Out of Tibet: Genomic Perspectives on the Evolutionary History of Extant Pikas

XiaoYun Wang, Dan Liang, Wei Jin, MingKun Tang, Shalayiwu, ShaoYing Liu, and Peng Zhang

1State Key Laboratory of Biocontrol, College of Ecology and Evolution, School of Life Sciences, Sun Yat-sen University, Guangzhou, China
2Xinjiang Locust Control and Rodent Control Command Office, Urumchi, China
3Xinjiang Locust Control and Rodent Control Command Office, Urumchi, China

*Corresponding authors: E-mails: shaoyliu@163.com; zhangp35@mail.sysu.edu.cn.

Abstract

Pikas are widely distributed in the Northern Hemisphere and are highly adapted to cold and alpine environments. They are one of the most complex and problematic groups in mammalian systematics, and the origin and evolutionary history of extant pikas remain controversial. In this study, we sequenced the whole coding sequences of 105 pika samples (29 named species and 1 putative new species) and obtained DNA data for more than 10,000 genes. Our phylogenomic analyses recognized four subgenera of extant pikas: Conothoa, Conothoa, Ochotona, and Pika. The interrelationships between the four subgenera were strongly resolved as (Conothoa, (Ochotona, Pika)), with the mountain group Conothoa being the sister group of all other pikas. Our divergence time and phylogeographic analyses indicated that the last common ancestor of extant pikas first occurred on in the middle Miocene, ~14 Ma. The emergence of opportunities related to the climate, food supply, and spreading paths in concert promoted the dispersal of pikas from the Qinghai-Tibetan Plateau (QTP) to other parts of Eurasia and North America. We found that the genes that were positively selected in the early evolution of pikas were most concentrated in functional categories related to cold tolerance. These results suggest that the QTP may have served as a training ground for cold tolerance in early pikas, which gives pikas a great advantage when the climate continued to cool after the middle Miocene. Our study highlights the importance of the QTP as a center of origin for many cold-adapted animals.

Key words: Ochotona, phylogeny, divergence time, biogeography, cold adaptation.

Introduction

Pikas (Lagomorpha: Ochotonidae) are small rabbit-like mammals, with short limbs and rounded ears. They used to be widely distributed and highly diversified across Eurasia, Africa, and New World but most ancient pikas went extinct, leaving only a single extant genus, Ochotona. Extant pikas prefer cold climates. They are iconic cold-adapted animals and can be found in various cold environments, such as plateaus and cold tundra and steppe environments. Pikas graze on grasses, flowers, and young stems and play an important role in these natural ecosystems to help stabilize vegetative communities and promote plant growth (Delibes-Mateos et al. 2011; Wilson and Smith 2015; Smith et al. 2018).

The extant pikas include 34 nominal species and are tentatively divided into 5 subgenera: Conothoa, Conothoa, Lagotona, Ochotona, and Pika (Hoffmann and Smith 2005; Lissovsky 2014; Liu et al. 2017; Smith et al. 2018). Conothoa and Ochotona are mainly distributed on the Qinghai-Tibetan Plateau (QTP) and in its vicinity and adapt well to life at higher elevations. Subgenus Conothoa mainly lives in the eastern margin of the QTP, a transition region from high altitude to low elevation. The other two subgenera, Lagotona and Pika, are distributed in cold mountain and steppe areas outside the QTP and include species that live at lower elevations.

Where extant pikas originated and how they formed the current distribution pattern are still under debate. The known fossil records suggest that extant pikas likely originated at lower elevations in Central Asia (Dawson 1967), subsequently dispersed to the QTP, and began species differentiation along with the gradual uplift of the QTP (Yu et al. 1992, 2000). However, some researchers hold a different point of view, arguing that the QTP was the origin center of extant pikas (Ge et al. 2012). The main basis of the latter view is that subgenus Ochotona, which is mainly distributed on the QTP, is the basal lineage in the pika phylogeny. However, the phylogenetic relationships among the five subgenera of extant pikas remain elusive. The phylogenies of Ochotona inferred from a few mitochondrial genes (Yu et al. 2000; Niu et al. 2004; Lanier and Olson 2009; Ge et al. 2013; Lissovsky 2014; Liu et al. 2017) or from multilocus data sets (12 nuclear genes: ~7,500 nt, Melo-Ferreira et al. 2015; 2 mitochondrial genes and 5 nuclear genes: ~7,000 nt, Koju et al. 2016)
are generally poorly resolved. This uncertainty about relationships has impeded the reconstruction of a robust time-calibrated phylogeny for extant pikas, which is necessary for understanding the macroevolutionary processes that underlie their diversity.

In this study, we used the whole-exome sequencing technique to collect molecular data of pikas. This approach allowed us to simultaneously analyze thousands of nuclear coding genes and mitochondrial genomes to study the evolutionary history of extant pikas. Our taxon sampling efforts included 105 pika specimens covering 30 Ochotona species (nominal or tentatively named) and all 5 subgenera, enlarging the taxon sampling for the poorly studied subgenus Alienauroa in particular. To our knowledge, this is the largest and most comprehensive data set ever used in the evolutionary study of pika. In addition, we also performed a combined analysis by incorporating previously published sequences so that pika species that were not sampled by us can be included. Our goal is to propose a robust hypothesis of phylogenetic relationships and divergence times of the major lineages and to provide genomic perspectives on the historical biogeography and the adaptive evolutionary history of extant pikas. These results will help us to better understand the role of the QTP in shaping the biodiversity of the extant pikas.

Results and Discussion
Taxon Sampling and Data Sets
The extant pikas taxonomy, informed by morphology and mitochondrial cytochrome b sequences, recognizes five subgenera: Alienauroa, Conothoa, Lagotona, Ochotona, and Pika (Lissovsky 2014; Liu et al. 2017). In this study, we collected 105 pika samples covering all 5 subgenera (fig. 1A and supplementary table S1, Supplementary Material online). For Alienauroa, our sampling included 17 samples (5 species) covering 100% of the species of this subgenus. For Conothoa, our sampling included 22 samples (11 species), covering 66% of the species of this subgenus. For Ochotona, we collected 57 samples (10 species), covering 100% of the species of this subgenus. For Pika, we collected seven samples (three species), covering 33% of the species of this subgenus. Subgenus Lagotona includes only one species (O. pusilla), which is largely distributed in Central Asia and Russia. We collected one sample of O. pusilla from Xingjiang Province, China, and confirmed its identity by mitochondrial gene sequencing and morphology. In addition, for widespread species such as O. thibetana and O. cansus, our sampling purposely included more samples collected from different elevations and distributions, which may help to reveal potential cryptic species.

We sequenced the whole exomes from the 105 pika samples by sequence capture. A total of 200 GB of sequence data were obtained, with an average of 1.9 GB data for each sample. Approximately 8.5% of the sequence data could be mapped to the 20,147 reference coding genes (see Materials and Methods). The individual-level data set contained 2,552 protein-coding genes from 106 pikas (plus the American pika Ochotona princeps) and 6 outgroups (3,686,673 nt in length and 46% complete). The species-level data set obtained by merging the data of all individuals from each nominal species included 6,607 protein-coding genes from 31 pika species and 6 outgroups (11,863,752 nt in length and 64% complete). To increase species coverage for subgenera, we downloaded 9 nuclear protein-coding genes of 34 additional pika samples from GenBank to supplement the species of Conothoa and Pika that were not sampled. Based on these public data and our data, we constructed an individual-level combined data set (146 samples, 2,561 genes, 3,793,698 nt, 36% complete) and a species-level combined data set (44 species, 393 genes, 755,901 nt, 70% complete). Finally, in addition to nuclear protein-coding genes, we extracted mitochondrial sequences from our NGS data and constructed a mitochondrial genome data set of 13,884 bp including 87 pikas and 2 outgroups.

Resolving the Phylogeny of Extant Pikas
Concatenated maximum likelihood (ML) analysis (IQ-TREE) of the individual-level data set yielded a highly resolved phylogeny: 84% of the nodes within pikas exhibited ultrafast bootstrap (UFBS) values ≥95% (fig. 1B). ML analysis of the individual-level combined data set (with more species added) produced a completely congruent phylogeny (supplementary fig. S1, Supplementary Material online), although with weaker support (70% of nodes have UFBS ≥ 95%), suggesting that the missing species had little influence on our phylogenetic inference. ML (IQ-TREE) and gene tree-based coalescent (ASTRAL) analyses of the species-level data set produced a more resolved phylogeny (92% of nodes with UFBS = 100% and ASTRAL bootstrap = 100%) that was identical to the backbone tree of pikas inferred from the individual-level data set (fig. 2A). This species-level phylogeny remained stable after adding more missing species (the species-level combined data set; fig. 2B and supplementary fig. S2, Supplementary Material online). In addition, we performed a series of data subsampling treatments on the species-level data set to explore the potential impact on phylogenetic inference from GC content, missing data, gene signals, and gene evolution rates. Nearly, all data subsets produced an identical phylogeny to the original data set, and the relationships among subgenera remained stable (supplementary fig. S3, Supplementary Material online). All these congruent results indicated that our obtained phylogeny was highly robust regardless of the data set and tree-building method used.

Our phylogenetic trees (figs. 1B and 3 and supplementary fig. S1, Supplementary Material online) largely supported the current species classification of pikas (Lissovsky 2014; Liu et al. 2017; Smith et al. 2018). Twenty-eight species can be affirmed, including Ochotona alpine, O. cansus, O. chinensis (previously O. roylii chinensis), O. collaris, O. coreana, O. curzoniae, O. dauurica, O. erythrotis, O. fatcalvariam, O. forresti, O. gloveri, O. hyperborea, O. himalayana, O. hsianglensis, O. macrotis, O. rubrica, O. princeps, O. pallasii, O. pusilla, O. qianglaensis, O. rufescens, O. roylii, O. rutila, O. sikimaria, O. thibetana, O. thomasi, and O. xunhuaensis (=O. syrinx). However, we also found that the genetic boundaries of some pika species were not clear, such as those between O. morosa (recently elevated as species level from O. cansus...
morosa; Lissovsky et al. 2019) and O. canus, between O. dabashanensis and O. xunhuaensis, between O. gloveri and O. muliensis, between O. gaoligongensis and O. nigritia, and between O. hoffmanni and O. matchurica. Their species validity needed further studies. Five species were not included in this study (O. argentata, O. ladacensis, O. iliensis, O. koslowi, and O. opaca) and their validities were not investigated here. Although we did not find cryptic species among the two

**Fig. 1.** The distribution and phylogeny of 105 pika specimens used in this study. (A) The collection localities of all pika specimens. (B) ML analysis result of the individual-level data set (112 taxa, 2,552 CDS, 3,686,673 bp). The tree was inferred by concatenation ML analysis (IQ-TREE) using a partitioned scheme (three codon position partitioning). Values beside nodes are IQ-TREE ultrafast bootstrap support.
widespread species, *O. thibetana* and *O. kansus*, we discovered a distinct lineage (sample XZ13027) that may represent a new pika species (fig. 1 and supplementary fig. S1, Supplementary Material online). A reliable taxonomic system for extant pikas is still on its way, and denser taxon sampling of individuals will no doubt be necessary. Our work provides valuable genetic data for further pika classification studies.

The placement of the monotypic subgenus *Lagotona* (whose only species is *Ochotona pusilla*) is one of the most compelling questions in pika phylogenetics. Traditionally, the steppe pika (*O. pusilla*) has been considered to be the most primitive lineage of extant pikas, based on its dental and mandibular morphology, karyotype, and relatively old fossil history (Erbajeva 1994; Fostowicz-Frelik et al. 2010). Most molecular phylogenetic studies of Ochotonidae based on mtDNA alone (Niu et al. 2004; Dahal et al. 2017; Liu et al. 2017) have also shown that the steppe pika is the sister group of all other extant pikas. However, based on similar mtDNA data, some authors (Ge et al. 2012, 2013) have argued that *O. pusilla* is nested within subgenus *Ochotona*. Recently, by analyzing a small number of nuclear genes, Melo-Ferreira et al. (2015) found that *Lagotona* was actually sister to *Pika*. The incongruence of the phylogenetic placement of *O. pusilla* among different studies has long caused controversies regarding the origin of extant pikas among evolutionary biologists. By using only mitochondrial data (the mitochondrial genome data set), our ML analysis showed that *O. pusilla* is a rather distinct lineage and is the sister group of all other extant pikas (fig. 3), which is consistent with the results of most previous studies based on mitochondrial genes. However, the nodes at the base of the mtDNA tree lacked strong support (fig. 3).
bootstrap 56–77%), suggesting that using mitochondrial data alone cannot reliably resolve the phylogeny of extant pikas. On the other hand, all of our phylogenetic analyses based on genome-scale nuclear data strongly support *O. pusilla* as the sister group of subgenus *Pika* (figs. 1 and 2 and supplementary figs. S1 and S2, Supplementary Material online). According to the phylogeny inferred from genome-scale nuclear data (fig. 2), the genetic distance between *O. pusilla* and other species of *Pika* is not far enough to establish *Lagotona* as a subgenus, as suggested by Erbajeva (1988). We suggest that *O. pusilla* should be moved into subgenus *Pika*. Our results showed that maternally inherited mtDNA presented a different genealogical history compared with the biparentally inherited nuclear genome during the early speciation of pikas, which was the cause of the incongruence in the placement of *O. pusilla* among previous studies. This finding highlights the necessity of using large nuclear data sets to investigate the phylogeny and taxonomy of pikas.

The trees that we constructed clearly recognized four distinct lineages of extant pikas that correspond to the previously proposed subgenera *Conothoa*, *Alienauroa*, *Ochotona*, and *Pika* (Lissovsky 2014; Liu et al. 2017) (fig. 2A). Overall, we support the ecotype division of extant pikas proposed by Yu et al. (2000) but with some additions. From the aspects of distribution and habitat, these four lineages represent four ecotypes. *Conothoa* species are mainly distributed around the Himalayan Mountains and in their vicinity (fig. 2C) and are mostly rock-talus dwellers (mountain group). *Alienauroa* species are distributed on the eastern edge of the QTP (fig. 2C) and are commonly found in humid forest areas with mosses and leaf litter (forest group). *Ochotona* species are mainly distributed in northern and northeastern areas of the QTP (fig. 2C) and are mostly shrub-steppe dwellers (shrub-steppe...
group). *Pika* species are distributed at high latitudes and have expanded into the New World (fig. 2C), comprising both rock dwellers and steppe dwellers (northern group). According to these results, the extant pikas comprise only four valid subgenera: *Conothoa*, *Alienauroa*, *Ochotona*, and *Pika*.

Except for the oddly highly supported phylogeny of *Ge et al.* (2012) based on only two mitochondrial genes, the relationships among the subgenera of pikas have not been reliably inferred by most previous studies (*Melo-Ferreira et al.* 2015; *Dahal et al.* 2017; *Koju et al.* 2017). Our study reliably resolved the subgenus-level relationships of pikas for the first time, with all nodes at the base of the tree receiving maximal statistical support (fig. 2A). The interrelationships between the four subgenera were strongly recovered as (*Conothoa, (Alienauroa, (Ochotona, Pika)*)], with the mountain group *Conothoa* being the sister group of all other pikas. Traditionally, extant pikas have been considered to have originated at lower elevations and to have subsequently dispersed to high-elevation areas (*Yu et al.* 2000). Our pika phylogeny does not support this view. Both *Conothoa* and *Alienauroa* do not comprise low-elevation species, and these subgenera are the first and second branching lineages in the pika tree, respectively. Low-elevation species (such as *O. daurica* and *O. manchurica*) are only found in the younger clades, such as *Ochotona* and *Pika*, suggesting that the extant pikas may have originated in mountain areas with high elevations rather than low-elevation areas.

In summary, our analyses corroborate many of the deeper or shallower nodes in the pika phylogenies inferred by other studies but with greater support. The placement of *O. pusilla* among pikas, for which previous studies have reported conflicting results, is now robustly resolved. The validity of subgenus *Lagotona* is rejected in this study. Our important findings include the position of the mountain group *Conothoa* as the earliest branching lineage of extant pikas and the different genealogical histories of mtDNA and nuclear DNA during the evolution of pikas.

### Divergence Times and Historical Biogeography of Extant Pikas

Our divergence time analyses were based on both the species-level data set (6,607 genes) and the species-level combined data set (393 genes) using a Bayesian relaxed clock method (MCMCTREE, *Yang* 2007). We did not calibrate any nodes within Ochotonidae because the phylogenetic positions of most fossil *Ochotona* taxa are uncertain (*Erbajeva et al.* 2015). We used three fossil calibration points outside of Ochotonidae to calibrate the timetree: the Rodentia–Lagomorpha divergence, the Leporidae–Ochotonidae split, and the Mus–Rattus divergence. Because the times of the three calibration points were somewhat controversial, we used two different schemes of values (schemes A and B; see Materials and Methods). The molecular age estimates from the two data sets and the two calibration schemes were similar (supplementary figs. S4 and S5, Supplementary Material online), and the average time deviation between comparable nodes among the four analyses was ~5%, suggesting that the time estimations were robust regardless of calibration and data.

For this discussion, we used the time estimates from the species-level data set and calibration scheme A as our primary time result (fig. 4). The origin of extant pikas (genus *Ochotona*) was estimated to be 13.75 Ma, in the middle Miocene, in agreement with the origins reported in previous studies based on a small number of genes (11.6–14.7 Ma; e.g., *Lanier and Olson* 2009; *Ge et al.* 2012; *Koju et al.* 2017). In terms of the fossil record, *Ochotona* first appeared unequivocally in the middle Miocene (~15 Ma; reviewed by *Ge et al.* [2013]), which is close to our time estimate. The divergence among the four subgenera started immediately after the origin time of pikas and occurred within a short period (11.98–13.75 Ma), suggesting that there was rapid diversification in the early evolution of extant pikas. The crown ages of the four extant subgenera *Conothoa, Alienauroa, Pika*, and *Ochotona* estimated in this study were 10.46, 6.81, 9.45, and 7.50 Ma, respectively. *Conothoa* is the oldest clade among the four extant pika subgenera, which is consistent with its basal position in the pika tree.

We performed a biogeographic analysis on the species-level combined data set using BioGeoBEARS (*Matzke* 2013). We used three models, dispersal-extinction cladogenesis, dispersal-vicariance analysis, and Bayesian inference of historical biogeography for discrete areas (BAYAREALIKE), to reconstruct the possible ancestral areas of extant pikas. All models with an additional *J* parameter performed significantly better than the original models (supplementary table S2, Supplementary Material online), indicating the importance of the *J* parameter, which models long-distance or “jump” dispersal. This result suggests that long-distance dispersal has been a common phenomenon during the diversification of pikas.

All three models produced similar biogeographic results (supplementary fig. S6, Supplementary Material online), and we used the result from the BAYAREALIKE + J model as our primary biogeographic result (fig. 4). The most recent common ancestor of extant pikas was distributed on the QTP and in its vicinity (99.4% probability; fig. 4). Remarkably, with the exception of *Pika*, the other three subgenera, *Conothoa, Alienauroa*, and *Ochotona*, were also strongly supported as having originated on the QTP and in its vicinity (fig. 4). It is customary to consider Central Asia (*Dawson* 1967) or northern China (*Fostowicz-Frelik et al.* 2010) as the center of origin of extant pikas; pikas later migrated from low elevations to high elevations on the QTP, and their differentiation in the Palearctic Region was closely related to the gradual uplift of the QTP (*Yu et al.* 2000). However, our biogeographic analyses strongly indicated the QTP and its vicinity as the most likely ancestral distribution areas of extant pikas, which was similar to the suggestions of *Ge et al.* (2012) but based on a more robust phylogeny, much larger data set and different analytical methods. This result suggests that the extant pikas may actually have originate in high-elevation areas of the QTP and later descend to lower elevation areas of Central Asia and northern China.
Out of Tibet: Possible Origin and Dispersal Scenario of Extant Pikas

A noteworthy phenomenon is that the origin of extant pikas (13.75 Ma) occurred immediately after the middle Miocene global cooling event (fig. 5A). The middle Miocene global cooling event is dated to ~13.9 Ma, which marked the Earth’s final transition into an “icehouse” climate (Zachos et al. 2001; Holbourn et al. 2005). This synchronization is unlikely to have occurred by chance because all extant pikas are characteristically adapted to cold environments. The members of the pika family exhibited a wide distribution across Africa, Eurasia, and North America from the late Eocene to the middle Miocene (Smith et al. 2018). During this period, they flourished when the Earth was warming, but their diversification slowed when the Earth was cool (Ge et al. 2012). The generic diversity of Ochotonidae greatly decreased after the middle Miocene when the global average temperature dropped quickly (Ge et al. 2012). Based on these observations, we speculate that most ancient pika lineages preferred warm habitats and became extinct due to the continuing cool-down after the middle Miocene, with only the cold-adapted genus Ochotona surviving to the present.
Notably, the ancestral branch connecting to the extant pikas is rather long: \( \approx 40 \) My (fig. 4B). The lack of other extant pika lineages originating from this stem branch corroborates our hypothesis that the global cooling event decimated these warmth-loving lineages after the middle Miocene.

All known species of extant pikas favor cool or cold environments, and our biogeographic analysis shows that their center of origin lies on the QTP. The cold environment of the QTP may have provided a training ground for cold adaptation in the ancestor of extant pikas well before the start of the global cooling event. This possibility is in agreement with the "Out-of-Tibet" hypothesis (Deng et al. 2011; Wang et al. 2014, 2016). However, Ge et al. (2012) proposed a "warm refuge hypothesis" for pika evolution, which claims that the ecological conditions of the QTP were largely tropical during the transition period from the Miocene to Pliocene and offered a refuge for the ancestor of extant pikas. Geologically, the uplift of the QTP started in its southern area (Himalayan Mountains) and gradually reached its current height (fig. 5B). From the Eocene to the late Oligocene period, the southern part of the QTP was high and cold, but the northern part was still below 2,000 m (Deng and Ding 2015; Deng et al. 2019) and had a warm and wet climate (Wang et al. 2015; Wu et al. 2017). Notably, our pika phylogeny indicates that the most basal clade of pikas is *Conothoa*, which is largely distributed among the Himalayan Mountains but not on the northern QTP, which is not congruent with the "warm refuge hypothesis." Large-scale patterns of species diversity are often driven by abiotic factors such as climate, landscape, or food supply (Benton 2009). The middle Miocene global cooling event provided climatic opportunities for the diversification of cold temperature-preferring pikas (fig. 5A). In terms of the landscape, the uplift of the QTP progresses from south to north (Mulch and Chamberlain 2006; Favre et al. 2015). In the middle Miocene \( \approx 15 \) Ma, the north part of the QTP and its surrounding mountain ranges such as the Tianshan Mountains and the Qilian Mountains, reached their current heights (fig. 5B, Mulch and Chamberlain 2006; Favre et al. 2015). In the middle Miocene the eastern edge of the QTP experienced a significant uplift after 10 Ma, forming the Hengduanshan Mountains east to the QTP (Mulch and Chamberlain 2006; Sun et al. 2011). These high mountains probably provided stepping
stones out of the QTP for pikas. The timing of the divergence among pika subgenera largely coincided with the timing of these orogenies of high mountain ranges, which is consistent with this hypothesis. In addition, to determine the dynamics of the food supply of pikas during this period, based on the evolutionary times estimated for angiosperms in China by Lu et al. (2018), we extracted the origin times of 126 angiosperm genera favored by extant pikas (supplementary table S3, Supplementary Material online). We found that the number of food plant genera consumed by pikas increased quickly after the global cooling event (fig. 5C); among these 126 plant genera, 90 (71%) occurred after the origin of extant pikas. The flourishing of forage plants provided substantial food resources and promoted the proliferation of pikas. Therefore, we argue that the opportunities provided by the climate, landscape, and food supply may together have provoked rapid diversification and outward dispersal in the early evolution of extant pikas.

To detect the gene flow and the direction of introgression among the four pika subgenera, we calculated $D_{FOL}$ statistics (Pease and Hahn 2015) for every quartet of pikas species using rabbits as outgroup. We performed three $D_{FOL}$ analyses based on three symmetric five-taxon phylogenies (fig. 6). For every five-taxon phylogeny, we performed $D_{FOL}$ test for all possible combinations of five species belonging to each of the five taxa (supplementary tables S4–S6, Supplementary Material online). In all the three analyses, gene flow can be detected in ~12–17% of species permutations, showing gene flow among pika subgenera exists but is not strong. In the analysis 1, among the detected 174 gene flow events (fig. 6), 103 (59%) are ancestral introgressions ($P_{12} \leq \leftarrow \rightarrow P_{3}$ and $P_{12} \leq \rightarrow \rightarrow P_{4}$) without distinguished direction. Remarkably, among the remaining 71 directional introgressions, most of the gene flow events (73.2%) are from the subgenus Conothoa to the subgenus Alienauroa and Pika ($P_{1} \rightarrow \rightarrow P_{2}$, $P_{1} \rightarrow \rightarrow P_{2} \rightarrow \rightarrow P_{9}$, and $P_{2} \rightarrow \rightarrow P_{4}$), instead of the opposite direction (fig. 6). In the other two $D_{FOL}$ analyses, gene flow from Conothoa to other subgenera are also more frequent than the opposite directions (fig. 6). Because Conothoa originated on the QTP and the other three subgenera are more or less distributed outside the QTP, our gene flow analyses imply that the main dispersal direction of early pikas is spreading outward from the QTP.

According to these results, we argue that the last common ancestor of extant pikas was likely restricted in the southern part of the QTP (Himalayan Mountains) and characterized by cold adaption (fig. 5D). The continuing cool-down of the climate after the middle Miocene resulted in a decline of other ancient warmth-prefering pika lineages, providing ecological niches for cold-adapted pikas. The occurrence of food plants and the connection of the QTP with other adjacent high mountain ranges further provoked the outward dispersal of pikas from the QTP. Different dispersal directions eventually resulted in divergence among the four subgenera of extant pikas. During their dispersal, the mountain taxon Conothoa mainly spread along the Himalayan Mountains and entered Central Asia through Hindukush; the northern taxon Pika spread northward through the Tian Shan Mountains, entered the Eurasian steppe and further dispersed into North America via the Bering Land Bridge; the forest taxon Alienauroa crossed the Hengduan Mountains eastward and reached central China; and the shrub-steppe group Ochotona fanned out on the QTP and reached Northeast Asia (fig. 5D).

**Cold Tolerance Was the Major Adaptive Direction in the Early Evolution of Extant Pikas**

Environmental stress has been a major driving force in the evolution of living organisms (Parsons 2005). In general, there are three major environmental stresses faced by animals that live in high elevations: hypoxia, strong UV radiation, and low temperatures. Previous studies of adaptive evolution in high-elevation animals such as snub-nosed monkeys (Yu et al. 2016), Tibetan frogs, and sand lizards (Sun et al. 2018) have often paid more attention to adaptation to hypoxia and UV radiation. In this study, we argue that the ancestor of extant pikas occurred in a high-altitude area of the Himalayas; the cold winters in high mountains might have provided a training ground for cold tolerance in early pikas, which gives them a great advantage when the Miocene global cooling event finally occurred. Similar scenarios have been proposed for other mammalian taxa such as woolly rhinos and arctic foxes, but these scenarios were exclusively based on morphological characters from fossil records (Deng et al. 2011; Wang et al. 2014). The case of pikas addressed herein provides a living example that can be used to test this hypothesis at the genetic level. We wonder whether the selection signals in the genome of the ancestor of extant pikas would be more concentrated in functional categories related to cold adaptation versus the other two adaptation directions.

Prior to this study, several other studies addressed the evolutionary genetics of high-elevation adaptation in pikas, but focused only on a few predefined genes (Henry and Russello 2013; Rankin et al. 2017). To perform a genome-wide investigation, we constructed a high-confidence orthologous gene set (13,573 genes) including at least 2 pika species and using rabbit, rat, mouse, hamster, guinea pig, human, gorilla, macaca, cow, goat, and pig as outgroups (fig. 7A). Among the 13,573 genes, 1,681 were related to hypoxia and positive (PSGs) and positively selected genes (PSGs) specific to the ancestral branch of extant pikas. Both REGs and PSGs are genes that have experienced special selection pressures caused by biotic or abiotic factors, and the genes that overlap between them are expected to be more reflective of major factors selected in the evolution of animals. The detailed evolutionary analysis results are provided in supplementary tables S7 and S8, Supplementary Material online.

We identified a total of 249 REGs and 673 PSGs in the ancestral branch of pikas with a false discovery rate $\leq 0.05$ (supplementary tables S9 and S10, Supplementary Material online). Among the 249 REGs, 45 were related to hypoxia, showing a significant enrichment pattern compared with the background (13,573 genes, 1,681 related to hypoxia) (hypergeometric test $P = 0.0056$). However, among the 673 PSGs, only 57 were related to hypoxia, which did not show a
significant enrichment pattern compared with the background (hypergeometric test $P = 0.9996$). There were 86 PSGs that overlapped with REGs (fig. 7B; supplementary table S11, Supplementary Material online). Importantly, among the 86 PSGs overlapping with REGs, 15 were related to hypoxia, which did not show a significant enrichment pattern compared with the background (249 REGs, 45 related to hypoxia) (hypergeometric test $P = 0.6370$). These results suggest that although hypoxia adaptation was apparent in the early evolution of extant pikas, it was not the major adaptation direction.

Functional enrichment analyses of the REGs and PSGs relative to the whole gene set identified a total of 533 and 490 significantly ($P < 0.05$) enriched Gene Ontology (GO) categories, respectively (supplementary tables S12 and S13, Supplementary Material online). These enriched GO terms were widely distributed among different biological processes such as immune, reproduction, metabolism, and biosynthesis processes and were not concentrated on specific biological processes, with only a small number of functional categories related to strong UV radiation (fig. 7C and D). These findings emphasize that the adaptation direction of early pikas lived on the QTP was multidimensional and that strong UV radiation imposed relatively low selective pressure on early pikas.

Functional enrichment analyses of the 86 overlapping PSGs relative to the REGs identified a total of 20 significantly ($P < 0.05$) enriched GO terms. Notably, these enriched GO terms were mainly related to three biological functions: morphological characters (six GO terms), metabolism (five GO terms), and biosynthetic functions (six GO terms) (fig. 7D). Among the 20 GO terms, 10 (50%) were related to the biosynthesis and metabolism of lipids and steroids, including terms such as lipid biosynthetic process (GO:0008610), steroid biosynthetic process (GO:0006694), and steroid metabolic process (GO:0008610). The significant enrichment of PSGs for functions related to the biosynthesis and metabolism of lipids and steroids in early pikas likely has biological implications. Biologically, pikas have evolved an unusually high resting metabolic rate and robust nonshivering thermogenesis (Li et al. 2001; Sheafor 2003) to ward off hypothermia at low temperatures, which is highly related to the synthesis and metabolism of fat. On the other hand, steroids can help maintain the fluidity of the cell membrane under cold conditions, which is also helpful for allowing pikas to withstand cold. Recently, Lynch et al. (2015)
**Fig. 7.** Evolutionary analyses of functions associated with adaptation to cold in extant pikas. (A) The reference phylogeny used in PAML (ancestral branch of pikas in purple). (B) There are 249 REGs and 673 PSGs identified in the ancestral branch of pikas, with 86 overlapping genes. Manhattan plot of GO terms enriched from (C) the 249 REGs relative to the background 13,573 genes, (D) the 673 PSGs relative to the background 13,573 genes, and (E) the 86 overlapping PSGs relative to the background 279 REGs. The $-\log_{10}$ (hypergeometric $P$ values) values are shown for each GO term; GO terms are grouped by related biological functions. Horizontal red line, $P = 0.05$. (F) The 20 significantly enriched GO terms from plot (E) and the corresponding detailed enrichment results. The ten GO terms related to the biosynthesis and metabolism of lipids, and steroids are highlighted in yellow.
suggested that PSGs associated with insulin signaling and lipid and steroid metabolism may have played an important role in the evolution of woolly mammoths and adaptation to arctic life. Interestingly, insulin-induced genes 1 and 2 (Insig-1 and Insig-2), which regulate cholesterol metabolism and lipogenesis, are repeatedly present in our functional enrichment analysis (fig. 7E), suggesting that pikas may have followed a similar evolutionary path to woolly mammoths in the adaptation to cold. These findings show that the major adaptive direction in the early history of extant pikas is cold tolerance. The cold climate of high-altitude areas of the QTP subjected the common ancestor of extant pikas to valuable environmental stress, resulting in their preadaptation to the cold before the Miocene global cooling and allowing them to successfully expand outward from the QTP with climatic change toward cooler conditions at the end of the middle Miocene.

Conclusion

Pikas are typically cold-adapted mammals. When and where the extant pikas arose and how they adapted to cold remains compelling questions. Our study resolved the phylogeny of extant pikas (Ochotona) via extensive sampling of both species and genes. We identified four strongly supported clades of pikas originating on the QTP \( \sim 14 \) Ma and then dispersed out of the QTP to other parts of Eurasia and North America. Genome-wide evolutionary analyses indicated that extant pikas originated on the QTP \( \sim 14 \) Ma and then dispersed out of the QTP to other parts of Eurasia and North America. Genome-wide evolutionary analyses suggested that the major adaptive direction of early pikas on the QTP was toward cold adaptation. Thus, our results support the “Out-of-Tibet” hypothesis indicating that many widespread, cold-adapted animal groups of the Northern Hemisphere originated on the QTP and preadapted to its harsh winter before migrating to other Northern Hemisphere regions. Finally, the genomic data that we have generated will be a useful resource for future studies on the biology, classification, and adaptive evolution of pikas.

Materials and Methods

Sample and Data Collection

In this study, we collected a total of 105 pika samples. Detailed information for these samples, such as the collection locality, latitude, longitude, and altitude, is provided in figure 1A and supplementary table S1, Supplementary Material online. For each sample, total genomic DNA was extracted from ethanol-preserved tissues (liver or muscle) using the TIANamp Genomic DNA kit (TIANGEN Inc, Beijing, China). A 200 ng sample of genomic DNA was randomly fragmented with NEBNext dsDNA fragmentase (New England Biolabs) to 200–400 bp and used for DNA library construction with the NEBNext Illumina DNA Library Prep Kit (New England Biolabs). We used the SureSelectXT Mouse All Exon Kit (Agilent Technologies) to enrich all coding sequences (CDSs) of each pika sample. The hybridization enrichment experiments were performed using the SureSelectXT Reagent Kit (G9611A) according to the manufacturer’s protocol. After enrichment, hybridized libraries were captured with Dynabeads MyOne Streptavidin C1 magnetic beads (Invitrogen) and amplified with index primers. The indexed captured libraries were pooled in equal quantities and sequenced on an Illumine HiSeqX10 sequencer. Finally, \( \sim 200 \) GB of Illumina 150-bp paired-end reads were obtained.

Data Processing and Data Set Construction

We first combined the CDSs of the American pika (O. princeps, OchPri2.0-Ens) and the rabbit (Oryctolagus cuniculus, OrCyNam2.0-GCA_000003625.1) as the reference sequences for data mapping. Briefly, the gene orthology between the CDSs of O. princeps and Or. cuniculus was determined via the mutual best-hit BLAST (expectation value \( \sim 1 \times 10^{-10} \)) approach. Three criteria were used to choose reference CDSs: 1) the missing data of a reference CDS should be \( < 1 \% \), 2) the length of a reference CDS should be more than 300 bp, and 3) the sequence of O. princeps has priority for use as reference. As a result, a total of 20,147 CDS (7,261 from O. princeps and 13,787 from Or. cuniculus) were obtained as reference sequences for data mapping.

For each pika sample, the sequence reads were aligned to the reference CDS using BWA (Li et al. 2010) under a stringent setting of -B 6. The alignments were converted to BAM files with SAMtools (Li et al. 2009), and consensus sequences were generated with BCFtools (Li 2011). In addition, seven published genomes or transcriptomes (O. princeps, Or. cuniculus, Lepus timidus, Chinchilla lanigera, Peromyscus maniculatus, Mus musculus, and Rattus norvegicus) were downloaded from GenBank (supplementary table S14, Supplementary Material online). We used a “pseudo read mapping” strategy to identify orthologous CDSs from these data. Briefly, the raw genome and transcriptome sequences were artificially fragmented into small “reads” of 150 bp at a 10 \( \times \) tiling density. These “pseudo” reads were mapped to the reference CDSs using BWA (Li et al. 2010) under a default setting of -B 4, and orthologous CDSs were generated as mentioned above. Finally, we obtained 20,147 orthologous CDSs for 112 samples (105 pikas and 7 outgroups). For each reference CDS, the obtained orthologous CDSs from different samples exhibited exactly the same length, and the coordinates of these CDSs exactly matched those of the reference sequence.

Based on the 20,147 orthologous CDSs of 112 samples, we constructed an individual-level data set using the following filtering criteria: 1) the percentage of ambiguous bases per CDS must be \( < 0.1 \% \), 2) the percentage of missing data per CDSs must be \( < 60 \% \), and 3) at least 90% of samples must be included per CDS. To reduce missing data and include more genes, we constructed a species-level data set by merging orthologous CDS data of all individuals from every nominal species. The nominal species and the individuals of those species are shown in figure 1A. After merging the data, we used the same criteria to filter the species-level orthologous CDSs. In addition, to increase species coverage, we downloaded the sequences of nine genes (DARC, TSHB, OXAl1, RAG1, RAG2, TTN, EHHADH, IGFI, and NOS2) from 34 pika samples in GenBank, covering 7 pika species that were not sampled in this study (supplementary table S14, supplementary table S15, Supplementary Material online). We used a “pseudo read mapping” strategy to identify orthologous CDSs from these data. Briefly, the raw genome and transcriptome sequences were artificially fragmented into small “reads” of 150 bp at a 10 \( \times \) tiling density. These “pseudo” reads were mapped to the reference CDSs using BWA (Li et al. 2010) under a default setting of -B 4, and orthologous CDSs were generated as mentioned above. Finally, we obtained 20,147 orthologous CDSs for 112 samples (105 pikas and 7 outgroups). For each reference CDS, the obtained orthologous CDSs from different samples exhibited exactly the same length, and the coordinates of these CDSs exactly matched those of the reference sequence.

Based on the 20,147 orthologous CDSs of 112 samples, we constructed an individual-level data set using the following filtering criteria: 1) the percentage of ambiguous bases per CDS must be \( < 0.1 \% \), 2) the percentage of missing data per CDSs must be \( < 60 \% \), and 3) at least 90% of samples must be included per CDS. To reduce missing data and include more genes, we constructed a species-level data set by merging orthologous CDS data of all individuals from every nominal species. The nominal species and the individuals of those species are shown in figure 1A. After merging the data, we used the same criteria to filter the species-level orthologous CDSs. In addition, to increase species coverage, we downloaded the sequences of nine genes (DARC, TSHB, OXAl1, RAG1, RAG2, TTN, EHHADH, IGFI, and NOS2) from 34 pika samples in GenBank, covering 7 pika species that were not sampled in this study (supplementary table S14, supplementary table S15, Supplementary Material online).
Supplementary Material online). These sequences were appended to the original 20,147 orthologous CDSs using the “pseudo read mapping” strategy. Based on this data source, we constructed an individual-level combined data set and a species-level combined data set using similar procedure to those indicated above. The filtering criteria for the former were the same as those used for the individual-level data set. For the species-level combined data set, more stringent criteria were used: 1) the percentage of ambiguous bases per CDS must be <0.1%, 2) the percentage of missing data per CDS must be <20%, and 3) 100% of samples should be included per CDS.

Finally, we constructed a mitochondrial genome data set from our sequencing data. For each pika sample, the reads were mapped to the *O. princeps* mitochondrial genome (NCBI: AJ537415) using BWA (Li et al. 2010) under a soft setting (-B 2). The mapped reads were assembled into an annotated mitochondrial genome (partial or complete) by using Mitoz (Meng et al. 2019). Only 83 mitochondrial genomes were complete and were retained for further analysis. In addition, six published mitochondrial genomes (from four pikas, *O. princeps, O. collaris, O. erythrotis*, and *O. curzoniae*, and two rabbits, *O. cuniculus* and *Lepus timidus*) were downloaded from GenBank and included in the mitochondrial analysis (supplementary table S14, Supplementary Material online). MUSCLE v.3.8 (Edgar 2004) was used to align mitochondrial sequences, and GBlock v.0.91b (Castresana 2000) was used to refine the alignment. A bioinformatic diagram summarizing the data processing and data set construction procedures is provided in supplementary figure S7, Supplementary Material online.

Phylogenetic Analyses

Phylogenetic trees were reconstructed with ML inference using IQ-TREE (Nguyen et al. 2015). PartitionFinder (Lanfear et al. 2012) was used to select the best-fit models and partitioning schemes according to the Bayesian information criterion. For the individual-level data set, species-level data set, individual-level combined data set, and species-level combined data set, the best partitioning scheme involved three separate partitions with separate GTR + I + G models for the three codon positions; for the mitochondrial genome data set, the best partitioning scheme involved six separate partitions with separate GTR + I + G models for all tRNAs combined, 12S rRNA, 16S rRNA, and three codon positions for all protein-coding genes. Branch support was estimated by using the ultrafast bootstrap algorithm (UFBoot) embedded in IQ-TREE with 10,000 replicates.

For the species-level data set, we also used coalescent-based species tree inference (ASTRAL; Mirarab et al. 2014) to estimate the species tree. We used a data binning strategy to improve multispecies coalescent analyses for handling gene trees with low phylogenetic signals (Jarvis et al. 2014). The 6,607 genes of the species-level data set were divided into 200 bins of similar sizes according to their evolutionary rates. For each data bin, 200 ML bootstrap trees and the final best-scoring ML tree were estimated using RAxML (Stamatakis 2014) under the GTR + GAMMA model. These best-scoring ML trees and bootstrapping trees were used as input files for ASTRAL with the option: “-i –b” to calculate the final species tree and branch support.

Data subsampling is an effective way to verify the robustness of phylogenetic inferences (Chen et al. 2017). Therefore, we generated a series of data subsets from the species-level data set according to the missing data, GC content, evolutionary rate, and phylogenetic signal of each gene. A total of 12 data subsets were generated: completeness data sets 1–3 required the percentages missing data per gene to be <24%, 31.4%, or 39.2%, respectively; GC data sets 1–3 required the average GC content per gene to be <48.1%, 54.4%, or 59.4%, respectively; rate data sets 1–3 required the average pairwise identity per gene to be <93.93%, 95%, or 95.79%, respectively; and resolution data sets 1–3 required the average bootstrap support of the ML tree per gene to be >55.5%, 46%, and 37.2%, respectively. ML trees were estimated for each data subset using both RAxML (Stamatakis 2014) and IQ-TREE (Nguyen et al. 2015) to evaluate the robustness of our phylogenetic inferences.

Divergence Time Analyses

Divergence time analyses were performed for both the species-level data set and the species-level combined data set using the MCMCTREE program in the PAML packages (Yang 2007). Two calibration schemes were used to calibrate the clock. In scheme A, according to Meredith et al. (2011), the root was set to 79.5 Ma. The Rodentia–Lagomorpha split, the Leporidae–Ochotonidae split, and the Mus–Rattus split were constrained at 71.5–94.1, 47.4–56.9, and 11–13 Ma, respectively. In scheme B, according to Bininda-Emonds et al. (2007), the root was set to 88.9 Ma. The Rodentia–Lagomorpha split, the Leporidae–Ochotonidae split, and the Mus–Rattus split were constrained at 87.8–901.6, 60–68.7, and 11–13 Ma, respectively. The calibration constraints were specified with soft boundaries by using 2.5% tail probabilities above and below their limits; this is a built-in function of MCMCTREE.

For each data set, the included genes were divided into 20 bins according to evolutionary rates. The ML estimates of branch lengths for each of the 20 bins were obtained by using the BASEML (in PAML) program under the GTR + G model. Based on the mean estimate from the 20 bins, the overall substitution rates (rge, gamma) were set at G (1, 9.4624) and G (1, 10.5831) for the two data sets. The prior for the rate-drift parameter (sigma2 gamma) was set at G (1, 4.5, 1). The independent rate model (clock = 2 in MCMCTREE) was used to specify the rate priors for internal nodes. The MCMC run was first executed for 10,000,000 generations as burn-in and then sampled every 150 generations until a total of 100,000 samples were collected. Two MCMC runs using random seeds were compared for convergence, and similar results were found.

Biogeographic Analyses

Based on the current distribution pattern of extant pikas, we defined six biogeographic areas: central Asia, Northeast Eurasia, the QTP and its vicinity, North America, Europe,
and central China. Biogeographic analyses were performed by using BioGeoBEARS. The current distribution of each pika species was assigned to each biogeographic area according to Smith et al. (2018). We used the chronogram generated from the species-level combined data set as the input timetree. Three models, dispersal-extinction cladogenesis, dispersal-vicariance analysis, and BAYAREALIKE, were used in the analyses. These models were compared with their variants (with an additional J parameter) to determine the influence of founder-event dispersal on biogeographic patterns.

Gene Flow Analyses
We used $D_{FOIL}$ statistics (Pease and Hahn 2015) to estimate the direction of introgression among the four pika subgenera. The $D_{FOIL}$ test requires a symmetric five-taxon phylogeny $((P_1, P_2), (P_3, P_4), O)$ in which the most recent common ancestor of $P_1$ and $P_2$ is younger than that of $P_3$ and $P_4$. According to the pikas time tree we obtained, we analyzed 3 five-taxon phylogenies as shown in figure 6. For every five-taxon phylogeny, we applied $D_{FOIL}$ test exhaustively for all combinations of five species belonging to each of the five taxa (supplementary tables S4–S6, Supplementary Material online). To do this, we wrote a custom Python script to extract subsets from the species-level data set. These sequences subsets were used by $D_{FOIL}$ to generate an AB-pattern count file and then calculate $D_{FOIL}$ statistics by using the default significance cutoff of 0.01.

Evolutionary Analyses
We downloaded the coding DNA sequences (CDSs) of 11 mammalian species (used as outgroups) from the following taxa from the ENSEMBL genome database (www.ensembl.org): Lagomorpha (Or. cuniculus), Rodentia (Cavia porcellus, Mesocricetus auratus, Mus musculus, and Rattus norvegicus), Primates (Gorilla gorilla, Homo sapiens, and Macaca mulatta), and Artiodactyla (Bos taurus, Ovis aries, and Sus scrofa). The gene orthology among these CDSs and our pika CDSs was determined via the mutual best-hit BLAST approach. All the orthologous CDSs for each gene were included in a separate FASTA file as an orthologous gene group (OGG). The obtained OGG files were aligned using the ClustalW algorithm implemented in MEGA v6 (Tamura et al. 2013) according to the translated protein sequences. The program PAL2NAL v14 (Suyama et al. 2006) was used to convert the protein sequence alignments into corresponding codon alignments for each OGG. To reduce the impact of missing data, all the OGG codon alignments were filtered to exclude species with more than 20% missing data. Only OGGs with more than 6 outgroups and 2 pika species were retained for further analysis, and a total of 13,573 OGGs were obtained.

Selection analyses were performed using the CODEML program of the PAML package (Yang 2007). For each OGG codon alignment, a separate user tree was generated by pruning the corresponding missing species from the reference tree (fig. 7A) and labeling the ancestral branch of the pika lineage as the foreground branch. To identify REGs in the ancestral branch of pikas, we used the branch model. We compared one-ratio model and two-ratio model for all 13,573 genes. The one-ratio model assumes that all branches in the phylogeny (fig. 7A) have the same $\omega$ value ($dN/dS$), whereas the two-ratio model assumes two different $\omega$ values within the phylogeny: one for the ancestral branch of pikas ($\omega_1$) and one for the remaining branches in the tree ($\omega_0$). If $\omega_1 > \omega_0$ and the two-ratio model is significantly better than the one-ratio model for a gene, the gene is considered to be evolving at a significantly faster rate in the ancestral branch of pikas. To identify PSGs in the ancestral branch of pikas, we used the branch-site model. For each OGG, we compared model A (“model = 2, NSsites = 2, fix_omega = 0, omega = 1.5”; some sites in the foreground branch are under positive selection) with model A null (“model = 2, NSsites = 2, fix_omega = 1, omega = 1”; all sites evolve neutrally) to check whether the gene was under positive selection in early pika. The significance of the compared LRTs was evaluated by the $\chi^2$ test, and multiple testing with the false discovery rate was used to correct the significance in $R$.

GO enrichment analyses were performed by using the R package GOSeq (Young et al. 2010). We applied enrichment analyses using all genes as the background and the REGs and PSGs as the foreground. These two analyses aimed to reveal the functional categories in which the REGs and PSGs were concentrated. In addition, we performed enrichment analysis for the overlapping genes between the REGs and PSGs relative to all REGs. This enrichment further revealed the functional categories in which the PSGs among the REGs are most concentrated. Then, the GO enrichment results were illustrated in a Manhattan plot by using the R package qman (Turner 2018). GO terms with $P$ values lower than 0.05 were considered significantly enriched.

Supplementary Material
Supplementary data are available at Molecular Biology and Evolution online.

Acknowledgments
We thank all of our lab members for help in experiments and data analyses. This work was supported by National Natural Science Foundation of China (Grant Nos. 31872205 and 31970399) and the National Youth Talent Support Program (W02070133). The authors declare they have no competing interests.

References
Benton MJ. 2009. The Red Queen and the Court Jester: species diversity and the role of biotic and abiotic factors through time. Science 323(5915):728–732.
Bininda-Emonds OR, Cardillo M, Jones KE, MacPhee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL, Purvis A. 2007. The delayed rise of present-day mammals. Nature 446(7135):507–512.
Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 17(4):540–552.
Chen MY, Liang D, Zhang P. 2017. Phylogenomic resolution of the phylogeny of Laurasiatherian mammals: exploring phylogenetic signals within coding and noncoding sequences. Genome Biol Evol 9(8):1998–2012.
Dahal N, Lissovsky AA, Lin Z, Solari K, Hadly EA, Zhan X, Ramakrishnan U. 2017. Genetics, morphology and ecology reveal a cryptic pika lineage in the Sikkim Himalaya. Mol Phylogenet Evol. 106:55–60.

Dawson MR. 1967. Lagomorph history and the stratigraphic record. Lawrence (KS): Kansas University.

Delibes-Mateos M, Smith AT, Slobodchikoff CN, Svensson JE. 2011. The paradox of keystone species persecuted as pests: a call for the conservation of abundant small mammals in their native range. Biol Conserv. 144(5):1335–1346.

Deng T, Ding L. 2015. Paleoaltimetry reconstructions of the Tibetan Plateau: progress and contradictions. Natl Sci Rev. 2(4):417–437.

Deng T, Wang XM, Fortelius M, Li Q, Wang Y, Tseng ZJ, Takeuchi GT, Saylor JE, Sakia UK, Xie G. 2011. Out of Tibet: Pliocene woolly rhino suggests high-plateau origin of Ice Age megaherbivores. Science. 333(6047):1285–1288.

Deng T, Wang XM, Wu F, Wang Y, Li Q, Wang S, Hou S. 2019. Implications of vertebrate fossils for paleo-elevations of the Tibetan Plateau. Glob Planet Change. 174:58–69.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32(5):1792–1797.

Erbajeva MA. 1988. Pishchukhi kaiozoia. Nauka Moskva. 224 p.

Erbajeva MA. 1994. Phylogeny and evolution of Ochotonidae with emphasis on Asian ochotons. Natl Sci Mus Monogr. 8:1–13.

Erbajeva MA, Flynn LJ, Alexeeva N. 2015. Late Cenozoic Asian Ochotonidae: taxonomic diversity, chronological distribution and biostatigraphy. Quat Int. 355:18–23.

Favre A, Packert M, Pauls SU, Jahnig SC, Uhl D, Michalak I, Muellner-Riehl E. 2015. Late Cenozoic Asian Ochotonidae: taxonomic diversity, chronological distribution and biostatigraphy. Quat Int. 355:18–23.

Favre A, Packert M, Pauls SU, Jahnig SC, Uhl D, Michalak I, Muellner-Riehl E. 2015. The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of the Tibetan biota. Biol Rev Camb Philos Soc. 90(1):236–253.

Fostowicz-Frelik Ł, Frelik GJ, Gasparik M. 2010. Morphological phylogeny of pikas (Lagomorpha: Ochotonidae), with a description of a new species from the Pliocene/Pliocene transition of Hungary. Proc Acad Nat Sci Phila. 159(1):97–119.

Ge DY, Wen Z, Xia L, Zhang Z, Erbajeva M, Huang C, Yang QS. 2013. Evolutionary history of lagomorphs in response to global environmental change. PLoS One 8(4):e69668.

Ge DY, Zhang Z, Xia L, Zhang Q, Ma Y, Yang QS. 2012. Did the expansion of C4 plants drive extinction and massive range contraction of micromammals? Inferences from food preference and historical biogeography of pikas. Palaeogeogr Palaeoclimatol Palaeoecol. 326:328–160–171.

Henry P, Russell M. 2013. Adaptive divergence along environmental gradients in a climate-change-sensitive mammal. Ecol Evol. 3(11):3906–3917.

Hoffmann RS, Smith AT. 2005. Order Lagomorpha. In: Wilson DE, Reeder DM, editors. In mammal species of the world: a taxonomic and geographic reference. Maryland: Johns Hopkins University Press.

Hobloun A, Kuhnt W, Schulz M, Erlenkeuser H. 2005. Impacts of orbital forcing and atmospheric carbon dioxide on Miocene ice-sheet expansion. Nature 438(7067):483–487.

Jarvis ED, Mirarab S, Reaz R, Bayzid MS, Zimmermann T, Swenson MS, Warnow T. 2016. Lagomorphs: Lagomorpha (Lagomorpha, Ochotonidae). In: Meredith RW, Janecka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E, Simao TLL, Stadler T, et al. 2011. Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. Science. 334(6055):521–524.

Koju NP, He K, Chalise MK, Ray C, Chen Z, Zhang B, Lan X, Chen S, Jiang X. 2017. Multilocus approaches reveal underestimated species diversity and inter-specific gene flow in pikas (Ochotona) from southwestern China. Mol Phylogenet Evol. 107:239–245.

Lanfear R, Calcott B, Ho SY, Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol Biol Evol. 29(6):1695–1701.

Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. Bioinformatics. 27(21):2987–2993.

Liu R, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. Bioinformatics. 25(16):2078–2079.

Li Q, Sun R, Huang C, Wang Z, Liu X, Hou J, Liu J, Cai L, Li N, Zhang S, et al. 2001. Cold adaptive thermogenesis in small mammals from different geographical zones of China. Comp Biochem Physiol Part A Mol Integr Physiol. 129(4):949–961.

Liu SY, Jin W, Liao R, Sun Z, Zeng T, Fu J, Liu Y, Wang X, Li P, Tang M. 2017. Phylogenetic study of Ochotonidae based on mitochondrial cyt b and morphology with a description of one new subgenus and five new species. Acta Theriol. 37(1):1–43.

Liu LM, Mao LF, Yang T, Ye JF, Liu B, Li HL, Sun M, Miller JT, Mathews S, Hu HH, et al. 2018. Evolutionary history of the angiosperm flora of China. Nature 554(7691):234–238.

Lynch VJ, Bedoya-Reina OC, Ratam A, Sulak M, Drautz-Moses DL, Perry GH, Miller W, Schuster SC. 2015. Elephantid genomes reveal the molecular bases of woolly mammoth adaptations to the Arctic. Cell Rep. 12(2):217–228.

Matzke NJ. 2013. BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis in R Scripts. R package, version 02 12013.

Melo-Ferreira J, de Matos AL, Areal H, Lissovsky AA, Carneiro M, Esteves PJ. 2015. The phylogeny of pikas (Ochotonidae) inferred from a multilocus coalescent approach. Mol Phylogenet Evol. 84:240–244.

Meng G, Li Y, Yang C, Liu S. 2019. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. Nucleic Acids Res. 47(11):e63.

Meredith RW, Janecka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E, Simao TLL, Stadler T, et al. 2011. Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. Science. 334(6055):521–524.

Nabholz B, Howard JT, et al. 2014. Whole-genome analyses resolve a cya nan dh i g ht h r o u g h p u t .

Neves CA, Oliver LA, et al. 2013. The phylogeny of pikas (Lagomorpha, Ochotonidae) inferred from mitochondrial cyt b sequences. Folia Zool-PRAHA 53(2):141–156.

Parsons PA. 2005. Environments and evolution: interactions between stress, resource inadequacy and energetic efficiency. Biol Rev. 80(4):589–610.

Pease JB, Hahn MW. 2015. Detection and polarization of introgression in a five-taxa phylogeny. Syst Biol. 64(4):651–662.

Rankin AM, Gablebreath KE, Teeter KC. 2017. Signatures of adaptive molecular evolution in American pikas (Ochotona princeps). J Mammal. 98(4):1156–1167.

Sheafr BA. 2003. Metabolic enzyme activities across an alitudinal gradient: an examination of pikas (genus Ochotona). J Exp Biol. 206(7):1241–1249.

Smith AT, Johnston CH, Alves PC, Hackländer K. 2018. Lagomorphs: pikas, rabbits, and hares of the world. Washington: Johns Hopkins University Press.
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312–1313.
Sun BN, Liu YS, Ding ST, Li XG, Xie SP, Yan DF, Lin ZC. 2011. Reconstructing Neogene vegetation and climates to infer tectonic uplift in western Yunnan, China. *Palaeogeogr Palaeoclimatol Palaeoecol*. 304(3–4):328–336.
Sun YB, Fu TT, Jin JQ, Murphy RW, Hillis DM, Zhang YP, Che J. 2018. Species groups distributed across elevational gradients reveal convergent and continuous genetic adaptation to high elevations. *Proc Natl Acad Sci U S A*. 115(45):E10634–E10641.
Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res*. 34(Web Server):W609–W612.
Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*. 30(12):2725–2729.
Turner SD. 2018. qqman: an R package for visualizing GWAS results using Q–Q and Manhattan plots. *J Open Source Softw*. 3:731.
Wang C, Zhao X, Liu Z, Lippert PC, Graham SA, Coe RS, Yi H, Zhu L, Liu S, Li Y. 2008. Constraints on the early uplift history of the Tibetan Plateau. *Proc Natl Acad Sci U S A*. 105(13):4987–4992.
Wang X, Wang Y, Li Q, Tseng ZJ, Takeuchi GT, Deng T, Xie G, Chang MM, Wang N. 2015. Cenozoic vertebrate evolution and paleoenvironment in Tibetan Plateau: progress and prospects. *Gondwana Res*. 27(4):1335–1354.
Wang XM, Li Q, Takeuchi GT. 2016. Out of Tibet: an early sheep from the Pliocene of Tibet, *Protavis himalayensis*, genus and species nov. (Bovidae, Caprini), and origin of Ice Age mountain sheep. *J Vertebr Paleontol*. 36(5):e1169190.
Wang XM, Tseng ZJ, Li Q, Takeuchi GT, Xie G. 2014. From ‘third pole’ to north pole: a Himalayan origin for the arctic fox. *Proc R Soc B*. 281(1787):20140893.
Wilson MC, Smith AT. 2015. The pika and the watershed: the impact of small mammal poisoning on the ecohydrology of the Qinghai-Tibetan Plateau. *AMBIO*. 44(1):16–22.
Young MD, Wakefield MJ, Smyth GK, Oshlack A. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol*. 11(2):R14.
Yu L, Wang G-D, Ruan J, Chen Y-B, Yang C-P, Cao X, Wu H, Liu Y-H, Du Z-L, Wang X-P, et al. 2016. Genomic analysis of snub-nosed monkeys (*Rhinopithecus*) identifies genes and processes related to high-altitude adaptation. *Nat Genet*. 48(8):947–952.
Yu N, Zheng C, Zhang YP, Li WH. 2000. Molecular systematics of pikas (genus *Ochotona*) inferred from mitochondrial DNA sequences. *Mol Phylogenet Evol*. 16(1):85–95.
Yu N, Zheng CL, Feng ZJ. 1992. The phylogenetic analysis of subgenus *Ochotona* of China. *Acta Theriol Sin*. 12(4):255–266.
Zachos J, Pagani M, Sloan L, Thomas E, Billups K. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*. 292(5517):686–693.