



Ultraconserved elements improve resolution of marmot phylogeny and offer insights into biogeographic history

Kendall K. Mills^{a,b,*}, Kathryn M. Everson^{b,c}, Kyndall B.P. Hildebrandt^b, Oleg V. Brandler^d, Scott J. Steppan^e, Link E. Olson^b

^a Department of Biology and Wildlife, University of Alaska Fairbanks, 982 Koyukuk Drive, Fairbanks, AK 99709, USA

^b Department of Mammalogy, University of Alaska Museum, 1962 Yukon Drive, Fairbanks, AK 99775, USA

^c Department of Integrative Biology, Oregon State University, 2701 SW Campus Way, Corvallis, OR 97331, USA

^d Koltzov Institute of Developmental Biology of Russian Academy of Sciences, Vavilova 26, Moscow, Russia

^e Department of Biological Science, Florida State University, Tallahassee, FL 32306, USA

ARTICLE INFO

Keywords:

Marmota
Pleistocene
Ultraconserved elements
Beringia
Holarctic
Cryptic species

ABSTRACT

Marmots (*Marmota* spp.) comprise a lineage of large-bodied ground squirrels that diversified rapidly in the Pleistocene, when the planet quickly transitioned to a drier, colder, and highly seasonal climate—particularly at high latitudes. Fossil evidence indicates the genus spread from North America, across Beringia, and into the European Alps over the course of only a few million years, beginning in the late Pliocene. Marmots are highly adapted to survive long and severely cold winters, and this likely favored their expansion and diversification over this time period. Previous phylogenetic studies have identified two major subgenera of marmots, but the timing of important speciation events and some species relationships have been difficult to resolve. Here we use ultraconserved elements and mitogenomes, with samples from all 15 extant species, to more precisely retrace how and when marmots came to inhabit a vast Holarctic range. Our results indicate marmots arose in North America in the Early Miocene (~16.3 Ma) and dispersed across the Bering Land Bridge in the Pliocene (~3–4 Ma); in addition, our fossil-calibrated timeline is suggestive of the rise and spread of open grasslands as being particularly important to marmot diversification. The woodchuck (*M. monax*) and the Alaska marmot (*M. broweri*) are found to be more closely related to the Eurasian species than to the other North American species. Paraphyly is evident in the bobak marmot (*M. bobak*) and the hoary marmot (*M. caligata*), and in the case of the latter the data are highly suggestive of a second, cryptic species in the Cascade Mountains of Washington.

1. Introduction

The onset of the Quaternary is marked by dramatic cooling and aridification of the planet, giving rise to the coldest period on earth for at least the last 65 million years (Zachos et al., 2001). Many taxa adapted to a warm, wet, and aseasonal planet went extinct, and the volatile glacial-interglacial cycles of the Pleistocene further contributed to widespread faunal turnover (Azzaroli et al., 1988; Cerling et al., 1997; Peter, 2004; Van Kolfshoten and Markova, 2005). Within this context flourished marmots, a lineage of Miocene ground squirrels that made only sporadic appearances in the fossil record prior to the Quaternary (Black, 1963). Marmots are readily distinguished from the rest of the world's squirrels (family Sciuridae) by their large body size, which, along with heightened sociality and long bouts of hibernation, is probably an adaptation

to seasonally cold and resource-scarce environments (Barash, 1974; Armitage, 1981; Arnold, 1992). These adaptations facilitated a rapid expansion across a vast Holarctic region in the Pleistocene (Zimina and Gerasimov, 1973) and gave rise to the 15 currently recognized extant species whose ranges, although diminished since the Holocene, still extend from eastern North America to the European Alps (Fig. 1).

After evolving a much larger body size, very little morphological evolution occurred during the radiation of marmots, consistent with the high degree of conservatism thought to characterize the sciurid body plan (Hafner, 1984). Despite this, marmots occupy a broad range of elevations and biomes. Several species reside at sea level, and, in some cases, even inhabit rocky seaside beaches (Armitage, 2000; Kryštufek and Vohralík, 2013), while others are found only above 3000 m in severely hypoxic conditions (Bai et al., 2019). However, all marmots are

* Corresponding author.

E-mail address: kendall.k.mills@gmail.com (K.K. Mills).

<https://doi.org/10.1016/j.ympev.2023.107785>

Received 11 October 2022; Received in revised form 1 March 2023; Accepted 13 April 2023

Available online 20 April 2023

1055-7903/© 2023 Elsevier Inc. All rights reserved.

restricted to environments with a season of resource scarcity and prolonged snow cover (Armitage, 2014). Social organization among marmots ranges from large, highly social colonies, to smaller multigenerational family groups, to solitary individuals that interact only to mate (Armitage, 2000). For this reason, and because they are diurnal and relatively amenable to human observation, marmots have featured prominently in sociobiology for decades—particularly since Barash's (1974) assertion that increased sociality and decreased aggression correlate with elevation in marmots and probably facilitate survival in harsh environments.

Marmots are members of the rodent family Sciuridae, a clade that emerged sometime in the Eocene but whose geographic origins are unknown (Heissig, 2003). Sciurids underwent a substantial radiation around 34 Ma (Mercer and Roth, 2003), coincident with the massive Eocene-Oligocene extinction event triggered by planetary cooling and earth's transition to the modern, temperate climate (Prothero, 1994). This radiation produced five major subfamilies widely distributed across Eurasia, Africa, and the Americas. Of these five, only Xerinae includes ground squirrels—lineages that departed from the arboreal lifestyle and came to spend most or all of their time on or beneath the ground. Terrestriality is thought to have evolved twice within Xerinae, once in the ancestor of Xerini (African and Asian ground squirrels) and once in the ancestor of Marmotini (Steppan et al., 2004; Emry and Korth, 1996). Extant members of Marmotini include marmots, chipmunks, prairie dogs, and the 11 genera of ground squirrel previously included in the paraphyletic *Spermophilus* (Harrison et al., 2003; Herron et al., 2004; Helgen et al., 2009). The earliest fossil appearances of Marmotini are found in North America near the Oligocene-Miocene boundary (Sinita, 2018), which is approximately when open-habitat grasslands became ecologically dominant in North America (Strömberg, 2011). Derived traits of Marmotini are largely adaptations to grassland ecosystems and burrowing behavior, such as grass-based diets, short tails relative to body mass, and enlargement of the clavobrachialis muscle (Hayssen, 2008; McLean et al., 2018; Zelditch et al., 2017).

The oldest fossils attributable to Marmotini are *†Protospermophilus* from the early Oligocene and *†Miospermophilus* from the late Oligocene (Black, 1963). It was previously assumed that marmots descended from *†Protospermophilus* while the other ground squirrels descended from *†Miospermophilus*, which would imply marmots diverged soon after the radiation of all Sciuridae and are distantly related to other Marmotini. However, Black (1963) noted that *Marmota* fossils are not known until the Late Miocene, and genetic analyses have consistently found that the common ancestor of *Marmota* emerged at the same time as other crown Marmotini (Thomas and Martin, 1993; McLean et al., 2019; Menéndez et al., 2021). Consistent with a more recent origin of marmots, dental and mandibular morphology suggests *†Protospermophilus* is basal to all Marmotini, whereas *†Miospermophilus* arose after the divergence of *Tamias* from this clade (Sinita, 2018).

Marmots are generally assumed to have originated in North America. Steppan et al. (1999) recognized two subgenera of marmots based on a phylogeny derived from cytochrome *b* (*cytb*) gene sequences: *Petromarmota*, which includes only Nearctic species, and *Marmota*, which includes both Nearctic and Palearctic species. The most parsimonious phylogeny based on their *cytb* analysis nests *Marmota* within *Petromarmota*, consistent with a North American origin, but reciprocal monophyly between the subgenera could not be rejected. The Tertiary fossil record of marmots is scarce but also suggests the genus evolved in North America in the mid-Miocene. Only four specimens are known from the first ~12 million years of their existence (Gordon and Czaplewski, 2000; Korth, 1994b). The oldest is *M. vetus* from the Clarendonian Land Mammal Age (LMA) of Nebraska (11–9.5 Ma) (Marsh, 1871), and the others are *M. minor* of the Hemphillian LMA of Nevada (10.3–4.9 Ma), a *Marmota* sp. from Oklahoma (Gordon and Czaplewski 2000), and *M. arizonae* from the Blancan LMA of Arizona (4.9–1.8 Ma; Korth, 1994b). However, Nikol'skii and Rumiantsev (2012) postulated that marmots originated in Asia because the greatest species diversity is centered on an epi-platfomal orogeny (consisting of the Hindu Kush, Tan Shan, Pamir, Altai, and Sayan mountains) that arose there in the

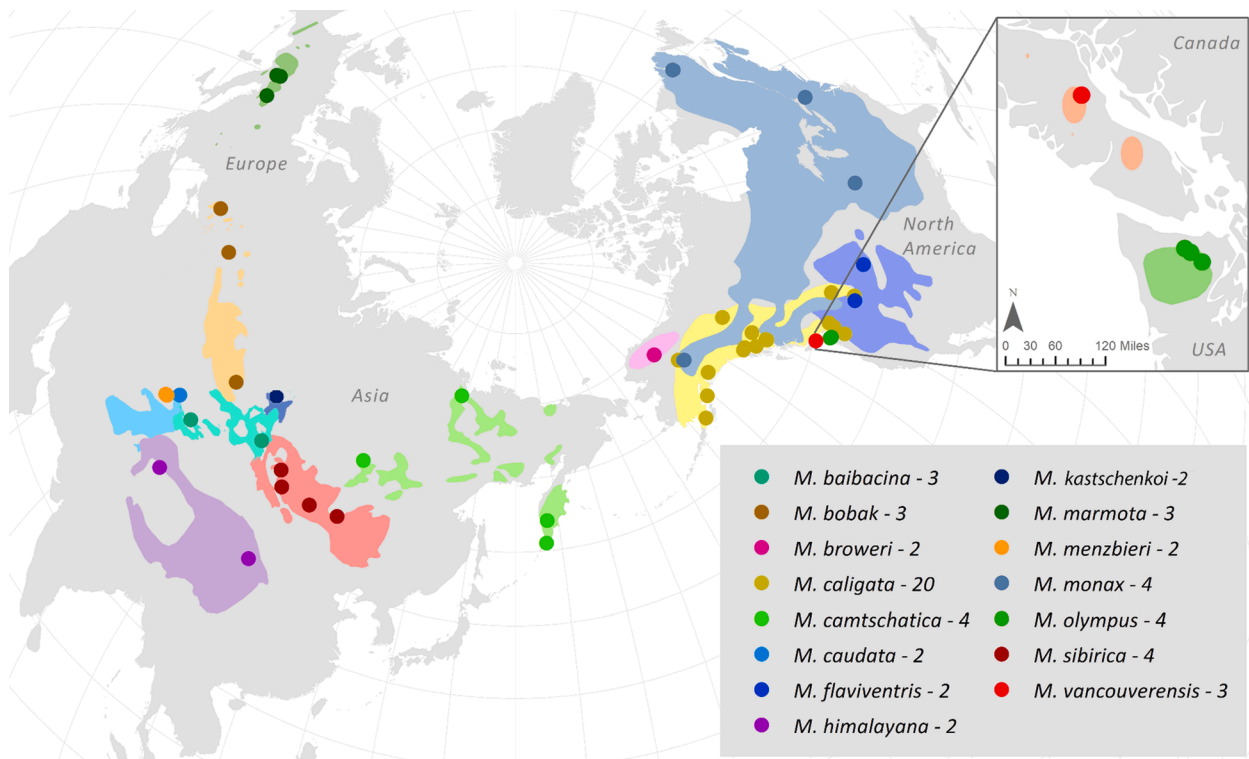


Fig. 1. Ranges of the 15 extant *Marmota* species and localities of the specimens used in this study. Numbers represent sample size for this study. All range maps are from the IUCN Red List database except for that of *M. kastschenkoi*, which is redrawn from Krystufek and Vohralík, 2013.

Miocene. In this view, marmots were widespread in Asia by the mid-Miocene, around which time mountain uplift began to promote geographic isolation and subsequent allopatric speciation.

The Holarctic distribution of marmots implies at least one dispersal across the Bering Land Bridge (BLB), but the timing and directionality are unclear because the geographic origin of marmots is disputed. It has been suggested that, if marmots originated in North America, the Alaska marmot (*Marmota broweri*), the woodchuck (*M. monax*), or both subsequently backcrossed the BLB into North America (Hoffmann et al., 1979; Steppan et al., 1999). Both species are found only in North America (Fig. 1), but genetic analyses based on *cytb* (Steppan et al., 1999) and another mitochondrial gene (*ND4*; Kruckenhauser et al., 1999), as well as a nuclear gene (*RAG1*; Steppan et al., 2004, 2011), place them within the subgenus *Marmota*, more closely related to the Eurasian species than to the other North American species. There is more support for a backcrossing by *M. broweri* because these studies found the split leading to the alpine marmot (*M. marmota*), found in the European Alps, to be basal to the split leading to *M. broweri*, suggesting *M. broweri* had a Palearctic ancestor. In contrast, data from non-coding regions of the nuclear genome indicate *M. broweri* and *M. monax* are found within *Petromarmota* and are therefore more closely related to the other Nearctic species (Brandler et al., 2010). Resolving interrelationships within the subgenus is therefore needed to clarify the history of BLB crossings by marmots.

Within the subgenus *Marmota*, a clade consisting of the Himalayan marmot (*M. himalayana*), the black-capped marmot (*M. camtschatica*), and the tarbagan (*M. sibirica*) is strongly supported, but the relationships within the clade are unresolved (Steppan et al., 1999; Brandler and

Lyapunova, 2009). Another strongly supported clade is that uniting the bobak marmot (*M. bobak*), the gray marmot (*M. baibacina*), and the forest-steppe marmot (*M. kastschenkoii*). As suggested by its name, the forest-steppe marmot occupies an unusual niche in that it is much more tolerant of forest and forest-edge habitats than other species. It is allopatric with and was previously considered a subspecies of *M. baibacina* but was elevated to species in 2003 due to its distinct karyotype ($2n = 36$ in *M. kastschenkoii* compared to $2n = 38$ in *M. baibacina*; Brandler, 2003).

The relationship between the long-tailed marmot (*Marmota caudata*) and Menzbier's marmot (*M. menzbieri*) of central Asia is unclear but is of particular interest evolutionarily. The two species are peripatric and highly morphologically distinct—*M. caudata* has the largest body mass and longest tail of any marmot, and *M. menzbieri* has a short tail and the second-smallest body size (Armitage, 2000). In addition, *M. caudata* has a distinct rufous pelage not found in any other species (Fig. 2). Alarm calls of the two species have some similarities and some differences: calls differ significantly in duration and the presence of full and deep amplitude modulation in *M. caudata* as opposed to *M. menzbieri*, but they are similar in that they lack a distinct low-frequency component in the beginning of each alarm call. The only other Eurasian species whose call lacks this low-frequency component is *M. marmota* (Nikol'skii, 2014). Previous studies recovered *M. menzbieri* and *M. caudata* as sister species in phylogenies based on mtDNA, which would imply rapid morphological evolution in both lineages. However, phylogenies based on nuclear loci do not recover the two as sister species, which raises the possibility they are more distantly related and have undergone mitochondrial introgression (Brandler et al., 2010; Steppan et al., 2011). A



Fig. 2. Top: *M. caudata* in Valley of the Surkhob (Vakhsh) River, Peter the Great Ridge, Pamir, Tajikistan (photo by Oleg Brandler). Bottom: *M. menzbieri* in Sairam-Ugam National Park, Turkestan region, Kazakhstan (photo by Georgiy Shakula).

larger nuclear dataset is needed to determine which of these hypotheses is best supported.

Within the subgenus *Petromarmota*, the only well-established placement is that of the yellow-bellied marmot (*M. flaviventris*), which is consistently recovered as sister to the remaining three species. The *M. caligata* superspecies complex consists of the hoary marmot (*M. caligata*), the Vancouver Island marmot (*M. vancouverensis*), and the Olympic marmot (*M. olympus*). Both *M. olympus* and *M. vancouverensis* are small, imperiled populations found on the periphery of the range of *M. caligata*. However, each is clearly differentiated from *M. caligata* by cranial and pelage morphology (Fig. 3; Cardini, 2003; Cardini et al., 2005; Cardini and Thorington, 2006; Armitage, 2009). *Marmota olympus* is further distinguished by karyotype: *M. caligata* and *M. vancouverensis* both have $2n = 42$ chromosomes (Hoffmann and Nadler, 1968), whereas

M. olympus has only $2n = 40$ (Rausch and Rausch, 1971). Relationships among these species have been particularly difficult to resolve, most likely because of intermittent secondary contact during Pleistocene glacial cycles (Kerhoulas et al., 2015). Mitochondrial data suggest *M. olympus* is an older offshoot of *M. caligata* from ~2.6 Ma, while *M. vancouverensis* only split from *M. caligata* 0.4–1.2 Ma (Steppan et al., 2004). However, phylogenies based on nuclear data are poorly resolved at these nodes. It is possible that the sister relationship of *M. vancouverensis* and *M. caligata* suggested by mitochondrial data is the product of recent interspecific gene flow, and that the original speciation event is considerably older (Kerhoulas et al., 2015). Moreover, Kerhoulas et al. (2015) identified two mitochondrial clades of *M. caligata* (deemed ‘coastal’ and ‘continental’), one of which is sister to *M. vancouverensis*. These clades are not present in nuclear-based phylogenies, and Kerhoulas et al. (2015) suggested the two clades represent two refugial populations from the Last Glacial Maximum—the continental clade deriving from a refugial population south of the Laurentide ice sheet, and the coastal clade deriving from a refugial population near the Cascade Mountains or Alexander Archipelago.

Here we use ultraconserved elements (UCEs), which have been successful in resolving many recalcitrant phylogenies (Faircloth et al., 2012; Meiklejohn et al., 2016), from all 15 extant marmot species to better elucidate the phylogenetic history of marmots. These sequences are highly conserved among all tetrapods, allowing thousands of homologous loci to be curated rapidly from across the genome (Faircloth et al., 2012). Because sequencing initiates in the conserved region and proceeds outward in both directions into increasingly variable regions, large stretches of sequence that contain genetic diversity informative across a broad evolutionary timescale are captured along with each UCE. We also recovered complete mitogenomes from a subset of species, which allowed us to compare phylogenetic signals detected in two distinct datasets. The results add considerable details to the biogeographic history of marmots and indicate there may be a 16th marmot species in the Cascade Mountains of the Pacific Northwest.

2. Material and methods

2.1. Data generation and phylogenetic analysis of ultraconserved elements

We obtained frozen tissues from 56 individuals representing all 15 currently recognized extant marmot species (Table S1). Genomic DNA was extracted following the PureGene kit protocol for animal tissue (Gentra Systems Inc.). Sequence capture of ultraconserved elements (UCEs) was performed on a pilot dataset of 15 individuals according to the original UCE capture and sequencing protocols (Faircloth et al., 2012). Briefly, genomic libraries were prepared using KAPA library kits (Kapa Biosystems) and barcoded adapters for multiplexing (Faircloth and Glenn, 2012). Enrichment was performed using the UCE Tetrapods 5kV1 probe set (<https://ultraconserved.org/>), and post-enrichment libraries were pooled and sequenced on one lane of an Illumina HiSeq 2000 instrument with 150 bp paired-end reads at the University of California Los Angeles Neuroscience Genomics Core. An additional 41 individuals were then sent to RAPiD Genomics (Gainesville, FL, USA) for library prep and sequencing using the same lab protocols. In the second batch, targeted reads were sequenced on an Illumina HiSeq 4000 platform with 150 bp paired-end reads. All indexed reads were demultiplexed with the bcltofastq script from Illumina CASAVA v1.8.4.

We followed the phyluce v1.6.2 bioinformatics pipeline (Faircloth, 2016) to process raw reads. Our newly generated reads were supplemented with additional raw UCE sequence data from ten *Marmota* individuals, which were provided by Bryan S. McLean as part of a recent phylogenetic study (library preparation and sequencing methods described in McLean et al. (2019); Table S1). Adapter sequences and low-quality base calls were removed using Trimmomatic (Bolger et al., 2014). The cleaned reads were assembled using velvet v1.2.10 (Zerbin and Birney, 2008),



Fig. 3. Top: *M. caligata* (photo by L. E. Olson). Middle: *M. vancouverensis* (photo by Oli Gardner). Bottom: *M. olympus* (photo by Richard Klawitter).

with optimal kmer values for each taxon estimated with VelvetOptimiser v2.2.6 (Gladman and Seemann, 2008). To supplement our dataset, all five published *Marmota* genomes (*M. himalayana*, GCA_005280165; *M. monax*, GCA_014533835; *M. marmota*, GCA_001458135; *M. flaviventris*, GCA_003676075; and *M. vancouverensis*, GCA_005458795) were retrieved from GenBank; for outgroups, we also retrieved genomes for the Cape ground squirrel (*Xerus inauris*, GenBank accession GCA_004024805), the arctic ground squirrel (*Urocitellus parryi*; GCA_003426925), Gunnison's prairie dog (*Cynomys gunnisoni*; GCA_011316645), the Daurian ground squirrel (*Spermophilus dauricus*; GCA_002406435), and the thirteen-lined ground squirrel (*Ictidomys tridecemlineatus*; GCA_016881025). Contigs containing UCE matches were extracted from both the velvet assemblies and the downloaded genomes using the *phyluce_assembly_match_contigs_to_probes* script, and each UCE locus was aligned with MAFFT 6.240 (Katoh and Standley, 2013) using the default settings and edge trimmed with GBLOCKS (Castresana, 2000). The dataset was then reduced to a 75% complete matrix by removing any loci with data from fewer than 75% of individuals.

To partition the dataset, we first used the entropy-based method SWSC-en, which is designed specifically for UCE loci (Tagliacollo and Lanfear, 2018). Most UCes are split into three partitions: one containing the middle, highly conserved core of the UCE; and two flanking ends containing the more variable positions outside the core. Because SWSC-en partitions each locus individually, the resulting partitions were input into PartitionFinder2 for further consolidation (Lanfear et al., 2017). In this step, for example, the middle partitions from many different loci might be combined into one. PartitionFinder2 was also used to identify the best substitution model for each partition based on AICc score. Maximum-likelihood trees were estimated for the concatenated dataset using IQ-TREE (Nguyen et al., 2015), with branch supports obtained using the ultrafast bootstrap method (UFboot; Hoang et al., 2017) and the SH-like approximate likelihood ratio test (SH-aLRT; Anisimova et al., 2011), both with 1000 replicates.

We also used a coalescent-based method to estimate a phylogeny. Gene trees are weighted equally when imputing a species tree, but many UCE loci contained only weak phylogenetic signal; to account for this, we sorted the UCE loci by number of parsimony-informative sites using the *phyluce_align_get_informative_sites* script, and the dataset was reduced to the top quartile of loci. A maximum-likelihood gene tree was then generated for each UCE locus in the reduced dataset using IQ-TREE, with branch supports obtained using UFboot and SH-aLRT. Because not all loci are expected to contain strong phylogenetic signal for every node in the tree, all branches with less than 10% bootstrap support were collapsed (Zhang et al., 2018). These gene trees were used to estimate species trees using ASTRAL-III (Rabiee et al. 2019). Because we found evidence in our IQ-TREE analysis that some species did not form monophyletic groups, we left individuals as OTUs (operational taxonomic unit) rather than categorizing individuals by species (i.e., we did not use the -a flag option).

2.2. Data generation and phylogenetic analysis of mitogenomes

The sequence-capture method for UCes occasionally yields partial or entire mitogenomes as a byproduct. Mitochondrial sequences were

extracted from the velvet assemblies by identifying the longest contigs in each assembly with GATK (McKenna et al. 2010). Of the 56 sequenced individuals, 22 complete and three partial mitogenomes were recovered (see Table S1 for accession information). For outgroups, mitogenomes from the thirteen-lined ground squirrel (*Ictidomys tridecemlineatus*), Gunnison's prairie dog (*Cynomys gunnisoni*), and white-tailed prairie dog (*Cynomys leucurus*) were retrieved from GenBank (Table S1).

Mitogenomes were aligned to the *Marmota himalayana* mitogenome (NC018367) using MAFFT in Geneious (v2021.0). We used GBLOCKS to mask hypervariable regions that may be saturated and poorly aligned (Castresana, 2000), and we manually inspected the trimmed alignment for incorrectly aligned and annotated bases. Genes were annotated in Geneious based on the available *M. himalayana* annotation. Each gene (including D-loop and the replication origin) was subsetted into individual partitions, except for protein-coding genes, which were each split into three partitions corresponding to codon position. PartitionFinder2 consolidated these into 13 partitions and estimated the best substitution model for each (Table S2). A maximum-likelihood tree was estimated in IQ-TREE with branch supports obtained with UFboot and SH-aLRT.

2.3. Divergence time estimation

We estimated divergence times from our UCE data using the program MCMCTree (Yang and Rannala, 2006), implemented in PAML v4.9j (Yang, 2007). Node ages were calibrated using a set of eight fossils (Table 1), which could be placed after adding three additional outgroups (*Aplodontia rufa*, GenBank accession GCA_004027875; *Sciurus carolinensis*, GCA_902686445; and *Sciurus vulgaris*, GCA_902686455) following the same bioinformatics pipeline outlined above. We used the concatenated UCE 75% complete alignment as the input for MCMCTree and fixed the tree topology to the IQ-TREE result of the 75% complete matrix. The root of the tree was constrained to a minimum age of 40.4 Ma (based on †*Douglassciurus oaxacensis*, the oldest known sciurid fossil; Ferrusquia-Villafranca et al., 2018) and a maximum age of 56 Ma (based on †*Acritoparamys*, the oldest stem member of Sciuromorpha; Ivy, 1990; Korth, 1994). All other fossils were used to place minimum dates on internal nodes (Table 1). The MCMCTree analysis was conducted in a two-step approach, following the guidelines for analyzing large datasets established by dos Reis et al. (2012). First, we obtained maximum-likelihood estimates of branch lengths along with information about the curvature of the likelihood surface using the *usedata* = 3 option. In the second step, we used the *usedata* = 2 option to estimate divergence times on the previously estimated branch lengths with uncorrelated rates across loci and a GTR + Γ model of nucleotide substitution. In both steps, we analyzed two independent MCMC chains sampled every 100 generations until 25,000 samples were collected after a burn-in period of 10,000 generations. We confirmed that the MCMC chains had converged and that effective sample sizes were large using Tracer v1.7.1 (Rambaut et al., 2018). After verifying that the two independent runs had produced similar divergence times, the runs were combined and summarized using the *print* = -1 setting in MCMCTree to calculate the mean posterior divergence times and their 95% highest posterior density credible intervals.

Table 1
Fossils used for calibration in divergence time analysis. Fossil ages are in millions of years (Ma).

Fossil species	Locality	Estimated min. age (Ma)	Calibrated Node (MCMCTree)	Reference
<i>Douglassciurus oaxacensis</i>	Oaxaca, Mexico	40.4	Sciuridae - Aplodontiidae	Ferrusquia-Villafranca et al. 2018
<i>Plesiartomys (Sciurus) spectabilis</i>	Egerkingen, Switzerland	37.2	<i>Sciurus</i> - other sciurids	Costeur and Schneider, 2011
<i>Palaeosciurus goti</i>	Quercy, France	28.4	<i>Xerus</i> - Marmotini	McKenna and Bell 1997
<i>Miospermophilus</i>	Colorado, USA	16.0	<i>Spermophilus</i> - <i>Marmota</i>	(Black, 1963)
<i>Marmota bobak</i>	Kantemirovka, Russia	0.5	<i>M. bobak</i> - <i>M. baibacina</i>	Gromov et al. (1965); Erbajeva and Alexeeva (2009)
<i>Marmota marmota</i>	Castile, Spain	1.2	<i>M. marmota</i> - sister clade	Gil and Sesé (1991)
<i>Marmota himalayana</i>	Shanxi, China	1.0	<i>M. himalayana</i> - <i>M. sibirica</i>	Gromov et al. 1965
<i>Marmota monax</i>	Maryland, USA	0.8	<i>M. monax</i> - sister clade	Kurten and Anderson (1980)

2.4. Reconstruction of ancestral ranges

To infer the geographic origins of marmots (Nearctic vs. Palearctic), we reconstructed ancestral ranges using the BioGeoBEARS package (Matzke, 2014) implemented in R. The UCE-based divergence time tree was used as the input, and we compared all six available models of ancestral range evolution: dispersal-extinction-cladogenesis (DEC), dispersal-vicariance (DIVA), BayArea, and each of these models with the addition of a “jump” parameter (DEC + J, DIVA + J, and BayArea + J). We coded each extant species as Nearctic or Palearctic based on known ranges, and the top model was selected using the Akaike Information Criterion (AIC). The model with the highest AIC weight was used to infer ancestral ranges. The initial round of analyses did not include outgroups. To determine if the exclusion of outgroups impacted this analysis, we repeated the above steps on the UCE-based divergence time tree with the eight outgroup taxa included (Table S3). With this dataset, each taxon was coded as Nearctic, Palearctic, Holarctic, or African.

3. Results

3.1. Data generation and phylogenetic analysis of ultraconserved elements

We obtained a mean of $\sim 2.1 \times 10^7$ paired-end reads per sample. Before data filtering, the concatenated alignment contained 4,427 UCE loci and 88,252 parsimony-informative sites; mean locus length was 587.8 bp, and mean number of informative sites per locus was 19.9. The 75% complete dataset contained 1,986 UCE loci with a mean of 22.5 parsimony-informative sites per locus. The dataset filtered for ASTRAL-III input contained 496 loci with a mean of 44.2 parsimony-informative sites per locus. Full alignment statistics are provided in Table 2, and nexus files for each of the 4,427 UCE alignments are available at <https://ddoi.org/10.5061/dryad.6hdr7sr56>. The maximum-likelihood tree estimated in IQ-TREE (Fig. 4) shows unambiguous support for the two subgenera, *Marmota* and *Petromarmota*, identified by Steppan et al. (1999) from *cytb* data. Monophyly was highly supported for all species except for *M. baibacina* and *M. caligata*. The *M. baibacina* individual from Kazakhstan clustered with *M. bobak*, and *M. caligata* was split into two groups in both trees with relatively high support. The three *M. caligata* from Washington state are paraphyletic with respect to *M. olympus*, while all of the other *M. caligata* (ranging from Montana to Alaska) form a monophyletic group that is sister to *M. vancouverensis*.

The species tree estimated in ASTRAL-III (Fig. 5) has several differences from the maximum-likelihood tree. First, many nodes within the *M. bobak* + *M. baibacina* + *M. kastschenkoi* clade are poorly resolved, and none of the three species are monophyletic. Second, support values are similarly low within the *M. caligata* + *M. olympus* + *M. vancouverensis* clade, but in this case the latter two species do form clades. Third, *M. menzbieri* is sister to *M. marmota* rather than *M. caudata*, and *M. caudata* is sister to all other Palearctic species.

3.2. Data generation and phylogenetic analysis of mitogenomes

Partial or complete mitogenomes were recovered in 45% of samples and 11 of 15 species, and new mitogenome sequences were accessioned for 7 species (*M. bobak*, *M. baibacina*, *M. caligata*, *M. kastschenkoi*, *M. menzbieri*, *M. olympus*, and *M. sibirica*). GenBank accession numbers

are provided in Table S1. The raw mitogenome alignment contained 44 sequences and was 16,914 bp in length. GBLOCKS masked 1,313 bp, and the trimmed alignment used in downstream analyses contained 15,601 bp. Both alignments are available at <https://doi.org/10.5061/dryad.6hdr7sr56>.

The ML tree inferred from mitogenomes differs from the UCE-based ML tree in that *M. marmota*, rather than *M. menzbieri* + *M. caudata*, is sister to all other Palearctic species (although this branch has relatively weak support; Fig. 6). While *M. caligata* is non-monophyletic in this tree as well, the topology is entirely different. Here, one *caligata* lineage is sister to *M. vancouverensis*, and the other lineage is sister to both of these. The two lineages correspond to the mitochondrial clades identified in Kerhoulas et al. (2015) as “coastal” and “continental”, with the coastal clade being sister to *M. vancouverensis*.

3.3. Divergence time estimation and reconstruction of ancestral ranges

Our time-calibrated phylogeny (Fig. 7) estimated the divergence of tribe Marmotini and tribe Xerini at 35.0 million years ago (Ma) [95% confidence interval (C.I.) = 33.8–36.7]. The divergence of the genus *Marmota* from other marmotines occurred 16.3 Ma (95% C.I. = 13.7–18.8), followed later by the crown diversification of *Marmota* at 5.71 Ma (95% C.I. = 3.95–7.62). Finally, the crown diversification of Palearctic *Marmota* occurred 3.57 Ma (95% C.I. = 2.45–4.72) while the crown diversification of the *M. caligata*-*olympus*-*vancouverensis* complex occurred 2.31 Ma (95% C.I. = 2.62–3.11).

In our analyses of ancestral range evolution, the DEC + J model (dispersal-extinction-cladogenesis with a jump parameter) was highly supported when outgroups were excluded, receiving 70% of the corrected AIC weight (Table S3). The DEC + J model also received the highest support when outgroups were included, but the corrected AIC weight was only 41.6%. In both analyses, the DEC + J model estimated a single transition to the Palearctic from an ancestor that occurred in the Nearctic (Fig. 8). Because the jump (J) parameter has been criticized in some literature (see Ree and Sanmartín, 2018), we also evaluated the top-ranking models that did not include J. This was the DEC model in both analyses (Table S3) and a North American ancestor of *Marmota* was still well supported (Figure S1).

4. Discussion

4.1. Origin of marmots

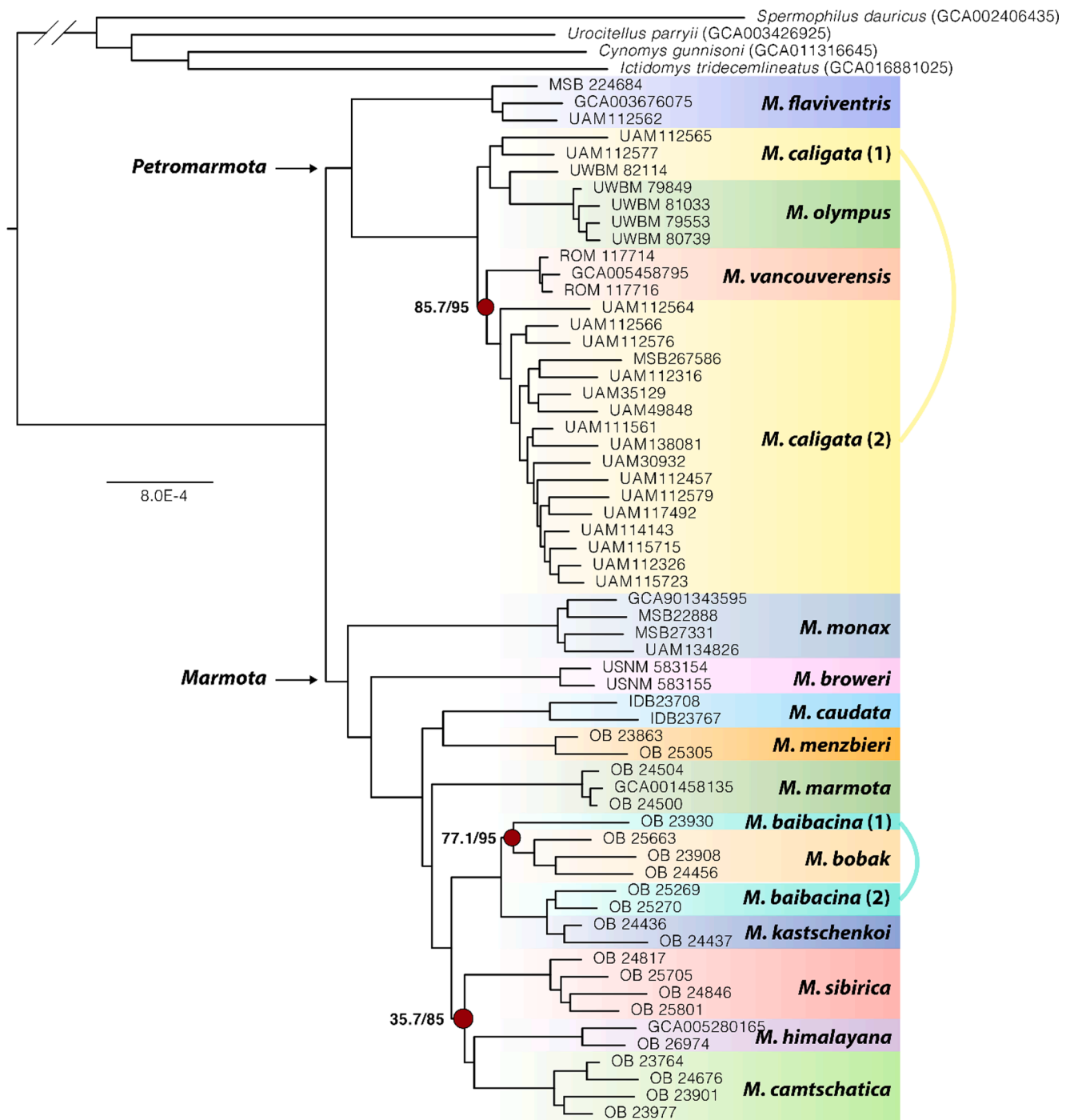
Our phylogenetic results show the Palearctic species unambiguously nested within the Nearctic species, providing strong support for a Nearctic origin of marmots. Our fossil-calibrated phylogeny suggests marmots diverged from other marmotines (chipmunks, prairie dogs, and other ground squirrels) in the Early to Middle Miocene, and the most recent common ancestor of the crown *Marmota* lived sometime in the Late Miocene (Fig. 7). This is contemporaneous with the other major ground squirrel radiations documented in the fossil record (Goodwin, 2008) and with the development of open grassland ecosystems in North America (Strömberg, 2005, 2011).

The world in which marmots first evolved was warmer, wetter, and much more heavily forested than today (Polly et al., 2015; Steinhorsdottir et al., 2021); the earliest open grasslands are thought to have

Table 2

Summary statistics for UCE alignments. The complete matrix is the full UCE dataset after assembly and cleanup. For analysis in IQ-TREE, the dataset was reduced to a 75% complete matrix by excluding any loci present in fewer than 75% of OTUs. The top quartile of these loci, in terms of number of parsimony-informative (PI) sites, was used as input for ASTRAL-III. Numbers in brackets represent the full range of locus lengths (bp) and number of PI sites per locus.

Dataset	UCE loci	Mean locus length (bp)	Mean PI sites/locus	Total concatenated length (bp)	Total PI sites
Complete matrix	4,427	587.8 [109–1152]	19.9 [0–173]	2,602,213	88,252
75% complete matrix (IQ-TREE input)	1,986	621.7 [340–1235]	22.5 [0–134]	1,234,679	44,604
Top quartile (ASTRAL-III input)	496	672.3 [362–1235]	44.2 [31–134]	333,483	21,905



emerged as part of a mosaic landscape of closed forests and “more open woodland-grassland elements” ((Strömberg, 2011), p. 535). If early marmotines were grassland specialists at this time, their habitat would have been fragmented by intervening forests—a situation that would have promoted allopatric speciation and may have spurred the radiations of marmots, chipmunks, prairie dogs, and other genera of ground squirrels. However, it is also possible that early marmotines were forest dwellers and that their radiation resulted from multiple, parallel specializations to open habitat. If marmots did evolve a grazing diet independently of the other ground squirrels, as is perhaps evinced by their

highly distinct cranial, mandibular, and dental morphology (McLean et al., 2019; Zelditch et al., 2017), it may reflect their independent colonization of grasslands in North America. Regardless, this patchy landscape gradually gave way to vast, unforested grasslands in the Middle to Late Miocene as the climate became cooler and drier (Strömberg, 2005), by which time the crown marmotines, including *Marmota*, were well established.

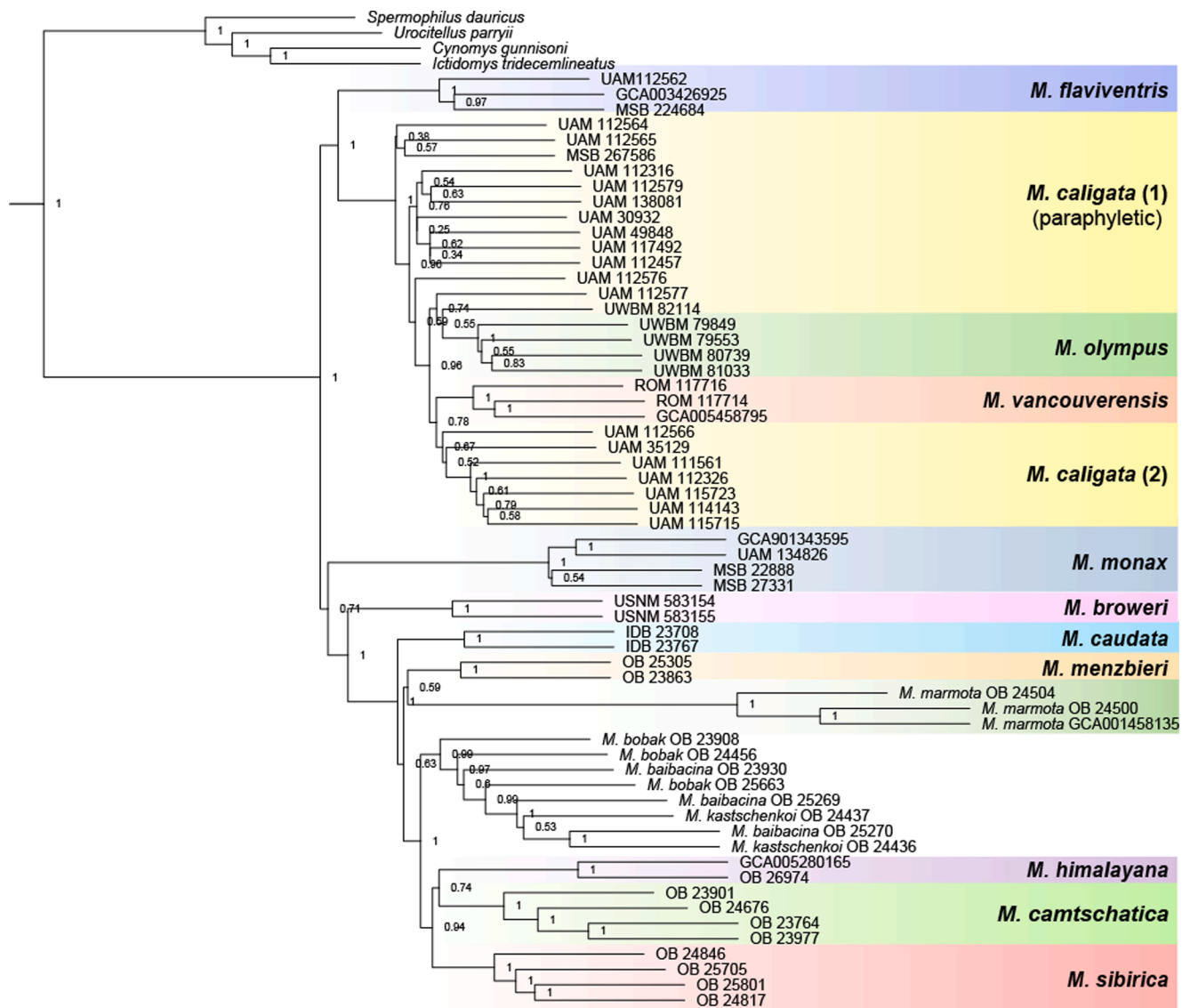


Fig. 5. Phylogeny estimated with ASTRAL-III using individual UCE loci as gene trees. Tip labels as in Fig. 4. Node values represent the local posterior probability of the quadripartition, which has high precision above a threshold of 0.7 (Sayyari and Mirarab 2016).

4.2. Dispersal across Beringia

The North American marmots eventually spread into Eurasia as part of the great faunal interchange across Beringia, which took place over several million years and was facilitated by the intermittent formation of the Bering Land Bridge (Elias et al., 1996; Hopkins, 1959; Sher, 1999). This interchange was bidirectional but consisted primarily of eastward expansions by Eurasian taxa into North America (Beard, 1998; Waltari et al., 2007). The native North American fauna were thought to be limited by the Laurentide ice sheet and a patchwork of Arctic landscapes, both of which impeded northward dispersal. Despite this, the marmotines made at least four westward dispersals. The chipmunk genus *Tamias* recolonized Eurasia in the Oligocene, very soon after its emergence in North America and when Beringia was still heavily forested (Ge et al., 2014; Mercer and Roth, 2003). The partially arboreal *Tamias* was probably the only marmotine group that could disperse across Beringia at this time; the strictly terrestrial genera did not access the Beringian corridor until much later. The land bridge was first submerged (creating the Bering Strait) at the onset of the Miocene, and, upon reemergence some 5–10 Ma, the denuded landscape was colonized by the newly dominant species of open-habitat grasses (Yeakel et al., 2013).

There are nine extant marmot species in Eurasia, and because most phylogenetic analyses have found them to form a monophyletic group, they have typically been considered to descend from a single trans-Beringian dispersal event (Steppan et al., 1999). However, the unstable placement of *M. broweri*, *M. monax*, and *M. marmota* have made this hypothesis somewhat tenuous (Brandler and Lyapunova, 2009). Our results provide strong support for a single westward dispersal event that gave rise to all extant Palearctic species. Furthermore, it has been proposed that *M. broweri*, which is endemic to northern Alaska (Gunderson et al., 2009), may represent a reinvasion of North America by the Palearctic clade (Hoffmann et al., 1979; Stepan et al., 1999). Our ancestral range reconstruction, however, suggests that *M. broweri* has always been in the Nearctic (Fig. 8).

The earliest fossil record of marmots in Eurasia is *M. tologoica* from Transbaikalia, Russia, dated to approximately 2.9 Ma (Erbajeva and Alexeeva, 2009). This species is thought to be ancestral or closely related to *M. sibirica* and *M. camtschatica* (Erbajeva, 2003; Gromov et al., 1965). In northern China, a fossil of *M. himalayana robusta* is dated to a similar time period (but was found in a higher stratum and is thus slightly younger than the oldest *M. tologoica*; Flynn et al., 1997). Our fossil-calibrated tree (Fig. 7) shows the Palearctic species diverging from

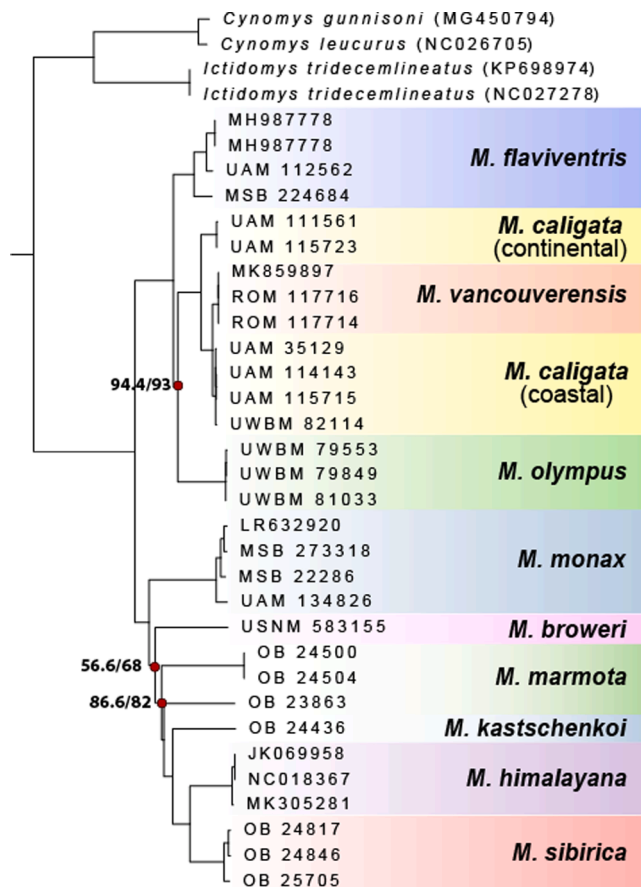


Fig. 6. Maximum-likelihood phylogeny based on mitogenomes estimated in IQ-TREE. Tip labels as in Fig. 4. Branch support values are UFboot/SH-aLRT and are greater than 97/100 unless labeled in red. Coastal and continental refer to mitochondrial clades of *M. caligata* as identified in Kerhoulas et al. (2015).

the Nearctic earlier than this, near the Miocene-Pliocene boundary ~5.3 Ma. This roughly coincides with the widespread transition from C3 to C4 grasses in North America, the latter of which came to dominate the vast Great Plains (Strömberg and McInerney, 2011). Plants using the C4 photosynthetic pathway generally outcompete those using the C3 pathway in warm, arid environments; C3 grasses retreated to higher elevations and to certain lowland steppe regions too cold to support C4 species (Ehleringer and Cerling, 2002; McInerney et al., 2011). Paleoeological evidence points to Beringia being one such lowland region (Kohn, 2010; Gaglioti et al., 2011). The diets of extant marmots consist almost entirely of C3 plants (Armitage, 2000), and presumably they have retained this ancestral preference since their emergence in the Miocene. The ancestral Palearctic lineage may therefore have tracked C3 grasslands northward as C4 species spread across central North America, eventually traversing Beringia and colonizing Eurasia. The patchiness of C3 grasslands may once again have promoted diversification in ground squirrels, though in this period the interceding habitat was open C4 grasslands rather than forests (Strömberg, 2011).

4.3. Palearctic species

The base of the Palearctic clade is entirely resolved in the ML tree estimated from UCEs (Fig. 4), whereas it was previously recovered as a large polytomy (Steppan et al., 1999; Brandler and Lyapunova, 2009). However, there are some topological disagreements between this tree, the species tree (Fig. 5), and the ML tree based on mitogenomes (Fig. 6). Discordance between ML and species trees is not uncommon in studies using UCE datasets (e.g., Esselstyn et al., 2017; Calderón-Acevedo et al.,

2022), probably because gene-tree estimation error is higher for UCEs due to how little phylogenetic signal each individual locus contains (Xi et al., 2015). This may be especially true for UCEs in tetrapods, which seem to exhibit even lower substitution rates than UCEs in, for example, Hymenoptera (Bossert et al., 2021). These nodes represent the deepest divergences occurring in Eurasia following dispersal across the Bering Land Bridge. The epi-platfomal orogeny probably promoted speciation as suggested by Nikol'skii and Rumiantsev (2012), but by filtering apart the ancestral lineage that crossed over from North America rather than the lineage ancestral to all marmots.

Marmota marmota, which is found only in southern Europe and is the westernmost marmot species in Eurasia, has been suggested to be sister to the rest of the Palearctic clade (Kryštufek and Vohralík, 2013). We found support for this in the mitogenomic data (Fig. 6), but the ML tree estimated from UCEs shows *M. menzbieri* + *M. caudata* as sister to the rest of the clade (Fig. 5), whereas the species tree estimated from UCEs shows *M. caudata* alone as sister to the rest of the clade (Fig. 6). The species tree also shows *M. marmota* and *M. menzbieri* to be sister species, but support for this node is weak. Regardless, *M. marmota* is a relatively ancient lineage that apparently traversed nearly all of Eurasia over a short period of time; there are *M. marmota* fossils from the Spanish Pyrenees dating to 1.2 Ma (Gil Bazán, 1997).

Previous studies recovered *M. caudata* and *M. menzbieri* as sister species in phylogenies based on mtDNA but not in those based on nuclear loci, which has been interpreted as suggesting they are not closely related species but underwent an ancient mitochondrial introgression event (Brandler et al., 2010). In contrast, our nuclear (UCE) phylogeny unequivocally recovers them as sister to each other. This is notable because the two species are peripatric and highly morphologically distinct—*M. caudata* has the largest body mass and longest tail of any marmot, and *M. menzbieri* has a relatively short tail and the second-smallest body size (Armitage, 2000).

The *M. himalayana* + *M. camtschatica* + *M. sibirica* relationships could not be resolved, consistent with other studies finding discordant or poorly supported topologies (Brandler and Lyapunova, 2009). This could reflect previous or ongoing gene flow among the three species, but given their allopatric ranges and relatively distinct habitat preferences, it seems more likely that rapid radiation across eastern Asia led to incomplete lineage sorting. Moreover, *M. sibirica* actively hybridizes with *M. baibacina* where the two are sympatric (Brandler et al., 2021), but the two species are clearly separated in all phylogenies.

A single *M. baibacina centralis* individual (IDB 23930) from the Tien Shen clustered with a monophyletic *M. bobak* group. Although *M. baibacina baibacina* has a narrow zone of contact with *M. bobak shagadensis* (Nikol'skii et al., 1983), the Tien Shen population of *M. baibacina centralis* is completely geographically isolated from all *M. bobak* (Fig. 1). The subspecies *M. baibacina centralis* is very similar to other *M. baibacina*, but it exhibits morphological features intermediate between *M. baibacina* and *M. bobak*. Galkina (1970) suggested that *M. b. centralis* should be reclassified as either a subspecies of *M. bobak* or elevated to species, and our results support either of these options over the current taxonomy. The other two *M. baibacina* individuals in this study are from Altai Mountains in Mongolia and are sister to *M. kastschenkoi*. *Marmota kastschenkoi* was previously considered a subspecies of *M. baibacina*, but was recently elevated to species based on its distinct diploid chromosome count (Brandler, 2003; Brandler et al., 2008).

4.4. Nearctic species

The Alaska marmot (*M. broweri*) and woodchuck (*M. monax*) are unambiguously recovered in the subgenus *Marmota* (Fig. 4). This topology implies divergence of the Palearctic and Nearctic clades occurred within North America and that dispersal across Beringia followed sometime after the *M. broweri* and *M. monax* lineages diverged. Given the distribution of *M. broweri*, it is likely that the common ancestor of *M. broweri* and the Palearctic species was already in or near Beringia

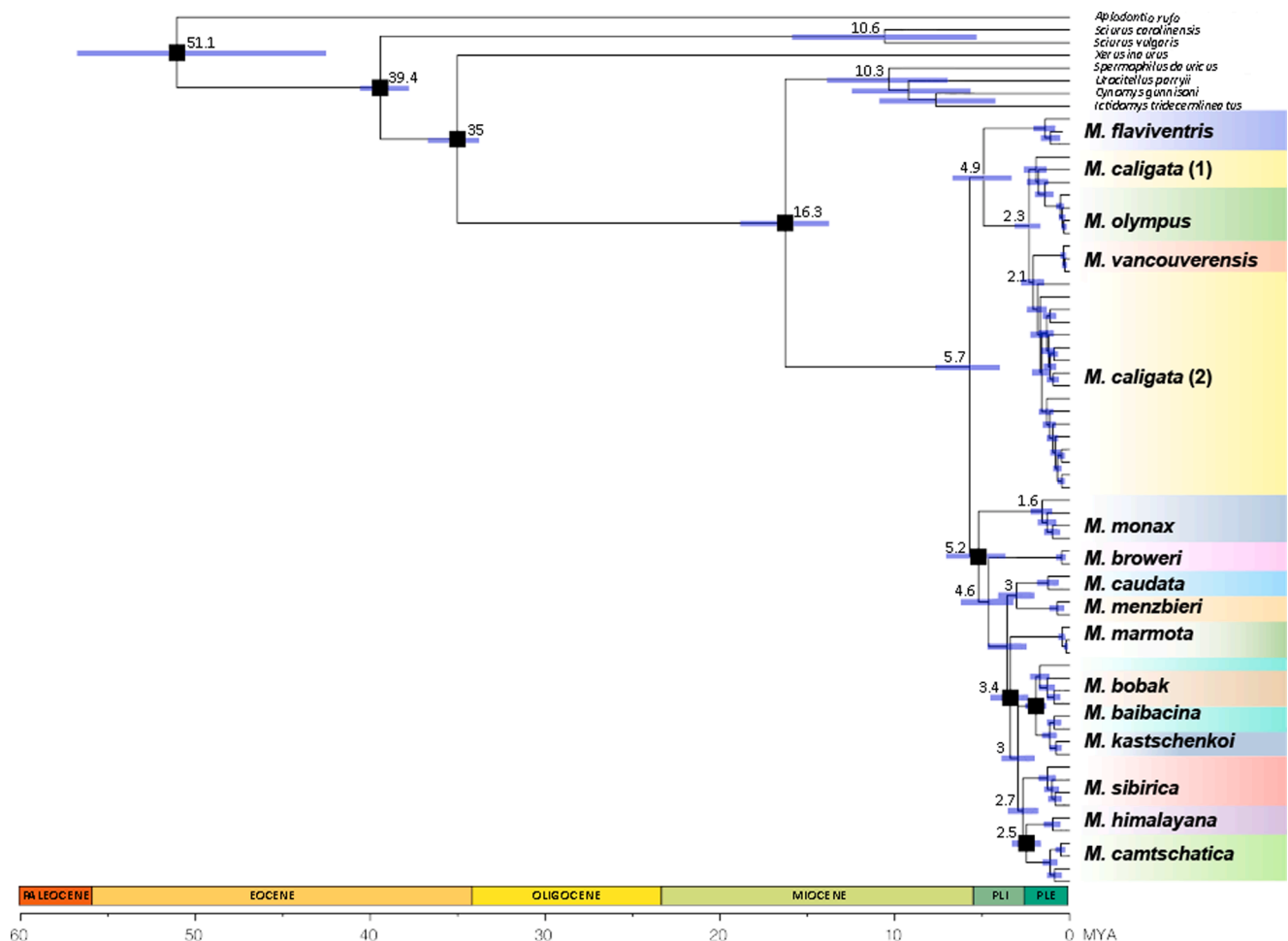


Fig. 7. Fossil-calibrated tree estimated in MCMCTree from the 75% complete UCE dataset. Blue bars represent 95% confidence intervals for divergence dates, and black squares correspond to fossil calibrations (Table 1). Dates on nodes <2 Ma have been removed for clarity.

prior to its divergence. This hypothesis is supported by our ancestral range reconstruction analysis (Fig. 8), although it does not conclusively reject the possibility of marmots in the Palearctic prior to the divergence of *M. broweri*. Moreover, speciation events around this time could have occurred in the vast regions of Pliocene Beringia that are now underwater and not classifiable as Nearctic or Palearctic. Sher (1999) emphasizes that Beringia was more a region unto itself than a simple bridge from one region to another, and that it was probably a center of diversification in its own right.

There is significant discordance in the *M. caligata* superspecies complex between the trees estimated from UCEs and mitogenomes (Figs. 4 and 6), and *M. caligata* is paraphyletic in both. We recovered the same mitochondrial clades identified by Kerhoulas et al. (2015) in our mitogenome reconstructions, with the coastal *M. caligata* clade being sister to *M. vancouverensis*, and the continental clade sister to *M. caligata* coastal + *M. vancouverensis*. This topology is not indicated in the UCE data, and it likely reflects relatively recent biogeographic events of the Pleistocene. Specifically, the two clades may represent separate refugial populations from the LGM—the continental clade deriving from a refugial population south of the Laurentide ice sheet, and the coastal clade deriving from a refugial population near the Cascade Mountains or Alexander Archipelago (Kerhoulas et al., 2015). That the latter is sister to *M. vancouverensis* suggests this refugial population (and not the other) interbred with *M. vancouverensis* during or soon after the LGM. In contrast, the UCE-based phylogeny shows two clusters of *M. caligata* unrelated to the mitochondrial clades (Fig. 4). The three *M. caligata* individuals from Washington state are paraphyletic with *M. olympus*,

whereas all the other *M. caligata* are sister to *M. vancouverensis*. Our time-calibrated phylogeny estimates these two groups diverged ~2.3 Ma (Fig. 7), near the onset of the Pleistocene. The timing of this divergence and the modern geographical species ranges suggests formation of the Laurentide ice sheet may have been a vicariant event that split the ancestral lineage into two. Delimiting the genetic and geographical range of *M. caligata*, and determining whether there is a cryptic species in the Cascade Mountains of Washington, is of interest given the conservation status of *M. vancouverensis*. This species numbers only about 200 individuals and is considered Critically Endangered by the IUCN; it has been suggested that interbreeding with *M. caligata* may be necessary to genetically rescue *M. vancouverensis* in the future (Lyapunova and Brandler, 2002). In this context it would be useful to better understand the genetic relationship between *M. vancouverensis* and *M. caligata* individuals from the coastal mitochondrial clade versus *M. caligata* individuals from the sister nuclear clade.

5. Conclusion

Consistent with fossil evidence, our genetic data indicate marmots most likely evolved in North America in the Middle Miocene. This is around the same time that other crown marmotine genera arose and is roughly coincident with the rise of grassland ecosystems in North America (i.e., the Great Plains). Marmots dispersed across the Bering Land Bridge ~3–4 Ma, and this may have been driven by the displacement of C3 grasses by C4 grasses into colder and drier regions. The placement of the Nearctic *M. broweri* and *M. monax* suggests divergence

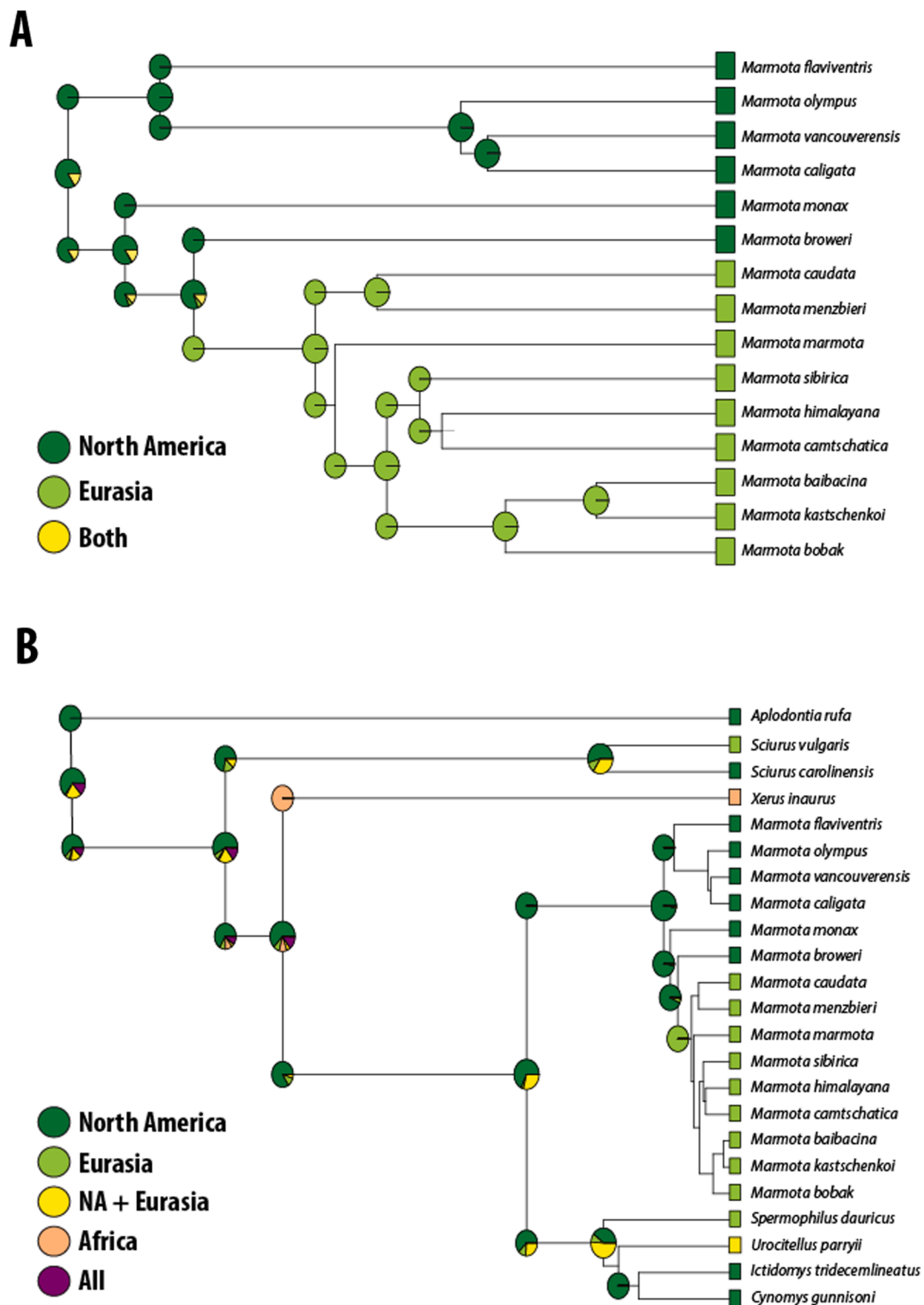


Fig. 8. Ancestral range reconstructions derived from the divergence time tree (Fig. 7) with outgroups excluded (A) and outgroups included (B). Range likelihoods from the highest ranked dispersal model (DEC + J; Table S3) are shown as pie charts at the nodes, and current distributions are shown as boxes on each tip. Pie charts on all nodes dated to <2 Ma have been removed for clarity.

of the two subgenera occurred in North America. We recovered non-monophyletic clusters of *M. baibacina* and *M. caligata*. For the latter, we found highly suggestive evidence of a second, cryptic species in the Cascade Mountains of Washington state that is paraphyletic with respect to *M. olympus*. More intensive taxon and geographic sampling is desirable to delimit the taxonomic and geographical boundaries of the two groups of *M. caligata*.

CRediT authorship contribution statement

Kendall K. Mills: Conceptualization, Methodology, Software, Funding acquisition, Project administration, Visualization, Writing – original draft, Writing – review & editing, Data curation, Formal analysis. **Kathryn M. Everson:** Conceptualization, Resources, Methodology, Software, Writing – original draft, Writing – review & editing,

Visualization, Formal analysis, Investigation. **Kyndall B.P. Hildebrandt:** Resources, Methodology. **Oleg V. Brandler:** Resources, Writing – review & editing, Investigation. **Scott J. Steppan:** Resources, Writing – review & editing, Investigation. **Link E. Olson:** Conceptualization, Resources, Supervision, Project administration, Funding acquisition, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

In addition to the many scientific collectors without whose efforts the specimens in our study would not have been available, we are grateful to the following individuals and institutions for providing tissues: Aren M. Gunderson (University of Alaska Museum, Fairbanks, AK); Joseph A. Cook and Jonathan L. Dunnum (Museum of Southwestern Biology, University of New Mexico, Albuquerque, NM); Sharlene E. Santana and Jeffrey E. Bradley (Burke Museum, University of Washington, Seattle, WA); Koltzov Institute of Developmental Biology (Moscow, Russia); Burton K. Lim (Royal Ontario Museum, Toronto, ON); and Neal Woodman, Alfred L. Gardner, Robert D. Fisher, and Suzanne C. Peurach (U.S. National Museum of Natural History, Smithsonian, Washington, DC).

Kirby L. Mills kindly created the map used in Figure 1. We also thank Bryan S. McLean for sharing then-unpublished UCE sequences and P. David Polly for providing an exhaustive directory of dated *Marmota* fossils that facilitated the time-calibrated phylogenies. For their assistance in generating and analyzing the UCE data, we thank Bryan S. McLean, Travis C. Glenn, Brant C. Faircloth, Michael J. Alfaro, and Janet C. Buckner. Aren M. Gunderson, Katie E. Rubin, Krystal R. Fales, and Nick J. Kerhoulas performed invaluable, foundational work in the long process of amassing the marmot tissues and sequence data necessary for this project. Devin M. Drown, Jesse N. Weber, Vadim B. Fedorov, and Diane E. Wolf provided valuable feedback on earlier versions of this manuscript, and we are particularly grateful for the helpful comments on the first submission provided by Conrad A. Matthee and an anonymous reviewer. This research was supported by Alaska Department of Fish and Game State Wildlife Grants T-1, T-6, and T-9 awarded to LEO; a University of Alaska Fairbanks Hopkins Fellowship and a Grant-in-Aid from the Society of Systematic Biologists awarded to KKM; Alaska EPSCoR NSF Award #EPS-0701898; the University of Alaska Museum; IDB RAS Government Basic Research Program [No. 0088-2021-0019]; and a generous gift from the Jay Pritzker Foundation. Services provided at the University of Alaska Institute of Arctic Biology's Genomics Core Lab were supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health [grant number 2P20GM103395].

Finally, we are deeply indebted to Ken Armitage, not only for his thoughtful and generous advice over the years, but also for his lifelong work with marmots that continues to influence the authors of this paper and inspire so many other marmotologists.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ympev.2023.107785>.

References

Anisimova, M., Gil, M., Dufayard, J.F., Dessimoz, C., Gascuel, O., 2011. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Syst. Biol.* 60, 685–699.

- Armitage, K.B., 1981. Sociality as a life-history tactic of ground squirrels. *Oecologia* 48, 36–49.
- Armitage, K.B., 2000. The evolution, ecology, and systematics of marmots. *Oecologia Montana* 9, 1–18.
- Armitage, K.B., 2009. Fur color diversity in marmots. *Ethol. Ecol. Evol.* 21, 183–194.
- Arnold, W., 1992. Adaptation to the cold. The physiology of marmot hibernation. Proceedings of the 1st international symposium on Alpine marmot and genera *Marmota*, Turin, 31–40.
- Azzaroli, A., De Giuli, C., Ficcarelli, G., Torre, D., 1988. Late Pliocene to early mid-Pleistocene mammals in Eurasia: faunal succession and dispersal events. *Palaeogeography, Palaeoclimatology, Palaeoecology* 66, 77–100.
- Bai, L., Liu, B., Ji, C., Zhao, S., Liu, S., Wang, R., Wang, W., Yao, P., Li, X., Fu, X., Yu, H., 2019. Hypoxic and cold adaptation insights from the Himalayan marmot genome. *Science* 11, 519–530.
- Barash, D.P., 1974. The evolution of marmot societies: a general theory. *Science* 185, 415–420.
- Beard, K.C., 1998. East of Eden: Asia as an important center of taxonomic origination in mammalian evolution. *Bulletin of the Carnegie Museum of Natural History* 34, 5–39.
- Black, C.C., 1963. A review of the North American Tertiary Sciuridae. *Bull. Mus. Comp. Zool.* 130, 159–166.
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.
- Bossert, S., Murray, E.A., Pauly, A., Chernyshov, K., Brady, S.G., Danforth, B.N., 2021. Gene tree estimation error with ultraconserved elements: an empirical study on *Pseudapis* bees. *Syst. Biol.* 70, 803–821.
- Brandler, O.V., Kapustina, S.Y., Nikol'skii, A.A., Kolesnikov, V.V., Badmaev, B.B., Adiya, Y., 2021. A study of hybridization between *Marmota baibacina* and *M. sibirica* in their secondary contact zone in Mongolian Altai. *Frontiers in Ecology and Evolution* 9, 1–16.
- Brandler, O.V., Lyapunova, E.A., 2009. Molecular phylogenies of the genus *Marmota* (Rodentia Sciuridae): comparative analysis. *Ethol. Ecol. Evol.* 21, 289–298.
- Brandler, O.V., Lyapunova, E.A., Boeskorov, G.G., 2008. Comparative karyology of Palearctic marmots (*Marmota*, Sciuridae, Rodentia). *Mammalia* 72, 24–34.
- Brandler, O.V., Lyapunova, E.A., Bannikova, A.A., Kramerov, D.A., 2010. Phylogeny and systematics of marmots (*Marmota*, Sciuridae, Rodentia) inferred from inter-SINE PCR data. *Russ. J. Genet.* 46, 283–292.
- Brandler, O.V., 2003. On species status of the forest-steppe marmot *Marmota kastschenkoi* (Rodentia, Marmotinae). *Zoologicheskii Zhurnal* (In Russian) 82, 1498–1505.
- Calderón-Acevedo, C.A., Bagley, J.C., Muchhala, N., 2022. Genome-wide ultraconserved elements resolve phylogenetic relationships and biogeographic history among Neotropical leaf-nosed bats in the genus *Anoura* (Phyllostomidae). *Mol. Phylogenet. Evol.* 167, 107356.
- Cardini, A., Thorington, R.W., 2006. Postnatal ontogeny of marmot (Rodentia, Sciuridae) crania: allometric trajectories and species divergence. *J. Mammal.* 87, 201–215.
- Cardini, A., 2003. The geometry of the marmot (Rodentia: Sciuridae) mandible: phylogeny and patterns of morphological evolution. *Syst. Biol.* 52, 186–205.
- Cardini, A., Hoffmann, R.S., Thorington, R.W., 2005. Morphological evolution in marmots (Rodentia, Sciuridae): size and shape of the dorsal and lateral surfaces of the cranium. *J. Zool. Syst. Evol. Res.* 43, 258–268.
- Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540–552.
- Cerling, T.E., Harris, J.M., MacFadden, B.J., Leakey, M.G., Quade, J., Eisenmann, V., Ehleringer, J.R., 1997. Global vegetation change through the Miocene/Pliocene boundary. *Nature* 389, 153–158.
- Costeur, L., Schneider, M., 2011. Catalogue of the Eocene mammal types of the Natural History Museum Basel. *Carnets de géologie (Notebooks on geology)*.
- Ehleringer, J.R., Cerling, T.E., 2002. C3 and C4 photosynthesis. *Encycl. Global Environ. Change* 2, 186–190.
- Elias, S.A., Short, S.K., Nelson, C.H., Birks, H.H., 1996. Life and times of the Bering land bridge. *Nature* 382, 60–63.
- Emry, R.J., Korth, W.W., 1996. The Chadronian squirrel “*Sciurus jeffersoni*” Douglass, 1901: a new generic name, new material, and its bearing on the early evolution of Sciuridae (Rodentia). *J. Vertebr. Paleontol.* 16, 775–780.
- Erbajeva, M.A., 2003. History, evolutionary development and systematics of marmots (Rodentia, Sciuridae) in Transbaikalia. *Russian Journal of Theriology* 2, 33–42.
- Erbajeva, M.A., Alexeeva, N.V., 2009. Pliocene–Recent Holarctic marmots: an overview. *Ethol. Ecol. Evol.* 22, 339–348.
- Esselstyn, J.A., Oliveros, C.H., Swanson, M.T., Faircloth, B.C., 2017. Investigating difficult nodes in the placental mammal tree with expanded taxon sampling and thousands of ultraconserved elements. *Genome Biol. Evol.* 9, 2308–2321.
- Faircloth, B.C. Illumiprocessor: a trimmomatic wrapper for parallel adapter and quality. *Faircloth, B.C., 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics* 32, 786–788.
- Faircloth, B.C., Glenn, T.C., 2012. Not all sequence tags are created equal: designing and validating sequence identification tags robust to indels. *PLoS ONE* 7, e42543.
- Faircloth, B.C., McCormack, J.E., Crawford, N.G., Harvey, M.G., Brumfield, R.T., Glenn, T.C., 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Syst. Biol.* 61, 717–726.
- Ferrusquia-Villafranca, I., Flynn, L.J., Ruiz-Gonzalez, J.E., Torres-Hernandez, J.R., Martinez-Hernandez, E., 2018. New Eocene rodents from northwestern Oaxaca, southeastern Mexico, and their paleobiological significance. *J. Vertebr. Paleontol.* 38, 1514615.
- Flynn, L.J., Wu, W.Y., Downs III, W.R., 1997. Dating vertebrate microfaunas in the late Neogene record of northern China. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 133, 227–242.

- Gaglioti, B.V., Barnes, B.M., Zazula, G.D., Beaudoin, A.B., Wooller, M.J., 2011. Late Pleistocene paleoecology of arctic ground squirrel (*Urocitellus parryi*) caches and nests from Interior Alaska's mammoth steppe ecosystem, USA. *Quat. Res.* 76, 373–382.
- Galkina, L. I. (1970). Ocenka nekotorykh sistematicheskikh osobennostei' surkov i geograficheskaya izmenchivost' gornoziatskogo surka (*Marmota baibacina* Kastschenko, 1899) [Assessment of some systematic characteristics of marmots and the geographical variation of Asian mountain marmot (*Marmota baibacina* Kastschenko, 1899)]. Fauna Sibiri, (Fauna of Siberia), Nobosibirsk: Nauka 1970, 267–279.
- Ge, D.Y., Liu, X., Lv, X.F., Zhang, Z.Q., Xia, L., Yang, Q.S., 2014. Historical biogeography and body form evolution of ground squirrels (Sciuridae: Xerinae). *Evolutionary Biology* 41, 99–114.
- Gil Bazán, E., 1997. Presencia de *Marmota marmota* en los relleños kársticos de Atapuerca (Burgos, España). *Geogaceta* 22, 71–72.
- Gladman, S., Seemann, T., 2008. Velvet Optimiser: for automatically optimising the primary parameter options for the Velvet de novo sequence assembler. Victorian Bioinformatics Consortium, Monash University, Melbourne, Australia.
- Goodwin, H.T., 2008. Sciuridae. In: Janis, C.M., Gunnell, G.F., Uhen, M.D. (Eds.), *Evolution of tertiary mammals of North America: small mammals, xenarthrans, and marine mammals*. Cambridge University Press, Cambridge (MA), pp. 355–376.
- Gordon, C.L., Czaplewski, N.J., 2000. A fossil marmot from the late Miocene of western Oklahoma. *Oklahoma Geology Notes* 60, 28–32.
- Gromov, I. M., Bibikov, D. I., Kalabuchov, N. I., & Mejer, M. N. (1965). Fauna SSSR.: Mlekopitaiushchie. Nazemny'e belichi' (Marmotinae) [Mammals. Ground squirrels (Marmotinae)]. Publishing house Nauka, Moscow-Leningrad.
- Gunderson, A.M., Jacobsen, B.K., Olson, L.E., 2009. Revised distribution of the Alaska marmot, *Marmota broweri*, and confirmation of parapatry with hoary marmots. *J. Mammal.* 90, 859–869.
- Hafner, D.J., 1984. Evolutionary relationships of the Nearctic Sciuridae. In: Murie, J.O., Michener, G.R. (Eds.), *The Biology of Ground-Dwelling Squirrels*. University of Nebraska Press.
- Harrison, R.G., Bogdanowicz, S.M., Hoffmann, R.S., Yensen, E., Sherman, P.W., 2003. Phylogeny and evolutionary history of the ground squirrels (Rodentia: Marmotinae). *J. Mamm. Evol.* 10, 249–276.
- Hayssen, V., 2008. Patterns of body and tail length and body mass in Sciuridae. *J. Mammal.* 89, 852–873.
- Heissig, K., 2003. Origin and early dispersal of the squirrels and their relatives. *Deinsea* 10, 277–286.
- Helgen, K.M., Cole, F.R., Helgen, L.E., Wilson, D.E., 2009. Generic revision in the Holarctic ground squirrel genus *Spermophilus*. *J. Mammal.* 90, 270–305.
- Herron, M.D., Castoe, T.A., Parkinson, C.L., 2004. Sciurid phylogeny and the paraphyly of Holarctic ground squirrels (*Spermophilus*). *Mol. Phylogenet. Evol.* 31, 1015–1030.
- Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2017. UFBoot2: Improving the ultrafast bootstrap approximation. *BioRxiv* 35, 518–522.
- Hoffmann, R.S., Koeppl, J.W., Nadler, C.F., 1979. The relationships of the amphiberian marmots (Mammalia, Sciuridae). Occasional papers of the Museum of Natural History, University of Kansas, 83, 1–56.
- Hoffmann, R.S., Nadler, C.F., 1968. Chromosomes and systematics of some North American species of the genus *Marmota* (Rodentia: Sciuridae). *Experientia* 24, 740–742.
- Hopkins, D.M., 1959. Cenozoic history of the Bering Land Bridge: The seaway between the Pacific and Arctic basins has often been a land route between Siberia and Alaska. *Science* 129, 1519–1528.
- Ivy, L.D., 1990. Systematics of Late Paleocene and Early Eocene Rodentia (Mammalia) from the Clarks Fork Basin, Wyoming. Contributions from the Museum of Paleontology, University of Michigan 28, 21–70.
- Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780.
- Kerhoulas, N.J., Gunderson, A.M., Olson, L.E., 2015. Complex history of isolation and gene flow in hoary, Olympic, and endangered Vancouver Island marmots. *J. Mammal.* 96, 810–826.
- Kohn, M.J., 2010. Carbon isotope compositions of terrestrial C3 plants as indicators of (paleo) ecology and (paleo) climate. *Proc. Natl. Acad. Sci. USA* 107, 19691–19695.
- Korth, W.W., 1994. The Tertiary record of rodents in North America. In: *Topics in Geobiology*, vol. 12. Springer Science & Business Media.
- Kruckenhauser, L., Pinsker, W., Haring, E., Arnold, W., 1999. Marmot phylogeny revisited: molecular evidence for a diphyletic origin of sociality. *J. Zool. Syst. Evol. Res.* 37, 49–56.
- Kryštufek, B., Vohralík, V., 2013. Taxonomic revision of the Palaearctic rodents (Rodentia). Part 2. Sciuridae: *Urocitellus*, *Marmota* and *Sciurotamias*. *Lynx series nova* 44, 27–138.
- Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2017. Partitionfinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* 34, 772–773.
- Lyapunova, E.A., Brandler, O.V., 2002. Genetic approaches for rehabilitation of Vancouver Marmot population. In: IV Marmot World Conference: Montreux, Switzerland, 15–17 September 2002: Abstracts. International Marmot Network, pp. 86–87.
- Marsh, O.C., 1871. Notice of some new fossil mammals and birds, from the Tertiary Formation of the West. *Am. J. Sci. Arts* 2, 120–127.
- Matzke, N.J., 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Syst. Biol.* 63, 951–970.
- McInerney, F.A., Strömberg, C.A., White, J.W., 2011. The Neogene transition from C3 to C4 grasslands in North America: stable carbon isotope ratios of fossil phytoliths. *Paleobiology* 37, 23–49.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303.
- McKenna, M.C., Bell, S.K., 1997. *Classification of Mammals: Above the Species Level*. Columbia University Press.
- McLean, B.S., Bell, K.C., Allen, J.M., Helgen, K.M., Cook, J.A., 2019. Impacts of inference method and data set filtering on phylogenomic resolution in a rapid radiation of ground squirrels (Xerinae: Marmotini). *Syst. Biol.* 68, 298–316.
- McLean, B.S., Helgen, K.M., Goodwin, H.T., Cook, J.A., 2018. Trait-specific processes of convergence and conservatism shape ecomorphological evolution in ground-dwelling squirrels. *Evolution* 72, 473–489.
- Meiklejohn, K.A., Faircloth, B.C., Glenn, T.C., Kimball, R.T., Braun, E.L., 2016. Analysis of a rapid evolutionary radiation using ultraconserved elements: evidence for a bias in some multispecies coalescent methods. *Syst. Biol.* 65, 612–627.
- Menéndez, I., Gómez Cano, A.R., Cantalapiedra, J.L., Peláez-Campomanes, P., Álvarez-Sierra, M.A., Hernández Fernández, M., 2021. A multi-layered approach to the diversification of squirrels. *Mammal Rev.* 51, 66–81.
- Mercur, J.M., Roth, V.L., 2003. The effects of Cenozoic global change on squirrel phylogeny. *Science* 299, 1568–1572.
- Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274.
- Nikol'skii A. A., 2014. Vidovaya spetsifika i geograficheskaya izmenchivost' zvukovogo signala surkov (*Marmota*, Sciuridae, Rodentia) Evrazi [Species specificity and geographical variation of Eurasian marmots' alarm call (*Marmota*, Sciuridae, Rodentia)]. *Zoologicheskyy Zhurnal*, 93, 1026–1026 (in Russian).
- Nikol'skii, A.A., Yanina, I.Y., Rutovskaia, M.V., Formozov, N.A., 1983. The variation of sound call of *Marmota bobac* and *M. baibacina* (Sciuridae, Rodentia) in the zone of secondary contact. *Zoologicheskyy Zhurnal* 62, 1258–1266.
- Polly, P.D., Cardini, A., Davis, E.B., Stepan, S., Cox, P.G., Hautier, L., 2015. Marmot evolution and global change in the past 10 million years. In: Cox, P.G., Hautier, L. (Eds.), *Evolution of the Rodents: Advances in Phylogeny, Paleontology and Functional Morphology*. Cambridge University Press, Cambridge, United Kingdom, pp. 246–276.
- Peter, B.D., 2004. African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth and Planetary Science Letters* 220, 3–24.
- Prothero, D.R., 1994. The late Eocene-Oligocene extinctions. *Annu. Rev. Earth Planet. Sci.* 22, 145–165.
- Rabiee, M., Sayyari, E., Mirarab, S., 2019. Multi-allele species reconstruction using ASTRAL. *Mol. Phylogenet. Evol.* 130, 286–296.
- Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67, 901–904.
- Rausch, R.L., Rausch, V.R., 1971. The somatic chromosomes of some North American marmots (Sciuridae), with remarks on the relationships of *Marmota broweri* Hall and Gilmore. *Mammalia* 35, 85–101.
- Ree, R.H., Sanmartín, I., 2018. Conceptual and statistical problems with the DEC+J model of founder-event speciation and its comparison with DEC via model selection. *J. Biogeogr.* 45, 741–749.
- Sayyari, E., Mirarab, S., 2016. Fast coalescent-based computation of local branch support from quartet frequencies. *Mol. Biol. Evol.* 33, 1654–1668.
- Sher, A., 1999. Traffic lights at the Beringian crossroads. *Nature* 397, 103–104.
- Sinitisa, M.V., 2018. Phylogenetic position of *Sinotamias* and the early evolution of Marmotini (Rodentia, Sciuridae, Xerinae). *J. Vertebr. Paleontol.* 38, 1419251.
- Steinhorsdottir, M., Coxall, H.K., De Boer, A.M., Huber, M., Barbolini, N., Bradshaw, C. D., Burls, N.J., Feakins, S.J., Gasson, E., Henderiks, J., Holbourn, A.E., 2021. The Miocene: The future of the past. *Paleoceanogr. Paleoclimatol.* 36 e2020PA004037.
- Steppan, S.J., Akhverdyan, M.R., Lyapunova, E.A., Fraser, D.G., Vorontsov, N.N., Hoffmann, R.S., Braun, M.J., 1999. Molecular phylogeny of the marmots (Rodentia: Sciuridae): tests of evolutionary and biogeographic hypotheses. *Syst. Biol.* 48, 715–734.
- Steppan, S.J., Kenagy, G.J., Zawadzki, C., Robles, R., Lyapunova, E.A., Hoffmann, R.S., 2011. Molecular data resolve placement of the Olympic marmot and estimate dates of trans-Beringian interchange. *J. Mammal.* 92, 1028–1037.
- Steppan, S.J., Storz, B.L., Hoffmann, R.S., 2004. Nuclear DNA phylogeny of the squirrels (Mammalia: Rodentia) and the evolution of arboreality from c-myc and RAG1. *Mol. Phylogenet. Evol.* 30, 703–719.
- Strömberg, C.A., McInerney, F.A., 2011. The Neogene transition from C3 to C4 grasslands in North America: assemblage analysis of fossil phytoliths. *Paleobiology* 37, 50–71.
- Strömberg, C.A., 2005. Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the Cenozoic of North America. *Proceedings of the National Academy of Sciences USA* 102, 11980–11984.
- Strömberg, C.A., 2011. Evolution of grasses and grassland ecosystems. *Annu. Rev. Earth Planet. Sci.* 39, 517–544.
- Tagliacollo, V.A., Lanfear, R., 2018. Estimating improved partitioning schemes for ultraconserved elements. *Mol. Biol. Evol.* 35, 1798–1811.
- Thomas, W.K., Martin, S.L., 1993. A recent origin of marmots. *Mol. Phylogenet. Evol.* 2, 330–336.
- Van Kolschoten, T., Markova, A.K., 2005. Response of the European mammalian fauna to the mid-Pleistocene transition. *Geological Society, London, Special Publications* 247, 221–229.
- Waltari, E., Hoberg, E.P., Lessa, E.P., Cook, J.A., 2007. Eastward Ho: phylogeographical perspectives on colonization of hosts and parasites across the Beringian nexus. *J. Biogeogr.* 34, 561–574.

- Xi, Z., Liu, L., Davis, C.C., 2015. Genes with minimal phylogenetic information are problematic for coalescent analyses when gene tree estimation is biased. *Mol. Phylogenet. Evol.* 92, 63–71.
- Yang, Z., Rannala, B., 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* 23, 212–226.
- Yang, Z., 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591.
- Yeakel, J.D., Guimarães Jr., P.R., Bocherens, H., Koch, P.L., 2013. The impact of climate change on the structure of Pleistocene food webs across the mammoth steppe. *Proc. R. Soc. B Biol. Sci.* 280, 20130239.
- Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* 292, 686–693.
- Zelditch, M.L., Ye, J., Mitchell, J.S., Swiderski, D.L., 2017. Rare ecomorphological convergence on a complex adaptive landscape: body size and diet mediate evolution of jaw shape in squirrels (Sciuridae). *Evolution* 71, 633–649.
- Zerbino, D.R., Birney, E., 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829.
- Zhang, C., Rabiee, M., Sayyari, E., Mirarab, S., 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19, 15–30.
- Zimina, R.P., Gerasimov, I.P., 1973. The periglacial expansion of marmots (*Marmota*) in Middle Europe during Late Pleistocene. *J. Mammal.* 54, 327–340.