Graphic Abstract

Molecular Phylogenetics and Evolution – Research Article

Revision of widespread red squirrels (genus: *Tamiasciurus*) highlights the complexity of speciation within North American forests

Andrew G. Hope^{*af}, Jason L. Malaney^{bg}, Kayce C. Bell^{ch}, Fernando Salazar-Miralles^{ci}, Andreas S. Chavez^{dj}, Brian R. Barber^{ek}, Joseph A. Cook^{cl}

^aDivision of Biology, Kansas State University, Manhattan, KS 66506 USA

^bDepartment of Biology, Austin Peay State University, Clarksville, TN 37044 USA

^cDepartment of Biology and Museum of Southwestern Biology, University of New Mexico, Albuquerque, NM 87131 USA

^dMuseum of Vertebrate Zoology and Department of Integrative Biology, University of California-Berkeley, Berkeley, CA 94720 USA

^eBiodiversity Institute, University of Wyoming, Laramie, WY 82071, USA

^fahope@ksu.edu

^gjmalaney@gmail.com

^hkayce.bell@gmail.com

ⁱf.salazarmiralles@gmail.com

^jaschavez@berkeley.edu

^kbbarber@uwyo.edu

^lcookjose@unm.edu

* Corresponding author (A.G. Hope). Fax: 785-532-6653. E-mail: ahope@ksu.edu. Postal:
Kansas State University, Division of Biology, 116 Ackert Hall, Manhattan, KS 66547.

ABSTRACT

Integration of molecular methods, ecological modeling, and statistical hypothesis testing are increasing our understanding of differentiation within species and phylogenetic relationships among species by revealing environmental connections to evolutionary processes. Within mammals, novel diversity is being discovered and characterized as more complete geographic sampling is coupled with newer multi-disciplinary approaches. North American red squirrels exemplify a forest obligate genus whose species are monitored as indicators of forest ecosystem condition, yet phylogenetic relationships reflecting evolutionary history within this genus remain tentative. Through testing of competing systematic and niche-based divergence hypotheses, we recognize three species, *Tamiasciurus douglasii*, *T. hudsonicus*, and *T. fremonti*. Our data provide evidence of regional differences in evolutionary dynamics and continental gradients of complexity that are important both for future management and for investigating multiple pathways that can lead to the formation of new species.

Keywords

Boreal biome, Quaternary, sky island, Southeast Alaska, speciation with gene flow, species tree estimation

1. Introduction

Understanding how the formation of species is shaped by common biogeographic and evolutionary processes (Hewitt, 2000; Taberlet et al., 1998) is a critical step for recognizing and conserving biodiversity (Riddle et al., 2008; Whittaker et al., 2005). Range-wide evolutionary assessments of species that incorporate both genetic and niche-based approaches provide critical perspectives on regional evolutionary and ecological differences in biodiversity. For instance, many vertebrate species associated with the boreal biome in North America have broad continent-wide distributions coincident with temperate and northern conifer forests (Arbogast and Kenagy, 2001), but these forests are now experiencing increased mortality (Allen et al., 2015), and regional declines (Coops and Waring, 2011; Gauthier et al., 2015) will likely continue to impact vertebrate communities. Geography and past climate have shaped the assembly of communities according to predictable biogeographic patterns that often reflect common evolutionary processes (Hampe and Petit, 2005). Northernmost contemporary populations frequently constitute a leading edge of expansion due to climate warming following the Last Glacial Maximum (LGM; ~ the last 12 kys). However, high-latitude populations may have also experienced spatial and temporal stability if they persisted through the Last Glacial period (~130-12 ka) within one or more northern refugia (Brubaker et al., 2005; Hewitt, 1999; Lessa et al., 2003). Mid-latitude forest-associated populations often have multiple distinct lineages distributed longitudinally across the continent reflecting repeated isolation in discrete areas south of glacial-phase continental ice-sheets. However, the genetic signatures of independent histories among these lineages are often obscured at their distributional edges where they now overlap geographically (Atkinson et al., 2007; Hope et al., 2014a). Signatures of these complex histories

of isolation and reconnection are often apparent along contact zones (Swenson and Howard, 2005) where dynamics of gene flow and interspecific interactions result in increased community complexity (Hewitt, 2000). Finally, peripherally isolated populations in southern and coastal regions generally occur in patches of discontinuous forest. Peripheral populations can reflect either long-term isolation (≥ 1 glacial phase; e.g., Cook et al., 2001, 2006) or more recent fragmentation of previously continuous habitat since the LGM (Galbreath et al., 2009). Such isolates often constitute important sources of genetic diversity for species (Channell and Lomolino, 2000) with high conservation value (Malaney and Cook, 2013).

A growing body of comparative genetic evidence indicates that evolutionary processes of diversification on a continental scale are repeatable both across taxonomic groups and through time, where cryptic diversity is discovered as additional groups are assessed, resulting in recognition of new species, or minimally, distinct evolutionary lineages (Hope et al., 2014a). Given that species descriptions are ongoing, even among mammals, current taxonomy frequently fails to adequately reflect the extent of evolutionary diversity or relationships among extant lineages (Patterson, 2000; Bickford et al., 2007), thus impeding effective management and conservation of fundamental components of biodiversity (Soltis and Gitzendanner, 2001; Hoberg et al., 2012). New toolsets are facilitating species discovery based on multi-locus sequence data (Knowles and Carstens, 2007). Concurrently, new statistical methods assess “species” limits, recognizing taxonomic or systematic arrangements that hold strongest support for actual relationships among distinct biological units (reviewed in Carstens et al., 2013). Both historical and ongoing gene flow between related species (e.g., Sullivan et al., 2014), however, present additional challenges to molecular systematic assessments.

The evolutionary history of North American red squirrels (genus *Tamiasciurus* Trouessart 1880) reflects the influence of long-term Quaternary climate cycles on continental patterns of diversification and community change (Arbogast and Kenagy, 2001). Periodic allopatry and divergence has been followed by episodes of reconnection and gene flow as climate cycled through glacial phases (Chavez et al., 2014). Simultaneously, some red squirrel populations, particularly peripheral ones, likely remained stable and isolated through extended glacial cycles. As such, systematic relationships, species limits, and taxonomy within this group remain tentative. Thorington and Hoffmann (2005) recognized three species: *Tamiasciurus douglasii* Bachman, 1839, distributed in forests from southern California and the Sierra Nevada northward through the Cascade and Coastal ranges; *Tamiasciurus hudsonicus* Erxleben, 1777, distributed across most of North America north of Mexico including multiple peripheral populations; and *Tamiasciurus mearnsi* Townsend, 1897, restricted within Baja California Norte, Mexico. Previous mitochondrial DNA (mtDNA) assessments (Arbogast et al., 2001; Barber, 2007) documented several divergent lineages, recounting a complex biogeographic history. Across these species, *T. hudsonicus* is paraphyletic with respect to *T. douglasii*, while *T. mearnsi* is minimally divergent from *T. douglasii* (Arbogast et al., 2001; Barber, 2007). In the American Southwest, populations of *T. hudsonicus*, including those on sky islands, potentially represent a third species (*Tamiasciurus fremonti*), based on genetic (Arbogast et al., 2001), behavioral, and morphological differences (Findley, 1961; Hardy, 1950). More recently, Chavez et al. (2014) largely corroborated these relationships with nuclear data, although mito-nuclear discord indicated historic introgression between populations of the Pacific Northwest. However, to date, no study has explicitly tested the validity of systematic relationships using a phylogenetic approach.

The North American red squirrel complex constitutes an excellent system for testing diversification hypotheses because their complex biogeographic history includes both physical (insular) and ecogeographic isolation of lineages, paraphyly, and gene flow. We present a multi-locus phylogenetic assessment of North American red squirrels and test three alternative hypotheses based on previous studies (Fig. S1). In the first hypothesis (H1 – Fig. 1) we apply the current taxonomy (Thorington and Hoffmann, 2005) of three species (*T. douglasii*, *T. hudsonicus*, and *T. mearnsi*). Considering that the specific status of *T. mearnsi* is tenuous (Arbogast et al., 2001), we test a second hypothesis (H2 – Fig. 2) proposed by Arbogast et al. (2001) that also posits three differently circumscribed species: *T. douglasii* (inclusive of *T. mearnsi*), *T. fremonti*, and *T. hudsonicus* (includes all currently recognized lineages except putative *T. fremonti*). Finally, we consider a third hypothesis (H3 – Fig. 3) treating nine well-supported mtDNA lineages as terminal taxa, all documented across independent studies (Arbogast et al., 2001; Barber, 2007; Chavez et al., 2014; this study). These three hypotheses require few assumptions and reflect realistic alternative relationships based on previous literature. Following genetic-based species delimitation, we incorporate niche modeling to assess geographic stability through time and ecological divergence among taxa. Finally, we use results from both multiple loci and climatic data to explore evolutionary complexity across North America, demonstrating different modes of speciation within this complex that are reflected more broadly among mammals. Critically, for forest ecosystem management, we identify areas that sustain key dynamic evolutionary processes as well as isolated populations needing further evaluation.

2. Methods

2.1. Study area, sampling and sequencing

We obtained specimens from museum collections that maximize geographic sampling through the known range of red squirrels (Fig. 1, 3), including samples from all species and localities coincident with the geographic distribution of 22 of 28 subspecies (Hall, 1981). We used 252 specimens (Appendix A) including 40 *T. douglasii*, 209 *T. hudsonicus*, and 3 *T. mearnsi* from 109 general collecting localities (Fig. 3; 153 specific localities; Appendix A). Sequences for 123 individuals for at least 1 gene were retrieved from GenBank. We then sequenced 129 more specimens for 1 to 4 independent loci where a subset of 60 samples (including 34 specimens from Chavez et al., 2014) were represented by both mtDNA and nuDNA. In total we obtained 3,192 bp of sequence data across all loci. Datasets included 771 base pairs (bp) of mtDNA cytochrome-*b* gene (Cytb; $n = 252$), 449 bp of the glutamate decarboxylase 2 gene (intron 1; GAD2; $n=60$), 666 bp of the ganglioside-induced differentiation-associated protein 1 gene (intron 1; GDAP1; $n=58$), 630 bp of the UPF0351 protein C9orf32 gene (intron 1; METTL11A; $n=59$), and 676 bp of the prolyl 4-hydroxylase alpha-2 subunit precursor gene (intron 1; P4HA2; $n=60$).

We extracted genomic DNA from tissue (stored at -80°C or ethanol preserved) through standard salt extraction, followed by PCR amplification and cycle sequencing with ABI BigDye v3.1 (Applied Biosystems, Foster City, California) following protocols of Platt et al. (2007). Primers for Cytb were MVZ05/MVZ16 (Smith and Patton, 1993). Other primers are listed in Chavez et al. (2014). We conducted automated sequencing of complimentary strands using the Applied Biosystems 3110 DNA sequencer at the University of New Mexico.

We edited and aligned sequences in Sequencher v4.8 (Genecodes, Ann Arbor, Michigan) and verified visually. We translated protein-coding sequences for Cytb to amino acids and examined them for internal stop codons, rates of transition/transversion changes, and relative 1st, 2nd, and 3rd position changes in codons that might indicate a Numt. None was found. We compared complementary strands of DNA and deposited contiguous sequences in GenBank (Appendix A). We inferred alleles of heterozygotes using Phase (Stephens et al., 2001), by accepting results with a probability > 90%, using a burn-in of 10,000 iterations and a run length of 10,000 iterations (McCormack et al., 2011). Among 5 runs, we retained results from the run with best goodness-of-fit to an approximate coalescent model, resulting in 2 nuclear haplotype sequences or alleles per individual, using only 1 allele per individual (chosen randomly) for analyses.

2.2. Multi-locus gene trees and evolutionary rates

We estimated independent genealogies and haplotype networks for all loci. We determined DNA substitution models for each locus using MrModeltest v2.3 (Nylander, 2004) under the Akaike Information Criterion. We performed Markov chain Monte Carlo (MCMC) searches in MrBayes v3.1 (Ronquist and Huelsenbeck, 2003), incorporating the short-branch method (Marshall, 2010). Stationarity of MCMC runs was assessed in Tracer v1.5.0 (Rambaut and Drummond, 2007). Phylograms were midpoint rooted and visualized with posterior probabilities >0.75 included in FigTree v1.3.1 (Rambaut, 2009). For all loci, we assigned haplotypes using DnaSP v5 (Librado and Rozas, 2009). Median-joining haplotype networks for all loci were constructed in Network v4.610 (Bandelt et al., 1999). To assess the timing of

cladogenesis and demographic changes among lineages, we calculated average and lineage-specific evolutionary substitution rates utilizing a time to coalescence of the entire genus as estimated by Chavez et al. (2014). In BEAST v1.6.1 (Drummond and Rambaut, 2007), we produced a chronogram based on the Cytb locus by fixing the time to most recent common ancestor ($T_{MRC A}$) for all samples at 0.4229 Ma (normal distribution, standard deviation 0.1; Chavez et al., 2014), applying a relaxed clock: uncorrelated log-normal, and set the substitution rate to be estimated. We used the HKY+G model of substitution, empirical base frequencies, partitioned into three codon positions, and a constant population size tree prior. Three separate runs of 100 million iterations were run, sampling trees to file every 1,000 iterations. To retrieve mean lineage-specific rates for each nominal lineage recovered from the Cytb chronogram, we weighted all branch specific rates by branch length (Hope et al. 2014b) according to the equation $\sum (\text{rate} \times \text{length}) / \sum (\text{length})$.

2.3. Mitochondrial diversity and demographics

Genetic diversity analyses for Cytb used all available red squirrel sequences, performing analyses separately for major lineage assignments (Table 1). In DnaSP, for each group we assessed genetic diversity including number of segregating sites (S), haplotypes (h), haplotype diversity (Hd), nucleotide diversity (π), and tested for demographic expansion by calculating Tajima's D (Tajima, 1989) and $R2$ (Ramos-Onsins and Rozas, 2002).

To calculate population size change through time, we implemented the coalescent Bayesian Skyline Plot (BSP) tree prior in BEAST. Results of Bayes factor tests to determine the best clock model indicated no significant difference between strict and relaxed clocks for any

lineage, so we utilized a strict clock, and ran MCMC analyses (length of chain = 100 M logging trees every 1,000), assigning lineage-specific rates calculated from the Cytb chronogram.

2.4. Systematic scenarios

Using a model comparison technique that applies Bayes Factors ($2\ln Bf$) or Akaike information criterion through Markov chain Monte Carlo analysis (AICM) to test between alternative scenarios of species delimitation (Aydin et al., 2014; Grummer et al., 2014) allows flexibility for alternative hypotheses and is robust to systems where historical introgression is suspected (Grummer et al., 2014). We performed model comparison tests of alternative systematic hypotheses among designated lineages of North American red squirrels considering all loci within a species-tree framework implemented using *BEAST (Heled and Drummond, 2010). Because *BEAST accommodates different numbers of gene copies for each taxon, we included all samples per lineage that were sequenced for more than a single locus.

We performed Bayes Factor and AICM pairwise tests of alternative hypotheses (Table 2; Fig. S1). All hypotheses considered samples ($n = 60$) representing various combinations of the nine well-supported monophyletic mtDNA (Cytb) lineages: Two widespread lineages (Continental and Northwestern), a Western lineage (combining *T. douglasii* and Baja California samples unless otherwise noted), Vancouver Island, Central Rockies, North Pacific Coast, Sacramento Mountains (NM), Southern Rockies, and Southwestern Sky Islands (Fig. 1). Despite a high likelihood of ongoing gene flow within and among certain lineages, all were included to assess relative nodal support for clade relationships considering multiple loci under three different hypotheses. Hypothesis H1 considered the currently assigned taxonomy (*T. douglasii*, *T. hudsonicus*, and *T. mearnsi*; Fig. 2). Hypothesis H2 considered three species hypothesized by

Arbogast et al. (2001): *T. douglasii* (including *T. mearnsi*), *T. fremonti* (Southern Rockies, Sacramento Mountains, and Southwestern Sky Island), and *T. hudsonicus* (Continental, Northwestern, Southeast Alaska, Vancouver Island and Central Rockies). Hypothesis H3 considered all Cytb lineages as independent “species”. All hypotheses were compared statistically using both all loci and only nuclear loci.

We produced a separate traits file indicating lineage assignments for all samples under each hypothesis. In BEAUti, for each hypothesis, we unlinked datasets for each locus across all partitions, assigned substitution models by locus, and applied empirical base frequencies. We applied a relaxed clock: uncorrelated log-normal to Cytb as the uclsd standard deviation of the rate for this gene when analyzed independently significantly deviated from zero, but assigned a strict clock to nuclear loci to reflect lower divergence over the evolutionary timeframe investigated. We chose a Yule tree prior with piecewise linear and constant root and assigned proper ploidy to each locus. Two independent runs for each hypothesis used MCMC chains of 100 million generations sampling trees every 10,000.

In Tracer, we ran model comparisons to identify the hypothesis that best explained the data considering marginal likelihood estimates of competing hypotheses using both $2\ln Bf$ and $2AICM$ scores. Scores were ranked for each hypothesis only with respect to the “best” hypothesis (that with the marginal likelihood estimate closest to zero). As described in Grummer et al. (2014), and based on the recommendations of Kass and Raftery (1995), the strength of support of $2\ln Bf$ for the best hypothesis (labeled as N/A) compared with competing hypotheses was as follows: 0-2 means no significant support; 2-6 means positive support for the best hypothesis (N/A) over this competing scenario; 6-10 means strong support for the best hypothesis; and >10 means decisive rejection of the competing hypothesis compared with the

best hypothesis. As in Grummer et al. (2014), we recognized distinct lineages from hypotheses (topologies) where $2\ln Bf$ support for the leading topology was >10 in pairwise comparisons with competing hypotheses. Similarly for AICM which equates using two times the sample mean log likelihood, we use a cutoff of >20 for $2AICM$ (multiplied by 2 for consistency with Bayes Factor methods).

2.5. Bioclimatic envelope modeling

We used bioclimatic envelope modeling (BEM, Araújo and Peterson, 2012) to characterize occupied distributions and reconstruct paleodistributions of each species. We used these niche-based tests as an independent perspective of ecological divergence among putative species using a three-step approach: 1) reconstruct contemporary distributions and project (transfer) the final model to ancestral climate conditions (i.e., mid-Holocene and LGM); 2) assess spatial overlap of predicted distributions for all periods; and 3) conduct niche identity tests using geographic-based approaches (Warren et al., 2008, 2010).

2.5.1. Occurrence and environmental data

BEMs are often constructed from two forms of data: occurrence records and environmental layers. We downloaded all available records of North American red squirrels from VertNet (accessed February 2015) and conducted a series of screening and filtering steps to reduce bias. We removed all samples with a georeferenced error $>5\text{km}$, those that failed to provide an estimate of error, and those that were outside the known distribution of *Tamiasciurus*. We also removed samples prior to 1940. To reduce spatial autocorrelation, we used a custom

aggregated reduction technique (Fourcade et al., 2014; Warren et al., 2014) by retaining a single random sample within 10km² radius and discarding other spatially redundant locations.

To assess lineage-based distributional overlap and test for recent niche divergence among evolutionarily independent groups we used mtDNA haplotypes to quantify areas occupied by lineages. First, we added a minimum convex polygon in ArcGIS v10.2 to each lineage-based set of points and applied a 200 km buffer to assign all points within each polygon to respective mtDNA lineages. We then reassigned mtDNA samples to putative species identified from the best model in tests of evolutionary hypotheses.

The geographic extent of species is an important consideration when using correlative models (Barve et al., 2011). However, the total extent of *Tamiasciurus* is likely too broad considering our goals of pairwise tests between geographically limited species. Based on the geographic distribution of putative species, we divided the study area using the three regional clades (systematic hypothesis H2; Table 1; Fig. S1) and delimited their extents to 100 km beyond the N/S and E/W extremes of each distribution.

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

We used 19 bioclimatic (temperature and precipitation) variables from the WorldClim database to quantify ecological tolerances of each lineage. These landscape-level data are often useful for assessing both current and paleodistributions (Malaney and Cook, 2013; Waltari and Guralnick, 2009; Waltari et al., 2007), and also for assessing Grinnellian-based niche divergence (McCormack et al., 2010). Then, we identified tightly correlated (≥ 0.80) and removed redundant variables, favoring those more temporally inclusive (e.g., bio11 – mean temperature of the coldest quarter versus bio6 – minimum temperature of the coldest month) or relevant to squirrel ecology (e.g., cone mast is often linked to seasonal conditions; bio15 – precipitation seasonality).

2.5.2. Correlative modeling

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

0.01 to 5.0 at 0.5 intervals. In sum, we tested 55 alternate models for each species (5 FCs x 11 RMs), identifying optimum model settings using corrected Akaike information criterion (AICc).

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

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

2.5.3. Paleodistributions

We reconstructed paleodistributions of individual lineages to understand the history of geographic distributions, frequently yielding insights of ancestral areas as well as regions of historic admixture (Pelletier et al., 2015). Specifically, we projected predictions for the mid-Holocene (contemporary models; i.e., Younger Dryas) and late-Pleistocene (i.e., Last Glacial Maximum – LGM). Considering many North American mammals have shifted distributions

(Malaney and Cook, 2013; Lessa et al., 2003), we expanded the spatial extents through past timeframes by 500 km to include areas nearby that might have been historically occupied by ancestral populations. Further, we validated paleodistribution reconstructions with fossil records obtained from FaunMap database (accessed March, 2015). We then assessed spatial overlap among clades using techniques analogous to contemporary overlap measurements.

2.5.4. Model-based niche divergence

To assess the degree of niche overlap among the major clades, we used the G-space niche-identity test using the D statistic (Schoener, 1968; Warren et al., 2008), which ranges from 0 (discordant) to 1.0 (identical). The niche-identity test is used to assess whether measured overlap is significantly different (one-tailed test) from an underlying null expectation. This approach randomized samples from each lineage pair using 100 pseudoreplicates for each pairwise test.

3. Results

3.1. Gene trees and evolutionary rates

A previously calculated age for the T_{MRCA} of all North American red squirrels of 0.4229 Myr (Chavez et al., 2014) provided an average mutation rate of 5.23% per million years (10.46% divergence rate) for the Cytb dataset (Table 1). Substitution rates ranging from 5.03-5.46% per million years were recovered for nominal lineages within the red squirrel Cytb genealogy (Table 1). The chronogram indicates nine well-supported, spatially-structured lineages (Fig. 3). Two

lineages exhibit broad distributions, often within previously glaciated areas, indicative of post-glacial population expansion. The Continental lineage is broadly distributed across North America whereas the Northwestern lineage is coincident with the Rocky Mountain cordillera. These two lineages together form a well-supported clade distinct from all other North American red squirrels. Other well-supported lineages include the Central Rockies with a limited distribution within the inland forests of Idaho, western Montana, and eastern Oregon. The North Pacific Coast (NPC) lineage is only distributed on islands of the Alexander Archipelago (Betton, Grant, Gravina, Kuiu, Mitkof, Revillagigedo, and Zarembo islands), and in British Columbia on Calvert and Swindle islands (Fig. 3). Mitochondrial signatures indicative of both NPC and Northwestern lineages occur on Gravina and Revillagigedo islands of the southern Alexander Archipelago. The Northwestern lineage also occurs on Admiralty (introduced), Tatoosh, and Wrangell islands. The Western mtDNA lineage is paraphyletic with respect to Vancouver Island (due to recognized mito-nuclear discord) and is distributed from Baja California (*T. mearnsi*) north to southern British Columbia. Finally, substructure within the well-supported Southwestern clade includes lineages associated with the Sacramento Mountains of New Mexico, the Southern Rockies of Colorado and northern New Mexico, and two paraphyletic lineages associated with the Southwestern Sky Island archipelago (Fig 3).

The initial dichotomy within the Cytb chronogram is dated to ~400 ka. Continental and Northwestern lineages were estimated to have split at 200 ka. Coalescence of all Southwestern lineages is coincident with the Illinoian glacial period (~160 ka) although initial divergence from other squirrel lineages is dated >250 ka. Coalescence of most terminal lineages is <100 ka.

Independent nuclear genealogies were less well resolved although Western haplotypes were consistently private across all loci (Fig. S2). All well-supported mtDNA lineages exhibited

at least one private nuclear haplotype. Continental and Northwestern haplotypes were closely related, as were those of all Southwestern lineages. Within the Southwestern Sky Island lineages, the federally endangered Mt. Graham (Pinaleno Mountains) population exhibited private haplotypes over all loci except one (Fig. 3; Fig. S2).

3.2. Diversity and demographics

Mitochondrial DNA haplotype and nucleotide diversity were lowest for the Vancouver Island lineage and highest for the Western lineage (Table 1). Within the Continental lineage, samples collected from western and eastern North America were divergent (Fig. 3). Only Continental and Northwestern lineages exhibited consistent signals of demographic expansion over multiple tests, a finding also supported by star network topologies (Table 1; Fig. S2). BSPs for these lineages indicated the most dramatic population growth coincident with post-glacial expansion (Fig. S3). BSPs for both the NPC and Western lineages also indicated significant population growth but non-significant demographic expansion, suggesting *in situ* growth through the late-glacial and Holocene timeframes. Skyline changes for Central Rockies, Sacramento Mountains, Southern Rockies and combined Southwestern Sky Islands lineages failed to reach the 95% highest probability distributions indicating minimal population growth (Fig. S3). Uncorrected pairwise mtDNA genetic distances between the 9 major lineages ranged from 0.84-3.48% sequence divergence with highest divergence between Northwestern and Southern Rockies lineages.

3.3. Systematic hypotheses

All species-tree topologies strongly supported the presence of multiple species and lack of reciprocal monophyly of *T. hudsonicus* as currently recognized. Each systematic hypothesis is reported in conjunction with the geographic ranges of nominal taxa according to respective hypotheses shown on maps, and these phylogenies reflect the results of species-tree analyses, including relative branch lengths, empirical topologies, and posterior probabilities at nodes (Fig. 1, 2, 3, S1). When Southwestern lineages were considered as a separate species, they were consistently less divergent from *T. douglasii* than from the remaining lineages of *T. hudsonicus*, although low posterior probabilities coupled with basal coalescences support the possibility of a radiation of three species through initial fragmentation, dating to the pre-Illinoian (~400 ka). Based on all loci, Bayes Factor model comparisons decisively supported Hypothesis H2 as the best systematic arrangement (Fig. S1; Table 2). AICM comparisons using only nuclear loci also supported H2 above other scenarios. Hypothesis H2 provides highest likelihood for true species relationships, considering all species tree estimates, and reflects three species. As also reflected by divergent ecological associations, we recognize the Douglas squirrel, *T. douglasii* consisting of the Western lineage, the southwestern red squirrel, *T. fremonti* Audubon and Bachman, 1853, consisting of the Southern Rockies, Sacramento Mountains, and Southwestern Sky Islands lineages, and the American red squirrel, *T. hudsonicus* consisting of the remaining 5 lineages.

3.4. Bioclimatic envelope modeling

3.4.1. Niche variables and occurrence points

We obtained 20 niche variables that exhibited varying levels of importance and correlation across the study region (Table S1). For occurrence points, we obtained a total of 1,160 *T. douglasii* (including *T. mearnsi*), and 3,413 *T. hudsonicus* records. After processing (screening, partitioning, thinning, and clipping), we identified 279 *T. douglasii* and 740 *T. hudsonicus* records for reconstructing distributions and conducting niche tests. We then assigned 59 of the *T. hudsonicus* records from the Southwest as *T. fremonti* using genetic data as a guide.

3.4.2. Model Performance

We found that optimized settings reflected a best-fit for our study system using model selection that increased the regularization penalty, simplified features, and reduced the overall dimensionality of final models. MaxEnt performed well with mean training AUC >0.85 for all analyses, often a positive indication of optimized models. Randomly withheld test data were predicted significantly better than random with AUC >0.90 for all analyses (Table 3) and AUC standard deviations all <0.05. Threshold-dependent approaches yielded results with low omissions and significantly better omission rates from threshold-dependent predictions than from random predictions.

3.4.3. Current distributions and overlap

Distributions of North American red squirrels were best predicted ($\geq 10\%$ variable importance) by three variables for each species and six variables total (Table 3). For *T. douglasii*, >73% of the variation was accounted for by winter precipitation (bio19 = 39.7%), precipitation seasonality (bio15 = 20.0%), and isothermality (bio3 = 13.5%). For *T. hudsonicus*, over 50% variable importance was due to isothermality (31.3%) and winter precipitation (21.5%), with

mean annual temperature (bio1 = 12.7%) the third most important, totaling 79.3%. The three most important variables (65.5%) for *T. fremonti* included average summer temperature (bio5 = 50.0%), temperature seasonality (bio4 = 13.5%), and like other species, winter precipitation (15.8%).

Current distributions predicted for each species overlap (Fig. 2), with *T. hudsonicus* sharing predicted range with *T. douglasii* in the Pacific Northwest (overlap = 77,754 km²) and with *T. fremonti* in the Southern Rocky Mountains (overlap = 71,408 km²).

3.4.4. Paleodistributions

We document distributional change for all species between the LGM, mid-Holocene, and present. Even after applying thresholds of currently occupied distributions and projecting to historical conditions (i.e., mid-Holocene and LGM), most fossil records (>80%, Fig. S4-S6) were accurately predicted. Predicted ancestral areas (LGM) for *T. douglasii* were centered on northern California (Fig S4) with isolated regions in southern California and northern Baja, and although generally further south, total predicted area was only slightly larger than mid-Holocene and today. Predicted distributions shifted to higher elevations and more northern latitudes during the Holocene and are most restricted in total range at present. For *T. fremonti*, LGM distributions were projected to be >2x the extent of present distributions, with a retraction to higher elevations during the Holocene rather than a shift across latitude (Fig. S5). For *T. hudsonicus*, latitudinal shifts appear to have been the dominant change from LGM to mid-Holocene and today (Fig. S6). Potential LGM occupation included ancestral regions along the North Pacific Coast and within eastern Beringia (Alaska).

3.4.5. Stable regions for squirrels.

By overlaying predictions from all models, we identified stable regions for all three species, although eco-geographic stability varied. For *T. douglasii*, three regions have remained relatively stable between the LGM and today, including medium elevations in the Sierra Nevada of eastern California, the Coast Ranges in northwestern California and Oregon, and high elevations of southern California and Baja Mexico. The total combined stable area for *T. douglasii* is ca. 69,000 km² (20.8% of the present distribution). For *T. fremonti*, we detected three temporally stable areas largely coincident in extent with the current range, including the Rocky Mountains of Colorado and New Mexico, the Sacramento Mountains in southeastern New Mexico, and the Mogollon Mountains of New Mexico and Arizona. When combined, these stable areas for *T. fremonti* totaled ca. 111,000 km² (62.9% of the present distribution). In contrast to *T. douglasii* and *T. fremonti*, *T. hudsonicus* had few consistently occupied regions between the LGM and today totaling <80,000 km², a very small proportion of the total potential range in any timeframe (2% of present distribution). Whereas historic areas are largely coincident with temporally stable regions for *T. fremonti*, and slightly less so for *T. douglasii*, past predictions for *T. hudsonicus* are largely independent of interglacial (recent) distributions.

3.4.6. Niche identity

Despite large areas of potential distributional overlap between species (Fig. 2), the background test of niche identity indicated that each squirrel species occupies distinct ecospace when compared to expected niche overlap from pooled samples. All three species have non-identical niches and distributions (Schoener's *D* values ≤ 0.50).

4. Discussion

This comprehensive phylogenetic assessment of the North American red squirrels provides a more complex interpretation of speciation within these forest obligate species and characterizes processes that also influence other species of this increasingly fragmented community. Congruence across species hints at fundamental ecogeographic variability on a continental scale that has predictable evolutionary consequences. Three main dynamics emerge from our analyses.

The first dynamic reflects evolutionary complexity across North America among forest dwelling species, with a dichotomy that contrasts minimal evolutionary structure in eastern North America (a single widespread lineage for red squirrels) with comparatively complex structure in the west (three species and multiple lineages). This east/west split reflects the Mississippi discontinuity as now reported across multiple other taxa (e.g., Soltis et al., 2006), with the Great Plains Ecoregion serving as a major barrier between east and west, and a gradient of increasing genetic complexity found within temperate forested biomes from east to west. Within the west, three geographic centers of diversity exist as reflected by *T. douglasii* in westernmost generally mesic forests, *T. fremonti* distributed through the Southwest, and *T. hudsonicus* with an extensive distribution through remaining interconnected forests.

The second dynamic reflects regional variability in spatial responses of species to long-term climate trends. Both genetic demographic analyses and niche predictions indicate relative stability through the southwest region with over 60% of current niche for *T. fremonti* consistently predicted as suitable through all recent (Late Quaternary) climate phases. By contrast, all lineages within *T. hudsonicus* except for the Central Rockies exhibit significant signals of recent demographic expansion and virtually no persistent niche space through timeframes. Dramatically

different histories for *T. hudsonicus* and *T. fremonti* reflect alternate modes of diversification. In the southwest, *T. fremonti* is the product of ancient peripatric isolation and subsequent divergence accompanied by periodic fragmentation/reconnection at local scales. In contrast, glacial and interglacial phase distributions of *T. hudsonicus* are non-overlapping, and such vast spatial displacement likely reflects repeated population extirpations as leading edge expansions and contractions predominate through time, resulting in species-wide range shifts. Across the continent-wide distribution of *T. hudsonicus*, broad-scale fragmentation and range contractions with subsequent differentiation in multiple isolated areas, generated high overall genetic diversity. *Tamiasciurus douglasii* exhibits an intermediate history with reduced glacial-interglacial phase displacement and medium-term persistence of populations.

The third major dynamic is relatively rapid radiation of three species despite repeated and ongoing gene flow, reflecting again high regional complexity in western North America and climate phase expansion-contraction dynamics. Similar episodes of rapid diversification are being uncovered in other forest-dwelling mammals with shared geographic distributions (e.g., Shafer et al., 2010). For instance, both the *Sorex cinereus* complex and *Sorex palustris* complex exhibit longitudinally distributed lineages that reflect a Mississippi discontinuity, rapid speciation, high genetic complexity in the west, and ongoing gene flow among species (Hope et al., 2012, 2014a). Similarities in diversification processes are also found among other sciurids. Chipmunks (multiple genera; Patterson and Norris, 2015) have a deeper history of speciation than red squirrels and provide a model for examining how gene flow has impacted the formation of species (Good et al., 2008; Sullivan et al., 2014). In addition, both intra- and inter-specific distributional histories among chipmunks reflect low diversity in eastern North America (*Tamias striatus* and a single lineage of *Neotamias minimus*), as opposed to 22 species that have arisen in

western regions (Hall, 1981) with extensive hybridization at range boundaries. Marmots (Genus *Marmota*) also exhibit inter-specific gene flow (Kerhoulas et al., 2015) and the history of diversification of multiple genera of ground squirrels has been accompanied by periodic episodes of introgressive hybridization (Helgen et al., 2009). Recent speciation, coupled with replicated episodes of gene flow, both spatially and temporally, among North American red squirrels should provide novel perspectives on the formation of Nearctic species (Abbott et al., 2013).

4.2. Systematic relationships

Given the extensive range of red squirrels, a comprehensive assessment of systematic relationships and evolutionary history within *Tamiasciurus* has been hampered by limited geographic sampling and associated multi-locus genetic perspectives. Hall (1981) identified substantial morphological variation, recognizing 4 subspecies within *T. douglasii* and 25 subspecies within *T. hudsonicus*. Subsequently, Lindsay (1981) raised *T. d. mearnsi* to species status based on distinctive morphospace among the three species, but lacking range-wide sampling. The first geographically widespread genetic assessment based on a single locus (Arbogast et al., 2001) found no genetic divergence between *T. douglasii* and *T. mearnsi* and instead recovered three independent lineages as western (combined *T. douglasii* and *T. mearnsi*), eastern (all *T. hudsonicus* except for Southwestern populations) and Southwestern (Arizona, New Mexico, and southern Colorado populations of *T. hudsonicus*). Additional mtDNA sequence data identified a paraphyletic relationship of *T. hudsonicus* with respect to *T. douglasii* (including *T. mearnsi*) and reported a distinct lineage located in the Alexander Archipelago of Southeast Alaska (Barber, 2007). Most recently, multiple nuclear loci suggest refugial isolation

on Vancouver Island (Chavez et al., 2014), and indicate a complex history of diversification including gene flow, as has been found elsewhere between *T. hudsonicus* and *T. douglasii* (Chavez et al., 2011).

Considering differentiation across multiple loci, climatic niche divergence, and previous morphological evidence, the taxonomy of red squirrels should be updated. The isolated population occurring in Baja California (previously *T. mearnsii*) is minimally divergent from the remainder of the Western lineage over all loci examined; however, a combined western clade representing the Douglas squirrel (*T. douglasii* and *T. mearnsii*) forms one of three species that are well supported over all analyses. The other two species are the southwestern red squirrel (*T. fremonti*; including Southern Rockies, Sacramento Mountains, and Southwestern Sky Island lineages), and the broadly distributed American red squirrel (*T. hudsonicus*) that includes five mtDNA lineages (Continental, Northwestern, Central Rockies, North Pacific Coast, and Vancouver Island). These three species better reflect the evolutionary history and ecology of boreal forest obligates and should improve management and conservation planning.

Although our methods allow for more rigorous representation of systematic relationships within *Tamiasciurus*, the revised taxonomy does not fully account for either incomplete lineage sorting or recent and historic admixture. In this regard, a closer appraisal of the nuclear evidence compared with matrilineal inheritance is informative (Fahey et al., 2014; Hung et al., 2012). Three mtDNA lineages are highly supported as unique among red squirrels, including Vancouver Island, Central Rockies, and NPC lineages; however, all exhibit strong evidence of introgression across different temporal scales. Vancouver Island has fixed mito-nuclear discord where the mitochondria are indicative of *T. douglasii*, but only three haplotypes among nuclear loci examined are private (most being shared with *T. hudsonicus*). Given the single mtDNA lineage

present on Vancouver Island, we concur with Chavez et al. (2014) that this population represents a refugial lineage that persisted through at least part of the last glacial, resulting from a small founder event by hybrid individuals of *T. douglasii*/*T. hudsonicus*. North Pacific Coast and Central Rockies mtDNA lineages are more difficult to interpret. Both exhibit geographically restricted and private nuclear haplotypes, but also share haplotypes with other populations, particularly within *T. hudsonicus* and the Vancouver Island lineage.

4.3. Niche variability among species

Consistent with updated systematic relationships, modeled distributions are indicative of ecological differences among all three species. However, we detected two areas of geographic overlap coincident with regions of hybridization. Overlaps are likely the product of incomplete ecological divergence that includes a narrow set of environmental conditions. For example, although all three species have overall discordant niches ($D = 0.35-0.49$), they all appear to respond similarly to winter precipitation (either the most important or second most important variable), where each is predicted in areas with between 200 to 1000 mm of winter precipitation. Therefore, incomplete ecological differences are likely because of a combination of winter precipitation, temperature evenness, and seasonal variation in temperature and/or precipitation, at distributional peripheries. As a result, in the Rocky Mountains, temperature seasonality differs less between *T. hudsonicus* and *T. fremonti*, allowing these genetically and ecologically divergent species to overlap in peripheral areas. Similarly, in the Pacific Northwest, isothermality (temperature evenness) is most consistent where *T. douglasii* and *T. hudsonicus* meet, allowing them to occupy peripheral areas. There are likely also temporal differences in

spatial overlap. For example, the current overlap in the Rocky Mountains may have formed only recently, whereas repeated episodes of isolation and reconnection in Northwestern North America likely permitted greater gene flow in this area.

4.4. Ecology and conservation of regional forest ecosystems

Isolated forest-associated species often reflect classic dynamics of land-bridge island biogeography (Brown, 1971; Patterson, 1982). Because of their age, size, and isolation (e.g., Frey et al., 2007) genetic diversity and effective population size (both at present and historically) likely dictate risk of extirpation due to warming and drying climatic trends compounded with stochastic events (e.g., fire – Koprowski et al., 2005). Conservation of forest peripheral isolates is critical considering that many populations among forest-dwelling species: 1) occur in small, contracting patches of habitat (Holocene refugia), 2) contribute disproportionately to overall genetic diversity, and 3) often exhibit unique adaptations due to their occurrence at the edge of a species' range (Channell and Lomolino, 2000; Hampe and Petit, 2005; Koprowski, 2005).

A clearer understanding of the historical biogeography and systematics of these squirrels enables us to better assess conservation implications for peripheral isolates, due to comprehensive geographic sampling and integrated eco-evolutionary methods (Hope et al., 2013). Peripheral isolates experience high risk of extirpation for several reasons. Small population sizes lead to reduced genetic diversity, although differential selection may contribute to increased rates of divergence and potentially local adaptation (Losos and Ricklefs, 2009). Isolates are also at increased risk of stochastic events coupled with marginal ecological conditions, as has been exemplified by the Mt. Graham red squirrels in southern Arizona

(Koprowski et al., 2005) and human-mediated habitat fragmentation. Finally, if isolation persists for extended periods, then subsequent admixture or genetic “rescue” from other populations may dilute recently evolved adaptations, and erode or swamp unique diversity through introgression (Tallmon et al., 2004).

Several populations within North American red squirrel species are under increased threat of extinction due to peripheral isolation. Within *T. douglasii*, the Baja California populations (*T. d. mearnsi*) typify a peripheral isolate, occurring in marginal pine/fir forest habitat through a narrow altitudinal range, with low population density, diagnosable morphological features, and unique haplotypes (De Grammont and Cuarón, 2008). In addition, these populations exhibit unique behavioral and ecological adaptation that are not seen further north (e.g., Koprowski et al., 2006). Furthermore, due to the effects of warming and drying climate trends, continued habitat fragmentation (e.g., wildfire, logging, grazing), and competition from introduced eastern gray squirrels, this southernmost refugial population should be closely monitored (De Grammont and Cuarón, 2008).

Within *T. hudsonicus*, multiple peripheral isolates exist. The hybrid population distributed on Vancouver Island and multiple satellite islands (Fig. 3) is distinct from others along the NPC, indicating the role of multiple isolated refugia in the Pacific Northwest during the last glacial phase. The NPC lineage has an extensive distribution through Southeast Alaska and coastal British Columbia, with six island populations appearing to represent a pure NPC mitochondrial lineage, and three islands where multiple mitochondrial lineages occur. More comprehensive sampling of the vast Pacific coastal archipelagos is needed to clarify the spatial extent of coastal lineages (Fig. 3). Nuclear loci provide a more complex interpretation that may reflect recent or past hybridization or incomplete lineage sorting. The NPC harbors endemic

diversity over multiple plant and animal taxa reflecting a long and repeated history of colonization interspersed with isolation on islands due to raised sea levels during interglacials, and isolation in refugia due to ice sheet barriers during glacial phases (Cook et al., 2001). For example, mammals such as ermine (Dawson et al., 2014), wolves (Weckworth et al., 2010, 2015) and marten (Small et al., 2003) show similar patterns of differentiation in this coastal region, likely reflecting prolonged isolation from continental populations. Considering multiple human-mediated introductions of squirrels to individual islands in Southeast Alaska (MacDonald and Cook, 1996), and other human impacts such as old growth timber harvest, squirrels endemic to particular islands may be at high risk. A clearer understanding of the geographic extent of the NPC lineage is imperative, including the dynamics and history of population admixture between discrete lineages. Multiple island populations through this region should provide a lucrative system for investigating hybrid dynamics and processes of local adaptation using genomic methods.

Tamiasciurus fremonti is also represented by multiple insular populations. Of high conservation concern, the Mt. Graham population of *T. fremonti* in the Pinaleno Mountains of Arizona is here assessed for the first time as part of a range-wide genetic evaluation of the species in context with its sister species. We confirm that this population is genetically distinct across loci, but most similar to populations from the Kaibab Plateau in northern Arizona and the Gila region of western New Mexico. In addition to the Pinaleno Mountains, squirrels of the San Mateo Range in central New Mexico appear to be genetically divergent, but our sampling is limited to a single individual. Perhaps most surprisingly, squirrels from the Sacramento Mountains in south-central New Mexico constitute a highly divergent lineage within *T. fremonti*, dated to roughly the last interglacial period (130 ka) and coincident with initial diversification

among all lineages within the Southwest. Little evidence of demographic expansion among any lineages of *T. fremonti*, and distinct diversity coincident with isolated mountain ranges indicates a potentially extinction driven system in the future, if lack of connectivity is coupled with further degradation of montane forested habitats. Limited current gene flow between populations indicates that this species is of high conservation concern. Within the Sacramento Mountains, isolation may be exacerbated by their “high and dry” nature. Red squirrels in the Southwest are restricted to higher elevation conifer forests although they readily utilize blue spruce (*Picea pungens*) corridors along riparian systems. During the last glacial phase, spruce likely extended more widely along riparian corridors, facilitating connections between higher elevation forested areas through the Southwestern Sky Island Archipelago. However, no major riparian corridors now exist between the Sacramento Mountains and other sky islands. Niche predictions support relative isolation of this region through multiple timeframes (Fig. S5).

A significant contribution of a multi-locus and integrated niche-prediction approach to conservation phylogenetics is that, in addition to considering unique peripheral diversity, we can evaluate evolutionary potential within more complex regions where multiple zones of contact exist between sibling species through time (Sullivan et al., 2014). Our niche models predict two broad regions of overlap between the three species of red squirrels, centered over the Southern Rockies and the Pacific Northwest/Interior Highlands ecosystems. Nuclear loci suggest ongoing gene flow. However, considering multiple divergent matrilineages, these regions of contact are likely representative of recent (Holocene) admixture, with hybridization an ongoing dynamic process as climate and environments continue to shift (Hamilton and Miller, 2015; Taylor et al., 2015). Considering that hybridization may either be detrimental to parental species or promote both adaptive potential and diversification (Brennan et al., 2015), these regions of complex

contemporary interactions between species, in addition to peripheral populations, constitute critical areas for future research, and conservation-based management.

4.5. Refugia and important ancestral areas of occupation

This multi-locus assessment provides valuable insight for the temporal and spatial development of continent-wide ecological and genetic diversity. Because genetic coalescent estimates and demographics coupled with lineage-specific BEMs can highlight regional responses (Hope et al., 2015), red squirrels provide a basis for identifying common processes shaping the biogeography of forest communities. From niche predictions through time for *T. hudsonicus* (Fig. S7) and multiple distinct matrilineages, we infer several ancestral areas of allopatry during the last glacial phase, generally coincident with those discussed broadly for other taxa (e.g., Swenson & Howard 2005). Unique nuclear diversity within mitogroups, although minimal, is consistent with these predictions. Major areas were centered upon the eastern and western U.S. south of continental ice sheets and within Beringia, the former two being corroborated by fossil evidence (Kurtén and Anderson, 1980). The Beringian refugium (specifically considering the Northwestern matrilineage of *T. hudsonicus*) was previously hypothesized (Chavez et al., 2014). Our expanded genetic sampling indicates very low historic population size and recent rapid demographic expansion within this lineage (Fig. S3), but with high relative genetic diversity at highest latitudes. Coupled with the current pervasive distribution through Southeast Alaska (not considering human reintroductions), the evidence suggests an early bout of colonization through this region, and likely southward expansion out of Beringia. There is growing evidence for persistence of conifer trees within Beringia through the

last glacial phase (Anderson et al., 2006; Brubaker et al., 2005; Zazula et al., 2006), a strict requirement for persistence of these squirrels. In addition, red squirrels are currently widespread in Alaska, occurring within the northernmost conifers south of Anaktuvuk Pass at 68°N (Rausch, 1951).

The NPC lineage of *T. hudsonicus* exhibits relatively high nucleotide diversity and a broad distribution through this extensive archipelago, indicating persistence in one or more coastal refugia. *Tamiasciurus hudsonicus* joins a growing list of taxa for which evidence of refugial persistence exists, and reflecting long-term intact communities along the NPC (Cook and MacDonald, 2013). The mixed distribution of haplotypes coupled with niche predictions within both *T. douglasii*, and independently within mainland populations of *T. hudsonicus* suggests connectivity across much of their current range is recent and glacial phase ranges were more fragmented, particularly through western areas, supporting the concept of “refugia within refugia” (Shafer et al., 2010). The pattern is reversed in *T. fremonti*, in that the fragmented archipelago of sky-island populations represent current refugia considering many of these populations experienced much greater connectivity during the last glacial. However, niche predictions indicate effectively persistent isolation among the three major matrilineages of this species (Fig. S5 – LGM). The Sacramento Mountains were most isolated, while the Southern Rockies were partially isolated from sites further south due to alpine glaciers in southwestern Colorado, a significant set of barriers also hypothesized for other taxa (Barber 2007).

5. Conclusions

Comprehensive molecular phylogenetic and ecological niche assessment of a widespread genus through North American forest ecosystems provides new insights on speciation across the Nearctic by highlighting how species diverge in response to long-term environmental variability. The obligate association of red squirrels with forests, coupled with relatively recent differentiation accompanied by gene flow, provides perspective on evolutionary responses to climate trends across broadly co-distributed taxa. Specifically within North America, southern peripheral populations of forest species may reflect long-term persistence through glacial cycling in core habitat areas that is coupled with periodic elevational shifts. In other populations across the continent-wide range, periodic expansion or retraction to predictable refugial areas occur through latitudinal (and longitudinal) shifts as climate changes. In addition to focusing on the management of isolated and potentially shrinking peripheral populations, as contemporary environmental perturbations accelerate, it is critical that we also conserve highly dynamic regions of phylogeographic complexity within forested systems where evolutionary processes associated with gene exchange are influencing biodiversity in ways that are as yet much less clear.

Acknowledgements

Funding was provided by the Beringian Coevolution Project (NSF0415668 and 1258010). USDA Forest Service, and Sitka Sound Science Center. Special thanks to S. O. MacDonald for help with field sampling. Photo credits for the graphic abstract: *T. hudsonicus* - <http://www.public-domain-image.com/free-images/fauna-animals/squirrel/red-squirrel-on-tree-tamiasciurus-hudsonicus/attachment/red-squirrel-on-tree-tamiasciurus-hudsonicus> (Author:

776 Donna Dewhurst, U.S. Fish and Wildlife Service); *T. douglasii* - Wikipedia image by
777 *Tamiasciurus_douglasii_6021.JPG*: Walter Siegmund (talk) derivative work: Walter Siegmund
778 (*Tamiasciurus_douglasii_6021.JPG*) [CC BY-SA 3.0 ([http://creativecommons.org/licenses/by-](http://creativecommons.org/licenses/by-sa/3.0)
779 [sa/3.0](http://creativecommons.org/licenses/by-sa/3.0)) or GFDL (<http://www.gnu.org/copyleft/fdl.html>)], via Wikimedia Commons; *T. fremonti*
780 - <http://www.mountgraham.org/content/mount-graham-red-squirrel-claire-zugmeyer>. In addition
781 to all museum collections listed in Appendix A that provided tissue loans, we thank all
782 contributing research collections that reported red squirrel specimen records on VertNet at the
783 time of access for use in niche predictions.

784

REFERENCES

- Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J.E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., Butlin, R.K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S.H., Hermansen, J.S., Hewitt, G., Hudson, A.G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-Rodriguez, P., Möst, M., Mullen, S., Nichols, R., Nolte, A.W., Parisod, C., Pfennig, K., Rice, A.M., Ritchie, G., Seifert, B., Smadja, C.M., Stelkens, R., Szymura, J.M., Väinölä, R., Wolf, J.B.W., Zinner, D., 2013. Hybridization and speciation. *J. Evol. Biol.* 26, 229-246.
- Allen, C.D., Breshears, D.D., McDowell, N.G., 2015. On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene. *Ecosphere* 6, 129.
- Anderson, L.L., Hu, F.S., Nelson, D.M., Petit, R.J., Paige, K.N., 2006. Ice-age endurance: DNA evidence of a white spruce refugium in Alaska. *Proc. Nat. Acad. Sci. USA* 103, 12447-12450.
- Anderson, R.P., Gonzalez, I.J., 2011. Species-specific tuning increases robustness to sampling bias in model of species distributions: An implementation with Maxent. *Ecol. Model.* 222, 2796-2811.
- Anderson, R.P., Raza, A., 2010. The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *J. Biogeog.* 37, 1378-1393.
- Araújo, M.B., Peterson, A.T., 2012. Uses and misuses of bioclimatic envelope modeling. *Ecology* 93, 1527-1539.
- Arbogast, B.S., Kenagy, G.J., 2001. Comparative phylogeography as an integrative approach to historical biogeography. *J. Biogeog.* 28, 819-825.

809 Arbogast, B.S., Browne, R.A., Weigl, P.D., 2001. Evolutionary genetics and Pleistocene
 810 biogeography of North American tree squirrels (*Tamiasciurus*). J. Mammal. 82, 302-319.
 811 Atkinson, R.J., Rokas, A., Stone, G.N., 2007. Longitudinal patterns in species richness and
 812 genetic diversity in European oaks and oak gallwasps, in: Weiss, S., Ferrand, N. (Eds.),
 813 Phylogeography of southern European refugia, Springer, Netherlands, pp. 127-151.
 814 Aydin, Z., Marcussen, T., Ertekin, A.S., Oxelman, B., 2014. Marginal Likelihood Estimate
 815 Comparisons to Obtain Optimal Species Delimitations in *Silene* sect. *Cryptoneuræ*
 816 (Caryophyllaceae). PLoS ONE 9, e106990.
 817 Bandelt, H.-J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific
 818 phylogenies. Mol. Biol. Evol. 16, 37-48.
 819 Barber, B.R., 2007. Comparative phylogeography of diverse Rocky Mountain fauna. Ph.D.
 820 dissertation, University of Minnesota, pp. 1-95.
 821 Barbet- Massin, M., Jiguet, F., Albert, C.H., Thuiller, W., 2012. Selecting pseudo- absences
 822 for species distribution models: how, where and how many? Methods Ecol. Evol. 3, 327-338.
 823 Barve, N., Barve, V., Jimenez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T.,
 824 Soberón, J., Villalobos, F., 2011. The crucial role of the accessible area in ecological niche
 825 modeling and species distribution modeling. Ecol. Model. 222, 1810-1819.
 826 Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K., Das,
 827 I., 2007. Cryptic species as a window on diversity and conservation. Trends Ecol. Evol. 22,
 828 148-155.
 829 Brennan, A.C., Woodward, G., Seehausen, O., Muñoz-Fuentes, V., Moritz, C., Guelmami, A.,
 830 Abbott, R.J., Edelaar, P., 2015. Hybridization due to changing species distributions: adding

problems or solutions to conservation of biodiversity during global change? *Evol. Ecol. Res.*,
16, 475-491.

Brown, J.H., 1971. Mammals on mountaintops: nonequilibrium insular biogeography. *Am. Nat.*
105, 467-478.

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

Brubaker, L.B., Anderson, P.M., Edwards, M.E., Lozhkin, A.V., 2005. Beringia as a glacial
refugium for boreal trees and shrubs: new perspectives from mapped pollen data. *J. Biogeog.*
32, 833-848.

Burnham, K.P., Anderson, D.R., 2002. Model selection and multimodel inference: a practical
information-theoretic approach. Springer, New York.

Carstens, B.C., Pelletier, T.A., Reid, N.M., Satler, J.D., 2013. How to fail at species delimitation.
Mol. Ecol. 22, 4369-4383.

Channell, R., Lomolino, M.V., 2000. Dynamic biogeography and conservation of endangered
species. *Nature* 403, 84-86.

Chavez, A.S., Saltzberg, C.J., Kenagy, G.J., 2011. Genetic and phenotypic variation across a
hybrid zone between ecologically divergent tree squirrels (*Tamiasciurus*). *Mol. Ecol.* 20,
3350-3366.

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.

852 Cook, J.A., Bidlack, A.L., Conroy, C.J., Demboski, J.R., Fleming, M.A., Runck, A.M., Stone,
 853 K.D., MacDonald, S.O., 2001. A phylogeographic perspective on endemism in the Alexander
 854 Archipelago of southeast Alaska. *Biol. Conserv.* 97, 215-227.

855 Cook, J.A., Dawson, N.G., MacDonald, S.O., 2006. Conservation of highly fragmented systems:
 856 the north temperate Alexander Archipelago. *Biol. Conserv.* 133, 1-15.

857 Cook, J.A., MacDonald, S.O., 2013. Island life: Coming to grips with the insular nature of North
 858 Pacific Coastal Forests, in: Orians, G.H., Schoen, J.W. (Eds.), *Conservation of North Pacific*
 859 *Coastal Forests*, Univ. Washington Press, Seattle, pp. 19-42.

860 Coops, N.C., Waring, R.H., 2011. Estimating the vulnerability of fifteen tree species under
 861 changing climate in Northwest North America. *Ecol. Model.* 222, 2119-2129.

862 De Grammont, P.C. Cuarón, A., 2008. *Tamiasciurus mearnsi*. The IUCN Red List of threatened
 863 Species 2008: e.T21378A9276538.
 864 <http://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T21378A9276538.en> . Downloaded on 29
 865 October 2015.

866 Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees.
 867 *BMC Evol. Biol.* 7, 214.

868 Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J.,
 869 Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion,
 870 G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J.,
 871 Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S., Wisz, M.S.,
 872 Zimmermann, N.E., 2006. Novel methods improve prediction of species' distributions from
 873 occurrence data. *Ecography* 29, 129-151.

874 Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011. A statistical
875 explanation of MaxEnt for ecologists. *Div. Distrib.* 17, 43-57.

876 Fahey, A.L., Ricklefs, R.E., Dewoody, J.A., 2014. DNA- based approaches for evaluating
877 historical demography in terrestrial vertebrates. *Biol. J. Linn. Soc.* 112, 367-386.

878 Findley, J.S., 1961. Geographic variation in New Mexican chickarees. *J. Mammal.* 42, 313-322.

879 Fourcade, Y., Engler, J.O., Rödder, D., Secondi, J., 2014. Mapping Species Distributions with
880 Maxent Using a Geographically Biased Sample of Presence Data: A Performance Assessment
881 of Methods for Correcting Sampling Bias. *PloS ONE* 9, e97122.

882 Frey, J.K., Bogan, M.A., Yates, T.L., 2007. Mountaintop island age determines species richness
883 of boreal mammals in the American Southwest. *Ecography* 30, 231-240.

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

887 Gauthier, S., Bernier, P., Kuuluvainen, T., Shvidenko, A.Z., Schepaschenko, D.G., 2015. Boreal
888 forest health and global change. *Science* 349, 819-822.

889 Good, J.M., Hird, S., Reid, N., Demboski, J.R., Stepan, S.J., Martin-Nims, T.R., Sullivan, J.,
890 2008. Ancient hybridization and mitochondrial capture between two species of chipmunks.
891 *Mol. Ecol.* 17, 1313-1327.

892 Grummer, J.A., Bryson, R.W., Reeder, T.W., 2014. Species delimitation using Bayes factors:
893 simulations and application to the *Sceloporus scalaris* species group (Squamata:
894 Phrynosomatidae). *Syst. Biol.* 63, 119-133.

895 Guillera- Arroita, G., Lahoz- Monfort, J.J., Elith, J., Gordon, A., Kujala, H., Lentini, P.E.,
 896 McCarthy, M.A., Tingley, R., Wintle, B.A., 2015. Is my species distribution model fit for
 897 purpose? Matching data and models to applications. *Global Ecol. Biogeog.* 24, 276-292.
 898 Hall, E.R., 1981. *The Mammals of North America*, 2nd edn. The Blackburn Press, New Jersey.
 899 Hamilton, J.A., Miller, J.M., 2015. Adaptive introgression as a resource for management and
 900 genetic conservation in a changing climate. *Cons. Biol.* In Press.
 901 Hampe, A., Petit, R.J., 2005. Conserving biodiversity under climate change: the rear edge
 902 matters. *Ecol. Lett.* 8, 461-467.
 903 Hardy, R., 1950. A new tree squirrel from central Utah. *Proc. Biol. Soc. Wash.* 63, 13-14.
 904 Heled, J., Drummond, A.J., 2010. Bayesian inference of species trees from multilocus data. *Mol.*
 905 *Biol. Evol.* 27, 570-580.
 906 Helgen, K.M., Cole, F.R., Helgen, L.E., Wilson, D.E., 2009. Generic revision in the Holarctic
 907 ground squirrel genus *Spermophilus*. *J. Mamm.* 90, 270-305.
 908 Hewitt, G.M., 1999. Post- glacial re- colonization of European biota. *Biol. J. Linn. Soc.* 68, 87-
 909 112.
 910 Hewitt, G.M., 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405, 907-913.
 911 Hoberg, E.P., Galbreath, K.E., Cook, J.A., Kutz, S.J., Polley, L., 2012. Northern Host-Parasite
 912 Assemblages: History and Biogeography on the Borderlands of Episodic Climate and
 913 Environmental Transition. *Adv. Parasit.* 79, 1.
 914 Hope, A.G., Speer, K.A., Demboski, J.R., Talbot, S.L., Cook, J.A., 2012. A climate for
 915 speciation: rapid spatial diversification within the *Sorex cinereus* complex of shrews. *Mol.*
 916 *Phylogenet. Evol.* 64, 671-684.

917 Hope, A.G., Waltari, E., Payer, D.C., Cook, J.A., Talbot, S.L., 2013. Future distribution of
 918 tundra refugia in northern Alaska. *Nat. Clim. Change*, 3, 931-938.

919 Hope, A.G., Panter, N., Cook, J.A., Talbot, S.L., Nagorsen, D.W., 2014a Multilocus
 920 phylogeography and systematic revision of North American water shrews (genus: *Sorex*). *J.*
 921 *Mammal.* 95, 722-738.

922 Hope, A.G., Ho, S.Y.W., Malaney, J.L., Cook, J.A., Talbot, S.L., 2014b. Accounting for rate
 923 variation among lineages in comparative demographic analyses. *Evolution* 68, 2689-2700.

924 Hope, A.G., Waltari, E., Malaney, J.L., Payer, D.C., Cook, J.A., Talbot, S.L., 2015. Arctic
 925 biodiversity: increasing richness accompanies shrinking refugia for a cold-associated tundra
 926 fauna. *Ecosphere* 6, 159.

927 Hung, C.M., Drovetski, S.V., Zink, R.M., 2012. Multilocus coalescence analyses support a
 928 mtDNA-based phylogeographic history for a widespread Palearctic passerine bird, *Sitta*
 929 *europaea*. *Evolution* 66, 2850-2864.

930 Kass, R.E., Raftery, A.E., 1995. Bayes factors. *J. Am. Stat. Assoc.* 90, 773–795.

931 Kerhoulas, N.J., Gunderson, A.M., Olson, L.E., 2015. Complex history of isolation and gene
 932 flow in hoary, Olympic, and endangered Vancouver Island marmots. *J. Mamm.* 96, 810-826.

933 Knowles, L.L., Carstens, B.C., 2007. Delimiting species without monophyletic gene trees. *Syst.*
 934 *Biol.* 56, 887-895.

935 Koprowski, J.L., 2005. The response of tree squirrels to fragmentation: a review and synthesis.
 936 *Anim. Conserv.* 8, 369-376.

937 Koprowski, J.L., Alanen, M.I., Lynch, A.M., 2005. Nowhere to run and nowhere to hide:
 938 response of endemic Mt. Graham red squirrels to catastrophic forest damage. *Biol. Conserv.*
 939 126, 491-498.

940 Koprowski, J.L., Ramos, N., Pasch, B.S., Zugmeyer, C.A., 2006. Observations on the ecology of
 941 the endemic Mearns's squirrel (*Tamiasciurus mearnsi*). Southwest. Nat. 51, 426-430.
 942 Kurtén, B., Anderson, E., 1980. Pleistocene mammals of North America. Columbia Univ. Press,
 943 New York.
 944 Lessa, E.P., Cook, J.A., Patton, J.L., 2003. Genetic footprints of demographic expansion in North
 945 America, but not Amazonia, during the Late Quaternary. Proc. Nat. Acad. Sci. USA 100,
 946 10331-10334.
 947 Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA
 948 polymorphism data. Bioinformat. 25, 1451-1452.
 949 Lindsay, S.L., 1981. Taxonomic and biogeographic relationships of Baja California chickarees
 950 (*Tamiasciurus*). J. Mammal. 62, 673-682.
 951 Losos, J.B., Ricklefs, R.E., 2009. Adaptation and diversification on islands. Nature 457, 830-
 952 836.
 953 MacDonald, S.O., Cook, J.A., 1996. The land mammal fauna of Southeast Alaska. Can. Field
 954 Nat. 110, 571-598.
 955 Malaney, J.L., Cook, J.A., 2013. Using biogeographic history to inform conservation: the case
 956 of Preble's meadow jumping mouse. Mol. Ecol. 22, 6000-6017.
 957 Marshall, D.C., 2010. Cryptic failure of partitioned Bayesian phylogenetic analyses: lost in the
 958 land of long trees. Syst. Biol. 59, 108-117.
 959 McCormack, J.E., Heled, J., Delaney, K.S., Peterson, A.T., Knowles, L.L., 2011. Species trees
 960 versus gene trees: implications for speciation history of *Aphelocoma* jays. Evolution 65, 184-
 961 202.

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

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

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

969 Patterson, B.D., 1982. Pleistocene vicariance, montane islands, and the evolutionary divergence
 970 of some chipmunks (genus *Eutamias*). *J. Mammal.* 63, 387-398.

971 Patterson, B.D., 2000. Patterns and trends in the discovery of new Neotropical mammals. *Div.*
 972 *Dist.* 6, 145-151.

973 Patterson, B.D., Norris, R.W., 2015. Towards a uniform nomenclature for ground squirrels: the
 974 status of the Holarctic chipmunks. *Mammalia* DOI: 10.1515/mammalia-2015-0004.

975 Pearson, R.G., Raxworthy, C.J., Nakamura, M., Peterson, A.T., 2007. Predicting species
 976 distributions from small numbers of occurrence records: a test case using cryptic geckos in
 977 Madagascar. *J. Biogeog.* 34, 102-117.

978 Pelletier, T.A., Crisafulli, C., Wagner, S., Zellmer, A.J., Carstens, B.C., 2015. Historical Species
 979 Distribution Models Predict Species Limits in Western Plethodon Salamanders. *Syst. Biol.*
 980 64, 909-925.

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

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

985 Platt, A.R., Woodhall, R.W., George, A.L., 2007. Improved DNA sequencing quality and
 986 efficiency using an optimized fast cycle sequencing protocol. *BioTech.* 43, 58.

987 Rambaut, A., 2009. FigTree v1.3.1. <http://tree.bio.ed.ac.uk/software/figtree/> Accessed 21
 988 December 2009.

989 Rambaut, A., Drummond, A.J., 2007. Tracer v1.5.0. <http://beast.bio.ed.ac.uk/Tracer>. Accessed 1
 990 December 2009.

991 Ramos-Onsins, S.E., Rozas, J., 2002. Statistical properties of new neutrality tests against
 992 population growth. *Mol. Biol. Evol.* 19, 2092-2100.

993 Rausch, R.L., 1951. Notes on the Nunamiut Eskimo and Mammals of the Anaktuvuk Pass
 994 Region, Brooks Range, Alaska. *Faculty Pub. Harold W. Manter Lab. Parasit.* 502.

995 Richmond, O.M.W., McEntee, J.P., Hijmans, R.J., Brashares, J.S., 2010. Is the climate right for
 996 pleistocene rewilding? Using species distribution models to extrapolate climatic suitability for
 997 mammals across continents. *PLoS ONE* 5, e12899.

998 Riddle, B.R., Dawson, M.N., Hadly, E.A., Hafner, D.J., Hickerson, M.J., Mantooth, S.J., Yoder,
 999 A.D., 2008. The role of molecular genetics in sculpting the future of integrative
 1000 biogeography. *Prog. Phys. Geog.* 32, 173-202.

1001 Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed
 1002 models. *Bioinformat.* 19, 1572-1574.

1003 Schoener, T.W., 1968. *Anolis* lizards of Bimini: Resource partitioning in a complex fauna.
 1004 *Ecology* 49, 704-726.

1005 Shafer, A., Cullingham, C.I., Cote, S.D., Coltman, D.W., 2010. Of glaciers and refugia: a decade
 1006 of study sheds new light on the phylogeography of northwestern North America. *Mol. Ecol.*
 1007 19, 4589-4621.

1008 Shcheglovitova, M., Anderson, R.P., 2013. Estimating optimal complexity for ecological niche
 1009 models: a jackknife approach for species with small sample sizes. *Ecol. Model.* 269, 9-17.
 1010 Small, M.P., Stone, K.D., Cook, J.A., 2003. American marten (*Martes americana*) population
 1011 structure across a landscape fragmented in time and space. *Mol. Ecol.* 12, 89-103.
 1012 Smith, M.F., Patton, J.L., 1993. The diversification of South American murid rodents: evidence
 1013 from mitochondrial DNA sequence data for the akodontine tribe. *Biol. J. Linn. Soc.* 50, 149-
 1014 177.
 1015 Soltis, P.S., Gitzendanner, M.A., 1999. Molecular systematics and the conservation of rare
 1016 species. *Cons. Biol.* 13, 471-483.
 1017 Soltis, D.E., Morris, A.B., McLachlan, J.S., Manos, P.S., Soltis, P.S., 2006. Comparative
 1018 phylogeography of unglaciated eastern North America. *Mol. Ecol.* 15, 4261-4293.
 1019 Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype
 1020 reconstruction from population data. *Am. J. Hum. Genet.* 68, 978-989.
 1021 Sullivan, J., Demboski, J.R., Bell, K.C., Hird, S., Sarver, B., Reid, N., Good, J.M., 2014.
 1022 Divergence with gene flow within the recent chipmunk radiation (*Tamias*). *Heredity* 113,
 1023 185-194.
 1024 Swenson, N.G., Howard, D.J., 2005. Clustering of contact zones, hybrid zones, and
 1025 phylogeographic breaks in North America. *Am. Nat.* 166, 581-591.
 1026 Taberlet, P., Fumagalli, L., Wust-Saucy, A.G., Cosson, J.F., 1998. Comparative phylogeography
 1027 and postglacial colonization routes in Europe. *Mol. Ecol.* 7, 453-464.
 1028 Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA
 1029 polymorphism. *Genetics* 123, 585-595.

1030 Tallmon, D.A., Luikart, G., Waples, R.S., 2004. The alluring simplicity and complex reality of
 1031 genetic rescue. *Trends Ecol. Evol.* 19, 489-496.

1032 Taylor, S.A., Larson, E.L., Harrison, R.G., 2015. Hybrid zones: windows on climate change.
 1033 *Trends Ecol. Evol.* 30, 398-406.

1034 Thorington, R.W.Jr., Hoffmann, R.S., 2005. Family Sciuridae, in: Wilson, D.E., Reeder, D.M.
 1035 (Eds), *Mammal species of the world: A taxonomic and geographic reference*, third edition,
 1036 John Hopkins University Press, Baltimore, Maryland, pp 754–818.

1037 Thuiller, W., Lafourcade, B., Engler, R., Araújo, M.B., 2009. BIOMOD—a platform for ensemble
 1038 forecasting of species distributions. *Ecography* 32, 369-373.

1039 Waltari, E., Guralnick, R.P., 2009. Ecological niche modelling of montane mammals in the Great
 1040 Basin, North America: examining past and present connectivity of species across basins and
 1041 ranges. *J. Biogeog.* 36, 148-161.

1042 Waltari, E., Hijmans, R.J., Peterson, A.T., Nyari, A.S., Perkins, S.L., Guralnick, R.P., 2007.
 1043 Locating Pleistocene refugia: comparing phylogeographic and ecological niche model
 1044 predictions. *PLoS ONE* 2, e563.

1045 Warren, D.L., Seifert, S.N., 2011. Ecological niche modeling in Maxent: the importance of
 1046 model complexity and the performance of model selection criteria. *Ecol. App.* 21, 335-342.

1047 Warren, D.L., Glor, R.E., Turelli, M. 2008. Environmental niche equivalency versus
 1048 conservatism: quantitative approaches to niche evolution. *Evolution* 62, 2868-2883.

1049 Warren, D.L., Glor, R.E., Turelli M., 2010. ENMTools: a toolbox for comparative studies of
 1050 environmental niche models. *Ecography* 33, 607-611.

1051 Warren, D.L., Wright, A.N., Seifert, S.N., Shaffer, H.B., 2014. Incorporating model complexity
 1052 and spatial sampling bias into ecological niche models of climate change risks faced by 90
 1053 California vertebrate species of concern. *Div. Dist.* 20, 334-343.
 1054 Weckworth, B.V., Dawson, N.G., Talbot, S.L., Cook, J.A., 2015. Genetic distinctiveness of
 1055 Alexander Archipelago wolves (*Canis lupus ligoni*): reply to Cronin et al. (2015). *J. Hered.*
 1056 106, 412-414.
 1057 Weckworth, B., Talbot, S., Cook, J.A., 2010. Phylogeography of wolves (*Canis lupus*) in the
 1058 Pacific Northwest. *J. Mamm.* 91, 363-375.
 1059 Whittaker, R.J., Araújo, M.B., Jepson, P., Ladle, R.J., Watson, J.E., Willis, K.J., 2005.
 1060 Conservation biogeography: assessment and prospect. *Div. Dist.* 11, 3-23.
 1061 Zazula, G.D., Telka, A.M., Harington, C.R., Schweger, C.E., Mathewes, R.W., 2006. New
 1062 spruce (*Picea* spp.) macrofossils from Yukon Territory: implications for Late Pleistocene
 1063 refugia in Eastern Beringia. *Arctic* 59, 391-400.
 1064

Fig. 1. Distribution and current taxonomy of North American tree squirrels (*Tamiasciurus*). Data points indicate localities of all samples considered for niche analyses. Inset provides the results of a *BEAST species tree analysis (including empirical topology and nodal support values) considering systematic hypothesis H1 with three species as *T. hudsonicus*, *T. douglasii*, and *T. mearnsi*. Downloaded from Wikipedia March 2015. Author: Cephas. This file is licensed under the Creative Commons Attribution-Share Alike 3.0. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled *GNU Free Documentation License*.

Fig. 2. Bioclimatic envelope models for North American red squirrels considering three species according to systematic hypothesis H2 as *T. douglasii* (green), *T. fremonti* (brown), and *T. hudsonicus* (dark blue). We used model selection and applied a minimum of 90% of training samples as our threshold for each species. Potential geographic overlap between species is depicted with intermediate colors (cyan between *T. hudsonicus* and *T. douglasii*, and gray between *T. hudsonicus* and *T. fremonti*). Left inset provides the results of *BEAST species tree analysis considering H2, including empirical topology and nodal support values. Top inset shows results of niche identity test, where all pairwise comparisons indicate significant ecological differentiation between species. The histogram provides expected niche overlap when pairwise species are pooled and randomized, and arrows indicate empirical niche overlap between species (Schoener's *D*).

Fig. 3. Left – Chronogram of evolutionary relationships among North American red squirrels (*Tamiasciurus*) based on 771bp of the mtDNA Cytochrome b gene with a root coalescence estimate from Chavez et al. (2014). Phylogeny estimation was performed in BEAST, providing both posterior probability nodal support values (left of node) and coalescence times (millions of years; italics, right of node). The topology is presented as ultrametric and proportional to an evolutionary timeline (bottom) that extends from the present (right) to the past (left). Tip labels indicate general collecting localities (Appendix A). Major lineages are color coded and coincident with geography. Tip labels indicate sampling localities. Right – Map of the study area indicating general specimen localities for samples of North American red squirrels included in the present study. Locality numbers correspond to those listed in Appendix A. Left inset provides finer detail of localities in the vicinity of the Alexander Archipelago. Bottom-right inset provides finer detail of localities within the Southwestern region. Localities are colored according to major mtDNA lineages. Circles with multiple colors indicate general localities where multiple mtDNA lineages are sympatric. For insets, insular populations are labeled by island or mountain range. Top inset provides the results of a *BEAST species tree analysis (including empirical topology and nodal support values) considering systematic hypothesis H3 with nine independent taxa color-coded according to mtDNA lineages.

Table 1. Genetic polymorphism and demographic statistics considering 771bp of the mtDNA Cytochrome b gene among major North American red squirrel lineages. Abbreviations include n =sample size; S = number of segregating sites; h =number of haplotypes; Hd =haplotype diversity; π = nucleotide diversity; D =Tajima's D (with associated P -value); $R2$ =Ramos-Onsins and Rozas $R2$ (with associated P -value); μ = lineage-specific substitution rate (percent per million years) considering rates weighted by branch lengths for each lineage from the Cytochrome b BEAST chronogram. The rate given for all lineages combined was calculated in BEAST using a root calibration of 0.4229 Myr (HPD: 0.2979–0.5538; Chavez et al., 2014). N/A = not applicable. Bold values are statistically significant at $P<0.05$.

	n	S	h	Hd	π	D	$P(D)$	$R2$	$P(R2)$	μ
All	252	134	142	0.982	0.0181	N/A	N/A	N/A	N/A	5.23
Central Rockies	11	4	4	0.600	0.0015	-0.542	0.324	0.142	0.107	5.14
Continental	60	41	34	0.866	0.0045	-2.167	0.001	0.046	0.000	5.20
Northwestern	52	45	32	0.924	0.0031	-2.593	0.000	0.022	0.000	5.35
North Pacific Coast	20	15	16	0.963	0.0040	-1.192	0.132	0.088	0.056	5.03
Sacramento Mountains	10	11	7	0.933	0.0035	-1.407	0.096	0.124	0.056	5.26
Southern Rockies	15	14	10	0.924	0.0035	-1.509	0.064	0.079	0.000	5.42

Southwestern Sky Islands	22	17	12	0.905	0.0037	-1.413	0.062	0.073	0.011	5.26
Vancouver Island	19	8	7	0.544	0.0012	-2.022	0.001	0.084	0.002	5.21
Western	42	35	26	0.962	0.0067	-1.325	0.074	0.065	0.059	5.25

1168

1169

Table 2. AICM and Bayes Factor ($2\ln Bf$) results considering marginal likelihood estimates from alternative species-tree hypotheses of relationships among major lineages of North American red squirrels (see Fig. S1 for scenario topologies). The hypothesis receiving the best marginal-likelihood score is indicated by N/A. 2AICM and $2\ln Bf$ scores for other scenarios reflect comparative strength of the best supported hypothesis, where for $2\ln Bf$ a score of 0–2 means “not worth more than a bare mention”, 2–6 means “positive” support for the best hypothesis over that alternative, 6–10 provides “strong” support, and >10 means “decisive” support in distinguishing between best and competing hypotheses (after Kass and Raftery, 1995). For 2AICM scores, we consider >20 as decisive, 12-20 as strong, and 4-12 as positive support for the best hypothesis over that alternative. Hypotheses were analyzed considering all five loci (both mtDNA and nuDNA) and using only the four nuclear loci.

Scenario	Description	All five loci		Four nuDNA loci	
		2AICM	$2\ln Bf$	2AICM	$2\ln Bf$
H1	3 species as currently described (<i>T. douglasii</i> , <i>T. hudsonicus</i> , <i>T. mearnsi</i>)	N/A	10.28	5.72	10.74
H2	3 species as per Arbogast et al. (2001; <i>T. douglasii</i> , <i>T. hudsonicus</i> , <i>T. fremonti</i>)	27.85	N/A	N/A	N/A
H3	All lineages considered separately	104.10	22.76	31.70	16.06

1180 **Table 3.** Relative importance of niche model variables. During modeling, we conducted 20 replicates, retained 20% of samples for
1181 training, and applied optimized parameters from model selection. The training, test, and standard deviation of AUC (=Area Under
1182 Curve) scores are measures of model performance. Below the AUC metrics are the relative contributions of variables. Variables
1183 shown contributed >10% to each model and were used for ecological interpretation, corresponding to the top three variables in each
1184 model. Symbols reflect results of the jackknife test of variable importance showing the highest gain when used in isolation (*), the
1185 variable that decreased the gain most when omitted (#), and the variable with the highest permutation importance (\$)
1186

AUC/Variable	<i>T. douglasii</i>	<i>T. fremonti</i>	<i>T. hudsonicus</i>
Test AUC	0.942	0.911	0.905
AUC SD	0.008	0.047	0.012
Annual Temperature (Bio1)			12.7 *
Isothermality (Bio3)	13.5		31.3 #,\$
Temperature Seasonality (Bio4)		13.5	
Summer Temperature (Bio5)		50.0 *,\$	

Precipitation Seasonality (Bio15) 20.0 ^{#, \$}

Winter Precipitation (Bio19) 39.7 * 15.8 [#] 21.5

1187

Figure 1

Figure 2

Figure 3