C. gapperi* IRBP and G6pd (ETS2 was not sequenced). *Clethrionomys rutilus* is more

closely associated with tundra habitat than other red-backed vole species, leading to a hypothetical adaptive mechanism for mitochondrial capture in *C. glareolus*, but this pattern could simply be due to demographic processes on the edge of expanding species ranges (Boratynski, Alves et al. 2011). Other northern animals show possible adaptive introgression of mitochondria related to the metabolic role of mtDNA genes (Doiron, Bernatchez et al. 2002; Alves, Melo-Ferreira et al. 2008). An ancient mitochondrial capture event for *A. lemminus* may have been adaptive, allowing this species to colonize the harsher environments of the far north. A similar transfer of mtDNA to *A. macrotis* from *C. centralis* is possible (similar Cyt b divergence value), and may account for this second perceived case of *Alticola* polyphyly (Lebedev, Bannikova et al. 2007). However, it should be noted that this pair of species do not currently overlap in range.

Within *Clethrionomys*, monophyly of the subgenus *Craseomys* (*C. rufocanus* and Korean and Japanese red-backed voles) and apparent deep divergence from other *Clethrionomys* was recognized in morphologic and molecular studies (Nadler, Zhurkevich et al. 1978; Iwasa and Suzuki 2002; Lebedev, Bannikova et al. 2007; Koh, Yang et al. 2010), with some concluding that *Craseomys* should be recognized as a new genus (Lebedev, Bannikova et al. 2007). We concur, as *Craseomys* was found to be reciprocally monophyletic to the remainder of the *Clethrionomys*-*Alticola* clade (with a divergence date at least 3 MYA). This revision would avoid placing *Alticola* within *Clethrionomys*.

Another controversial genus, *Caryomys*, is endemic to China and distinctive from other species in our mtDNA analysis, but only one of the two species (*C. eva*) was included in nuclear analyses. Based on external morphology and previous molecular

work, *Caryomys* shares some characters with *Clethrionomys* but others with *Eothenomys*, making its taxonomic position uncertain (Carleton and Musser 2005). In all of our analyses, *Caryomys* is consistently basal to *Eothenomys* or shares a basal position to all other tribe genera (Figures 2 and 3). Our species tree reconstructions indicate divergence occurred between 2.9-3.8 MYA. Average sequence divergence is equivocal (Table 3). Future work should further test the validity of this taxon by including representatives of remaining species, *C. inez*, in multilocus analyses.

4.1.2 Newly sequenced species

Clethrionomyini is a diverse tribe with several species occurring in difficult to access regions of Asia that previously prevented their inclusion in phylogenetic studies. Inclusion of *Hyperacrius fertilis* as a member of Clethrionomyini is not supported by analysis of Cyt b and G6pd; a finding concordant with morphological analyses aligning the species with a basal arvicolid, *Prometheomys* (Galewski, Tilak et al. 2006) instead of an offshoot of *Alticola* (Phillips 1969). In the past, *Hyperacrius* has been incorrectly affiliated with Clethrionomyini (Carleton and Musser 2005), probably as a result of convergent morphology related to rocky, arid habitats (Phillips 1969). This result emphasizes the need to integrate morphology and molecules to correctly resolve the placement of *Hyperacrius* in Arvicolinae.

Both *A. montosa* and *A. albicaudus* are supported as members of *Alticola*. *A. albicaudus* is closely related to *A. argentatus*, with low pairwise divergence values at all loci and forming a monophyletic group with strong support (Table 2; Figures 2 and 3). *A. montosa* also appears related to *A. argentatus* based on the Cyt b tree, with a strongly supported clade composed of *A. argentatus, A. montosa*, and *A. albicaudus* (Figure 2A,

D). Nesting of *A. albicaudus* individuals among *A. argentatus* for IRBP and Cyt b indicates that boundaries among these three species and their relationships should be examined in more detail.

4.1.3 Intraspecific diversity

While a single molecular marker should not be the basis for phylogenetic relationships, the mtDNA Cyt b has proven highly informative for initial exploration of phylogeography (Avise 2000). We sampled multiple individuals per species from across broad geographic areas to explore intraspecific diversity and uncover potential cryptic diversity (Figure 2, Appendices). Three species (*C. gapperi*, *C. eva*, *A. argentatus*) in our analysis exhibit deep intraspecific structure, raising the possibility of cryptic species (Figure 2A). Deep structure previously reported for *C. gapperi* (Cook, Runck et al. 2004) appears to coincide with isolation due to glacial coverage in North America during the Pleistocene. Nuclear loci do not show structure, but the presence of reciprocally monophyletic mtDNA groups warrants further investigation. Similarly, individuals of *C. eva* form a well-supported clade but deep intraspecific mtDNA divergence in *C. eva* suggests the possible existence of a cryptic species of *Caryomys* (Figure 2A). *A. argentatus* has an unusual, disjunct, and extensive distribution across central Asia (Figure 1). Deep divergence and lack of monophyly and among samples of putative *A. argentatus* point toward the possibility of unrecognized diversity or hybridization. *E. chinensis* do not cluster in the Cyt b tree, but these GenBank sequences included one sample hypothesized to be a new species by Liu et al. (2012). Expanded sampling and a comprehensive multilocus treatment of *Eothenomys* should help clarify taxonomic limits in this diverse genus.

To a lesser extent, phylogeographic structure is evident in other species. For example, *C. rutilus*, which has the largest distribution of Clethrionomyini members, consists of 3 well-defined clades (Kohli, Chapter 2). Similarly, *C. rufocanus* exhibits a deep split between populations on Hokkaido Island, Japan and those from mainland Asia (Figure 2A).

4.2. Evolutionary history and biogeography

Arvicoline evolution has been characterized as a series of pulses of diversification over the past 5-10 MY. Pulses might be expected for taxa inhabiting the high latitudes of the Northern Hemisphere, a region heavily impacted by repeated episodes of environmental change (Repenning, Fejfar et al. 1990; Conroy and Cook 1999). This analysis of Clethrionomyini represents a first attempt to characterize the mode and tempo of diversification in one of the most diverse groups within Arvicolinae. For this tribe, there is limited support for pulses of diversification, but instead a relatively constant rate of lineage accumulation characterizes the 4 MY history of this group (Figure 4). However, this result does not completely dismiss the hypothesis that episodic environmental events contributed to diversity in the group. For example, the LTT plot shows two brief periods of rapid lineage accumulation occurring at about 3 MYA and 1.5-1.35 MYA, which correspond to the origin of Clethrionomyini genera and subsequent diversification within *Alticola*, *Clethrionomys*, and *Craseomys* (Figure 3A), respectively. These periods may correspond to rapid diversification related to Late Cenozoic glacial cycles, but this preliminary pattern should be more rigorously tested. Although this is the largest multilocus Clethrionomyini dataset analyzed to date and is corrected for

incomplete taxon sampling, more extensive (>80%) taxon sampling is recommended to confirm results from MCCR tests (Cusimano and Renner 2010).

Simultaneous diversification, as suggested by polytomies observed in the species tree, is intriguing from an evolutionary standpoint despite the overall tribe not conforming to the pulse model of diversification. *C. rutilus* represents an example of episodic diversification at the intraspecific level and perhaps at an interspecific level as well (Kohli, Chapter 2). Unresolved relationships among *C. rutilus*, *C. gapperi*, and *C. glareolus* (Figures 2 and 3) (Cook, Runck et al. 2004; Lebedev, Bannikova et al. 2007) mirror the 3 historically isolated lineages within *C. rutilus*, with central, eastern, and western distributions relative to one another (Kohli, Chapter 2). *C. gapperi* was located south of the ice sheets in North America (Runck and Cook 2005), *C. glareolus* persisted and diversified in several refugia in and near Europe (Deffontaine, Libois et al. 2005; Kotlik, Deffontaine et al. 2006), and *C. rutilus* was limited to regions between the Ural Mountains and East Beringia, occupying the central distribution between *C. gapperi* and *C. glareolus*. Both sets of taxa (intra- and interspecific) collapse to polytomies, even in multilocus analyses (Figures 2 and 3). Given the cyclical nature of environmental change in high-latitude ecosystems since the Pliocene, it is reasonable to hypothesize that the repeatability of evolutionary outcomes over time as represented at multiple phylogenetic levels can result from similar biogeographic processes.

Following the origin of modern genera of Clethrionomyini, diversification of each of these clades was more gradual than previously suspected. Idiosyncratic biotic responses to changes associated with glacial cycles may have contributed to the lack of significantly punctuated diversification of these predominantly Asian voles. Although

most of Asia was relatively ice-free during glacial periods, climate cycles dramatically altered community distributions and compositions in northern Asia (Tarasov, Volkova et al. 2000; Harrison, Yu et al. 2001; Velichko and Wright 2005). *Alticola* and *Clethrionomys* would have been subject to these environmental changes. In contrast to their more northern relatives, diversification within *Eothenomys* is thought to be tied to the uplift of the Tibetan Plateau (Luo, Yang et al. 2004; Liu, Liu et al. 2012), as hypothesized for other regional species (Zhang and Ge 2007). Major episodes of uplift occurred 3 times in roughly 1 MY intervals between 3.6-1.5 MYA (Li and Fang 1999). Multilocus data for *Eothenomys* are limited but do not support the correlation between uplift and increased diversification. The basal position occupied by the *Eothenomys* clade contrasts with the fossil record that shows *Clethrionomys* as the earliest genus, appearing 2.5-3 MYA in western Asia and in the Lake Baikal region (Chaline and Graf 1988; Repenning, Fejfar et al. 1990). The earliest fossil of *Eothenomys* is from China and dates to about 2 MYA (Zheng and Li 1990). Although the recent fossil record of arvicoline rodents and Clethrionomyini is substantial, deep fossils are not common and molecular estimates of the origin of tribe predate all known fossils (Lebedev, Bannikova et al. 2007).

Clethrionomys was probably most affected by glacial cycles leading to divergence in multiple refugia across the Holarctic, a biogeographic process that was influenced by the repeated connection of North America to Eurasia via the Bering Land Bridge. Species of *Clethrionomys* dispersed eastward through Beringia, at least twice and possibly on more occasions (Repenning 2001; Cook, Runck et al. 2004; Runck, Matocq et al. 2009). Close affinity of *C. rufocanus* with Korean and Japanese red-backed voles

("Craseomys") supports their independent diversification in east Asia, which agrees with previous studies suggesting the isolation of *C. rufocanus* in the Asian Far East by the mid-Pleistocene (~0.7 MYA) (Chaline and Graf 1988). Our t_{MRCA} estimate of ~1.4 MYA for *Craseomys* significantly predates the fossil evidence.

Alticola is unique among Clethrionomyini in their association with arid, rocky environments. Adaptations to new ecological conditions may have allowed *Alticola* to invade and diversify in arid geographic regions in Asia. However, the distribution of the genus, centered on central Asian mountain chains, invokes the possible role of periodic glaciations leading to vicariance and speciation. Discovery of mtDNA polyphyly of two *Alticola* species within *Clethrionomys* led to the hypothesis that the "*Alticola* ecotype" had independently evolved multiple times from *Clethrionomys* forms, presumably because of its adaptive advantage for colonizing rocky alpine habitats (Lebedev, Bannikova et al. 2007). Multilocus data support a more parsimonious evolutionary history in which a single evolutionary transition occurred from a forest-associated ancestor to one that inhabited rocky sites and subsequently diversified into 12 species. The basal position of *A. lemminus* among *Alticola* support a morphological assessment that characterized it and *A. macrotis* (formerly considered a single species) as "transitional" between *Alticola* and *Clethrionomys* (Gromov and Polyakov 1977).

4.3. Conclusions

By using a multilocus approach and sampling all 5 genera of Clethrionomyini, standing taxonomic issues were resolved and cryptic diversity was uncovered that requires further study. Monophyly of *Alticola* is supported by nuclear sequences, with cases of ancient mitochondrial introgression now detected, a finding that falsifies

previous hypotheses of polyphyly of species of *Alticola*. In contrast, *Clethrionomys*, as currently defined (Musser and Carleton 2005), is paraphyletic, so we propose elevating *Craseomys* to genus. This phylogenetic analysis also supports excluding *Hyperacrius* from the tribe, leaving 5 extant genera: *Alticola*, *Clethrionomys*, *Craseomys*, *Caryomys*, and *Eothenomys*. Pulses are not the primary pattern of diversification within the tribe over the last 4 MY. However, this preliminary assessment of Clethrionomyini should be tested with increased taxonomic and genomic sampling for these broadly distributed and ecologically diverse northern organisms.

References

Abbott, R. J. and C. Brochmann (2003). "History and evolution of the arctic flora: in the footsteps of Eric Hulten." Molecular Ecology **12**(2): 299-313.

Abramson, N. I., V. S. Lebedev, et al. (2009). "Supraspecies Relationships in the Subfamily Arvicolinae (Rodentia, Cricetidae): An Unexpected Result of Nuclear Gene Analysis." Molecular Biology **43**(5): 834-846.

Alves, P. C., J. Melo-Ferreira, et al. (2008). "The ubiquitous mountain hare mitochondria: multiple introgressive hybridization in hares, genus *Lepus*." Philosophical Transactions of the Royal Society B: Biological Sciences **363**(1505): 2831-2839.

Avise, J. C. (2000). Phylogeography: The history and formation of species. Cambridge, Massachusetts, Harvard University Press.

Barraclough, T. G. and S. Nee (2001). "Phylogenetics and speciation." Trends in Ecology & Evolution **16**(7): 391-399.

Boratynski, Z., P. C. Alves, et al. (2011). "Introgression of mitochondrial DNA among *Myodes* voles: consequences for energetics?" BMC evolutionary biology **11**(1): 355.

Buzan, E. V., B. Krystufek, et al. (2008). "Mitochondrial phylogeny of Arvicolinae using comprehensive taxonomic sampling yields new insights." Biological Journal of the Linnean Society **94**(4): 825-835.

Carleton, M. D. and G. G. Musser (2005). Order Rodentia. Mammal Species of the World. D. E. Wilson and D. M. Reeder. Baltimore, Maryland, The Johns Hopkins University Press.

Chaline, J., P. Brunet-Lecomte, et al. (1999). Anatomy of the arvicoline radiation (Rodentia): palaeogeographical, palaeoecological history and evolutionary data. Annales Zoologici Fennici, Helsinki: Suomen Biologian Seura Vanamo, 1964-.

Chaline, J. and J. D. Graf (1988). "Phylogeny of the Arvicolidae (Rodentia): biochemical and paleontological evidence." Journal of Mammalogy: 22-33.

Conroy, C. J. and J. A. Cook (1999). "MtDNA evidence for repeated pulses of speciation within arvicoline and murid rodents." Journal of Mammalian Evolution **6**(3): 221-245.

Cook, J. A., A. M. Runck, et al. (2004). "Historical biogeography at the crossroads of the northern continents: molecular phylogenetics of red-backed voles (Rodentia : Arvicolinae)." Molecular Phylogenetics and Evolution **30**(3): 767-777.

Cusimano, N. and S. S. Renner (2010). "Slowdowns in diversification rates from real phylogenies may not be real." Systematic Biology **59**(4): 458-464.

Deffontaine, V., R. Libois, et al. (2005). "Beyond the Mediterranean peninsulas: evidence of central European glacial refugia for a temperate forest mammal species, the bank vole (*Clethrionomys glareolus*)." Molecular Ecology **14**(6): 1727-1739.

Doiron, S., L. Bernatchez, et al. (2002). "A comparative mitogenomic analysis of the potential adaptive value of arctic charr mtDNA introgression in brook charr populations (*Salvelinus fontinalis* Mitchill)." Molecular biology and evolution **19**(11): 1902-1909.

Drummond, A. J. and A. Rambaut (2007). "BEAST: Bayesian evolutionary analysis by sampling trees." BMC Evolutionary Biology **7**.

Edwards, S. V., L. Liu, et al. (2007). "High-resolution species trees without concatenation." Proceedings of the National Academy of Sciences **104**(14): 5936.

Ehlers, J. and P. L. Gibbard (2007). "The extent and chronology of Cenozoic global glaciation." Quaternary International **164**: 6-20.

Galewski, T., M.-k. Tilak, et al. (2006). "The evolutionary radiation of Arvicolinae rodents (voles and lemmings): relative contribution of nuclear and mitochondrial DNA phylogenies." Bmc Evolutionary Biology **6**(80): 1-17.

Gromov, I. M. and I. Y. Polyakov (1977). Voles (Microtinae), Fauna of the USSR, Mammals Washington, D.C., Smithsonian Institution Libraries and the National Science Foundation.

Hadly, E. A., U. Ramakrishnan, et al. (2004). "Genetic response to climatic change: insights from ancient DNA and phylochronology." PLoS Biology **2**(10): e290.

Harrison, S., G. Yu, et al. (2001). "Palaeovegetation (Communications arising): diversity of temperate plants in east Asia." Nature **413**(6852): 129-130.

Heled, J. and A. J. Drummond (2010). "Bayesian Inference of Species Trees from Multilocus Data." Molecular Biology and Evolution **27**(3): 570-580.

Hewitt, G. M. (2004). "The structure of biodiversity–insights from molecular phylogeography." Frontiers in Zoology **1**(4): 1-16.

Hope, A. G., E. Waltari, et al. (2010). "High-latitude diversification within Eurasian least shrews and Alaska tiny shrews (Soricidae)." Journal of Mammalogy **91**(5): 1041-1057

Huelsenbeck, J. P. and F. Ronquist (2001). "MRBAYES: Bayesian inference of phylogenetic trees." Bioinformatics **17**(8): 754-755.

Iwasa, M. A. and H. Suzuki (2002). "Evolutionary significance of chromosome changes in northeastern Asiatic red-backed voles inferred with the aid of intron 1 sequences of the G6pd gene." Chromosome Research **10**(5): 419-428.

Kaneko, Y. (1992). "Identification and Morphological Characteristics of *Clethrionomys rufocanus*, *Eothenomys shanseius*, *E. inez* and *E. eva* from the USSR, Mongolia,

and Northern and Central China." Journal of the Mammalogical Society of Japan **16**(2): 71-95.

Kaneko, Y. (1996). "Morphological variation, and latitudinal and altitudinal distribution of *Eothenomys chinensis*, *E. wardi*, *E. custos*, *E. proditor*, and *E. olitor* (Rodentia, Arvicolidae) in China." Mammal Study **21**(2): 89-114.

Koh, H. S., B. K. Yang, et al. (2010). "Genetic distinctiveness of the Korean red-backed vole (*Myodes regulus*) from Korea, revealed by mitochondrial cytochrome *b* gene sequences." Biochemical Genetics: 1-8.

Kotlik, P., V. Deffontaine, et al. (2006). "A northern glacial refugium for bank voles (*Clethrionomys glareolus*)." Proceedings of the National Academy of Sciences **103**(40): 14860-14864.

Lebedev, V. S., A. A. Bannikova, et al. (2007). "Molecular phylogeny of the genus Alticola (Cricetidae, Rodentia) as inferred from the sequence of the cytochrome b gene." Zoologica Scripta **36**(6): 547-563.

Li, J. and X. Fang (1999). "Uplift of the Tibetan Plateau and environmental changes." Chinese Science Bulletin **44**(23): 2117-2124.

Librado, P. and J. Rozas (2009). "DnaSP v5: a software for comprehensive analysis of DNA polymorphism data." Bioinformatics **25**(11): 1451-1452.

Lister, A. M. (2004). "The impact of Quaternary Ice Ages on mammalian evolution." Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **359**(1442): 221.

Liu, S., Y. Liu, et al. (2012). "Phylogeny of Oriental Voles (Rodentia: Muridae: Arvicolinae): Molecular and Morphological Evidence." Zoological Science **29**(9): 610-622.

Luo, J., D. Yang, et al. (2004). "Molecular phylogeny and biogeography of Oriental voles: genus *Eothenomys* (Muridae, Mammalia)." Molecular phylogenetics and evolution **33**(2): 349-362.

Lyons, L. A., T. F. Laughlin, et al. (1997). "Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes." Nature Genetics **15**(1): 47-56.

Maddison, W. and D. Maddison (2011). Mesquite: a modular system for evolutionary analysis. Version 2.75.

Maddison, W. P. (1997). "Gene trees in species trees." Systematic biology **46**(3): 523.

Miller, G. H., J. Brigham-Grette, et al. (2010). "Temperature and precipitation history of the Arctic." Quaternary Science Reviews **29**(15): 1679-1715.

Nadler, C. F., N. Zhurkevich, et al. (1978). "Biochemical relationships of the Holarctic vole genera (*Clethrionomys*, *Microtus*, and *Arvicola* (Rodentia: Arvicolinae))." Canadian journal of zoology **56**(7): 1564-1575.

Ni, J., G. Yu, et al. (2010). "Palaeovegetation in China during the late Quaternary: Biome reconstructions based on a global scheme of plant functional types." Palaeogeography, Palaeoclimatology, Palaeoecology **289**(1-4): 44-61.

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

Phillips, C. J. (1969). "Review of Central Asian Voles of the Genus *Hyperacrius*, with Comments on Zoogeography, Ecology, and Ectoparasites." Journal of Mammalogy: 457-474.

Platt, A. R., R. W. Woodhall, et al. (2007). "Improved DNA sequencing quality and efficiency using an optimized fast cycle sequencing protocol." BioTechniques **43**(1): 58.

Pybus, O. and A. Rambaut (2002). "GENIE: estimating demographic history from molecular phylogenies." Bioinformatics **18**(10): 1404-1405.

Pybus, O. G. and P. H. Harvey (2000). "Testing macro-evolutionary models using incomplete molecular phylogenies." Proceedings of the Royal Society of London. Series B: Biological Sciences **267**(1459): 2267-2272.

Repenning, C. A. (2001). "Beringian climate during intercontinental dispersal: a mouse eye view." Quaternary Science Reviews **20**(1-3): 25-40.

Repenning, C. A., O. Fejfar, et al. (1990). Arvicolid rodent biochronology of the Northern Hemisphere. *International Symposium on the Evolution and Phylogenetic Biostratigraphy of Arvicolids (Rodentia, Mammalia)*, Prague, Pfeil-Verlag.

Robovsky, J., V. Ricankova, et al. (2008). "Phylogeny of Arvicolinae (Mammalia, Cricetidae): utility of morphological and molecular data sets in a recently radiating clade." Zoologica Scripta **37**(6): 571-590.

Rosenberg, N. A. (2002). "The probability of topological concordance of gene trees and species trees." Theoretical Population Biology **61**(2): 225-247.

Runck, A. M. and J. A. Cook (2005). "Postglacial expansion of the southern red-backed vole (*Clethrionomys gapperi*) in North America." Molecular Ecology **14**(5): 1445-1456.

Runck, A. M., M. D. Matocq, et al. (2009). "Historic hybridization and persistence of a novel mito-nuclear combination in red-backed voles (genus *Myodes*)." BMC evolutionary biology **9**(114): 1-15.

Sanchez-Gracia, A. and J. Castresana (2012). "Impact of Deep Coalescence on the Reliability of Species Tree Inference from Different Types of DNA Markers in Mammals." PloS one **7**(1): e30239.

Serdyuk, N. V. and A. S. Tesakov (2006). "New form of rhizodont voles (Rodentia, Arvicolinae, Clethrionomyini) from Pleistocene of Central Altai (Russia)." Russian Journal of Theriology **5**(2): 79-83.

Silberfeld, T., J. W. Leigh, et al. (2010). "A multi-locus time-calibrated phylogeny of the brown algae (Heterokonta, Ochrophyta, Phaeophyceae): Investigating the evolutionary nature of the "brown algal crown radiation"." Molecular phylogenetics and evolution **56**(2): 659-674.

Stanhope, M. J., M. R. Smith, et al. (1996). "Mammalian evolution and the interphotoreceptor retinoid binding protein (IRBP) gene: convincing evidence for several superordinal clades." Journal of Molecular Evolution **43**(2): 83-92.

Stephens, M., N. J. Smith, et al. (2001). "A new statistical method for haplotype reconstruction from population data." American Journal of Human Genetics **68**(4): 978-989.

Tamura, K., D. Peterson, et al. (2011). "MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods." Molecular biology and evolution **28**(10): 2731-2739.

Tarasov, P. E., V. S. Volkova, et al. (2000). "Last glacial maximum biomes reconstructed from pollen and plant macrofossil data from northern Eurasia." Journal of Biogeography **27**(3): 609-620.

Tegelström, H. (1987). "Transfer of mitochondrial DNA from the northern red-backed vole (*Clethrionomys rutilus*) to the bank vole (*C. glareolus*)." Journal of Molecular Evolution **24**(3): 218-227.

Tesakov, A. S., V. S. Lebedev, et al. (2010). "*Clethrionomys* Tilesius, 1850 is the valid generic name for red-backed voles and *Myodes* Pallas, 1811 is a junior synonym of *Lemmus* Link, 1795." Russian Journal of Theriology **9**(2): 83-86.

Tolley, K. A., B. M. Chase, et al. (2008). "Speciation and radiations track climate transitions since the Miocene Climatic Optimum: a case study of southern African chameleons." Journal of Biogeography **35**(8): 1402-1414.

Velichko, A. A. and H. E. Wright (2005). Cenozoic climatic and environmental changes in Russia, Geological Society of Amer.

Webb III, T. and P. Bartlein (1992). "Global changes during the last 3 million years: climatic controls and biotic responses." Annual Review of Ecology and Systematics **23**: 141-173.

Zhang, Y. and S. Ge (2007). "Molecular evolution study in China: progress and future promise." Philosophical Transactions of the Royal Society B: Biological Sciences **362**(1482): 973-986.

Zheng, S. and C. Li (1990). Comments on fossil arvicolids of China. Int. Symp. Evol. Phyl. Biostr. Arvicolids.

Zhisheng, A., J. E. Kutzbach, et al. (2001). "Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since Late Miocene times." Nature **411**(6833): 62-66.

Figures and Tables

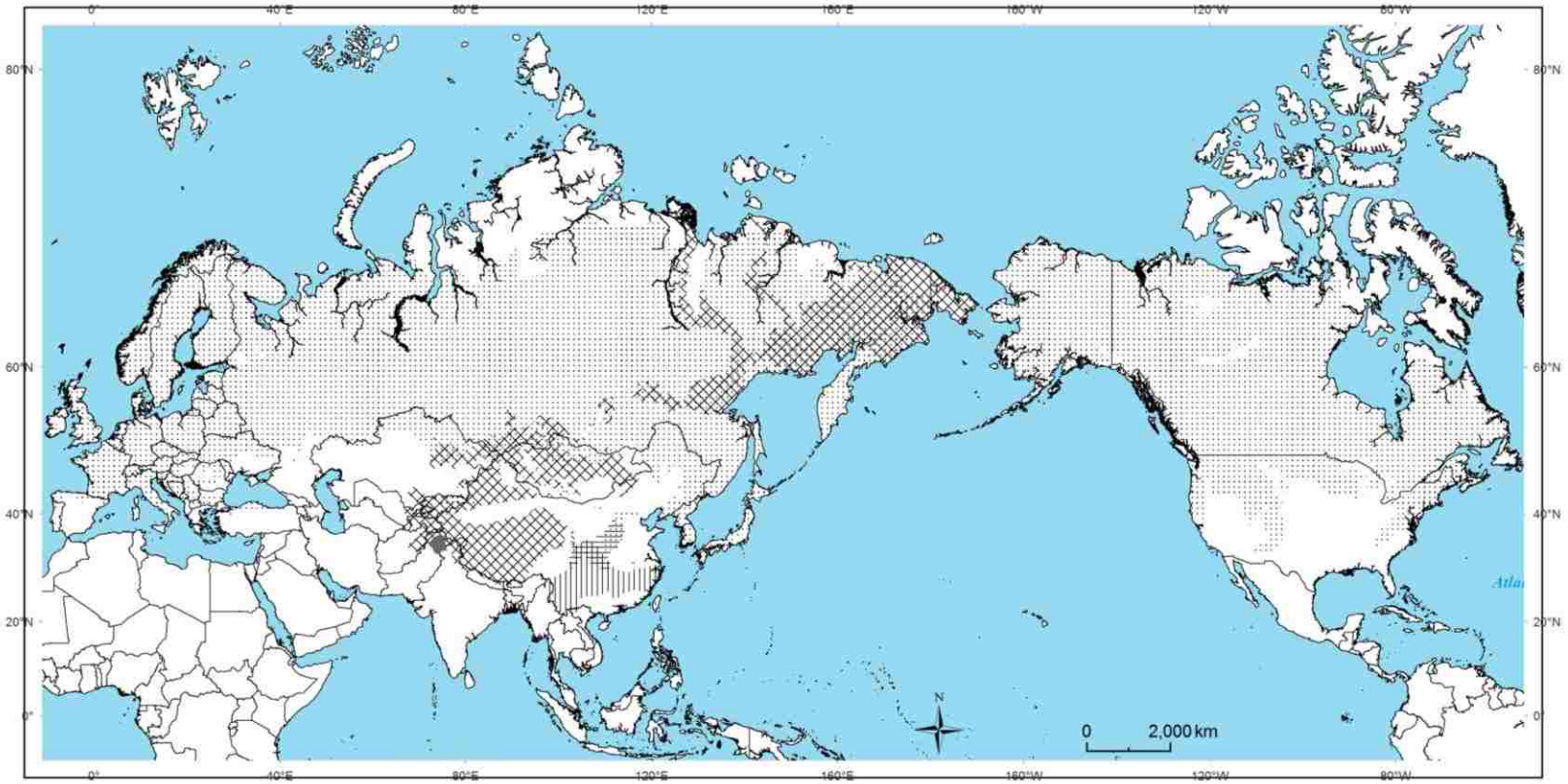
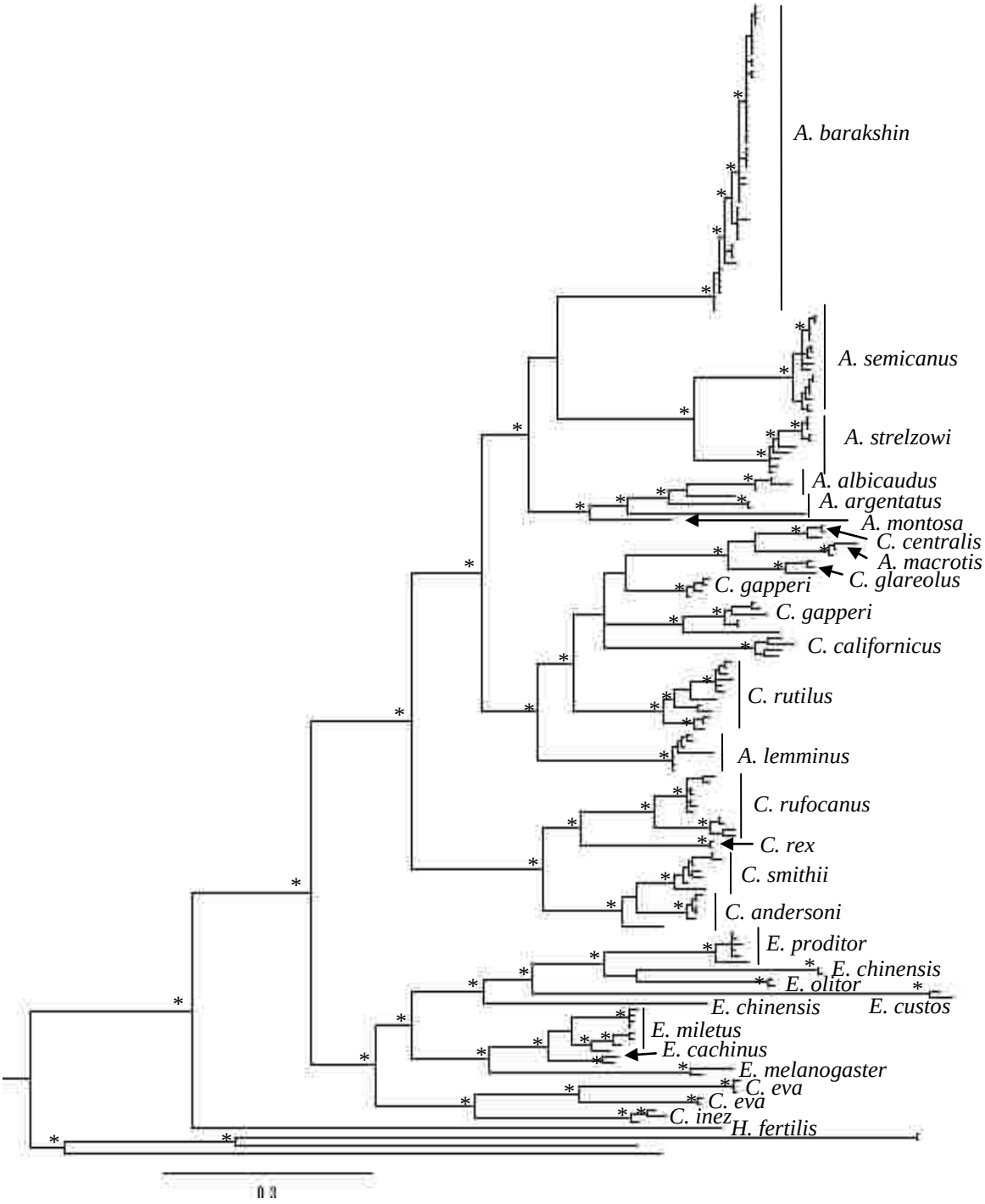


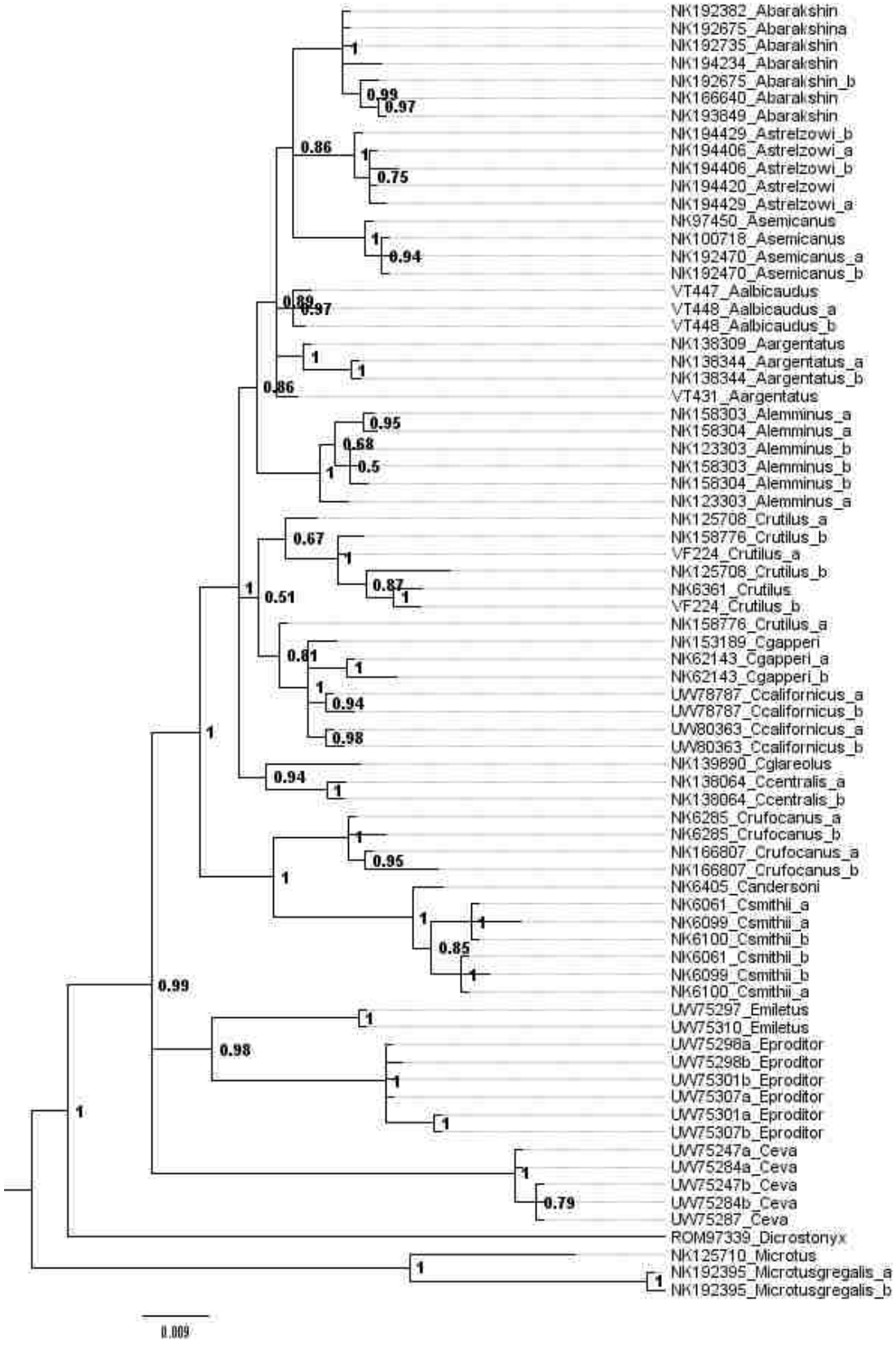
Figure 1. Distribution of the 5 nominal genera of Clethrionomyini (Source: IUCN): *Alticola* (cross-hatching), central to northeast Asia; *Clethrionomys* (stippling), Europe to North America; *Eothenomys* (vertical lines), south-central China; *Caryomys* (horizontal lines), central China; and *Hyperacrius* (gray shading), northern India and Pakistan.



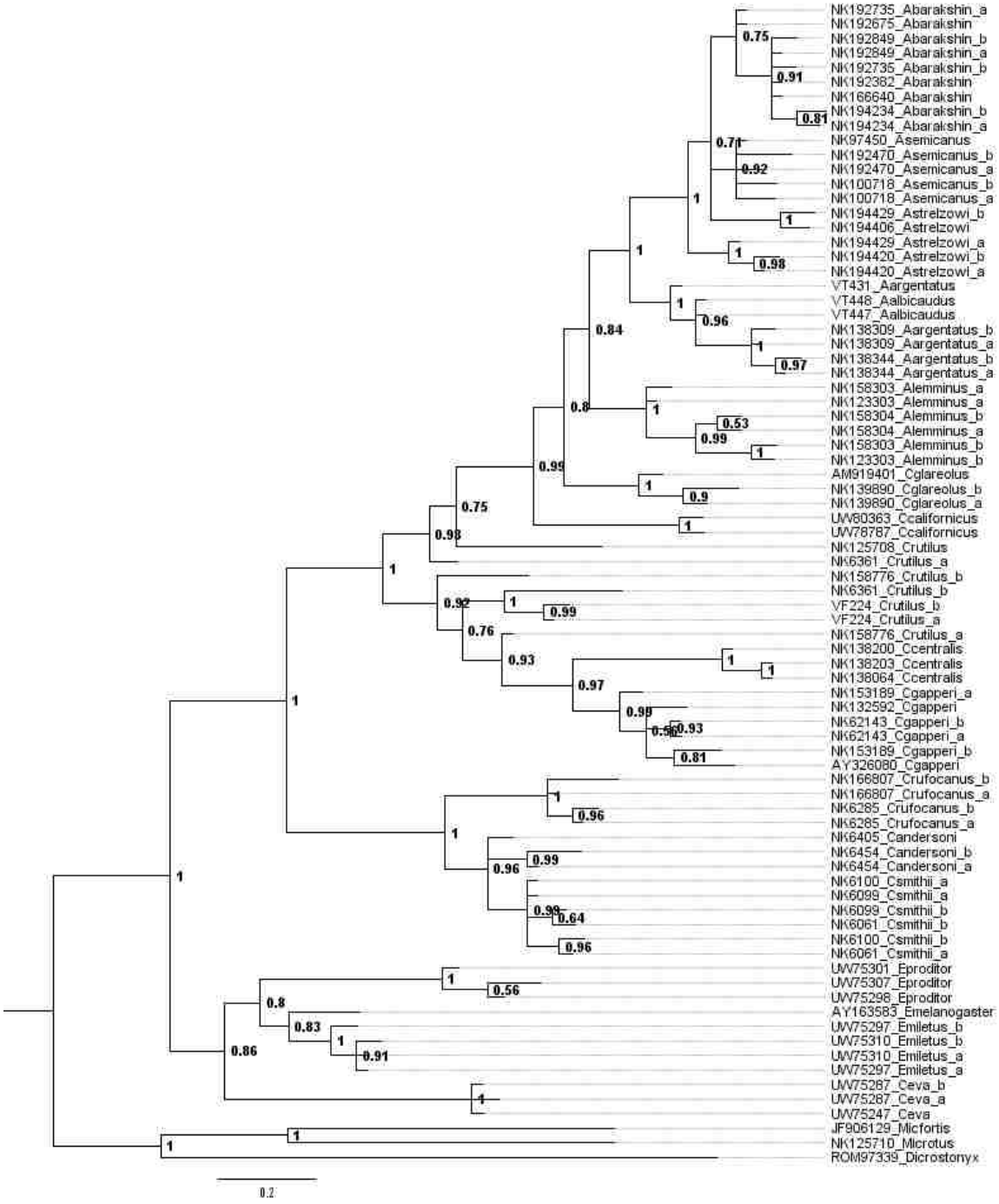
A) Cyt b

Figure 2. Bayesian gene trees for four independent loci from up to 28 species of Clethrionomyini. Cyt b was partitioned by codon position. An asterisk indicates posterior probability of a node >0.9 if values are not presented at nodes. A) Cytochrome b, B) ETS2, C) IRBP, D) G6pd.

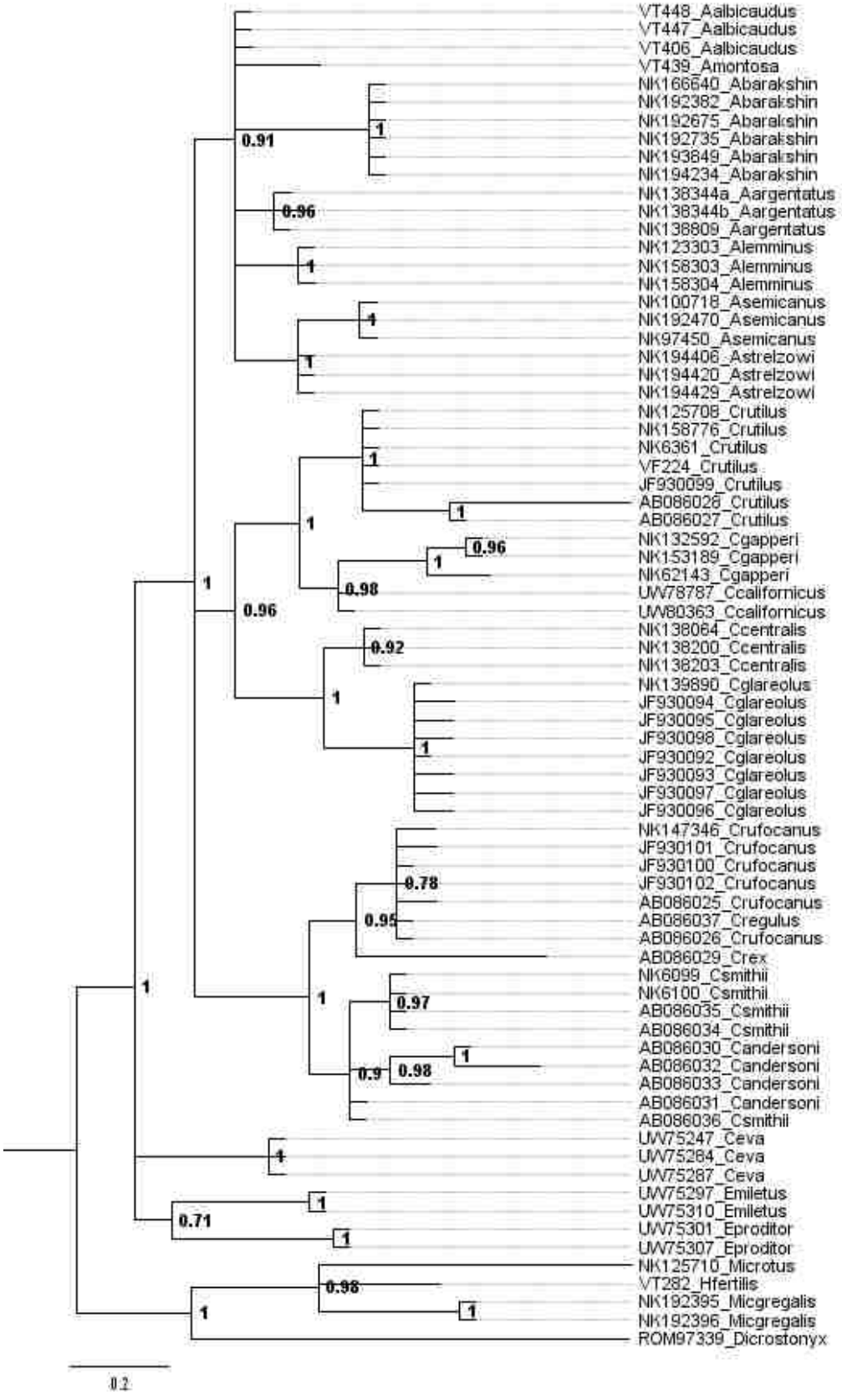
Figure 2 cont. B) ETS2



Figured 2 cont. C) IRBP



Figured 2 cont. D) G6pd



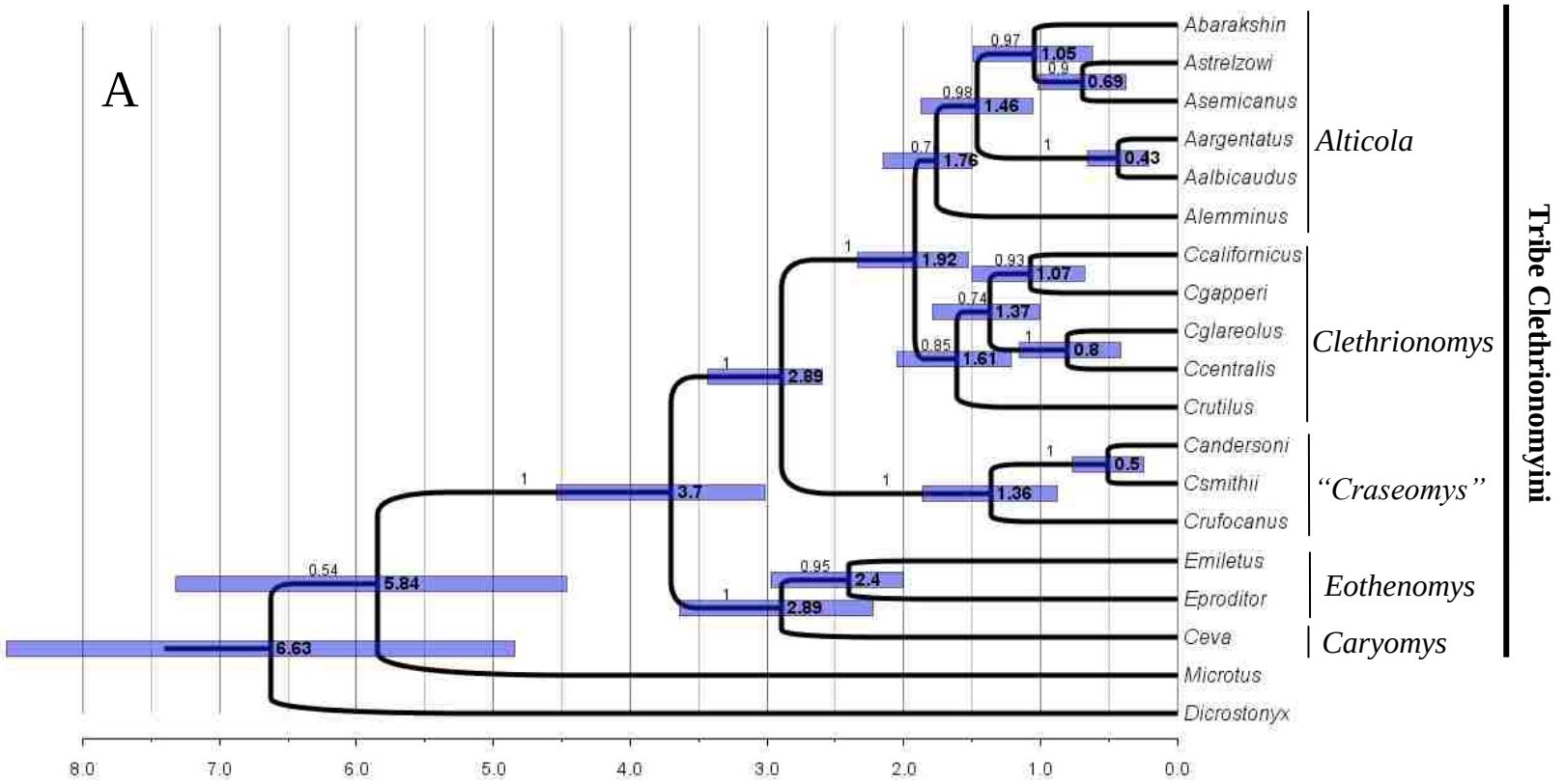
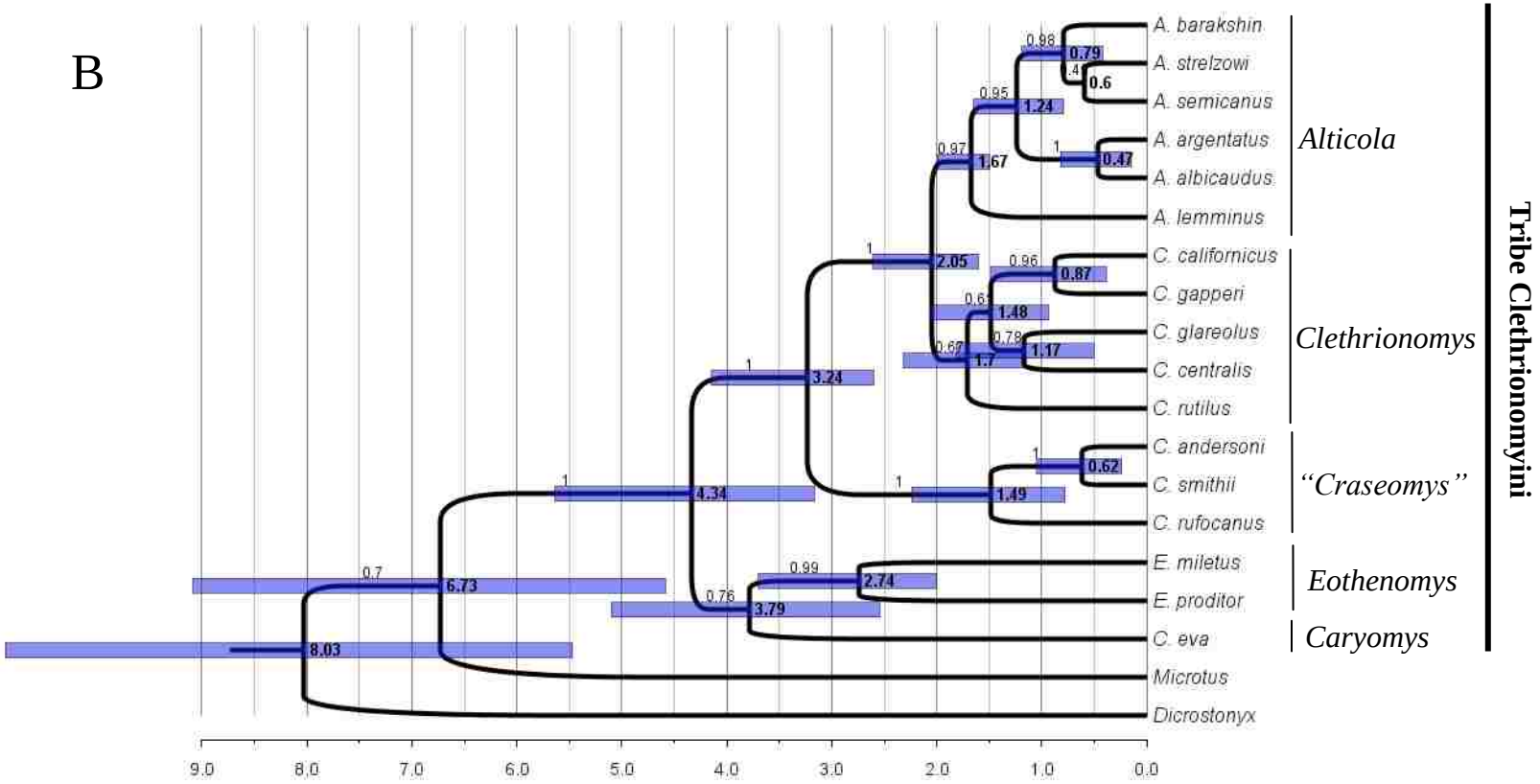


Figure 3. Species tree for 17 species of Clethrionomyini based on A) 3 nuclear genes and Cyt b or B) 3 nuclear genes, constructed using *BEAST to estimate divergence dates. *Microtus* and *Dicrostonyx* are included as outgroups. A time scale is included showing past to present (left to right) in millions of years. Divergence dates are shown to the right of nodes within the 95% confidence intervals for date estimates (blue bars). Posterior probability is reported above branches.

Figure 3 cont.



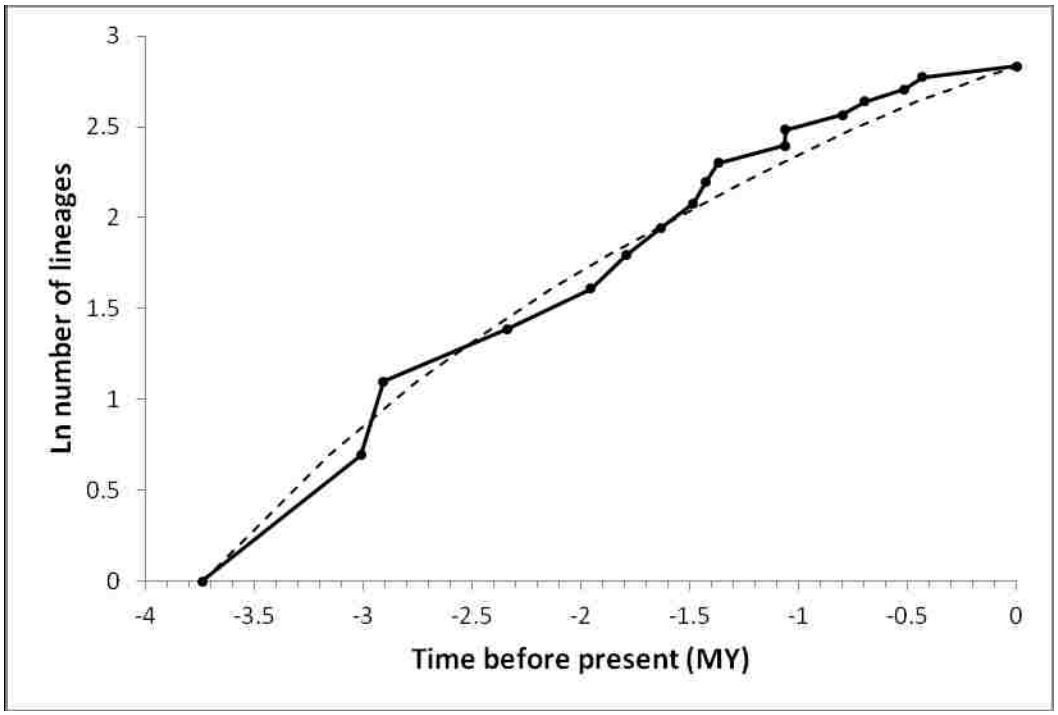


Figure 4. Lineage through time plot. The solid line and points represent the rate of lineage accumulation over time for Clethrionomyini as derived from a species tree that included only ingroup taxa. The dotted line represents the expected constant rate curve, as determined by 1000 simulated trees generated by a process that accounts for missing taxa.

Table 1. Genetic locus information and sampling. Sample size (Clethrionomyini only, N), number of segregating sites (S), number of haplotypes (H), haplotype diversity (Hd), nucleotide diversity (π), and model of evolution as determined by ModelTest are included for each gene for all species.

Locus	**N (gene trees)**	**N (species trees)**	**length (bp)**	**Genomic location**	**S**	**H**	**Hd**	**π**	**Model of Evolution**
G6pd	69	58	587	X chromosome	39	26	0.953	0.0176	K80 + Γ
ETS2	45	44	893	nuclear exon	102	46	0.983	0.0263	HKY+Γ
IRBP	53	50	1059	nuclear exon	58	42	0.978	0.0228	HKY+I+Γ
Cyt b	189	65	671-1143	MtDNA	229	124	0.992	0.0838	GTR+I+Γ
pos1									SYM+I+Γ
pos2									HKY+I+Γ
pos3									GTR+I+Γ

Table 2. Range of average pairwise divergence among species, as compared within and between genera of Clethrionomyini. Gene tree sampling sets for each of the four genes studied were used in the analysis. A dash indicates that less than two species of the genus were successfully sequenced. *Clethrionomys* excludes what we propose as "*Craseomys*"; this newly proposed genus is shown separately. *Hyperacrius* is excluded from within genera comparisons because only one species was included in our analysis. All standard errors less than 0.01 except in cases involving *H. fertilis*.

Average divergence comparison	Cyt b	G6pd	ETS2	IRBP
Within genera				
Alticola	0.038-0.112	0.003-0.016	0.008-0.018	0.003-0.014
Caryomys	0.086	-	-	-
"Craseomys"	0.026-0.064	0.001-0.020	0.006-0.027	0.004-0.012
Clethrionomys	0.049-0.078	0.008-0.024	0.007-0.020	0.014-0.025
Eothenomys	0.028-0.117	0.019	0.032	0.012
Between genera				
Alticola/Caryomys	0.096-0.122	0.015-0.023	0.045-0.056	0.030-0.034
Alticola/Clethrionomys	0.051-0.104	0.009-0.025	0.011-0.022	0.012-0.032
Alticola/"Craseomys"	0.074-0.110	0.012-0.031	0.027-0.037	0.027-0.034
Alticola/Eothenomys	0.089-0.139	0.016-0.026	0.029-0.042	0.028-0.039
Clethrionomys/Caryomys	0.078-0.113	0.019-0.027	0.050-0.055	0.028-0.036
Clethrionomys/Eothenomys	0.085-0.132	0.021-0.030	0.031-0.042	0.025-0.040
Clethrionomys/"Craseomys"	0.076-0.100	0.017-0.034	0.027-0.040	0.024-0.033
Caryomys/Eothenomys	0.092-0.123	0.019-0.021	0.055-0.057	0.021-0.031
Caryomys/"Craseomys"	0.078-0.102	0.020-0.031	0.057-0.061	0.37
Eothenomys/"Craseomys"	0.085-0.134	0.022-0.035	0.038-0.049	0.028-0.039
Hyperacrius/Alticola	0.119-0.150	0.029-0.037	-	-
Hyperacrius/Clethrionomys	0.125-0.142	0.035-0.043	-	-
Hyperacrius/"Craseomys"	0.122-0.137	0.037-0.044	-	-
Hyperacrius/Caryomys	0.118-0.126	0.033	-	-
Hyperacrius/Eothenomys	0.120-0.150	0.037-0.045	-	-

Table 3. Uncorrected pairwise sequence divergence of four genes (A. Cyt b, B. G6pd, C. ETS2, D. IRBP) between species of *Alticola* and *Clethrionomys*. Solid lines demarcate current genus boundaries; dotted lines demarcate proposed genus boundaries (*Craseomys*). Species are named with the first three letters of their specific epithet, as identified in the Appendix.

A.

Cyt b	A alb.	A arg.	A bar.	A lem.	A mac.	A mon.	A sem.	A str.	C cal.	C cen.	C gap.	C gla.	C rut.	Cr rex	Cr ruf.	Cr and.	Cr smi.
A alb.																	
A arg.	.038																
A bar.	.069	.078															
A lem.	.088	.090	.088														
A mac.	.098	.103	.102	.094													
A mon.	.060	.058	.078	.081	.091												
A sem.	.082	.086	.086	.091	.097	.078											
A str.	.080	.089	.088	.095	.112	.096	.052										
C cal.	.100	.097	.097	.074	.090	.090	.096	.102									
C cen.	.098	.096	.087	.078	.051	.096	.088	.104	.064								
C gap.	.082	.087	.083	.067	.083	.078	.084	.093	.069	.068							
C gla.	.092	.095	.091	.075	.062	.086	.090	.103	.078	.049	.064						
C rut.	.092	.092	.088	.060	.093	.089	.096	.098	.068	.068	.064	.074					
Cr rex	.088	.081	.074	.089	.098	.076	.076	.087	.076	.085	.077	.088	.087				
Cr ruf.	.092	.091	.090	.081	.108	.083	.092	.102	.095	.090	.084	.093	.088	.059			
Cr and.	.091	.096	.087	.092	.106	.096	.101	.105	.093	.100	.092	.095	.093	.061	.063		
Cr smi.	.089	.093	.081	.091	.110	.095	.099	.098	.090	.097	.092	.099	.092	.057	.064	.026	

Table 3 cont.

B.

G6pd	A alb.	A arg.	A bar.	A lem.	A mon.	A sem.	A str.	C cal.	C cen.	C gap.	C gla.	C rut.	Cr reg.	Cr rex	Cr ruf.	Cr and.	Cr smi.
A alb.																	
A arg.	.003																
A bar.	.009	.011															
A lem.	.004	.006	.012														
A mon.	.005	.008	.014	.009													
A sem.	.008	.009	.016	.007	.010												
A str.	.004	.006	.012	.007	.007	.003											
C cal.	.009	.011	.017	.012	.014	.016	.012										
C cen.	.011	.013	.019	.014	.016	.017	.014	.012									
C gap.	.015	.019	.025	.020	.022	.023	.020	.008	.020								
C gla.	.015	.016	.022	.017	.019	.020	.017	.017	.008	.024							
C rut.	.014	.016	.022	.017	.019	.021	.017	.008	.017	.015	.022						
Cr reg.	.013	.014	.021	.016	.017	.019	.016	.017	.019	.025	.022	.022					
Cr rex	.022	.025	.031	.026	.028	.029	.026	.028	.029	.034	.032	.033	.014				
Cr ruf.	.013	.015	.021	.016	.018	.020	.016	.018	.020	.026	.023	.023	.001	.015			
Cr and.	.015	.017	.024	.019	.020	.022	.019	.020	.022	.028	.025	.025	.010	.020	.010		
Cr smi.	.012	.014	.020	.015	.017	.019	.015	.017	.019	.025	.022	.022	.007	.017	.007	.006	

Table 3 cont.

C.

ETS2	A alb.	A arg.	A bar.	A lem.	A sem.	A str.	C cal.	C cen.	C gap.	C gla.	C rut.	Cr ruf.	Cr and.	Cr smi.
A alb.														
A arg.	0.008													
A bar.	0.009	0.011												
A lem.	0.011	0.013	0.016											
A sem.	0.012	0.014	0.014	0.018										
A str.	0.011	0.012	0.013	0.017	0.015									
C cal.	0.011	0.011	0.016	0.015	0.018	0.017								
C cen.	0.015	0.013	0.018	0.018	0.021	0.020	0.016							
C gap.	0.014	0.014	0.018	0.018	0.019	0.019	0.007	0.018						
C gla.	0.015	0.016	0.021	0.020	0.021	0.022	0.018	0.016	0.019					
C rut.	0.012	0.013	0.017	0.016	0.020	0.019	0.011	0.016	0.014	0.020				
Cr ruf.	0.028	0.027	0.030	0.030	0.028	0.031	0.027	0.030	0.030	0.032	0.029			
Cr and.	0.030	0.031	0.034	0.036	0.032	0.035	0.033	0.031	0.035	0.036	0.035	0.024		
Cr smi.	0.032	0.033	0.035	0.037	0.033	0.036	0.035	0.035	0.037	0.040	0.036	0.027	0.006	

D.

IRBP	A alb.	A arg.	A bar.	A lem.	A sem.	A str.	C cal.	C cen.	C gap.	C gla.	C rut.	Cr ruf.	Cr and.	Cr smi.
A alb.														
A arg.	0.003													
A bar.	0.009	0.012												
A lem.	0.011	0.013	0.014											
A sem.	0.009	0.012	0.005	0.014										
A str.	0.009	0.012	0.007	0.012	0.007									
C cal.	0.016	0.019	0.020	0.017	0.019	0.019								
C cen.	0.027	0.030	0.032	0.029	0.032	0.031	0.029							
C gap.	0.024	0.027	0.029	0.026	0.029	0.028	0.025	0.014						
C gla.	0.012	0.014	0.014	0.014	0.014	0.015	0.017	0.023	0.024					
C rut.	0.019	0.022	0.023	0.020	0.022	0.022	0.019	0.020	0.015	0.018				
Cr ruf.	0.030	0.033	0.034	0.031	0.033	0.033	0.033	0.030	0.029	0.027	0.027			
Cr and.	0.027	0.030	0.031	0.029	0.030	0.030	0.030	0.029	0.028	0.024	0.025	0.012		
Cr smi.	0.028	0.031	0.032	0.029	0.031	0.031	0.031	0.030	0.028	0.024	0.026	0.012	0.004	

Appendices

Appendix A. Specimens examined, their museum catalog numbers, specimen ID numbers, and associated location information. Abbreviations for collections are as follows: MSB, Museum of Southwestern Biology; UWBM, University of Washington Burke Museum; VF, Vadim Fedorov personal collection; ROM, Royal Ontario Museum.

Species	Collection	Catalog number	Sample ID	Country	State/ Province	Locality	Latitude	Longitude
Alticola albicaudus	Vermont U	VT406	VT406	Pakistan		Northern Areas, Trunken, 3550 m	36.8	75.13
Alticola albicaudus	Vermont U	VT447	VT447	Pakistan		Northern Areas, Trunken, 3500 m	36.8	75.13
Alticola albicaudus	Vermont U	VT448	VT448	Pakistan		Northern Areas, Trunken, 3500 m	36.8	75.13
Alticola argentatus	MSB	158760	NK138344	China	Xinjiang	3.9 km N., 28.3 km E. Narati	43.3671	84.3595
Alticola argentatus	MSB	158797	NK138309	China	Xinjiang	5.6 km N., 29.9 km E. Narati	43.3822	84.3794
Alticola argentatus blandfordi (?)	Vermont U	VT431	VT431	Pakistan		NWFP, Chitral District, Shandur Pass, 3700m	36.06	72.46
Alticola barakshin	MSB	199767	NK165544	Mongolia	Omnogovi	Gobi Gurvan Saikhan National Park; 4.0 km N, 5.0 km W of Yolin Am	43.52085	104.02997
Alticola barakshin	MSB	215004	NK165614	Mongolia	Omnogovi	Gobi Gurvan Saikhan National Park; 4.0 km N, 5.0 km W of Yolin Am	43.52085	104.02997
Alticola barakshin	MSB	215005	NK165609	Mongolia	Omnogovi	Gobi Gurvan Saikhan National Park; 4.0 km N, 5.0 km W of Yolin Am	43.52085	104.02997
Alticola barakshin	MSB	215032	NK166613	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Tsagaan Ovoo Uul	43.61791667	103.75455
Alticola barakshin	MSB	215074	NK166773	Mongolia	Bayanhongor Prov.	Ikh Bogd Forest; Tsagaan Khoshoot; camp	44.98051667	100.4407333
Alticola barakshin	MSB	215124	NK166818	Mongolia	Bayanhongor Prov.	Ikh Bogd Forest	44.97755	100.3342
Alticola barakshin	MSB	215134	NK166769	Mongolia	Bayanhongor Prov.	Ikh Bogd Forest; Tsagaan Khoshoot; camp	44.98051667	100.4407333
Alticola barakshin	MSB	215153	NK165665	Mongolia	Omnogovi	Gobi Gurvan Saikhan National Park; 4.0 km N, 5.0 km W of Yolin Am	43.52085	104.02997
Alticola barakshin	MSB	215216	NK165571	Mongolia	Omnogovi	Gobi Gurvan Saikhan National Park; 4.0 km N, 5.0 km W of Yolin Am	43.52085	104.02997
Alticola barakshin	MSB	215225	NK165569	Mongolia	Omnogovi	Gobi Gurvan Saikhan National Park; 4.0 km N, 5.0 km W of Yolin Am	43.52085	104.02997
Alticola barakshin	MSB	215227	NK166584	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Tsagaan Ovoo Uul	43.61791667	103.75455
Alticola barakshin	MSB	215236	NK166568	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Tsagaan Ovoo Uul	43.61791667	103.75455
Alticola barakshin	MSB	215278	NK165667	Mongolia	Omnogovi	Gobi Gurvan Saikhan National Park; 4.0 km N, 5.0 km W of Yolin Am	43.52085	104.02997
Alticola barakshin	MSB	215280	NK166617	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Tsagaan Ovoo Uul	43.61791667	103.75455

Alticola barakshin	MSB	215287	NK165689	Mongolia	Omnogovi	Gobi Gurvan Saikhan National Park; 4.0 km N, 5.0 km W of Yolin Am	43.52085	104.02997
Alticola barakshin	MSB	215442	NK166638	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Baruun Saikhan Vul	43.82028333	103.2809167
Alticola barakshin	MSB	215489	NK166640	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Baruun Saikhan Vul	43.82028333	103.2809167
Alticola barakshin	MSB	215558	NK166611	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Tsagaan Ovoo Uul	43.61791667	103.75455
Alticola barakshin	MSB	215759	NK166605	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Tsagaan Ovoo Uul	43.61791667	103.75455
Alticola barakshin	MSB	215760	NK166601	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Tsagaan Ovoo Uul	43.61791667	103.75455
Alticola barakshin	MSB	215859	NK166637	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Baruun Saikhan Vul	43.82028333	103.2809167
Alticola barakshin	MSB	215951	NK165715	Mongolia	Omnogovi	Gobi Gurvan Saikhan National Park; 1.6 km E of Yolin Am	43.48547	104.0804
Alticola barakshin	MSB	215997	NK166642	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Baruun Saikhan Vul	43.82028333	103.2809167
Alticola barakshin	MSB	216024	NK166641	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Baruun Saikhan Vul	43.82028333	103.2809167
Alticola barakshin	MSB	216051	NK166639	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Baruun Saikhan Vul	43.82028333	103.2809167
Alticola barakshin	MSB	216133	NK165711	Mongolia	Omnogovi	Gobi Gurvan Saikhan National Park; 1.6 km E of Yolin Am	43.48547	104.0804
Alticola barakshin	MSB	216196	NK166517	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Tsagaan Ovoo Uul	43.61791667	103.75455
Alticola barakshin	MSB	216202	NK166723	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Baruun Saikhan Vul	43.82028333	103.2809167
Alticola barakshin	MSB	216206	NK166727	Mongolia	Omnogovi Prov.	Gobi Gurvan Saikhan National Park; Baruun Saikhan Vul	43.82028333	103.2809167
Alticola barakshin	MSB	227105	NK192359	Mongolia	Gobi Altai	Gichgene Nuruu, Davaany Ovor, Emeeltseg Nuruu.	45.30331	97.70625
Alticola barakshin	MSB	227164	NK192360	Mongolia	Gobi Altai	Gichgene Nuruu, Davaany Ovor, Emeeltseg Nuruu.	45.30331	97.70625
Alticola barakshin	MSB	230764	NK192382	Mongolia	Gobi Altai	Gichgene Nuruu, Davaany Ovor, Emeeltseg Nuruu.	45.30331	97.70625
Alticola barakshin	MSB	233789	NK192628	Mongolia	Shinejinst Soum	Tsagaan Bogd Uul; Suuj Bulag	42.882111	98.818417
Alticola barakshin	MSB	248930	NK192629	Mongolia	Shinejinst Soum	Tsagaan Bogd Uul; Suuj Bulag	42.882111	98.818417
Alticola barakshin	MSB	233804	NK192630	Mongolia	Shinejinst Soum	Tsagaan Bogd Uul; Suuj Bulag	42.882111	98.818417
Alticola barakshin	MSB	233842	NK192674	Mongolia	Altai Soum	Aj Bogd Uul; Bayan Gol	44.83346	95.27166
Alticola barakshin	MSB	233813	NK192675	Mongolia	Altai Soum	Aj Bogd Uul; Bayan Gol	44.83346	95.27166
Alticola barakshin	MSB	233658	NK192677	Mongolia	Altai Soum	Aj Bogd Uul; Bayan Gol	44.83346	95.27166
Alticola barakshin	MSB	233659	NK192679	Mongolia	Altai Soum	Aj Bogd Uul; Bayan Gol	44.83346	95.27166
Alticola barakshin	MSB	233812	NK192735	Mongolia	Altai Soum	Aj Bogd Uul; Bayan Gol	44.83346	95.27166

Alticola barakshin	MSB	233759	NK192736	Mongolia	Altai Soum	Aj Bogd Uul; Bayan Gol	44.83346	95.27166
Alticola barakshin	MSB	233696	NK192737	Mongolia	Altai Soum	Aj Bogd Uul; Bayan Gol	44.83346	95.27166
Alticola barakshin	MSB	233686	NK192741	Mongolia	Altai Soum	Aj Bogd Uul; Bayan Gol	44.83346	95.27166
Alticola barakshin	MSB	233760	NK192742	Mongolia	Altai Soum	Aj Bogd Uul; Bayan Gol	44.83346	95.27166
Alticola barakshin	MSB	233761	NK192743	Mongolia	Altai Soum	Aj Bogd Uul; Bayan Gol	44.83346	95.27166
Alticola barakshin	MSB	233694	NK193531	Mongolia	Bugat Soum	Bayangol Havtsal	45.50748	94.33577
Alticola barakshin	MSB	233836	NK193849	Mongolia	Bugat Soum	Bij Gol	45.570333	93.820222
Alticola barakshin	MSB	233674	NK193892	Mongolia	Altai Soum	Baitag Bogd Uul; Nariin Hargait Gol	45.26446	91.06402
Alticola barakshin	MSB	233821	NK193997	Mongolia	Altai Soum	Baitag Bogd Uul; Nariin Hargait Gol	45.26446	91.06402
Alticola barakshin	MSB	233811	NK194234	Mongolia	Altai Soum	Baitag Bogd Uul; Buduun Hargait	45.26821	90.95323
Alticola lemminus	MSB	144872	NK123303	Russia	Republic of Sakha	Stokovo station	61.8472	147.662
Alticola lemminus	MSB	149891	NK158303	Russia	Chukotka Autonomous Okrug	Upper Anadyr River	66.76666667	169.5666667
Alticola lemminus	MSB	149892	NK158304	Russia	Chukotka Autonomous Okrug	Upper Anadyr River	66.76666667	169.5666667
Alticola lemminus	MSB	149895	NK158305	Russia	Chukotka Autonomous Okrug	Upper Anadyr River	66.76666667	169.5666667
Alticola lemminus	MSB	149897	NK158306	Russia	Chukotka Autonomous Okrug	Upper Anadyr River	66.76666667	169.5666667
Alticola montosa	Vermont U	VT439	VT439	Pakistan		NWFP, Hazar District, Besal, 3400 m	35.07	73.95
Alticola semicanus	MSB	94156	NK100600	Mongolia	Ovorkhangay	Ulaan Tsutgalan; 46D 47' 13" N, 101D 57' 47" E.	46.786944	101.963056
Alticola semicanus	MSB	94157	NK100602	Mongolia	Ovorkhangay	Ulaan Tsutgalan; 46D 47' 13" N, 101D 57' 47" E.	46.786944	101.963056
Alticola semicanus	MSB	94170	NK100683	Mongolia	Ovorkhangay	Ulaan Tsutgalan; 46D 47' 13" N, 101D 57' 47" E.	46.786944	101.963056
Alticola semicanus	MSB	94171	NK100684	Mongolia	Ovorkhangay	Ulaan Tsutgalan; 46D 47' 13" N, 101D 57' 47" E.	46.786944	101.963056
Alticola semicanus	MSB	94184	NK100716	Mongolia	Ovorkhangay	Khetsuugiin Ovor; 46D 35' 56" N, 102D 43' 07" E.	46.598889	102.718611
Alticola semicanus	MSB	94186	NK100718	Mongolia	Ovorkhangay	Khetsuugiin Ovor; 46D 35' 56" N, 102D 43' 07" E.	46.598889	102.718611
Alticola semicanus	MSB	98508	NK97450	Mongolia	Hovsgol	Hovsgol National park; 1 km S. of Hangard tourist camp; 50D 30'32"N, 100D 09'19"E	50.508889	100.155278
Alticola semicanus	MSB	98508	NK97450	Mongolia	Hovsgol	Hovsgol National park; 1 km S. of Hangard tourist camp; 50D 30'32"N, 100D 09'19"E	50.508889	100.155278

Alticola semicanus	MSB	98510	NK97452	Mongolia	Hovsgol	Hovsgol National park; 1 km S. of Hangard tourist camp; 50D 30'32"N, 100D 09'19"E	50.508889	100.155278
Alticola semicanus	MSB	98511	NK97453	Mongolia	Hovsgol	Hovsgol National park; 1 km S. of Hangard tourist camp; 50D 30'32"N, 100D 09'19"E	50.508889	100.155278
Alticola semicanus	MSB	98512	NK97454	Mongolia	Hovsgol	Hovsgol National park; 1 km S. of Hangard tourist camp; 50D 30'32"N, 100D 09'19"E	50.508889	100.155278
Alticola semicanus	MSB	230791	NK192470	Mongolia	Tov	40 km SW Altanbulag, Ivgeelt Mtn.	47.43276	106.0416
Alticola semicanus	MSB	230813	NK192487	Mongolia	Tov	40 km SW Altanbulag, Ivgeelt Mtn.	47.43276	106.0416
Alticola semicanus	MSB	230902	NK192468	Mongolia	Tov	40 km SW Altanbulag, Ivgeelt Mtn.	47.43276	106.0416
Alticola semicanus	MSB	230903	NK192490	Mongolia	Tov	40 km SW Altanbulag, Ivgeelt Mtn.	47.43276	106.0416
Alticola strelzovi	MSB	233729	NK194355	Mongolia	Bulgan Soum	Chachirt Belchir	46.67569	91.40286
Alticola strelzovi	MSB	233671	NK194406	Mongolia	Bulgan Soum	Chachirt Belchir	46.67569	91.40286
Alticola strelzovi	MSB	233670	NK194420	Mongolia	Bulgan Soum	S of Chachirt Belchir	46.67425	91.396472
Alticola strelzovi	MSB	233693	NK194422	Mongolia	Bulgan Soum	S of Chachirt Belchir	46.67425	91.396472
Alticola strelzovi	MSB	233728	NK194429	Mongolia	Bulgan Soum	Chachirt Belchir	46.67569	91.40286
Caryomys eva	UWBM	75247	UW75247	China	Sichuan Sheng	Jin-Hou-Feng, Xi-Ling-Xue-Shan		
Caryomys eva	UWBM	75284	UW75284	China	Sichuan Sheng	Hong-Shi-Bao, Xi-Ling-Xue-Shan		
Caryomys eva	UWBM	75287	UW75287	China	Sichuan Sheng	Hong-Shi-Bao, Xi-Ling-Xue-Shan		
Clethrionomys andersoni	MSB	45692	NK6405	Japan	Honshu	3.2 km E Zao Onsen	38.16667	140.43207
Clethrionomys andersoni	MSB	45695	NK6404	Japan	Honshu	3.2 km E Zao Onsen	38.16667	140.43207
Clethrionomys andersoni	MSB	45699	NK6454	Japan	Honshu	3.2 km E Zao Onsen	38.16667	140.43207
Clethrionomys californicus	UWBM	78787	UW78787	USA	Oregon	Near Bend, Cascade Lakes Hwy, 3.7 mi W of FR370; creek to Devils Lk		
Clethrionomys californicus	UWBM	80363	UW80363	USA	Oregon	Mt. Hood, Government Camp, FR2645 2.7mi SW of Timberline Hwy jct; Still Cr.		
Clethrionomys californicus	UWBM	80365	UW80365	USA	Oregon	Mt. Hood, Government Camp, FR2645 2.7mi SW of Timberline Hwy jct; Still Cr.		
Clethrionomys centralis	MSB	158500	NK138200	China	Xinjiang	24.5 km E., 7.9 km S. Narati	43.2439	84.31
Clethrionomys centralis	MSB	158508	NK138203	China	Xinjiang	24.3 km E., 8 km S. Narati	43.2426	84.3078
Clethrionomys gapperi	MSB	70688	NK62143	United States	Pennsylvania	Allegheny Portage Railrd. 40D27'30", 78D33'00". 85; 40D27'78D33'	40.45833	-78.55
Clethrionomys gapperi	MSB	147281	NK132592	Canada	British Columbia	Cassiar Highway (Hwy 37), Devil Creek Crossing	57.08133333	-130.26443

Clethrionomys gapperi	MSB	193266	NK153189	United States	New Mexico	Santa Fe Mtns., 1.5 km N. Elk Mountain	35.78184	-105.55959
Clethrionomys glareolus	MSB	193928	NK139890	Russia	Saratov oblast	Slavianka	51.56666667	46.03333333
Clethrionomys rufocanus	MSB	45581	NK147346	Japan	Hokkaido	Rumoi Prov, 9 km N Toikanbetsu, Hokkaido U Experim. Forest	44.99488	142.02389
Clethrionomys rufocanus	MSB	153616	NK166807	Russia		Khabarovsk	48.3	135.05
Clethrionomys rufocanus	MSB	215593	NK6285	Mongolia	Bayanhongor	Ikh Bogd Forest	44.97755	100.3342
Clethrionomys rutilus	MSB	45628	NK6361	Japan	Hokkaido	Rumoi Prov, 8.2 km N, 1.3 km W Toikanbetsu, Hokkai	44.98768	142.00743
Clethrionomys rutilus	MSB	149289	NK125708	Canada	Yukon Territory	Campell Hwy (Hwy #4); 3km N and 7.5km E of Carmacks	62.1159	-136.14425
Clethrionomys rutilus	MSB	193200	NK158776	Russia	Khabarovskii krai	Amur River; Dachnaya	48.46666667	134.9333333
Clethrionomys rutilus	VF	VF224	VF224	Russia	Chelyabinsk District	Vil. Metlino, Kozhakul Lake	55.74	60.87
Clethrionomys smithii	MSB	45312	NK6061	Japan	Honshu	2 km E Haramura	35.96444	138.23801
Clethrionomys smithii	MSB	45313	NK6099	Japan	Honshu	2 km E Haramura	35.96444	138.23801
Clethrionomys smithii	MSB	45314	NK6100	Japan	Honshu	2 km E Haramura	35.96444	138.23801
Dicrostonyx groenlandicus	ROM	97339	ROM97339	Canada	Nunavut			
Eothenomys miletus	UWBM	75297	UW75297	China	Sichuan Sheng	Muli County First Forestry Station, Seedling Garden		
Eothenomys miletus	UWBM	75310	UW75310	China	Sichuan Sheng	Muli County First Forestry Station, Seedling Garden		
Eothenomys proditor	UWBM	75298	UW75298	China	Sichuan Sheng	Muli County First Forestry Station, Seedling Garden		
Eothenomys proditor	UWBM	75301	UW75301	China	Sichuan Sheng	912 Forestry Station, Seedling Garden		
Eothenomys proditor	UWBM	75307	UW75307	China	Sichuan Sheng	Muli County First Forestry Station, Lu Jin Gou Stream		
Hyperacrius fertilis	Vermont U	VT282	VT282	Pakistan		NWFP, Sari	34.41	73.28
Microtus gregalis	MSB	230797	NK192395	Mongolia	Gobi Altai	Khar Azarga Nuruu, Khaalga Gol.	45.83021667	96.09326667
Microtus gregalis	MSB	230804	NK192396	Mongolia	Gobi Altai	Khar Azarga Nuruu, Khaalga Gol.	45.83021667	96.09326667
Microtus sp.	MSB	149290	NK125710	Canada	Yukon Territory	Campell hwy (hwy #4); 3km N and 7.5km E of Carmacks	62.1159	-136.14425

Appendix B. Specimens examined by gene.

Gene	Cytb		ETS2		G6pd		IRBP	
	Sample ID #	Species	Sample ID #	Species	Sample ID #	Species	Sample ID #	Species
	NK100600	Asemicanus	NK123303	Alemminus	NK100718	Asemicanus	NK100718	Asemicanus
	NK100602	Asemicanus	NK125708	Crutilus	NK123303	Alemminus	NK123303	Alemminus
	NK100683	Asemicanus	NK125710	Microtus	NK125708	Crutilus	NK125708	Crutilus
	NK100684	Asemicanus	NK138064	Ccentralis	NK125710	Microtus	NK125710	Microtus
	NK100716	Asemicanus	NK138309	Aargentatus	NK132592	Cgapperi	NK132592	Cgapperi
	NK100718	Asemicanus	NK138344	Aargentatus	NK138064	Ccentralis	NK138064	Ccentralis
	NK123303	Alemminus	NK139890	Cglareolus	NK138200	Ccentralis	NK138200	Ccentralis
	NK125708	Crutilus	NK153189	Cgapperi	NK138203	Ccentralis	NK138203	Ccentralis
	NK125710	Microtus sp.	NK158303	Alemminus	NK138344	Aargentatus	NK138309	Aargentatus
	NK132592	Crutilus	NK158304	Alemminus	NK138809	Aargentatus	NK138344	Aargentatus
	NK138200	Ccentralis	NK158776	Crutilus	NK139890	Cglareolus	NK139890	Cglareolus
	NK138203	Ccentralis	NK166640	Abarakshin	NK147346	Crufocanus	NK153189	Cgapperi
	NK138309	Aargentatus	NK166807	Crufocanus	NK153189	Cgapperi	NK158303	Alemminus
	NK138344	Aargentatus	NK192382	Abarakshin	NK158303	Alemminus	NK158304	Alemminus
	NK139890	Cglareolus	NK192395	Microtusgregalis	NK158304	Alemminus	NK158776	Crutilus
	NK147346	Crufocanus	NK192470	Asemicanus	NK158776	Crutilus	NK166640	Abarakshin
	NK153189	Cgapperi	NK192675	Abarakshin	NK166640	Abarakshin	NK166807	Crufocanus
	NK158303	Alemminus	NK192675	Abarakshina	NK192382	Abarakshin	NK192382	Abarakshin
	NK158304	Alemminus	NK192735	Abarakshin	NK192396	Microtusgregalis	NK192470	Asemicanus
	NK158305	Alemminus	NK193849	Abarakshin	NK192470	Asemicanus	NK192675	Abarakshin
	NK158306	Alemminus	NK194234	Abarakshin	NK192675	Abarakshin	NK192735	Abarakshin
	NK158776	Crutilus	NK194406	Astrelzowi	NK192735	Abarakshin	NK192849	Abarakshin
	NK165544	Abarakshin	NK194420	Astrelzowi	NK193849	Abarakshin	NK194234	Abarakshin
	NK165569	Abarakshin	NK194429	Astrelzowi	NK194234	Abarakshin	NK194406	Astrelzowi
	NK165571	Abarakshin	NK6061	Csmithii	NK194406	Astrelzowi	NK194420	Astrelzowi
	NK165609	Abarakshin	NK6099	Csmithii	NK194420	Astrelzowi	NK194429	Astrelzowi
	NK165614	Abarakshin	NK6100	Csmithii	NK194429	Astrelzowi	NK6061	Csmithii
	NK165665	Abarakshin	NK62143	Cgapperi	NK6099	Csmithii	NK6099	Csmithii
	NK165667	Abarakshin	NK6285	Crufocanus	NK6100	Csmithii	NK6100	Csmithii
	NK165689	Abarakshin	NK6361	Crutilus	NK62143	Cgapperi	NK62143	Cgapperi
	NK165711	Abarakshin	NK6405	Candersoni	NK6361	Crutilus	NK6285	Crufocanus
	NK165715	Abarakshin	NK97450	Asemicanus	NK97450	Asemicanus	NK6361	Crutilus
	NK166517	Abarakshin	ROM97339	Dicrostonyx	ROM97339	Dicrostonyx	NK6405	Candersoni
	NK166568	Abarakshin	UW75247	Ceva	UW75247	Ceva	NK6454	Candersoni
	NK166584	Abarakshin	UW75284	Ceva	UW75284	Ceva	NK97450	Asemicanus
	NK166601	Abarakshin	UW75287	Ceva	UW75287	Ceva	ROM97339	Dicrostonyx
	NK166605	Abarakshin	UW75297	Emiletus	UW75297	Emiletus	UW75247	Ceva
	NK166611	Abarakshin	UW75298	Eproditor	UW75301	Eproditor	UW75287	Ceva
	NK166613	Abarakshin	UW75301	Eproditor	UW75307	Eproditor	UW75297	Emiletus
	NK166617	Abarakshin	UW75307	Eproditor	UW75310	Emiletus	UW75298	Eproditor
	NK166637	Abarakshin	UW75310	Emiletus	UW78787	Ccalifornicus	UW75301	Eproditor
	NK166638	Abarakshin	UW78787	Ccalifornicus	UW80363	Ccalifornicus	UW75307	Eproditor
	NK166639	Abarakshin	UW80363	Ccalifornicus	VF224	Crutilus	UW75310	Emiletus
	NK166640	Abarakshin	VF224	Crutilus	VT282	Hfertilis	UW78787	Ccalifornicus
	NK166641	Abarakshin	VT431	Aargentatus	VT406	Aalbicaudus	UW80363	Ccalifornicus

NK166642	Abarakshin	VT447	Aalbicaudus	VT439	Amontosa	VF224	Crutilus
NK166723	Abarakshin	VT448	Aalbicaudus	VT447	Aalbicaudus	VT431	Aargentatus
NK166727	Abarakshin	NK100718	Asemicanus	VT448	Aalbicaudus	VT447	Aalbicaudus
NK166769	Abarakshin			NK192395	Microtusgregalis	VT448	Aalbicaudus
NK166773	Abarakshin						
NK166807	Crufocanus						
NK166818	Abarakshin						
NK192359	Abarakshin						
NK192360	Abarakshin						
NK192382	Abarakshin						
NK192395	Microtusgregalis						
NK192396	Microtusgregalis						
NK192468	Asemicanus						
NK192470	Asemicanus						
NK192487	Asemicanus						
NK192490	Asemicanus						
NK192628	Abarakshin						
NK192629	Abarakshin						
NK192630	Abarakshin						
NK192674	Abarakshin						
NK192675	Abarakshin						
NK192677	Abarakshin						
NK192679	Abarakshin						
NK192735	Abarakshin						
NK192736	Abarakshin						
NK192737	Abarakshin						
NK192741	Abarakshin						
NK192742	Abarakshin						
NK192743	Abarakshin						
NK193531	Abarakshin						
NK193849	Abarakshin						
NK193892	Abarakshin						
NK193997	Abarakshin						
NK194234	Abarakshin						
NK194355	Astrelzowi						
NK194406	Astrelzowi						
NK194420	Astrelzowi						
NK194422	Astrelzowi						
NK194429	Astrelzowi						
NK6061	Csmithii						
NK6099	Csmithii						
NK6100	Csmithii						
NK62143	Cgapperi						
NK6285	Crufocanus						
NK6361	Crutilus						
NK6404	Candersoni						
NK6405	Candersoni						
NK6406	Candersoni						
NK6454	Candersoni						
NK97450	Asemicanus						

NK97452	Asemicanus
NK97453	Asemicanus
NK97454	Asemicanus
ROM97339	Dicrostonyx
UW75247	Ceva
UW75284	Ceva
UW75287	Ceva
UW75297	Emiletus
UW75298	Eproditor
UW75301	Eproditor
UW75307	Eproditor
UW75310	Emiletus
UW80363	Ccalifornicus
UW80365	Ccalifornicus
VF224	Crutilus
VT282	Hfertilis
VT406	Aalbicaudus
VT431	Aargentatus
VT439	Amontosa
VT447	Aalbicaudus
VT448	Aalbicaudus

Appendix C. GenBank Accession numbers of sequences used. No GenBank sequences were available for ETS2. Names may reflect updated taxonomy that does not match original GenBank information.

Gene	Cytochrome b		G6pd		IRBP	
	Accession #	Species	Accession #	Species	Accession #	Species
	AB037281	Candersoni	AB086025	Crufocanus	AM919401	Cglareolus
	AB037296	Candersoni	AB086026	Crufocanus	AY163583	Emelanogaster
	AB037305	Csmithii	AB086027	Crutilus	AY326080	Cgapperi
	AB037316	Csmithii	AB086028	Crutilus	JF906129	Micfortis
	AB104505	Candersoni	AB086029	Crex		
	AB104508	Csmithii	AB086030	Candersoni		
	AB565454	Crex	AB086031	Candersoni		
	AB565456	Crex	AB086032	Candersoni		
	AF119273	Alemminus	AB086033	Candersoni		
	AF272631	Crutilus	AB086034	Csmithii		
	AF272632	Crutilus	AB086035	Csmithii		
	AF272635	Cgapperi	AB086036	Csmithii		
	AF272636	Cgapperi	AB086037	Cregulus		
	AF272638	Crutilus	JF930092	Cglareolus		
	AF272639	Cgapperi	JF930093	Cglareolus		
	AF272640	Crufocanus	JF930094	Cglareolus		
	AY309411	Alemminus	JF930095	Cglareolus		
	AY309412	Crufocanus	JF930096	Cglareolus		
	AY309413	Crufocanus	JF930097	Cglareolus		
	AY309414	Crufocanus	JF930098	Cglareolus		
	AY309415	Crufocanus	JF930099	Crutilus		
	AY309416	Crufocanus	JF930100	Crufocanus		
	AY309417	Crufocanus	JF930101	Crufocanus		
	AY309418	Crufocanus	JF930102	Crufocanus		
	AY309420	Cglareolus				
	AY309421	Cglareolus				
	AY309422	Ccalifornicus				
	AY309423	Ccalifornicus				
	AY309424	Crutilus				
	AY309425	Crutilus				
	AY309426	Crutilus				
	AY309427	Crutilus				
	AY309429	Cgapperi				
	AY309430	Cgapperi				
	AY309431	Cgapperi				
	AY309432	Cgapperi				
	AY309434	Cgapperi				
	AY426675	Ecachinus				
	AY426676	Ecustos				
	AY426677	Ecustos				
	AY426678	Eeleusis				
	AY426679	Eeleusis				
	AY426680	Efidelis				
	AY426681	Emelanogaster				
	AY426682	Emelanogaster				
	AY426683	Emiletus				
	AY426686	Emiletus				
	AY426687	Eolitor				
	AY426690	Eolitor				
	AY426691	Eproditor				
	DQ845185	Ccentralis				
	DQ845186	Aargentatus				
	DQ845187	Asemicanus				
	DQ845188	Abarakshin				
	DQ845189	Amacrotisvinogradovi				
	DQ845190	Astrelzowi				
	DQ845191	Astrelzowi				
	DQ845192	Asemicanus				
	DQ845193	Asemicanus				
	DQ845194	Abarakshin				
	DQ845195	Amacrotisvinogradovi				
	DQ845196	Amacrotisvinogradovi				
	HM165383	Ecachinus				

HM165384	Ecachinus
HM165385	Cinez
HM165386	Cinez
HM165387	Cinez
HM165398	Echinensis
HM165403	Ceva
HM165405	Ceva
HM165433	Echinensis
HM165434	Echinensis
HM165440	Eproditor
HM165441	Eproditor
JN032747	Astrelzowistrelzowi
JN032754	Astrelzowidesertorum
JN032755	Astrelzowidesertorum

Chapter 4

Thesis Summary

My Master's thesis highlights the dynamic history of a group of northern mammals and the processes responsible for diversification and geographic structure. It also reinforces the importance of applying multiple independent inference methods and broad sampling (taxonomically, geographically, and genetically) in biogeographic inference. This thesis is the first comprehensive phylogeographic investigation of the widespread northern red-backed vole, *Clethrionomys rutilus*, and, at a higher taxonomic level, provides new insight into phylogenetic relationships and species limits of 28 of the 36 members of the tribe Clethrionomyini. The addition of multilocus data facilitated the resolution of numerous long-standing taxonomic issues within the species and tribe and revealed problems not previously recognized by mtDNA or morphological studies. The Holarctic perspective that this species and tribe provide is essential for understanding the history of a climatically and geologically dynamic region of the world (Repenning, Fejfar et al. 1990) and I begin to develop a more complete understanding of evolution and diversification at high latitudes in the Northern Hemisphere.

In Chapter 2, I characterized the dynamic and idiosyncratic biogeographic history of northern red-backed voles, which emphasizes the role of distinct Asian refugia in Holarctic diversification and recolonization of terrestrial organisms during the Late Pleistocene. By about 100-200KYA, populations of *C. rutilus* became isolated in three refugia, two in Asia and the third in Beringia. From these regions ancestral red-backed voles expanded to their current widespread distribution during the late Pleistocene but prior to the Last Glacial Maximum. Thus, the species does not conform to the typical

phylogeographic pattern exhibited by previously examined boreal forest-associated species, which show limited structure throughout their distributions (Fedorov, Goropashnaya et al. 2008; Korsten, Ho et al. 2009). I identified secondary contact zones and refugial locations that are commonly recognized for other species but shed new light on the importance of these refugia for the recolonization of the Holarctic.

In Chapter 3, I performed the first multilocus genetic analysis of Clethrionomyini using broad taxonomic sampling that included over half of the species in the tribe. Apparent paraphyly between *Alticola* and *Clethrionomys*, as indicated by mtDNA, was not corroborated by nuclear loci. Similarly, this first molecular assessment of *Hyperacrius* showed that it has been incorrectly assigned to the tribe. These analyses supported the elevation of a new genus, *Craseomys,* consisting of *C. rufocanus* and Korean and Japanese red-backed voles, a deeply divergent group based on several independent lines of evidence. Cryptic diversity or significant phylogeographic structure was uncovered in several species, most notably *Clethrionomys gapperi*, *Caryomys eva*, *Alticola argentatus*, and should be investigated in more detail. Contrary to previous molecular and paleontological work on other arvicolines, preliminary characterization of evolution in the tribe does not indicate discrete pulses of diversification as the dominant process. However, episodic environmental changes during the Quaternary throughout Asia still may have been critical factors for some groups of species in Clethrionomyini and may have led to repeated patterns of diversification that are manifested at multiple taxonomic levels. This intriguing process is seen in the approximately simultaneous divergence among three *C. rutilus* clades (western, central, and eastern) and also among three species (*C. glareolus*, *C. rutilus*, and *C. gapperi*) that presumably occupied these

relative geographical locations during a previous glacial cycle that led to (or reinforced) speciation.

In addition to the knowledge gained about the tribe and species' history and evolution, phylogeographic and phylogenetic studies are critical for a vast array of other research endeavors, ranging from ecology to parasitology. Making the link between genetic variation and demography within a Holarctic mammal provides the necessary foundation for addressing questions concerning potential response to environmental change and species/habitat management and conservation. My work illustrates the consequences of dynamic distribution changes and sheds light on the importance of multiple refugia in Asia that were significant centers of diversification and sources for recolonization throughout the Late Cenozoic. Thus, it illustrates the need to preserve species in critical areas that could serve as refugia or that have generated significant diversity. This is especially critical in the northern high-latitudes and central Asia, which are projected to rapidly change due to climate change and development (Parmesan 2006; Cruz, Harasawa et al. 2007).

Finally, widespread species such as *C. rutilus* have a variety of ecological interactions that are dynamic in time and space and are excellent candidates for investigating coevolution (Thompson 2005). According to the geographic mosaic model of coevolution, species interactions vary among populations in different parts of the species' distribution due to spatial and temporal variation in environmental pressures, genetic makeup/variability, and co-occurring species. The resulting mosaic of coevolutionary forces influences biological diversity in numerous ways. One of the fundamental components of the mosaic model is describing spatial genetic variation

across a species' range (Thompson 2005), one of the primary outcomes of this thesis. Hence, this thesis provides an evolutionary framework for further integrative research involving *C. rutilus* in the future.

References

Cruz, R., H. Harasawa, et al. (2007). Asia. In Climate Change 2007: impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.(Eds ML Parry, OF Canziani, JP Palutikof, PJ van der Linden, CE Hanson) pp. 469-506, Cambridge University Press: Cambridge, UK.

Fedorov, V. B., A. V. Goropashnaya, et al. (2008). "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.

Korsten, M., S. Y. W. Ho, et al. (2009). "Sudden expansion of a single brown bear maternal lineage across northern continental Eurasia after the last ice age: a general demographic model for mammals?" Molecular Ecology **18**(9): 1963-1979.

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

Repenning, C. A., O. Fejfar, et al. (1990). Arvicolid rodent biochronology of the Northern Hemisphere. *International Symposium on the Evolution and Phylogenetic Biostratigraphy of Arvicolids (Rodentia, Mammalia)*, Prague, Pfeil-Verlag.

Thompson, J. N. (2005). The Geographic Mosaic of Coevolution, University of Chicago Press.