Spring 4-15-2017

# Pattern and Process in the Radiation of Ground-dwelling Squirrels

Bryan S. McLean
*University of New Mexico - Main Campus*

Bryan Stephen McLean
*Candidate*

Department of Biology
*Department*

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

*Approved by the Dissertation Committee:*

Dr. Joseph Cook , Chairperson

Dr. Steven Poe

Dr. Felisa Smith

Dr. James Degnan

Dr. Kristofer Helgen

**PATTERN AND PROCESS IN THE RADIATION OF GROUND-DWELLING SQUIRRELS**

**by**

**BRYAN STEPHEN MCLEAN**

B.S., Biology, Appalachian State University, 2004
M.S., Biology, University of North Carolina Wilmington, 2011

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

**Doctor of Philosophy**

**Biology**

The University of New Mexico
Albuquerque, New Mexico

**May, 2017**

## ACKNOWLEDGEMENTS

This dissertation and the degree to which it contributes represent the culmination of six years of fieldwork, labwork, travels, ups and downs, life and love. I owe the largest debt of gratitude to my major advisor Joe Cook, who gave me the freedom to pursue my own ideas and interests about ground squirrels while keeping me from going too far astray. Joe excels as an advisor and leader, but over the past years I had the privilege to witness his talent as a mentor as well. I also thank the additional members of my committee, Drs. Steven Poe, Felisa Smith, James Degnan, and Kristofer Helgen, for their time, insight, and valuable criticisms on the research contained herein.

Being part of the exceptionally rich history of mammalogy at University of New Mexico has been a highlight of my graduate career. I thank the staff, scientists, and recent students at Museum of Southwestern Biology Division of Mammals with whom I have interacted over the past several years, in particular Jon Dunnum, Adrienne Raniszewski, and Mariel Campbell. I thank my past and present labmates Kayce Bell, Donavan Jackson, Schuyler Liphardt, Dianna Krejsa, Steve Greiman, Jocie Colella, Carlos Carrion, Andrew Hope, Jason Malaney, Jessica Weber, and Yadeeh Sawyer for support and lots of great times, especially in the field. My development as a scientist in general also greatly benefitted from interactions with Kayce Bell, Grey Gustafson, Libby Beckman, Mason Ryan, Dan Colman, Chris Witt, Andy Johnson, and Levi Gray.

I spent one of the past six years as a fellow at the National Museum of Natural History. I thank my coadvisors there, Drs. Kris Helgen and Gene Hunt, for that valuable and constructive opportunity. I thank staff in the NMNH Division of Mammals for additional help and guidance, particularly Al Gardner, Darrin Lunde, Esther Langan, and

Nicole Edmison. While at NMNH, I was able to greatly expand my thinking about ground squirrel evolution and the amount of data included in this dissertation, and I benefitted from interactions with a wide array of resident and visiting researchers. Of course, this dissertation relied on specimens contained in many museums other than MSB and NMNH; in particular, I wish to thank Bob Timm (KU), Jim Patton, Eileen Lacey, and Chris Conroy (MVZ), Priscilla Tucker and Cody Thompson (UMMZ), and Sharlene Santana and Jeff Bradley (UWBM) for access to materials under their care.

Finally, I thank my parents, Steve and Bonnie McLean, and my sister, Dara McLean, for their long-time support, but especially over these past six years. There were times when visits were few and far between, and when birthday wishes were one or two days late, but without your help this dissertation would not have been possible. I owe a special debt of gratitude to my wife, Sally, a Tennessee girl who I was lucky enough to meet in New Mexico. Sally has been a fellow adventurer, a constant source of support and advice, and a last-minute editor on too many drafts and proposals to remember. Finally, to the many people who contributed in intangible ways, but are not mentioned here, thank you.

**PATTERN AND PROCESS IN THE RADIATION OF GROUND-DWELLING SQUIRRELS**

Bryan S. McLean

B.S. Biology, Appalachian State University
M.S. Biology, University of North Carolina Wilmington
Ph.D Biology, University of New Mexico

ABSTRACT

What causes biological diversity to be unevenly apportioned across the Tree of Life? The pattern is widespread and well-characterized; but our understanding of processes underlying the taxonomic, phenotypic, and ecological disparities of clades remains incomplete. At least some of this disparity is due to clade-specific differences in the ability to respond to ecological opportunity, whereby access to, and exploitation of, different resources in ecological time drives evolutionary divergence and adaptive radiation. However, not all clades respond equivalently to ecological opportunity, and considerable heterogeneity therefore exists in diversification patterns across radiations.

This dissertation focuses on patterns and processes of diversification in ground-dwelling squirrels of the tribe Marmotini. It seeks to infer phylogeny and describe variation in ecological and phenotypic traits (patterns) and, ultimately, to relate those to the developmental, environmental, and evolutionary factors shaping them (processes). It integrates molecular, morphological, and environmental datasets derived from museum specimens at two taxonomic levels (across the entire tribe and within the genus *Urocitellus*). Results at each level are evaluated in the context of current evolutionary theory and practice, and these are used to determine whether evolutionary themes exist in marmotine radiation transcending taxonomic and phylogenetic scales.

# INTRODUCTION

Rapid radiations – the multiplication of one lineage into many in a geologically short period of time – have long captured the interest of biologists. Rapid radiations are an exception to the gradual nature of lineage accumulation that defines most clades and the Tree of Life as a whole (Hedges et al. 2015). As such, they beg explanations as to potential external influences, internal drivers, and consequences for related clades. What are the environmental and ecological contexts that allow some clades to radiate, but others not? Do all clades stand equal chances of radiating, given sufficient time and opportunity? Why do radiations differ so much in form? Can rapid radiation occur without key innovations and adaptation? Answers to these questions are still incompletely known for most lineages, but they are fundamental for understanding how biological diversity has arisen, why it is apportioned unevenly across the Tree of Life, and even how it will respond in an era of unprecedented environmental change.

This dissertation is an attempt to understand evolutionary pattern and process in a radiation of ground-dwelling squirrels, tribe Marmotini (Sciuridae: Xerinae). As presently recognized, Marmotini includes ~93 species allocated to 13 genera and 9 subgenera (Wilson and Reeder 2005, Helgen et al. 2009), although understanding of species limits and systematics below the tribal level is continually evolving (Gunduz et al. 2007, Álvarez-Castañeda and Cortés-Calva 2011, Hoisington-Lopez et al. 2012, Mantooth et al. 2013, Phuong et al. 2014, Patterson and Norris 2016). Marmotines occupy a vast distribution across middle- and high-latitudes of the Holarctic, but their core taxonomic and ecological diversity is in North America, where the group comprises nearly 25% of generic-level rodent diversity north of Mexico (Bradley et al. 2014).

Marmotine species can be found in most biomes throughout this range, including desert, scrubland, woodland, arid and montane grasslands, forests, and alpine and arctic tundra. Many are charismatic and easily observable, and some have been the subjects of substantial scientific attention from ecological, physiological, behavioral, and life history perspectives (Murie and Michener 1984, Boyer and Barnes 1999, Blumstein and Armitage 1998, Armitage 2012).

The concept of Marmotini as a rapid evolutionary radiation significantly antedates the field of molecular systematics (Howell 1938, Bryant 1945, Simpson 1945, Moore 1959, Black 1963). Signatures of marmotine radiation are particularly evident in the fossil record, where nearly simultaneous origins of multiple, genus-level crown lineages is documented in the middle Miocene (see Goodwin 2008 for a review). Ongoing work and new tools have reinforced the view that Marmotini was a rapid radiation (Mercer and Roth 2003, Fabre et al. 2012), established the general sequence of diversification events in the tribe (Thomas and Martin 1993, Giboulet et al. 1997, Harrison et al. 2003), and begun to assess macroevolutionary patterns of phenotypic variation (Hafner 1984, Goodwin 2009, Zelditch et al. 2015). Still, many questions have persisted regarding the exact patterns and processes of taxonomic, phenotypic, and ecological diversification in Marmotini.

The hierarchical nature of evolution guarantees that answers to some of these questions are encoded in the patterns of diversity that surround us. That is the premise of the comparative approach, a research program that has grown rapidly over the past several decades and one that I employ throughout this dissertation. This growth can be attributed to advances in the broadly complementary fields of molecular phylogenetics

and phylogenomics (which are allowing the Tree of Life to be reconstructed in finer detail; Carstens et al. 2012) and comparative phylogenetic methods (which facilitate analysis of diversity patterns in evolutionarily explicit contexts; Glor 2010, O'Meara 2012). The comparative approach, however, is susceptible to numerous biases and assumptions, some of which have plagued previous analyses of marmotine evolution. These include incomplete taxonomic frameworks, inaccurate phylogenetic hypotheses, use of qualitative or strongly allometric morphological data (i.e., in fossil taxa), and use of quantitative morphological data sampled from small numbers of traits. I strive to avoid potential pitfalls by taking a holistic approach in each chapter, incorporating genomic, phenotypic, and environmental datasets and assessing sensitivity of results as often as possible.

*Chapter 1* seeks to resolve phylogeny of Marmotini using a genomic-scale dataset of >3,900 ultraconserved element loci (UCEs). It samples taxa from all 13 marmotine genera as well as 8 of 9 subgenera as currently recognized within the tribe, focusing specifically on resolving the sequence of higher-level diversification events that are the core of marmotine morphological, ecological, and life history diversification. Significant focus is also given to assessing sensitivity of phylogenetic estimates to methodological factors known to bias inferences when speciation times are short; specifically, the taxonomic completeness and phylogenetic informativeness of loci and the choice of inference method. This chapter clarifies a number of previously ambiguous higher-level relationships, providing new context for understanding marmotine radiation. It also reveals clear and substantial impacts of methodological factors on results and, in doing so, identifies persistent uncertainties in need of resolution.

*Chapter 2* assesses patterns and processes of phenotypic evolution in tribe Marmotini. It uses 3 taxonomically comprehensive trait datasets (a novel 3D cranial shape dataset and previously published mandibular and molariform tooth shape datasets) to explore macroevolutionary dynamics in each trait; including inferred evolutionary modes, patterns of disparity in skull shapes among ecotypes, and the frequency of convergent shifts in size and shape. Simulations are also employed to generate null distributions of trait values for comparison with observed patterns. Consistent with the mosaic nature of phenotypic evolution, this work reveals substantial evolutionary heterogeneity among traits, and thus additional focus is given to exploring the extent of these differences and their consequences for macroevolutionary inference in Marmotini.

*Chapter 3* includes a finer-scale phylogenetic analysis of Holarctic ground squirrels (*Urocitellus*), a derived clade of arctic-adapted marmotines whose systematics and evolutionary history is still incompletely understood. It uses a comprehensive multilocus dataset sampled from all 12 species (and 33 of 36 subspecies) to infer phylogeny and test longstanding taxonomic and systematic hypotheses. To distinguish among incomplete lineage sorting and introgression as causes of mitonuclear discordance, it employs phylogenetically-based posterior predictive tests. This work resolves many uncertainties in *Urocitellus* phylogeny and provides robust evidence for multiple ancient, high-latitude introgression events in the *parryii-richardsonii-elegans* clade, adding to a growing list of such documented events in Marmotini and more generally in high-latitude regions.

The goal of the final chapter, *Chapter 4*, is to assess niche evolution, phenotypic adaptation, and modes of speciation in *Urocitellus* by presenting a new and robustly

resolved phylogenetic hypothesis of *Urocitellus* based on >3,500 UCEs. These analyses employ environmental data and an extensive database of >10,000 digitized specimen records to examine evolution of niche mean and breadth in *Urocitellus,* and to test for correlations between environment and phenotype using multiple trait datasets (body size, body shape, and cranial shape). Results reveal expanding niche breadth in *Urocitellus* and evidence of adaptation in all traits analyzed, including correlated changes in body and cranial shape. Finally, the adaptive value of body size and shape changes is evaluated in the context of commonly cited ecogeographic rules to ask whether they may play a role in speciation in *Urocitellus*, thereby linking lower-level diversification conceptually and mechanistically to patterns and processes inferred at higher phylogenetic levels (i.e., *Chapters 1 + 2*).

G. G. Simpson, the architect of much of the modern conceptual framework surrounding evolutionary radiations, commented that "in adaptive radiation… all the modes and all the factors of evolution are inextricably woven. The total process cannot be made simple, but it can be analyzed in part" (Simpson 1953). The goal of this dissertation is to disentangle a small number of those factors and modes by interrogating genomic, phenotypic, and environmental datasets at 2 different taxonomic scales, where the mechanisms operating to foster rapid radiation and adaptation can be compared and contrasted. This approach is integrative and, as a result, has opened the door to added layers of biological complexity, much like those Simpson alluded to. Nevertheless, it is hoped that the conclusions pave the way to a richer and more nuanced understanding of evolutionary pattern and process in ground-dwelling squirrels and other radiations.

REFERENCES

Álvarez-Castañeda, S. T., and P. Cortés-Calva. 2011. Genetic evaluation of the Baja California rock squirrel. Zootaxa 51:35–51.

Armitage, K. B. 2012. Sociality, individual fitness and population dynamics of yellow-bellied marmots. Molecular Ecology 21:532–540.

Black, C. C. 1963. A review of the North American Tertiary Sciuridae. Bulletin of the Museum of Comparative Zoology of Harvard University 130:109–248. Cambridge.

Blumstein, D. T., and K. B. Armitage. 1998. Life history consequences of social complexity: a comparative study of ground-dwelling sciurids. Behavioral Ecology 9:8–19.

Boyer, B. B., and B. M. Barnes. 1999. Molecular and metabolic aspects of mammalian hibernation. Bioscience 49:713–724.

Bradley, R. D., L. K. Ammerman, R. J. Baker, L. C. Bradley, J. A. Cook, R. C. Dowler, C. Jones, D. J. Schmidly, F. B. Stangl, R. A. Van Den Bussche, and B. Wursig. 2014. Revised checklist of North American mammals north of Mexico, 2014. Occasional Papers of the Museum of Texas Tech University 327:1–27.

Bryant, M. D. 1945. Phylogeny of Nearctic Sciuridae. American Midland Naturalist 33:257–390.

Carstens, B., A. R. Lemmon, and E. M. Lemmon. 2012. The promises and pitfalls of next-generation sequencing data in phylogeography. Systematic Biology 61:713–715.

Fabre, P.-H., L. Hautier, D. Dimitrov, and E. J. P. Douzery. 2012. A glimpse on the pattern of rodent diversification: a phylogenetic approach. BMC Evolutionary Biology 12:88.

Giboulet, O., P. Chevret, R. Ramousse, and F. Catzeflis. 1997. DNA-DNA hybridization evidence for the recent origin of marmots and ground squirrels (Rodentia: Sciuridae). Journal of Mammalian Evolution 4:271–284.

Glor, R. E. 2010. Phylogenetic insights on adaptive radiation. Annual Review of Ecology, Evolution, and Systematics 41:251–270.

Goodwin, H. T. 2009. Odontometric patterns in the radiation of extant ground-dwelling squirrels within Marmotini (Sciuridae: Xerini). Journal of Mammalogy 90:1009–1019.

Goodwin, H. T. 2008. Sciuridae. Pp. 355–376 *in* Janis, Christine M., G. F. Gunnell, and M. D. Uhen, eds. Evolution of Tertiary Mammals of North America: Small Mammals, Xenarthrans, and Marine Mammals. Cambridge University Press, Cambridge.

Gündüz, I., M. Jaarola, C. Tez, C. Yeniyurt, P. D. Polly, and J. B. Searle. 2007. Multigenic and morphometric differentiation of ground squirrels (*Spermophilus*, Scuiridae, Rodentia) in Turkey, with a description of a new species. Molecular Phylogenetics and Evolution 43:916–35.

Hafner, D. J. 1984. Evolutionary Relationships of the Nearctic Sciuridae. Pp. 3–23 *in* J. O. Murie and G. R. Michener, eds. The Biology of Ground-Dwelling Sciurids. University of Nebraska Press, Lincoln.

Harrison, R. G., S. M. Bogdanowicz, R. S. Hoffmann, E. Yensen, and P. W. Sherman. 2003. Phylogeny and evolutionary history of the ground squirrels (Rodentia: Marmotinae). Journal of Mammalian Evolution 10:249–276.

Hedges, S. B., J. Marin, M. Suleski, M. Paymer, and S. Kumar. 2015. Tree of life reveals clock-like speciation and diversification. Molecular Biology and Evolution 32:835–845.

Helgen, K. M., F. R. Cole, L. E. Helgen, and D. E. Wilson. 2009. Generic revision in the Holarctic ground squirrel genus *Spermophilus*. Journal of Mammalogy 90:270–305.

Hoisington-Lopez, J. L., L. P. Waits, and J. Sullivan. 2012. Species limits and integrated taxonomy of the Idaho ground squirrel (*Urocitellus brunneus*): genetic and ecological differentiation. Journal of Mammalogy 93:589–604.

Howell, A. H. 1938. Revision of the North American Ground Squirrels, with a Classification of the North American Sciuridae. North American Fauna 56:1–256. Washington, D.C.

Mantooth, S. J., D. J. Hafner, R. W. Bryson, and B. R. Riddle. 2013. Phylogeographic diversification of antelope squirrels (*Ammospermophilus*) across North American deserts. Biological Journal of the Linnean Society 109:949–967.

McCormack, J. E., S. M. Hird, A. J. Zellmer, B. C. Carstens, and R. T. Brumfield. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Molecular Phylogenetics and Evolution 66:526–538.

Mercer, J. M., and V. L. Roth. 2003. The effects of Cenozoic global change on squirrel phylogeny. Science 299:1568–72.

Moore, J. C. 1959. Relationships among the living squirrels of the Sciurinae. Bulletin of the American Museum of Natural History 118:155–206.

Murie, J. O., and G. R. Michener (eds). 1984. The Biology of Ground-Dwelling Squirrels: Annual Cycles, Behavioral Ecology and Sociality. University of Nebraska Press, Lincoln.

O'Meara, B. C. 2011. Evolutionary inferences from phylogenies: a review of methods. Annual Review of Ecology, Evolution and Systematics 43:267–285.

Patterson, B. D., and R. W. Norris. 2016. Towards a uniform nomenclature for ground squirrels: the status of the Holarctic chipmunks. Mammalia 80:241–251.

Phuong, M. A., M. C. W. Lim, D. R. Wait, K. C. Rowe, and C. Moritz. 2014. Delimiting species in the genus *Otospermophilus* (Rodentia: Sciuridae), using genetics, ecology, and morphology. Biological Journal of the Linnean Society 113:1136–1151.

Simpson, G. G. 1953. The Major Features of Evolution. Columbia University Press, New York.

Simpson, G. G. 1945. The principles of classification and a classification of mammals. Bulletin of the American Museum of Natural History 85:1–350.

Thomas, W. K., and S. L. Martin. 1993. A recent origin of marmots. Molecular Phylogenetics and Evolution 2:330–336.

Wilson, D. E., and D. M. Reeder (eds). 2005. Mammal Species of the World. 3rd ed. Johns Hopkins University Press, Baltimore, Maryland.

Zelditch, M. L., J. Li, L. A. P. Tran, and D. L. Swiderski. 2015. Relationships of diversity, disparity, and their evolutionary rates in squirrels (Sciuridae). Evolution 69:1284–1300.

# CHAPTER 1

## Impacts Of Inference Method And Dataset Filtering On Phylogenomic Resolution In A Rapid Mammalian Radiation

AUTHORS
Bryan S. McLean[1], Kayce C. Bell[2], Julie M. Allen[3], Kristofer M. Helgen[4], Joseph A. Cook[1]

[1] Department of Biology and Museum of Southwestern Biology, 1 University of New Mexico, MSC03-2020, Albuquerque, NM 87131 USA
[2] Department of Invertebrate Zoology, Smithsonian Institution National Museum of Natural History, P.O. Box 37012, MRC 163, Washington, DC 20013-7012 USA
[3] Illinois Natural History Survey, University of Illinois, Urbana-Champaign, Il 61820 USA
[4] School of Biological Sciences, University of Adelaide, Adelaide, SA 5005 Australia

ABSTRACT

Large phylogenomic datasets are illuminating many problematic nodes in the Tree of Life. However, size of these datasets alone may be insufficient to resolve the most extreme of radiations. This inadequacy is because inferences in zones of extraordinarily low phylogenetic signal such as rapid radiations can be highly sensitive to taxonomic sampling, locus type and informativeness, and the philosophy and method of inference. We used a dataset of >3,900 ultraconserved element (UCE) loci from a classic mammalian radiation, ground-dwelling squirrels of the tribe Marmotini (Sciuridae: Xerinae), to assess the sensitivity of phylogeny reconstruction across different inference methods (RAxML concatenated, ASTRAL, NJst) and a range of filtering scenarios (taxonomic completeness and phylogenetic informativeness of loci). Topological discordance and bootstrap support variation was found between concatenation- and coalescent-based inferences; within the coalescent-based framework but between methods; and within each method in response to different filtering scenarios. Contrary to some recent UCE-based studies, filtering with either criterion did not promote among-method topological concordance, although ASTRAL and NJst did agree for the single, strictest filtering scenario. Phylogenetic uncertainty in Marmotini is localized to several short internodes and consistent with incomplete lineage sorting and, in at least 1 case, ancient isolation of lineages. Our results reiterate the complexities of resolving ancient radiations even with robust taxon and character sampling, and suggest that sensitivity analyses are crucial when inferring phylogeny in clades known to have experienced rapid diversification.

KEYWORDS

UCEs, Radiation, Incomplete Lineage Sorting, Missing Data, Concatenation, Coalescent

## INTRODUCTION

Rapid radiations are among the most difficult phylogenetic problems to resolve. This difficulty may stem from an incomplete fossil record, a paucity of molecular and morphological synapomorphies, or, in some recent radiations, past and/or ongoing gene flow. Increased genomic coverage of molecular datasets is now clarifying some of the most recalcitrant radiations, including within mammals (Janecka et al. 2007, Song et al. 2012, McCormack et al. 2012), birds (McCormack et al. 2013, Jarvis et al. 2014), and fish (Wagner et al. 2013, Gilbert et al. 2015). However, despite these advances, other major radiations remain incompletely resolved despite availability of thousands, or even tens of thousands, of loci (e.g., solenoid plants, Pease et al. 2016; iguanian lizards, Leache et al. 2015, Streicher et al. 2016). These cases reiterate the fact that, in zones of low phylogenetic signal such as rapid radiations, inferences can be significantly affected by issues other than dataset size, including taxon sampling (Huelsenbeck 1991, Poe 1998), character type and informativeness (Huelsenbeck 1991, Townsend 2007, Salichos and Rokas 2013), and the method of inference used (Huelsenbeck 1995, Maddison and Knowles 2006, Kubatko and Degnan 2007). Parsing the contributions of such methodological factors to phylogenomic estimates is paramount for accurate resolution of rapid radiations and a more fully resolved Tree of Life.

Two core characteristics of molecular datasets that can significantly impact historical inferences are taxonomic completeness and phylogenetic informativeness. We define taxonomic completeness as the proportion of taxa represented *per locus* (instead of the total number of taxa in a dataset). Taxonomic completeness (hereafter, TC) is thus a special case of "missing data" wherein missing data are due nearly or completely to missing taxa. The impacts of TC, and of missing data in general, have been amply discussed in phylogenetic literature (Huelsenbeck 1991, Poe 1998, Wiens 2006, Wiens and Morrill 2011, Roure et al. 2013). These concepts have received renewed attention due to the commonality of missing data in phylogenomic datasets, owing either to stochasticity or taxon- and sample-specific biases in capture or sequencing. Researchers opting for high-TC datasets therefore often face the prospect of excluding many potentially informative loci from phylogenetic analysis due to incomplete taxon representation (e.g., Huang and Knowles 2016). Conversely, an increased tolerance for missing data may not hinder, and can even improve, phylogenetic estimates (Wiens 2006, Wiens and Morrill 2011, Phillipe et al. 2004, Streicher et al. 2016). Yet the potential for improved accuracy depends on other factors as well (e.g., speciation times, mutation rate, the distribution of branch lengths); potential interactions of missing data with these factors have rarely been systematically investigated with genomic-scale datasets.

Phylogenomic datasets can often include loci with minimal variation, but phylogenetic informativeness (PI) has received even less attention as a filtering criterion than TC. As with TC, we define PI on a per-locus basis, as the ratio of parsimony informative sites to total locus length. Exclusion of low-PI loci should decrease computation time and improve estimation of some parameters from the data, as required in maximum likelihood and Bayesian inference (Roure et al. 2013). Filtering loci by PI may also improve phylogeny estimation under the coalescent model, particularly for summary methods (e.g., STAR, MP-EST, ASTRAL, and others). This is because coalescent-based methods use gene trees as input and should be sensitive to inclusion of

imprecise topologies inferred from low-PI loci (e.g., Mirarab et al. 2014, Meiklejohn et al. 2016, Manthey et al. 2016). Finally, exclusion of low-PI loci should improve inferences in the presence of missing taxa, as higher-PI loci will guide placement of poorly represented taxa (Wiens 2006, Wiens and Morrill 2011, Roure et al. 2013). Conversely, total evidence approaches can benefit from inclusion of all loci regardless of PI, as even low-PI loci are still informative for inferring topology and parameter values within a likelihood framework. To date, effects of varying PI for gene- and species-tree inference have been incompletely evaluated for phylogenomic datasets.

We investigated how choice of inference method as well as dataset filtering by TC and PI impact phylogenomic resolution of a classic mammalian radiation, ground-dwelling squirrels of the tribe Marmotini (chipmunks, ground squirrels, marmots, prairie dogs; >91 species in 13 genera; Wilson and Reeder 2005, Helgen et al. 2009, Thorington et al. 2012). Marmotine ground squirrels represent a rapid radiation that began in the middle to late Miocene (~7-9 Ma), spurred by new ecological opportunities associated with expansion of grassland ecosystems across the vast Holarctic region. At this time, the fossil record documents nearly simultaneous origins of multiple, genus-level crown lineages of ground squirrels (Black 1963, 1972). However, despite extensive exploration of different morphological and molecular datasets, phylogenetic uncertainty at deeper levels in Marmotini has persisted due to the rapidity of lineage diversification, morphological conservatism and convergence (Bryant 1945, Hafner 1984, Goodwin 2008, Zelditch et al. 2015) and recurrent hybridization in some clades (e.g., McLean et al. 2016).

We generated the first genomic-scale sequence dataset from Marmotini comprised of >4000 ultraconserved element (UCE) loci, focusing specifically on resolving the sequence of generic- and subgeneric-level diversification events that represent the core of marmotine morphological, ecological, and behavioral evolution (e.g., Helgen et al. 2009). Ultraconserved elements are well suited for this question because they contain highly conserved core regions that can be targeted across taxonomic levels (Bejerano et al. 2004, Faircloth et al. 2012) as well as flanking regions sufficiently variable to be informative at a range of temporal scales (Gilbert et al. 2015), making them effective markers in resolving vertebrate radiations (e.g., McCormack et al. 2012, McCormack et al. 2013). However, a wide range of filtering approaches have been employed in UCE-based studies to date, and the effects of filtering by TC and PI have rarely been comprehensively investigated for these popular markers. We used our dataset to simultaneously reconstruct marmotine phylogeny and assess the sensitivity of these inferences across a broadly representative range of filtering scenarios and under the commonly used inference philosophies (i.e., concatenation, coalescent-based) and methods (RAxML, ASTRAL, NJst).

## METHODS

### *Samples and Sequencing*

We sampled tissues of 67 specimens from 35 species representing all 13 genera, as well as 8 of 9 subgenera, currently recognized within Marmotini (Wilson and Reeder 2005, Helgen et al. 2009; Appendix 1). This sampling scheme includes the single

currently recognized chipmunk genus (*Tamias*), but we note the recent proposal (Patterson and Norris 2016) that the 3 chipmunk subgenera (*Tamias*, *Eutamias*, *Neotamias*) be recognized at the generic level. Although we do not adopt this proposed change here, we note that our dataset includes representatives of all subgenera and may thus provide additional evidence for their genus-level distinctiveness. Ingroup sampling also included Père David's Rock Squirrel (*Sciurotamias davidianus*), which has been variously included in the tribes Tamiasciurini and Ratufini, but whose inclusion in Marmotini is currently supported by limited nuclear data (Steppan et al. 2004). In most cases, tribal sampling included multiple individuals per species (32 of 35) and multiple species per genus (9 of 11 polytypic genera) to maximize chances of observing coalescent events and thus increasing gene tree accuracy. We also sampled 3 unambiguous outgroup taxa: 2 distantly related tree squirrels (*Sciurus aberti*, *Tamiasciurus hudsonicus*) and the primitive sciuromorph, *Aplodontia rufa* (Family Aplodontiidae).

Genomic DNA was extracted from frozen tissues, dried muscle, skin clips, or toe pads using a standard salt extraction protocol. Dried samples (muscle, skin clips, toe pads) were first cleaned by removing hair and debris, cut into sub-centimeter sized pieces, soaked in 70% ethanol overnight, with 3 changes of ethanol. Samples were then rinsed in STE buffer for 24 h under refrigeration, vortexing intermittently. Final extractions were quantified flourometrically using a Qubit (Life Technologies Corporation). Aliquots of 0.5-4 micrograms of genomic DNA were submitted to RapidGenomics, LLC (Gainesville, Florida) for library prep and sequencing. UCE loci were captured using the UCE-5Kv1 probe set (http://www.mycroarray.com/mybaits/mybaits-UCEs.html), libraries were enriched, and sequencing was performed using 2x100 paired-end technology on the Illumina HiSeq platform.

UCE reads were quality trimmed in the FASTX Toolkit v0.0.14 (unpublished; available at http://hannonlab.cshl.edu/fastx_toolkit/) by removing the first 5 bases with consistently lower scores from the 5' end of the sequence. All reads were then quality trimmed from the 3' end to remove bases with a phred score less than 27 using a sliding window of 1nt. Finally, any trimmed reads with fewer than 35 nt were removed from the dataset. UCE loci were assembled in aTRAM (Allen et al. 2015) using the 5041 UCE probes (downloaded from ultraconserved.org) as targets. Loci were assembled with Trinity (Grabherr et al. 2011) over 5 iterations. Following assembly, custom scripts (available at https://github.com/juliema/aTRAM_UCE_pipeline) were used to retrieve the longest contig from the aTRAM best files, generate a consensus sequence for the UCE loci with multiple probes, and combine sequences for all samples by UCE locus. Loci were then aligned in MAFFT version 7.2 using default settings (Katoh and Standley 2013). When summary statistics excluded outgroups, these were removed and realignments performed. We next trimmed UCE flanks using the *deleteGaps* function in the ips package (Heibl 2014) in R v3.2.3 (R Core Team 2015); specifically, we trimmed to positions where at least half of the samples contained data. That script also served as a general tool to clean poorly represented nucleotide positions within core UCEs, which has the potentially negative effect of removing indels. However, indels have rarely been reported in UCEs, and potential removal is unlikely to alter inferences. Finally, we used the script *phyluce_align_get_informative_sites.py* from the phyluce pipeline (Faircloth et

al. 2012, Faircloth 2015) to identify and remove UCE loci with >50% informative sites ($N = 51$). Based on current knowledge of UCE variability (e.g., Bejerano et al. 2004), these may represent errors in locus capture or assembly, so their removal is a conservative step.

*Filtering Loci by Completeness and Informativeness*

Following bioinformatics processing, we generated 42 nested UCE datasets of variable TC and PI. We used the phyluce script *phyluce_align_get_only_loci_with_min_taxa.py* to filter loci by taxonomic representation, specifically, ≥95, 90, 80, 70, 60, and 50% of taxa. This range is representative of that used in simulation- and empirically-based studies (Wiens and Morrill 2011, Roure et al. 2013) as well as UCE-based studies (McCormack et al. 2013, Hosner et al. 2015, Meiklejohn et al. 2016, Streicher et al. 2016). Each of the taxonomically-variable datasets was then further filtered into bins by PI by first calculating relative informativeness (PI = numbers of informative sites / locus length) using output from the phyluce script *phyluce_align_get_informative_sites.py* and extracting target loci. We used all loci as well as 6 filters of ≥1, 2.5, 5, 7.5, 10, and 20% PI, a range that is both statistically representative of our dataset (median 9.7% informative sites, first and third quartiles 5.5 and 15.8%, calculated for the ≥50% taxonomically complete dataset) and broadly representative of the few UCE-based studies that have used PI as a filter (e.g., Hosner et al. 2015, Hawkins et al. 2016, Meiklejohn et al. 2016).

*Phylogenetic Inference*

We inferred phylogeny from each dataset using a maximum likelihood (ML) concatenated approach in RAxML (Stamatakis 2014) as well as a summary coalescent approach in both ASTRAL (Mirarab et al. 2014) and NJst (Liu and Yu 2011). These are among the most commonly used methods in phylogenomic studies; however, they span a major philosophical dichotomy in multilocus phylogenetic inference (total evidence vs. coalescent-based). The latter 2 methods are statistically consistent under the coalescent and therefore better accommodate the gene tree heterogeneity that can be pervasive in rapid radiations (Degnan and Rosenberg 2006, Kubatko and Degnan 2007). However, ASTRAL and NJst are algorithmically distinct and may perform differently for different empirical systems and filtering scenarios. On the other hand, concatenation may perform as well or better than summary coalescent methods in a minority of situations, such as when loci have limited variation or when ILS is low (e.g., Gatesy and Springer 2014, Mirarab et al. 2014). Performance of these 3 methods has rarely been compared, and never across the breadth of filtering regimes imposed here. We did not employ a Bayesian total evidence approach due to finite access to computing resources, and because additional biases are possible in Bayesian inference via interaction of missing data and priors (Lemmon et al. 2009).

For concatenated ML analyses, we generated matrices for each of the 42 datasets described above using the Python package *amas.py* (Borowiec 2016). Inferences were performed in RAxML v8.2.8 (Stamatakis 2014) using a GTR substitution model with

gamma-distributed rate heterogeneity (-m GTRGAMMA option) with 25 categories (the default) and support was assessed by performing a rapid bootstrap analysis with 100 replicates (-f a option), followed by a thorough ML tree search. We conducted inferences under the GTRGAMMA heterogeneity model instead of the GTRCAT model to obtain more interpretable estimates of rate heterogeneity across datasets, because many of our datasets contained few (<50) taxa, and because preliminary analyses using GTRCAT yielded inconsistent support values for a small minority of problematic branches. All RAxML analyses were conducted on the CIPRES Science Gateway (www.phylo.org).

Coalescent-based methods ASTRAL (Mirarab et al. 2014) and NJst (Liu and Yu 2011) were used to infer species trees from the same 42 datasets. ASTRAL computes a species tree that agrees with the highest number of taxon quartets induced by a set of given gene trees, whereas NJst infers a species tree by first constructing a distance matrix describing average distances (in number of internodes) between taxon pairs across gene trees (i.e., gene-tree internode distances; Liu and Yu 2011). NJst is closely related to the STAR method (Liu et al. 2009) but applicable to unrooted trees (Allman et al. 2016). Both ASTRAL and NJst accept unrooted gene trees as input as well as missing data and can immediately accommodate the varying levels of TC used here. We first inferred gene trees in RAxML v8.1.17, assuming a GTR substitution model with gamma-distributed rate heterogeneity for each locus. Gene tree support was assessed from a rapid bootstrap analysis with 100 replicates. We used the gene tree with the highest likelihood score (RAxML bestTree) as the preferred gene tree for input in ASTRAL and NJst. We ran ASTRAL v4.10.8 with default settings and we ran NJst with default settings in the R package phybase (Liu 2014).

To assess support for ASTRAL and NJst inferences, we ran bootstrap analyses for each of the 42 datasets (100 replicates per dataset). We employed the "site-only" bootstrap method in ASTRAL, each time drawing at random a gene tree bootstrap replicate for each locus and running the method with default settings. We used custom scripts to perform an identical procedure in NJst (randomization of gene tree bootstraps, each time drawing 1 replicate per locus). This approach was preferable because it allowed us to use all loci present in each filtered dataset. Finally, we used the addConfidences function in the R package phangorn (Schliep et al. 2016) to compute bootstrap support values for each dataset, using the ASTRAL and NJst trees inferred from the RAxML best trees above as target trees. All gene tree inferences and bootstrapping analyses were conducted using clusters at the University of New Mexico Center for Advanced Computing (www.carc.unm.edu).

*Quantifying Discordance and Summarizing Trees*

We used the R packages ape (Paradis et al. 2016) and phangorn as well as custom scripts to quantify bootstrap support and precision of gene trees and of species trees, and to assess concordance among species trees. Gene tree/species tree support and precision were quantified as the mean and coefficient of variation (CV) of bootstrap support values, respectively. The effects of TC and PI on gene tree support and precision were assessed using a standard ANOVA. Species tree discordance among the 42 datasets was quantified with the Robinson-Foulds distance metric. We used default settings in the CONSENSE

module in PHYLIP (Felsenstein 2005) to summarize frequencies with which certain clades were recovered in the separate inference methods across filtering scenarios.

## RESULTS

Sequencing on the Illumina HiSeq resulted in a total of 414,288,196 paired-end reads (average 5,675,180 per sample, range 1,551,160 - 14,007,738). aTRAM assemblies averaged 4,901,561 reads per sample (range 597,507- 12,378,998). There were 4,402 UCE alignments, similar to that observed in some other vertebrate studies that utilized approximately the same number of probes (e.g., Hosner et al. 2015, Hawkins et al. 2016, Streicher et al. 2016, Manthey et al. 2016). Median length of trimmed loci was 590.5bp and median taxon representation per locus was 66 (of 73 total samples; Table 1). Exclusion of 6 outgroup samples (from 3 outgroup species) did not alter average locus length, but did result in significant decreases in both the number and proportion of variable and phylogenetically informative sites per locus (Wilcoxon rank sum tests, $P < 0.01$ for all; Table 1). Of the 4,402 UCE loci, 483 were excluded from further analyses because they contained less than 50% of taxa. Systematic filtering of remaining loci resulted in datasets that varied by over an order of magnitude in size, from 203 loci (≥95% TC & ≥20% PI) to 3,919 loci (≥50% TC & all levels of PI; Fig. 1a). Similarly, concatenated matrices composed of these loci varied in length by over an order of magnitude, from 140,511bp (95% TC & ≥20% PI) to 2,365,391bp (≥50% TCs & all levels of PI).

As expected, TC and PI were significant predictors of mean gene tree bootstrap support, as was their interaction term ($P << 0.01$ for all). TC and PI were each significant predictors of gene tree precision as well (quantified as the coefficient of variation of tree-wide bootstrap values; $P << 0.01$ for all), although their interaction was marginally insignificant ($P = 0.07$). However, this difference in effect sizes for PI was much larger than for TC in both tests ($F = 5457$ and 242 for mean branch support, $F = 1888$ and 90 for CV of branch support). This was manifested as a nearly 50% increase in bootstrap support and 40% increase in precision across levels of PI, whereas only minor changes in these metrics were found across levels of TC (Fig. 1b,c). Filtering by PI thus appears to minimize the component of gene tree heterogeneity that is due to gene trees of low information (but not necessarily due to other causes such as deep coalescence).

The topology inferred from RAxML concatenated analysis of the largest dataset (Fig. 2; ≥50% TC and all levels of PI) is also the tree from which RF distances of other species trees were measured. We refer to this as the preferred concatenated phylogeny because it is based on the largest dataset that we analyzed (as in Roure et al. 2013) and is also topologically identical to results from concatenated analysis of 13 additional, filtered datasets (Fig. 4). This tree is largely consistent with existing knowledge of marmotine relationships and, like all methods and datasets, it supports the apportionment of ground squirrel diversity into 4 major groups: chipmunks (*Tamias*), which form the sister clade to all other marmotines; antelope and tropical ground squirrels (*Ammospermophilus* and *Notocitellus*, respectively); marmots and other ecologically generalist ground squirrels (*Marmota, Otospermophilus, Callospermophilus*); and a morphologically and ecologically diverse clade of 6 derived ground squirrel genera (*Spermophilus, Urocitellus, Poliocitellus, Xerospermophilus, Ictidomys, Cynomys*).

There were persistent differences across the 42 topologies resulting from each of 3 inference methods (Figs. 2-3). However, each method displayed sensitivity to varying TC and PI (Fig. 4). Concatenation was by far the most sensitive method based on overall variation in Robinson-Foulds distances as measured from the preferred concatenated tree. Specifically, topologies inferred from concatenated analyses varied in symmetric distance by more than twice those of ASTRAL and NJst (Fig. 4). RF distances of concatenated trees from the preferred concatenated tree were directly correlated with dataset size (Fig. 4), and most of the largest filtered datasets produced identical topologies (e.g., top right quadrant in Fig. 4). The most discordant concatenated topologies resulted from datasets filtered to include loci with ≥20% PI, which were also the smallest datasets. This pattern of extreme discordance was observed regardless of levels of TC, and thus was probably due to the relatively steep loss in characters when an extreme filtering regime is imposed. The absence of a similar sensitivity in ASTRAL and NJst appears to reflect the disproportionate influence that filtering has in concatenated analyses owing to loss of many characters versus exclusion of whole gene trees for summary methods.

Because ASTRAL and NJst are statistically consistent under the coalescent model, we predicted they would result in similar inferences from identical datasets, and may also respond similarly across the filtering gradient. While ASTRAL and NJst trees were more similar on average than either were to concatenated trees, they were rarely identical. ASTRAL and NJst were concordant only under the most extreme filtering scenario (≥95% TC & ≥20% PI), wherein only 203 loci contributed to inferences (Figs. 3-4, Supp. Fig. 1). ASTRAL and NJst also differed in the form of response to filtering. In general, as more data were added (by relaxing either or both filter(s)), ASTRAL trees diverged from the NJst-consistent tree and converged on the preferred concatenated tree (Fig. 5). The form of the response in ASTRAL and RAxML was thus broadly similar, although addition of more data led to a decrease in bootstrap support in ASTRAL, which is illustrated in Figure 3. Conversely, the form of the response of NJst across this filtering gradient was less clear. NJst and concatenated trees showed highest concordance under the most extreme filtering scenarios, but increases in dataset size did not lead to convergence of NJst trees with those of either RAxML or ASTRAL (Fig. 4).

In general, concatenated analyses resulted in species tree topologies with higher bootstrap support and greater precision compared to coalescent-based methods (Supp Fig. 2). The 2 classes of methods displayed opposite responses to increasing dataset size, however, with addition of more loci leading to increased support and precision in concatenated trees (Fig. 5a,d), but decreased support and precision in ASTRAL and NJst (Fig. 5b-c,d-f). In the coalescent-based methods considered here, PI appears to be a more important driver of bootstrap support values than TC, consistent with the relatively greater impact of this filtering criterion on gene tree precision that we described above. Thus, results of bootstrap analyses suggest that the ability of coalescent-based methods to accommodate gene tree heterogeneity also leads to a more accurate representation of true uncertainty in species tree estimates when increasing numbers of low-TC and low-PI loci are added.

Our estimates have several discrepancies with existing phylogenetic hypotheses, and these are localized to short internal branches that subtend taxa originating deep in the tree, or that subtend splits of major species complexes, subgenera, or genera (Figs. 2-3). These types of scenarios are known to have high potential to complicate phylogenetic

inference (e.g., Degnan and Rosenberg 2006). First, substantial discordance existed among methods and datasets in placement of the East Asian genus *Sciurotamias*. This genus historically was allocated to sciurid tribes other than Marmotini, reflecting its lack of morphological as well as ecological continuity with the tribe. As in previous multilocus analyses (Mercer and Roth 2003, Steppan et al. 2004), we were unable to consistently place *Sciurotamias*, which we take to indicate a deep history of geographic and genomic isolation from other extant marmotines. The remainder of discordance, however, was localized to short internal branches subtending different pulses of marmotine diversification. For example, split(s) of the marmot subgenera *Marmota* and *Petromarmota* were not recovered in concatenated analyses when the most stringent filters were applied. Splits of *Poliocitellus* and *Urocitellus* from the remainder of derived ground squirrel genera were also inconsistent; the majority of ASTRAL and NJst runs actually supported non-monophyly of *Urocitellus* (Fig. 6). These issues appear to reflect the severity of ILS within different ground squirrel radiations, particularly that of derived ground squirrels. Conversely, we recovered full support for a sister relationship between prairie dogs (*Cynomys*) and lined ground squirrels (*Ictidomys*) to the exclusion of pygmy ground squirrels (*Xerospermophilus*), which differs from previous mtDNA studies that supported a ((*Cynomys*, *Xerospermophilus*), *Ictidomys*) relationship (Harrison et al. 2003). Thus, ILS may not be equally severe across marmotine phylogeny, even within derived ground squirrels.

## DISCUSSION

Interrogation of a novel empirical dataset allowed us to assess the sensitivity of phylogenetic estimates to different inference methods and dataset characteristics in a rodent radiation. Our results confirm that, in topological zones of low phylogenetic signal such as those which characterize many rapid radiations, phylogenomic inference can be sensitive to both inference method and the TC and PI of datasets used. Although this sensitivity is not unexpected, few empirical studies have evaluated its severity. The discordance we document has important implications for other phylogenomic studies. First, *post hoc* dataset filtering (i.e., after topological discordance has been discovered) as a means to achieve among-method concordance, and therefore identify a global phylogenetic hypothesis, may be an unproductive and potentially misleading approach in some clades. Instead, it may be more informative to consider multiple filtered datasets at the outset of phylogenetic analyses. Second, and relatedly, the best filtering strategy will likely depend on the clade of interest and the inference method used. Thus, there may be no globally optimal filtering strategy for all phylogenetic questions.

### *The Nature of Discordance in Marmotine Phylogeny*

The majority of topological discordance that we observed between distinct philosophical approaches to tree-building (concatenation vs. coalescent-based) can be attributed to inability of the concatenation approach to fully accommodate gene tree heterogeneity. Specifically, most of these topological differences were localized to short, internal branches subtending various pulses of ground squirrel diversification, all of which were problematic in previous gene-by-gene studies (Figs. 2-3; *Tamias*, Reid et al.

2012; *Marmota*, Steppan et al. 1999, 2011; *Urocitellus*, Harrison et al. 2003, McLean et al. 2016). The concatenation approach was also more sensitive to dataset filtering than other methods, which is likely related to handling of gene tree heterogeneity as well. Filtering contributed to high instability in concatenated topologies (particularly for <1,000 locus datasets) due to disproportionately high loss of information that occurs (many fewer single nucleotides) relative to summary approaches (fewer gene trees). Conversely, for larger and more inclusive datasets, concatenation may converge on a stable estimate but displays inflated confidence in the form of uniform and high bootstrap support. These findings reinforce the multiple susceptibilities of concatenation-based inferences when gene trees are heterogeneous due to both lack of information and lack of lineage sorting (Salichos and Rokas 2013).

Nevertheless, topological discordance was also found within the summary coalescent framework, suggesting that the above discrepancies are not solely due to the inability of concatenation to account for distinct locus histories. Resolution of clades can also be hampered by paucity of shared, derived characters caused by ancient isolation, even in genomic-scale datasets. Indeed, marmotine history has been marked by repeated colonization of and diversification within varied habitats of temperate North America and Asia, creating opportunities for rapid evolution of lineages. The East Asian lineage *Sciurotamias*, which is morphologically and ecologically distinct among marmotines, is an ostensible example of this. Placement of *Sciurotamias* differed not only among the 3 inference methods, but also within them (Figs. 2-4), suggesting that lineage was isolated early in the history of modern ground squirrels and lacks sufficient synapomorphies in UCEs to be placed with confidence. An alternative possibility is that some of our analyses were biased by long-branch attraction (LBA). LBA due to mutational saturation is unlikely given that UCEs have depressed mutation rates and thus a relatively low likelihood of saturation (McCormack et al. 2012, Gilbert et al. 2015). Our analyses confirm placement of *Sciurotamias* outside Tamiasciurini (as shown by Steppan et al. 2004), but we remain unable to assign it to Marmotini with high certainty. Alternatively, *Sciurotamias* may ultimately need to be accommodated taxonomically within a tribe of its own, Sciurotamiini, as advocated by Kryštufek and Vohralík (2012).

Roure et al. (2013) manipulated a series of empirical phylogenomic datasets and showed that missing data can impact inferences in the total evidence framework, specifically due to impaired parameter estimation, detection of multiple substitutions, and biased branch support. They suggested that smaller and less incomplete datasets are optimal in this philosophical context; specifically, complete removal of loci with missing data was preferable despite the increased loss of characters that this entails. However, the optimality of "smaller" and "less incomplete" matrices likely depends on the dataset at hand, and it is necessary to test the generality of these assertions with other empirical datasets (Roure et al. 2013). Our results differ from those of Roure et al. (2013) in that no significant changes in parameter estimation (e.g., tree length, gamma) in the likelihood framework were observed across the majority of datasets. Also, because UCE cores are highly conserved, it is unlikely that filtering impacts the detection of multiple substitutions. Finally, concatenated analyses based on increasing amounts of missing data tend to produce more similar and better-supported phylogenies (Fig. 4-5). Although not all branches supported in our concatenated analyses may be accurate, our results suggest that filtering UCE datasets to include only highly taxonomically complete loci, as has

often been done, could greatly erode available phylogenetic signal, especially in recent radiations.

*Disagreement among Coalescent-based Methods*

Coalescent-based methods remain preferable to concatenation when gene tree heterogeneity is severe, but our study confirms that disagreement among the former class of methods is possible and, occasionally, pervasive. The nearly ubiquitous lack of topological concordance among our coalescent-based estimates contrasts with a number of previous UCE-based studies in which topological reconciliation was achieved using various filtering approaches (e.g., Hosner et al. 2015, Manthey et al. 2016). This suggests that accurate phylogenetic reconstruction in problematic clades is dependent not just on philosophical stance, but also on understanding how different factors such as TC and PI affect gene tree accuracy and downstream species tree inference.

Simulations have demonstrated that the accuracy of many summary coalescent-based methods improves with increasing numbers of loci (Mirarab et al. 2014, 2016; DeGiorgio and Degnan 2014; Giarla and Esselstyn 2015), but in a way that is sensitive to speciation times and population genetic parameters (i.e., $\theta$, $\mu$). As in Mirarab et al. (2014), we found direct responses of ASTRAL to increasing numbers of loci as tolerances for TC and PI were relaxed. That response involved divergence from the only tree in which ASTRAL and NJst showed concordance (inferred from the smallest dataset, Fig. 3-4) toward a topology more similar to the preferred concatenated tree (inferred from the largest dataset, Fig. 2). But which of those ASTRAL topologies is most accurate? Conversely, Liu and Yu (2011) and DeGiorgio and Degnan (2014) found NJst/STAR to be capable of high accuracy despite relatively low (e.g., 50-100) numbers of loci, less than half the size of our smallest dataset (203 loci). Unfortunately, because dataset size in our study is conflated with TC and PI, comparison to these and other simulation studies (where loci may be more or less uniformly complete and informative) provide only partial insight into this question. Also, simulations are often conducted on small datasets and tree sizes under simplistic evolutionary scenarios, and thus applicability to phylogenomic problems may be limited (Roure et al. 2013).

Alternatively, comparisons to other empirical studies might allow broader insights, especially with popular genomic markers that are highly conserved and may display similar evolutionary features across different taxonomic groups (e.g., UCEs). In their study of the galliform radiation, Hosner et al. (2015) found ASTRAL and other summary methods to be more sensitive to filtering by PI than by TC, consistent with our results (Fig. 4-5). However, those authors believed ASTRAL performed optimally when TC was low but PI was high, which is inconsistent with our results. In a more focused study of gallopheasants, Meiklejohn et al. (2016) found better performance from ASTRAL when uninformative or low-information loci were excluded. Likewise, in a study of tanagers, Manthey et al. (2016) found positive responses of both ASTRAL and STAR when filtering by informativeness. Those empirical examples support previous assessments (e.g., Mirarab et al. 2014) that ASTRAL and other summary methods are sensitive to gene tree precision (and therefore PI), particularly when speciation times are short. We note that our inclusion of 2 individuals for most species could have muted the effects of varying TC on species trees, effectively increasing the chances that 1 sample

per species will be placed correctly on any given gene tree. Still, it is evident that filtering by PI can provide important advantages when studying rapid radiations with UCEs and, possibly, other phylogenomic markers.

While the importance of PI seems intuitive, more attention has historically been given to TC, likely due to technological limits on the size of traditional empirical sequence datasets (and therefore the range of PI). This focus has persisted into the phylogenomic era, with TC being the most commonly applied filter in UCE-based studies to date. Streicher et al. (2016) and Streicher and Wiens (2016) reiterated previous arguments (Phillipe et al. 2004, Wiens 2006, Wiens and Morrill 2011) that a relaxed tolerance for missing taxa can improve phylogenomic inferences. While that approach maximizes the number of loci available for analysis, often leading to increased branch support, the latter metric is not necessarily correlated with accuracy and would thus be a poor measure of quality of inference. For example, we found increased topological concordance between RAxML and ASTRAL (RF distance = 8) in the largest datasets, but significant divergence in branch support values. Similar trends in support values are obtained from NJst (e.g., in Streicher et al. (2016) and this study). These reductions in bootstrap values are an appropriate reflection by coalescent-based methods of uncertainty as more incomplete loci are added (e.g., Giarla and Esselstyn 2015). We suggest that robust conclusions in particularly challenging radiations are more likely to emerge from a process of dataset interrogation (wherein data are manipulated and zones of methodological bias identified) than from seeking to maximize dataset size and/or support values.

*Old and New Uncertainties in Marmotine Phylogeny*

The radiation of marmotine ground squirrels represents a particularly problematic region of the rodent Tree of Life (e.g., Mercer and Roth 2003, Fabre et al. 2012). Nevertheless, it remains well established that marmotine diversification was extremely rapid (Black 1963, Hafner 1984, Zelditch et al. 2015), an apparent evolutionary response to expansion of grass-dominated ecosystems across vast tracts of North America and Asia in the mid-Miocene (Jacobs et al. 1999, Strömberg 2005). The potential for this paleoenvironmental shift to drive the marmotine radiation is supported by numerous other examples of rodent diversification in response to emerging ecological opportunities (e.g., muroids, Schenk et al. 2013, Rowe et al. 2016; octodontoids, Upham and Patterson 2012; sciurids, Mercer and Roth 2003, Hawkins et al. 2016).

Although methodological factors affected phylogenetic estimation in Marmotini, our analyses provide robust support for a number of previous single-locus hypotheses such as the deep split separating chipmunks and all other marmotines, interrelationships among higher chipmunk taxa (i.e., ((*Eutamias*, *Tamias*), *Neotamias*)), and the deep origin of the aridland-adapted antelope ground squirrels (*Ammospermophilus*) and tropical ground squirrels (*Notocitellus*). Historically, however, most conflict in marmotine phylogeny has been concentrated in 2 sequential diversification pulses (Figs. 2-3); the first is a radiation that includes marmots (*Marmota*), rock squirrels (*Otospermophilus*), and golden mantled ground squirrels (*Callospermophilus*); and the second is an ecologically diverse radiation of 6 derived ground squirrel genera (*Spermophilus, Urocitellus, Poliocitellus, Xerospermophilus, Ictidomys,* and *Cynomys*). Our data resolve

some of these conflicts but illuminate new zones of uncertainty that are in need of resolution.

First, most concatenated analyses supported a sister relationship between *Marmota* and (*Otospermophilus*, *Callospermophilus*), which is consistent with mtDNA (Harrison et al. 2003, Zelditch et al. 2015). Conversely, some ASTRAL and NJst topologies favor a closer relationship of (*Otospermophilus*, *Callospermophilus*) with derived marmotines to the exclusion of *Marmota*, which is more consistent with signals from fossil (Black 1963), morphological (Bryant 1945, Hafner 1984), and electrophoretic (Hafner 1984) data. Resolving the sequence of this radiation is important for understanding the evolution of marmots in particular, which are among the most distinctive genera in the tribe in morphology and socioecology (Hafner 1984, Blumstein and Armitage 1998). Second, UCEs provide new insight into the radiation of derived ground squirrels, robust supporting the arrangement (*Xerospermophilus*, (*Ictidomys*, *Cynomys*)), which differs from morphological (Bryant 1945, Hafner 1984), allozymic (Hafner 1984), and mtDNA data (Harrison et al. 2003). Such high level of support for this relationship among datasets and inference methods were unexpected given fossil evidence for the rapid Pleistocene evolution of prairie dogs (*Cynomys*; Black 1963, Goodwin 1995), and our results provide a new framework for understanding the biology of this genus. However, conflict persists in the placement of both *Urocitellus* (Holarctic ground squirrels) and the monotypic *Poliocitellus* (Franklin's ground squirrel). Most remarkable is support for non-monophyly of *Urocitellus* in the majority of ASTRAL and NJst trees (Figs. 3,6). Although that genus is informally divided into "big-eared" and "small-eared" groups based on major differences in body size, habitat, and ecology (Howell 1938, Helgen et al. 2009, McLean et al. 2016), there is little evidence to date suggesting that *Urocitellus* is not a monophyletic assemblage. These findings demand further attention and possibly additional genomic markers to resolve.

## CONCLUSIONS

Taken together, our results echo the complexities of resolving ancient, rapid radiations, even when thousands of genomic markers are available. When phylogenetic signal is low due to lack of synapomorphies caused by extreme ILS, ancient isolation of lineages, or both, sensitivity of historical inferences to methodological factors can be high. Even with a robust taxon and locus sampling strategy, methodological biases can accrue from dataset filtering, in the estimation of gene trees, and by strict adherence to a particular inference philosophy or software/method. In clades known to have experienced rapid radiation, it is critical to build confidence not only through evaluating concordance of point estimates of phylogeny and/or branch support, but also via simulations on known species phylogenies (e.g., Giarla and Esselstyn 2015), integration across a sufficiently wide variety of filtering parameters and inference methods (e.g., Hosner et al. 2015, Manthey et al. 2016, this study), or other means. Doing so is imperative for identifying zones of suboptimal methodological performance and ultimately for a more accurately resolved Tree of Life.

ACKNOWLEDGEMENTS

This work was funded by the National Science Foundation (NSF 1258010), Smithsonian Institution (Peter Buck Predoctoral Fellowship to B.S.M.), American Museum of Natural History (Theodore Roosevelt Memorial Research Grant to B.S.M.), American Society of Mammalogists (Grant-in-Aid of Research to B.S.M. and ASM Fellowship to K.C.B.), and the University of New Mexico Department of Biology and Graduate and Professional Student Association. We thank many museum curators and collection managers for access to important specimens, particularly Jonathan Dunnum and Mariel Campbell (MSB) and Darrin Lunde and Chris Huddleston (USNM). We thank John McCormack and Whitney Tsai for help and advice on the initial phase of UCE library preparation, James Degnan for discussion of the ideas presented here and for comments on an earlier draft of this manuscript, and Siavash Mirarab for advice on the bootstrapping procedure implemented in ASTRAL.

REFERENCES

Allen, J.M., Huang, D.I., Cronk, Q.C., Johnson, K.P., 2015. aTRAM - automated target restricted assembly method: a fast method for assembling loci across divergent taxa from next-generation sequencing data. BMC Bioinformatics 16:98.

Allman E.S., Degnan J.H., Rhodes J.A. 2016. Species tree inference from gene splits by Unrooted STAR methods. EEE/ACM Trans. Computational Biology and Bioinformatics In press.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S., Haussler, D., 2004. Ultraconserved elements in the human genome. Science 304:1321–1325.

Black, C.C., 1972. Holarctic Evolution and Dispersal of Squirrels (Rodentia: Sciuridae), in: Dobzhansky, T., Hecht, M.K., Steere, W.C. (Eds.), Evolutionary Biology. Appleton-Century-Crofts, Meredith Corporation, New York, pp. 305–322.

Black, C.C., 1963. A review of the North American Tertiary Sciuridae. Bulletin of the Museum Comparative Zoology of Harvard University 130:109–248.

Blumstein D.T., Armitage K.B. 1998. Life history consequences of social complexity: a comparative study of ground-dwelling sciurids. Behavioral Ecology 9:8–19.

Borowiec, M.L., 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. PeerJ 4, e1660.

Bryant, M.D., 1945. Phylogeny of Nearctic Sciuridae. American Midland Naturalist 33:257–390.

DeGiorgio, M., Degnan, J.H., 2014. Robustness to divergence time underestimation when inferring species trees from estimated gene trees. Systematic Biology 63:66–82.

Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely gene trees. PLoS Genet. 2:e68.

Edwards, S.V, 2009. Is a new and general theory of molecular systematics emerging? Evolution 63:1–19.

Fabre, P.-H., Hautier, L., Dimitrov, D., Douzery, E.J.P., 2012. A glimpse on the pattern of rodent diversification: a phylogenetic approach. BMC Evolutionary Biology 12:88.

Faircloth, B.C., 2015. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics 32:786–788.

Faircloth, B.C., McCormack, J.E., Crawford, N.G., Harvey, M.G., Brumfield, R.T., Glenn, T.C., 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Systematic Biology 61:717–726.

Felsenstein, J., 2005. PHYLIP (Phylogeny Inference Package).

Gatesy, J., Springer, M.S., 2014. Phylogenetic analysis at deep timescales: unreliable gene trees, bypassed hidden support, and the coalescence/concatalescence conundrum. Molecular Phylogenetics and Evolution 80:231–266.

Giarla, T.C., Esselstyn, J.A., 2015. The challenges of resolving a rapid, recent radiation: empirical and simulated phylogenomics of Philippine shrews. Systematic Biology 64:727–740.

Gilbert, P.S., Chang, J., Pan, C., Sobel, E.M., Sinsheimer, J.S., Faircloth, B.C., Alfaro, M.E., 2015. Genome-wide ultraconserved elements exhibit higher phylogenetic informativeness than traditional gene markers in percomorph fishes. Molecular Phylogenetics and Evolution 92:140–146.

Goodwin, H.T., 2008. Sciuridae, in: Janis, Christine M., Gunnell, G.F., Uhen, M.D. (Eds.), Evolution of Tertiary Mammals of North America: Small Mammals, Xenarthrans, and Marine Mammals. Cambridge University Press, Cambridge, pp. 355–376.

Goodwin, H.T., 1995. Pliocene-Pleistocene biogeographic history of prairie dogs, genus *Cynomys* (Sciuridae). Journal of Mammalogy 76:100–122.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology 29:644–652.

Hafner, D.J., 1984. Evolutionary Relationships of the Nearctic Sciuridae, in: Murie, J.O., Michener, G.R. (Eds.), The Biology of Ground-Dwelling Sciurids. University of Nebraska Press, Lincoln, pp. 3–23.

Harrison, R.G., Bogdanowicz, S.M., Hoffmann, R.S., Yensen, E., Sherman, P.W., 2003. Phylogeny and evolutionary history of the ground squirrels (Rodentia: Marmotinae). Journal of Mammalian Evolution 10:249–276.

Hawkins, M.T.R., Leonard, J.A., Helgen, K.M., McDonough, M.M., Rockwood, L.L., Maldonado, J.E., 2016. Evolutionary history of endemic Sulawesi squirrels constructed from UCEs and mitogenomes sequenced from museum specimens. BMC Evolutionary Biology 16:80.

Heibl, C. 2015. ips v0.0-7. https://cran.r-project.org/web/packages/ips/index.html.

Helgen, K.M., Cole, F.R., Helgen, L.E., Wilson, D.E., 2009. Generic revision in the Holarctic ground squirrel genus *Spermophilus*. Journal of Mammalogy 90:270–305.

Hosner, P.A., Faircloth, B.C., Glenn, T.C., Braun, E.L., Kimball, R.T., 2016. Avoiding missing data biases in phylogenomic inference: an empirical study in the landfowl (Aves: Galliformes). Molecular Biology and Evolution 33:1110–1125.

Howell, A.H., 1938. Revision of the North American Ground Squirrels, with a Classification of the North American Sciuridae. North American Fauna 56, 1–256.

Huang, H., Knowles, L.L., 2016. Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of rad sequences. Systematic Biology 65:357–365.

Huelsenbeck, J., 1991. When are fossil better than extant taxa in phylogenetic analysis? Systematic Zoology 40:458–469.

Huelsenbeck, J.P., 1995. Performance of phylogenetic methods in simulation. Systematic Biology 44:17–48.

Jacobs, B.F., Kingston, J.D., Jacobs, L.L., 1999. The origin of grass-dominated ecosystems. Annals of the Missouri Botanical Gardens 86:590–643.

Janečka, J.E., Miller, W., Pringle, T.H., Wiens, F., Zitzmann, A., Helgen, K.M., Springer, M.S., Murphy, W.J., 2007. Molecular and genomic data identify the closest living relative of primates. Science 318:792–794.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30:772–780.

Krystufek B., Vohralik V. 2013. Taxonomic revision of the Palaearctic rodents (Rodentia). Part 2. Sciuridae: *Urocitellus*, *Marmota* and *Sciurotamias*. Lynx. 44:27–138.

Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Systematic Biology 56:17–24.

Leaché, A.D., Chavez, A.S., Jones, L.N., Grummer, J.A., Gottscho, A.D., Linkem, C.W., 2015. Phylogenomics of phrynosomatid lizards: conflicting signals from sequence capture versus restriction site associated DNA sequencing. Genome Biology and Evolution 7:706–719.

Lemmon A.R., Brown J.M., Stanger-Hall K., Lemmon E.M. 2009. The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and Bayesian inference. Systematic Biology 58:130–145.

Liu, L., Yu, L., 2011. Estimating species trees from unrooted gene trees. Systematic Biology 60:661–667.

Liu, L., Yu, L., Pearl, D.K., Edwards, S. V, 2009. Estimating species phylogenies using coalescence times among sequences. Systematic Biology 58:468–77.

Maddison, W.P., 1997. Gene trees in species trees. Systematic Biology 46:523–536.

Maddison, W.P., Knowles, L.L., 2006. Inferring phylogeny despite incomplete lineage sorting. Systematic Biology 55:21–30.

Manthey, J.D., Campillo, L.C., Burns, K.J., Moyle, R.G., 2016. Comparison of target-capture and restriction-site associated DNA sequencing for phylogenomics: a test in cardinalid tanagers (Aves, Genus: Piranga). Systematic Biology 65:640–650.

McCormack, J.E., Faircloth, B.C., Crawford, N.G., Gowaty, P.A., Brumfield, R.T., Glenn, T.C., 2012. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. Genome Research 22:746–754.

McCormack, J.E., Harvey, M.G., Faircloth, B.C., Crawford, N.G., Glenn, T.C., Brumfield, R.T., 2013. A phylogeny of birds based on over 1,500 loci collected by target enrichment and high-throughput sequencing. PLoS One 8:e54848.

McLean, B.S., Jackson, D.J., Cook, J. a., 2016. Rapid divergence and gene flow at high latitudes shape the history of Holarctic ground squirrels (*Urocitellus*). Molecular Phylogenetics and Evolution 102:174–188.

Meiklejohn, K.A., Faircloth, B.C., Glenn, T.C., Kimball, R.T., Braun, E.L., 2016. Analysis of a rapid evolutionary radiation using ultraconserved elements: evidence for a bias in some multispecies coalescent methods. Systematic Biology 65:612–627.

Mercer, J.M., Roth, V.L., 2003. The effects of Cenozoic global change on squirrel phylogeny. Science 299:1568–1572.

Mirarab, S., Bayzid, M.S., Warnow, T., 2016. Evaluating summary methods for multilocus species tree estimation in the presence of incomplete lineage sorting. Systematic Biology 65:366–380.

Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S., Warnow, T., 2014. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics 30:i541–i548.

Paradis, E., B. Bolker, J. Claude, H. S. Cuong, R. Desper, B. Durand, J. Dutheil, O. Gascuel, C. Heibl, D. Lawson, V. Lefort, P. Legendre, J. Lemon, Y. Noel, J. Nylander, R. Opgen-Rhein, A.-A. Popescu, K. Schliep, K. Strimmer, de Vienne, D. 2015. ape: analyses of phylogenetics and evolution v4.1. https://cran.r-project.org/web/packages/ape/index.html.

Patterson B.D., Norris R.W. 2016. Towards a uniform nomenclature for ground squirrels: the status of the Holarctic chipmunks. Mammalia. 80:241–251.

Pease, J.B., Haak, D.C., Hahn, M.W., Moyle, L.C., 2016. Phylogenomics reveals three sources of adaptive variation during a rapid radiation. PLoS Biol. 14:e1002379.

Philippe, H., Snell, E.A., Bapteste, E., Lopez, P., Holland, P.W.H., Casane, D., 2004. Phylogenomics of eukaryotes: impact of missing data on large alignments. Molecular Biology and Evolution 21:1740–1752.

Piaggio, A., Spicer, G.S., 2001. Molecular phylogeny of the chipmunks inferred from mitochondrial cytochrome b and cytochrome oxidase II gene sequences. Molecular Phylogenetics and Evolution 20:335–350.

Poe, S., 1998. Sensitivity of phylogeny estimation to taxonomic sampling. Systematic Biology 47:18–31.

Reid N., Demboski J.R., Sullivan J. 2012. Phylogeny estimation of the radiation of western North American chipmunks (*Tamias*) in the face of introgression using reproductive protein genes. Systematic Biology 61:44–62.

Roure B., Baurain D., Philippe H. 2013. Impact of missing data on phylogenies inferred from empirical phylogenomic data sets. Molecular Biology and Evolution 30:197–214.

Rowe K.C., Achmadi A.S., Esselstyn J.A. 2016. Repeated evolution of carnivory among Indo-Australian rodents. Evolution 70:653–665.

Salichos L., Rokas A. 2013. Inferring ancient divergences requires genes with strong phylogenetic signals. Nature 497:327–331.

Schliep, K., Paradis, E., Potts, A., Kendall, M., 2016. phangorn v2.0.4. https://cran.r-project.org/package=phangorn

Schenk, J.J., Rowe, K.C., Steppan, S.J., 2013. Ecological opportunity and incumbency in the diversification of repeated continental colonizations by muroid rodents. Systematic Biology 62:837–864.

Song, S., Liu, L., Edwards, S.V., Wu, S., 2012. Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. Proceedings of the National Academy of Sciences 109:14942–14947.

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

Steppan, S., Kenagy, G., Zawadzki, C., Robles, R., Lyapunova, E., Hoffmann, R.S., 2011. Molecular data resolve placement of the Olympic marmot and estimate dates of trans-Beringian interchange. Journal of Mammalogy 92:1028–1037.

Steppan, S.J., Akhverdyan, M.R., Lyapunova, E.A., Fraser, D.G., Vorontsov, N.N., Hoffmann, R.S., Braun, M.J., 1999. Molecular phylogeny of the marmots (Rodentia: Sciuridae): tests of evolutionary and biogeographic hypotheses. Systematic Biology 48:715–734.

Steppan, S.J., Storz, B.L., Hoffmann, R.S., 2004. Nuclear DNA phylogeny of the squirrels (Mammalia: Rodentia) and the evolution of arboreality from c-myc and RAG1. Molecular Phylogenetics and Evolution 30:703–719.

Streicher, J.W., Schulte, J.A., Wiens, J.J., 2016. How should genes and taxa be sampled for phylogenomic analyses with missing data? An empirical study in iguanian lizards. Systematic Biology 65:128–145.

Streicher, J.W., Wiens, J.J., 2016. Phylogenomic analyses reveal novel relationships among snake families. Molecular Phylogenetics and Evolution 100:160–169.

Strömberg, C.A.E., 2005. Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the Cenozoic of North America. Proceedings of the National Academy of Sciences 102:11980–11984.

Thorington, R.W., Koprowski, J.L., Steele, M.L., Whatton, J.F., 2012. Squirrels of the World. The Johns Hopkins University Press, Baltimore.

Townsend, J.P., 2007. Profiling phylogenetic informativeness. Systematic Biology 56:222–231.

Upham, N.S., Patterson, B.D., 2012. Diversification and biogeography of the Neotropical caviomorph lineage Octodontoidea (Rodentia: Hystricognathi). Molecular Phylogenetics and Evolution 63:417–429.

Wagner, C.E., Keller, I., Wittwer, S., Selz, O.M., Mwaiko, S., Greuter, L., Sivasundar, A., Seehausen, O., 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. Molecular Ecology 22:787–798.

Wiens, J.J., 2006. Missing data and the design of phylogenetic analyses. Journal of Biomedical Informatics 39:34–42.

Wiens, J.J., Morrill, M.C., 2011. Missing data in phylogenetic analysis: Reconciling results from simulations and empirical data. Systematic Biology 60:719–731.

Wilson, D.E., Reeder, D.M. (Eds.), 2005. Mammal Species of the World, 3rd ed. Johns Hopkins University Press, Baltimore, Maryland.

Ye, C., Liang, S., Yan, Z., Zepeda, M.L., Campos, P.F., Missael, A., Velazquez, V., Samaniego, J.A., Avila-arcos, M., Martin, M.D., Barnett, R., Ribeiro, A.M., Mello, C. V, Lovell, P. V, Almeida, D., Maldonado, E., Pereira, J., Sunagar, K., Philip, S., Dominguez-bello, M.G., Bunce, M., Lambert, D., Brumfield, R.T., Sheldon, F.H., Holmes, E.C., Gardner, P.P., Steeves, T.E., Stadler, P.F., Burge, S.W., Li, C., Ho, S.Y.W., Faircloth, B.C., Nabholz, B., 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science 346:1320–1331.

Zelditch, M.L., Li, J., Tran, L.A.P., Swiderski, D.L., 2015. Relationships of diversity, disparity, and their evolutionary rates in squirrels (Sciuridae). Evolution 69:1284–1300.

FIGURE CAPTIONS

Figure 1. Effects of filtering approaches used in this paper on dataset characteristics: a) dataset size (in numbers of UCE loci), b) mean bootstrap support of gene trees (in %), and c) precision of gene trees (expressed as the coefficient of variation (CV; σ/μ) of bootstrap support values). Note all datasets are nested subsets of the largest filtered dataset (≥50% taxonomic completeness, ≥0 parsimony informative sites). See text for further details.

Figure 2. Phylogenetic hypothesis of tribe Marmotini based on concatenated analysis of the largest filtered dataset (3,919 UCE loci, ≥50% taxonomic completeness, ≥0 parsimony informative sites). The tree was inferred in RAxML under a GTRGAMMA model of evolution and support was assessed from 100 rapid bootstraps. All splits received high support (>90%) except 1 (placement of *Sciurotamias*, indicated on figure). Colors on tips indicate continent of origin; Urocitellus parryii is distributed in both North America and Asia.

Figure 3. Phylogenetic hypotheses of Marmotini inferred in ASTRAL under different filtering scenarios. a) the topology inferred under the strictest filtering regime and the only tree in which ASTRAL and NJst were concordant b) the topology inferred under the most inclusive filtering regime, i.e., that used to construct the tree in Fig. 2. Bootstrap support values are listed and branches are colored according to strength of support to aid visualization (stronger support = darker coloration).

Figure 4. Effects of dataset filtering on phylogenomic inferences in b) RAxML, c) ASTRAL, and d) NJst. The schematic in a) depicts general changes in total taxonomic completeness and phylogenetic informativeness with the applicaton of different filters. Colors on heatmaps indicate Robinson-Foulds distances from the concatenated tree shown in Figure 2. Note the different scales for each heatmap.

Figure 5. Effects of dataset filtering on bootstrap support values in RAxML (top row), ASTRAL (middle row), and NJst (bottom row). a-c) mean bootstrap values, d-f) topological precision expressed as coefficient of variation (σ/μ) of bootstrap support values. Note slightly different scales for each heatmap.

Figure 6. Conflict in inferred relationships among derived marmotine ground squirrel genera. Each tree is the majority rule consensus of 42 trees inferred on nested datasets of varying taxonomic completeness and phylogenetic informativeness, as described in text. All branches were supported across 100% of trees unless otherwise indicated.

TABLES

Table 1. Summary statistics for UCE alignments, computed prior to filtering. Data are presented as medians with first and third quantiles where applicable. Top: outgroups and ingroups, Bottom: ingroup taxa only.

| | Samples | Loci | Locus Length | Taxa per Locus | Percent Variable Sites | Percent Informative Sites |
|---|---|---|---|---|---|---|
| Marmotini + outgroups | 73 | 4402 | 590.5 (475, 685) | 66 (55, 70) | 16.2 (10.2, 24.2) | 9.0 (4.7, 15.1) |
| Marmotini only | 67 | 4401 | 594 (479, 689) | 61 (50, 64) | 12.5 (7.9, 18.4) | 6.7 (3.6, 11.1) |

FIGURES

Figure 1.

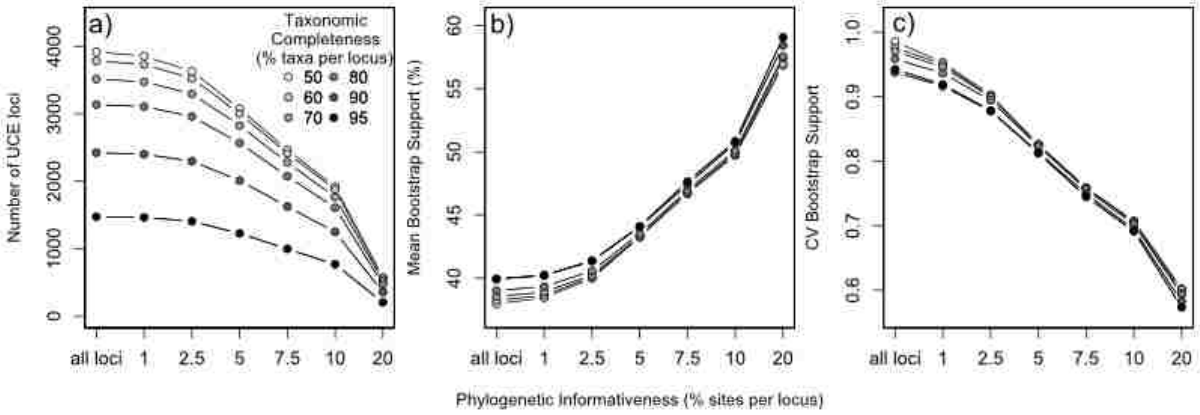


Figure 2.

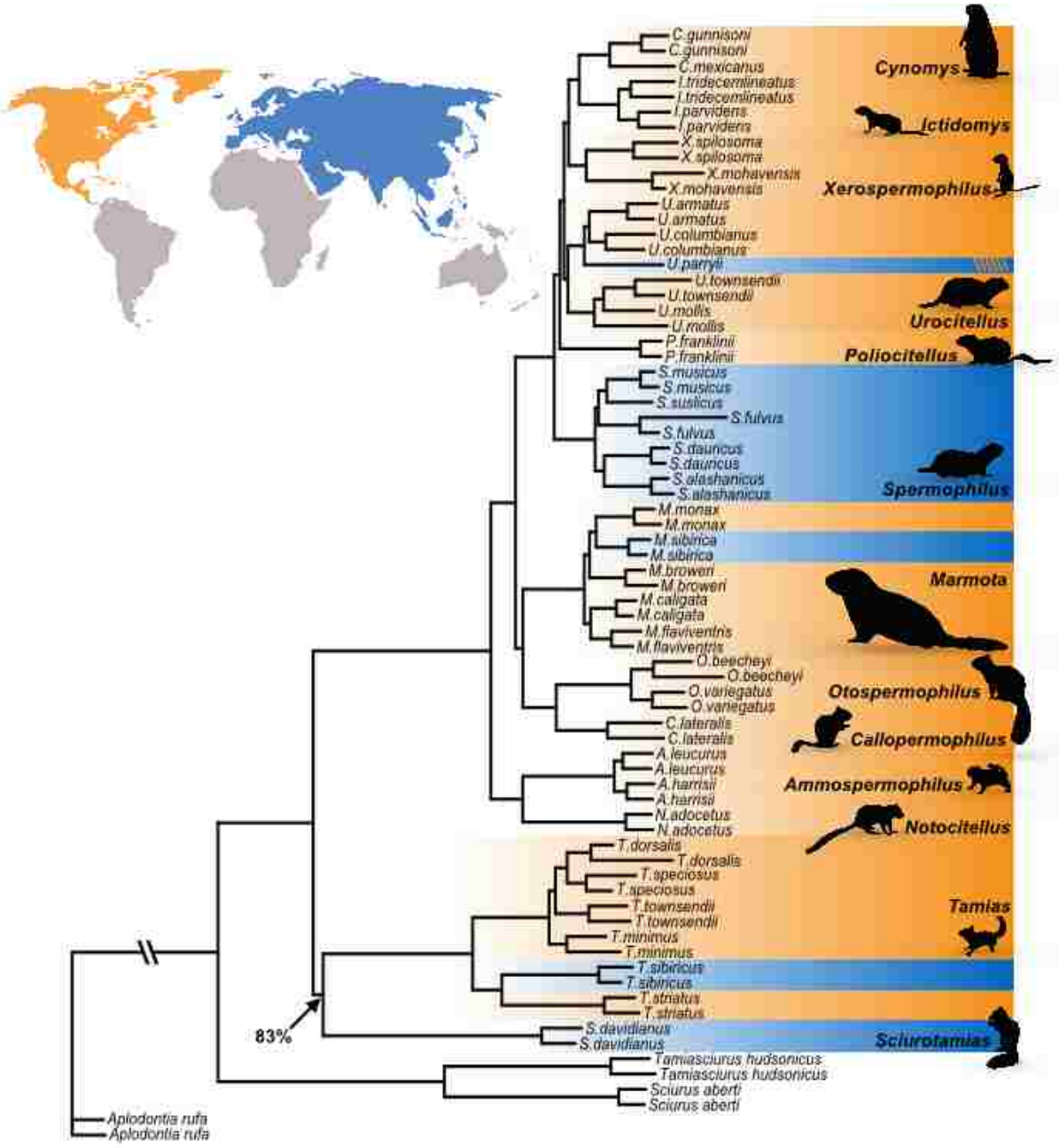

C.gunnisoni
C.gunnisoni
C.mexicanus
I.tridecemlineatus
I.tridecemlineatus
I.parvidens
I.parvidens
X.spilosoma
X.spilosoma
X.mohavensis
X.mohavensis
U.armatus
U.armatus
U.columbianus
U.columbianus
U.parryii
U.townsendii
U.townsendii
U.mollis
U.mollis
P.franklinii
P.franklinii
S.musicus
S.musicus
S.suslicus
S.fulvus
S.fulvus
S.dauricus
S.dauricus
S.alashanicus
S.alashanicus
M.monax
M.monax
M.sibirica
M.sibirica
M.broweri
M.broweri
M.caligata
M.caligata
M.flaviventris
M.flaviventris
O.beecheyi
O.beecheyi
O.variegatus
O.variegatus
C.lateralis
C.lateralis
A.leucurus
A.leucurus
A.harrisii
A.harrisii
N.adocetus
N.adocetus
T.dorsalis
T.dorsalis
T.speciosus
T.speciosus
T.townsendii
T.townsendii
T.minimus
T.minimus
T.sibiricus
T.sibiricus
T.striatus
T.striatus
S.davidianus
S.davidianus
Tamiasciurus hudsonicus
Tamiasciurus hudsonicus
Sciurus aberti
Sciurus aberti
Aplodontia rufa
Aplodontia rufa
83%
Cynomys
Ictidomys
Xerospermophilus
Urocitellus
Poliocitellus
Spermophilus
Marmota
Otospermophilus
Callopermophilus
Ammospermophilus
Notocitellus
Tamias
Sciurotamias

Figure 3.

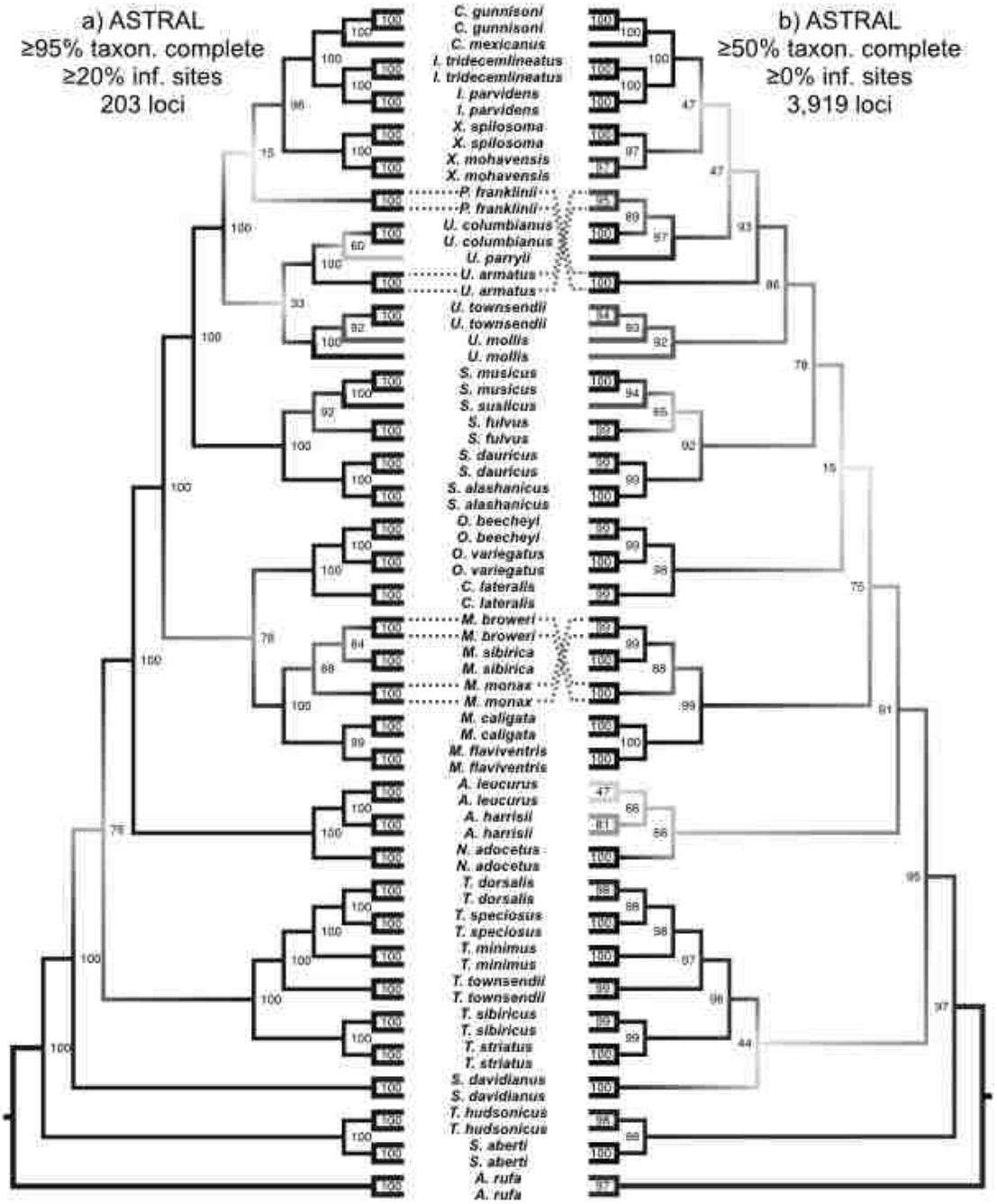

a) ASTRAL
≥95% taxon. complete
≥20% inf. sites
203 loci
b) ASTRAL
≥50% taxon. complete
≥0% inf. sites
3,919 loci
C. gunnisoni
C. gunnisoni
C. mexicanus
I. tridecemlineatus
I. tridecemlineatus
I. parvidens
I. parvidens
X. spilosoma
X. spilosoma
X. mohavensis
X. mohavensis
P. franklinii
P. franklinii
U. columbianus
U. columbianus
U. parryii
U. armatus
U. armatus
U. townsendii
U. townsendii
U. mollis
U. mollis
S. musicus
S. musicus
S. suslicus
S. fulvus
S. fulvus
S. dauricus
S. dauricus
S. alashanicus
S. alashanicus
O. beecheyi
O. beecheyi
O. variegatus
O. variegatus
C. lateralis
C. lateralis
M. broweri
M. broweri
M. sibirica
M. sibirica
M. monax
M. monax
M. caligata
M. caligata
M. flaviventris
M. flaviventris
A. leucurus
A. leucurus
A. harrisii
A. harrisii
N. adocetus
N. adocetus
T. dorsalis
T. dorsalis
T. speciosus
T. speciosus
T. minimus
T. minimus
T. townsendii
T. townsendii
T. sibiricus
T. sibiricus
T. striatus
T. striatus
S. davidianus
S. davidianus
T. hudsonicus
T. hudsonicus
S. aberti
S. aberti
A. rufa
A. rufa

Figure 4.

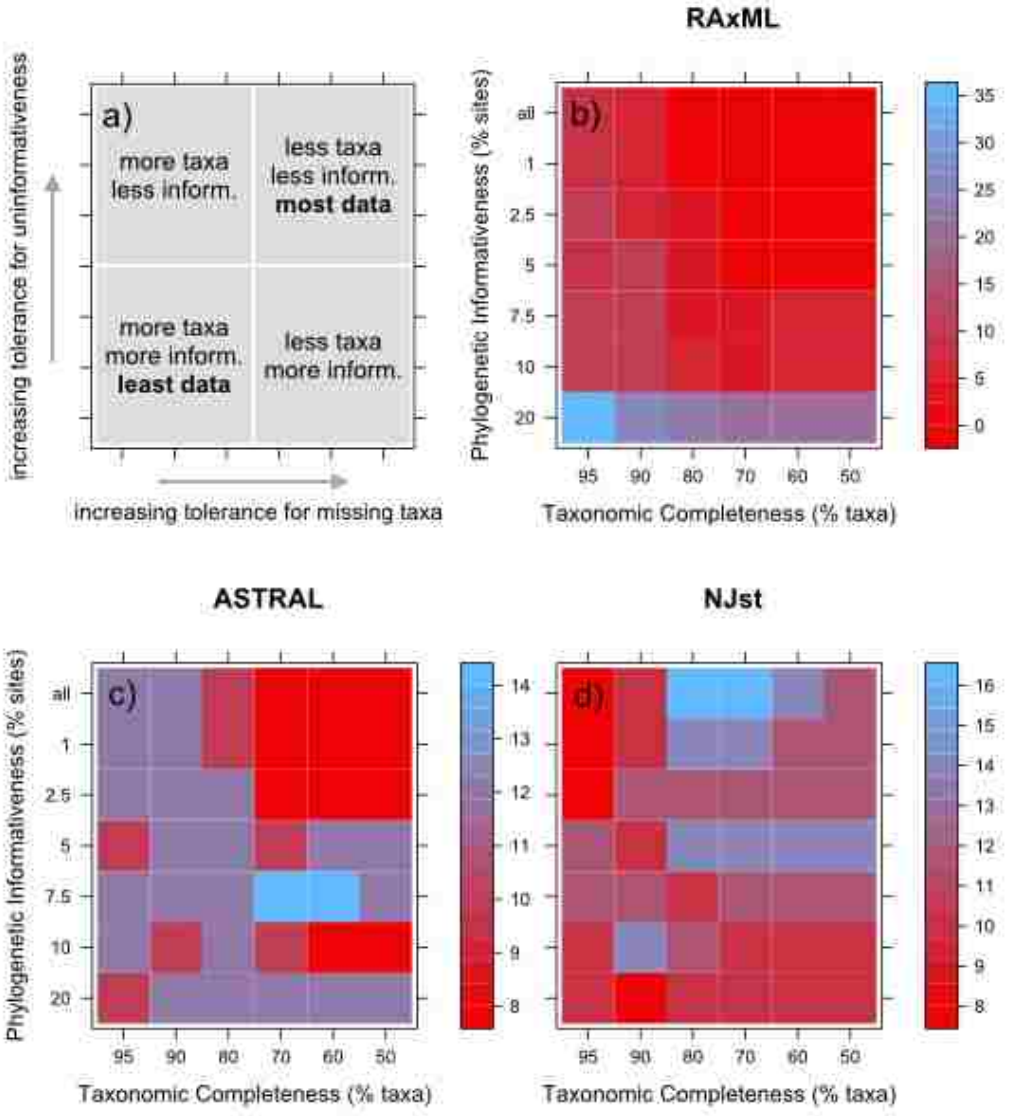

RAxML
ASTRAL
NJst
a)
more taxa
less inform.
less taxa
less inform.
most data
more taxa
more inform.
least data
less taxa
more inform.
increasing tolerance for uninformativeness
increasing tolerance for missing taxa
b)
c)
d)
Phylogenetic Informativeness (% sites)
all
1
2.5
5
7.5
10
20
95
90
80
70
60
50
Taxonomic Completeness (% taxa)

Figure 5.

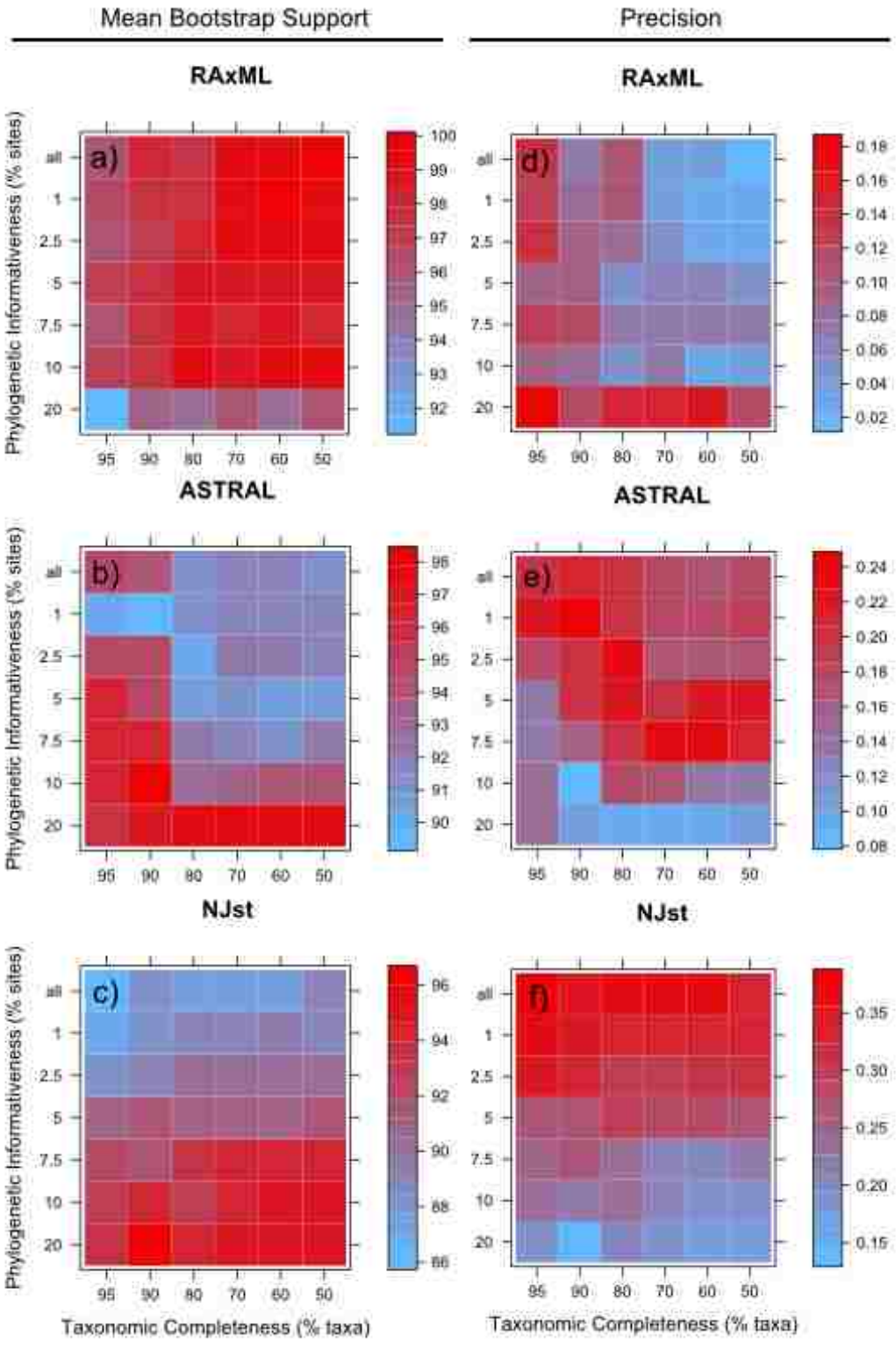
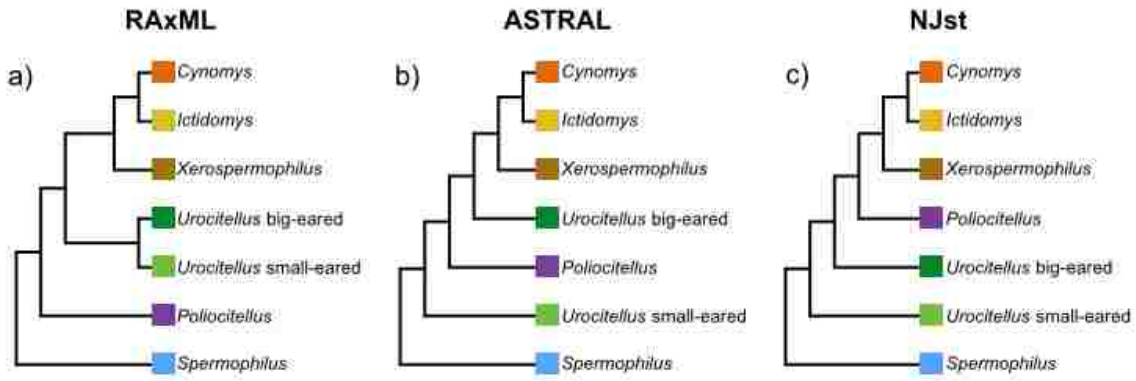

Mean Bootstrap Support
Precision
RAxML
RAxML
ASTRAL
ASTRAL
NJst
NJst
a)
b)
c)
d)
e)
f)
Phylogenetic Informativeness (% sites)
Taxonomic Completeness (% taxa)
all
1
2.5
5
7.5
10
20
95
90
80
70
60
50

Figure 6.


RAxML
a)
Cynomys
Ictidomys
Xerospermophilus
Urocitellus big-eared
Urocitellus small-eared
Poliocitellus
Spermophilus
ASTRAL
b)
Cynomys
Ictidomys
Xerospermophilus
Urocitellus big-eared
Poliocitellus
Urocitellus small-eared
Spermophilus
NJst
c)
Cynomys
Ictidomys
Xerospermophilus
Poliocitellus
Urocitellus big-eared
Urocitellus small-eared
Spermophilus

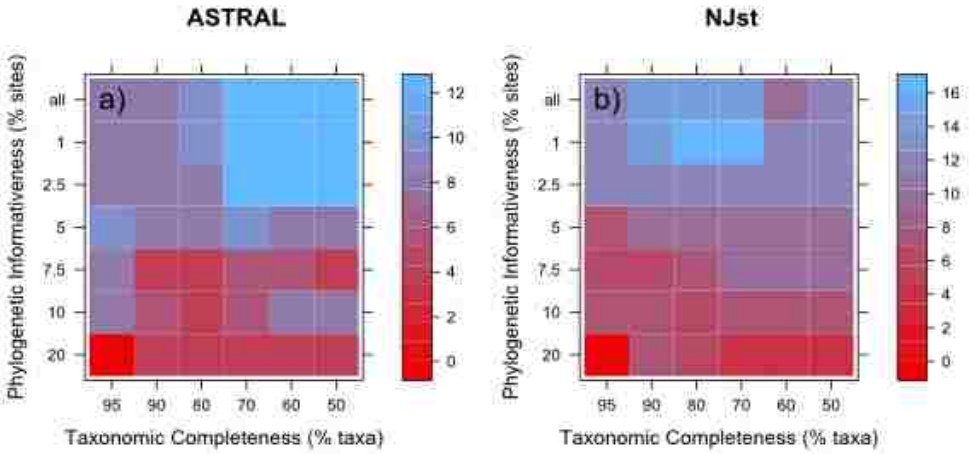
SUPPLEMENTARY FIGURE CAPTIONS

Supplementary Figure 1. Effects of dataset filtering on phylogenomic inference in a) ASTRAL and b) NJst. For each of the 42 datasets, Robinson-Foulds distances were computed between reconstructions and the only tree in which the 2 methods were concordant (i.e., the strictest filtering scenario; lower left corner in heatmaps; see Fig 3a) as well). Note the slightly different scales for each heatmap.

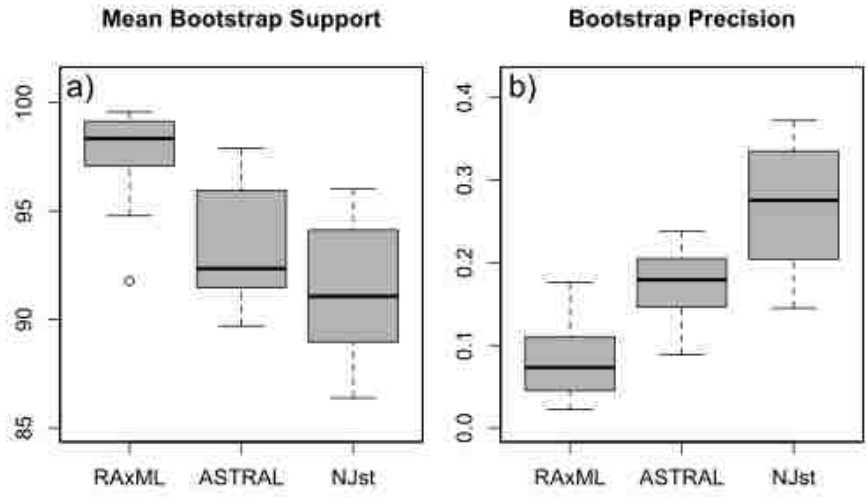
Supplementary Figure 2. Effects of different inference methods on a) mean bootstrap support and b) overall precision (coefficient of variation of bootstrap values) for trees inferred from the 42 nested UCE datasets of varying taxonomic completeness and phylogenetic informativeness.

SUPPLEMENTARY FIGURES

Supplementary Figure 1.



Supplementary Figure 2.



# CHAPTER 2

## Heterogeneity in evolutionary mode and macroevolutionary dynamics among ground squirrel morphological traits

AUTHORS
Bryan S. McLean[1], Kristofer M. Helgen[2], Joseph A. Cook[1]

[1] Department of Biology and Museum of Southwestern Biology, 1 University of New Mexico, MSC03-2020, Albuquerque, NM 87131 USA
[2] School of Biological Sciences, University of Adelaide, Adelaide, SA 5005 Australia

ABSTRACT

Our understanding of the dynamics of phenotypic evolution in deep time is often constrained to inferences drawn from one or a few traits. Yet, phenomes of complex organisms include many traits that may differ in evolutionary mode due to distinct selective pressures, genetic or developmental constraints, and/or covariation with other traits. The potential consequences of this among-trait heterogeneity for macroevolutionary inference are significant, yet largely unexplored for most clades. We used a novel phylogenomic hypothesis and multiple morphological datasets to assess the extent of evolutionary heterogeneity among 4 ecomorphological traits (cranial, mandibular, molariform-tooth shape, and body size) in an extant radiation of ground-dwelling squirrels (Xerinae: Marmotini). Despite significant covariation among functionally related skull traits, we found evidence for evolution via different modes in cranial and mandibular shape (OU with 2 adaptive regimes) relative to molariform-tooth shape and body size (random walk). However, even skull traits evolving under a similar evolutionary model (OU) displayed divergent dynamics on the macroevolutionary landscape, including differences in the frequency of convergent shifts in morphospace. Among-trait variation of this magnitude reiterates the mosaic nature of phenotypic evolution and suggests extreme caution is necessary when drawing macroevolutionary conclusions from small numbers of traits assumed to be involved in the process of taxonomic and ecophenotypic diversification.

KEYWORDS

Radiation, ecomorphology, convergence, morphological integration, geometric morphometrics, covariance ratio

## INTRODUCTION

A basic goal in evolutionary biology is to understand the origins of phenotypic diversity and its apportionment across the Tree of Life. Many investigations focused on describing the processes that have produced patterns of whole-organism phenotypic variation nevertheless employ relatively small numbers of traits, a sampling bias attributable to the autocorrelations that often exist among larger numbers of traits (e.g., due to allometry); basic constraints on researcher time and effort, especially when studying large clades; or access to morphologically complete specimens (e.g., for rare extant lineages or fossil lineages with poor preservation). Yet, the phenomes of complex organisms are comprised of many traits whose evolution is shaped by within- and among-trait processes operating at the genetic, developmental, or functional level(s), which may generate substantial heterogeneity in tempo and mode of trait evolution. This among-trait heterogeneity is beginning to be more fully understood for fossil lineages (Cheetham 1987, Hunt 2007, Hopkins and Lidgard 2012, Hopkins and Smith 2015, Hunt et al. 2015); however, the true extent of this heterogeneity and its potential consequences for macroevolutionary inference remain largely unexplored for most clades of organisms, including extant groups.

Evolutionary heterogeneity among morphological traits might arise from their unique functional roles, which subject them to distinct – and potentially dynamic – selective pressures. For example, Harmon et al. (2005) demonstrated that different morphological character "systems" in *Anolis* are each related to environmental variables (i.e., are adaptive), but vary with respect to each other and in their ability to distinguish ecomorphs. However, trait heterogeneity can arise from processes other than adaptation (e.g., Gould and Lewontin 1979, Futuyma 2010), including genetic drift, differences in genetic or developmental constraints, or integration among different sets of genes (e.g., pleiotropy) and traits (e.g., allometry, shared function). Integration in particular can impose significant conditionality on phenotypic evolution, structuring morphospace so that evolutionary change is restricted to a developmentally or functionally possible subset of changes (Pigliucci 2003, Rolian et al. 2010, Gerber 2013, Conner et al. 2014, Klingenberg 2014). Integration is also capable of producing greater disparity than that resulting from non-integrated traits over macroevolutionary timescales, including convergence on advantageous phenotypes (Gould and Lewontin 1979, Goswami et al. 2014). Understanding the complex mosaic of phenotypic evolution therefore requires parsing the contributions of within- as well as among-trait processes.

The skull is perhaps the most widely studied morphological structure with regard to evolutionary pattern and process in vertebrates. Although comprising only a small portion of the vertebrate corpus, the skull is a structure of manifold ecological significance, housing hard and soft structures that are essential in sensory perception, procurement and processing of food, respiration, defense, and cognition, each of which varies greatly within and among vertebrate clades. Such multifunctionality means that skull evolution is potentially a product of multiple evolutionary forces (e.g., adaptation, constraint, drift) operating on traits of varying function, as well as the magnitude and form of integration among them (e.g., Cheverud 1982, Marroig and Cheverud 2001, Hallgrimsson et al. 2007, Monteiro and Nogueira 2010). Capturing the evolutionary heterogeneity among skull components requires description of increasing numbers of

traits in richer detail, e.g., with the tools of landmark-based geometric morphometrics (Pigliucci 2003, Cardini and Elton 2008). Unfortunately, there is often a trade-off between number of traits and the detail in which they are sampled, and thus many recent macroevolutionary studies in mammals have been limited to use of only 1 or 2 structures (e.g., cranium and/or mandible).

We examined patterns of evolutionary variation and covariation in body size as well as 3 different skull traits (cranial, mandibular, and molariform tooth shape) sampled from a classic mammalian radiation, ground-dwelling squirrels of the tribe Marmotini (chipmunks, ground squirrels, marmots, prairie dogs; >91 species in 13 genera; Wilson and Reeder 2005, Helgen et al. 2009). Members of this clade occur across temperate, subarctic, and arctic biomes; vary by nearly 2 orders of magnitude in body size (Hayssen 2008); and display important differences in ecological modality ranging from granivorous and scansorial forms to herbivorous, burrowing, and strictly terrestrial forms (Hafner 1984, Helgen et al. 2009, Thorington et al. 2012). Nevertheless, ecomorphological diversity in extinct and extant Marmotini has generally been interpreted as low, owing to within-clade constraint as well as among-clade convergence (Black 1963, 1972; Hafner 1984, Goodwin 2008, 2009; Zelditch et al. 2015). Unfortunately, decades of previous studies have been limited by a number of potential methodological biases including incomplete taxon sampling, use of qualitative or strongly allometric morphological characters (i.e., in fossil taxa), analysis of single skull structures, and the lack of a robust multilocus phylogenetic hypothesis for Marmotini.

We generated a new phylogenetic hypothesis for Marmotini using the largest mitochondrial DNA dataset compiled to date, which we constrained with a recent backbone phylogenomic hypothesis of all marmotine genera and subgenera derived from >3,900 ultraconserved element loci (UCEs; McLean et al. in review). We compiled a novel 3D geometric morphometric dataset of cranial shape and combined this with existing datasets describing mandibular and molariform tooth shapes, and body size. In a phylogenetic context, we tested the association(s) of each skull trait with dietary ecology and body size, quantified patterns of evolutionary integration among skull traits, and explored trait-specific differences in evolutionary mode and macroevolutionary dynamics, including the frequency of convergent shifts. We document evolutionary heterogeneity among traits, as well as within traits but among ecotype classes; identify possible drivers of these patterns; and discuss our results in the context of current study design in macroevolutionary research.

## METHODS

### *Ecological Classifications*

We classified extant marmotine species *a priori* as “grazers” or “non-grazers” based on accounts of diet and natural history in published monographs (Howell 1938, Bryant 1945, Hafner 1984, Thorington et al. 2012), species accounts (www.mammalogy.org/publications/mammalian-species), and personal field observations. Diets of many ground squirrel species are broad, and vary with microhabitat, local plant phenology, and the shifting physiological demands of breeding, nursing, and hibernation. Many ground squirrels also opportunistically consume

invertebrates, other vertebrates (including conspecifics), and fungi. Our 2-category classification thus captures the most trenchant dietary differences that exist between these otherwise generalist ecological modes. We note that the non-grazing classifier will lump granivores and nut-eaters; however, seeds and nuts are not equivalent and may thus require very different masticatory apparatuses. Similarly, the grazing classifier was applied to taxa that regularly consume forbs in addition to grass, which would thus not be considered grazers in the narrow sense (Clauss et al. 2008). However, this classifier would still emphasize that grass (considered a tougher and more abrasive food source) comprises at least some component of the diet. Our approach differs from, and is more conservative than, some recent analyses that applied narrower dietary categorizations of marmotines (Casanovas-Vilar and van Dam 2013, Zelditch et al. 2017); we discuss implications of these differences on our results and conclusions throughout.

*Phylogenetic Inference*

We inferred phylogeny using the most comprehensive sequence dataset assembled from Marmotini to date. We constrained the backbone of marmotine phylogeny using a dataset consisting of 3,919 ultraconserved element (UCE) loci (total 2,365,391 bp) from 32 marmotine species, which included representatives of all 13 genera and 8 of 9 subgenera (McLean et al. in review). To achieve species-level sampling, we assembled the most comprehensive mitochondrial (mtDNA) dataset assembled for the tribe to date including 3 regions (cytochrome b [CYTB], control region [CR], cytochrome c oxidase subunit 3 [COIII]; 3120bp total). The mtDNA dataset includes previously published sequences (all from GenBank) and unpublished sequences from ongoing projects (Appendix 1). The full mtDNA dataset contains a total of 65 species (~72% of recognized marmotine species diversity).

Our study is primarily focused on understanding patterns of phenotypic evolution in Marmotini exclusive of *Tamias* (chipmunks, 25 spp. total), which we excluded from this study based on their small cranial size, extreme morphological conservatism, and incomplete taxon sampling across the skull trait datasets. However, our dataset does include 2 chipmunk species from different subgenera (*T. sibiricus* and *T. striatus*, subgenus *Eutamias* and *Tamias*, respectively), which recent workers have proposed to be elevated to generic status (Patterson and Norris 2016). Five other marmotine species were also omitted from our sampling due to lack of access to morphologically complete cranial specimens (*Marmota bobak, M. camtschatica, M. menzbieri, Spermophilus brevicauda, S. ralli*). In compiling the mtDNA and cranial shape datasets, we followed recent redescription of *Ammospermophilus insularis* (Mantooth et al. 2013), taxonomic elevation of *Otospermophilus douglasii* (Phuong et al. 2014), and synonymization of *Otospermophilus atricapillus* with *O. beecheyi* (Álvarez-Castañeda & Cortés-Calva 2011, Phuong et al. 2014). Within Holarctic ground squirrels (*Urocitellus*), we included 1 taxon that warrants specific status but which is currently recognized at the subspecific level (*U. mollis idahoensis;* McLean et al. 2016).

We accommodated uncertainty in higher-level phylogenetic relationships in Marmotini by retaining a distribution of UCE backbone topologies from McLean et al. (in review), which were obtained from different inference methods applied to different filtered datasets. We assembled this distribution by accessing results of RAxML

concatenated analysis of the largest UCE dataset, which assessed branch support with 100 bootstrap replicates, using scripts from the package ape (Paradis et al. 2016) in R v3.3.2 (R Core Team 2016) to read bootstrapped topologies, trim these to binary trees, and extract all unique backbone topologies (*N*=4; including the tree with the highest likelihood score (RAxML bestTree) and 3 deviations from that tree). We considered 2 additional backbone constraints obtained from coalescent-based analysis of UCEs in ASTRAL (Mirarab et al. 2014), which we trimmed and processed as above. One of these backbone trees was inferred from the same, largest UCE dataset as above and the other was inferred on a filtered dataset including only UCEs with ≥95% taxonomic completeness and ≥20% phylogenetically informative sites (see McLean et al. in review for details). We thus retained 6 backbone topologies capturing the majority of uncertainty in marmotine phylogeny at the generic and subgeneric levels (1 based on the RAxML best tree, 3 based on RAxML bootstraps, and 2 based on ASTRAL analyses).

We performed alignments for each mtDNA region individually in MUSCLE v3.7 (Edgar 2004) using default settings as implemented on the CIPRES science gateway (www.phylo.org; Miller et al. 2010) and concatenated alignments using the Python package *amas.py* (Borowiec 2016). We then inferred the optimal partitioning scheme for the mtDNA matrix in PartitionFinder 2 (Lanfear et al. 2016). We performed 6 separate species-level phylogenetic analyses from the mtDNA dataset using RAxML v8.2.9 (Stamatakis 2014) on CIPRES, each time using a distinct UCE-based topology as backbone constraint (-r option). All analyses otherwise used identical settings, including the single optimal partitioning scheme inferred in PartitionFinder. Support in each run was assessed using 1000 rapid bootstraps (-f a -x option) and branch lengths were optimized and printed for each replicate (-k option). We extracted the tree with the highest log likelihood from the analysis constrained by the preferred UCE backbone topology and considered this the best point estimate of marmotine phylogeny. We also extracted 20 unique bootstrapped topologies from each of the other 5 analyses (the best tree and 19 other trees). This procedure resulted in 100 unique trees that we carried through most subsequent morphological analyses to assess sensitivity of our inferences to topological variation. We quantified overall variation in the mtDNA topologies using Robinson-Foulds distances calculated in the R package phangorn v2.1.1 (Schliep et al. 2016).

*Morphological Data Collection and Compilation*

We collected 3D landmark data from 136 crania representing 65 marmotine species as described in *Phylogenetic Inference* (Appendix 1). This represents the first geometric morphometric dataset to describe any marmotine trait in 3D. For species subject to recent taxonomic changes, specimens were selected following geographic range descriptions of previous authors, similar to that described above for mtDNA dataset compilation. We sampled 2-3 adult specimens per species (64 of 65 species, or >98%). Our criterion for classifying specimens as adults was complete eruption and development of upper premolars 3 and 4 ($P^3$ and $P^4$, respectively). To avoid potential confounding contributions of sexual dimorphism to overall cranial shape variation, only females were used when possible (126 of 136 specimens, or 92%).

Fifty 3D landmarks (35 traditional and 15 semilandlarks; Supp. Fig 1) were collected on dorsal and ventral surfaces of crania using a Microscribe G2X digitizer and associated Microscribe Utility software (Solution Technologies, Inc.). Semilandmarks were collected along the midline of the cranium using the Scan Planes function in the Microscribe Utility software by first establishing a 3D plane corresponding to following axes: 1) a line connecting the anterolateral edges of nasal bones, 2) a line connecting the tip of the nasals to the posteriormost projection of the occipital (=lambdoidal) crest, and 3) a dorsoventral line orthogonal to lines 1 and 2. All landmarks were collected 3 times to ensure precision, these replicates were averaged, and dorsal and ventral aspects were stitched together based on 4 common landmarks using the software DVLR (http://pages.nycep.org/nmg/programs.html). Cranial data were subjected to Procrustes superimposition in the R package geomorph v3.0.3 (Adams et al. 2016) with semilandmarks slid to minimize bending energy.

We used the mandibular dataset of Zelditch et al. (2015; accessed at http://datadryad.org/) consisting of 98 2D landmarks (14 landmarks and 84 semilandmarks) collected from 52 marmotine species (total after incorporating recent taxonomic rearrangements and excluding species unsampled for cranial shape). We also used the dataset of Goodwin (2009) consisting of 16 linear measurements collected on upper and lower molariform teeth of 58 marmotine species after accounting for the taxonomy adopted here (2 original measurements on upper premolar 3 ($P^3$) were excluded as this tooth is absent in 2 marmotine species). Forty-nine species were shared across all 3 skull trait datasets, with higher numbers shared in most pairwise trait comparisons (52 for crania/mandibles, 58 for crania/teeth, 49 for mandibles/teeth). Finally, we used Hayssen's (2008) dataset of sciurid head-body length as a metric for body size, which we updated for the few taxonomic changes mentioned above by compiling head-body lengths from museum specimens accessed in VertNet (http://vertnet.org/; $N \geq 10$ adult specimens per species). The logarithm of head-body length was used in PGLS analyses but untransformed lengths were used in all other analyses. All trait datasets were standardized to a common coding scheme and species means were computed in R and used for all further analyses.

*Trait Variation and Correlation with Phylogeny, Body Size, and Ecology*

We performed principal components analysis (PCA) in geomorph to identify major axes of variation and patterns of morphospace occupation in different shape traits as well as to provide reduced-dimensionality variables for further analyses. We assessed the level of trait covariation with marmotine phylogeny by computing phylogenetic signal for each trait as the multivariate extension of Blomberg's $K$ ($K_{mult}$; Adams 2014). $K_{mult}$ quantifies phylogenetic covariance relative to expectations under Brownian motion (BM) and is thus a useful metric for identifying deviations from that null model, as might arise from conservatism or convergence (Blomberg et al. 2003). Significance of $K_{mult}$ was assessed via 999 permutations of Procrustes data among tips of the preferred phylogenetic tree, and this was repeated for all additional phylogenetic hypotheses. We assessed the extent to which body size and dietary ecology contribute to shape traits using multivariate phylogenetic generalized least squares (PGLS; Grafen 1989) implemented in the procD.pgls function in geomorph. Body size, dietary ecology, and their interaction

term were used as model factors. Significance was assessed from 999 permutations on the preferred phylogeny and the test was repeated for all 100 additional topologies.

*Trait Covariation*

We explored evolutionary covariation among skull traits in Marmotini by testing for integration in all pairwise trait comparisons. To confirm we were using appropriate units for analyses, we first tested whether each of the 3 skull traits can be considered a single, self-similar evolutionary module or, alternatively, if modular structure exists within structures. For crania, we tested a simple 2-module hypothesis with distinct partitions corresponding to the rostrum and braincase, as well as a 3-module hypothesis that further recognizes developmental discreteness of the neurocranium and basicranium within the braincase (Hallgrimsson et al. 2007). For mandibles, we tested a previously supported 2-module hypothesis with partitions corresponding to the alveolar region and ascending ramus, respectively (Cheverud et al. 1997, Klingenberg et al. 2003). Finally, several testable hypotheses of modularity exist within and among partitions of mammalian molariform teeth, none of which have been extensively investigated at genetic, developmental, and evolutionary levels and across the taxonomic diversity of mammals (e.g., see Gomez-Robeles and Polly 2012). Our data only contain molariform teeth, but we tested a hypothesis of upper (maxillary) and lower (mandibular) molariform tooth modularity. All tests were based on the covariance ratio (CR; Adams 2016), which was calculated using the phylo.modularity function in geomorph.

Levels of integration among evolutionary modules in the skull were assessed using partial least squares (PLS) as implemented in the phylo.integration function in geomorph. That function estimates covariation between partitions of shape variables while accounting for phylogenetic nonindependence and assuming a BM model of evolution (Adams and Felice 2014). To separate the integrating effects of allometry from those of other forces, such as mechanical function, we also assessed integration in the non-allometric component(s) of shape variation. Size-free shape variables for this analysis were obtained from a multivariate regression of Procrustes data on the logarithm of head-body length. We performed pairwise PLS for all module combinations using both raw and size-free data and assessed significance at the $p = 0.05$ level following 999 permutations of the original data matrices. Finally, we compared the strength of integration (PLS correlations) across these trait combinations using a method that is insensitive to differences in sample size and number of shape variables (Adams and Collyer 2016).

*Trait Evolutionary Modes*

We inferred trait evolutionary mode(s) by fitting common macroevolutionary models to all skull traits and body size. For skull traits, models were fit to PC axes that cumulatively explained ~80% or more of the total shape variation (cranial PC 1-5, mandibular PC 1-4, and molariform tooth PC 1-3). We considered 3 general classes of evolutionary models: those describing trait evolution as a random walk (i.e., BM), as a directional random walk towards selective optima (i.e., Ornstein-Uhlenbeck, or OU), and as subject to decreasing rates through time (e.g., as may occur as niches become filled;

Early Burst, or EB). We fit models that assumed homogeneity of parameters across all of Marmotini (BM1, EB) as well as, in the case of the OU model, existence of a single selective optimum (OU1). We also fit a more complex BM model allowing evolutionary parameters to vary between the 2 dietary groups (BM2), as well as an OU model with 2 selective regimes corresponding to dietary groups (OU2). For BM2 and OU2 models, we painted selective regimes (i.e., dietary group membership) on phylogenies using the make.simmap function in phytools v0.5-38 (Revell 2016) in R. All model fits were conducted in a multivariate framework in the package mvMORPH v1.0.8 (Clavel et al. 2015) using PC-specific parameters, and model fit was assessed using Akaike's Information Criterion corrected for small sample sizes (AICc) and the ΔAICc metric.

*Trait-specific Tests of Convergence*

Values of $K_{mult}$ <1 indicate that species resemble each other more than expected under BM, a pattern that may arise via several processes, including convergence (Blomberg et al. 2003, Revell et al. 2008, Adams 2014). However, convergence is difficult to distinguish from conservatism because both processes can result in similar patterns of phenotypic variation. To begin to assess the consistency of our data with convergence towards grazing and non-grazing ecomorphologies, we compared within-group disparities to expectations under BM. This approach is similar to that used in previous studies (Mahler et al. 2013, Ingram and Kai 2014) with the exception that the test statistic is the mean of squared among-species distances within dietary groups (i.e., disparity, or cohesion in morphospace) rather than the similarities between individual species pairs presumed to be convergent within those groups, which is more appropriate for our dataset since the majority of morphological and ecological diversity occurs above the species level in Marmotini. Using the same sets of PC axes described above, we calculated empirical within-group disparities for each trait using the function disparity in geiger v2.0.6 (Harmon et al. 2008). We then used the optimal BM rate parameters estimated above (*Trait Evolutionary Modes*) to simulate 999 trait datasets on the preferred topology, calculating within-group disparities for each. We tested the hypothesis that empirical within-group disparities were lower than expected under BM using a one-tailed t-test and assessed significance at the p=0.05 level, and repeated this for each of the 100 unique topologies. To isolate allometric and ecological contributions to disparity patterns, we repeated these tests for size-corrected skull trait datasets as well.

The above approach compares disparity of dietary groups to that expected under a BM process, but does not fully distinguish between convergence and constraint as causes of disparity patterns. We conducted more explicit tests of convergence using one method of Stayton (2015). Briefly, the method ('convnum') assesses the frequency with which lineages on a phylogenetic tree have entered a region of morphospace presumed to harbor convergent taxa. We used prespecified dietary groups to delineate 'grazing' and 'non-grazing' regions of morphospace for each skull trait based on the empirical PC datasets used above, then calculated the number of convergent shifts occurring in each dietary group using the convnum function in the R package convevol v1.0, using default settings. We compared observed numbers of convergent shifts to expectations under a single-state OU model (OU1; representing a situation where phenotypic diversity is constrained) and a single-state BM model (BM1; a random walk which may induce convergent events by

chance, and thus provides a more conservative test of convergence). We simulated 999 datasets under each model for each skull trait using the optimal parameters inferred above. For each simulation, we tested whether the number of observed convergent shifts was greater than expected using a one-tailed t-test. The procedure was repeated across the 100 topologies to assess sensitivity of results. Finally, to distinguish if patterns and intensity of convergence were due to allometry, adaptive shape change, or both, we repeated the above test for size-corrected skull trait datasets for the single preferred topology.

## RESULTS

### *Phylogeny of Marmotini*

Phylogenetic analyses in RAxML were straightforward, with higher-level phylogenetic uncertainties propagated through shallower-level mtDNA tree searches. There was significant uncertainty among species-level mtDNA estimates; all 100 bootstrapped topologies that we sampled randomly from RAxML analyses were unique. The median pairwise Robinson-Foulds (symmetric) distances for trees in this distribution was 26 (range 2-48). This level of uncertainty is consistent with previous mtDNA analyses (Harrison et al. 2003) and suggests that sensitivity analyses are a critical component of any comparative analysis of this group.

### *Cranial Shape PCA*

In all cases, our ecological classifications were identical among congeners, consistent with the ecological distinctiveness that is most pronounced at the genus level in Marmotini (Bryant 1945, Hafner 1984, Helgen et al. 2009, Zelditch et al. 2017). PCA revealed significant dimensionality in cranial shape, with cranial PC axes 1-5 accounting for 79.5% of total variation (Supp. Fig 2). The relatively high dimensionality in this structure is similar to results of Caumul and Polly (2005) for marmots. Two broad similarities emerged from the cranial shape PCA. First, marmots (*Marmota*) are clearly differentiated from all other genera along PC1, consistent with the unique size and cranial morphology of this genus. However, although cranial PC1 appears to be largely associated with body size, a test of evolutionary allometry across Marmotini is not statistically significant (see below); thus, PC1 must capture unique cranial shape differences in marmots that are not captured by the evolutionary allometric relationships in the remainder of marmotine taxa. Second, subsequent axes (PC 2-3) of cranial morphospace tend to separate marmotine genera by dietary groups rather than by phylogenetic relationships, suggesting substantial ecological signal in this trait (e.g., Fig. 1, Supp. Fig. 2).

### *Trait-specific Patterns of Variation and Covariation*

Estimates of phylogenetic signal ($K_{mult}$, Fig. 2) on the preferred topology varied widely among crania (0.22), mandibles (0.51), molariform teeth (0.92), and body size (0.65). These estimates are notable because they span much of the range reported for

morphological traits in animals (e.g., Blomberg et al. 2003) and are also more variable than reported for suites of skull traits in many mammal clades (Renaud et al. 2007, Cardini and Elton 2008, Alvarez et al. 2015, Maestri et al. 2017), although these papers have used different metrics to quantify trait correlation with phylogeny. Phylogenetic signal in mandible shape is similar to, but slightly lower than, that calculated by Zelditch et al. (2015) across all Sciuridae, likely reflecting different evolutionary dynamics operating at these 2 scales and use of different phylogenetic hypotheses. Phylogenetic signal in molariform tooth shape is qualitatively consistent with results of Goodwin (2009), who worked at the same phylogenetic scale as here and recovered a linear relationship between odontometric distances and time since divergence. Interestingly, however, patterns of phylogenetic signal in skull traits are largely opposite of those found by Caumul and Polly (2005) for Eurasian *Marmota*. All estimates of $K_{mult}$ were sensitive to topological variation, with trait-specific ranges of 0.11-0.34 (crania), 0.37-0.60 (mandibles), 0.68-1.12 (teeth), and 0.39-0.85 (body size). Nevertheless, ranges for the traits were statistically significant from one another ($P < 0.01$ for all comparisons; Fig. 2), supporting distinct strengths of phylogenetic covariance.

Deviations from Brownian expectations can occur via a variety of processes. PGLS confirmed that variation in skull traits is at least partly related to dietary adaptation, as shape was significantly related to dietary ecology in each trait (Table 1). Mandibles and molariform teeth also displayed significant allometry, while crania did not, although diet had a larger effect size than size in every trait (Table 1). A notable idiosyncrasy was that mandible shape was the only trait for which the interaction term (diet x size) was significant, a result that is in agreement with results of Zelditch et al. (2017) for mandible shape across Sciuridae. This indicates that mandible evolution is more constrained than that of crania and teeth, with dietary-related adaptations being effectively constrained to size-related axes. This is opposite from the pattern seen in molariform teeth, where these axes are orthogonal. Results were largely robust to topological uncertainty, although the statistical significance of an allometric relationship did change in ~5% of trees from that reported in Table 1 for both crania and mandibles.

Contrary to some previous analyses in other mammalian taxa, we found no support for evolutionary modularity within crania, mandibles, or molariform teeth corresponding to the hypotheses considered above. Covariance ratios (CR) for cranial 2-module and 3-module hypotheses were 1.09 and 1.04 ($P > 0.95$ for both), suggesting within-module variation is no different than that among modules (e.g., the case where CR = 1). CR for the mandible 2-module hypothesis was higher than for crania (1.33, $P = 1$), suggesting even greater integration across this structure. CR calculated for upper and lower molariform toothrows was likewise nonsignificant (1.07, $P = 0.55$), and a *post hoc* test revealed that toothrows are extremely highly integrated (r-PLS = 0.99, $P < 0.01$). Thus, skull traits can be considered self-similar for the purposes of our analyses.

Significant levels of evolutionary integration were recovered between crania and mandibles (r-PLS = 0.75), crania and molariform teeth (r-PLS = 0.84), and mandibles and molariform teeth (r-PLS = 0.61; $P < 0.01$ for all, Table 2). Integration was strongest between crania and molariform teeth ($Z = 5.53$), exceeding that of crania and mandibles ($Z = 3.29$) and mandibles and molariform teeth ($Z = 3.15$), although the latter test was only marginally significant ($P = 0.04$ and 0.07, respectively). Significant integration was also detected in all comparisons of size-corrected shape traits (crania-mandibles, r-PLS =

0.76; crania-teeth, r-PLS = 0.74; mandibles, teeth, r-PLS = 0.57; $P < 0.01$ for all), but there were no significant differences in the strength of integration across comparisons for size-corrected traits (Table 2). Thus, evolutionary integration was found among skull traits and is due to allometry and non-allometric factors, but different strengths of integration are only found when the allometric component is included.

*Trait Evolutionary Dynamics*

Cranial and mandibular shape evolution were best described by an OU process with selective regimes corresponding to prespecified dietary groups (OU2 model; Table 3), while evolution of molariform tooth shape and body size were most consistent with a 2-state BM model (BM2). These results are consistent with estimates of $K_{mult}$, which were closer to 1 in molariform teeth and body size but lower for crania and mandibles. The most complex models (defined solely in number of parameters) were favored for crania and mandibles, but this was not the case for body size and tooth shape (Table 3). Early Burst models received among the lowest support across all traits, consistent with the apparent rarity of this mode in many phylogenetic comparative datasets of animals (Harmon et al. 2010, Slater 2015a,b).

Comparisons to simulations permitted the differences among taxa and traits to be more finely parsed. In both cranial and mandibular shape, non-grazers showed significantly lower disparity than expected under a BM process (Fig. 3), while disparity in these traits in grazers was consistent with BM. When the highly distinctive genus *Marmota* was excluded from analysis, however, among-grazer disparity was lower than BM expectations, as was observed in non-grazers. Patterns in molariform teeth were more divergent. While non-grazers again showed significantly less disparity than expected, grazer molariform tooth shape was either higher than BM expectations (all genera) or consistent with BM (exclusion of *Marmota*; Fig. 3). Thus, non-grazers showed significantly less disparity than expected under BM for all skull traits, while grazers only showed this pattern for crania and mandibles when *Marmota* was excluded. Repeating this analysis on size-corrected traits resulted in significantly lower disparity than expected for all groups in crania and mandibles, regardless of inclusion of *Marmota*, indicating that the allometric component of those traits causes increased shape disparity. For size-corrected tooth shape, only non-grazers showed lower disparity than expected (Supp. Fig. 3). Together, these results support strong ecomorphological cohesiveness within *a priori* dietary groups, a result that we would not be expected if our dietary classifications were erroneously broad (e.g., if unrecognized levels of ecomorphological variation were lumped together).

Frequency-based tests of convergence indicated that the above disparity patterns are due to processes of both convergence and constraint. Numbers of convergent shifts observed in crania and molariform teeth were significantly higher than expected under an OU process in non-grazing taxa (N=5 and 7 shifts, respectively, $P < 0.02$ for both, Fig. 4), while numbers of shifts in these traits in grazers ranged from 1-3 but were not higher than expected, a result that was insensitive to inclusion of *Marmota* (Fig. 4). Notably, we found no evidence of convergence in mandible shape for any dietary grouping; interestingly, this result appears consistent with a broader analysis of mandible shape across Sciuridae (Zelditch et al. 2017). Our results suggest that reduced within-group

disparity in mandible shape is therefore due not to convergence, but to evolutionary constraint. All of the above results held under the more conservative comparison with Brownian motion expectations, where increased numbers of convergent shifts are expected to occur by random entrance of convergent zones in morphospace (Supp. Fig 4). Qualitatively identical results were also obtained when the same analyses were performed for size-corrected shape traits, which indicates that patterns of convergence and constraint are not driven by allometric scaling but by dietary adaptation, although the actual numbers of shifts varied and were sometimes lower for size-corrected traits (Supp. Fig 5-6). Finally, these tests were largely (but not completely) robust to our estimates of topological uncertainty; the largest discrepancy was found in the number of statistically significant convergent shifts in mandible shape of non-grazers (Fig. 4, Supp. Fig. 4). *Post hoc* inspection revealed these discrepancies to be concentrated within (but not limited to) topologies with divergent placement of *Otospermophilus* +*Callospermophilus* (not sister to *Marmota*) and *Poliocitellus* (basal with respect to big-eared *Urocitellus*, rendering the latter genus paraphyletic), which we cannot rule out but are inconsistent with most current information on marmotine phylogeny.

## DISCUSSION

### *Patterns of Phenotypic Diversification in Marmotini*

Significant focus has been placed over the past 2 decades on understanding phenotypic dynamics in evolutionary radiations such as Marmotini. As perhaps expected, substantial variation in these dynamics has been uncovered, even among vertebrate clades (e.g., Adams et al. 2009, Harmon et al. 2010, Derryberry et al. 2011, Mahler et al. 2013, Slater 2015a, Zelditch et al. 2015, Maestri et al. 2017, Cantalapiedra et al. 2017), thus complicating the search for general processes that underlie and potentially fuel instances of exceptional lineage diversification. However, the comparative phylogenetic approach, which is often used for studies of extant clades, is known to be sensitive to several confounding factors (e.g., McLean 2017). Our results demonstrate the magnitude of effect that 1 factor (trait choice) on macroevolutionary inferences. Specifically, we found evolutionary heterogeneity among traits (evolutionary mode and intensity of convergence) that otherwise display strong evolutionary integration due to functional linkage (i.e., the procurement and processing of plant foods). It is important to note that our skull datasets vary both in detail (linear measurements vs. landmark configurations) and dimensionality (2D vs. 3D), properties that by themselves could introduce heterogeneity into patterns of trait variation. Nevertheless, our results do not appear dependent on the richness with which traits are described (e.g., cranial shape, which was captured in highest resolution, displayed greater evolutionary flexibility than mandibles). Such heterogeneity is consistent with classic intuitions of, and a growing amount of quantitative evidence for, the mosaic nature of phenotypic evolution.

At least some of the heterogeneity we document among skull traits can be attributed to differences in the strength of allometry. In a morphological context, allometry is defined as that variation in shapes that is correlated with variation in body size, e.g., to maintain proportions and optimize biological function (Gould 1966; Klingenberg 1996, 2016). Molariform tooth shape is strongly allometric in Marmotini

(Table 1; Goodwin 2009) and, consequently, best described by the same evolutionary process that describes body size (2-state BM model, Table 3). Although tooth morphology is a common metric of dietary adaptation in many mammals, sampling body size and/or tooth shape alone would lead to different conclusions about mechanisms of phenotypic evolution in Marmotini than if only cranial and/or mandibular shape were sampled. An emerging consensus is that body size has been exceptionally labile over macroevolutionary timescales (Smith et al. 2010, Evans et al. 2012, Slater 2015a), with size-related traits thus experiencing different evolutionary dynamics than shape-related traits (e.g., Hunt 2007). Moreover, size-related traits may not be causal agents in the ecomorphological differentiation that defines most adaptive radiations. Particular care is therefore necessary when drawing inferences of evolutionary dynamics in radiations solely from body size or other size-related trait data.

On the other hand, even traits evolving via a similar evolutionary mode can display divergent dynamics on the macroevolutionary landscape. An OU process with selective peaks corresponding to grazing and non-grazing ecotypes best described patterns of cranial and mandibular shape evolution. However, when considering the frequency of convergent shifts, only crania were exceptional compared to expectations under common constant rate models. This striking difference in the frequency of convergence between 2 integrated structures appears due to greater developmental or functional constraints on mandible evolution. Mandibles were unique in our study in that allometric and ecological components of shape variation were non-orthogonal (Table 1), meaning adaptive shape change is at least partly constrained to size-related dimensions in this structure. Conversely, taxa of varying body sizes have converged on similar regions of cranial ecological shape space (e.g., cranial PC2 in our analysis; Supp. Fig 2). Greater evolutionary flexibility in crania is consistent with both the lower phylogenetic signal (Fig. 2) and higher clade-wide disparity (results not shown) in this structure relative to mandibles. An alternative explanation for the latter pattern is that crania integrate a higher diversity of biological functions than mandibles, and that more varied selective pressures lead to greater overall levels of cranial shape disparity. Regardless, our results are consistent with different levels of evolutionary constraint in cranial and mandibular shape, reiterating the key distinction that must be made in evolutionary studies between character- and whole-organism constraint (e.g., Futuyma 2010). If our results are representative of broader, trait-specific phenomena, they have important implications for recent studies that have drawn inferences of macroevolutionary process from mandibles alone (Zelditch et al. 2017).

When placed in context of recent literature, our results speak to another potential confounding aspect of the comparative approach as well; that is, the scale of analysis. Specifically, scale-dependent differences often exist in speciation and extinction rates, developmental or functional constraints on trait evolution, and strength of the phenotype-environment correlation, and these may negatively impact macroevolutionary inferences (e.g., McLean 2017). Caumul and Polly (2005) analyzed trait datasets that are qualitatively similar to ours, but at a much shallower phylogenetic scale (5 species of Eurasian *Marmota*). As in our study, they documented heterogeneity in contributions of phylogeny, allometry, and ecology to skull traits; however, while some patterns are shared across these 2 analyses (e.g., depressed phylogenetic signal in mandibles), others are opposed (e.g., they found low association of mandibles with size and diet; they also

found phylogenetic signal to be high in crania but low in molariform tooth shape). Scale-dependency in macroevolutionary analyses has not been adequately addressed to date but, like the choice of trait(s) sampled, it is a crucial confounding factor to consider (Hopkins and Smith 2015).

*Ecotype Dynamics*

Squirrels in general display relatively limited phenotypic complexity, with early workers delineating only a few ecotypes (Howell 1938, Bryant 1945, Black 1963, 1972; Thorington 1997). Black (1972), inspired by patterns of skull and dental evolution in the fossil record, argued that just 2 main adaptive types exist within Sciuridae – tree and ground squirrels – but that additional subtypes exist within each type. Our trait data support at least 2 broadly defined adaptive subtypes within marmotine ground squirrels (grazers and non-grazers; but see discussion of *Marmota* below), with convergent shifts towards each subtype across the history of the clade detectable in at least some skull traits. The repeated origins of grazing and non-grazing subtypes across this time may be mechanistically similar (although more temporally limited) to the "subzonal radiations" described in North American fossil canids by Slater (2015a). On the other hand, such limited ecophenotypic diversity in Marmotini is perhaps surprising given the multiple intercontinental dispersals that have occurred to Eurasia, which may have opened new ecological opportunities to this tribe.

Several of our findings were sensitive to classification of *Marmota* as a grazer (Figs. 3-4). Marmots were once considered the sister taxon to all other marmotines (Bryant 1945, Black 1963) due to their extreme morphological, ecological, and behavioral distinctiveness; only relatively recently has molecular data confirmed their nestedness within Marmotini (Thomas and Martin 1993, Giboulet et al. 1997, Harrison et al. 2003). *Marmota* is thus an example of rapid evolution towards a novel ecological zone (Goodwin 2009, Polly et al. 2015). While our broad dietary classifications (grazer, non-grazer) capture much of the limited ecological disparity in Marmotini, sensitivity of results to inclusion of *Marmota* suggests these classifications could be overly simplistic. *Marmota* occupies unique regions of morphospace in all traits considered here, and their inclusion as grazers distorts the total volume of morphospace occupied by that guild, impacting downstream inferences of the extent and frequency of convergence. Conversely, analysis of size-corrected data appears more insensitive to inclusion of *Marmota* as a grazer, suggesting extreme uniqueness in cranial and molariform-tooth shape is largely or completely due to allometric effects. Sampling of additional classes of morphological traits may help to identify whether marmots indeed inhabit a unique adaptive subzone that could be poorly delineated by dietary classifications alone.

There is a polarity in marmotine ecomorphological evolution similar to that of other rodent radiations. Basal marmotines are small- to medium-sized, dietary generalists, and predominately scansorial, which is likely the ancestral condition for this clade (Black 1972, Goodwin 2008). A hypothesis of generalist ancestors giving rise to ecologically specialist taxa is also accepted for other major rodent clades, including Old World murids (Renaud et al. 2007, Rowe et al. 2016), Neotropical sigmodontines (Maestri et al. 2017), and hystricognaths (Hautier et al. 2012). Nevertheless, available ecological and ecomorphological data suggest generalist rodent phenotypes are extremely evolutionarily

successful, and generalists tend to account for the majority of species diversity in these clades (Renaud et al. 2007, Hautier et al. 2012, Maestri et al. 2017). Similarly, the majority of marmotine species are non-grazers (including *Tamias*), and exceptional levels of convergence are only observed towards non-grazing morphologies, not vice-versa (at least in crania and molariform teeth). A lack of similarly exceptional numbers of convergent grazing lineages could be due to the narrower niche requirements of this ecotype and/or current constraints on available niche space in the Holarctic (grazing genera are geographically widespread in North America), both of which would limit evolution of new grazing lineages.

*Integration and Modularity in Rodent Evolution*

Integration is believed to play an important role in shaping macroevolutionary patterns of phenotypic diversity by constraining the directionality of trait evolution. While integration limits the total range of phenotypic possibilities, it also facilitates movement between selectively advantageous phenotypes, resulting in convergence at macroevolutionary scales (Armbruster et al. 2014, Goswami et al. 2014, Conner et al. 2014). There is significant integration in the marmotine skull due to allometry as well as non-allometric factors (Table 2), suggesting that integration may have played a role in convergence of skull traits within dietary groups. Strong integration among skull traits has been demonstrated in many other vertebrate groups (Cheverud 1982, Marroig and Cheverud 2001, Goswami 2006, Conner et al. 2014), including some rodent radiations (tree squirrels, Roth 1996; hystricognath rodents, Hautier et al. 2012, Alvarez et al, 2015), although not all studies have attempted to isolate allometric and non-allometric integrating factors.

Functional demands of biting or mastication are possible non-allometric integrating factors in marmotine skull traits; however, additional forces can favor integration among skull structures as well. For example, Goswami (2006) found that integration in mammalian carnivores was not always related to diet. Tsuboi et al. (2014) showed that head shape in cichlid fish is related to feeding ecology and brain size (but not body size), which in turn suggests that selection for both ecological as well as cognitive functionality can interact to favor integration. Marmotine species vary in degree of sociality from solitary to highly colonial (Blumstein and Armitage 1998), and it is possible that cognitive demands impact cranium size and/or shape evolution, although we recovered no evidence of that in our limited study. Alternatively, or perhaps in addition, axes of the behavioral niche such as methods of burrow construction or postures involved in predator vigilance could exert selective pressures on skull traits and favor integration among individual components.

Because organisms must function as a whole, evolutionary integration should be detectable at the whole-organism level in some traits (e.g., Cheverud 1982; but see Pigliucci 2003). Skull trait variation within many clades may thus be phylogenetically coincident with that of skull musculature, sense organs, or hard and soft postcranial traits. In Marmotini specifically, grazing and non-grazing ecotypes are at least partially indicative of locomotory mode, which varies from highly scansorial (e.g., *Tamias* and other generalized ground squirrel genera) to burrowing and highly terrestrial (e.g., *Marmota*, *Cynomys*). Skull traits could thus be integrated with locomotory traits such as

limb or digit shape. Conversely, integration among cranial and postcranial traits appears low in some other rodent clades such as Dipodoidea (birch mice, jumping mice, jerboas). Wu et al. (2014) showed that dental and locomotory traits are decoupled over dipodid evolution, and Moore et al. (2015) demonstrated an additional lack of integration among postcranial traits related to locomotion in this superfamily. Such partitioning of phenotypic covariation into increasingly independent blocks of traits broadens the range of possible morphological variation, potentially enabling access to new ecological opportunities. On the other hand, modularity may be an exception rather than rule in other rodents, where generalist morphologies are both successful and phylogenetically widespread. Better understanding of the relative importance of integration versus modularity in rodent evolution remains an important research avenue.

*Accounting for Trait Heterogeneity in Macroevolution*

A major breakthrough in the life sciences has been the widespread comprehension that different portions of the genome integrate distinct evolutionary histories due to locus-specific processes of selection, horizontal transfer, and/or coalescent stochasticity. In a similar way, evolutionary heterogeneity should be an *a priori* expectation among phenotypic traits, as these are ultimately functions of the genes (or, more often, groups of interacting genes) that encode them. Although trait heterogeneity is not a new idea (e.g., Gould and Lewontin 1979, Cheetham 1987, Pigliucci 2003), the design of many modern comparative studies, where 1 or a few traits alone contribute to inferences of macroevolutionary processes, suggests the idea is an important one to reinforce. Just as the difference between gene histories and species histories is now clearer (Maddison 1997), so too must a better distinction be drawn when considering the histories of phenotypic traits and those of whole organisms.

The potential for trait heterogeneity to be widespread both within phenomes and across the Tree of Life begs the question: how many traits must be sampled for robust macroevolutionary inference? That question is difficult to answer at present due to the insufficient numbers of traits that have been captured within lineages at high enough resolution to quantify the extent of differences among them. Hopkins and Lidgard (2012) have demonstrated that the likelihood of discordant evolutionary mode(s) is a function of the number of traits analyzed. Many (but not all) traits analyzed in that study are linear measurements or ratios, which often fail to capture the complex variation present in many morphological structures. Higher-resolution shape data are preferable, but it is not generally known whether higher resolution might lead to more or less concordance among trait evolutionary modes. Better insights into the magnitude and extent of trait heterogeneity may be increasingly possible via combination of landmark- and eigenshape-based morphometric approaches with ongoing efforts at high-resolution digitization of museum and herbarium specimens worldwide.

ACKNOWLEDGEMENTS

Portions of this work were funded by the National Science Foundation (NSF 1258010), Smithsonian Institution (Peter Buck Predoctoral Fellowship to B.S.M.), American Museum of Natural History (Theodore Roosevelt Memorial Research Grant to B.S.M.), T&E, Inc. of New Mexico, American Society of Mammalogists (Grant-in-Aid of Research to B.S.M.), and the University of New Mexico Department of Biology and Graduate and Professional Student Association. We thank many museum curators and collection managers for access to important specimens, especially Darrin Lunde, Suzy Peurach, Esther Langan, and Nicole Edmison (NMNH). Matt Carrano (NMNH) graciously loaned the Microscribe used in cranial shape data collection. We thank Gene Hunt (NMNH) for valuable comments on an earlier draft of this manuscript.

REFERENCES

Adams, D.C., M. Collyer, E. Sherratt. 2016. geomorph v3.0.3. https://cran.r-project.org/web/packages/geomorph/index.html.

Adams, D.C. 2014. A generalized *K* statistic for estimating phylogenetic signal from shape and other high-dimensional multivariate data. Systematic Biology 63:685–697.

Adams, D.C. 2016. Evaluating modularity in morphometric data: challenges with the RV coefficient and a new test measure. Methods in Ecology and Evolution 7:565–572.

Adams, D.C., C.M. Berns, K.H. Kozak, J.J. Wiens. 2009. Are rates of species diversification correlated with rates of morphological evolution? Proceedings of the Royal Society B 276:2729–2738.

Adams, D.C., M.L. Collyer. 2016. On the comparison of the strength of morphological integration across morphometric datasets. Evolution 70:2623–2631.

Adams, D.C., R.N. Felice. 2014. Assessing trait covariation and morphological integration on phylogenies using evolutionary covariance matrices. PLoS One 9:e94335.

Alvarez, A., S.I. Perez, D.H. Verzi. 2015. The role of evolutionary integration in the morphological evolution of the skull of caviomorph rodents (Rodentia: Hystricomorpha). Evolutionary Biology 42:312–327.

Álvarez-Castañeda, S.T., P. Cortés-Calva. 2011. Genetic evaluation of the Baja California rock squirrel. Zootaxa 51:35–51.

Armbruster, W.S., C. Pelabon, G.H. Bolstad, T.F. Hansen. 2014. Integrated phenotypes: understanding trait covariation in plants and animals. Philosophical Transactions of the Royal Society B 369:20130245.

Black, C.C. 1963. A review of the North American Tertiary Sciuridae. Bull. Museum Comp. Zool. Harvard Univ. 130:109–248. Cambridge.

Black, C.C. 1972. Holarctic Evolution and Dispersal of Squirrels (Rodentia: Sciuridae). Pp. 305–322 *in* T. Dobzhansky, M. K. Hecht, and W. C. Steere, eds. Evolutionary Biology. Appleton-Century-Crofts, Meredith Corporation, New York, NY.

Blomberg, S.P., T. Garland, A.R. Ives. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57:717–745.

Blumstein, D.T., K.B. Armitage. 1998. Life history consequences of social complexity: a comparative study of ground-dwelling sciurids. Behavioral Ecology 9:8–19.

Borowiec, M.L. 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. PeerJ 4:e1660.

Bryant, M.D. 1945. Phylogeny of Nearctic Sciuridae. American Midland Naturalist 33:257–390.

Cantalapiedra, J.L., J.L. Prado, M. Hernández Fernández, M.T. Alberdi. 2017. Decoupled ecomorphological evolution and diversification in Neogene-Quaternary horses. Science 355:627–630.

Cardini, A., S. Elton. 2008. Does the skull carry a phylogenetic signal? Evolution and modularity in the guenons. Biological Journal of the Linnean Society 93:813–834.

Casanovas-Vilar, I., J. van Dam. 2013. Conservatism and adaptability during squirrel radiation: what is mandible shape telling us? PLoS One 8:e61298.

Caumul, R., P.D. Polly. 2005. Phylogenetic and environmental components of morphological variation: skull, mandible, and molar shape in marmots (*Marmota*, Rodentia). Evolution 59:2460–2472.

Cheetham, A.H. 1987. Tempo of evolution in a Neogene Bryozoan: are trends in single morphologic characters misleading? Paleobiology 13:286–296.

Cheverud, J.M. 1982. Phenotypic, genetic, and environmental morphological integration in the cranium. Evolution 36:499–516.

Cheverud, J.M., E.J. Routman, D.J. Irschick. 1997. Pleiotropic effects of individual gene loci on mandibular morphology. Evolution 51:2006–2016.

Clavel, J., G. Escarguel, G. Merceron. 2015. mvMORPH: an R package for fitting multivariate evolutionary models to morphometric data. Methods Ecol. Evol. 6:1311–1319.

Clauss, M., T. Kaiser, J. Hummel. 2008. The Morphophysiological Adaptations of Browsing and Grazing Mammals. Pp. 47–88 *in* I. J. Gordon and H. H. T. Prins, eds. The Ecology of Browsing and Grazing. Springer-Verlag, Heidelberg.

Conner, J.K., I.A. Cooper, R.J. La Rosa, S.G. Pérez, A.M. Royer. 2014. Patterns of phenotypic correlations among morphological traits across plants and animals. Philosophical Transactions of the Royal Society B 369:20130246.

Derryberry, E.P., S. Claramunt, G. Derryberry, R.T. Chesser, J. Cracraft, A. Aleixo, J. Pérez-Emán, J.V. Remsen, R.T. Brumfield. 2011. Lineage diversification and

morphological evolution in a large-scale continental radiation: the neotropical ovenbirds and woodcreepers (Aves: Furnariidae). Evolution 65:2973–2986.

Edgar, R.C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797.

Evans, A.R., D. Jones, A.G. Boyer, J.H. Brown, D.P. Costa, S.K.M. Ernest, E.M.G. Fitzgerald, M. Fortelius, J.L. Gittleman, M.J. Hamilton, L.E. Harding, K. Lintulaakso, S.K. Lyons, J.G. Okie, J.J. Saarinen, R. M. Sibly, F.A. Smith, P.R. Stephens, J.M. Theodor, M.D. Uhen. 2012. The maximum rate of mammal evolution. Proceedings of the National Academy of Sciences 109:4187–90.

Futuyma, D.J. 2010. Evolutionary constraint and ecological consequences. Evolution 64:1865–1884.

Gerber, S. 2013. On the relationship between the macroevolutionary trajectories of morphological integration and morphological disparity. PLoS One 8:e63913.

Giboulet, O., P. Chevret, R. Ramousse, F. Catzeflis. 1997. DNA-DNA hybridization evidence for the recent origin of marmots and ground squirrels (Rodentia: Sciuridae). Journal of Mammalian Evolution 4:271–284.

Gomez-Robles, A., and P.D. Polly. 2012. Morphological integration in the hominin dentition: evolutionary, developmental, and functional factors. Evolution 66:1024–1043.

Goodwin, H.T. 2008. Sciuridae. Pp. 355–376 *in* Janis, Christine M., G.F. Gunnell, M.D. Uhen, eds. Evolution of Tertiary Mammals of North America: Small Mammals, Xenarthrans, and Marine Mammals. Cambridge University Press, Cambridge.

Goodwin, H.T. 2009. Odontometric patterns in the radiation of extant ground-dwelling squirrels within Marmotini (Sciuridae: Xerini). Journal of Mammalogy 90:1009–1019.

Goswami, A. 2006. Morphological integration in the carnivoran skull. Evolution 60:169–183.

Goswami, A., J.B. Smaers, C. Soligo, and P.D. Polly. 2014. The macroevolutionary consequences of phenotypic integration: from development to deep time. Philosophical Transactions of the Royal Society B 369:20130254.

Gould, S.J. 1966. Allomery and size in ontogeny and phylogeny. Biological Reviews 41:587–640.

Gould, S.J., and R.C. Lewontin. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proceedings of the Royal Society B 205:581–598.

Grafen, A. 1989. The phylogenetic regression. Philosophical Transactions of the Royal Society B 326:119–157.

Hafner, D.J. 1984. Evolutionary Relationships of the Nearctic Sciuridae. Pp. 3–23 *in* J. O. Murie and G.R. Michener, eds. The Biology of Ground-Dwelling Sciurids. University of Nebraska Press, Lincoln.

Hallgrimsson, B., D.E. Lieberman, W. Liu, A.F. Ford-Hutchinson, F.R. Jirik. 2007. Epigenetic interactions and the structure of phenotypic variation in the cranium. Evolution and Development 9:76–91.

Harmon, L.J., J.J. Kolbe, J.M. Cheverud, J.B. Losos. 2005. Convergence and the multidimensional niche. Evolution 59:409–421.

Harmon, L.J., J.B. Losos, T. Jonathan Davies, R.G. Gillespie, J.L. Gittleman, W. Bryan Jennings, K.H. Kozak, M.A. McPeek, F. Moreno-Roark, T.J. Near, A. Purvis, R.E. Ricklefs, D. Schluter, J.A. Schulte, O. Seehausen, B.L. Sidlauskas, O. Torres-Carvajal, J.T. Weir, A.T. Mooers. 2010. Early bursts of body size and shape evolution are rare in comparative data. Evolution 64:2385–2396.

Harmon, L.J., J.T. Weir, C.D. Brock, R.E. Glor, W. Challenger. 2008. GEIGER: investigating evolutionary radiations. Bioinformatics 24:129–131.

Harrison, R.G., S.M. Bogdanowicz, R.S. Hoffmann, E. Yensen, P.W. Sherman. 2003. Phylogeny and evolutionary history of the ground squirrels (Rodentia: Marmotinae). Journal of Mammalian Evolution 10:249–276.

Hautier, L., R. Lebrun, P.G. Cox. 2012. Patterns of covariation in the masticatory apparatus of hystricognathous rodents: implications for evolution and diversification. Journal of Morphology 273:1319–1337.

Hayssen, V. 2008. Patterns of body and tail length and body mass in Sciuridae. Journal of Mammalogy 89:852–873.

Helgen, K.M., F.R. Cole, L.E. Helgen, D.E. Wilson. 2009. Generic revision in the Holarctic ground squirrel genus *Spermophilus*. Journal of Mammalogy 90:270–305.

Hopkins, M.J., S. Lidgard. 2012. Evolutionary mode routinely varies among morphological traits within fossil species. Proceedings of the National Academy of Sciences 109:20520–20525.

Hopkins, M.J., A.B. Smith. 2015. Dynamic evolutionary change in post-Paleozoic echinoids and the importance of scale when interpreting changes in rates of evolution. Proceedings of the National Academy of Sciences 112:3758–3763.

Howell, A.H. 1938. Revision of the North American Ground Squirrels, with a Classification of the North American Sciuridae. North American Fauna 56:1–256. Washington, D.C.

Hunt, G. 2007. The relative importance of directional change, random walks, and stasis in the evolution of fossil lineages. Proceedings of the National Academy of Sciences 104:18404–18408.

Hunt, G., M.J. Hopkins, S. Lidgard. 2015. Simple versus complex models of trait evolution and stasis as a response to environmental change. Proceedings of the National Academy of Sciences 112:4885–4890.

Ingram, T., Y. Kai. 2014. The geography of morphological convergence in the radiations of Pacific *Sebastes* rockfishes. The American Naturalist 184:E115–31.

Klingenberg, C.P. 1996. Multivariate allometry. Pp. 23–49 *in* L. F. Marcus, M. Corti, A. Loy, G. J. P. Naylor, and D. E. Slice, eds. Advances in Morphometrics. Plenum Press, New York.

Klingenberg, C.P. 2016. Size, shape, and form: concepts of allometry in geometric morphometrics. Development Genes and Evolution 226:113–137.

Klingenberg, C.P. 2014. Studying morphological integration and modularity at multiple levels: concepts and analysis. Philosophical Transactions of the Royal Society B 369:20130249.

Klingenberg, C.P., K. Mebus, J.C. Auffray. 2003. Developmental integration in a complex morphological structure: how distinct are the modules in the mouse mandible? Evolution and Development 5:522–531.

Lanfear, R., P.B. Frandsen, A.M. Wright, T. Senfeld, B. Calcott. 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Molecular Biology and Evolution 34:772–773.

Maddison, W. P. 1997. Gene trees in species trees. Systematic Biology 46:523–536.

Maestri, R., L.R. Monteiro, R. Fornel, N.S. Upham, B.D. Patterson, T.R.O. de Freitas. 2017. The ecology of a continental evolutionary radiation: is the radiation of sigmodontine rodents adaptive? Evolution *in press*.

Mahler, D.L., T. Ingram, L.J. Revell, J.B. Losos. 2013. Exceptional convergence on the macroevolutionary landscape in island lizard radiations. Science 341:292–295.

Mantooth, S.J., D.J. Hafner, R.W. Bryson, B.R. Riddle. 2013. Phylogeographic diversification of antelope squirrels (*Ammospermophilus*) across North American deserts. Biological Journal of the Linnean Society 109:949–967.

Marroig, G., and J. M. Cheverud. 2001. A comparison of phenotypic variation and covariation patterns and the role of phylogeny, ecology, and ontogeny during cranial evolution of New World monkeys. Evolution 55:2576–2600.

McLean, B. 2017. Digest: Splendid (continental) radiations. Evolution 71:802-803.

McLean, B.S., D.J. Jackson, J.A. Cook. 2016. Rapid divergence and gene flow at high latitudes shape the history of Holarctic ground squirrels (*Urocitellus*). Molecular Phylogenetics and Evolution 102:174–188.

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

Mirarab, S., R. Reaz, M.S. Bayzid, T. Zimmermann, M.S. Swenson, T. Warnow. 2014. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics 30:i541–i548.

Monteiro, L.R., M.R. Nogueira. 2010. Adaptive radiations, ecological specialization, and the evolutionary integration of complex morphological structures. Evolution 64:724–744.

Moore, T.Y., C.L. Organ, S.V. Edwards, A.A. Biewener, C.J. Tabin, F.A. Jenkins, K.L. Cooper. 2015. Multiple phylogenetically distinct events shaped the evolution of limb skeletal morphologies associated with bipedalism in the jerboas. Current Biology 25:2785–2794.

Paradis, E., B. Bolker, J. Claude, H.S. Cuong, R. Desper, B. Durand, J. Dutheil, O. Gascuel, C. Heibl, D. Lawson, V. Lefort, P. Legendre, J. Lemon, Y. Noel, J. Nylander, R. Opgen-Rhein, A.-A. Popescu, K. Schliep, K. Strimmer, D. de Vienne. 2015. ape: analyses of phylogenetics and evolution v4.1. https://cran.r-project.org/web/packages/ape/index.html.

Patterson, B.D., R.W. Norris. 2016. Towards a uniform nomenclature for ground squirrels: the status of the Holarctic chipmunks. Mammalia 80:241–251.

Phuong, M.A., M.C.W. Lim, D.R. Wait, K.C. Rowe, C. Moritz. 2014. Delimiting species in the genus *Otospermophilus* (Rodentia: Sciuridae), using genetics, ecology, and morphology. Biological Journal of the Linnean Society 113:1136–1151.

Pigliucci, M. 2003. Phenotypic integration: studying the ecology and evolution of complex phenotypes. Ecology Letters 6:265–272.

Polly, P.D., A. Cardini, E.B. Davis, S.J. Steppan. 2015. Marmot evolution and global change in the past 100 million years. Pp. 246–276 *in* P. G. Cox and L. Hautier, eds. Evolution of the Rodents: Advances in Phylogeny, Palaeontology and Functional Morphology. Cambridge University Press, Cambridge.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Renaud, S., P. Chevret, J. Michaux. 2007. Morphological vs. molecular evolution: ecology and phylogeny both shape the mandible of rodents. Zoologica Scripta 36:525–535.

Revell, L., L.J. Harmon, D. Collar. 2008. Phylogenetic signal, evolutionary process, and rate. Systematic Biology 57:591–601.

Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution 3:217–223.

Rolian, C., D.E. Lieberman, B. Hallgrímsson. 2010. The coevolution of human hands and feet. Evolution 64:1558–1568.

Roth, V.L. 1996. Cranial integration in the Sciuridae. American Zoologist 36:14–23.

Rowe, K.C., A.S. Achmadi, J.A. Esselstyn. 2016. Repeated evolution of carnivory among Indo-Australian rodents. Evolution 70:653–665.

Schliep, K., E. Paradis, A. Potts, M. Kendall. 2016. phangorn v2.1.1. https://cran.r-project.org/package=phangorn

Slater, G.J. 2015a. Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution. Proceedings of the National Academy of Sciences 112:4897–4902.

Slater, G.J. 2015b. Not-so-early bursts and the dynamic nature of morphological diversification. Proceedings of the National Academy of Sciences 112:3595–3596.

Smith, F.A, A.G. Boyer, J.H. Brown, D.P. Costa, T. Dayan, S.K.M. Ernest, A.R. Evans, M. Fortelius, J.L. Gittleman, M.J. Hamilton, L.E. Harding, K. Lintulaakso, S.K. Lyons, C. McCain, J.G. Okie, J.J. Saarinen, R.M. Sibly, P.R. Stephens, J. Theodor, M.D. Uhen. 2010. The evolution of maximum body size of terrestrial mammals. Science 330:1216–1219.

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

Stayton, C.T. 2015. The definition, recognition, and interpretation of convergent evolution, and two new measures for quantifying and assessing the significance of convergence. Evolution 69:2140–2153.

Thomas, W.K., S.L. Martin. 1993. A recent origin of marmots. Molecular Phylogenetics and Evolution 2:330–336.

Thorington, R.W., K. Darrow, A.D.K. Betts. 1997. Comparative myology of the forelimb of squirrels (Sciuridae). Journal of Morphology 234:155–182.

Thorington, R.W., J.L. Koprowski, M.L. Steele, J.F. Whatton. 2012. Squirrels of the World. The Johns Hopkins University Press, Baltimore.

Tsuboi, M., A. Gonzalez-Voyer, N. Kolm. 2014. Phenotypic integration of brain size and head morphology in Lake Tanganyika cichlids. BMC Evolutionary Biology 14:39.

Wilson, D.E., and D.M. Reeder (eds). 2005. Mammal Species of the World. 3rd ed. Johns Hopkins University Press, Baltimore, Maryland.

Wu, S., F. Zhang, S.V. Edwards, W. Wu, J. Ye, S. Bi, X. Ni, C. Quan, J. Meng, C.L. Organ. 2014. The evolution of bipedalism in jerboas (Rodentia: Dipodoidea): origin in humid and forested environments. Evolution 68:2108–2118.

Zelditch, M.L., J. Li, L.A.P. Tran, D.L. Swiderski. 2015. Relationships of diversity, disparity, and their evolutionary rates in squirrels (Sciuridae). Evolution 69:1284–1300.

Zelditch, M.L., J. Ye, J.S. Mitchell, D.L. Swiderski. 2017. Rare ecomorphological convergence on a complex adaptive landscape: body size and diet mediate evolution of jaw shape in squirrels (Sciuridae). Evolution *in press*.

FIGURE CAPTIONS

Figure 1. Preferred phylogenetic hypothesis of Marmotini based on partitioned analysis of 3 mtDNA regions and constrained with a backbone of 3,919 UCE loci from 32 taxa. Scores for the ecologically-associated principal components (PC) axis with highest variability in each trait are plotted for all available taxa. Non-grazing and grazing species are indicated by black and grey colors on PC axes, respectively.

Figure 2. Estimates of phylogenetic signal (multivariate Blomberg's *K*) for 3 skull traits in Marmotini. For each trait, *K* was calculated for 100 unique topologies that represent uncertainty in phlogenetic relationships (see text for futher details). Lines represent median values, edges of boxes represent 1st and 3rd quantiles, and whiskers represent 95% confidence intervals.

Figure 3. Observed within-group disparity (red arrows) compared to expectations under Brownian motion (grey histograms; 999 multivariate trait simulations) for 3 skull traits in different dietary ecotypes in Marmotini. Effects of excluding the distinctive genus *Marmota* from the grazing group are shown in middle column. P-values of significance tests are listed for each plot (top text), as well as the number of times the alternative hypothesis was accepted across 100 unique topological hypotheses (bottom text).

Figure 4. Observed number of convergent shifts in shape traits within dietary groups (red arrows) compared to that found in 999 trait simulations under an Ornstein-Uhlenbeck process (grey histograms). Effects of excluding the distinctive grazing genus *Marmota* are shown in middle column. P-values of significance tests are listed for each plot (top text), as well as the number of times the alternative hypothesis was accepted across 100 unique topological hypotheses (bottom text).

TABLES

Table 1. Results of PGLS analysis assessing contributions of dietary ecology and body size to skull shape traits. Statistically significant associations indicated in bold and asterisks.

| Factor | Df | SS | MS | $R^2$ | F | Z | P |
|---|---|---|---|---|---|---|---|
| Crania | | | | | | | |
| diet | 1 | 0.044 | 0.044 | 0.027 | 1.944 | 14.263 | **0.001** |
| log (body size) | 1 | 0.178 | 0.178 | 0.108 | 7.762 | 1.438 | 0.154 |
| diet x log (body size) | 1 | 0.021 | 0.021 | 0.012 | 0.918 | 0.664 | 0.468 |
| Residuals | 61 | 1.402 | 0.022 | | | | |
| Total | 64 | 1.646 | | | | | |
| Mandibles | | | | | | | |
| diet | 1 | 0.060 | 0.060 | 0.099 | 5.813 | 19.774 | **0.001** |
| log (body size) | 1 | 0.025 | 0.025 | 0.042 | 2.456 | 2.702 | **0.017** |
| diet x log (body size) | 1 | 0.020 | 0.020 | 0.032 | 1.917 | 3.017 | **0.011** |
| Residuals | 48 | 0.502 | 0.010 | | | | |
| Total | 51 | 0.608 | | | | | |
| Molariform Teeth | | | | | | | |
| diet | 1 | 414.64 | 414.64 | 0.237 | 46.560 | 31.109 | **0.001** |
| log (body size) | 1 | 846.05 | 846.05 | 0.483 | 95.003 | 29.321 | **0.001** |
| diet x log (body size) | 1 | 7.75 | 7.75 | 0.004 | 0.870 | 0.585 | 0.322 |
| Residuals | 54 | 480.89 | 8.91 | | | | |
| Total | 57 | 1749.34 | | | | | |

Table 2. Strength of integration among skull structures in Marmotini. For each pairwise trait comparison, the PLS correlation (r) and standardized effect size are listed. Values above the diagonal were computed from raw shape data and values below the diagonal were computed from size-corrected data. All pairwise tests of integration were significant ($P < 0.01$).

| | Crania | Mandibles | Molariform Teeth |
|---|---|---|---|
| Crania | - | r = 0.76<br>Z = 3.56 | r = 0.78<br>Z = 2.18 |
| Mandibles | r = 0.75<br>Z = 3.23 | - | r = 0.58<br>Z = 2.27 |
| Molariform Teeth | r = 0.84<br>Z = 5.53 | r = 0.61<br>Z = 3.15 | - |

Table 3. Fit of models of phenotypic evolution to 3 skull traits and body size in Marmotini (BM1, BM2=Brownian motion model with 1 and 2 states; OU1, OU2=Ornstein-Uhlenbeck model with 1 and 2 selective peaks; EB=Early burst). Best-fit models (based on ΔAICc values) are indicated in bold (logLik = log likelihood, AICc = sample size-corrected AIC, No. Param = total number of model parameters).

| | | Crania | | |
|---|---|---|---|---|
| Model | logLik | AICc | ΔAICc | No. Param. |
| OU2 | **965.26** | **-1838.96** | **-** | **40** |
| OU1 | 937.92 | -1797.13 | 41.83 | 35 |
| BM2 | 926.62 | -1774.52 | 64.44 | 35 |
| BM1 | 898.68 | -1754.59 | 84.37 | 20 |
| EB | 891.94 | -1738.82 | 100.14 | 21 |
| | | Mandibles | | |
| | logLik | AICc | ΔAICc | No. Param. |
| OU2 | **592.97** | **-1120.86** | **-** | **28** |
| OU1 | 583.41 | 1112.26 | 8.60 | 24 |
| BM2 | 571.35 | -1088.15 | 32.71 | 24 |
| BM1 | 554.88 | -1079.59 | 41.27 | 14 |
| EB | 550.18 | -1067.86 | 53.00 | 15 |
| | | Molariform Teeth | | |
| | logLik | AICc | ΔAICc | No. Param. |
| OU2 | -125.78 | 291.98 | 18.99 | 18 |
| OU1 | -143.06 | 319.16 | 46.18 | 15 |
| BM2 | **-119.97** | **272.99** | **-** | **15** |
| BM1 | -154.75 | 328.59 | 55.60 | 9 |
| EB | -154.75 | 330.84 | 57.86 | 10 |
| | | Body Size | | |
| | logLik | AICc | ΔAICc | No. Param. |
| OU2 | -343.40 | 695.47 | 12.58 | 4 |
| OU1 | -346.98 | 700.35 | 17.47 | 3 |
| BM2 | **-338.24** | **682.88** | **-** | **3** |
| BM1 | -347.61 | 699.42 | 16.53 | 2 |
| EB | -347.61 | 701.62 | 18.73 | 3 |

FIGURES

Figure 1.

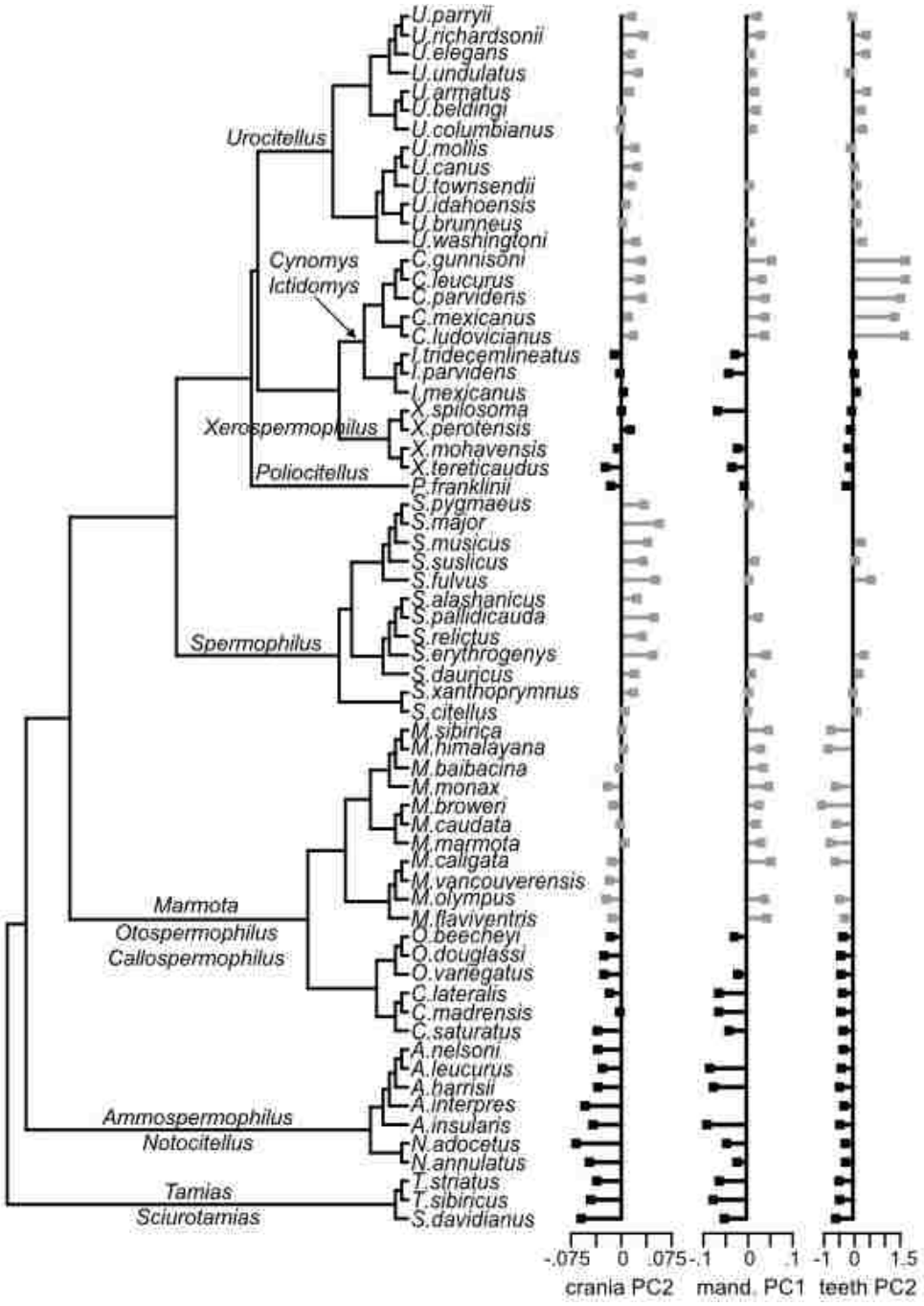


Figure 2.

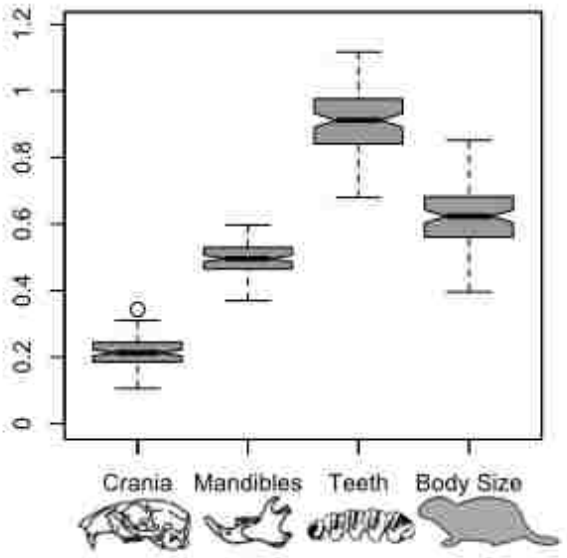


Figure 3.

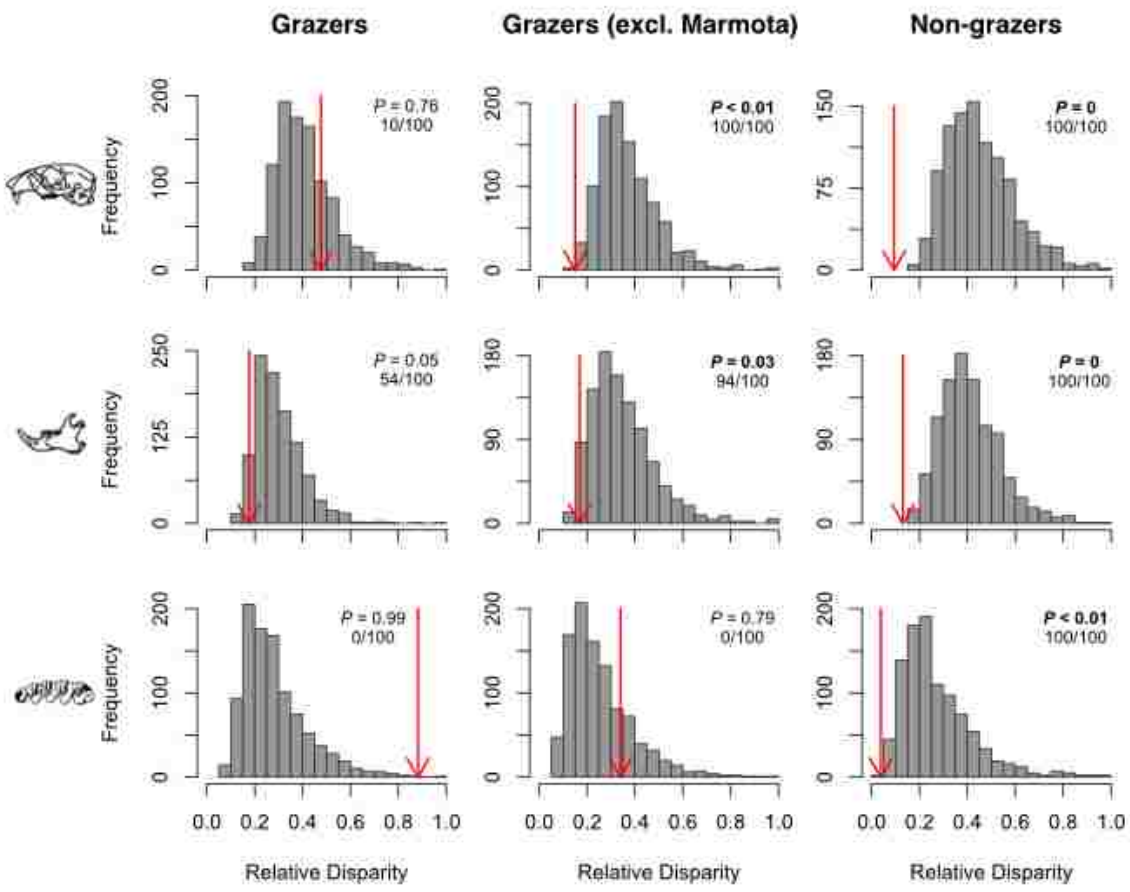
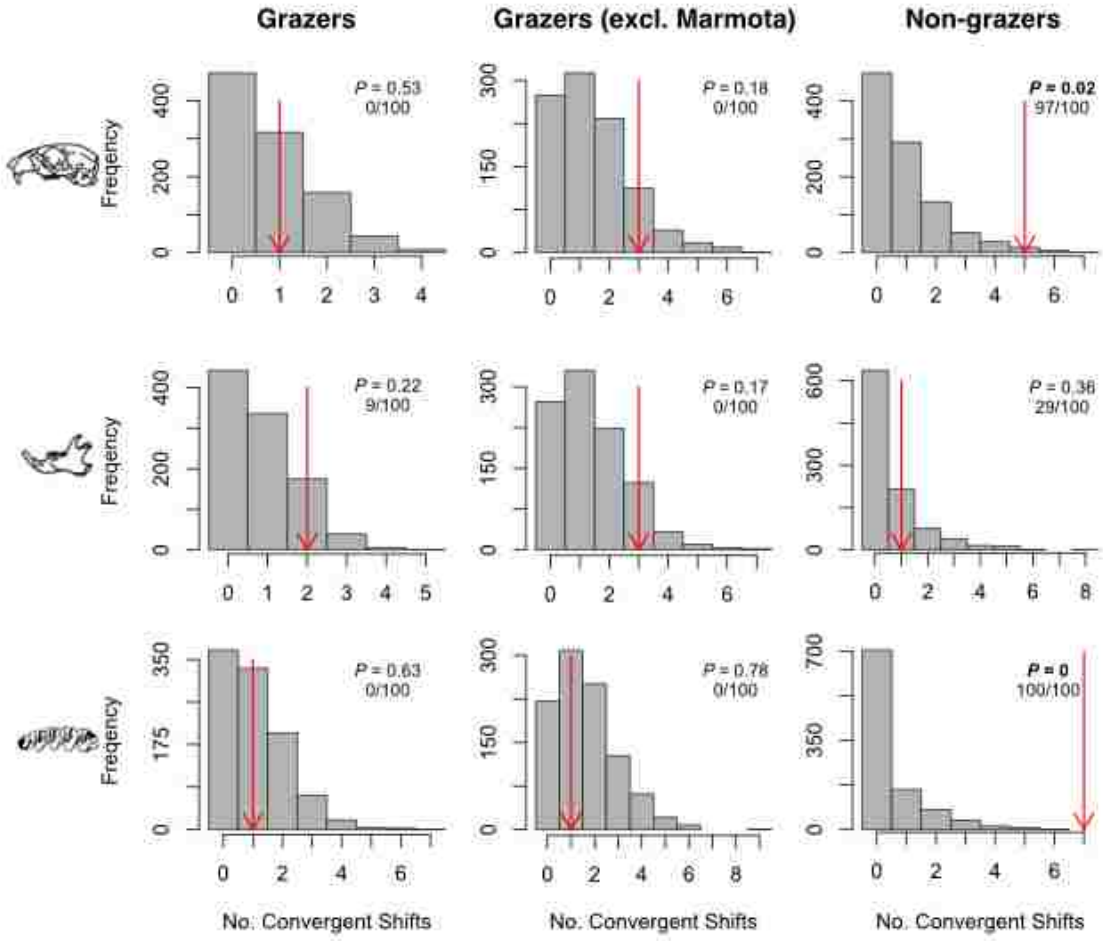

Grazers
Grazers (excl. Marmota)
Non-grazers
Frequency
P = 0.76
10/100
P < 0.01
100/100
P = 0
100/100
P = 0.05
54/100
P = 0.03
94/100
P = 0
100/100
P = 0.99
0/100
P = 0.79
0/100
P < 0.01
100/100
Relative Disparity
Relative Disparity
Relative Disparity

Figure 4.


Grazers
Grazers (excl. Marmota)
Non-grazers
Freqency
P = 0.53
0/100
P = 0.18
0/100
P = 0.02
97/100
P = 0.22
9/100
P = 0.17
0/100
P = 0.36
29/100
P = 0.63
0/100
P = 0.78
0/100
P = 0
100/100
No. Convergent Shifts
No. Convergent Shifts
No. Convergent Shifts

SUPPLEMENTARY FIGURE CAPTIONS

Supplementary Figure 1. Cranium of *Cynomys parvidens* (USNM 158060, female) in a) dorsal, b) lateral, and c) ventral views showing positions of 35 anatomical landmarks. Each landmark is shown in only 1 aspect. Fifteen semilandmarks were also digitized along the dorsal midline of the cranium, indicated by the dotted line in a). Scale bar equals 1cm.

Supplementary Figure 2. Results of principal components analysis of cranial shape in tribe Marmotini based on 50 3D landmarks, with non-grazing genera shown in warmer colors and grazing genera shown in cooler colors. a) PC1 vs PC2, b) PC2 vs PC3. See text and Supp. Fig. 1 for details of landmark data acquisition and analysis.

Supplementary Figure 3. Observed within-group disparity (red arrows) compared to expectations under Brownian motion (grey histograms; 999 multivariate trait simulations) for 3 size-corrected skull traits in different dietary ecotypes in Marmotini. Effects of excluding the distinctive genus *Marmota* from the grazing group are shown in middle column. P-values of significance tests are listed for each plot (top text), as well as the number of times the alternative hypothesis was accepted across 100 unique topological hypotheses (bottom text).

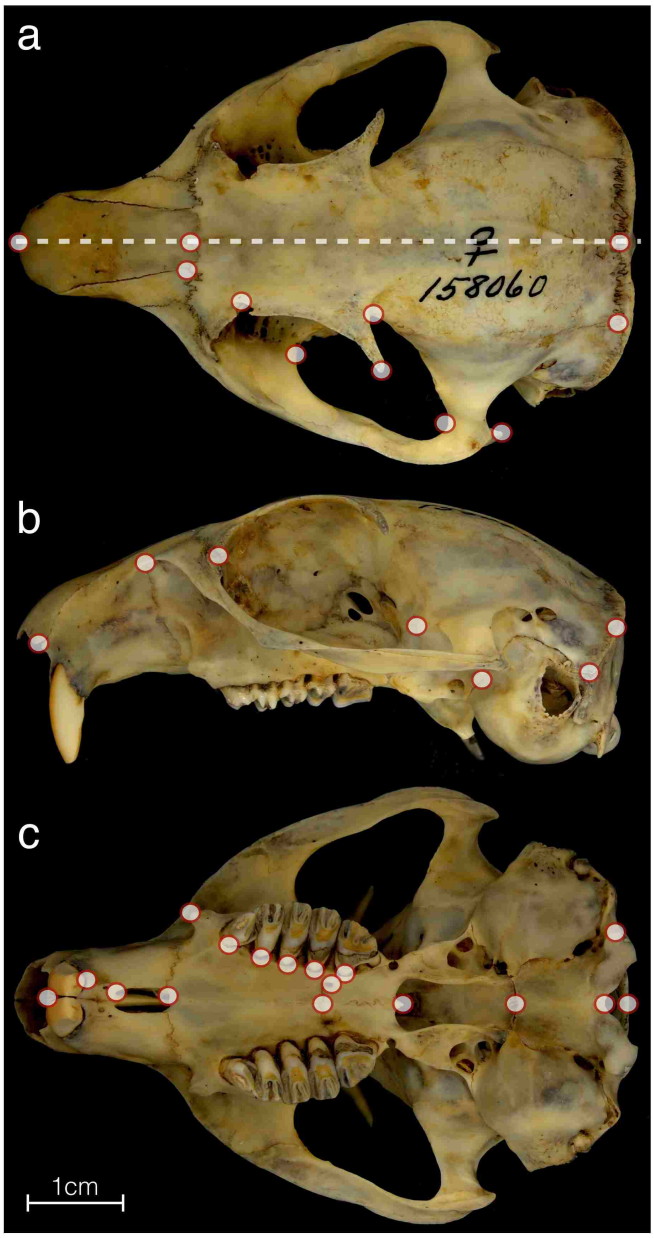
Supplementary Figure 4. Observed number of convergent shifts in size-corrected shape traits within dietary groups (red arrows) compared to that found in 999 trait simulations under a Brownian motion process (grey histograms). Effects of excluding the distinctive grazing genus *Marmota* are shown in middle column. P-values of significance tests are listed for each plot (top text), as well as the number of times the alternative hypothesis was accepted across 100 unique topological hypotheses (bottom text).

Supplementary Figure 5. Observed number of convergent shifts in size-corrected shape traits within dietary groups (red arrows) compared to that found in 999 trait simulations under an Ornstein-Uhlenbeck process (grey histograms). Effects of excluding the distinctive grazing genus *Marmota* are shown in middle column. P-values of significance tests are listed for each plot (top text), as well as the number of times the alternative hypothesis was accepted across 100 unique topological hypotheses (bottom text).

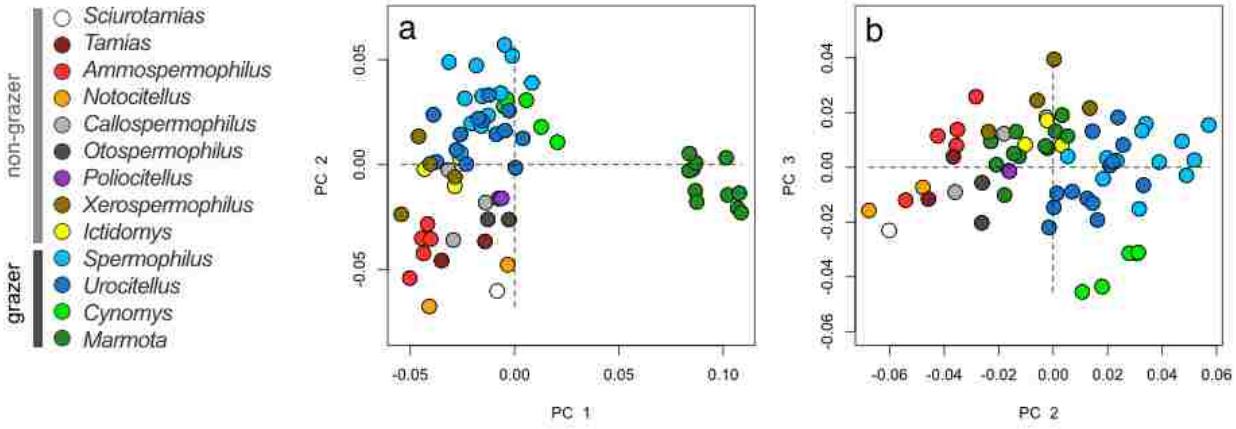
Supplementary Figure 6. Observed number of convergent shifts in size-corrected shape traits within dietary groups (red arrows) compared to that found in 999 trait simulations under a Brownian motion process (grey histograms). Effects of excluding the distinctive grazing genus *Marmota* are shown in middle column. P-values of significance tests are listed for each plot (top text), as well as the number of times the alternative hypothesis was accepted across 100 unique topological hypotheses (bottom text).

SUPPLEMENTARY FIGURES

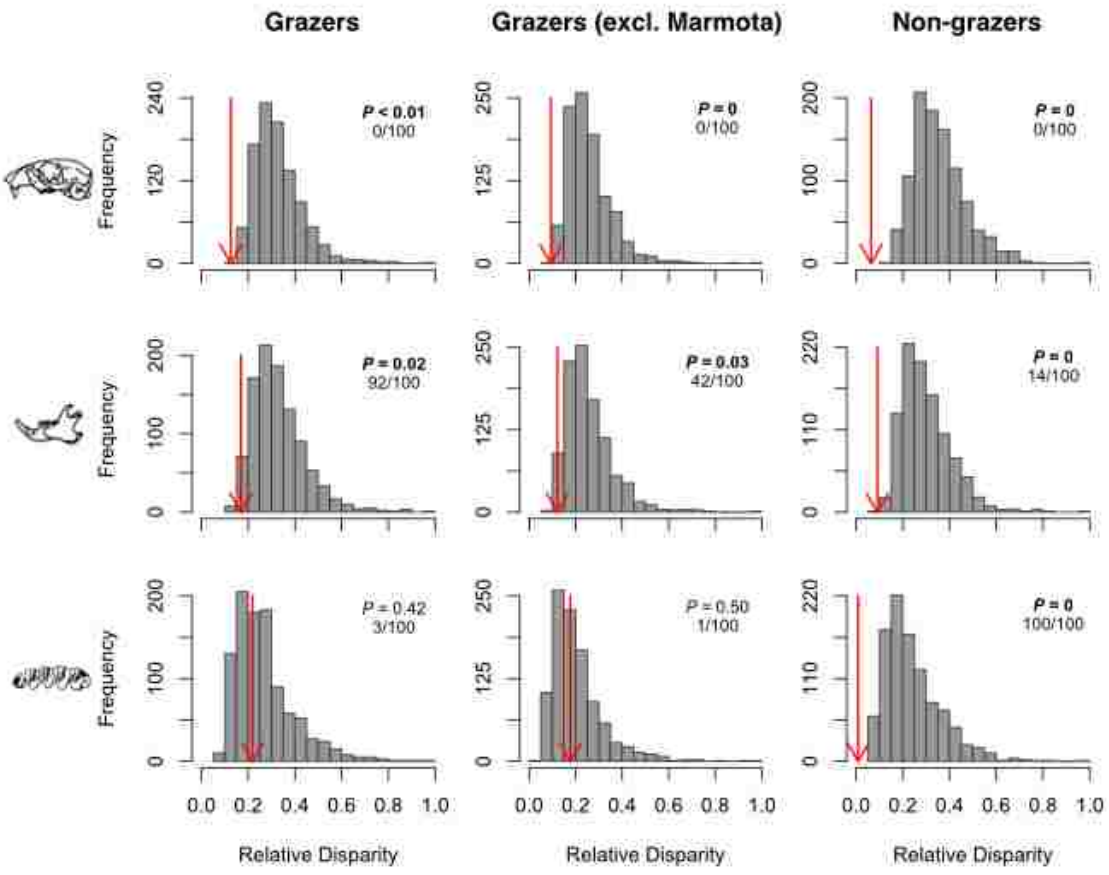
Supplementary Figure 1.



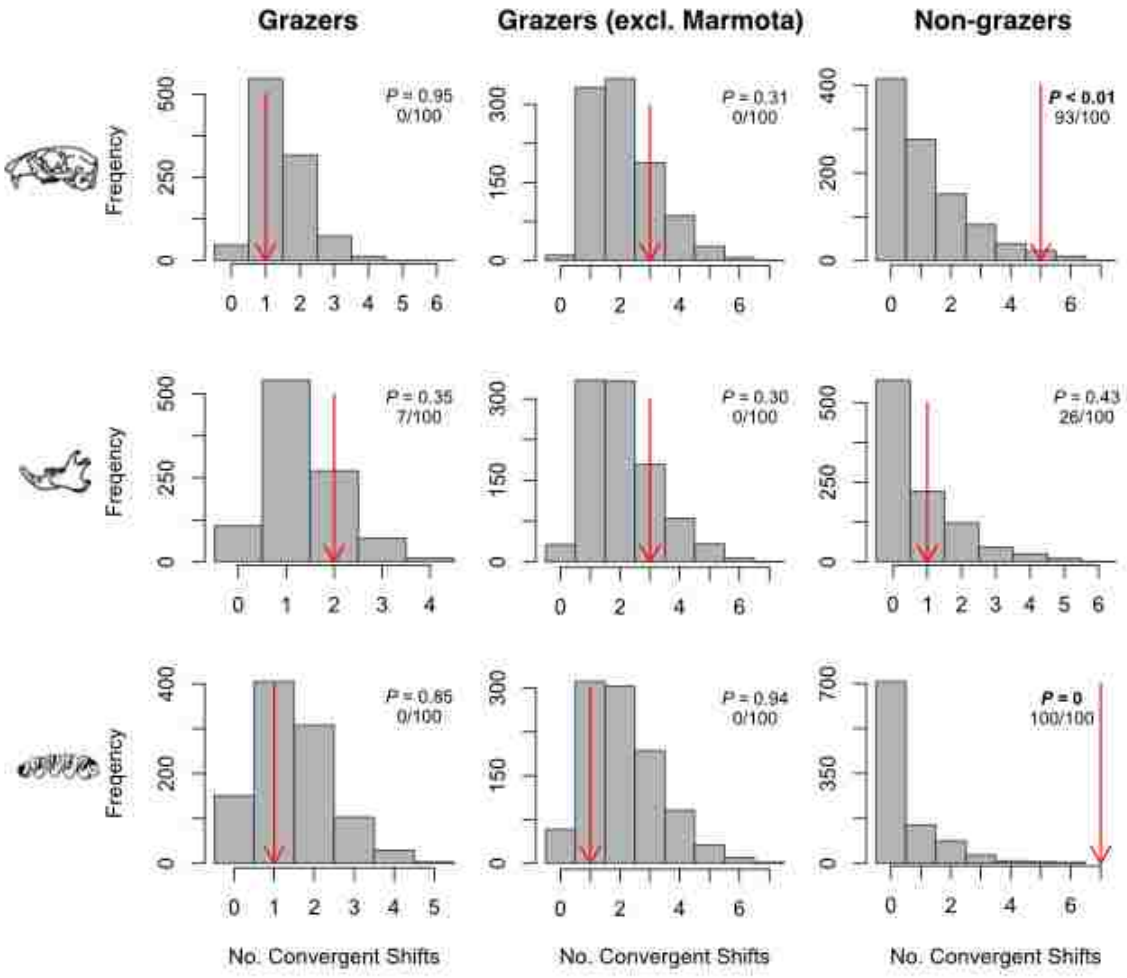
Supplementary Figure 2.



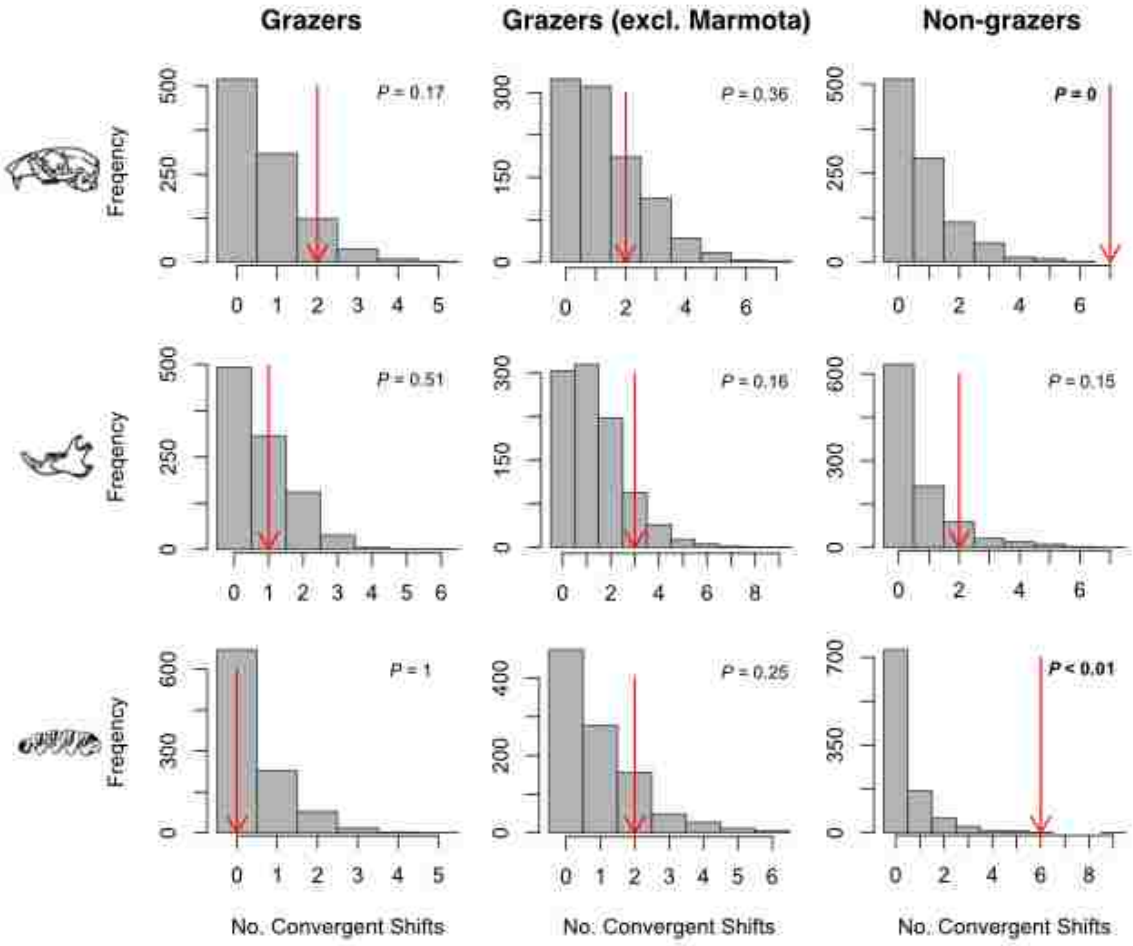
Supplementary Figure 3.



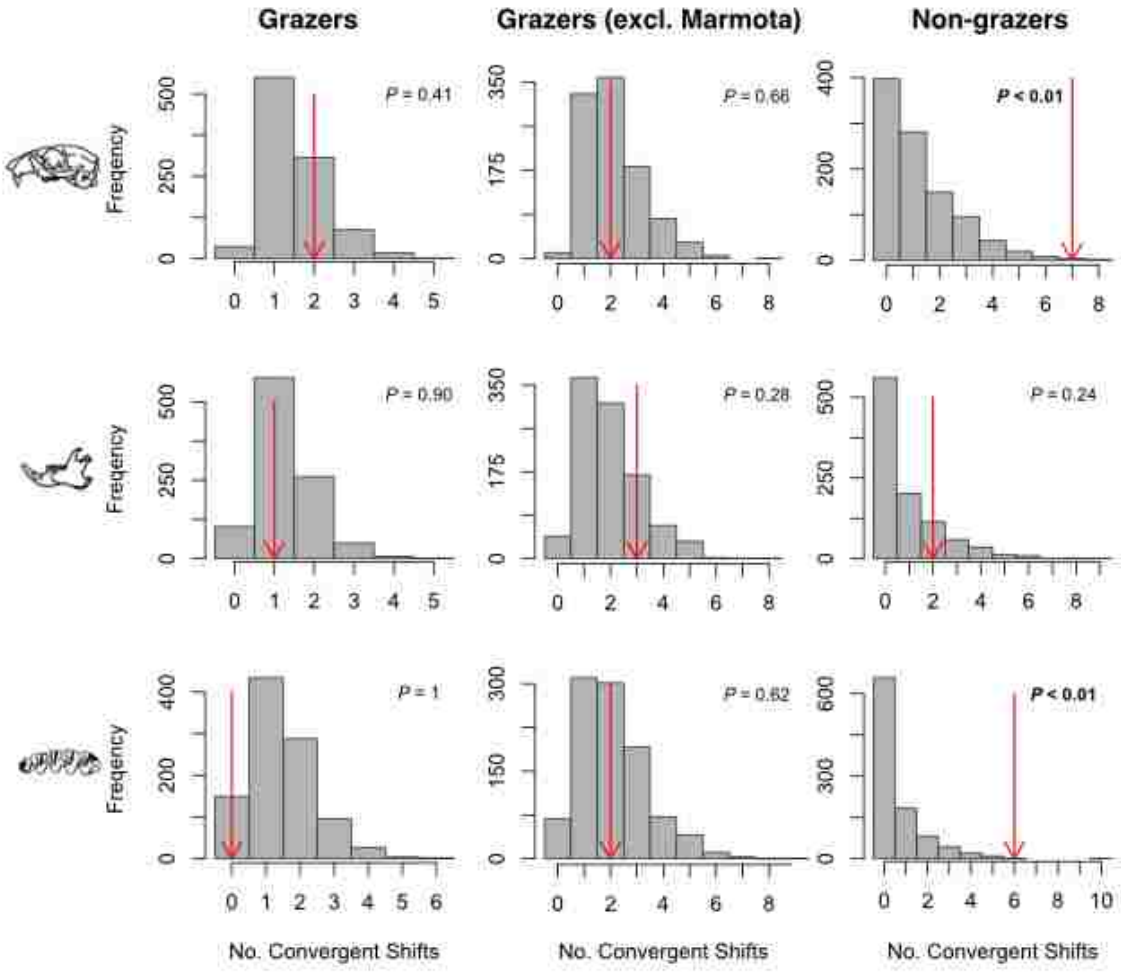
Supplementary Figure 4,



Supplementary Figure 5.



Supplementary Figure 6.



# CHAPTER 3

## Rapid divergence and gene flow at high latitudes shape the history of Holarctic ground squirrels (*Urocitellus*)

AUTHORS
Bryan S. McLean[1], Donavan J. Jackson[1], Joseph A. Cook[1]

[1] Department of Biology and Museum of Southwestern Biology, 1 University of New Mexico, MSC03-2020, Albuquerque, NM 87131 USA

ABSTRACT

Across the animal tree of life, the prevalence and evolutionary role(s) of hybridization remain incompletely understood. Rapidly radiating clades can serve as important systems for investigating these issues; however, such groups are often characterized by additional, widespread sources of gene tree discordance (e.g., incomplete lineage sorting). In this paper, we employed a multilocus dataset, Bayesian gene tree inference, and multiple species tree reconstruction methods to infer phylogeny of Holarctic ground squirrels (*Urocitellus*). We tested phylogenetic hypotheses based on previous morphological, cytological and single-locus datasets, and began to parse the causes of pervasive gene tree discordance that was observed. There is widespread incomplete lineage sorting in Urocitellus, consistent with rapid diversification embedded within the larger radiation of marmotine ground squirrels. We also recovered strong support for 2 instances of mitonuclear discord due to ancient hybridization among members of the high-latitude *parryii-richardsonii-elegans* clade. These results add to a growing number of documented hybridization events in ground squirrels, suggesting their radiation is a fertile system for understanding the interplay of diversification and hybridization in animal evolution.

KEYWORDS

Incomplete lineage sorting, Hybridization, Mitonuclear, Posterior predictive simulation, Radiation, Beringia

## INTRODUCTION

The importance of hybridization in animal diversification has become a central issue in evolutionary biology. Long known to be a common feature of plant evolution (Arnold 2006), it is only recently that hybridization has been recognized as widespread in animals (Mallet 2005). However, a multitude of studies now demonstrate gene flow at various points along the animal speciation continuum (Hedrick 2013, Toews and Brelsford 2012). In mammals specifically, divergence in the face of gene flow has been documented across multiple taxonomic orders (e.g., Evans et al. 2001, Koepfli et al. 2015, Melo-Ferreira et al. 2012, Sullivan et al. 2014). Still, our understanding of the true prevalence of hybridization as well as its primary role(s) across the animal Tree of Life remains incomplete.

Rapidly radiating clades are ideal systems for investigating the interplay of diversification and hybridization. Some have posited that hybridization can drive radiations, occurring early in divergence and serving primarily to transfer adaptive variation (Joyce et al. 2011, Seehausen 2004), while others have viewed hybridization as a byproduct of radiation, when speciation rates simply outpace the evolution of reproductive incompatibilities (Wiens et al. 2006). However, due to short times between speciation events, rapid radiations also carry high probabilities of incomplete lineage sorting (ILS) owing to failure of gene copies to coalesce within genetic lineages (Degnan and Rosenberg 2006, Maddison 1997). Identifying hybridization in the presence of ILS is problematic because both phenomena can result in similar phylogenetic patterns (Holder et al. 2001) as well as in gene tree discordance that is difficult to parse even with large numbers of loci (e.g., Leache et al. 2014). Moreover, hybridization is often unpredictable in its' extent, being limited to organellar genome flow in some systems (e.g., Good et al. 2015, Melo-Ferreira et al. 2011) but involving extensive nuclear gene flow in others (e.g., Cui et al. 2013, Martin et al. 2013).

Squirrels (Family Sciuridae, 58 genera) provide one example of apparently widespread hybridization in a mammal clade. In northern temperate latitudes alone, hybridization has been documented in flying squirrels (Garroway et al. 2010) and tree squirrels (Chavez et al. 2013), but most often in ground-dwelling squirrels of the tribe Marmotini (Hafner and Yates 1983, Kerhoulas et al. 2015, Nadler et al. 1971, Spiridinova et al. 2006, Thompson et al. 2015, Tsvirka et al. 2006). Marmotine ground squirrels represent a rapid, ongoing radiation (Hafner 1984, Zelditch et al. 2015) that consists of approximately 91 species in 12 genera distributed across North America and Asia (Helgen et al. 2009, Wilson and Reeder 2005). The dynamics of marmotine hybridization have been most thoroughly investigated in *Tamias* chipmunks, where mtDNA introgression has occurred repeatedly in parapatrically distributed, non-sister species across western North America (Good et al. 2008, Reid et al. 2010, Sullivan et al. 2014). Unfortunately, our understanding of core Marmotini phylogeny remains incomplete, obfuscating the extent to which similar dynamics characterize the remainder of this clade.

We reconstructed the evolutionary history of Holarctic ground squirrels (*Urocitellus*), a clade of medium-bodied, diurnal, and relatively social marmotine ground squirrels distributed across western North America, Beringia, and central Asia (Figs. 1 and 2). *Urocitellus* species have been the subject of extensive ecological, physiological and life history research over the past several decades (Barnes 1989, Boyer and Barnes

1999, Karels and Boonstra 2000, Murie and Michener 1984); still, a taxonomically comprehensive and well-resolved phylogeny of *Urocitellus* is lacking. We generated a multilocus dataset for *Urocitellus* that included all 12 species (and 33 of 36 subspecies) and used Bayesian gene tree inference as well as 3 species tree reconstruction methods to infer phylogeny and test longstanding phylogenetic hypotheses concerning this high-latitude genus (Section 2.1). We identified gene tree discordance due to both ILS and hybridization, and used posterior predictive simulations to assess consistency of our data with the latter process. We discuss these results in the broader context of the taxonomically widespread hybridization observed to date in ground squirrels, and attempt to identify potential drivers of these phenomena.

## METHODS

### *Phylogenetic Hypotheses*

We tested 6 hierarchical phylogenetic hypotheses based on previous morphological, cytological and single-locus DNA sequence datasets. *Urocitellus* has traditionally been divided into "small-eared" and "big-eared" species groups (Figs. 1 and 2), so we first tested if these groups are reciprocally monophyletic (Table 1). The small-eared group consists of 5 small bodied, xeric adapted species distributed in and around the Great Basin, USA (*U. washingtoni*, *U. brunneus*, *U. canus*, *U. mollis*, *U. townsendii*). The big-eared group contains 7 species of larger bodied and more mesic adapted squirrels widespread in the Intermountain West, Northern Rockies, Beringia and central Asia (*U. armatus*, *U. beldingi*, *U. columbianus*, *U. elegans*, *U. richardsonii*, *U. parryii*, *U. undulatus*). We included the Idaho ground squirrel (*U. brunneus*) in the small-eared group, but note its' previous alliance with the large-eared group based on several morphological characters (Davis 1939, Yensen and Sherman 1997).

Systematic relationships within the small-eared species group have been historically problematic. In particular, Townsend's ground squirrel (*U. townsendii*) has been considered conspecific with Merriam's ground squirrel (*U. canus*) and the Piute ground squirrel (*U. mollis*), an assemblage termed the '*townsendii* complex' (Howell 1938, Nadler et al. 1982; reviewed by Rickart 1987). These are currently considered 3 distinct species based on chromosomal variation (Cole and Wilson 2009, Nadler et al. 1982, Rickart 1985) but monophyly of this complex has not been established; thus it is unclear whether the morphological similarity of these species is due to convergence or common ancestry. We therefore tested monophyly of the *townsendii* complex (Table 1).

Previous sequence-based analyses of *Urocitellus* used a single mitochondrial (mtDNA) locus (CYTB; Harrison et al. 2003) and failed to recover reciprocally monophyletic big-eared and small-eared groups as well as a monophyletic *townsendii* complex. That analysis also recovered multiple instances of species non-monophyly. Specifically, Richardson's and Wyoming ground squirrels (*U. richardsonii* and *U. elegans*, respectively) were nested within an Arctic ground squirrel (*U. parryii*) clade, rendering *U. elegans* polyphyletic and *U. parryii* paraphyletic with respect to the former two taxa (also see Galbreath et al. 2011). Similarly, in the small-eared group, *U. mollis* was recovered as paraphyletic with respect to *U. canus* (Harrison et al. 2003). We tested

the monophyly of each of these currently recognized species (*U. elegans, U. parryii, U. mollis*; Table 1) using our multilocus dataset.

*Sample Selection and Sequencing*

We used 2 sampling approaches based on museum specimens (Appendix 1). The first was an expanded taxon sampling scheme, wherein 98 *Urocitellus* and 2 outgroup marmotines (*Ammospermophilus leucurus*, *Otospermophilus beecheyi*) were sequenced for up to 2 mtDNA and 1 nuclear locus (mtDNA CYTB and control region [CTRL]; partial von Willebrand factor [VWF]). The second approach aimed to expand character sampling: 44 of these individuals (including 2 outgroups) were sequenced at 4 additional nuclear loci (breast and ovarian cancer susceptibility 1 [BRCA], beta fibrinogen [FGB], glucosidase, beta, acid [GBA], growth hormone receptor [GHR]; Supplementary Table 2). Nuclear loci were chosen based on their utility for resolving rodent relationships at a variety of temporal scales (Adkins et al. 2001, Huchon et al. 1999). Several mtDNA sequences were also obtained from GenBank (Appendix 1).

Genomic DNA was extracted from frozen or ethanol-preserved liver or muscle tissues using Omega E.Z.N.A. extraction kits (Omega Bio-tek, Inc.) following manufacturer protocols. Ethanol-preserved samples were rinsed for 12-24 hours in dd$H_2O$ under refrigeration prior to extraction. A number of museum skins (ventral skin clips or footpads) were also sampled for one or both mtDNA loci. These were first washed in 95% ethanol (24 hours) and then STE buffer (24 hours) under refrigeration, with regular vortexing every few hours. Extractions proceeded using the same protocols as above. All extracts were quantified on a NanoDrop 2000 (Thermo Fisher Scientific).

Standard PCR amplification took place in 25µL reactions using primer combinations given in Table 1. All sequences were generated using Big Dye Terminator 3.1 technology (Applied Biosystems) on an ABI 3130 automated DNA sequencer in the Molecular Biology Facility in the Department of Biology at University of New Mexico. Sequences were manually edited in Sequencher v5.3 (Gene Codes Corp., Michigan, USA) and aligned using MUSCLE v3.7 (Edgar 2004) using default settings on the CIPRES science gateway (www.phylo.org). Raw sequence assemblies were deposited in GenBank (KX278443-KX278683, KX290205-290292).

*Tests of Selection*

We used the software Mega v.6.06 (Tamura et al. 2013) to conduct Z-tests of hypotheses of positive selection for each exonic locus in our dataset (CYTB, VWF, BRCA, GBA, GHR; FGB is largely intronic and was omitted from tests). For this test, GBA was trimmed to include only coding regions. The test statistic (Z) was calculated as the difference between nonsynonymous and synonymous substitutions (i.e., $d_N - d_S$) divided by the square of combined variances of each type of substitution. Hypothesis tests were conducted in a pairwise manner after first binning sequences by species. The Nei-Gojobori substitution model was employed, 500 bootstrap replicates were used, and statistical significance was assessed at the p=0.05 level.

*Phylogenetic Analysis*

Gene tree inference was conducted separately on mtDNA (partitioned by locus) and single nuclear locus datasets in MrBayes v3.2.3 (Ronquist and Huelsenbeck 2003) on CIPRES. Reversible jump MCMC (Huelsenbeck et al. 2004) was used to sample model space, allowing topological inference while incorporating uncertainty in substitution parameters. A gamma distribution with four categories was used to approximate among-site rate variation. Rate priors and other parameters (except topology and branch length) were allowed to vary among mtDNA partitions; default values were used for other priors in all analyses. Times to convergence varied significantly, with MCMC runs of 25 million (mtDNA), 50 million (BRCA, FGB, GHR), or 150 million (GBA, VWF) generations. Proper mixing was assessed by viewing trace plots in Tracer v1.6.0 (Rambaut et al. 2014), and we required effective sample sizes >200 for all estimated parameters. Posterior distributions of gene trees were summarized either as maximum clade credibility (MCC) phylograms (in text) or majority-rule consensus phylograms (Fig. 1) following removal of a burn-in of the first 20% of trees using TreeAnnotator v1.8.2.

Species trees were inferred from the expanded character dataset (42 *Urocitellus* + outgroups) using 3 different methods. First, we used *BEAST v1.8.1 (Drummond et al. 2012), which simultaneously infers gene trees and their encasing species tree in a hierarchical coalescent framework. *BEAST has been shown to have power in species tree inference even if ILS exists (e.g., Lanier et al. 2014). However, *BEAST does not model horizontal transfer as a source of gene tree discord, and results could be misleading if horizontal transfer goes undetected. *BEAST analyses were conducted using all loci and nuclear loci only (5), both via the CIPRES gateway. Analyses were run for 750 million generations, sampling posterior distributions every 2500 generations and assessing proper mixing in Tracer as described above. A burn-in of 50% of trees was removed and posterior distributions were summarized as MCCs as above.

The software STAR (Liu et al. 2009) and MP-EST (Liu et al. 2010) were also used to infer species trees from the expanded character dataset. Inputs for both are gene trees (which are assumed to be known without error). Of the variety of summary species tree methods recently developed, 2 broad categories exist: those that use only gene tree topologies and those that also use the coalescence times inferred on gene trees (DeGiorgio and Degnan 2014). STAR and MP-EST are in the former category (i.e., use topologies). Briefly, STAR estimates species trees based on a distance matrix describing the average rank coalescences of all possible taxon pairs (Liu et al. 2009), while MP-EST estimates species trees by maximizing a pseudo-likelihood function which takes into account all possible rooted taxon triples in gene trees that could be encoded by a species tree (Liu et al. 2010).

We chose STAR and MP-EST for 2 reasons. First, because they do not model ploidy, the influence of mtDNA relative to nuclear loci is reduced. It has been shown that mtDNA can have disproportionately strong influence on species trees in *BEAST due to its high variability and because assumptions of lower ploidy (which is modeled by *BEAST) are not always met (Jockusch et al. 2015). This is especially true in organisms with female-biased sex ratios and/or high male migration, such as occurs in *Urocitellus* (Byrom and Krebs 1999, Karels and Boonstra 2000). Second, and relatedly, we used these methods because both may be sensitive to possible instances of mtDNA

introgression, which has been commonly documented in ground squirrels. We note that introgression violates the assumptions of all 3 methods used here; however, confident identification of introgressed loci is not always possible at shallow phylogenetic levels or when ILS is pervasive. STAR and MP-EST may therefore decrease the effects of such events if they have occurred (Liu et al. 2010).

Most coalescent-based summary methods (including STAR and MP-EST) assume gene trees are known without error, which is rarely the case empirically. Bootstrapping approaches aim to incorporate gene tree uncertainty into species tree estimates (Seo 2008); however, using the uncertainty from Bayesian posterior distributions, where gene trees are represented in approximate proportion to their likelihood, is a better approach (Liu et al. 2009). This method may be intractable with very large numbers of loci (due to time required for Bayesian gene tree analysis), but it is a reasonable alternative to bootstrapping using smaller numbers of loci. To accomplish this, we used custom scripts in R (R Core Team 2015) to randomly sample 1000 trees from each gene tree posterior from MrBayes runs after removing 20% burn-in. We ran both STAR and MP-EST for 1000 iterations, each time drawing a single tree from each of the randomly downsampled gene tree distributions without replacement (thus inputs for every iteration of the species tree reconstruction methods contained 6 randomly sampled gene trees, 1 per locus). We summarized the results of both analyses (total of 1000 species tree reconstructions each) as MCC cladograms in TreeAnnotator, and performed the routine 3 separate times to ensure congruence.

*Multidimensional Scaling*

To compare and visualize variation in individual gene tree distributions, we used the Tree Set Visualization package (Tree Set Vis; Hillis et al. 2005) in Mesquite v3.03 (Maddison and Maddison 2015). Tree Set Vis performs multidimensional scaling on a pairwise matrix of Robinson-Foulds distances among topologies. Analyses were conducted on 500 trees selected randomly from the 6 MrBayes posterior distributions (mtDNA + nuclear loci; $N$=3000 trees total), and all were run until the stress function stabilized. We repeated each analysis 3 times to ensure congruence.

*Haplotype Network*

Upon identifying possible instances of mitonuclear discord in *Urocitellus*, we assembled an extended CYTB dataset based the dataset generated in this study combined with additional published and unpublished sequences (Appendix 1). We did this to visualize in greater detail the distribution of haplotypic diversity for all subspecies of *U. parryii, U. richardsonii* and *U. elegans,* as well as just those taxa involved in putative hybridization events (i.e., *U. elegans elegans, U. richardsonii, U. parryii parryii, and U. p. kennicottii*). We computed both haplotype networks using the package ape (Paradis et al. 2013) in R. Default settings were used for all parameters; the default distance parameter is the uncorrected distance among sequences under an infinite-sites model.

*Hypothesis Tests*

We tested the 6 phylogenetic hypothesis referenced above (Section 2.1, Table 1) using nodal support values as a significance criterion. We considered nodal support (posterior probabilities [PP], or "bootstrap" values in the case of STAR and MP-EST) greater than or equal to 0.95 to constitute support for a particular hypothesis; positive or negative results of all tests are reported relative to this threshold.

However, there are two potential issues with this approach. First, although 0.95 PP is commonly considered high nodal support, PP is not directly interpretable in a frequentist framework (i.e., recovering a clade with 0.95 PP support is not statistically equivalent to rejecting a hypothesis of non-monophyly under a p-value of 0.05). Moreover, this approach can lead to inflated Type I error rates (Alfaro and Holder 2006). We maintain the PP=0.95 threshold but also refer readers to actual PP or nodal support values.

Second, PP from *BEAST and support values from our STAR and MP-EST bootstrap procedure are not directly comparable. The same would apply to comparisons of PP and traditional (nucleotide site) bootstrapping, although studies have shown that bootstrap proportions and PP sometimes have relatively high correlation (Alfaro et al. 2003). However, we note that support values for STAR/MP-EST summary trees provide a measure of clade support that incorporates uncertainty in gene tree distributions and model parameters, and thus remain more directly interpretable than support values generated using nucleotide and/or locus bootstrapping.

*Posterior Predictive Checking*

Following identification of putative mitonuclear discord in *Urocitellus*, we assessed this possibility more rigorously using the posterior predictive method of Joly et al. (2009). That method (implemented as JML; Joly 2012) compares genetic distances observed between taxon pairs with the minimum genetic distances found from simulation on a set of species trees, the latter of which have been inferred under the coalescent with no migration (e.g., in *BEAST). If empirical genetic distances are significantly lower than those from simulations (i.e., if they are lower than expected based on the time since speciation and given ILS), a hypothesis of discordance due to ILS is rejected, supporting hybridization as a potential cause of gene tree discordance.

We performed posterior predictive checks using the CYTB expanded taxon dataset described previously (Section 2.2). Although the CTRL dataset should also encode a signature of mtDNA introgression, it has been shown that JML has reduced power to distinguish hybridization from ILS as sequence length decreases (Joly et al. 2009). Given that property, we also removed 4 samples from the CYTB dataset with only partial sequences (representing *U. canus canus* and *U. c vigilis*). We used species tree files generated by *BEAST from the expanded character dataset (mtDNA + nuclear) as input for JML, removing a burn-in of 50% of trees as described previously. We employed a heredity scalar of 0.5, a GTR model of sequence evolution, and mean values for sequence parameters (base frequencies, rate matrix, and mutation rate) extracted from original *BEAST .log files. We ran a second JML analysis using median values for the above parameters to ensure consistency. Default values were used for all other input parameters. We ran simulations on 10000 species trees and assessed significance at the p=0.1 level (the default setting).

## RESULTS

### *Gene Tree Concordance and Discordance in Urocitellus*

The partitioned mtDNA dataset with extended taxon sampling yielded a robustly supported tree (Fig. 3). The majority of nodes received >0.95 PP, including all basal nodes except 3 (i.e., those depicting interspecies relationships), which represents an improvement on single mtDNA gene analyses (Harrison et al. 2003). Monophyly of the big-eared group is strongly supported while monophyly of the small-eared group is not (Table 1), and the latter is true whether or not *U. brunneus* is included in the small-eared group. The mtDNA topology also failed to support monophyly of the *townsendii* complex. All 3 instances of species mtDNA non-monophyly encountered previously (Galbreath et al. 2011, Harrison et al. 2003) were also recovered in our tree (Fig. 3, Table 1). Specifically, *U. mollis* is recovered as paraphyletic with respect to *U. canus*, and *U. richardsonii* and *U. elegans elegans* are both nested within an Arctic ground squirrel (*U. parryii*) clade, rendering the latter 2 taxa non-monophyletic. We note that the 2 remaining subspecies of *U. elegans* (*U. e. aureus* + *U. e. nevadensis*) are not nested within *U. parryii*, but positioned basal to it.

The VWF tree with expanded taxon sampling failed to support a big-eared clade or a small-eared clade at 0.95 PP (Fig. 4, Table 1); however, both groups were recovered with >0.80 PP. Support for a monophyletic *townsendii* complex was poor, with strong support (>0.95 PP) instead obtained for the inclusion of *U. washingtoni* within this group. As expected, the VWF tree differs from mtDNA in overall resolution; 7 of 12 *Urocitellus* species are recovered as non-monophyletic in VWF. Interestingly, however, *U. parryii* and *U. elegans* are not among them, although *U. mollis* is (Table 1).

VWF and the 4 additional nuclear loci analyzed herein had distinct mutational profiles, including differences in mutation rate and patterns of among-site variation (Supplementary Table 2). Still, no nuclear locus strongly supported monophyly of big-eared or small-eared groups or of the *townsendii* complex (Table 1, Fig. 1). Multidimensional scaling of Robinson-Foulds distances in Tree Set Vis revealed that posterior distributions of inferred gene trees (mtDNA + 5 nuclear) were completely non-overlapping in tree space (Fig. 5). This inability of multiple independent markers to track deep divergences is consistent with ILS as a result of rapid diversification, which also characterizes the tribe Marmotini as a whole (Hafner 1984, Zelditch et al. 2015). In such situations, many loci are expected to have low probabilities of coalescence between speciation events (Degnan and Rosenberg 2006).

Conversely, single nuclear loci appeared to support robust tests of the 3 hypotheses of species monophyly (*U. elegans, U. parryii, U. mollis*). This is not surprising because, if radiations are old enough, sufficient variation may be accumulated to discriminate species despite the fact that ILS obscures deeper relationships. Monophyly of *U. parryii* and *U. elegans* was supported by 4 of 5 nuclear loci each (Table 1, Fig. 1), providing strong evidence that both are valid species, concordant with morphological and cytological assessments as well as current taxonomic arrangements (Howell 1938, Nadler et al. 1984, Robinson and Hoffmann 1975). However, unlike in *U.*

*parryii* and *U. elegans*, no nuclear loci support *U. mollis* monophyly (Table 1, Figs. 4 and 1).

All tests of positive selection in coding loci (CYTB, VWF, BRCA, GBA, GHR) were negative at the p=0.05 level (Appendix 2). Thus, recent episodes of selection are not supported as contributors to gene tree discordance among these particular loci.

*Species Tree Reconstructions*

Despite low phylogenetic resolution in individual loci, we recovered stronger support for interspecies relationships using 3 species tree reconstruction methods based on the expanded character dataset. When all loci (mtDNA + nuclear) were included, monophyly of a big-eared group was consistently supported at 0.95 PP, whereas monophyly of a small-eared group was supported only in *BEAST analyses (Fig. 6, Table 1). Support for monophyly of the *townsendii* complex was supported by 2 of 3 methods (*BEAST and STAR). Conversely, when only nuclear loci were included, species trees failed to support any of these hypotheses at the 0.95 level except for in one instance (small-eared monophyly was supported by *BEAST). Nevertheless, we note that monophyly of the big-eared group, small-eared group, and *townsendii* complex as traditionally defined were each supported at >0.75 across methods and datasets (Fig. 6).

We found general topological correspondence among species trees and datasets, despite some variation in nodal support (Fig. 6). Inclusion of mtDNA increased support for hypotheses across species tree reconstruction methods, and impacted topology only minimally. We initially predicted that effects of mtDNA inclusion would be more significant, as the mtDNA tree differs from nuclear gene trees and the former marker can have disproportionate influence in *BEAST due to higher variability and/or violated assumptions related to its lower ploidy (e.g., Jockusch et al. 2015). On the contrary, *BEAST trees using all loci versus nuclear loci were topologically identical and, although nodal support differed, there were no consistent trends in these differences. Conversely, STAR and MP-EST produced topologies that differed by 1 branch each among datasets (*U. undulatus* in STAR, *U. mollis* in MP-EST), suggesting greater but still subtle influence of mtDNA in those methods.

Relationships among small-eared species were generally better resolved than among big-eared species (Fig. 6). Poor support for big-eared species relationships is driven by instability in placement of *U. armatus, U. beldingi,* and *U. columbianus.* A sister relationship of *U. armatus* and *U. beldingi* (parapatric species in western North America; Fig. 1) was strongly supported by mtDNA (Fig. 3) but not recovered in any species tree. Instead, the most common (but weakly supported) relationship was *U. beldingi* as sister to *U. columbianus.* The Columbian ground squirrel (*U. columbianus*) has sometimes been considered sister to the Arctic ground squirrel (*U. parryii*) based on morphological (Howell 1938) and chromosomal (Nadler 1966, Nadler et al. 1974, Nadler et al. 1984) data, but we found no support for that arrangement, consistent with previous work (Harrison et al. 2003, Robinson and Hoffmann 1975). Finally, placement of the long-tailed ground squirrel (*U. undulatus*) is inconsistent*;* this species is exclusively Palearctic in distribution and may represent a long branch that is difficult to place phylogenetically.

*Support for Hybridization*

Based on phylogenetic inferences (summarized in Table 1), we hypothesize that mitonuclear discord in the *parryii-richardsonii-elegans* (*PRE*) clade is the result of hybridization leading to mtDNA capture. This is based on the fact that a majority of nuclear loci support monophyly of *U. elegans* and *U. parryii*, and lineage sorting in mtDNA is expected to occur faster than in nuclear loci based on lower ploidy and higher mutation rate (e.g., SupplementaryTable 2). However, a single hybridization event between *U. parryii* and the most recent common ancestor of *U. richardsonii* and *U. elegans* (likely sister species, see 3.2 above) is unsupported because remaining subspecies of *U. elegans* (*U. e. aureus* + *U. e. nevadensis*) carry ancestral, non-introgressed haplotypes (Fig. 3). We therefore propose that 2 ancient, sequential hybridization events occurred in the *PRE* clade, first from high arctic populations of *U. parryii* (*U. p. kennicottii* or *U. p. parryii*) to a single subspecies of *U. elegans* (*U. e. elegans*), and second with the latter taxon acting as a mitochondrial vector, transmitting a recently acquired *U. parryii* mtDNA genome to its sister species *U. richardsonii*.

This sequential hybridization scenario is supported by the CYTB haplotype network (Fig. 7), calculated using specimens of the 4 relevant taxa (Table 1), as well as by genetic distances between those taxa returned from JML analyses (Table 2, and below). However, distances between *U. parryii-U. e. elegans* and *U. parryii-U. richardsonii* differ only minimally, and the directionality of gene flow we propose must be interpreted with caution until additional nuclear data are available to resolve species phylogeny. The haplotype network for all PRE taxa is also ambiguous with respect to the most parsimonious scenario of mtDNA gene flow (Supp. Fig. 2). Thus, we cannot at present reject an alternate scenario of introgression (*U. parryii -> U. richardsonii -> U. elegans*), nor the possibility that this phylogroup carries an ancestral *U. richardsonii* mtDNA genome, such that *U. richardsonii* hybridized with *U. parryii* (ancient) and *U. elegans* (more recently).

Results of posterior predictive checking in JML support the assertion that mitonuclear discordance in *Urocitellus* is due to a process other than ILS. JML runs recovered 5 interspecific comparisons (of 91 total) with observed mtDNA distances significantly lower than expected under a coalescent model (Table 2). Notably, these include all 3 pairwise species comparisons in the *PRE* clade. Inspection of JML output confirmed that these results are driven solely by comparisons among *U. richardsonii, U. parryii parryii*, *U. parryii kennicottii*, and *U. elegans elegans*, consistent with the hybridization scenario we outline above. Table 2 also lists the number of significant comparisons at the subspecific level as a percentage of all such possible comparisons, highlighting the taxonomically isolated nature of introgression.

In addition to the *PRE* clade, JML recovered comparisons between *U. beldingi* and each of *U. mollis* and *U. canus* as deviating from a strict coalescent model. This is interesting because no evidence for mtDNA introgression exists in these taxa (Fig. 3), and minimum genetic distances are much higher than within the *PRE* clade (Table 2). Moreover, no comparisons were significant at the p=0.05 level, which possibly is reflective of power issues. We return to both of these points in the Discussion.

*New phylogeographic insights in the Arctic ground squirrel*

The expanded taxon mtDNA dataset included complete taxon sampling at the subspecific level for *U. parryii*, including 2 subspecies never before sequenced (*U. p. parryii* and *U. p. stejnegeri*). We recovered 4 major mtDNA clades as observed by previous workers (Eddingsaas et al. 2004, Galbreath et al. 2011; Figs. 3 and 8): Arctic, Southeast (SE Alaska, NW Canada), Southwest (SW Alaska, including Alaskan Peninsula and Aleutian Arc), and Beringian (Siberia, Seward Peninsula, interior Alaska). Our reconstructions placed *U. p. parryii* within the Arctic clade, sister to *U. p. kennicottii*, and placed *U. p. stegnejeri* (Kamchatka, Russia) in the Southwest clade. These results demonstrate that 2 distinct *U. parryii* mtDNA clades currently have amphiberingian distributions.

## DISCUSSION

### *The Complex Genomic History of Urocitellus*

To date, a comprehensive and well-resolved phylogeny of *Urocitellus* has proven elusive. This is despite extensive research using chromosomal and karyotypic (Nadler 1966, Nadler et al. 1971, Nadler et al. 1984, Rickart 1985), protein electrophoretic (Nadler and Hoffmann 1977, Nadler et al. 1974) and morphological (Robinson and Hoffmann 1975) data. While these previous studies greatly clarified issues of systematics and biogeography, the lack of a model-based phylogenetic framework for analyzing these data, incomplete taxon sampling, and a paucity of molecular research in intervening years have left important questions unanswered (e.g., Table 1). We addressed this using the most taxonomically complete sequence dataset available for *Urocitellus*, evaluating mtDNA and nuclear loci independently, and utilizing 3 methodologically distinct species tree methods that should also differ in their sensitivity to mtDNA inclusion as well as to isolated introgression events.

A crucial first step in understanding *Urocitellus* diversification lies in fully resolving the phylogenetic status of big-eared and small-eared groups. Small-eared *Urocitellus* are small-bodied, adapted to life in arid environments and geographically narrowly restricted to the Great Basin and surrounding regions of western North America (Fig. 2). In contrast, big-eared species are up to an order of magnitude larger in size (in *U. parryii*), more mesic-adapted, occupy larger geographic ranges, and have a much wider latitudinal distribution that also spans multiple continents (Fig. 1). Some big-eared species also display extreme physiological adaptations for hibernation (Barnes 1989, Boyer and Barnes 1999). Reciprocal monophyly of these groups would therefore presumably represent an important adaptive dichotomy within the genus.

All species tree reconstruction methods (*BEAST, STAR, MP-EST) indicate this dichotomy exists, but support for monophyly of the groups varies from moderate to very high depending on the dataset and species tree reconstruction method. We also recovered moderate to high support (0.79-0.99 nodal support; Fig. 6) for monophyly of a '*townsendii* complex' within the small-eared group. These small-eared taxa are similar in morphology and protein electrophoretic characters (Nadler et al. 1974), but divergent in karyotype (2N=36-46; Nadler 1966, Rickart 1985). Thus, our results suggest that these

latter taxa are sister species that have undergone instances of significant chromosomal evolution in spite of their morphological conservatism.

We submit that our inability to accept the phylogenetic hypotheses above is due to pervasive ILS owing to rapid diversification of *Urocitellus*, which itself is reflective of the larger radiation of marmotine ground squirrels. That scenario is supported by the poor resolution within, as well as topological disparity among, gene trees (Figs. 5 and 1), each of which point to ILS concentrated deep in *Urocitellus* phylogeny. While this resulted in negative tests for many hypotheses, we note that such results may not be consistent with formal rejection of phylogenetic hypotheses. Instead, failure to obtain robust clade-level support may be due to low statistical power, directly resulting from insufficient information in nuclear loci. Nevertheless, such inconclusive results were surprising given the amount of data available (5270bp), variability present in locus-specific mutational profiles (Supplementary Table 2), and evidence from previous studies demonstrating the power of the coalescent function in *BEAST (Lanier et al. 2014) and the performance of STAR and MP-EST (Degiorgio and Degnan 2014) in spite of moderate ILS. In the future, higher statistical power may be found with genome-scale sampling of *Urocitellus*.

*Sources of mtDNA Non-Monophyly in Urocitellus*

ILS, hybridization and incorrect species delimitation are 3 common explanations for non-monophyly in gene trees of vertebrate species (McKay and Zink 2010). While hybridization and ILS result in true non-monophyly of gene copies, incorrect species delimitation can lead to discordance between gene trees and recognized taxonomic boundaries (McKay and Zink 2010). As we discuss, ILS is apparent at the interspecific level in *Urocitellus*, particularly in nuclear loci. Yet, despite this, species delimitation remains possible in many cases (Table 1 and Fig. 1). Our taxonomically comprehensive multilocus dataset therefore allowed robust evaluation of current systematic and taxonomic arrangements at and below the species level; this includes the latter 3 phylogenetic hypotheses outlined above (i.e., monophyly of *U. parryii*, *U. elegans* and *U. mollis*; Table 1).

Phylogenetic inferences, as well as posterior predictive simulations, support hybridization leading to introgression as the source of mtDNA non-monophyly in *U. parryii* and *U. elegans*. We hypothesize that mtDNA gene flow occurred twice: first, from the Arctic clade of *U. parryii* (*U. p. kennicottii* or *U. p. parryii*) to the subspecies *U. e. elegans*, and second, from *U. e. elegans* to *U. richardsonii* (Fig. 8). Our study demonstrates the importance of a posterior predictive approach in particular when attempting to distinguish hybridization from ILS as explaining mitonuclear discordance. While we were unable to reject a role for ILS at the p=0.05 level, we note that the statistical power of the JML approach is sensitive to misspecification of substitution model parameters, length of the locus under investigation, and the time elapsed between events of speciation and hybridization (Joly et al. 2009). Thus, close relationship of members of the *PRE* clade as well as the small length of CYTB (1140bp) could explain why we attained only marginal statistical significance in rejecting ILS in this instance. Additional mtDNA data (to increase sequence length) as well as nuclear data (to further increase precision of species tree estimates) would be useful to further test our proposed hybridization hypothesis.

Conversely, there is no support for either ILS or hybridization as a driver of mtDNA non-monophyly in *U. mollis*. ILS is unlikely due to the expected faster sorting of mtDNA relative to nuclear loci, coupled with the fact that nuclear loci often delimit other *Urocitellus* species (but never *U. mollis*; Table 1). Hybridization (with *U. canus*) was also unsupported by JML, despite the performance of the method for closely related species in the *PRE* clade (we note that *U. c. canus* was not included in JML runs, however). We therefore hypothesize that current taxonomy inaccurately reflects species boundaries in this taxon. Indeed, multiple species were originally described from within *U. mollis* (e.g., Merriam 1913), and electrophoretic differences were reported among subspecies (Nadler et al. 1974). Expanded sampling is therefore warranted for accurate species delimitation in *U. mollis*. Future work should focus on range-wide sampling, especially along the Snake River, Idaho, where ranges of subspecies *U. m. artemesiae* and *U. m. idahoensis* abut that of *U. m. mollis* (Fig. 2).

Finally, it is interesting that posterior predictive checking rejected a role for ILS in explaining patterns of mtDNA genetic distance between *U. beldingi* and *U. mollis* (the *U. beldingi* - *U. canus* comparison was also significant at the p=0.1 level; Table 2). There is no evidence for hybridization among these taxa in mtDNA. Instead of hybridization, our results may result from the large discrepancies in *U. beldingi* placement within the big-eared clade existing between mtDNA and species trees (Figs. 3 and 6). Specifically, uncertainty in *BEAST species tree posterior distributions could lead to an inflation of simulated mtDNA distances between *U. beldingi* and small-eared taxa such as *U. mollis* relative to that observed in empirical data. Such a scenario could result in increased susceptibility of the posterior predictive approach to Type I error in cases where species trees are poorly resolved, but mtDNA (or other single locus) trees display high and conflicting support. Application of the method to similar situations in additional clades is therefore warranted.

*The Biogeography of Diversification in Urocitellus*

Differences in the distribution and habitat preferences of small-eared and big-eared *Urocitellus* have long been recognized (Howell 1938, Davis 1939, Durrant and Hansen 1954), but new insights as well as questions emerge when viewed in light of our results. A clear feature of *Urocitellus* evolution is colonization of high latitudes and altitudes by members of the big-eared group (Fig. 1). Four of these species inhabit latitudes higher than 50°N: *U. columbianus* (northern Rocky Mountains), *U. richardsonii* (northern Great Plains), *U. parryii* (North American high arctic and Siberia), and *U. undulatus* (central Asia and Siberia). Our results suggest multiple high latitude colonizations occurred in the history of the genus, which is notable given the demands of arctic and subarctic existence on homeothermic animals. However, the number and timing of arctic colonizations in particular remains uncertain. The exclusively Palearctic *U. undulatus* must have colonized Asia via Beringia just as *U. parryii* has done more recently (Galbreath et al. 2011); if *U. undulatus* shares a most recent common ancestor with the *PRE* clade (as suggested by mtDNA, *BEAST and STAR analyses), then colonization of arctic habitats must have occurred only once. However, additional data are needed to rule out an alternative scenario of 2 distinct arctic colonization events.

The concept of high-latitude diversification and adaptation is perhaps best embodied by the Arctic ground squirrel (*U. parryii*), which is the northernmost species of squirrel in the world. Intraspecific diversification of *U. parryii* within Beringia has been strongly shaped by Quaternary climate change (Eddingsaas et al. 2004, Galbreath et al. 2011, Nadler and Hoffmann 1977), and our data shed additional light on that history. First, we show that the Arctic clade of *U. parryii* is widespread across the North American high arctic (Fig. 8), such that it likely expanded from a Nearctic glacial refugium following the Last Glacial Maximum. We also demonstrate the occurrence of 2 separate mtDNA clades in core Beringia during the late Pleistocene, each of which currently persists in North American and Asian refugia. Based on that result, these clades (Beringia and Southwest [Alaska] clades *sensu* Cook et al. 2010, Eddingsaas et al. 2004, Galbreath et al. 2011) are best considered the "Northern Beringia" and "Southern Beringia" clades (Fig. 8; for the purposes of this study, we include the Southeast clade of previous authors in the "Northern Beringia" clade). The alliance of *U. p. stejnegeri* with the southwestern Alaska mtDNA clade (*U. p. ablusus, U. p. kodiacensis, U. p. nebulicola*) to comprise the Southern Beringia clade reveals a previously unrecognized level of mtDNA diversity in the Palearctic, and we note that this fails to support recent taxonomic treatments subsuming all Palearctic *U. parryii* into a single subspecies (*U. p. leucostictus*; Kryštufek and Vohralík 2013).

The phylogeographic patterns discussed above are intriguing because late Pleistocene Beringia is often reconstructed as an open steppe environment. However, Guthrie (2001) proposed a Beringian "mesic buckle" as an ecological barrier for some Pleistocene fauna, comprised of mesic tundra or forest and stretching north-south in the region of the modern-day Bering Strait. If inhospitable enough, this or other features may have fostered the isolation of *U. parryii* mtDNA clades in core Beringia, even if they were somewhat porous in space or time. In contrast, nuclear data have been interpreted as telling a story of high migration between the Northern Beringia and Southern Beringia clades (Galbreath et al. 2011). Those authors hypothesized that the significant structuring of mtDNA relative to nuclear loci could be driven by female philopatry and highly male-biased dispersal (Byrom and Krebs 1999, Karels and Boonstra 2000), which our data further support. In the future, it will be important to investigate whether a narrative of elevated mtDNA diversification, driven by demographic biases and superimposed on contrasting nuclear backgrounds, characterizes additional *Urocitellus* across the Holarctic.

The biogeographic patterns we outline for the big-eared group stand in contrast to those of small-eared *Urocitellus*, which are geographically limited to areas in and surrounding the Great Basin. Range sizes of small-eared taxa are also relatively small (except *U. m. mollis*). However, to date, the processes that have promoted diversification of low-elevation taxa (such as *Urocitellus*) in this region remain poorly understood (Riddle et al. 2014). Figure 2 highlights the geographic apportionment of diversity in small-eared *Urocitellus*, which appears demarcated in several instances by major river drainages. We therefore consider it possible that the evolutionary narrative of small-eared taxa is one of long-term endemism and *in situ* allopatric diversification within the Great Basin shrub-steppe ecosystem, potentially induced by shifts in the presence and location of glacial lakes and ancient/modern river systems. Testing this hypothesis will be possible with ongoing geographic sampling and taxonomic reconsideration of the group.

Finally, full resolution of *Urocitellus* phylogeny and completion of species-level phylogeographies will be paramount for effective management of imperiled taxa. This is particularly true for small-eared species, whose disproportionate representation on conservation lists is at least partly related to their restricted geographic distribution. Under the U.S. Endangered Species Act, 1 subspecies (*U. brunneus brunneus*) is currently listed as Threatened, 1 species (*U. washingtoni*) is listed as Candidate, and another species is Of Concern (*U. townsendii*). Anecdotal evidence for population-level declines also exists in *U. canus* (Yensen and Sherman 2003). Perhaps the most pressing result of our analysis is recovery of the Idaho ground squirrel (*U. brunneus*) as basal to the remainder of small-eared species, a relationship not supported by mtDNA or even most nuclear genes individually (Figs. 3 and 1; Harrison et al. 2003). *U. brunneus* has the most restricted geographic distribution of any *Urocitellus* (5 counties in western Idaho; Fig. 2), and our data suggest the possibility that it is an ancient and divergent lineage. This places renewed importance on continued monitoring of natural and anthropogenic threats to this taxon.

*What Drives Hybridization in Marmotini?*

Rapidly radiating clades are ideal systems for examining many evolutionary phenomena, including the interaction of hybridization and diversification. Maddison (1997) and others have suggested that phylogeny is not only a depiction of historical splitting events, but also a history of changing probabilities of interbreeding. Mallet (2005) calculated that roughly 10% of animal species hybridize, and it is now apparent that gene flow can proceed at various points along the speciation continuum without interrupting or reversing speciation itself (e.g., Evans et al. 2001, Koepfli et al. 2015, Melo-Ferreira et al. 2012, Sullivan et al. 2014). Yet some clades appear more prone to hybridization than others, and conclusive examples of hybridization now exist in at least 6 of 11 recognized marmotine genera containing >1 species (Hafner and Yates 1983, Kerhoulas et al. 2015, Nadler et al. 1971, Spiridinova et al. 2006, Sullivan et al. 2014, Thompson et al. 2015). But what are the factors facilitating such widespread gene flow?

We identify 3 potential drivers of hybridization in *Urocitellus* that are not mutually exclusive and may also extend to other ground squirrel genera. First, squirrels in general and ground squirrels in particular are morphologically conserved (Casanovas-Vilar and van Dam 2013, Emry and Thorington 1984, Zeldtich et al. 2015). Morphological diversification in *Urocitellus* has been reconstructed as among the lowest of all marmotine clades (Zelditch et al. 2015), and a significant amount of standing morphological variation in this genus may simply be due to allometry (Pearson 1981, Robinson and Hoffmann 1975, BSM unpublished data). A similar pattern of conservatism in those morphological traits directly involved in prezygotic incompatibility could facilitate widespread hybridization.

Second, conservatism in ecologically relevant traits, such as phenology and habitat preference, could also drive hybridization. There is overlap in both of these traits within each of the small-eared and big-eared groups of *Urocitellus*. For example, most big-eared species occur in parapatry across mesic and semi-arid grasslands of western North America (Fig. 1). They are known to partition habitat in zones of sympatry, but not in allopatry, suggesting a lack of ecological differentiation (Durrant and Hansen 1954).

Additionally, habitat partitioning in sympatry can be incomplete, resulting in existence of species pairs in syntopy (Davis 1939, Durrant and Hansen 1954, McLean personal observation). Such a lack of differentiation could facilitate hybridization if other major isolating mechanisms are absent, similar to what has been documented in non-sister species of *Tamias* chipmunks (Good et al. 2008, Reid et al. 2010, Sullivan et al. 2014).

Third, Quaternary climate change has repeatedly reorganized faunal and floral communities, creating significant opportunities for secondary contact and gene flow. This is particularly true at high northern latitudes (Hewitt 2000, Lessa et al. 2003), which are themselves emerging as prime venues for investigating the interplay of rodent diversification and hybridization (Chavez et al. 2014, Kerhoulas et al. 2015, Runck et al. 2009). The power of past climate change to create unique and unpredictable opportunities for gene flow is evidenced in the *PRE* clade; not only do all 3 taxa now occur in allopatry, but the directionality of gene flow we propose is unparsimonious with reference to current distributions. First, hybridization occurred between *U. parryii* and *U. e. elegans* despite the fact that the range of *U. richardsonii* currently lies between them. Second, according to our proposed scenario, it was the Arctic clade of *U. parryii* (not the geographically proximate *U. p. plesius*) that came into contact with *U. e. elegans.* Third, secondary introgression occurred between *U. e. elegans* and *U. richardsonii* despite the fact that *U. richardsonii* is currently sympatric only with *U. e. aureus*, with which it is documented to hybridize (Nadler et al. 1971) but has not shared mtDNA. Thus, while the lack of morphological or ecological isolating mechanisms sets a biological "stage" for hybridization, climatic and geographic contingencies can act as potent drivers of this process. A rigorous historical perspective will be crucial to future assessments of hybridization dynamics in other northern taxa.

*The Possibility of Adaptive Gene Flow in Urocitellus*

The ultimate role of hybridization in animal diversification remains incompletely understood (but see Hedrick 2013). Wiens et al. (2006) suggested that, in rapidly diversifying clades, hybridization might be a consequence of radiation, with speciation outpacing the evolution of reproductive incompatibilities. Seehausen (2004) suggested that hybridization might also drive radiations through transfer of adaptive variation. Both models accommodate hybridization as adaptive, but that of Seehausen (2004) posits a causal role for adaptive introgression in radiations.

Many authors have ascribed instances of mtDNA introgression to adaptive advantage of mtDNA haplotypes (Toews and Brelsford 2012), but evidence for this is slim, largely because it requires linking mtDNA functionality to individual fitness consequences (Storz and Wheat 2010). However, evidence is mounting that mtDNA variation does underlie important respiration-related adaptations (Shen et al. 2010, Scott et al. 2015, Toews et al. 2013, Wilson et al. 2013), and that differential fitness consequences can also result from interactions of mtDNA superimposed on different nuclear backgrounds (Ellison and Burton 2006, Hill et al. 2015, Levin et al. 2014).

Was hybridization adaptive in *Urocitellus*? Or was this a nonadaptive result of secondary contact between reproductively compatible species, driven by Quaternary climate cycling and/or demographic change? All *Urocitellus* are obligate hibernators (Thorington et al. 2012), but *U. parryii* is considered an extreme hibernator, capable of

enduring body temperatures below 0° C for several weeks (Barnes 1989, Buck and Barnes 1999). Mitochondria are main sites for cellular energy production and play crucial roles during hibernation and intermittent arousals in mammals, such as maintaining body temperatures above subzero arctic soil temperatures and functioning in non-shivering thermogenesis (Staples 2014, Staples and Brown 2008). The potential for mtDNA introgression in *Urocitellus* to be driven by adaptive advantage of *U. parryii* mtDNA should therefore continue to be explored.

## CONCLUSION

We investigated phylogeny of *Urocitellus*, a geographically widespread clade of ground squirrels whose evolutionary history has heretofore been incompletely understood. In doing so, we found the genomic history of this genus to be complex, characterized not only by pervasive ILS but also ancient hybridization among high-latitude taxa. These results represent important new insights into dynamics of *Urocitellus* diversification, but additional data are needed to fully parse these multiple drivers of gene tree discordance and thus resolve several recalcitrant phylogenetic and systematic questions. Application of genome-scale data to those issues, as well as to the issue of whether there is adaptive significance to the widespread hybridization that has been documented across Marmotini, is an important research direction to pursue.

ACKNOWLEDGMENTS

Funding for this research was provided by the National Science Foundation (NSF 1258010), UNM Department of Biology (Joseph Gaudin Scholarship), UNM Graduate and Professional Student Association, and UNM Biology Graduate Student Association. The following curators provided access to important museum specimens: Kristofer Helgen (NMNH); Jim Patton, Eileen Lacey, Michael Nachman (MVZ); Eric Rickart (UMNH); Sharlene Santana (UWBM); Larry Heaney, Bruce Patterson (FMNH); Duke Rogers (MLBM); Burton Lim (ROM); John Demboski (DMNS). We acknowledge important insights of Kurt Galbreath, Kayce Bell, and additional members of the Cook Lab at UNM that significantly improved the quality of this manuscript.

REFERENCES

Adkins, R.M., Gelke, E.L., Rowe, D., Honeycutt, R.L. 2001. Molecular phylogeny and divergence time estimates for major rodent groups: evidence from multiple genes. Molecular Biology and Evolution 18:777–791.

Alfaro, M.E., Holder, M.T. 2006. The posterior and the prior in Bayesian phylogenetics. Annual Reviews of Ecology, Evolution, and Systematics 37:19–42.

Alfaro, M.E., Zoller, S., Lutzoni, F., 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov Chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Molecular Biology and Evolution 20:255–266.

Arnold, M.L. 2006. Evolution Through Genetic Exchange. Oxford University Press, New York.

Barnes, B.M. 1989. Freeze avoidance in a mammal: body temperatures below 0 degrees C in an Arctic hibernator. Science 244:1593–1595.

Boyer, B.B., Barnes, B.M. 1999. Molecular and metabolic aspects of mammalian hibernation. Bioscience 49:713–724.

Byrom, A.E., Krebs, C.J. 1999. Natal dispersal of juvenile arctic ground squirrels in the boreal forest. Canadian Journal of Zoology 77:1048–1059.

Casanovas-Vilar, I., van Dam, J. 2013. Conservatism and adaptability during squirrel radiation: what is mandible shape telling us? PLoS One 8:e61298.

Chavez, A.S., Maher, S.P., Arbogast, B.S., Kenagy, G.J. 2014. Diversification and gene flow in nascent lineages of island and mainland North American tree squirrels (*Tamiasciurus*). Evolution 68:1094–1109.

Cole, F.R., Wilson, D.E. 2009. *Urocitellus canus*. Mammalian Species 834:1–8.

Cook, J.A., Eddingsaas, A.A., Loxterman, J.L., Ebbert, S., MacDonald, S.O. 2010. Insular arctic ground squirrels (*Spermophilus parryii*) of the North Pacific: indigenous or exotic? Journal of Mammalogy 91:1401–1412.

Cui, R., Schumer, M., Kruesi, K., Walter, R., Andolfatto, P., Rosenthal, G.G. 2013. Phylogenomics reveals extensive reticulate evolution in *Xiphophorus* fishes. Evolution 67:2166–2179.

Davis, W.B. 1939. The Recent Mammals of Idaho. The Caxton Printers, Ltd., Caldwell, Idaho.

DeGiorgio, M., Degnan, J.H. 2014. Robustness to divergence time underestimation when inferring species trees from estimated gene trees. Systematic Biology 63:66–82.

Degnan, J.H., Rosenberg, N.A. 2006. Discordance of species trees with their most likely gene trees. PLoS Genet. 2:e68.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution 29:1969–1973.

Durrant, S.D., Hansen, R.M. 1954. Distribution patterns and phylogeny of some western ground squirrels. Systematic Zoology 3:82–85.

Eddingsaas, A.A., Jacobsen, B.K., Lessa, E.P., Cook, J.A. 2004. Evolutionary history of the arctic ground squirrel (*Spermophilus parryii*) in Nearctic Beringia. Journal of Mammalogy 85:601–610.

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

Ellison, C.K., Burton, R.S. 2008. Genotype-dependent variation of mitochondrial transcriptional profiles in interpopulation hybrids. Proceedings of the National Academy of Sciences 105:15831–15836.

Emry, R.J., Thorington, R.W. 1982. Descriptive and comparative osteology of the oldest fossil squirrel, *Protosciurus* (Rodentia: Sciuridae). Smithsonian Contributions to Paleobiology 1–35.

Evans, B.J., Supriatna, J., Melnick, D.J. 2001. Hybridization and populations genetics of two macaque species in Sulawesi, Indonesia. *Evolution* 55:1886–1702.

Felsenstein, J., Churchill, G.A. 1996. A Hidden Markov Model approach to variation among sites in rate of evolution. Molecular Biology and Evolution 13:93–104.

Galbreath, K.E., Cook, J.A., Eddingsaas, A.A., Dechaine, E.G. 2011. Diversity and demography in Beringia: multilocus tests of paleodistribution models reveal the complex history of arctic ground squirrels. Evolution 65:1879–96.

Garroway, C.J., Bowman, J., Cascaden, T.J., Holloway, G.L., Mahan, C.G., Malcolm, J.R., Steele, M.A., Turner, G., Wilson, P.J. 2010. Climate change induced hybridization in flying squirrels. Global Change Biology 16:113–121.

Good, J.M., Hird, S., Reid, N., Demboski, J.R., Steppan, S.J., Martin-Nims, T.R., Sullivan, J. 2008. Ancient hybridization and mitochondrial capture between two species of chipmunks. Molecular Ecology 17:1313–1327.

Good, J.M., Vanderpool, D., Keeble, S., Bi, K. 2015. Negligible nuclear introgression despite complete mitochondrial capture between two species of chipmunks. Evolution 69:1961–1972.

Guthrie, R.D. 2001. Origin and causes of the mammoth steppe: a story of cloud cover, woolly mammal tooth pits, buckles, and inside-out Beringia. Quaternary Science Reviews 20:549–574.

Hafner, D.J. 1984. Evolutionary Relationships of the Nearctic Sciuridae, in: Murie, J.O., Michener, G.R. (Eds.), The Biology of Ground-Dwelling Sciurids. University of Nebraska Press, Lincoln, pp. 3–23.

Hafner, D.J., Yates, T.L. 1983. Systematic status of the Mojave ground squirrel, *Spermophilus mohavensis* (subgenus Xerospermophilus). Journal of Mammalogy 64:397–404.

Harrison, R.G., Bogdanowicz, S.M., Hoffmann, R.S., Yensen, E., Sherman, P.W. 2003. Phylogeny and evolutionary history of the ground squirrels (Rodentia: Marmotinae). Journal of Mammalian Evolution 10:249–276.

Hedrick, P.W. 2013. Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. Molecular Ecology 22:4606–4618.

Helgen, K.M., Cole, F.R., Helgen, L.E., Wilson, D.E. 2009. Generic revision in the Holarctic ground squirrel genus *Spermophilus*. Journal of Mammalogy 90:270–305.

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

Hill, G.E. 2015. Mitonuclear ecology. Molecular Biology and Evolution 32:1917–1927.

Hillis, D.M., Heath, T.A., St. John, K. 2005. Analysis and visualization of tree space. Systematic Biology 54, 471–482.

Holder, M.T., Anderson, J.A., Holloway, A.K. 2000. Difficulties in detecting hybridization. Systematic Biology 50:978–982.

Howell, A.H. 1938. Revision of the North American Ground Squirrels, with a Classification of the North American Sciuridae. North American Fauna 56.

Huchon, D., Catzeflis, F.M., Douzery, E.J.P. 1999. Molecular evolution of the nuclear von Willebrand factor gene in mammals and the phylogeny of rodents. Molecular Biology and Evolution 16:577–589.

Huelsenbeck, J.P., Larget, B., Alfaro, M.E. 2004. Bayesian phylogenetic model selection using reversible jump markov chain monte carlo. Molecular Biology and Evolution 21:1123–1133.

Jockusch, E.L., Martinez-Solano, I., Timpe, E.K. 2015. The effects of inference method, population sampling, and gene sampling on species tree inferences: an empirical study in slender salamanders (Plethodontidae: *Batrachoseps*). Systematic Biology 64:66–83.

Joly, S. 2012. JML: Testing hybridization from species trees. Molecular Ecology Resources 12:179–184.

Joly, S., McLenachan, P.A., Lockhart, P.J. 2009. A statistical approach for distinguishing hybridization and incomplete lineage sorting. The American Naturalist 174:E54–E70.

Joyce, D.A., Lunt, D.H., Genner, M.J., Turner, G.F., Bills, R., Seehausen, O. 2011. Repeated colonization and hybridization in Lake Malawi cichlids. Current Biology 21:R108–R109.

Karels, T.J., Boonstra, R. 2000. Concurrent density dependence and independence in populations of arctic ground squirrels. Nature 408:460–463.

Kerhoulas, N.J., Gunderson, A.M., Olson, L.E. 2015. Complex history of isolation and gene flow in hoary, Olympic, and endangered Vancouver Island marmots. Journal of Mammalogy 96:810–826.

Koepfli, K.-P., Pollinger, J., Godinho, R., Robinson, J., Lea, A., Hendricks, S., Schweizer, R.M., Thalmann, O., Silva, P., Fan, Z., Yurchenko, A.A., Dobrynin, P., Makunin, A., Cahill, J.A., Shapiro, B., Álvares, F., Brito, J.C., Geffen, E., Leonard, J.A., Helgen, K.M., Johnson, W.E., O'Brien, S.J., Van Valkenburgh, B., Wayne, R.K. 2015. Genome-wide evidence reveals that African and Eurasian golden jackals are distinct species. Current Biology 25:2158–2165.

Krystufek, B., Vohralik, V. 2013. Taxonomic revision of the Palaearctic rodents (Rodentia). Part 2. Sciuridae: *Urocitellus, Marmota* and *Sciurotamias*. Lynx 44:27–138.

Lanier, H.C., Huang, H., Knowles, L.L. 2014. How low can you go? The effects of mutation rate on the accuracy of species-tree estimation. Molecular Phylogenetics and Evolution 70:112–119.

Leache, A.D., Harris, R.B., Rannala, B., Yang, Z. 2014. The influence of gene flow on species tree estimation: a simulation study. Systematic Biology 63:17–30.

Lessa, E.P., Cook, J.A., Patton, J.L. 2003. Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary. Proceedings of the National Academy of Sciences 100:10331–10334.

Levin, L., Blumberg, A., Barshad, G., Mishmar, D. 2014. Mito-nuclear co-evolution: the positive and negative sides of functional ancient mutations. Frontiers in Genetics 5:448.

Liu, L., Yu, L., Edwards, S.V. 2010. A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. BMC Evolutionary Biology 10:302.

Liu, L., Yu, L., Pearl, D.K., Edwards, S.V. 2009. Estimating species phylogenies using coalescence times among sequences. Systematic Biology 58:468–77.

Maddison, W.P. 1997. Gene trees in species trees. Systematic Biology 46:523–536.

Maddison, W.P., Maddison, D.R. 2015. Mesquite: a modular system for evolutionary analysis.

Mallet, J. 2005. Hybridization as an invasion of the genome. Trends in Ecology and Evolution 20:229–237.

Martin, S.H., Dasmahapatra, K.K., Nadeau, N.J., Salazar, C., Walters, J.R., Simpson, F., Blaxter, M., Manica, A., Mallet, J., Jiggins, C.D. 2013. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. Genome Reseach 23:1817–1828.

McKay, B.D., Zink, R.M. 2010. The causes of mitochondrial DNA gene tree paraphyly in birds. Molecular Phylogenetics and Evolution 54, 647–650.

Melo-Ferreira, J., Boursot, P., Carneiro, M., Esteves, P.J., Farelo, L., Alves, P.C. 2012. Recurrent introgression of mitochondrial DNA among hares (*Lepus* spp.) revealed by species-tree inference and coalescent simulations. Systematic Biology 61:367–381.

Merriam, C.H. 1913. Six new ground squirrels of the *Citellus mollis* group from Idaho, Oregon, and Nevada. Proceedings of the Biological Society of Washington 26:135–138.

Murie, J.O., Michener, G.R. (Eds.) 1984. The Biology of Ground-Dwelling Squirrels: Annual Cycles, Behavioral Ecology and Sociality. University of Nebraska Press, Lincoln.

Nadler, C.F. 1966. Chromosomes and systematics of American ground squirrels of the subgenus *Spermophilus*. Journal of Mammalogy 47:579–596.

Nadler, C.F., Hoffmann, R.S. 1977. Patterns of evolution and migration in the arctic ground squirrel, *Spermophilus parryii* (Richardson). Canadian Journal of Zoology 55:748–758.

Nadler, C.F., Hoffmann, R.S., Greer, K.R. 1971. Chromosomal divergence during evolution of ground squirrel populations (Rodentia: *Spermophilus*). Systematic Zoology 20:298–305.

Nadler, C.F., Hoffmann, R.S., Vorontsov, N.N., Koeppl, J.W., Deutsch, L., Sukernik, R.I. 1982. Evolution in ground squirrels: II. Biochemical comparisons in Holarctic populations of *Spermophilus*. Zeitschrift Fur Saugetierkunde 47:198–215.

Nadler, C.F., Lyapunova, E.A., Hoffmann, R.S., Vorontsov, N.N., Shaitarova, L.L., Borisov, Y.M. 1984. Chromosomal evolution in Holarctic ground squirrels II. Giemsa-band homologies of chromosomes and the tempo of evolution. Zeitschrift Fur Saugetierkunde 49:78–90.

Nadler, C.F., Sukernik, R.I., Hoffmann, R.S., Vorontsov, N.N., Nadler Jr., C.F., Fomichova, I.I. 1974. Evolution in ground squirrels I. Transferrins in Holarctic populations of *Spermophilus*. Comparative Biochemistry and Physiology A 47:663–681.

Paradis, E., B. Bolker, J. Claude, H. S. Cuong, R. Desper, B. Durand, J. Dutheil, O. Gascuel, C. Heibl, D. Lawson, V. Lefort, P. Legendre, J. Lemon, Y. Noel, J. Nylander, R. Opgen-Rhein, A.-A. Popescu, K. Schliep, K. Strimmer, D. de Vienne. 2015. ape: analyses of phylogenetics and evolution v4.1. https://cran.r-project.org/web/packages/ape/index.html.

Pfeifer, B., Wittelsbuerger, U., Ramos-Onsins, S.E., Lercher, M.J. 2014. PopGenome: an efficient Swiss army knife for population genomic analyses in R. Molecular Biology and Evolution 31:1929–1936.

R Core Team. 2015. R: A language and environment for statistical computing.

Rambaut, A., Suchard, M.A., Xie, W., Drummond, A.J. 2014. Tracer.

Reid, N., Hird, S., Schulte-Hostedde, A., Sullivan, J. 2010. Examination of nuclear loci across a zone of mitochondrial introgression between *Tamias ruficaudus* and *T. amoenus*. Journal of Mammalogy 91:1389–1400.

Rickart, E.A. 1987. *Spermophilus townsendii*. Mammalian Species 268:1–6.

Rickart, E.A., Hoffmann, R.S., Rosenfield, M. 1985. Karyotype of *Spermophilus townsendii artemesiae* (Rodentia:Sciuridae) and chromosomal variation in the

*Spermophilus townsendii* complex. Mammalian Chromosomes Newsletter 26:94–102.

Riddle, B.R., Jezkova, T., Hornsby, A.D., Matocq, M.D. 2014. Assembling the modern Great Basin mammal biota: insights from molecular biogeography and the fossil record. Journal of Mammalogy 95:1107–1127.

Robinson, J.W., Hoffmann, R.S. 1975. Geographical and interspecific cranial variation in big-eared ground squirrels (*Spermophilus*): a multivariate study. Systematic Zoology 24:79–88.

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

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

Scott, G.R., Hawkes, L.A., Frappell, P.B., Butler, P.J., Bishop, C.M., Milsom, W.K. 2015. How Bar-Headed Geese fly over the Himalayas. Physiology 30:107–115.

Seehausen, O. 2004. Hybridization and adaptive radiation. Trends in Ecology and Evolution 19:198–207.

Seo, T.-K. 2008. Calculating bootstrap probabilities of phylogeny using multilocus sequence data. Molecular Biology and Evolution 25:960–971.

Shen, Y.-Y., Liang, L., Zhu, Z.-H., Zhou, W.-P., Irwin, D.M., Zhang, Y.-P. 2010. Adaptive evolution of energy metabolism genes and the origin of flight in bats. Proceedings of the National Academy of Sciences 107, 8666–8671.

Spiridonova, L.N., Chelomina, G.N., Tsuda, K., Yonekawa, H., Starikov, V.P. 2006. Genetic evidence of extensive introgression of short-tailed ground squirrel genes in a hybridization zone of *Spermophilus major* and *S. erythrogenys*, inferred from sequencing of the mtDNA cytochrome b gene. Russian Journal of Genetics 42:802–809.

Staples, J.F. 2014. Metabolic suppression in mammalian hibernation: the role of mitochondria. Journal of Experimental Biology 217:2032–2036.

Staples, J.F., Brown, J.C.L. 2008. Mitochondrial metabolism in hibernation and daily torpor: a review. Journal of Comparative Physiology B 178:811–827.

Storz, J.F., Wheat, C.W. 2010. Integrating evolutionary and functional approaches to infer adaptation at specific loci. Evolution 64:2489–2509.

Sullivan, J., Demboski, J.R., Bell, K.C., Hird, S., Sarver, B., Reid, N., Good, J.M. 2014. Divergence with gene flow within the recent chipmunk radiation (*Tamias*). Heredity 113:185–194.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution 30, 2725–2729.

Thompson, C.W., Stangl, F.B., Bradley, R.D. 2015. Ancient hybridization and subsequent mitochondrial capture in ground squirrels (genus *Ictidomys*). Occasional Papers of the Museum of Texas Tech University 331.

Thorington, R.W., Koprowski, J.L., Steele, M.L., Whatton, J.F. 2012. Squirrels of the World. The Johns Hopkins University Press, Baltimore.

Toews, D.P.L., Brelsford, A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. Molecular Ecology 21:3907–3930.

Toews, D.P.L., Mandic, M., Richards, J.G., Irwin, D.E. 2013. Migration, mitochondria, and the yellow-rumped warbler. Evolution 68:241–255.

Tsvirka, M. V., Chelomina, G.N., Korablev, V.P. 2006. Genetic evidence of hybridization between paletailed *Spermophilus pallidicauda* Satunin, 1903 and Alashanic *S. alaschanicus* Büchner, 1888 ground squirrels in Mongolia. Russian Journal of Genetics 42:421–428.

Wiens, J.J., Engstrom, T.N., Chippindale, P.T. 2006. Rapid diversification, incomplete isolation, and the "speciation clock" in North American salamanders (genus *Plethodon*): testing the hybrid swarm hypothesis of rapid radiation. Evolution 60:2585–2603.

Wilson, D.E., Reeder, D.M. (Eds.). 2005. Mammal Species of the World, 3rd ed. Johns Hopkins University Press, Baltimore.

Wilson, R.E., Peters, J.L., McCracken, K.G. 2013. Genetic and phenotypic divergence between low- and high-altitude populations of two recently diverged Cinnamon Teal subspecies. Evolution 67:170–184.

Yensen, E., Sherman, P.W. 1997. *Spermophilus brunneus*. Mammalian Species 560:1-5.

Yensen, E., Sherman, P.W. 2003. Ground-dwelling Squirrels of the Pacific Northwest. U.S. Fish and Wildlife Service, U.S. Bureau of Land Management, Boise.

Zelditch, M.L., Li, J., Tran, L.A.P., Swiderski, D.L. 2015. Relationships of diversity, disparity, and their evolutionary rates in squirrels (Sciuridae). Evolution 69:1284–1300.

FIGURE CAPTIONS

Figure 1. Geographic distribution of the big-eared species group of *Urocitellus* in the Palearctic (left) and Nearctic (right). Different species are indicated by numbers as follows: (1) *U. undulatus*, (2) *U. parryii*, (3) *U. columbianus*, (4) *U. richardsonii*, (5) *U. beldingi*, (6) *U. armatus*, and (7) *U. elegans*.

Figure 2. Geographic distribution of the small-eared species group of Urocitellus. Different species are indicated by different shapes. Subspecies are indicated by numbers as follows: (1) *U. townsendii nancyae*, (2) *U. t. townsendii*, (3) *U. washingtoni*, (4) *U. brunneus brunneus*, (5) *U. b. endemicus*, (6) *U. canus canus*, (7) *U. c. vigilis*, (8) *U. mollis idahoensis*, (9) *U. m. artemesiae*, (10) *U. m. mollis*. State abbreviations are CA = California, ID = Idaho, MT = Montana, NV = Nevada, OR = Oregon, UT = Utah, WA = Washington. Major rivers are indicated by white text.

Figure 3. Maximum clade credibility phylogram of *Urocitellus* based on Bayesian partitioned analysis of mtDNA (CYTB and control region, 1662 bp total). Nodes receiving >0.95 posterior probability (PP) support are indicated with closed circles. Three poorly-supported internal nodes are labeled specifically to highlight interspecies relationships. Bars at right are colored to represent recovery of species monophyly (black) or non-monophyly (gray). Sample locations (state/province) are indicated in underlined text at right of clades; all states in USA unless otherwise abbreviated (CAN = Canada, RUS = Russia, MON = Mongolia). Outgroups are *Ammospermophilus leucurus* and *Otospermophilus beecheyi*.

Figure 4. Maximum clade credibility phylogram of *Urocitellus* based on Bayesian analysis of partial von Willebrand's factor (VWF, 860 bp). Nodes receiving >0.95 posterior probability (PP) support are indicated with closed circles (as in Fig. 3). Bars at right are colored to represent species monophyly (black) or non-monophyly (gray). Outgroups as in Fig. 3.

Figure 5. Multidimensional scaling plot of Robinson-Foulds distances among gene trees. Trees were selected randomly from gene tree posterior distributions (N = 500 per locus; 3000 total).

Figure 6. Species tree reconstructions of *Urocitellus* using 3 different methods (*BEAST, STAR, MP-EST) based on all loci (top row) and nuclear loci only (bottom row). Trees are maximum clade credibility cladograms (i.e., w/o branch lengths). Outgroups as in Fig. 3.

Figure 7. Haplotype network based on mtDNA CYTB (1140 bp; sequences compiled from this and previous studies). Taxa included are *U. parryii parryii*, *U. parryii kennicottii*, *U. elegans elegans* and *U. richardsonii*. Labels indicate the state/province of collection (AK = Alaska, CO = Colorado, MT = Montana, WY = Wyoming; AB = Alberta, MB = Manitoba, NT = Northwest Territories, NU = Nunavut, SK =

Saskatchewan). Hash marks represent single mutations. Number of individuals is proportional to the size of the circle (see legend).

Figure 8. Geographic distribution of the entire *PRE* group with major mtDNA clades indicated by color/shading. Subspecies of *U. elegans* and *U. richardonii* are indicated with text, and subspecies of *U. parryii* are numbered as follows: (1) *U. p. ablusus*, (2) *U. p. kennicottii*, (3) *U. p. kodiacensis*, (4) *U. p. leucostictus*, (5) *U. p. lyratus*, (6) *U. p. nebulicola*, (7) *U. p. osgoodi*, (8) *U. p. parryii*, (9) *U. p. plesius*, (10) *U. p. stejnegeri*.

TABLES

Table 1. Phylogenetic hypotheses tested in this study (column 1) and their basis (column 2; M = morphology, K = karyotypic, mt = mitochondrial DNA; references are those most relevant to the hypotheses presented). Columns 3–11 list the results of hypothesis tests using single gene trees (mtDNA, 5 nuclear) as well as species trees reconstructed using 3 different methods. Species tree results are in the format: all loci/nuclear loci only. The ‘‘*townsendii* complex” consists of *U. canus*, *U. mollis*, and *U. townsendii*. See text for further details.

| **Hypothesis** | **Basis** | **mt** | ***VWF*** | ***BRCA*** | ***FGB*** | ***GBA*** | ***GHR*** | ***BEAST** | **STAR** | **MP-EST** |
|---|---|---|---|---|---|---|---|---|---|---|
| big-eared monophyly | K - Nadler 1966<br>K - Nadler 1984 | Y | N | N | N | N | N | Y/N | Y/N | Y/N |
| small-eared monophyly | K - Nadler 1966<br>mt - Harrison et al. 2003 | N | N | N | N | N | N | Y/Y | N/N | N/N |
| “*townsendii* complex” monophyly | M - Howell 1938<br>K - Nadler 1982<br>K - Nadler 1984 | N | N | N | N | N | N | Y/N | Y/N | N/N |
| *elegans* monophyly | M - Robinson and Hoffmann 1975<br>K - Nadler 1966<br>K - Nadler 1971 | N | Y | Y | N | Y | Y | - | - | - |
| *parryii* monophyly | M - Robinson and Hoffmann 1975<br>K - Nadler 1966<br>K - Nadler 1984 | N | Y | Y | Y | N | Y | - | - | - |
| *mollis* monophyly | M - Davis 1939<br>K - Rickart 1985 | N | N | N | N | N | N | - | - | - |

Table 2. Species comparisons from posterior predictive checking in JML recovered as significant at the $P = 0.1$ level. Listed are minimum empirical genetic distances found for samples within each significant species comparison and the exact significance level. The final column is a tabulation of the percentage of all possible subspecies combinations within each species comparison for which significance was found.

| **Species Comparison** | **Minimum Genetic Distance** | ***P*** | **Percentage of Subspecific Comparisons Significant** |
|---|---|---|---|
| *U. elegans – U. richardsonii* | 0.011 | 0.06 | 1/3 (33%) |
| *U. parryii – U. richardsonii* | 0.025 | 0.06 | 2/10 (20%) |
| *U. parryii – U. elegans* | 0.025 | 0.09 | 2/30 (6%) |
| *U. mollis – U. beldingi* | 0.066 | 0.04 | 6/9 (66%) |
| *U. canus – U. beldingi* | 0.071 | 0.07 | 3/3 (100%) |

FIGURES

Figure 1.

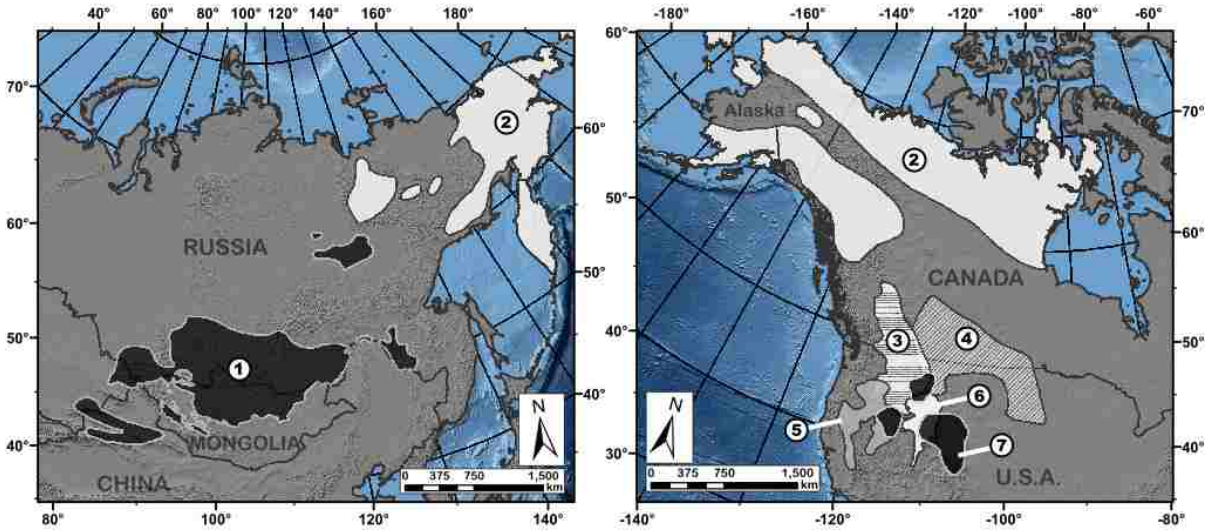


Figure 2.

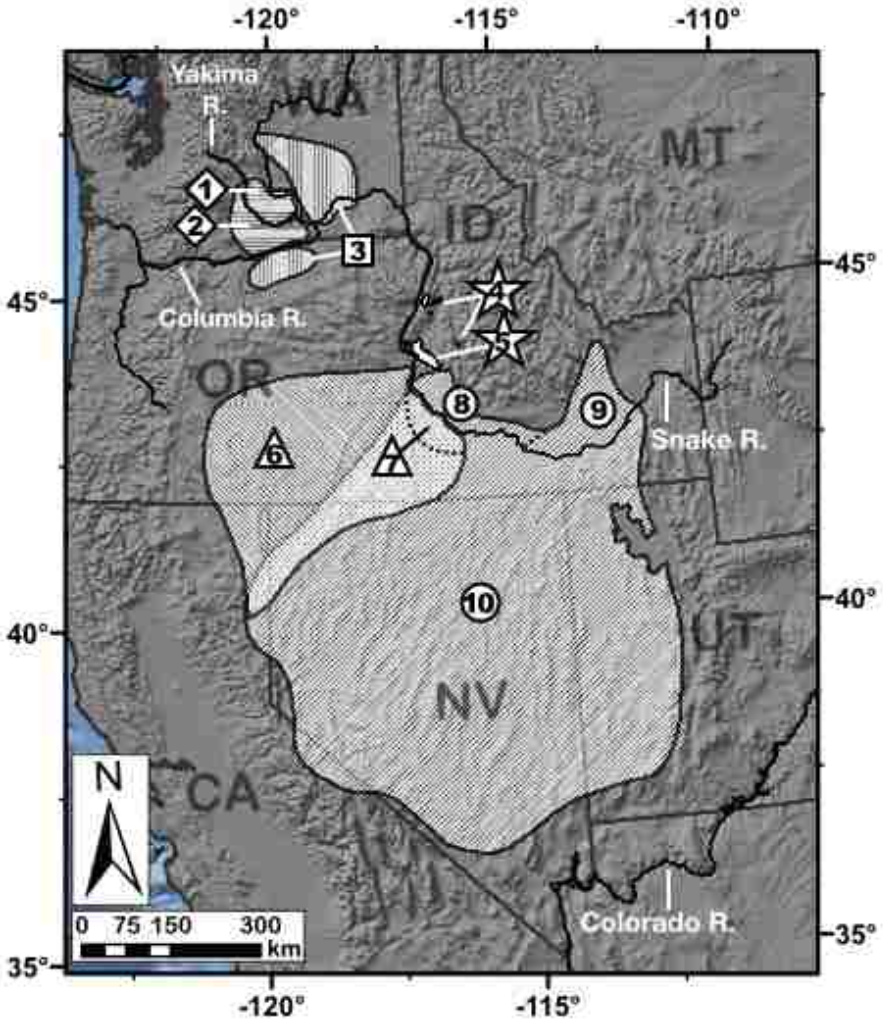

-120°
-115°
-110°
Yakima R.
WA
MT
ID
1
2
3
45°
Columbia R.
OR
8
9
Snake R.
6
7
10
40°
UT
NV
CA
N
0 75 150 300
km
Colorado R.
35°
-120°
-115°

Figure 3.

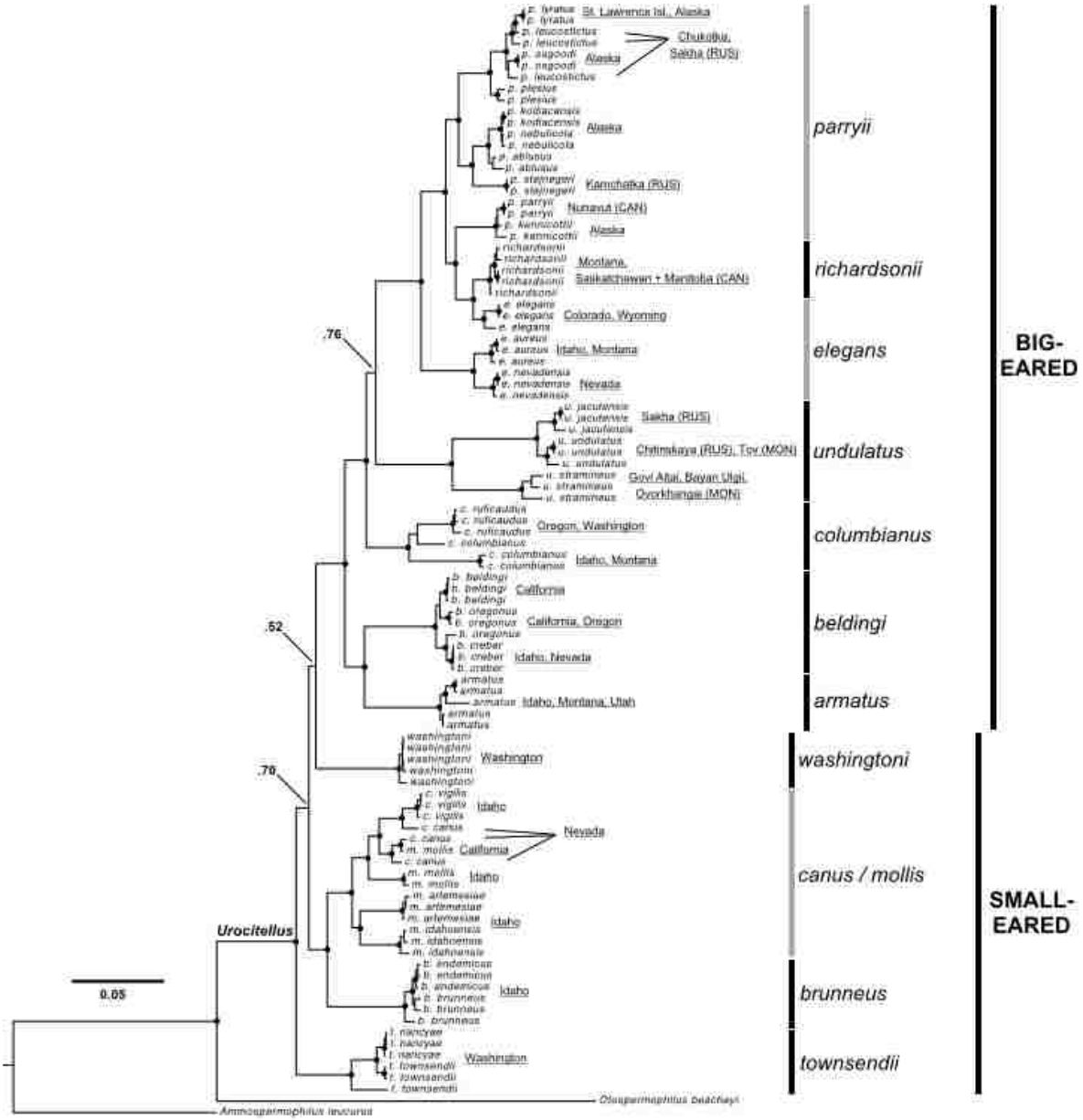

St. Lawrence Isl., Alaska
Chukotka, Sakha (RUS)
Alaska
Alaska
Kamchatka (RUS)
Nunavut (CAN)
Alaska
Montana, Saskatchewan + Manitoba (CAN)
Colorado, Wyoming
Idaho, Montana
Nevada
Sakha (RUS)
Chitinskaya (RUS), Tov (MON)
Govi Altai, Bayan Ulgii, Ovorkhangai (MON)
Oregon, Washington
Idaho, Montana
California
California, Oregon
Idaho, Nevada
Idaho, Montana, Utah
Washington
Idaho
Nevada
California
Idaho
Idaho
Idaho
Washington
parryii
richardsonii
elegans
undulatus
columbianus
beldingi
armatus
washingtoni
canus / mollis
brunneus
townsendii
BIG-EARED
SMALL-EARED
.76
.52
.70
Urocitellus
0.05
Ammospermophilus leucurus
Otospermophilus beecheyi

Figure 4.

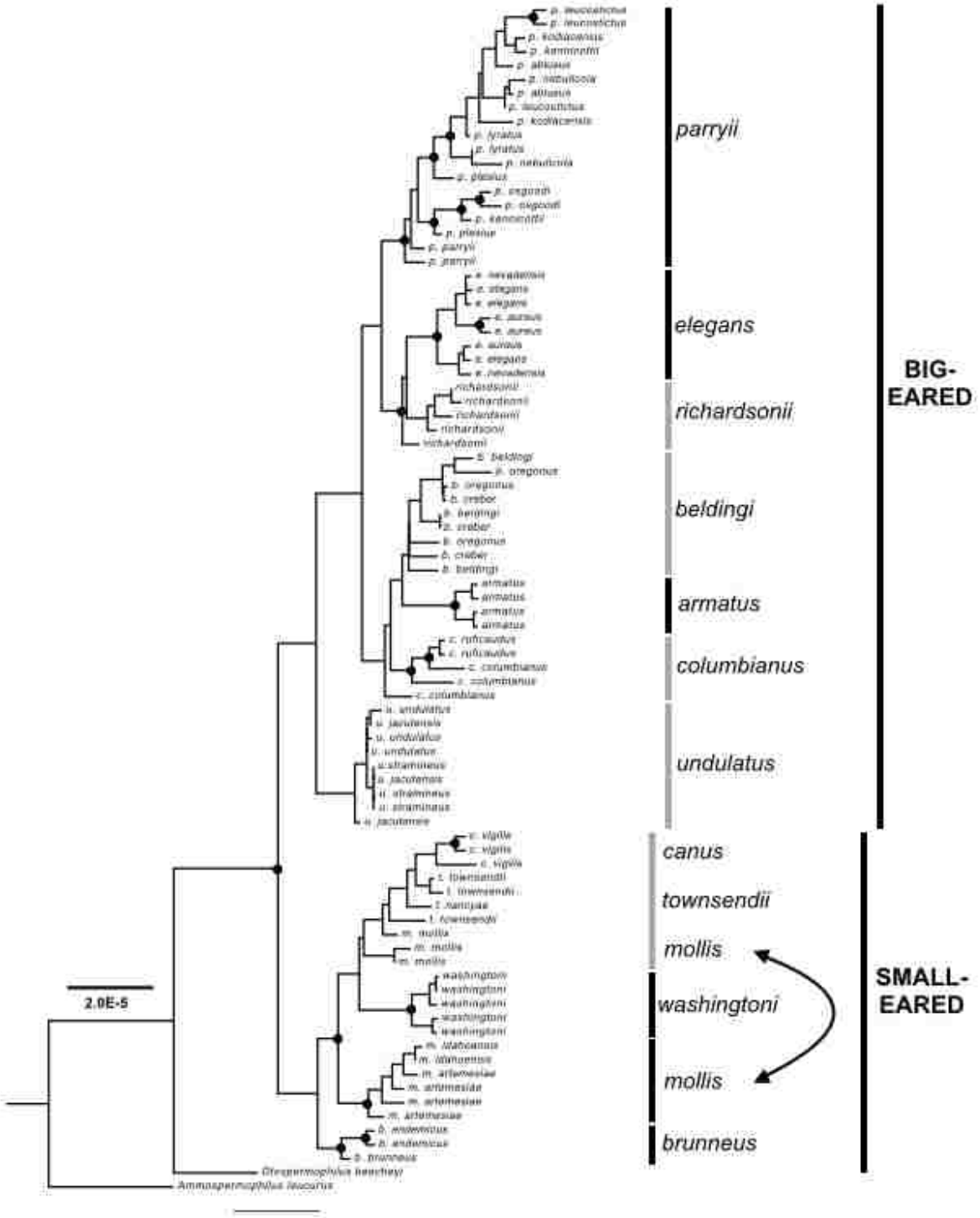

parryii
elegans
richardsonii
beldingi
armatus
columbianus
undulatus
canus
townsendii
mollis
washingtoni
mollis
brunneus
BIG-EARED
SMALL-EARED
2.0E-5
Dipodomys heermanni
Ammospermophilus leucurus

Figure 5.

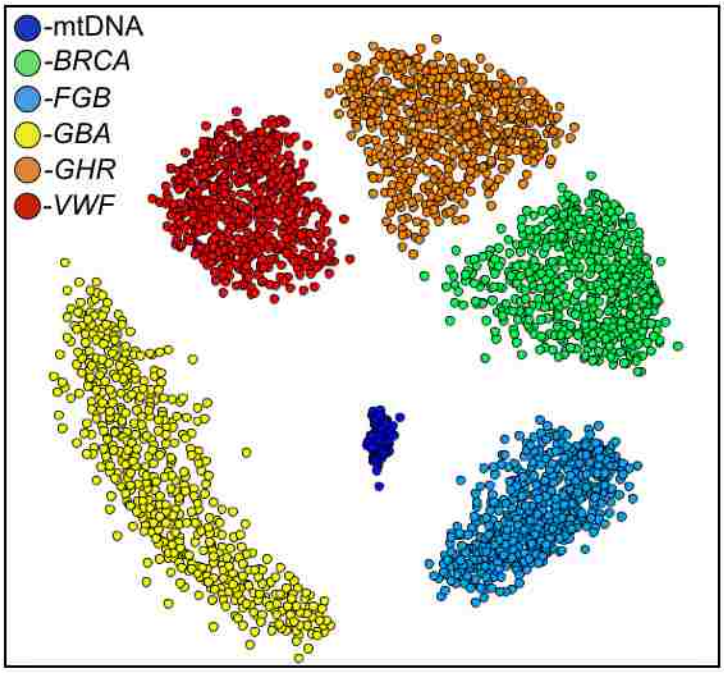


Figure 6.

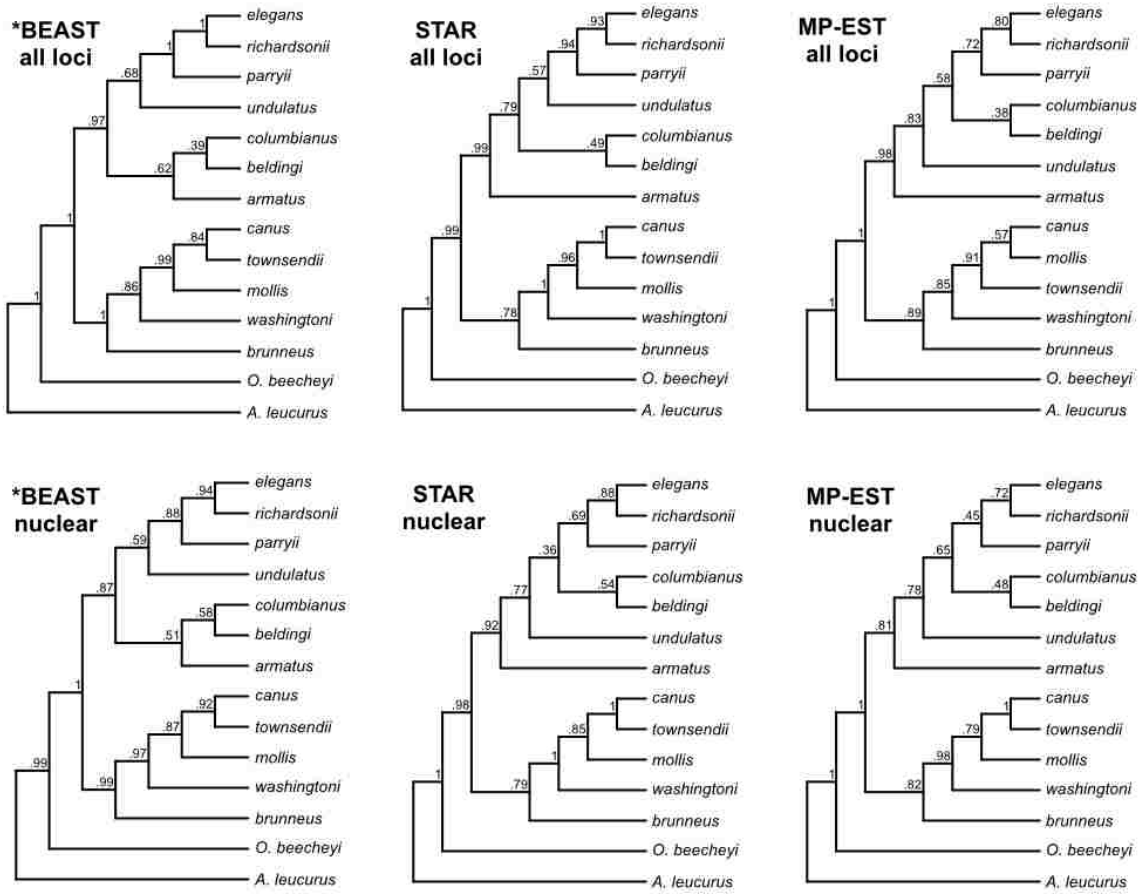

*BEAST all loci
elegans
richardsonii
parryii
undulatus
columbianus
beldingi
armatus
canus
townsendii
mollis
washingtoni
brunneus
O. beecheyi
A. leucurus
STAR all loci
elegans
richardsonii
parryii
undulatus
columbianus
beldingi
armatus
canus
townsendii
mollis
washingtoni
brunneus
O. beecheyi
A. leucurus
MP-EST all loci
elegans
richardsonii
parryii
columbianus
beldingi
undulatus
armatus
canus
mollis
townsendii
washingtoni
brunneus
O. beecheyi
A. leucurus
*BEAST nuclear
elegans
richardsonii
parryii
undulatus
columbianus
beldingi
armatus
canus
townsendii
mollis
washingtoni
brunneus
O. beecheyi
A. leucurus
STAR nuclear
elegans
richardsonii
parryii
columbianus
beldingi
undulatus
armatus
canus
townsendii
mollis
washingtoni
brunneus
O. beecheyi
A. leucurus
MP-EST nuclear
elegans
richardsonii
parryii
columbianus
beldingi
undulatus
armatus
canus
townsendii
mollis
washingtoni
brunneus
O. beecheyi
A. leucurus

Figure 7.

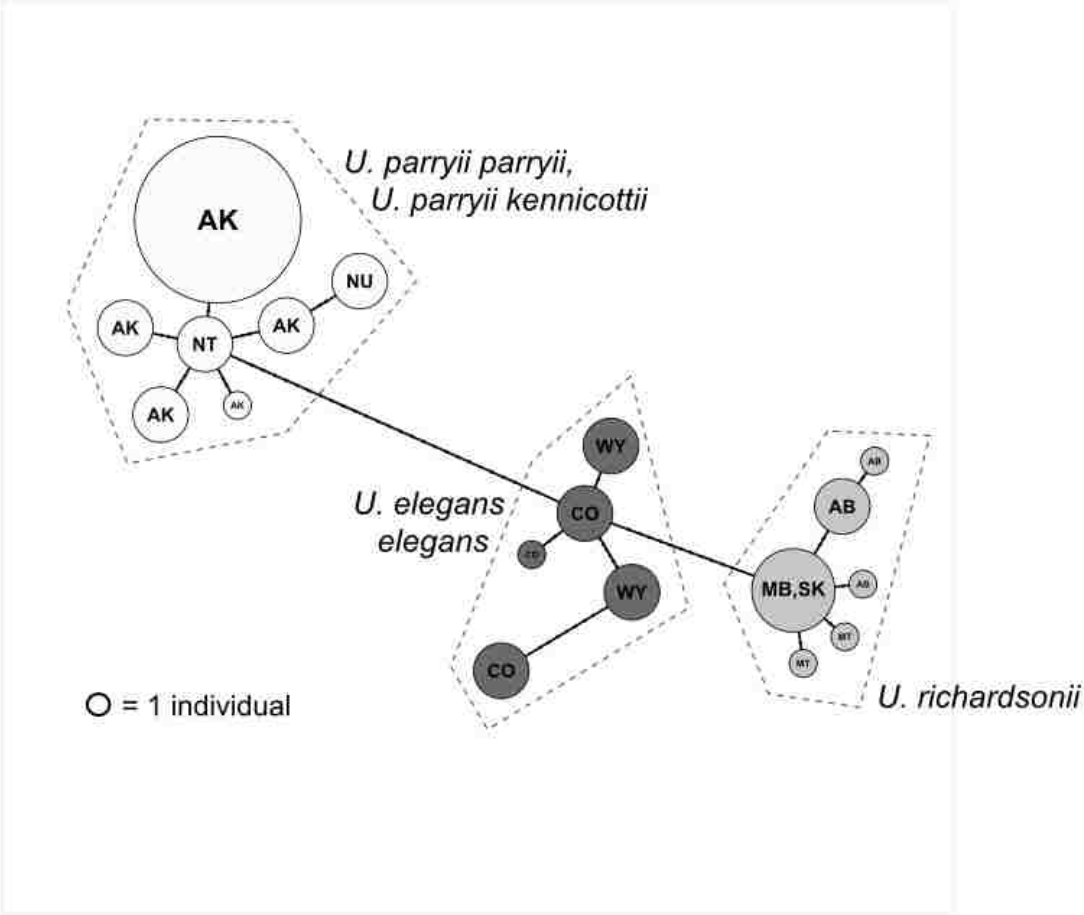
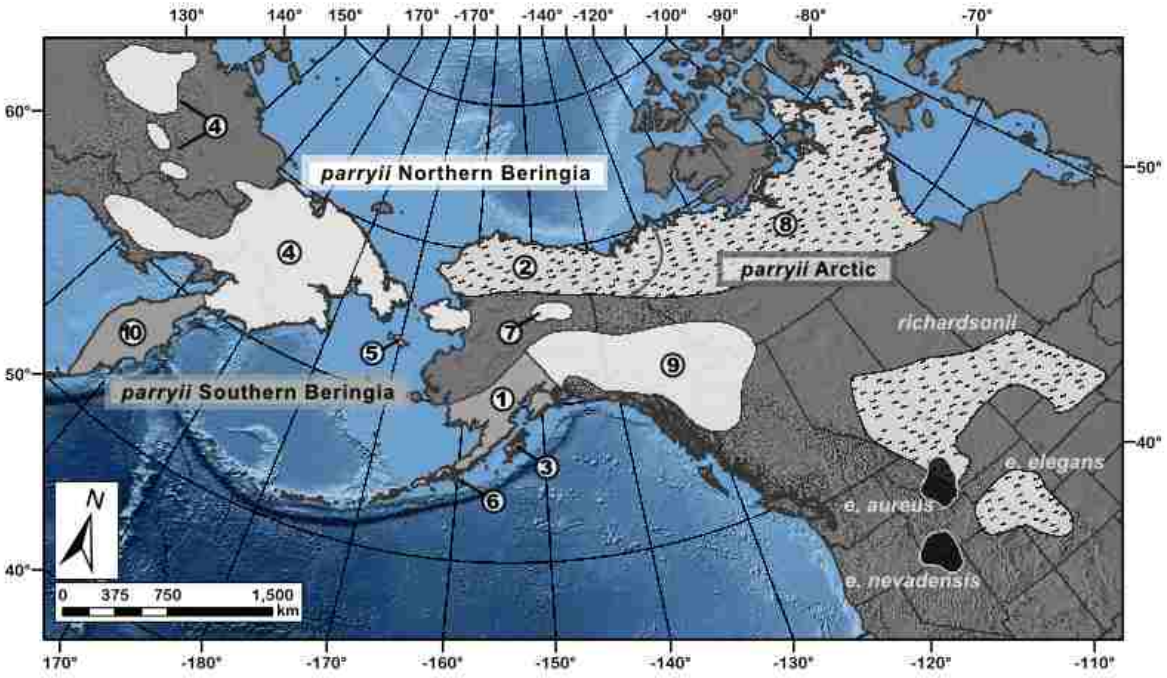

U. parryii parryii,
U. parryii kennicottii
AK
NU
AK
AK
NT
AK
AK
U. elegans
elegans
WY
CO
WY
CO
AB
MB,SK
U. richardsonii
O = 1 individual

Figure 8.


130°
140°
150°
170°
-170°
-140°
-120°
-100°
-90°
-80°
-70°
60°
50°
40°
170°
-180°
-170°
-160°
-150°
-140°
-130°
-120°
-110°
parryii Northern Beringia
parryii Arctic
parryii Southern Beringia
richardsonii
e. elegans
e. aureus
e. nevadensis
N
0 375 750 1,500
km
1
2
3
4
5
6
7
8
9
10

SUPPLEMENTARY FIGURE CAPTIONS

Supplementary Figure 1. Majority-rule consensus phylograms for each of 4 nuclear loci analyzed in the expanded character dataset (BRCA, FGB, GBA, GHR). Nodes receiving >0.95 posterior probability (PP) support are indicated with closed circles, and all are rooted with the outgroup *Ammospermophilus leucurus*. Museum catalog numbers or personal collector numbers are indicated to the right of taxon names, corresponding to specimens listed in Appendix 1.

Supplementary Figure 2. Haplotype network of the *parryii-richardsonii-elegans* (*PRE*) clade based on mtDNA cytochrome b (CYTB, 1140bp; sequences compiled from this and previous studies). All currently recognized subspecies are included. The haplotype network was computed in R. Hash marks represent single mutations, and sample size is proportional to the size of the circle (see legend). Default settings were used for all parameters, including a default distance parameter of uncorrected distances among sequences assuming an infinite-sites model.

SUPPLEMENTARY TABLES

Supplementary Table 1.

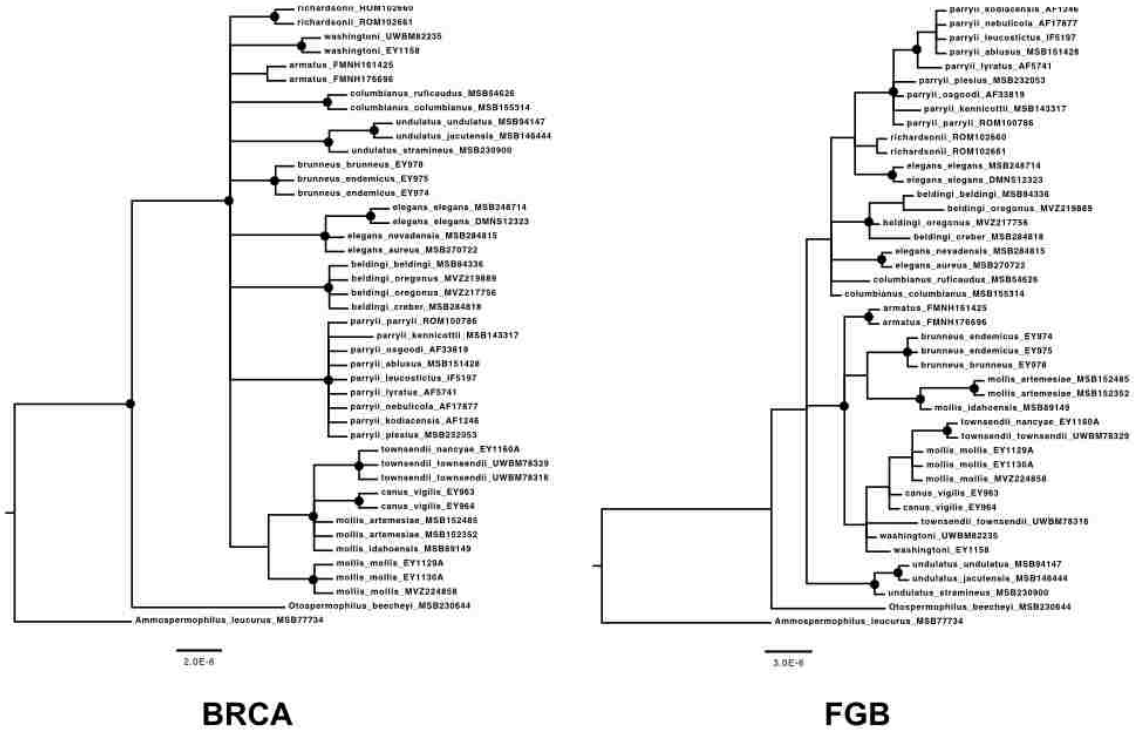
| | ***CTRL*** | ***CYTB*** | ***VWF*** | ***BRCA*** | ***FGB*** | ***GBA*** | ***GHR*** |
|---|---|---|---|---|---|---|---|
| Length (bp) | 522 | 1140 | 860 | 891 | 697 | 300 | 860 |
| S | 178 | 379 | 56 | 26 | 46 | 11 | 45 |
| rate | 1 | .54 (.07) | .08 (.01) | .03 (.01) | .06 (.01) | .04 (.01) | .06 (.02) |
| rate s.d. | .48 (.09) | .35 (.09) | .26 (.19) | .18 (.15) | .22 (.18) | .21 (.20) | .62 (.29) |
| alpha | .20 (.02) | .22 (.02) | .06 (.04) | .21 (.24) | .46 (.32) | .45 (.46) | .06 (.05) |

Supplementary Table 2.

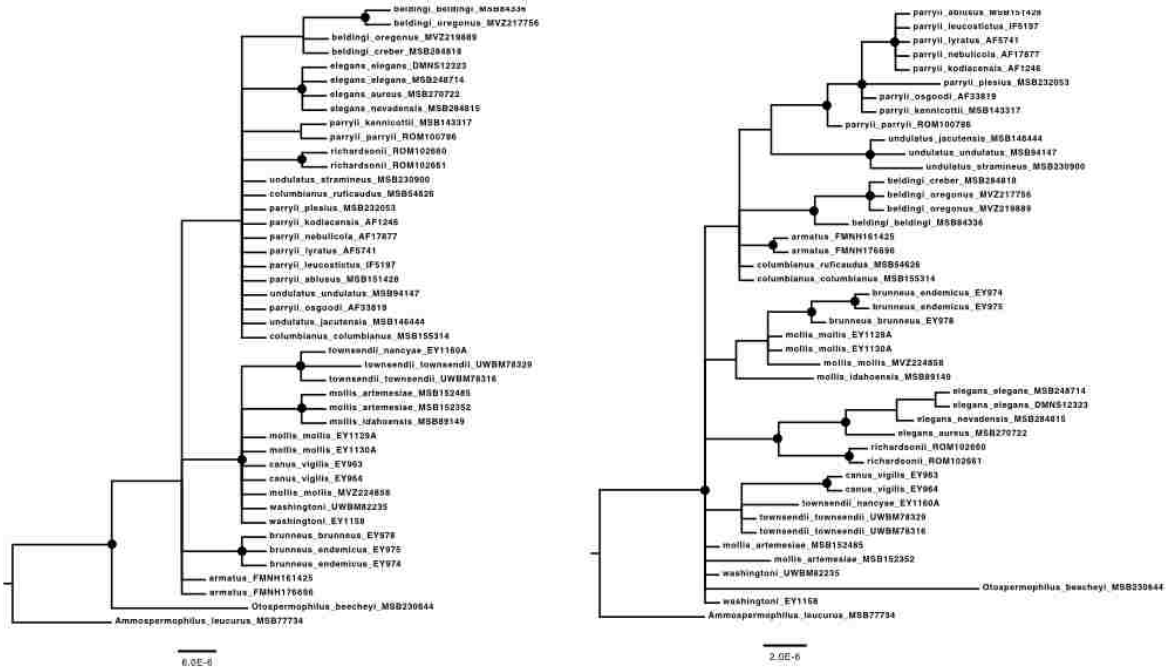
| | *parryii* | *elegans* | *richardsonii* |
|---|---|---|---|
| *parryii* | - | 28.5 (26-30) | 30.3 (27-33) |
| *elegans* | .025 (.023-.027) | - | 14.1 (11-19) |
| *richardsonii* | .027 (.024-.029) | .012 (.009-.017) | - |

SUPPLEMENTARY FIGURES

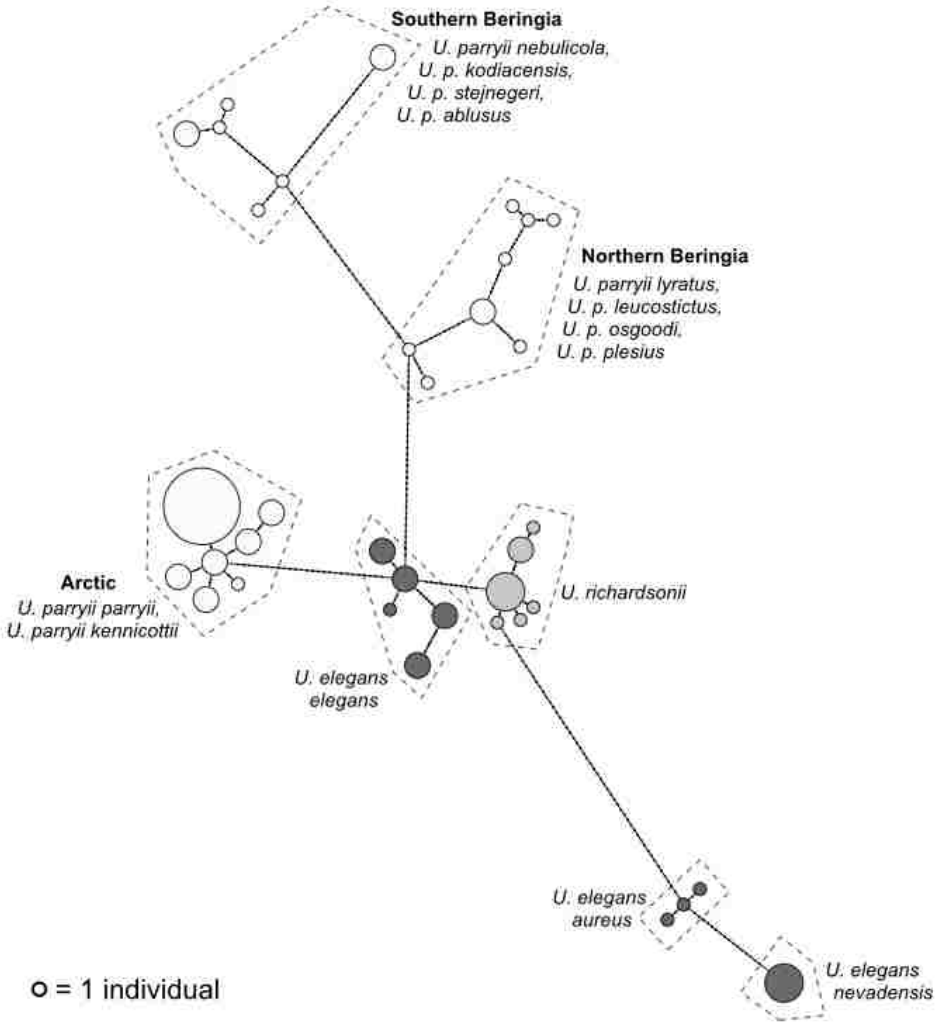
Supplementary Figure 1.



Supplementary Figure 1. (continued)



Supplementary Figure 2.



# CHAPTER 4

## Ecological Signatures of Diversification in Holarctic Ground Squirrels (*Urocitellus*)

AUTHORS
Bryan S. McLean[1], Kayce C. Bell[2], Julie M. Allen[3], Joseph A. Cook[1]

[1] Department of Biology and Museum of Southwestern Biology, 1 University of New Mexico, MSC03-2020, Albuquerque, NM 87131 USA
[2] Department of Invertebrate Zoology, Smithsonian Institution National Museum of Natural History, P.O. Box 37012, MRC 163, Washington, DC 20013-7012 USA
[3] Illinois Natural History Survey, University of Illinois, Urbana-Champaign, Il 61820 USA

ABSTRACT

Instances of rapid species proliferation are known in clades across the Tree of Life, but little consensus exists on the role of ecological factors in shaping this process. Lineages whose divergence has been ecologically mediated should display divergence in niche, modifications in traits used to exploit those niches, and evidence for divergent selection pressures with respect to close relatives. We tested for signatures of niche and trait differentiation in *Urocitellus* (12 spp.), a rapidly diversifying clade of ground squirrels distributed across the Holarctic region. We first resolved phylogeny using >3,750 ultraconserved element loci and then assessed variation in niche, body size, and body and cranial shape in this comparative context, leveraging a database of nearly 10,000 digitized biodiversity records. We found continuous but modest shifts in mean niche of *Urocitellus* species, but a strong relationship between niche differentiation and niche expansion; the latter pattern is difficult to reconcile with non-ecological forces and may be consistent with ongoing exploitation of ecological opportunity across North America and the broader Holarctic region. Furthermore, all phenotypic traits showed significant correlations with niche characteristics, suggesting the signature of divergent environments. However, paired analyses of specific body traits related to thermal conservation at inter- and intraspecific levels reveal complex patterns of ecophenotypic variation, precluding a direct link between microevolutionary processes of phenotypic variation and involved in the speciation process. Further study of the potential role of ecological opportunity in speciation are necessary for *Urocitellus* and other high-latitude organisms, and increasingly tractable using large, distributed biodiversity datasets such as that we employ here.

KEYWORDS

Allen's Rule, Bergmann's Rule, Ecological opportunity, Niche breadth, Phenotype-environment correlation, Speciation

## INTRODUCTION

Species richness varies greatly across the Tree of Life, a pattern that can be partially ascribed to differences in ability of clades to encounter and exploit new ecological opportunities (Sobel et al. 2009, Schluter 2016). Ecological opportunity is broadly defined as "a wealth of evolutionarily accessible resources little used by competing taxa" (Schluter 2000, pg. 69), and it can emerge via colonization of novel habitats, release from competition, appearance of new resources within the current range, or an evolved ability to access newly available resources (i.e., key innovations). Ecological opportunity might contribute to the speciation process indirectly (e.g., geographic isolation following broadening or shifting resource use) or directly (divergent selection against intermediate ecotypes). If such opportunities are available at broad geographic scales, they may foster rapid accumulation of lineages, thus forming a conceptual link between origin of species and patterns of diversity observed at higher levels. Unfortunately, the evidence for ecological opportunity as a causal mechanism in speciation is rare.

High latitude ecosystems are important venues for testing potential ecological bases of rapid diversification. This is due to 2 primary reasons. First, high latitude systems were significantly impacted by climatic and environmental fluctuations of the Quaternary (2.58mya to present), including growth and recession of massive continental glaciers. High-latitude environments were more heavily affected than lower latitudes (Flannery 2002, Sandel et al. 2011), with correspondingly greater impacts on the distribution, persistence, population sizes, and community structure of high-latitude species (Hewitt 2000, Lessa et al. 2003, Sandel et al. 2011, Hope et al. 2012, Weir et al. 2016). Many of these perturbations should have had evolutionary ramifications because they decreased competition (due to community disassembly, extirpation, or extinction during glacial periods), drove rapid expansion of existing niches (during glacial retreats and interglacial phases), and opened new resource bases (during either glacials or interglacials). Indeed, studies indicate that recent (~10mya to present) high latitude speciation rates are higher than at tropical latitudes in birds and mammals (Weir and Schluter 2007), a pattern that has been partially ascribed to the extreme gradients of ecological opportunity in recent high-latitude biomes (e.g., Schluter 2016).

A second reason that high latitudes may provide insight into the ecology of speciation is that they are characterized by reduced ecological dimensionality relative to lower-latitude systems. High latitudes generally display reduced biological productivity, greater seasonality, depressed species richness, and less complex trophic webs than lower latitude systems (Hawkins et al. 2003, Willig et al. 2003, Gillman et al. 2014), which may aid in identifying relevant axes of ecological divergence among closely related lineages. For example, stickleback lineages inhabiting postglacial Northern Hemisphere lakes have diverged repeatedly into just 2 forms, benthic and limnetic, with correspondingly straightforward adaptations in mouth shape, gill raker number, and other traits (Schluter and McPhail 1992). Additional examples of apparent or incipient speciation have identified in other high-latitude fish groups as well (Lu and Bernatchez 1999, Gislason et al. 1999, Olafsdottir et al. 2007, Harrod et al. 2010), but these examples are limited both taxonomically and geographically. Generalizations about the role of ecological

opportunity in high-latitude speciation and higher-level diversity patterns therefore remain difficult (e.g., Schluter 2016).

Holarctic ground squirrels (genus *Urocitellus*; 12 spp.) are a recent and rapid radiation distributed across 33° of latitude in North America and Asia (Helgen et al. 2009, McLean et al. 2016). They occur from the southern Great Basin to northernmost North America and Asia, and inhabit a range of biomes including desert, shrubland, xeric and montane grasslands, and Arctic tundra. Historically, diversification within *Urocitellus* has been explicitly or implicitly attributed to non-ecological forces, including climate change; fragmentation of habitats by forests or rivers, which limit dispersal of ground squirrels; isolation of lineages via intermittent inundation of the Bering land bridge; and instances of rapid chromosomal evolution (Nadler 1966, Nadler et al. 1974, Nadler and Hoffmann 1977, Harrison et al. 2003). Consequently, the potential contributions of ecology to the speciation process in *Urocitellus* have never been thoroughly investigated. Such an omission is at odds with existing hypotheses of marmotine ground squirrel radiation as a macroevolutionary response to ecological opportunity (Harrison et al. 2003, Goodwin 2008, Zelditch et al. 2015), and limits our ability to link broader patterns of lineage and phenotypic diversity to the population-level processes that have generated them.

In this paper, we used an integrative and phylogenetically-informed approach to investigate patterns and processes of niche and phenotype evolution in *Urocitellus.* To do this, we employed a novel phylogenomic hypothesis inferred from >3,500 ultraconserved element (UCE) loci, a database of thousands of digitized and georeferenced biodiversity records, and datasets describing multiple ecologically relevant traits (body size, body and cranial shape). We tested whether patterns of mean and breadth of realized niches were different than expected by chance, tested for the presence of a phenotype-environment correlation for all traits, and then compared adaptive mechanisms operating at the species level to those within species.

## METHODS

### *UCE Sequencing and Phylogenetic Inference*

We sampled tissues of 28 specimens from all 12 *Urocitellus* species for phylogenetic inference (Appendix 1). We also included 6 samples from 3 closely related outgroup species (*Poliocitellus franklinii*, *Spermophilus musicus*, *Spermophilus alashanicus*). Ethanol- or EDTA-preserved samples were first rinsed in STE buffer for 24 hours under refrigeration, with intermittent vortexing. Genomic DNA was extracted using a standard salt/EtOH protocol. DNA content was quantified using a Qubit (Life Technologies Corporation). Aliquots of 0.5-4 micrograms of genomic DNA were submitted to RapidGenomics, LLC (Gainesville, Florida) for library preparation using the UCE-5Kv1 probe set (http://www.mycroarray.com/mybaits/mybaits-UCEs.html). UCE libraries were sequenced using 2x100bp paired-end read technology on the Illumina HiSeq platform.

We performed QC and assembly of UCE reads as described in McLean et al. (in review); detailed methodological guidelines and code are also available on GitHub (https://github.com/juliema/aTRAM_UCE_pipeline). We aligned UCE assemblies in

MAFFT v7.2 using default settings (Katoh and Standley 2013), trimmed gapped regions using the *deleteGaps* function in the ips package (Heibl 2014) in R v3.2.3 (R Core Team 2015), and excluded a small number of anomalous UCE loci with high variability (>50% informative sites; $N = 12$). Based on our knowledge of typical UCE variability (McCormack et al. 2011, Gilbert et al. 2015), these likely represent assembly errors and their exclusion is a conservative step. Finally, we used scripts from the phyluce pipeline (Faircloth 2015) to obtain all remaining loci with ≥ 50% of taxa (i.e., ≥ 14 samples). We computed summary statistics for this dataset using additional scripts from the phyluce pipeline. The final dataset used for analysis consisted of 3,531 UCE loci (2,146,015bp total).

We inferred phylogeny using a maximum likelihood (ML) concatenated approach in RAxML v8.2.8 (Stamatakis 2014). We generated a matrix for the full UCE dataset using the Python package *amas.py* (Borowiec 2016) and used this as input for RAxML analysis using the default GTR substitution model with gamma-distributed rate heterogeneity (-m GTRGAMMA option) and 25 categories, assessing support using a rapid bootstrap analysis with 100 replicates (-f a option). We did not perform a model-testing approach to estimating substitution parameters due to the generally low variability of individual UCE loci in *Urocitellus*. All RAxML analyses were conducted on the CIPRES Science Gateway (www.phylo.org; Miller et al. 2010). We did not employ a coalescent-based species tree approach for this dataset because previous analyses of marmotine phylogeny have identified *Urocitellus* as a zone of particularly low phylogenetic signal in which such methods perform inconsistently (McLean et al. in review), a result we believe is due to the imprecision in many UCE gene trees. To identify potential biases in results of our concatenated analysis, we compared our results to previous mitochondrial and multilocus analyses (McLean et al. 2016).

*Environmental Data*

We accessed 9,767 records of *Urocitellus* from 40 biodiversity collections distributed through VertNet (http://vertnet.org/; accessed February-March 2017). Geographic coordinates were associated with the majority of records. We further expanded the spatial scope of the data by georeferencing approximately 750 records that previously lacked coordinates, but nevertheless contained associated locality data. Georeferencing was performed in the GEOLocate web application (http://www.museum.tulane.edu/geolocate/) as well as the USGS Geographic Names Information System (for historic/ambiguous place names only; https://geonames.usgs.gov/domestic/). Data records were modified to reflect current taxonomy (Helgen et al. 2009), quality checked for spatial accuracy by comparison with existing range limits, and questionable records removed. For the purposes of niche analyses, we also excluded all specimen records with duplicate coordinates ($N = 7,390$), which would bias estimates of both niche mean and breadth. The final database included 2,808 records and is taxonomically representative (mean 254, range 65-738 records, calculated for all species except *U. brunneus*), temporally expansive (spanning 150 years), and spatially heterogeneous (records reasonably well distributed across species ranges). Low numbers of records for the Idaho ground squirrel (*U. brunneus*; $N = 25$) in the final dataset reflect the highly restricted distribution of this taxon as well as its poor

representation in museum collections. However, exploratory analyses using similarly low numbers of specimens drawn at random for other species did not significantly affect our major results and conclusions.

Niche mean and breadth of *Urocitellus* species was quantified along 3 relevant axes of environmental variation (climate, precipitation, and biophysical characteristics). We obtained standard bioclimatic variables from the WorldClim global climate database v1.4 (http://www.worldclim.org/bioclim; Hijmans et al. 2005), which contains 19 temperature and precipitation metrics widely recognized as biologically relevant and commonly employed in species distribution modeling. Bioclimatic variables represent global averages for the period 1960-1990 and were accessed at 0.4 degree (2.5 minute) resolution. Next, we obtained an optical metric of vegetation greenness called enhanced vegetation index (EVI2; Jiang et al. 2008), derived from remote sensing data. EVI2 has been shown to be more robust than the NDVI and EVI metrics, which are more susceptible to noise and uncertainties driven by different atmospheric properties. EVI2 is derived directly from NASA Moderate Resolution Imaging Spectroradiometer (MODIS) data and available as part of the MEaSUREs project (https://lpdaac.usgs.gov/dataset_discovery/measures/measures_products_table/vip30_v004). We employed EVI2 data in 2 forms, similar to the approach of McCormack et al. (2009). The first was mean EVI2 (a measure of overall greenness and productivity) and the second was the annual standard deviation of EVI2, which captures the seasonality of productivity (computed as the mean of the annual standard deviations). Because snow cover impacts EVI2 quantification and *Urocitellus* species are inactive in extreme winter months due to hibernation, we excluded EVI2 data for Jan.Feb. and Nov.-Dec.; both EVI2 metrics therefore represent 8-month datasets (Table 2). EVI2 data were accessed as monthly averages for the period 2001-2014 at 0.05 degree resolution. Both EVI2 and bioclimatic variables are derived from long-term observations (≥14 years) and therefore robust to interannual-scale climatic and vegetation anomalies.

We used the R package raster (Hijmans et al. 2016) to extract environmental data for all georeferenced specimen records. Over half of the records had associated spatial uncertainties, which we incorporated by using verbatim values as buffers, extracting values for all cells falling within each buffer, and computing average values. For records without associated uncertainties ($N = 1161$), we assumed a 5km uncertainty radius, which is significantly greater than the median uncertainty for specimens with spatial uncertainty values, but not unreasonably large with respect to the resolution of the environmental data. To account for the spatial autocorrelation among bioclimatic variables, and to improve interpretability of niche axes, we computed pairwise Pearson's correlations among all extracted bioclimatic variables and excluded those producing correlation values > 0.8 ($N = 6$). The final dataset contained 15 total environmental variables (7 temperature, 6 precipitation, and 2 biophysical metrics).

*Niche Evolution*

Axes of maximum among-species niche variation were explored using linear discriminant analysis (LDA) of the full environmental dataset in R, using species as grouping factors. LDA is most appropriate for this goal as it maximizes among-group variation and minimizes within-group variation; the latter is not identically distributed

among *Urocitellus* species because it is higher in those with large geographic ranges (e.g., *U. parryii*) and lower in those with restricted ranges (*U. brunneus, U. townsendii*). Transformed niche variables were plotted and visualized using convex hulls. Jackknife reclassifications were performed for all specimens to assess niche distinctiveness based on the environmental data.

Patterns of niche evolution (both niche mean and niche breadth) were examined in an explicitly phylogenetic context using species means for all environmental variables and subjecting them to principal components analysis (PCA). All variables were centered and scaled to unit variance prior to PCA. We summarized niche variation within species by computing species standard deviations on all environmental variables as metrics for species niche breadths. However, to reduce the dimensionality of niche breadth data, we also performed a PCA for species values. Exploratory analyses revealed this procedure provides an acceptable summary of the magnitude and form of variation in niche breadth among species across the 15 raw environmental variables. We retained PC axes 1-5 from both PCAs as they each explained >3% and, cumulatively, >95% of the total variation in each analysis.

We assessed the extent of phylogenetic signal in niche mean and niche breadth by calculating the multivariate implementation of Blomberg's *K* (Adams 2014), a metric implemented in the geomorph v3.0.3 morphometrics package in R (Adams et al. 2016). Significance of phylogenetic signal was assessed using 9,999 randomizations of the original PC data among tips of the phylogeny. We evaluated the evolutionary relationship between niche mean and niche breadth by performing a Mantel test on Euclidean distance matrices computed from the niche PCs described above. Because the major axes from these different analyses may be combinations of different groups of raw variables, we also performed Mantel tests on distance matrices calculated from the raw data those PCAs are based on (species means and standard deviations on all environmental variables). Statistical significance of both Mantel tests was assessed using the phylogenetic permutation method of Lapointe and Garland (2001) implemented in the R language (Harmon and Glor 2010). The latter method incorporates phylogenetic relatedness into matrix permutations so that the extent of character permutation is proportional to the phylogenetic distance among tips. Because Mantel tests suffer from low statistical power, we used a second distance-based approach to compare evolution in niche mean and breadth. The approach is based on Stayton (2015) and uses multivariate trait data and ancestral state reconstruction to calculate the magnitude of trait evolution (change in Euclidean distance) for all ancestor-descendent pairs on a phylogeny. We performed this analysis on PC axes as well as raw data describing niche mean and breadth, as above for Mantel tests, and performed a linear regression of the data.

*Body Shape*

Drawing from the raw biodiversity data that were used to construct the niche database, we assembled a comprehensive dataset of external body shape data for *Urocitellus* that included 3 standard mammalian external measurements: tail, hindfoot, and ear lengths. We extracted values for head-body length (a proxy for body size) as well, which was calculated by subtracting tail length from total length. Both procedures were enabled by recent updates to the VertNet platform that encode presence/absence of

body size information (Guralnick et al. in review). We did not use body weight (a fifth standard measurement for mammals) as a character in either dataset as it is highly labile within species, even varying significantly within years, and is also highly correlated with head-body length (although see Meiri and Dayon 2003 for comparison of the 2 traits). The body size and body shape datasets together comprise a small number of measurements, but are relevant for understanding adaptation to different environments and latitudes and they correspond to well-known ecogeographic rules (Bergmann's Rule and Allen's Rule, respectively). We did not remove spatially unique specimen records because, unlike environmental data, data from multiple individuals within sites is still informative with respect to overall species phenotypic means and variation.

We curated the body shape database by removing all specimens with missing data as well as specimens labeled as juveniles or immature. However, because only a minority of specimens contained information on age class, we also chose to exclude all specimens with head-body lengths falling below the $10^{th}$ quantile in each species. Head-body length is the most appropriate metric for identifying non-adults because other metrics have high positive allometries and thus reach adult proportions earlier (Kiell and Miller 1978, Koeppl and Hoffmann 1981). Our use of a single cutoff assumes sampling efforts have targeted all age classes equally through time, as well as that ontogenetic allometries of head-body length are homogeneous across species. There is empirical evidence for this latter assumption in some, but not all, *Urocitellus* species (Koeppl and Hoffmann 1981), and heterogeneity likely exists due to varying phenology and life history demands (length of active season, litter size, environmental productivity; e.g., Kiell and Miller 1978). We cannot directly assess either of the above possibilities (equal representation of age classes, similarity in ontogenetic trajectories), but exploratory analyses using stricter cutoffs did not significantly affect our results, suggesting our data and analyses are likely robust to these potential inconsistencies.

Finally, we used functions in the R package rrcovHD (Todorov 2016) to identify and extract outliers for each species based on Euclidean distances calculated in a preliminary PCA. The final body shape dataset included 3,105 specimen records with measurements (mean 280 and range 49-1010; calculated for all species except *U. brunneus*,). Twenty-one adult specimens were available with complete measurement data for *U. brunneus*. All measurements were log-transformed and species means were computed for each trait. We also compiled a size-corrected shape dataset by regressing species means for the latter 3 measurements (tail, hindfoot, ear) on mean head-body length and extracting residual values for each trait. Raw and size-corrected datasets were subjected to PCA and the first 3 PC axes used for further analyses (note the size-corrected PCA had a total of 3 axes).

*Cranial Shape*

We collected 2D geometric morphometric data from crania of 454 museum specimens representing all currently recognized species and subspecies of *Urocitellus* (Appendix 1). We included a minimum of 20 adult specimens of each species, with the exception of *U. undulatus*, which was represented by 11 specimens. Specimens were identified as adults using the criterion of complete eruption and development of upper premolars 3 and 4 ($P^3$ and $P^4$, respectively, as in McLean et al. in prep).

All crania were photographed in ventral aspect using a mounted Nikon D90 DSLR camera fitted with a Nikon AF-S 60mm macro autofocus lens and a standardized imaging procedure. To ensure precision of landmarks, we used Helicon Remote software (http://www.heliconsoft.com/) to obtain 15-25 high-resolution images throughout the depth of field of each specimen (depending on cranial proportions). These images were then stacked using Helicon Focus software. Twenty-four 2D landmarks (Fig. 3a) were digitized on final images using the software tpsDig v2 (Rohlf 2006). The default procedure was to digitize landmarks on the left side of the ventral cranium, but we also used only the right side in a minority of cases due to damaged or incomplete crania. For the purposes of this study, which is focused only on the symmetric component of shape variation, we assume any effects of fluctuating asymmetry to be small relative to other factors examined.

Landmark data were subjected to Procrustes superimposition in the R package geomorph. We subjected all specimens to LDA using functions in the R package Morpho (Schlager and Jefferis 2016) and plotted these to visualize major axes of variation among species. We also calculated species mean shapes from raw Procrustes configurations and performed a PCA on these using the plotTangentSpace function in geomorph. To isolate the non-allometric component of ventral cranial shape for use in further analyses, we performed a multivariate regression of species mean shapes on the logarithm of head-body length (calculated from the body size dataset above) using the geomorph function procD.lm. From this, we extracted residuals and performed PCA as above. For both raw and size-corrected PCAs, we retained PC axes 1-5 for further analyses as they explained 90 and 89% of the total cranial shape variation, respectively.

*Phenotype - Environment Correlations and the Adaptive Basis of Trait Evolution*

We assessed whether observed phenotypic variation of all traits among species was consistent with adaptation to divergent niches. First, relationships between niche data and 5 phenotypic datasets (head-body length, raw and size-free body shape, raw and size-free cranial shape) were assessed using Mantel tests with 9,999 phylogenetic permutations as previously described. Inputs for tests were Euclidean distance matrices calculated from species mean values in each respective trait space, using the same PC axes previously extracted (niche PC1-5, body shape PC1-3, cranial shape PC 1-5).

Second, we tested whether adaptive signals revealed by Mantel tests were attributable specifically to ecogeographic rules (Bergmann's and Allen's) that predict larger body size and smaller appendage size in colder climates, respectively. Although both rules are typically formulated for intraspecies variation, they may also describe thermal adaptive mechanisms manifested at other scales as well (Millien et al. 2006), such as among related species distributed across major latitudinal gradients. To do this, we first extracted mean annual temperatures (MAT) for all georeferenced localities in the body shape dataset and calculated species mean values. We then calculated phylogenetic independent contrasts for species mean head-body lengths (our proxy for size) and relative tail, hindfoot, and ear lengths (proxies for appendage size), and regressed each on contrasts for MAT. Contrasts were calculated using the pic function in the R package ape (Paradis et al. 2016).

Finally, to begin to link macroevolutionary patterns in phenotypic evolution to potential population-level adaptive mechanisms, we tested for conformation to Bergmann's and Allen's Rules within each species using the body shape database by performing linear regressions of the body measurement data described above on MAT in a species-wise fashion. We made no attempt to exclude specimens from identical localities, or those falling within similar thermal ranges, as these are still informative for within-species patterns. All trait data used in this analysis and the among-species analysis were raw, untransformed measurements. To visualize patterns of support for ecogeographic rules, we summarized correlation coefficients for all species regressions as barplots on a trait-wise basis.

## RESULTS

### *Phylogenetic Inference*

Mean length of assembled UCEs prior to filtering was 563bp (range 207-2163bp), similar to that reported from other UCE-based studies of vertebrates (Hawkins et al. 2016, Streicher and Weins 2016) including analyses within Marmotini (McLean et al. in review). Summary statistics indicated that our dataset was characterized by relatively low information content on a per-locus basis; median percentage of phylogenetically informative sites per locus was less than 1.1% (~6 sites per locus) when excluding outgroups (Table 1). This finding is generally consistent with the relatively modest mutation rates documented in UCE cores and flanking regions (McCormack et al. 2012) but is likely exacerbated by rapid diversification in *Urocitellus*.

Phylogenetic analyses in RAxML resulted in a robustly supported tree (Fig. 1). With respect to species relationships, the topology is largely congruent with the multilocus analysis of McLean et al. (2016). A single exception is placement of the Columbian ground squirrel (*U. columbianus*), a species which we recover as sister to the *parryii-undulatus-richardsonii-elegans* clade, but which previously was recovered with low support as sister to *U. beldingi* by McLean et al. (2016). Placement of *U. columbianus* in the UCE-based topology is more consistent with mitochondrial relationships (McLean et al. 2016) as well as electrophoretic, biogeographic, and host-parasite data (Nadler 1966, Nadler and Hoffmann 1977, Nadler et al. 1982). We thus accept the concatenated topology as the current best estimate of *Urocitellus* phylogeny.

### *Niche Evolution*

LDA of the environmental dataset revealed three broad species groups in niche space: *U. parryii + U undulatus*, higher-latitude big-eared species (*U. columbianus, U. richardsonii, U. elegans*), and the remainder of *Urocitellus* (Fig. 3a,b). However, broad overlap existed for most species except for *U. parryii + U undulatus*, reflecting the similarity in general climatic niche that characterizes most species occurring across western North America. Nevertheless, jackknife reclassifications in niche space achieved >70% overall accuracy for 8 of the 12 species (exceptions were *U. armatus*, 60.6%; *U. canus*, 52.3%, *U. columbianus*, 69%, and *U. townsendii*, 69.7%).

PCA of species means for environmental variables summarized major niche differences among species. In this niche space, PC1 is largely a thermal niche axis, with bioclimatic variables related to temperature loading heavily (Table 2). PC1 thus separates species roughly based on latitudinal range; small-eared species and *U. beldingi* load most positively and the high-latitude *U. parryii*, *U. undulatus*, and *U. richardsonii* load most negatively. Conversely, PC2 is a precipitation axis, separating species found in more equitable environments (most big-eared species and *U. brunneus*) from those in drier environments such as the Great Basin (small-eared clade excepting *U. brunneus*) and central Asia (*U. undulatus*). Along the first 2 PC axes, the variables derived from vegetation indices (mean and annual standard deviation of EVI2) loaded most similarly to bioclimatic variables describing precipitation of the coldest quarter and warmest quarter, respectively. Mean niche (quantified by PC1-5) displayed significant phylogenetic signal ($K_{mult}$ = 1.12, $P$ < 0.001), consistent with our knowledge of the general biology of *Urocitellus*.

In the PCA of species standard deviations on all environmental variables, PC1 (53% of total variation) largely separated the species with the broadest, most environmentally heterogeneous range (*U. parryii*) from all others. However, higher axes (PC2-3, 18 and 13% of total variation, respectively) tend to separate remaining species based on meaningful variations in niche breadth including those whose ranges have high variation in temperature (*U. undulatus*, *U. richardsonii*) and high variation in precipitation metrics (*U. beldingi*). The metric of vegetation seasonality that we derived (EVI2_stdev) loaded positively on PC2 and captured important variation in niche breadth, with high latitude species displaying highest values (*U. parryii*, *U. undulatus*, *U. richardsonii*) and low-latitude, arid-dwelling species displaying lowest values (small-eared clade minus *U. brunneus*). Niche breadth as summarized by PCA displayed significant phylogenetic signal ($K_{mult}$ = 1.06, $P$ = 0.001); however, when PC1 was excluded from the niche matrix, we recovered no phylogenetic signal in niche breadth ($K_{mult}$ = 0.91, $P$ = 0.19). Use of a more direct measure of niche breadth not reliant on further dimensionality reduction (total niche volume in original PCA space) likewise displayed a lack of phylogenetic signal ($K$ = 0.98, $P$ = 0.11).

A Mantel test with phylogenetic permutation recovered a highly significant correlation between niche mean and niche breadth (Z = 3551, r = 0.53, $P$ = 0.002; Table 3), suggesting these 2 properties of the climatic niche are related across *Urocitellus* phylogeny. A separate Mantel test using just the raw species means and standard deviations in the same axes of niche space was also significant (r = 0.50, $P$ = 0.001). When an alternate metric of association was computed (change in Euclidean distances between ancestor-descendent pairs), the magnitude of per-branch shifts in niche mean and niche breadth were highly correlated across phylogeny ($r^2$ = 0.75, $P$ << 0.001).

*Trait Variation and Covariation*

Body size displayed significant phylogenetic signal in *Urocitellus* and was less variable than expected under Brownian motion ($K$ = 2.11, $P$ < 0.001). This is consistent with trenchant differences in body size between big-eared and small-eared clades, but less variation within clades, especially within the small-eared clade (Fig. 4a). Raw body shape and cranial shape also displayed significant phylogenetic signal ($K_{mult}$ = 1.96 and

1.11, respectively, $P < 0.001$ for both); however, both displayed significant allometry as well ($r^2 = 0.49$, $P = 0.001$ for body shape and $r^2 = 0.39$, $P = 0.001$ for crania). When corrected for size-related effects, body and cranial shapes did not display phylogenetic signal ($K_{mult} = 0.90$ and 0.87, $P = 0.26$ and 0.45 respectively), suggesting they may vary in ways possibly related to ecological adaptation.

The LDA of raw cranial data largely separated groups according to size and phylogeny, with clusters corresponding to a) *U.parryii*/*U. undulatus*, b) the remainder of big-eared species, and c) small-eared species. However, specimens could be assigned with >90% accuracy in all species except *U. townsendii* and *U. washingtoni* (83 and 85% accuracy for those taxa, respectively). The pattern of cranial morphospace occupation was similar to PCA for that same trait, which were both similar to PCA for body shape, as expected from their shared allometries. Size-corrected body shape separated species based on relative tail and ear lengths; hindfoot loaded weakly on all PC axes, however. Size-corrected cranial shape variation was less interpretable from PCA. Both analyses tended to separate *U.parryii* + *U. undulatus* on PC1, and to strongly separate *U. brunneus* from the remainder of small-eared species.

Mantel tests for niche and phenotype correlations were all statistically significant (Table 3, $P < 0.01$), suggesting that the traits sampled here are adaptive in different climatic niches. Significant correlations between size-corrected phenotypic traits and niche data were particularly striking, and indicate shape-related environmental adaptations in body shape and cranial shape (Fig. 4b,c). A hypothesis of adaptive change is consistent with the lack of phylogenetic signal observed in both traits when controlling for size. Moreover, we observed similar patterns of shape change in these 2 traits, as previously suggested by similar patterns of morphospace occupation; a *post hoc* Mantel test recovered high correlation between size-corrected body and cranial shapes ($r = 0.63$, $P = 0$). Significance of that association was confirmed by performing regressions of independent contrasts for the major axes of variation in each trait (PC1 contrasts $r^2 = 0.58$, $P = 0.003$; PC2 contrasts $r^2 = 0.38$, $P = 0.02$).

*Ecophenotypic Variation Among and Within Species*

Our tests of whether patterns of phenotypic change at the interspecific level correspond with commonly observed thermal adaptations yielded mixed results. There was strong support for a Bergmann's-like pattern, with species mean body size significantly correlated with MAT (regression of independent contrasts; $r^2 = 0.54$, $P = 0.005$; Fig. 5). Relative tail length was marginally insignificant ($r^2 = 0.18$, $P = 0.10$), as were relative hindfoot and ear lengths (hindfoot $r^2 = 0.09$, $P = 0.18$; ear $r^2 = -0.04$, $P = 0.47$). Exploration of trends within species yielded conflicting results for the generality of ecogeographic rules, and thus whether they are involved in ecologically-mediated speciation. Similar numbers of species were found to exhibit an inverse Bergmann's Rule (*U. columbianus*, *U. elegans*, *U. mollis*) as a traditional Bergmann's pattern (*U. undulatus, U. richardsonii, U. beldingi, U. parryii*). Similarly, even numbers of species displayed longer relative tail and hindfoot lengths in response to increasing MAT as shorter tail and hindfoot lengths (Fig. 5e-h). Relative ear lengths were more consistent across species in showing a negative slope, counter to predictions of Allen's Rule. This

latter trend in ear length could reflect constraints on how large ears can be due to the physical constraints of burrowing (Nevo 1979; Fig. 5h).

## DISCUSSION

### *Patterns of Niche Evolution*

Population divergence driven by adaptation to different niches is the core mechanism of ecological speciation. Exploring patterns of niche variation is an important first step towards identifying whether specific diversification events might have an ecological basis (Peterson et al. 1999, Kozak and Wiens 2006, McCormack et al. 2009). We found measureable niche divergence among most *Urocitellus* species along axes of temperature, precipitation, and local vegetation; classification accuracy of individuals based on environmental variables was 70% or greater for 10 of 12 species. These results suggest continuous expansion to new regions of niche space and are inconsistent with strict phylogenetic niche conservatism (Peterson et al. 1999, Wiens and Graham 2005), whereby niches of relatives are more similar than expected by chance (Pyron et al. 2015). However, they are still indistinguishable from a Brownian motion model, suggesting ancestor-descendent pairs usually display some level of niche similarity and that major niche shifts have been rare. This is consistent with our understanding of the natural history and biology of the genus (Howell 1938, Durrant and Hansen 1954, Rickart 1987, McLean et al. 2016).

When viewed in isolation, patterns of niche evolution are consistent with speciation driven by non-ecological processes, such as via allopatric or parapatric speciation. However, these latter modes of speciation are not expected to produce the association between niche mean and niche breadth on macroevolutionary scales observed in this system (Fig. 2, Table 3). Our finding that the magnitude of divergence in niche is correlated with an expansion of niche breadth is more consistent with ecological opportunity; larger niche shifts reveal ecological opportunity in the form of unused resources or reduced competition, allowing both ecological and geographic expansion. This correlation may be driven in large part by colonization of vast high latitude ranges by big-eared species, both in the Nearctic and Palearctic. Exposure of vast tracts of northern North America and northwestern Asia following glacial periods are known to have generated gradients of ecological opportunity by appearance of new resource bases as well as reduced competition due to southward faunal compression or extinction. However, the relationship holds for some niche shifts at lower latitudes, indicating that the phenomenon of expanding niche breadth is not restricted to high-latitude species.

Competing hypotheses exist as to the general pattern of niche breadth evolution when ecological opportunity is driving speciation (Schluter 2000, Ch. 3). These ideas are relevant to our assertion of ecologically-mediated speciation in *Urocitellus*. On one hand, if ecological opportunity is found via increasingly finer partitioning of ecological space, then niche shifts should be apparent and also accompanied by decreasing niche breadth. This may occur on islands, where both geography and competition presumably impose strong constraints on niche breadth. Conversely, if ecological opportunity is found by continuous expansion to (and exploitation of) new niche space, as might occur in continental settings, no general decrease in niche breadth should be expected if ecological

opportunities are high. Instead, geographically widespread resource bases may even facilitate population persistence across broader ecological and geographical zones. Schluter (2000) reanalyzed a series of datasets and concluded that stable or expanding niche breadth is indeed the more common response to ecological opportunity in radiating clades from continents and islands, suggesting this could be a common feature of ecologically-driven radiations.

*Signals of Environmental Adaptation*

While the link between niche mean and breadth may be evidence that *Urocitellus* diversification could be ecologically mediated, linking such patterns to the speciation process requires evidence of association between niches and the traits conferring fitness benefits in those niches (i.e., phenotype-environment correlations). This is because adaptive divergence in mean phenotype along ecological axes is capable of driving assortative mating and therefore the evolution of reproductive isolation (Schluter 2000, Coyne and Orr 2004). However, documenting phenotype-environment correlations can be difficult, as it requires *a priori* identification of relevant ecological axes, adaptive traits, and the precise measurement and quantification of both. This process may be more readily identifiable in high latitude systems if starker ecological gradients also give rise to more straightforward phenotypic changes, as observed in other taxa (e.g., Schluter and McPhail 1993, Schluter 2016).

At a comparative level, we recovered a strong association between environmental variables and all measured phenotypic traits (Table 3). This breadth of significant associations is partly mediated by body size, as body size is strongly associated with thermal niche, and body proportions and cranial shape each display strong allometry. Nevertheless, significant associations of size-corrected shape data with niche variables also exist, indicating additional axes of shape adaptation that are also themselves highly correlated (Fig. 4e). Concordance between body and cranial shape is surprising given that these traits do not appear functionally linked, and diversity often exists between different trait classes (e.g., Harmon et al. 2005, Hopkins and Lidgard 2012, McLean et al. in prep).

Some body shape adaptations may follow ecogeographic rules, because recovered a strong linear relationship between body size and MAT which, together with high values of Blomberg's $K$ for body size ($K = 1.96$), suggests this trait is of substantial adaptive value across *Urocitellus*. This result is in agreement with Bergmann's Rule as borne out among closely related species (i.e., as originally described by Bergmann; Blackburn et al. 1999). Body size is known to contribute to assortative mating in some clades (Jiang et al. 2013) and could be important for evolution of reproductive isolation in ground squirrels. Conversely, evidence for past hybridization among big-eared species of differing body sizes (including *U. parryii*) suggests size is not a direct physical prezygotic isolating mechanism. Nevertheless, other potential behavioral, ecological, or phenological isolating mechanisms important in *Urocitellus* may also scale with size (Jiang et al. 2013, Richardson et al. 2014).

Conversely, there was no significant association of any shape trait with MAT at the interspecific level. Allen's Rule predicts that reduced appendage size should accompany life in colder climates as a means of thermal conservation. However, this rule has received less critical evaluation than Bergmann's Rule and is unsupported by some

datasets (e.g., Scholander 1955). As a result, the potential role of appendage length in thermal adaptation and diversification of mammals is poorly characterized. Combined with the significant but modest strength of correlation between body size and MAT (Fig. 5a), these results suggest that physiological and/or behavioral mechanisms other than body size and appendage length mediate adaptation to different thermal niches (e.g., Fristoe et al. 2015).

Patterns of body size and shape variation within species likewise showed mixed results, indicating that any potential link between macroevolutionary patterns and speciation processes is complex. Of the 10 species with reasonable sampling in our body measurement dataset, 4 conformed to Bergmann's Rule while 3 others displayed an inverse Bergmann's pattern (Fig. 5d). Similar inconsistency was found in a recent, more taxonomically expansive analysis (Riemer et al. 2017). Relative appendage lengths also showed mixed patterns of association with MAT, with nearly even numbers of species showing support for Allen's Rule and its inverse in each trait (Fig. 5f-h). These results do not support ecogeographic rules as general axes of within-species adaptation in *Urocitellus*; however, thermal ranges of *Urocitellus* species vary in both their mean and variance (e.g., Supp. Fig 1), such that conformation to ecogeographic predictions may thus not be expected in all cases. Those species occurring across the widest environmental gradients typically demonstrate significant relationships associations with thermal environment (*U. beldingi, U. parryii, U. richardsonii, U. undulatus*), but finer-scale sampling is necessary to understand whether this is a methodological artifact (broader latitudinal range, larger sample sizes, or both) or, alternatively, if it might signify ecogeographic adaptations are only manifest in the most extreme environments.

*The Nature of Adaptive Diversification in Urocitellus*

Clades undergoing rapid diversification mediated by ecology should display niche shifts and adaptations enabling enhanced fitness in those niches. We were able to employ novel phylogenetic, morphometric, and ecological datasets to test these predictions at a comparative level. Some results are consistent with a role for ecology in *Urocitellus* speciation, but further population-scale sampling will ultimately be necessary to confirm this hypothesis, as observed differences along ecological and phenotypic axes may have evolved subsequent to speciation, for example, in allopatry or parapatry (Ackerly et al. 2006, Keller and Seehausen 2012). For the few phenotypic traits for which extensive inter- and intraspecific sampling was available, there is little evidence for similarity across scales that could implicate them as involved causally in ecological and taxonomic diversification.

It is unlikely that ecological speciation in clades is always linked to morphological traits; physiological adaptations related to phenology and diet may likewise be adaptive in clades with broad geographic distributions. For example, small-eared *Urocitellus* species are geographically restricted to the Great Basin and its periphery, a desert ecosystem that presents harsh challenges for many animals. These species are known to display important adaptations to these environments including urine-concentrating capabilities (Rickart 1989) and shifts towards an earlier active season than big-eared *Urocitellus* (Feb/Mar – July/August depending on species). Conversely, the Arctic ground squirrel (*U. parryii*), which exists on the opposite end of the

environmental spectrum (e.g., Fig. 2), displays major physiological adaptations related to hibernation (Barnes 1989, Boyer and Barnes 1999, McLean in press). A better comprehension of the exact adaptive components involved in *Urocitellus* might therefore be gained through fieldwork focused on characterizing additional morphological and physiological traits and their relationship to finer-scale environmental properties.

## CONCLUSIONS

We analyzed phylogenomic, ecological, and phenotypic datasets to assess the potential ecological bases of rapid radiation of *Urocitellus* ground squirrels. Expansions in niche breadth, correlations between multiple phenotypic traits and environmental variables, and adherence of some of these correlations to common ecogeographical rules are consistent with a role for ecological opportunity in the speciation process. These results set the stage for more detailed explorations of environment-phenotype correlations in high latitude organisms. In addition, our approach demonstrates how biodiversity data in combination with novel empirical datasets can help to parse patterns of ecological and phenotypic divergence at different taxonomic scales. Such approaches in combination with fine-scale natural history and field-based studies are essential for identifying the drivers of assortative mating and thus the ecology of speciation in high latitude organisms.

ACKNOWLEDGEMENTS

First and foremost, we thank the museum curators and collection managers who loaned tissues for sequencing, crania for imaging, and who have made their specimen records and associated data available online in VertNet. For access to crania in particular, we thank Robert Timm, Jon Dunnum, Kristofer Helgen, Darrin Lunde, Suzy Peurach, Esther Langan, Nicole Edmison. This work was funded in part by the National Science Foundation (NSF 1258010), Smithsonian Institution (Peter Buck Predoctoral Fellowship to B.S.M.), American Museum of Natural History (Theodore Roosevelt Memorial Research Grant to B.S.M.), and T&E of New Mexico, Inc. Jocie Colella provided valuable comments on an earlier draft of this manuscript.

REFERENCES

Ackerly, D.D., D.W. Schwilk, C.O. Webb. 2006. Niche evolution and adaptive radiation: testing the order of trait divergence. Ecology 87:S50–S61.

Adams, D.C. 2014. A generalized *K* statistic for estimating phylogenetic signal from shape and other high-dimensional multivariate data. Systematic Biology 63:685–697.

Adams, D.C., M. Collyer, E. Sherratt. 2016. geomorph v3.0.3. https://cran.r-project.org/web/packages/geomorph/index.html.

Barnes, B.M. 1989. Freeze avoidance in a mammal: body temperatures below 0 degrees C in an Arctic hibernator. Science 244:1593–1595.

Blackburn, T.M., K.J. Gaston, N. Loder. 1999. Geographic gradients in body size: A clarification of Bergmann's rule. Diversity and Distributions 5:165–174.

Borowiec, M.L. 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. PeerJ 4:e1660.

Boyer, B.B., B.M. Barnes. 1999. Molecular and metabolic aspects of mammalian hibernation. Bioscience 49:713–724.

Coyne, J.A., A.H. Orr. 2004. Speciation. Sinauer Associates, Inc., Sunderland, Massachusetts.

Durrant, S.D., R.M. Hansen. 1954. Distribution patterns and phylogeny of some western ground squirrels. Systematic Zoology 3:82–85.

Faircloth, B.C. 2015. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics 32:786–788.

Flannery, T. 2001. The Eternal Frontier: An Ecological History of North America and Its Peoples. Atlantic Monthly Press, New York.

Fristoe, T.S., J.R. Burger, M.A. Balk, I. Khaliq, C. Hof, J.H. Brown. 2015. Metabolic heat production and thermal conductance are mass-independent adaptations to thermal environment in birds and mammals. Proceedings of the National Academy of Sciences 112:15934–15939.

Gilbert, P.S., J. Chang, C. Pan, E.M. Sobel, J.S. Sinsheimer, B.C. Faircloth, M.E. Alfaro. 2015. Genome-wide ultraconserved elements exhibit higher phylogenetic informativeness than traditional gene markers in percomorph fishes. Molecular Phylogenetics and Evolution 92:140–146.

Gillman, L.N., S.D. Wright, J. Cusens, P.D. McBride, Y. Malhi, R.J. Whittaker. 2015. Latitude, productivity and species richness. Global Ecology and Biogeography 24:107–117.

Gíslason, D., M.M. Ferguson, S. Skúlason, S.S. Snorrason. 1999. Rapid and coupled phenotypic and genetic divergence in Icelandic Arctic char (*Salvelinus alpinus*). Canadian Journal of Fisheries and Aquatic Sciences 56:2229–2234.

Goodwin, H.T. 2008. Sciuridae. Pp. 355–376 *in* Janis, Christine M., G. F. Gunnell, and M. D. Uhen, eds. Evolution of Tertiary Mammals of North America: Small Mammals, Xenarthrans, and Marine Mammals. Cambridge University Press, Cambridge.

Guralnick, R.P., P.F. Zermoglio, J. Wieczorek, R. LaFrance, D. Bloom, L. Russell. 2016. The importance of digitized biocollections as a source of trait data and a new VertNet resource. Database: baw158.

Harmon, L.J., R.E. Glor. 2010. Poor statistical performance of the mantel test in phylogenetic comparative analyses. Evolution 64:2173–2178.

Harmon, L.J., J.J. Kolbe, J.M. Cheverud, J.B. Losos. 2005. Convergence and the multidimensional niche. Evolution 59:409–421.

Harrison, R.G., S.M. Bogdanowicz, R.S. Hoffmann, E. Yensen, P.W. Sherman. 2003. Phylogeny and evolutionary history of the ground squirrels (Rodentia: Marmotinae). Journal of Mammalian Evolution 10:249–276.

Harrod, C., J. Mallela, K.K. Kahilainen. 2010. Phenotype-environment correlations in a putative whitefish adaptive radiation. Journal of Animal Ecology 79:1057–1068.

Hawkins, B.A., R. Field, H.V. Cornell, D.J. Currie, J.F. Guégan, D.M. Kaufman, J.T. Kerr, G.G. Mittelbach, T. Oberdorff, E.M. O'Brien, E.E. Porter, J.R.G. Turner. 2003. Energy, water, and broad-scale geographic patterns of species richness. Ecology 84:3105–3117.

Hawkins, M.T.R., J.A. Leonard, K.M. Helgen, M.M. McDonough, L.L. Rockwood, J.E. Maldonado. 2016. Evolutionary history of endemic Sulawesi squirrels constructed from UCEs and mitogenomes sequenced from museum specimens. BMC Evolutionary Biology 16:80.

Heibl, C. 2015. ips v0.0-7. https://cran.r-project.org/web/packages/ips/index.html.

Helgen, K.M., F.R. Cole, L.E. Helgen, D.E. Wilson. 2009. Generic revision in the Holarctic ground squirrel genus *Spermophilus*. Journal of Mammalogy 90:270–305.

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

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones, A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25:1965–1978.

Hijmans, R.J., J. van Etten, J. Cheng, M. Mattiuzzi, M. Sumner, J.A. Greenberg, O.P. Lamigueiro, A. Bevan, E. Racine, A. Shortridge. 2016. raster v2.5-8. https://cran.r-project.org/web/packages/raster/index.html.

Hope, A G., N. Takebayashi, K.E. Galbreath, S.L. Talbot, J.A. Cook. 2012. Temporal, spatial and ecological dynamics of speciation among amphi-Beringian small mammals. Journal of Biogeography 40:415–429.

Hopkins, M.J., S. Lidgard. 2012. Evolutionary mode routinely varies among morphological traits within fossil species. Proceedings of the National Academy of Sciences 109:20520–20525.

Howell, A.H. 1938. Revision of the North American Ground Squirrels, with a Classification of the North American Sciuridae. North American Fauna 56:1–256. Washington, D.C.

Jiang, Y., D.I. Bolnick, M. Kirkpatrick. 2013. Assortative mating in animals. The American Naturalist 181:E125–E138.

Jiang, Z., A.R. Huete, K. Didan, T. Miura. 2008. Development of a two-band enhanced vegetation index without a blue band. Remote Sensing of Environment 112:3833–3845.

Katoh, K., D.M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30:772–780.

Kiell, D.J., and J.S. Millar. 1978. Growth of juvenile arctic ground squirrels (*Spermophilus parryii*) at Mcconnell River, N.W.T. Canadian Journal of Zoology 56:1475–1478.

Koeppl, J.W., R.S. Hoffmann. 1981. Comparative postnatal growth of four ground squirrel species. Journal of Mammalogy 62:41–57.

Kozak, K.H., J.J. Wiens. 2006. Does niche conservatism promote speciation? A case study in North American salamanders. Evolution 60:2604–2621.

Lapointe, F.-J., T. Garland. 2001. A generalized permutation model for the analysis of cross-species data. Journal of Classification 18:109–127.

Lessa, E.P., J.A. Cook, J.L. Patton. 2003. Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary. Proceedings of the National Academy of Sciences 100:10331–10334.

Lu, G., L. Bernatchez. 1999. Correlated trophic specialization and genetic divergence in sympatric lake whitefish ecotypes (*Coregonus clupeaformis*): support for the ecological speciation hypothesis. Evolution 53:1491–1505.

McCormack, J.E., B.C. Faircloth, N.G. Crawford, P.A. Gowaty, R.T. Brumfield, T.C. Glenn. 2012. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. Genome Research 22:746–754.

McCormack, J.E., A.J. Zellmer, L.L. Knowles. 2009. Does niche divergence accompany allopatric divergence in *Aphelocoma* jays as predicted under ecological speciation?: Insights from tests with niche models. Evolution 64:1231–1244.

McLean, B.S., D.J. Jackson, J.A. Cook. 2016. Rapid divergence and gene flow at high latitudes shape the history of Holarctic ground squirrels (*Urocitellus*). Molecular Phylogenetics and Evolution 102:174–188.

Meiri, S., T. Dayan. 2003. On the validity of Bergmann's rule. Journal of Biogeography 30:331–351.

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

Millien, V., S. Kathleen Lyons, L. Olson, F.A. Smith, A.B. Wilson, Y. Yom-Tov. 2006. Ecotypic variation in the context of global climate change: revisiting the rules. Ecology Letters 9:853–869.

Nadler, C.F. 1966. Chromosomes and systematics of American ground squirrels of the subgenus *Spermophilus*. Journal of Mammalogy 47:579–596.

Nadler, C.F., R.S. Hoffmann. 1977. Patterns of evolution and migration in the arctic ground squirrel, *Spermophilus parryii* (Richardson). Canadian Journal of Zoology 55:748–758.

Nadler, C.F., R.S. Hoffmann, N.N. Vorontsov, J.W. Koeppl, L. Deutsch, R.I. Sukernik. 1982. Evolution in ground squirrels: II. Biochemical comparisons in Holarctic populations of *Spermophilus*. Zeitschrift Fur Saugetierkunde 47:198–215.

Nadler, C.F., R.I. Sukernik, R.S. Hoffmann, N.N. Vorontsov, C.F. Nadler Jr., I.I. Fomichova. 1974. Evolution in ground squirrels I. Transferrins in Holarctic

populations of *Spermophilus*. Comparative Biochemistry and Physiology A 47:663–681.

Nevo, E. 1979. Adaptive convergence and divergence of subterranean mammals. Annual Reviews of Ecology and Systematics 10:269–308.

Ólafsdóttir, G.Á., S.S. Snorrason, M.G. Ritchie. 2007. Postglacial intra-lacustrine divergence of Icelandic threespine stickleback morphs in three neovolcanic lakes. Journal of Evolutionary Biology 20:1870–1881.

Paradis, E., B. Bolker, J. Claude, H.S. Cuong, R. Desper, B. Durand, J. Dutheil, O. Gascuel, C. Heibl, D. Lawson, V. Lefort, P. Legendre, J. Lemon, Y. Noel, J. Nylander, R. Opgen-Rhein, A.-A. Popescu, K. Schliep, K. Strimmer, D. de Vienne. 2015. ape: analyses of phylogenetics and evolution v4.1. https://cran.r-project.org/web/packages/ape/index.html.

Peterson, A.T., J. Soberon, V. Sanchez-Cordero. 1999. Conservatism of ecological niches in evolutionary time. Science 285:1265–1267.

Pyron, R.A., G.C. Costa, M.A. Patten, F.T. Burbrink. 2015. Phylogenetic niche conservatism and the evolutionary basis of ecological speciation. Biological Reviews 90:1248–1262.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Richardson, J.L., M.C. Urban, D.I. Bolnick, D.K. Skelly. 2014. Microgeographic adaptiation and the spatial scale of evolution. Trends in Ecology and Evolution 29:165–176.

Rickart, E.A. 1987. *Spermophilus townsendii*. Mammalian Species 268:1–6.

Rickart, E.A. 1989. Variation in renal structure and urine concentrating capacity among ground squirrels of the *Spermophilus townsendii* complex (Rodentia: Sciuridae). Comparative Biochemistry and Physiology A 92:531–534.

Riemer, K., R.P. Guralnick, E. White. 8 March 2017. No general relationship between mass and temperature in endotherm species. bioRxiv:1–27.

Rohlf, F.J. 2006. tpsDig, version 2.10. Department of Ecology and Evolution, State University of New York, Stony Brook.

Sandel, B., L. Arge, B. Dalsgaard, R.G. Davies, K.J. Gaston, W.J. Sutherland, J.-C. Svenning. 2011. The influence of late Quaternary climate-change velocity on species endemism. Science 334:660–664.

Schlager, S., G. Jefferis. 2016. Morpho v2.4.1. https://cran.r-project.org/web/packages/Morpho/index.html.

Schluter, D. 2016. Speciation, ecological opportunity, and latitude. The American Naturalist 187:1–18.

Schluter, D. 2000. The Ecology of Adaptive Radiation. Oxford University Press, Oxford.

Schluter, D., J.D. McPhail. 1992. Ecological character displacement and speciation in sticklebacks. The American Naturalist 140:85–108.

Schluter, D., J.D. McPhail. 1993. Character displacement and replicate adaptive radiation. Trends in Ecology and Evolution 8:197–200.

Scholander, P.F. 1955. Evolution of climatic adaptation in homeotherms. Evolution 9:15–26.

Sobel, J.M., G.F. Chen, L.R. Watt, D.W. Schemske. 2009. The biology of speciation. Evolution 64:295–315.

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

Stayton, C.T. 2015. The definition, recognition, and interpretation of convergent evolution, and two new measures for quantifying and assessing the significance of convergence. Evolution 69:2140–2153.

Streicher, J.W., J.J. Wiens. 2016. Phylogenomic analyses reveal novel relationships among snake families. Molecular Phylogenetics and Evolution 100:160–169.

Todorov, V. 2016. rrcovHD v0.2-5. https://cran.r-project.org/web/packages/rrcovHD/index.html.

Weir, J.T., O. Haddrath, H.A. Robertson, R.M. Colbourne, A.J. Baker. 2016. Explosive ice age diversification of kiwi. Proceedings of the National Academy of Sciences 113:E5580–E5587.

Weir, J.T., D. Schluter. 2007. The latitudinal gradient in recent speciation and extinction rates of birds and mammals. Science 315:1574–1576.

Wiens, J.J., C.H. Graham. 2005. Niche conservatism: integrating evolution, ecology, and conservation biology. Annual Reviews of Ecology, Evolution, and Systematics 36:519–539.

Willig, M.R., D.M. Kaufman, R.D. Stevens. 2003. Latitudinal gradients of biodiversity: patterns, scale, and synthesis. Annual Reviews of Ecology, Evolution, and Systematics 34:273–309.

Zelditch, M.L., J. Li, L.A.P. Tran, D.L. Swiderski. 2015. Relationships of diversity, disparity, and their evolutionary rates in squirrels (Sciuridae). Evolution 69:1284–1300.

FIGURE CAPTIONS

Figure 1. a) Geographic distribution of *Urocitellus*, with clades/species colored to illustrate major disparities in range size and location (orange = small-eared clade, green = big-eared species, blue = *U. parryii*, pink = *U. undulatus*). b) Phylogeny of *Urocitellus* and outgroups inferred from concatenated analysis of 3,531 UCE loci (2,146,015bp) in RAxML. All branches were highly supported (≥99% bootstrap support). Some additional samples sequenced were trimmed from the tree for clarity.

Figure 2. Patterns of variation in realized niches of *Urocitellus* species as visualized by linear discriminant analysis. Niches were quantified based on 15 environmental variables (13 temperature and precipitation metrics and 2 vegetation indices) extracted for 2,808 georeferenced specimen records. a) LD1 versus LD2, b) LD1 versus LD3. Total variation explained by each function is listed on axis. Species are colored as in Fig. 1 and labeled inside plots (ar = armatus, be = beldingi, br = brunneus, ca = canus, co = columbianus, el = elegans, mo = mollis, pa = parryii, ri = richardsonii, to = townsendii, un = undulatus, wa = washingtoni).

Figure 3. Cranial shape variation among *Urocitellus* species visualized by linear discriminant analysis. Data represent ventral cranial shape quantified with 24 2D landmarks (a)). b) LD1 vs LD2, c) LD1 versus LD3. Species are colored by major group as in Figs. 1 and 2. Scale bar in a) equals 1cm.

Figure 4. Summary plot depicting variation in 3 ecophenotypic traits in *Urocitellus* ground squirrels: a) body size (expressed as logarithm of head-body length), b) body shape (PC1 from size-corrected data), and c) cranial shape (PC1 from size-corrected data). The plot in e) is a regression of phylogenetic independent contrasts for cranial shape PC1 (40% of total variation) on contrasts for body shape PC1 (59.4% of total variation), with points labeled by node as shown in d). A regression of contrasts for PC2 of both traits is likewise significant ($r2 = 0.41$, $P = 0.013$, results not shown).

Figure 5. Tests for conformation to Bergmann's Rule and Allen's Rule at the among- and within-species levels. a-d) Linear regressions of contrasts for traits on contrasts for mean annual temperature (MAT). Trait data used in b-d) were residuals from regressions of species mean values on species head body lengths. e-h) Results of within-species linear regressions of the same traits used above on MAT for 10 species (2 species lacked sufficient museum records and were not included). Each barplot shows the magnitude and direction of the slope. Linear relationships significant at the $p < 0.05$ level are indicated by red asterisks.

TABLES

Table 1. Summary statistics for trimmed UCE alignments, computed prior to filtering. Data are presented as medians with first and third quantiles where applicable. Top: outgroups and ingroups, Bottom: ingroup taxa only.

| | **Samples** | **Loci** | **Locus Length** | **Taxa per Locus** | **Percent Variable Sites** | **Percent Informative Sites** |
|---|---|---|---|---|---|---|
| *Urocitellus* + outgroups | 34 | 3752 | 560 (461,677) | 30 (20,32) | 4.1 (2.4,6.5) | 1.6 (0.7,2.8) |
| *Urocitellus* only | 28 | 3727 | 562 (453,679) | 24 (17,27) | 3.2 (1.8,5.1) | 1.1 (0.4,2.0) |

Table 2. Environmental variables used in the quantification of mean niche of Urocitellus species, their descriptions, and loadings in a principal components analysis (PCA) of species mean values (PC1-3 shown only). The 6 variables loading most strongly on each PC axis (3 most positive and 3 most negative) are indicated in bold to facilitate interpretability of niche axes. The variation explained by each PC as well as cumulatively is also shown.

| Variable | Description | PC1 | PC2 | PC3 |
|---|---|---|---|---|
| BIO1 | Annual Mean Temperature | **0.3103** | **-0.1529** | 0.0539 |
| BIO2 | Mean Diurnal Temp Range | 0.2184 | -0.0806 | **0.4489** |
| BIO3 | Isothermality (BIO2/BIO7) | **0.3293** | -0.0080 | 0.1707 |
| BIO4 | Temperature Seasonality | **-0.3454** | -0.0834 | -0.0363 |
| BIO7 | Temperature Annual Range (BIO5-BIO6) | **-0.2973** | -0.1064 | 0.2310 |
| BIO8 | Mean Temperature of Wettest Quarter | -0.2909 | **-0.1242** | **0.3072** |
| BIO9 | Mean Temperature of Driest Quarter | **0.3429** | -0.0512 | -0.1218 |
| BIO12 | Annual Precipitation | 0.0464 | **0.5389** | 0.0138 |
| BIO13 | Precipitation of Wettest Month | -0.1140 | **0.4425** | **-0.2521** |
| BIO14 | Precipitation of Driest Month | 0.0392 | **0.4268** | **0.4586** |
| BIO15 | Precipitation Seasonality | -0.2599 | **-0.1271** | **-0.4557** |
| BIO18 | Precipitation of Warmest Quarter | **-0.3130** | 0.2039 | 0.2101 |
| BIO19 | Precipitation of Coldest Quarter | 0.2413 | 0.3483 | **-0.2634** |
| EVI_mean | Enhanced Vegetation Index (8 month Mean) | 0.1383 | 0.1292 | -0.0785 |
| EVI_stdev | Enhanced Veg. Index Seasonality (8 month Std. Dev.) | -0.2794 | 0.2639 | 0.0333 |
| Total (and Cumulative) % Among-Species Variation Explained | | 52.4 | 21.8 (74.2) | 8.8 (83.1) |

Table 3. Results of Mantel tests for correlations between niche mean and niche breadth, and between niche mean and species means on 5 phenotypic datasets. Exact variables used in each distance matrix are shown. Body shape and cranial shape residuals are residuals from a regression of original datasets on logarithm of head-body length. Significance for all tests was assessed using the phylogenetic permutation method of Lapointe and Garland (1992), with 9999 iterations.

| **Matrix 1 (variables)** | **Matrix 2 (variables)** | **Z** | ***P*** | **r** |
|---|---|---|---|---|
| Niche Mean (PC1-5) | Niche Breadth (PC1-5) | 3551.76 | 0.002 | 0.53 |
| Niche Mean (PC 1-5) | Head-body Length | 143.66 | 0.003 | 0.37 |
| Niche Mean (PC1-5) | Body Shape (PC1-3) | 436.07 | 0.001 | 0.42 |
| Niche Mean (PC1-5) | Body Shape residuals (PC1-3) | 161.53 | 0.003 | 0.40 |
| Niche Mean (PC1-5) | Cranial Shape (PC1-5) | 25.12 | 0 | 0.63 |
| Niche Mean (PC1-5) | Cranial Shape residuals (PC1-5) | 19.09 | 0.009 | 0.36 |

FIGURES

Figure 1.

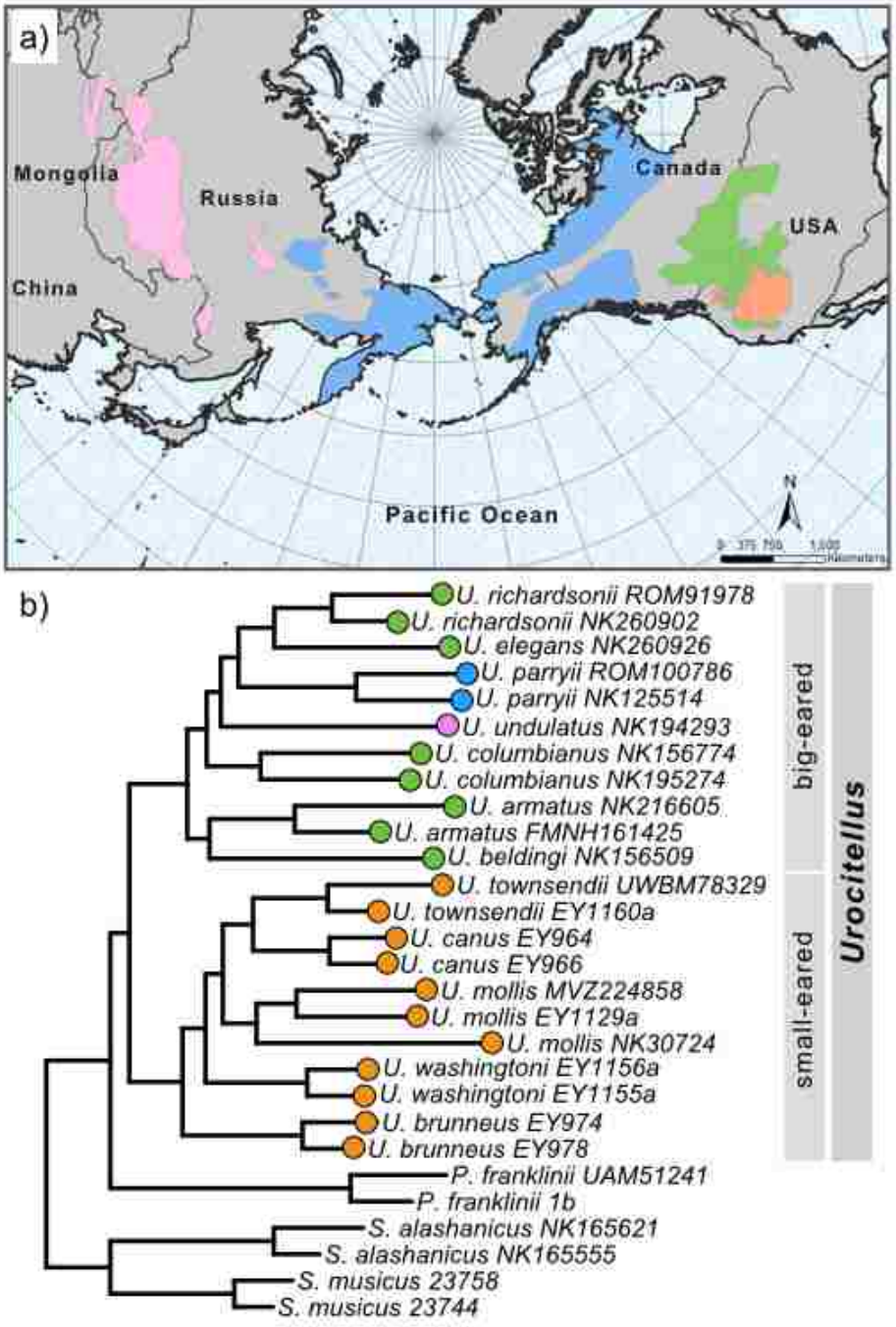


Figure 2.

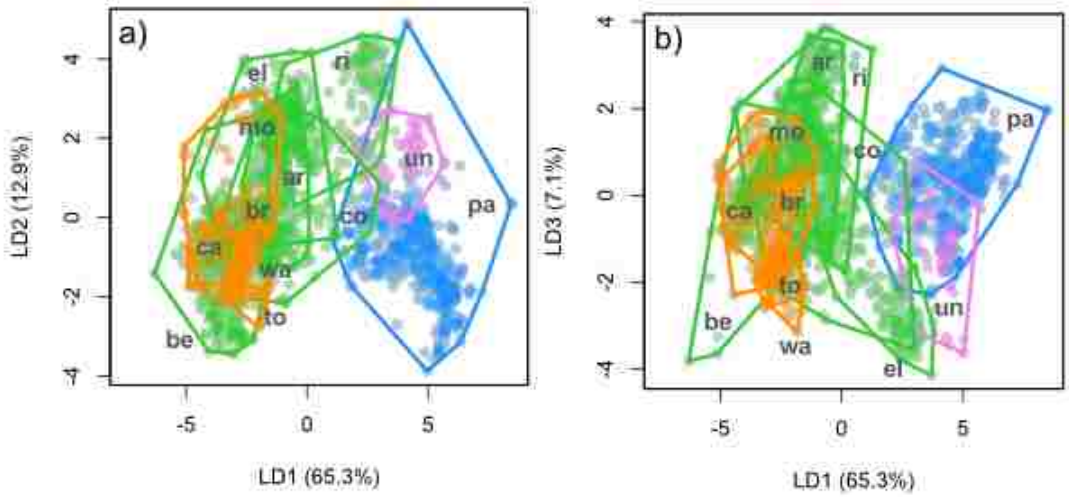

a)
el
ri
mo
un
ar
pa
br
co
ca
wa
to
be
LD2 (12.9%)
LD1 (65.3%)
b)
ar
ri
pa
mo
co
ca
br
to
un
be
wa
el
LD3 (7.1%)
LD1 (65.3%)

Figure 3.

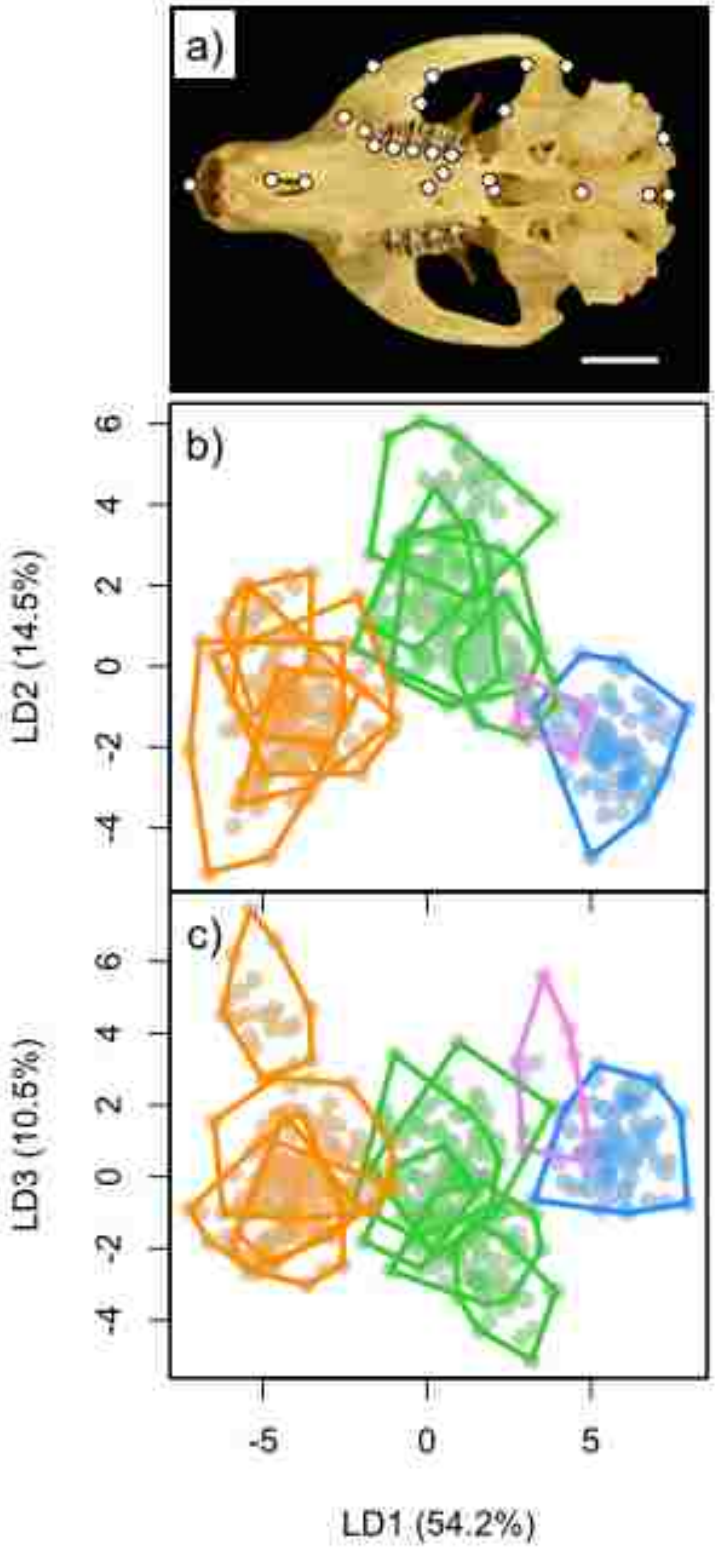


Figure 4.

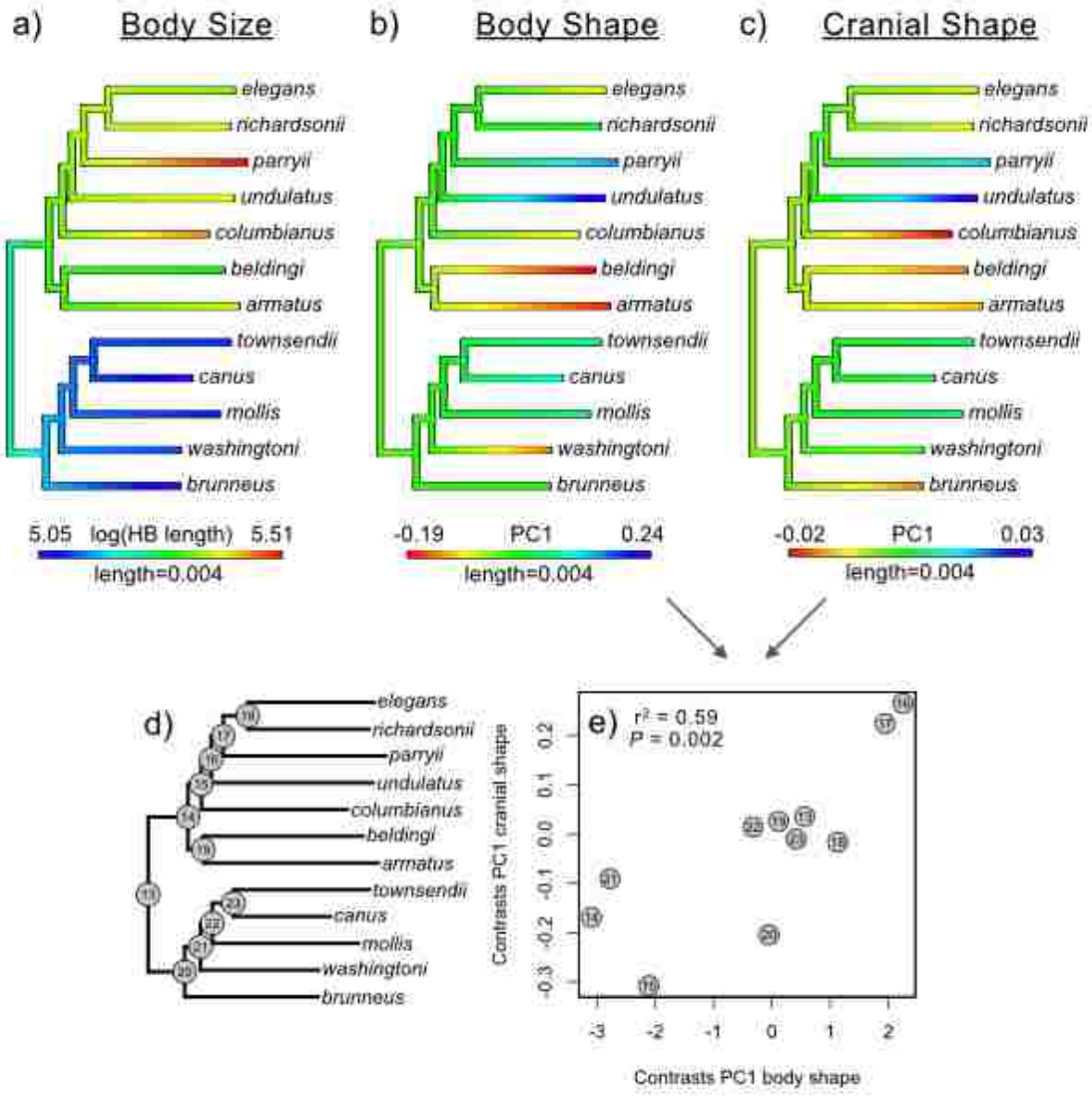
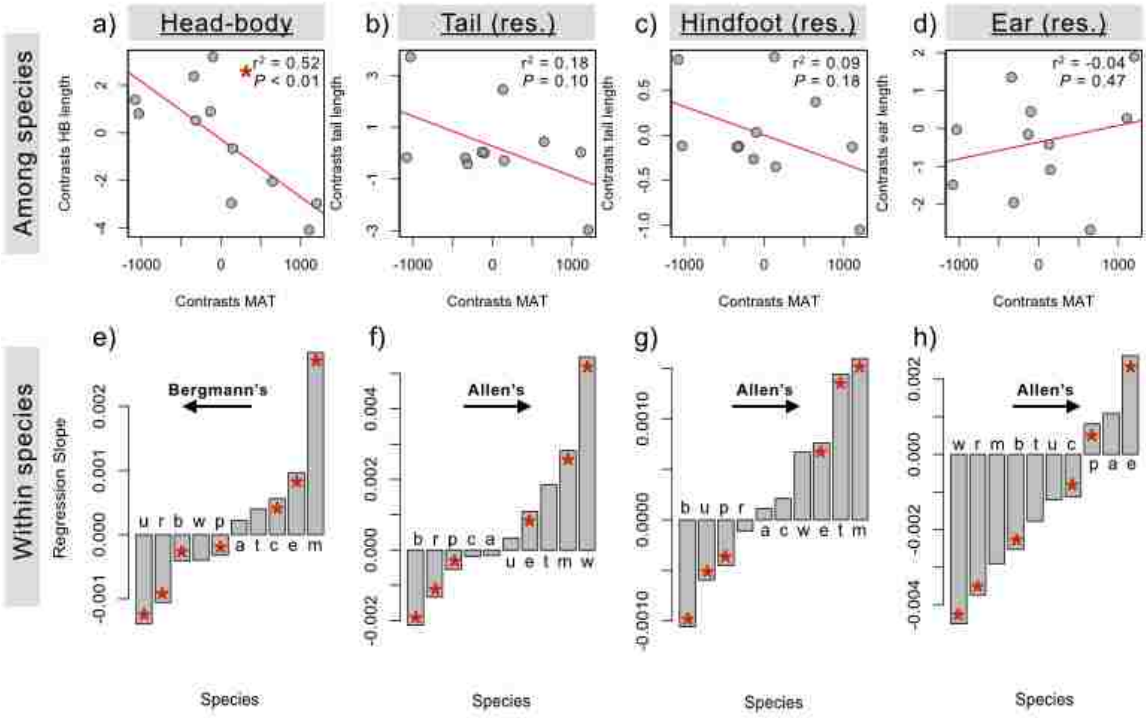

a) Body Size
b) Body Shape
c) Cranial Shape
elegans
richardsonii
parryii
undulatus
columbianus
beldingi
armatus
townsendii
canus
mollis
washingtoni
brunneus
5.05 log(HB length) 5.51
length=0.004
-0.19 PC1 0.24
length=0.004
-0.02 PC1 0.03
length=0.004
d)
e)
r² = 0.59
P = 0.002
Contrasts PC1 cranial shape
Contrasts PC1 body shape

Figure 5.


Among species
Within species
a) Head-body
b) Tail (res.)
c) Hindfoot (res.)
d) Ear (res.)
r² = 0.52
P < 0.01
r² = 0.18
P = 0.10
r² = 0.09
P = 0.18
r² = -0.04
P = 0.47
Contrasts HB length
Contrasts tail length
Contrasts ear length
Contrasts MAT
e)
f)
g)
h)
Bergmann's
Allen's
Regression Slope
u r b w p a t c e m
b r p c a u e t m w
b u p r a c w e t m
w r m b t u c p a e
Species

SUPPLEMENTARY FIGURE CAPTIONS

Supplementary Figure 1. Tests for conformation to Bergmann's Rule (column 1) and Allen's Rule (columns 2-4) for 5 selected *Urocitellus* species, chosen based on large available sample size (>100) and uniqueness in PCAs of multispecies data. For each species (=row), the relationship of head-body, tail length, hindfoot length, and ear length with mean annual temperature was assessed using linear regression. The latter 3 measurements were residuals from a regression on head-body length. Linear relationships significant at the $p < 0.05$ level are indicated by red asterisks. Linear models that were marginally insignificant ($p < 0.06$) are indicated with black asterisks.

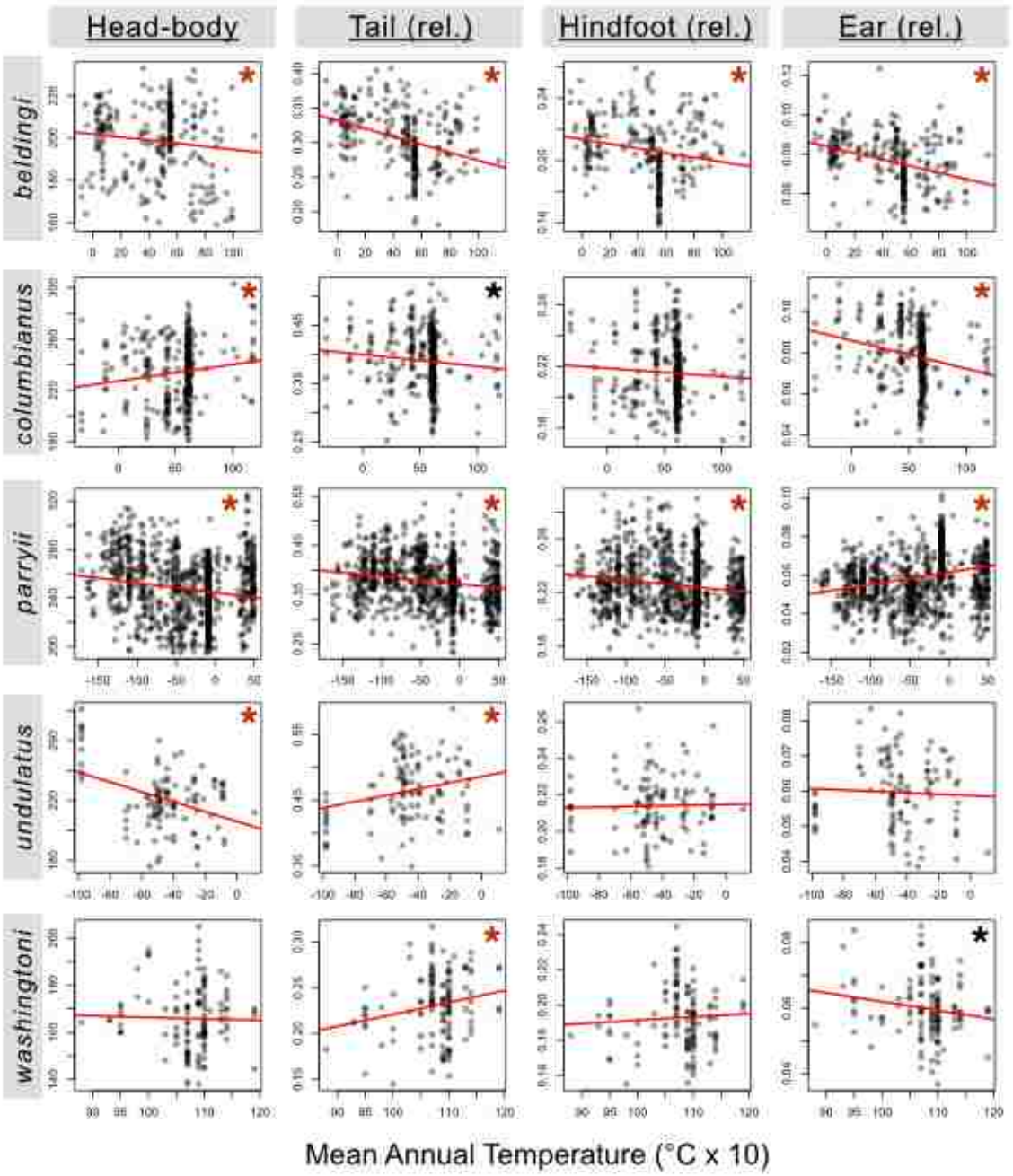
SUPPLEMENTARY TABLES

Supplementary Table 1. Environmental variables used in the quantification of niche breadth of *Urocitellus* species, their descriptions, and loadings in a principal components analysis (PCA) of species standard deviations for each variable (PC1-3 shown only). The 6 variables loading most strongly on each PC axis (3 most positive and 3 most negative) are indicated in bold to facilitate interpretability of major axes of niche variation. The variation among species breadths explained by each PC as well as cumulatively is also shown.

| Variable | Description | PC1 | PC2 | PC3 |
|---|---|---|---|---|
| BIO1 | Annual Mean Temperature | **0.3321** | -0.1019 | 0.0667 |
| BIO2 | Mean Diurnal Temp Range | **0.3185** | -0.1172 | 0.0902 |
| BIO3 | Isothermality (BIO2/BIO7) | **0.1608** | **0.3998** | -0.1665 |
| BIO4 | Temperature Seasonality | 0.2636 | **0.3097** | **-0.2701** |
| BIO7 | Temperature Annual Range (BIO5-BIO6) | **0.3159** | 0.1725 | -0.2141 |
| BIO8 | Mean Temperature of Wettest Quarter | **0.1076** | -0.0552 | **0.6270** |
| BIO9 | Mean Temperature of Driest Quarter | 0.2444 | 0.0434 | **0.4985** |
| BIO12 | Annual Precipitation | 0.2743 | **-0.3456** | -0.1192 |
| BIO13 | Precipitation of Wettest Month | 0.1987 | **-0.4353** | **-0.2289** |
| BIO14 | Precipitation of Driest Month | 0.3085 | 0.0158 | **0.1781** |
| BIO15 | Precipitation Seasonality | 0.2746 | 0.0258 | -0.0339 |
| BIO18 | Precipitation of Warmest Quarter | 0.2899 | 0.2313 | 0.0964 |
| BIO19 | Precipitation of Coldest Quarter | **0.1757** | **-0.5121** | -0.0719 |
| EVI_mean | Enhanced Vegetation Index (8 month Mean) | 0.1927 | -0.0562 | **-0.2921** |
| EVI_stdev | Enhanced Veg. Index Seasonality (8 month Std. Dev.) | 0.2907 | **0.2380** | -0.0034 |
| Total (and Cumulative) % Among-Species Variation Explained | | 53.3 | 18.1 (71.4) | 13.4 (84.9) |

SUPPLEMENTARY FIGURES

Supplementary Figure 1.



# CONCLUSIONS

There is considerable uncertainty in our understanding of diversification processes across the Tree of Life. While historical and biogeographical contingencies undoubtedly contribute to the observed taxonomic, phenotypic, and ecological disparities among clades, a role for ecological opportunity in these patterns is also widely acknowledged. This dissertation focused on describing phenotypic and ecological diversity in Marmotini and identifying some of the mechanisms by which these patterns have arisen. It resolved many key nodes in marmotine phylogeny, revealed evolutionary dynamics and idiosyncrasies across a suite of measured traits, and ultimately asked whether there might be processes in marmotine radiation that transcend taxonomic scales. Below, I review major conclusions of this work and identify needs for future empirical and theoretical research. I end by discussing the continued importance of a specimen-based approach for understanding diversification patterns and processes.

A major conclusion of this dissertation is that, while signatures of rapid speciation and ecophenotypic adaptation are detectable at all levels in Marmotini, the evolutionary processes characterizing radiation at each level are unique. Within *Urocitellus*, there has been continuous niche spread and expansion, a pattern consistent with response to ecological opportunity across temperate and high latitude biomes. Niches shifts have been accompanied by adaptations in body size, cranial shape, and body shape. Still, like many radiations of young age, phenotypic disparity in *Urocitellus* is low and reproductive isolation among close relatives remains incomplete. Radiation at and below the genus level in Marmotini may be analogous to movement of lineages around ridges of an adaptive subzone, with responses to local ecological opportunity but also morphological and ecological conservatism. Conversely, at a higher taxonomic level, traversals of the adaptive landscape are observable across multiple traits, reflecting colonizations of multiple adaptive subzones. Diversification at this level is more similar to subzonal radiation within the confines of a broader, yet topologically simple, adaptive zone (i.e., that favoring all ground squirrels).

Although taxonomic assignments are somewhat subjective in all clades, the above comparisons suggest that scale-dependency in evolutionary processes in mammals may be most evident below and above the genus level, respectively. Such scale-dependency has implications for how the comparative approach is employed to study radiations. Specifically, these dependencies indicate that the extent to which any given comparative study can be contextualized within current theoretical frameworks is limited. Not all clades may be radiating simultaneously and in comparable fashion. Moreover, traits presumed to be of adaptive importance may be evolving differently even among radiating subclades, as each clade perceives and exploits ecological opportunity differently. Comparative inferences at higher levels in particular may also be susceptible to historical biogeographical contingencies, heterogeneous speciation or extinction rates through time, and developmental or genetic constraints on phenotypic evolution. Nevertheless, results of this study are comparable to some other studies of continental radiations in demonstrating a decoupling of taxonomic and phenotypic diversity. This suggests a common mechanism or bias may be responsible for the lack of conformation to theoretical expectations, and points the way forward for identifying its basis.

In the context of broader radiation theory, there is a need to reconsider how climate and geography might interact with ecological opportunity over time to shift adaptive landscapes and contribute to diversification patterns. For example, the expansion of temperate and high latitude aridland biomes during the Neogene clearly played a role in marmotine diversification. Did these events greatly expand an existing adaptive peak, or create one anew? Are evolutionary patterns and processes in other groups inhabiting these same temperate biomes comparable to those observed in Marmotini, and, if not, why? Such reconsiderations should specifically address the roles of late Quaternary climate fluctuations in generating high latitude ecological opportunity and promoting adaptive diversification on recent timescales. In other words, to what extent is postglacial ecological opportunity being exploited by other high latitude organisms? Convincing assessments of these processes using genomic data and suites of ecological and phenotypic traits are necessary in many terrestrial organisms.

At the finest scale, there remains a pressing need to understand the speciation process in Marmotini and other continental radiations. How has lineage diversification been maintained at high rates despite relatively low divergence in ecology and phenotype? The comparative approach is powerful for describing variation in lineage and phenotypic diversification across phylogenies, but building support for particular speciation hypotheses is not possible from comparative approaches alone. Fruitful research avenues include analyses of differential diet and resource use across environmental gradients; character displacement in zones of sympatry or parapatry; assessments of the form and extent of gene flow across nuclear genomes; reassessment of the role that chromosomal evolution plays in reproductive isolation; and consideration of additional physiological and life history traits potentially involved in adaptive divergence but not commonly investigated in such contexts.

Finally, this dissertation demonstrates the importance of a natural history collections-based perspective in addressing pressing questions in evolutionary biology. Data used in this dissertation were drawn from a spatially, temporally, and taxonomically broad resource base of thousands of natural history specimens and specimen parts. Most of the specimens used were historic, and a few were well over a century old. Others were relatively new to the collective museum knowledge base, often being derived from expeditions conducted by the Museum of Southwestern Biology, and these filled crucial sampling gaps of the type that always exist when relying on historic material alone. Both old and new specimens are the foundation for describing how biological diversity has, and continues to, evolve. If fully utilized, these specimens are capable of fueling insights far beyond the magnitude of those presented in a single dissertation. Those insights will depend on a renewed commitment to collections growth and maintenance, a commitment that is also a social, political, and cultural imperative.