Running head: Introgression in Hares

Title: The Legacy of Recurrent Introgression during the Radiation of Hares

Mafalda S. Ferreira[1,2,3], Matthew R. Jones[3], Colin M. Callahan[3], Liliana Farelo[1], Zelalem Tolesa[4], Franz Suchentrunk[5], Pierre Boursot[6], L. Scott Mills[7,8,9], Paulo C. Alves[1,2,7], Jeffrey M. Good[3,7,*], José Melo-Ferreira[1,2,*]

[1]*CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Universidade do Porto, Vairão, Portugal.*

[2]*Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Porto, Portugal.*

[3]*Division of Biological Sciences, University of Montana, Missoula, Montana, USA.*

[4]*Department of Biology, Hawassa University, Hawassa, Ethiopia.*

[5]*Research Institute of Wildlife Ecology, University of Veterinary Medicine Vienna, Vienna, Austria.*

[6]*Institut des Sciences de l'Évolution Montpellier (ISEM), Université Montpellier, CNRS, IRD, EPHE, France.*

[7]*Wildlife Biology Program, College of Forestry and Conservation, University of Montana, Missoula, Montana, USA.*

[8]*Office of Research and Creative Scholarship, University of Montana, Missoula, Montana, USA.*

[9]*Fisheries, Wildlife, and Conservation Biology Program, Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, North Carolina, USA.*

*Shared senior authorship.

Correspondence: Mafalda S. Ferreira, mafaldasferreira@cibio.up.pt; Jeffrey M. Good, jeffrey.good@umontana.edu; José Melo-Ferreira, jmeloferreira@cibio.up.pt

ABSTRACT

Hybridization may often be an important source of adaptive variation, but the extent and long-term impacts of introgression have seldom been evaluated in the phylogenetic context of a radiation. Hares (*Lepus*) represent a widespread mammalian radiation of 32 extant species characterized by striking ecological adaptations and recurrent admixture. To understand the relevance of introgressive hybridization during the diversification of *Lepus*, we analyzed whole exome sequences (61.7 Mb) from 15 species of hares (1- 4 individuals per species), spanning the global distribution of the genus, and two outgroups. We used a coalescent framework to infer species relationships and divergence times, despite extensive genealogical discordance. We found high levels of allele sharing among species and show that this reflects extensive incomplete lineage sorting and temporally layered hybridization. Our results revealed recurrent introgression at all stages along the *Lepus* radiation, including recent gene flow between extant species since the last glacial maximum, but also pervasive ancient introgression occurring since near the origin of the hare lineages. We show that ancient hybridization between northern hemisphere species has resulted in shared variation of potential adaptive relevance to highly seasonal environments, including genes involved in circadian rhythm regulation, pigmentation, and thermoregulation. Our results illustrate how the genetic legacy of ancestral hybridization may persist across a radiation, leaving a long-lasting signature of shared genetic variation that may contribute to adaptation.

KEYWORDS

Ancient introgression; Hybridization; Adaptation; Phylogenomics; *Lepus*

Species radiations are often accompanied by extensive gene flow between nascent lineages (e.g., Lamichhaney et al. 2015; Árnason et al. 2018; Malinsky et al. 2018; Li et al. 2019; Barth et al. 2020). Genetic signatures of hybridization between several closely related species could either represent recent or ongoing introgressive hybridization (Eaton et al. 2015), or the remnants of hybridization among ancestral populations that remain shared among contemporary species (Malinsky et al. 2018; Li et al. 2019). Although these alternatives can be difficult to differentiate in large radiations (Eaton et al. 2015; Malinsky et al. 2018; Vanderpool et al. 2020), both ancient and contemporary introgression has been linked to local adaptation in several systems (e.g., Liu et al. 2015; Gittelman et al. 2016; Meier et al. 2017; Barlow et al. 2018; Giska et al. 2019; Svardal et al. 2020). Thus, unraveling the tempo and contribution of introgression to standing genetic variation within and among species remains a critical step in understanding the overall importance of introgression to evolution.

Reconstructing the history of hybridization between several closely related species requires inferring evolutionary relationships among species while considering the two primary processes – incomplete lineage sorting and gene flow – that may cause sharing of genetic variation among populations (Malinsky et al. 2018). The network multispecies coalescent (NMSC) model (Than et al. 2011; Solís-Lemus et al. 2017; Degnan 2018) offers one promising framework that appears to resolve species relationships in the face of multiple reticulation events and rapid speciation (Kozak et al. 2018; Edelman et al. 2019). However, the NMSC is still prohibitive for large datasets and relies on the user to choose the number of migration events based on *a priori* hypotheses (Yu and Nakhleh 2015). Alternatively, site-based summary statistics based on tree asymmetries (e.g., Green et al. 2010; Pease and Hahn

2015), or admixture proportions (e.g., Reich et al. 2009; Martin et al. 2015; Malinsky et al. 2018) are simpler to implement, but offer less power for localizing the timing and number of introgression events when recurrent hybridization is layered across a phylogeny (Malinsky et al. 2018). A combination of methods is thus most appropriate to infer a species tree that may have layered events of hybridization throughout time (e.g., Kozak et al. 2018; Malinsky et al. 2018; Edelman et al. 2019; Li et al. 2019).

Hares and jackrabbits comprise a group of 32 currently classified species (genus *Lepus*; Smith et al. 2018) whose common ancestor likely originated in North America and spread throughout most of the Northern Hemisphere and Africa presumably in the last 4-6 million years (Yamada et al. 2002; Matthee et al. 2004; Melo-Ferreira et al. 2012). Hares are primarily associated with open grasslands, but can be found across a broad range of biomes (e.g., desert, forest, or arctic) and elevations (e.g., from sea level to the Himalayan or Ethiopian plateau; Smith et al., 2018). The *Lepus* radiation also provides multiple case studies of hybridization and introgression, with admixture detected among several extant pairs of species (e.g., Liu et al. 2011; Melo-Ferreira et al. 2012; Tolesa et al. 2017; Jones et al. 2018; Seixas et al. 2018; Lado et al. 2019; Kinoshita et al. 2019). Selection on introgressed variation has been hypothesized to have aided the range expansion of the Iberian hare (Seixas et al. 2018), and has been directly linked to convergent adaptive evolution of non-white winter coats in populations of two species that change the color of their pelage seasonally (Jones et al. 2018, 2020a; Giska et al. 2019). These studies suggest that the relatively recent exchange of genetic variation among extant *Lepus* species has provided an important source of adaptive variation. However, the phylogenetic relationships among *Lepus* species remain poorly resolved (Halanych et al. 1999; Matthee et al. 2004; Melo-Ferreira et al. 2012; Melo-Ferreira and Alves 2018), and the contribution of ancient gene flow to the *Lepus* evolutionary history remains unknown.

Here, we use exome-wide data to infer the evolutionary history of 15 *Lepus* species and show that hybridization between lineages has likely occurred since the origin of the radiation. The combination of incomplete lineage sorting and these temporally layered events of hybridization have resulted in extremely high levels of shared genetic variation among extant species, including species that currently occur on different continents. We then use the case of ancient admixture among northern latitude species that occupy highly seasonal environments to investigate the gene content and possible functional relevance of introgressed genomic regions. Our work demonstrates that recurrent introgression throughout evolutionary history has made a substantial contribution to genetic variation within and among species of this widespread mammalian radiation.

MATERIALS AND METHODS

Taxon Sampling and Exome Sequencing

We generated new genome-wide resequencing data targeting 207,691 exonic and non-coding regions [totaling 61.7 Megabases (Mb) total] from 14 hare species (30 individuals) and the outgroup pygmy rabbit (*Brachylagus idahoensis*; 2 individuals). We combined these data with published whole exomes from four snowshoe hares [*Lepus americanus*; NCBI Sequence Read Archive BioProject PRJNA420081 from Jones et al. (2018, 2020b)] and extracted data from the reference genome of the European rabbit (*Oryctolagus cuniculus*; OryCun2.0; Carneiro *et al.*, 2014) to use as a second outgroup. Our total sample of 15 hare species (34 individuals, 1 to 4 individuals per species) and 2 outgroup species (3 individuals) included species from all major regions of the *Lepus* native distribution: Africa (3 species), Africa and Eurasia (1 species), Eurasia (6 species) and North America (5 species, see Fig. 1 and Supplementary Table S1 available on Dryad; doi:10.5061/dryad.bzkh18967).

Exome capture experiments were performed following the procedures outlined in Jones et al. (2018) and in the Supplementary Material and Methods. Briefly, we obtained samples as tissue or extracted DNA including samples from previous studies or through loans from collaborators (Supplementary Table S1). Depending on the sample, genomic DNA was isolated using a saline extraction method (Sambrook et al. 1989) or DNeasy Blood & Tissue Kit (Qiagen) (Supplementary Table S1 and Supplementary Materials and Methods). We prepared Illumina sequencing libraries for each sample following Meyer and Kircher (2010) with minor modifications [see Supplementary Material and Methods and Jones et al. (2018)]. Sequencing libraries were then enriched using NimbleGen SeqCap EZ v.4.3 protocol and a custom capture design consisting of 213,164 probes targeting ~25 Mb of protein-coding exons, ~28 Mb of untranslated regions, and ~9 Mb of intron/intergenic regions (Jones et al. 2018). Hybridization reactions were performed in two separate equimolar pools of indexed libraries (31 and 29 libraries, Supplementary Table S1), together with samples used for other studies. The target enriched pools were each sequenced across two lanes of an Illumina HiSeq1500 sequencer (125 bp paired-end reads) at CIBIO-InBIO's New-Gen sequencing platform, Portugal.

Read Processing and Genotyping

We trimmed adapters, low-quality bases, merged overlapping reads, and removed PCR duplicates from raw reads using the *expHTS* pipeline (v.0.Mar112016; https://github.com/msettles/expHTS). We then used *pseudo-it* (v1; Sarver et al. 2017) to generate pseudo-reference exomes for each species by iteratively mapping (four iterations and allowing ambiguity codes) cleaned single and paired-end reads from one individual per species (Supplementary Table S1) to the European rabbit reference genome (OryCun2.0) (Carneiro et al. 2014). For snowshoe hares and black-tailed jackrabbits (*L. californicus*), we

used pseudo-references generated by Jones et al. (2018). The genomes of hares and European rabbits are broadly syntenic despite two known large scale fusions/fissions (Robinson et al. 2002) and sufficiently closely related that iterative mapping to the rabbit reference is not expected to bias genotyping (Sarver et al. 2017; Marques et al. 2020).

We then mapped data from all individuals to each species-specific pseudo-reference using *bwa-mem* (v.0.7.12-r1039; Li 2013) with default options. Mapped reads were sorted with *samtools* (v1.4; Li et al. 2009), assigned to read groups, filtered for duplicates (*Picard* v1.140; http://broadinstitute.github.io/picard/), and realigned for insertion-deletion length variation using GATK (v3.4.46; Van der Auwera et al. 2013). We calculated coverage statistics and capture efficiency using *CalculateHSMetrics* from *Picard*. For each individual, we called and filtered genotypes using the *bcftools* (v1.4; Li 2011) mpileup, call and filter pipeline. We excluded indels and filtered single nucleotide variant genotypes with low mapping (MQ < 20) or phred-scaled quality (QUAL < 20), low sequencing depth (DP < 6), excess sequencing depth (3x individual coverage, see Supplementary Table S2), sites less than 10 bases from an indel (--SnpGap 10), and low quality non-reference homozygous or heterozygous genotypes (GQ <= 20). We used the filtered genotypes to construct consensus exome fasta sequences in the OryCun2.0 coordinate system.

Species Tree Inference

We used a concatenated alignment without missing data to estimate a single bifurcating phylogeny using a maximum likelihood (ML) search and rapid bootstrapping run under the GTR+Γ model of sequence evolution (autoMRE option) in *RAxML* (v8.2.10; Stamatakis 2018). We then used two complementary methods to infer multispecies-coalescent trees while accounting for local variation in phylogenetic histories along the genome. First, we extracted 50 kilobase (kb) alignments from the exome capture targeted regions plus 200

base pairs (bp) of flanking sequences using *msa_split* (*phast* (1.4; http://compgen.cshl.edu/phast/) and *bedtools* (1.9; Quinlan and Hall 2010), considering a balance between the expected extent of linkage disequilibrium in hares (10-20 kb; Jones et al. 2018) and the retention of information for phylogenetic analysis (alignment length > 1 kb). For each window, we filtered positions with missing data for > 30% of individuals using *TriSeq* (*TriFusion 1.0.0;* http://odiogosilva.github.io/TriFusion/), excluded windows smaller than 1 kb, and used *RAxML* to estimate local maximum likelihood gene trees (GTR+Γ, 100 bootstraps). For each gene tree, we used the corresponding bootstraps to calculate a tree certainty score based on the sum of certainty scores for all internodes of a tree (-L MRE option in *RAxML*; Salichos and Rokas 2013; Salichos et al. 2014; Kobert et al. 2016). The internode certainty score weighs the support of the bipartition represented by a given internode in the gene tree against the support of the second most prevalent conflicting bipartition (Salichos et al. 2014) present in the bootstrap trees. In our case, the maximum theoretical value of the tree certainty score is 31, or k – 3 with k equal to the number of taxa (Salichos et al. 2014). Only trees with certainty score above 5 were used in the species tree inference. We unrooted the gene trees using R package *ape* (Paradis et al. 2004) and estimated a consensus species tree using *ASTRAL-III* (5.6.3; Zhang et al. 2018).

We also estimated coalescent species trees using only variable sites with *SVDquartets* (Chifman and Kubatko 2014) implemented in *PAUP** (4a163; Swofford 2003). For the analyses based on variable sites, we recovered single nucleotide variants (SNVs) distanced 10 kb along the genome (within targeted regions and 200 bp flanking regions) using *snp-sites* (v2.3.3; Page et al. 2016), a custom script and *bedtools intersect*, and excluded sites with missing information for >30% of the individuals using a custom script (available at https://github.com/evochange). For both *SVDquartets* and *ASTRAL* analyses, species trees were estimated with and without assigning species identities and using sites/intervals genome-

wide or only from the X chromosome. The European and/or pygmy rabbits were included for all analyses requiring outgroups. Additional details on the phylogenetic analyses are provided in the Supplementary Materials and Methods.

Bayesian Divergence Time Inference

We performed Bayesian inference of divergence times in the inferred species tree using an approximate maximum likelihood method and assuming an autocorrelated relaxed molecular clock, implemented in *MCMCtree* (*PAML v.4.9*; Yang 2007) and described in dos Reis and Yang (2011, 2019). For one individual per species (Supplementary Table S1), we extracted the coding sequence for all genes included in our capture design (18,798 genes in the OryCun2.0 ENSEMBLE 94 database) with *bedtools getfasta*, after selecting the longest transcript per gene using R package *biomaRt* (v2.34.2; Durinck et al. 2005, 2009). We excluded alignments with more than 20% missing data using *AMAS* (Borowiec 2016) (see Supplementary Materials and Methods available on Dryad for details). With these, we constructed a concatenated alignment with three partitions, corresponding to the three codon positions. We assumed GTR+Γ for the model of sequence evolution and we used the prior of 3.33 for the average substitution rate per site per 100 million years, following Matthee et al. (2004). *Lepus* is poorly represented in the fossil record. The earliest hare record dates to the early Pleistocene [2.5 million years ago (Ma); White 1991; Lopez-Martinez 2008], which is much more recent than molecular estimates for the genus extrapolated from deeper fossil calibrations (4-6 Ma; Yamada et al. 2002). Therefore, we used two different calibrations for the species tree. We either (1) used molecular estimates of 4-6 million years (myr) for the most recent common ancestor (TMRCA) of *Lepus* extrapolated from deep fossil record calibrations of the order Lagomorpha (Yamada et al. 2002) and 9.7-14.5 myr for the TMRCA of *Oryctolagus-Lepus* divergence (Matthee et al. 2004) or (2) used fossil estimates of 2.5 myr

for the lower bound of *Lepus* diversification and constrained the root to be no older than 4.8 myr, which is when the fossil record suggests that the common ancestor of rabbits and hares existed (Hibbard 1963; White 1991). We discarded the first 1,000,000 samples as burn-in and ran the program until we gathered 1,000,000 samples from the posterior, sampling every 10 iterations, and repeated the analysis twice to ensure convergence. We checked for lack of convergence between the two runs by confirming a linear correlation between posterior times, trendless trace plots, and high effective sample size values (ESS > 200) following dos Reis and Yang (2019). Finally, we also checked for a linear relationship between posterior times and confidence interval widths in infinite sites plots (Inoue et al. 2010).

Ancestral Range Reconstruction

We used the R package *BioGeoBEARS* (Matzke 2013, 2014a, 2014b) to estimate ancestral ranges for the species in our dated phylogenies. For this analysis, we assigned each species to North America, Eurasia or Africa. Note that the Cape hare *sensu lato* is distributed in Africa and Eurasia, but Lado et al. (2019) showed deep divergence and non-monophyly of African and Eurasian lineages. Because our Cape hare samples represent the African lineages, we assigned the Cape hare distribution to Africa. In *BioGeoBEARS*, we implemented a maximum likelihood framework to estimate ancestral ranges under three biogeographical models: DEC (dispersal-extinction-cladogenesis; Ree and Smith 2008), DIVA (dispersal-vicariance analysis; Ronquist 1997) and BayArea (Bayesian Inference of Historical Biogeography for Discrete Areas; Landis et al. 2013). For each method, we tested each model with and without founder-event speciation (*j* parameter; Matzke 2014b), resulting in six models in total: DEC, DEC+J, DIVALIKE, DIVALIKE+J, BAYAREALIKE and BAYAREALIKE+J. The best fitting model was assessed with the likelihood ratio test (LRT), the Akaike Information criterion (AIC) and corrected AIC (AICc).

250

251 *Gene Tree Discordance and Phylogenetic Networks*

252 We performed a series of analyses to explore the amount and effect of gene tree
253 discordance in our dataset. To minimize misinterpreting poor phylogenetic resolution as
254 discordance, we retained only the filtered gene trees used as input for *ASTRAL.* These trees
255 were used to estimate a *SplitsTree4* network (v4.14.6; Huson and Bryant 2006) with the
256 option "Consensus Network with distances as means" and a 5% weight threshold to draw a
257 split, and a majority rule consensus tree with *RAxML* (-L MRE option), including internode
258 certainty scores (Salichos et al. 2014). We then used *DiscoVista* (Sayyari et al. 2018) to plot
259 *ASTRAL* quartet frequencies around nodes of interest. Finally, we calculated Robinson-Foulds
260 normalized distances between gene trees and the *ASTRAL* species tree where individuals are
261 not assigned to species, and among gene trees using the function *RF.dist()* from the R package
262 *phangorn* (2.4.0; Schliep 2011). This metric varies between 0 (no discordance between trees)
263 and 1 (complete discordance).

264 We used *PhyloNet* (v3.6.6; Yu and Nakhleh 2015) to model species relationships
265 under the network multispecies coalescent model, using all local genealogies (tree certainty
266 scores > 5). Given our extensive dataset, we applied the pseudo-maximum likelihood
267 inference of species networks (Yu and Nakhleh 2015). We ran *InferNetwork_MPL* (Yu and
268 Nakhleh 2015) with 0 up to 4 migration events (due to computational limitations), associating
269 individuals to species (option -a), and optimizing branch lengths and inheritance probabilities
270 to compute likelihoods for each proposed network (option -o). We used the best likelihoods
271 per run to calculate BIC and AICc following Yu et al. (2012, 2014) to compare the resulting
272 networks. Networks were visualized with *IcyTree* (https://icytree.org; last accessed July
273 2019).

Given that the pseudo-maximum likelihood network inference is computationally intensive (Yu and Nakhleh 2015), network support could not be inferred with bootstraps. Therefore, we additionally used *Treemix* and 'f-branch' statistics (see below) to determine if ancestral reticulations were consistently inferred across methods. We reconstructed ancestral population graphs with *TreeMix* (v1.13, options -global, -noss and -se; Pickrell and Pritchard 2012) following the steps used to generate an SNV dataset for *SVDQuartets* but only including *Lepus* species. We randomly phased the SNV alignments with a custom script (available at https://github.com/evochange) and created the *TreeMix* input using the script *TreeMix_from_nex.py* (https://github.com/mgharvey). We allowed 0 to 9 migration events and used the white-sided jackrabbit (*L. callotis*) as the outgroup relative to all other *Lepus* species (see *Results*).

Genetic Diversity, Divergence, and Admixture

We used the *genomics general* toolkit (https://github.com/simonhmartin/; last accessed January 14, 2019) to estimate pairwise genetic distances (d_{xy}) between species and nucleotide diversity (π) within species, and a custom script (available at https://github.com/evochange) to calculate the number of heterozygous sites per individual and the subset shared between at least one individual of each pair of species (used as proxy for shared variation). All diversity estimates were based on a genome-wide concatenated alignment, where we excluded sites with missing information for > 30% of the individuals.

We then used *genomics general* and custom scripts (available at https://github.com/evochange) to calculate several variants of the D-statistics (Green et al. 2010) from the informative sites in the same filtered alignment, treating the European rabbit sequence as the ancestral state (additional details are provided in the Supplementary Materials and Methods). Briefly, we calculated the minimum absolute value of D (D_{min}) (Malinsky et al.

2018) for all possible species trios. We calculated z-scores for each D value using a 1 Mb block jackknife approach. After finding the minimum D per trio, D values with Bonferroni-corrected $P \leq 0.05$ were considered significantly different from zero. We then calculated the 'f-branch' statistic ($f_b(C)$) (Malinsky et al. 2018). The 'f-branch' statistic measures admixture proportion between species *C* and branch *b* by calculating admixture proportion among all possible *f*(A,B,C,O) combinations where *A* are all descendants of branch *a* (sister to *b*), *B* are all descendants of branch *b*, and *C* is the donor taxa. $f_b(C)$ is the minimum *f* value across all possible *B* and the median across all possible *A*. A significant $f_b(C)$ value means that all descendants *B* of branch *b* share alleles with *C,* which is more parsimoniously explained by an event of ancestral introgression from *C* to *b* (Malinsky et al. 2018). Using the inferred species tree, we determined all conformations (A,B;C,O) needed to calculate $f_b(C)$ for all pairs of *C* species and *b* branches following Malinsky et al. (2018) with custom scripts (available at https://github.com/evochange) and R package *treeman* (1.1.3; Bennett et al. 2017). For each conformation, we calculated 'admixture proportion' (f_G) as defined in Martin et al. (2015) and Malinsky et al. (2018) and z-scores with 1 Mb block jackknife approach following Malinsky et al. (2018). 'f-branch' values with Bonferroni-corrected $P \leq 0.05$ were considered significantly different from zero. We also calculated f_{hom} (Martin et al. 2015) between black-tailed jackrabbits (P3) and several snowshoe hare populations (as P1 and P2) to evaluate levels of admixture estimated with $f_b(C)$ for this species pair (see *Results* and *Discussion*).

We also used *ABBABABAwindows.py* from *genomics general* to estimate the fraction of admixture (f_d) (Martin et al. 2015) across 50 kb genomic sliding windows (≥ 100 sites, 5 kb steps), to localize tracts of introgression in the genomes of northern latitude species. We performed three scans testing introgression between snowshoe hares (*L. americanus*) as P3 and Alaskan hares (*L. othus*), mountain hares (*L. timidus*) or white-tailed jackrabbits (*L. townsendii*) as alternative P2, using the Iberian hare (*L. granatensis*) as P1. Windows of top

0.5% f_d were considered significant. Following Liu et al. (2015), we considered that significant windows in all three tests reflected introgression between snowshoe hares and the ancestral lineage of white-tailed jackrabbits/mountain hares/Alaskan hares, while significant windows in only one test result from recent introgression between the focal extant species. We obtained the annotation of genes in these windows from the European rabbit reference using *biomaRt* and used unique annotations to perform an enrichment analysis in *g:Profiler* (accessed September 2019; Raudvere et al. 2019) using default parameters. We also calculated d_{xy} for f_d outlier windows and the exome-wide d_{xy} distribution between the focal pair of P2-P3 species using *popgenWindows.py* from *genomics general* to calculate d_{xy} in windows of 50 kb ($\geq$ 2000 sites, 5 kb steps).

RESULTS

Whole Exome Sequencing Data

Custom DNA captures showed high efficiency (32.3 average fold-enrichment) and specificity (average 10% of sequenced bases off-target; Supplementary Table S2). Mapping cleaned reads onto species-specific pseudo-references resulted in an average target sequencing coverage of 16× (5-35× on average per sample; Supplementary Table S2) with 57.8 million genotyped sites per individual (Supplementary Table S2). Two lower coverage individuals (one hare individual and one pygmy rabbit individual) and one locality duplicate (one hare) were removed from the final dataset to maximize data quality and avoid geographic redundancy (Supplementary Table S1). Combining new and published data, all analyses were performed on a dataset of 15 hare species (32 individuals), one pygmy rabbit, and the European rabbit reference genome, unless otherwise noted (see Supplementary Table S1).

Phylogenetic Relationships among Hares

The overall topologies of the concatenated ML phylogeny (11,949,529 positions with no missing data) and the multispecies-coalescent species trees of *ASTRAL* (8,889 gene trees estimated from 50 kb genomic intervals; alignment lengths between 1 kb and 29 kb) and *SVDquartets* (45,779 unlinked SNPs) were largely concordant (Supplementary Figs. S1, S2 and S3). Most branching relationships were highly supported in general (*ASTRAL* posterior probabilities = 1, *SVDquartets* bootstraps > 90; Fig. 1 and Supplementary Figs. S1, S2, S3 and S4) with most species recovered as monophyletic (Supplementary Figs. S1, S2a and S3a). The only exceptions involved very closely related species with disputed taxonomy, such as paraphyly of the mountain hare with the Alaskan hare, the broom hare (*L. castroviejoi*) with the Corsican hare (*L. corsicanus*) (Alves et al. 2008; Melo-Ferreira et al. 2012), or paraphyly among a lineage showing deep intraspecific divergence, the Cape hare (*L. capensis*), with the Ethiopian hare (*L. fagani*) (Lado et al. 2019) (Supplementary Figs. S1, S2 and S3 and Supplementary Tables S3 and S4).

Our two calibration strategies recovered overlapping 95% high posterior density (HPD) intervals of divergence times for relatively recent splits (e.g., diversification of Eurasian and African species; Supplementary Table S5), but the *Lepus*-based fossil calibrations suggested more recent ages for deeper nodes (Supplementary Table S5 and Fig. S5). For instance, the hare radiation was estimated at ~5.83 million year ago (95% HPD 6.17-5.34 Ma) using fossil calibrations outside *Lepus*, while *Lepus* fossil calibration sets it at ~4.05 Ma (95% HPD 5.00-3.18 Ma). All species tree analyses were consistent in showing that the deepest branching events involved North American species: white-sided jackrabbit, black-tailed jackrabbit, and snowshoe hare (Fig. 1b). In accordance, ancestral reconstruction of biogeographic distributions based on the best fitting DIVALIKE+J model supported a North America origin, with subsequent colonization of Eurasia (2.85 and 1.99 Ma for deeper and

374 *Lepus*-fossil calibration, respectively) and Africa (1.85 and 1.33 Ma) (Fig. 1c and
375 Supplementary Table S6). We also found support for one or more re-colonization events of
376 North America from Eurasia, represented by white-tailed jackrabbits and Alaskan hares (Fig.
377 1 and Supplementary Fig. S6 and Table S7).
378
379 *Incomplete Lineage Sorting and Introgression*
380 We recovered a highly supported species tree across phylogenetic methods, albeit with
381 considerable phylogenetic discordance among sequenced regions [average Robinson-Foulds
382 (RF) pairwise distance between local trees was 0.73]. No local tree completely recovered the
383 species tree topology (minimum RF distance between gene and species tree was 0.13)
384 (Supplementary Fig. S7) and the majority rule consensus tree showed low internode certainty
385 (Supplementary Fig. S8). Phylogenetic discordance was also apparent in the *ASTRAL* species
386 tree with all but five branches showing quartet scores below 0.6 (Fig. 1b, Supplementary Figs.
387 S2, S4 and Tables S8 and S9). The inference of a concordant species tree across *ASTRAL* and
388 *SVDquartets* suggests that the low quartet scores are not highly impacted by gene tree
389 estimation error (Molloy and Warnow 2018). Splits network analysis of individual gene trees
390 also supported many alternative relationships (represented by cuboid structures connecting
391 alternative topologies) particularly involving deeper branches (Fig. 2). We estimated similar
392 levels of discordance between X-linked and autosomal genealogies and the inferred species
393 trees (Supplementary Fig. S7). Furthermore, a species tree inferred with X-linked data
394 differed from the genome-wide species tree and showed lower overall branch support
395 (Supplementary Fig. S9 and Table S10). Nucleotide diversity was relatively low within
396 species (0.13-0.63%) and overlapped with estimates of absolute genetic divergence between
397 species (d_{xy} 0.17-1.11%; Supplementary Tables S3 and S4). However, these ranges reflect
398 some taxonomic uncertainties. For example, some instances of low interspecific divergence

399 concern species with debated species-level status (e.g., the Corsican-broom hare and
400 mountain-Alaskan hare complexes; Alves et al. 2008; Melo-Ferreira et al. 2012), while high
401 diversity may reflect intraspecific cryptic divergence (e.g., snowshoe hares, Melo-Ferreira et
402 al. 2014; Cape hares, Lado et al. 2019).

403 On average, 49% of heterozygous sites where shared between species (Supplementary
404 Table S11). Considerable phylogenetic discordance among gene trees combined with such
405 high levels of shared polymorphism between species could be explained by incomplete
406 lineage sorting, secondary introgression, or a combination of both processes. We found that
407 88% of the minimum absolute D-statistics (D_{min}) (Malinsky et al. 2018) for all trios of species
408 in our dataset were significantly different from zero (Bonferroni-corrected $P < 0.05$;
409 Supplementary Fig. S10), providing overwhelming support for gene flow either between
410 extant species pairs or between ancestral lineages (Malinsky et al. 2018). Our inferences of
411 the multispecies network coalescent implemented (Supplementary Fig. S11), ancestral
412 population graph reconstruction (based on 30,709 biallelic SNVs; Supplementary Figs. S12
413 and S13), and estimates of admixture proportions among species based on the 'f-branch'
414 metric ($f_b(C)$) (Fig. 3) were all consistent with recurrent gene flow between species layered
415 across the diversification of *Lepus*. With *fb(C)* (Fig. 3) and network analysis (Supplementary
416 Figs. S11 and S12), we detected introgression among extant species pairs within all of the
417 major geographic regions that are currently sympatric, suggesting ongoing or recent
418 hybridization. For example, strong gene flow was found between black-tailed jackrabbits and
419 snowshoe hares in North America ($f_b(C)$ = 19%; P = 2.04E-169), or between European brown
420 hares and mountain hares from Eurasia ($f_b(C)$ = 18%; P = 1.73E-233; Fig. 3 and
421 Supplementary Table S12). In general, we found decreased admixture proportions with
422 increased genetic divergence between species, although this correlation was only significant
423 when considering species with non-overlapping distributions (Fig. 4). Several species pairs

with current contact zones showed admixture even when genetic divergence was relatively high, such as snowshoe hares and black-tailed jackrabbits (d_{xy} = 0.97%; TMRCA ~ 4.8 Ma, Supplementary Fig. S5 and Tables S4 and S5).

We also found evidence for introgression between ancestral populations, which has likely affected deeper branches of the species tree. These ancestral events were detected as network reticulations, as significant *fb(C)* values among ancestral branches (y-axis in Fig. 3) or similar admixture levels shared among species within clades (x-axis in Fig. 3). We detect ancestral events of introgression connecting major clades within Eurasia (e.g., European hares and the ancestor of the mountain hare/Alaskan hare/white-tailed jackrabbit clade), African and Eurasian lineages (e.g., ancestral of all African species and the Corsican and broom hare ancestral), and Eurasian and North American lineages (e.g., snowshoe hares and the ancestral lineage of the mountain hare/Alaskan hare/white-tailed jackrabbit; Fig. 3, 4, Supplementary Figs. S11 and S12, and Table S12). Finally, we found significant introgression among species from different continents, such as North American hares and species from Africa and Western Europe (Fig. 3 and 4), which was also suggested by a network with two reticulations and by ancestral population graph reconstruction (Supplementary Figs. S11c and S12). These results suggest that introgression affected the very early branches of the *Lepus* radiation, and that the genetic legacy of these gene flow events persists in the gene pool of descendant species today (Fig. 3 and 4).

Genes affected by Ancestral Introgression

We consistently detected ancestral introgression between the ancestors of coat color changing species in our dataset: white-tailed jackrabbits, mountain hares, Alaskan hares, and snowshoe hares (Fig. 3 and Supplementary Fig. S11). Motivated by recent work showing that introgressive hybridization has shaped local adaptation in seasonal coat color changing

449 species (Jones et al. 2018, 2020a; Giska et al. 2019), we examined the contribution of gene
450 flow to standing variation and local adaptation in these species. We detected 119 putative
451 windows of ancient introgression across all major chromosomes, highlighting the genome-
452 wide contribution of ancestral introgression (Fig. 5; Supplementary Table S13). The f_d outlier
453 windows of ancestral introgression contained 54 annotated genes (Supplementary Table S14).
454 This set of genes was enriched for the gene ontology term "E-Box binding" (3 of the 54
455 genes; Supplementary Table S15), a DNA motif found in the promoters of many genes,
456 suggesting that genomic regions affected by ancient introgression may be enriched for
457 transcription factors involved in *trans*-regulation of gene expression. Among these
458 transcription factors we found the circadian clock related gene *ARNTL2* (Sasaki et al. 2009)
459 and pigmentation related gene *TCF4* (Furumura et al. 2001; Le Pape et al. 2009) (Fig. 5). In
460 addition, the list of 54 genes includes a gene involved in brown fat differentiation (*EBF2*;
461 Rajakumari et al. 2013), and a photoreceptor-related gene (*PDE6H*; Kohl et al. 2012).
462
463 **DISCUSSION**
464 We used phylogenetic analyses of whole exome data to tease apart signatures of
465 stochastic lineage sorting and admixture across the evolutionary history of the *Lepus*
466 radiation. By accounting for these sources of phylogenetic discordance, we were able to
467 detect pervasive introgression across the evolution of this recent and rapid mammalian
468 radiation. Below we discuss the biogeographic and evolutionary implications of our analyses,
469 focusing on the long-term impacts of temporally layered hybridization in shaping patterns of
470 shared genetic variation within and among extant species.
471
472 *The Effect of Persistent Gene Flow on Phylogenetic Inference*

We present a resolved genome-wide phylogeny for the genus *Lepus* despite extensive incomplete lineage sorting and pervasive gene flow. Our genome-wide analysis covered all major lineages across the worldwide range of hares, substantially extending previous analyses based on more limited genetic sampling (Halanych et al. 1999; Melo-Ferreira et al. 2012; Ge et al. 2013; Tolesa et al. 2017). A systematic evaluation of species limits and hare taxonomy is beyond the scope of our work, as it would require genome-wide data from an expanded inter and intraspecific sample. However, our sampling should not bias the estimated phylogenetic relationships and may even underestimate species diversity. Thus our inferences of shared variation across species are likely to be conservative.

To estimate species relationships, we combined species tree inferences that do not account for gene flow with network-based inferences that explicitly consider introgression. The resulting topologies were generally consistent between methods when considering up to three reticulations (Fig. 1 and Supplementary Fig. S11). However, there were limits to this approach. For example, the alternative placement of European, Corsican, and broom hares, as closer to the Eurasian or African *Lepus* clades depended on the number of reticulations considered (Supplementary Fig. S11). This uncertainty likely reflects long-term admixture among Eurasian and African lineages, including the European hare (Fig. 3 and Supplementary Fig. S11) whose range overlaps with species from both continents (Fig. 1a). Another example is the alternative sister relationship of snowshoe hares and black-tailed jackrabbits in the network with no reticulation (Supplementary Fig. S11), which could result from ancestral introgression between these species (see below). While an increased number of reticulation events could better represent the widespread gene flow uncovered in our work, allowing for more than three reticulations resulted in increased branch compression and therefore did not help resolve gene-to-gene incongruences (Supplementary Fig. S11; Yu and Nakhleh 2015; Wen et al. 2016).

497 We found that gene trees constructed from X-linked loci or autosome-linked loci
498 showed similar levels of discordance with the exome-wide species tree (Supplementary Fig.
499 S7). In general, the X chromosome could be expected to show less phylogenetic discordance
500 due to its smaller effective population size and faster lineage sorting, and a tendency to
501 accumulate hybrid incompatibilities (i.e., the large X-effect; Fontaine et al. 2015; Edelman et
502 al. 2019; Li et al. 2019). Although there is evidence for reduced X-linked gene flow between
503 some hybridizing European lineages (Seixas et al. 2018), very little is known about the
504 genetic architecture of reproductive isolation between hare species. Moreover, less
505 phylogenetic discordance of loci involved in reproductive isolation may not be always
506 expected, particularly when speciation events are clustered in time (Wang and Hahn 2018) as
507 is the case in *Lepus* (Supplementary Fig. S5).
508
509 *The Timing and Biogeography of the Lepus Radiation*
510 Our results suggest that the *Lepus* diversification followed major climatic shifts that
511 occurred during the late Miocene, Pliocene and Pleistocene, similar to other terrestrial
512 mammals (Simpson 1947; Ge et al. 2013). We inferred a North American origin between 5.83
513 Ma and 4.05 Ma [deeper or *Lepus* fossil calibrations, respectively, in agreement with Hibbard
514 (1963), Halanych et al. (1999), Melo-Ferreira et al. (2012) and Ge et al. (2013)], coinciding
515 with a global cold and dry period of the late Miocene that favored the expansion of grasslands
516 worldwide (Osborne and Beerling 2006; Ge et al. 2013). The establishment of the Bering land
517 bridges during late Pliocene and Pleistocene glacial periods (Simpson 1947; Hopkins 1959;
518 Cook et al. 2016) may have allowed the subsequent colonization of Eurasia (~2.77-1.99 Ma).
519 Finally, we inferred secondary colonization of North America from Eurasia within the last 1
520 million years (Fig. 1 and Supplementary Figs. S5 and S6; Halanych et al. 1999), in agreement

521 with the Eurasia-North America exchange of cold-adapted fauna during Pleistocene glacial
522 periods (Simpson 1947; Hoberg et al. 2012).
523 We note that our finding of pervasive introgression may distort aspects of these
524 biogeographic inferences (Leaché et al. 2014; Solís-Lemus et al. 2016; Long and Kubatko
525 2018; Li et al. 2019; Jiao et al. 2020). If the inferred topology of the *Lepus* phylogeny is
526 correct, as discussed above, then the general biogeographic reconstructions are probably
527 robust. However, the timing of these events remains tentative given potential estimation bias
528 caused by ancestral gene flow (Leaché et al. 2014; Li et al. 2019) and a lack of varied,
529 independent, and reliable calibration points.
530
531 *The Legacy of Introgression during the Rapid Radiation of Hares*
532 Our inferences indicate that lineages often hybridized when they came into contact
533 during the worldwide expansion of *Lepus*. Some of the inferred reticulation events overlap
534 with known recent introgression, such as between mountain hares and European hares
535 (Levänen et al. 2018) or between snowshoe hares and black-tailed jackrabbits (Jones et al.
536 2018, 2020a). However, our analysis also revealed prevailing signatures of deeper
537 hybridization between ancestral populations, suggesting a persistent contribution of secondary
538 introgression during the diversification of hares. These past hybridization events have resulted
539 in extensive shared polymorphism among extant species (Fig. 3 and Supplementary Tables
540 S11, S12 and Figs. S10, S11 and S12), with significant admixture still detected among species
541 with non-overlapping distributions, even from different continents (Fig. 3 and Fig. 4).
542 Our analyses highlight how ancient gene flow can obscure accurate detection of
543 contemporary hybridization. Similar to other systems (Malinsky et al. 2018; Edelman et al.
544 2019; Li et al. 2019), we estimated high levels of introgression based on D-statistics (e.g.,
545 88% of D_{min} values across all possible species trios were significant), but less reticulation

after taking phylogeny into account (e.g., 33% of 'f-branch' statistics were significant). These findings suggest that phylogenetic correlation causes non-independence of summary-statistics and can thus lead to false pairwise inferences of gene flow between species (Eaton et al. 2015; Malinsky et al. 2018; Li et al. 2019). Given these results, signatures of hybridization among closely related species should be interpreted in the context of broader phylogenetic relationships whenever possible.

We also detected some discrepancies between the magnitude of gene flow inferred here and in previous works, which underscores the challenges of quantifying introgression across a reticulating radiation. For instance, we inferred substantial admixture proportions between black-tailed jackrabbits and snowshoe hares ($f_b(C)$ = 19%, Fig. 3 and Supplementary Table S12). These estimates are one order of magnitude higher than recent studies suggesting that ~2-3% of genomic variation in the Pacific Northwest snowshoe hare populations descends from hybridization with black-tailed jackrabbits in the last ~10,000 generations (Jones et al. 2018, 2020a). Ancient introgression persisting in all snowshoe hare populations could reconcile this discrepancy. Indeed, we recovered consistent results (f_{hom} ~3%) when we infer admixture proportions among black-tailed jackrabbits and different snowshoe hare populations (Supplementary Table S16), which likely reflects recent and geographically localized introgression. However, our estimates of overall admixture proportions between these species should be taken with caution as they depend on the accurate reconstruction of a short internal branch not fully supported across methods, and a species (snowshoe hare) that has been involved in multiple instances of introgression with different hare lineages (Fig. 3, 5 and Supplementary Table S12 and Fig. S11).

Introgression between species is often limited by purifying selection against hybrid incompatibilities (Schumer et al. 2018; Edelman et al. 2019), which agrees with the predominantly negative consequences of hybridization (Mayr 1963). Nonetheless,

hybridization is also expected to produce novel allelic combinations that increase phenotypic variation (Grant and Grant 2019; Marques et al. 2019). If coincident with ecological opportunity, introgressed variation could broadly facilitate adaptation (Grant and Grant 2019; Taylor and Larson 2019). Recent work has shown at least two instances where introgression between *Lepus* species has driven local adaptation (Jones et al. 2018; Giska et al. 2019), and standing introgressed variation may have contributed even more generally to adaptation during the radiation. We found that genetic variation introduced by ancient hybridization can persist through several speciation events (Fig. 3 and 4).

While much of large reservoir of shared variation may reflect stochastic sorting of neutral variation, some may have been maintained by selection (Guerrero and Hahn 2017; Jamie and Meier 2020) and helped facilitate colonization of the diverse habitats currently inhabited by hare species, from desert to arctic environments (Ge et al. 2013; Smith et al. 2018). In this respect, introgression between the ancestor of mountain hares/Alaskan hares/white-tailed jackrabbits and snowshoe hares (or an ancestral lineage) is particularly intriguing. These four species have adapted to highly seasonal environments through striking forms of phenotypic plasticity (e.g., seasonal coat color change; Mills et al. 2013; Zimova et al. 2018) that have been at least partially shaped by adaptive introgression at the *Agouti* pigmentation gene from non-color changing species (*ASIP*; Jones et al. 2018; Giska et al. 2019). Here we estimated that a pulse of ancient introgression occurred at least 0.71 Ma (*Lepus* fossil calibration; Supplementary Fig. S5 available on Dryad) and affected genomic regions containing genes associated with circadian rhythm regulation (*ARNTL2*; Sasaki et al. 2009), pigmentation (*TCF4*; Furumura et al. 2001; Le Pape et al. 2009), thermoregulation (*EBF2*; Rajakumari et al. 2013) and visual perception (*PDE6H*; Kohl et al. 2012). While *ASIP* would also be a likely candidate for adaptive introgression between these lineages (Jones et

595 al. 2018), our exome sequencing coverage of this region was too sparse for detailed window-
596 based analysis.

597 The functions of these introgressed genes overlap with common physiological
598 adaptations of northern latitude animals to seasonal conditions, such as higher metabolic
599 rates, regulation of body temperature and non-shivering thermogenesis (Hart et al. 1965; Feist
600 and Rosenmann 1975; Rogowitz 1990; Pyörnilä et al. 2008; Sheriff et al. 2009), seasonal
601 camouflage (Grange 1932; Hewson 1958; Hansen and Bear 1963; Mills et al. 2018; Zimova
602 et al. 2018), and visual acuity in response to dim winter light in northern latitudes (Stokkan et
603 al. 2013). Furthermore, the functions of two of these genes, pigmentation and circadian
604 rhythms, are linked to pathways activated during seasonal coat color changing molts (Ferreira
605 et al. 2017, 2020). The functional relevance of these candidates to local adaptation must await
606 further testing. Nonetheless, substantial introgression along the rapid diversification of a
607 group of organisms, as we describe here, may bolster genetic variation within species and
608 have a greater role in local adaptation than previously anticipated (Grant and Grant 2019;
609 Taylor and Larson 2019). However, we also cannot exclude that some of these shared variants
610 have been maintained by long-term balancing selection rather than secondary introgression
611 (Supplementary Fig. S14; Smith and Kronforst 2013; Liu et al. 2015; Guerrero and Hahn
612 2017). Regardless of origin, the *Lepus* radiation provides an intriguing system by which to
613 test the long-term evolutionary importance of shared genetic variation across a rapid
614 radiation.

615

616 **SUPPLEMENTARY MATERIAL**

617 Raw sequence reads used in this article are available in the National Center for
618 Biotechnology Information (NCBI) Sequence Read Archive (SRA) under accession numbers
619 SRR12020579 to SRR12020510 and BioProject number PRJNA639005. Supplementary

materials, including data files and/or online-only appendices, can be found in the Dryad Digital Repository (doi:10.5061/dryad.bzkh18967). The pipeline and custom scripts used for this study are available at https://github.com/evochange.

FUNDING

Funding and support for this research was provided by Fundação para a Ciência e a Tecnologia (FCT) (project grant "CHANGE" – PTDC/BIA-EVF/1624/2014, Portuguese National Funds) and the National Science Foundation (NSF) (EPSCoR OIA-1736249 and DEB-1907022 grant). MSF was supported by POPH-QREN funds from ESF and Portuguese MCTES/FCT (PD/BD/108131/2015 PhD grant in the scope of BIODIV PhD programme at Faculty of Sciences, University of Porto), FCT (PTDC/BIA-EVF/1624/2014), and NSF (OIA-1736249). MRJ was supported by an NSF Graduate Research Fellowship (DGE-1313190). JM-F was supported by an FCT CEEC contract (CEECIND/00372/2018). Instrumentation, laboratory, and computational support was provided by CIBIO NEWGEN sequencing platform, supported by European Union's Seventh Framework Program for research, technological development and demonstration under grant agreement no. 286431, by the University of Montana Genomics Core, supported by a grant from the M.J. Murdock Charitable Trust. A grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R01HD073439) supported the development of exome capture protocols that were utilized in the current study. Additional support was obtained from European Union's Horizon 2020 Research and Innovation Programme under grant Agreement Number 857251, the Laboratoire International Associé (LIA) "Biodiversity and Evolution" funded by InEE (CNRS, France) and FCT (Portugal), COMPETE2020, PORTUGAL2020, and ERDF (POCI-01-0145-FEDER-022184), and from Portugal-United States of America

644 Research Networks Program funds from Fundação Luso-Americana para o Desenvolvimento
645 (FLAD) to PCA and MSF.

646

647 **ACKNOWLEDGEMENTS**

648 We thank Janet Rachlow, Gloria Portales, Armando Geraldes, José Carlos Brito,
649 Jérôme Letty, Conrad Matthee, Alexei Kryukov, Ettore Randi, Christian Gortazar, Rafael
650 Villafuerte, Fernando Ballesteros, Klaus Hackländer and Zbyszek Boratynski for assistance
651 with sample collection. The Alaskan hare samples used in this work were generously
652 provided by the University of Alaska Museum (voucher codes UAM42143 and UAM45545).
653 We thank Fernando Seixas, Tiago Antão, Hannes Svardal, Nathaniel Edelman, the
654 EVOCHANGE lab, the Good lab, and the UNVEIL network for helpful discussions on the
655 generation and analysis of data.

656

657 **REFERENCES**

658 Alves P.C., Melo-Ferreira J., Branco M., Suchentrunk F., Ferrand N., Harris D.J. 2008.
659 Evidence for genetic similarity of two allopatric European hares (*Lepus corsicanus* and
660 *L. castroviejoi*) inferred from nuclear DNA sequences. Mol. Phylogenet. Evol. 46:1191–
661 7.
662 Árnason Ú., Lammers F., Kumar V., Nilsson M.A., Janke A. 2018. Whole-genome
663 sequencing of the blue whale and other rorquals finds signatures for introgressive gene
664 flow. Sci. Adv. 4.
665 Van der Auwera G.A., Carneiro M.O., Hartl C., Poplin R., del Angel G., Levy-Moonshine A.,
666 Jordan T., Shakir K., Roazen D., Thibault J., Banks E., Garimella K. V, Altshuler D.,
667 Gabriel S., DePristo M.A. 2013. From fastQ data to high-confidence variant calls: The
668 genome analysis toolkit best practices pipeline. Curr. Protoc. Bioinforma. 11.

669 Barlow A., Cahill J.A., Hartmann S., Theunert C., Xenikoudakis G., Fortes G.G., Paijmans
670 J.L.A., Rabeder G., Frischauf C., Grandal-d'Anglade A., García-Vázquez A.,
671 Murtskhvaladze M., Saarma U., Anijalg P., Skrbinšek T., Bertorelle G., Gasparian B.,
672 Bar-Oz G., Pinhasi R., Slatkin M., Dalén L., Shapiro B., Hofreiter M. 2018. Partial
673 genomic survival of cave bears in living brown bears. Nat. Ecol. Evol. 2:1563–1570.
674 Barth J.M.I., Gubili C., Matschiner M., Tørresen O.K., Watanabe S., Egger B., Han Y.S.,
675 Feunteun E., Sommaruga R., Jehle R., Schabetsberger R. 2020. Stable species
676 boundaries despite ten million years of hybridization in tropical eels. Nat. Commun.
677 11:1–13.
678 Bennett D.J., Sutton M.D., Turvey S.T. 2017. Treeman: An R package for efficient and
679 intuitive manipulation of phylogenetic trees. BMC Res. Notes. 10:30.
680 Borowiec M.L. 2016. AMAS: A fast tool for alignment manipulation and computing of
681 summary statistics. PeerJ. 2016:e1660.
682 Carneiro M., Rubin C.J., Palma F. Di, Albert F.W., Alföldi J., Barrio A.M., Pielberg G.,
683 Rafati N., Sayyab S., Turner-Maier J., Younis S., Afonso S., Aken B., Alves J.M.,
684 Barrell D., Bolet G., Boucher S., Burbano H.A., Campos R., Chang J.L., Duranthon V.,
685 Fontanesi L., Garreau H., Heiman D., Johnson J., Mage R.G., Peng Z., Queney G.,
686 Rogel-Gaillard C., Ruffier M., Searle S., Villafuerte R., Xiong A., Young S., Forsberg-
687 Nilsson K., Good J.M., Lander E.S., Ferrand N., Lindblad-Toh K., Andersson L. 2014.
688 Rabbit genome analysis reveals a polygenic basis for phenotypic change during
689 domestication. Science. 345:1074–1079.
690 Chifman J., Kubatko L. 2014. Quartet inference from SNP data under the coalescent model.
691 Bioinformatics. 30:3317–3324.
692 Cook J.A., Galbreath K.E., Bell K.C., Campbell M.L., Carrière S., Colella J.P., Dawson N.G.,
693 Dunnum J.L., Eckerlin R.P., Fedorov V., Greiman S.E., Haas G.M., Haukisalmi V.,

694 Henttonen H., Hope A.G., Jackson D., Jung T.S., Koehler A. V, Kinsella J.M., Krejsa
695 D., Kutz S.J., Liphardt S., MacDonald S.O., Malaney J.L., Makarikov A., Martin J.,
696 McLean B.S., Mulders R., Nyamsuren B., Talbot S.L., Tkach V. V, Tsvetkova A.,
697 Toman H.M., Waltari E.C., Whitman J.S., Hoberg E.P., Cook J., Bell K., Colella J.,
698 Jackson D., Krejsa D., Liphardt S., McLean B., Galbreath K., Haas G., Toman H.,
699 Campbell M., Dunnum J., MacDonald S., Carrière S., Mulders R., Dawson N., Fedorov
700 V., Greiman S., Jung T., Koehler A., Kinsella J., Kutz S., Malaney J., Makarikov A.,
701 Martin J., Nyamsuren B., Waltari Aaron E., Hoberg Animal E. 2016. The Beringian
702 Coevolution Project: holistic collections of mammals and associated parasites reveal
703 novel perspectives on evolutionary and environmental change in the North. Arct. Sci.
704 3:585–617.
705 Degnan J.H. 2018. Modeling Hybridization Under the Network Multispecies Coalescent.
706 Syst. Biol. 67:786–799.
707 Durinck S., Moreau Y., Kasprzyk A., Davis S., De Moor B., Brazma A., Huber W. 2005.
708 BioMart and Bioconductor: a powerful link between biological databases and microarray
709 data analysis. Bioinformatics. 21:3439–3440.
710 Durinck S., Spellman P.T., Birney E., Huber W. 2009. Mapping identifiers for the integration
711 of genomic datasets with the R/ Bioconductor package biomaRt. Nat. Protoc. 4:1184–
712 1191.
713 Eaton D.A.R., Hipp A.L., González-Rodríguez A., Cavender-Bares J. 2015. Historical
714 introgression among the American live oaks and the comparative nature of tests for
715 introgression. Evolution. 69:2587–2601.
716 Edelman N.B., Frandsen P.B., Miyagi M., Clavijo B., Davey J., Dikow R.B., García-Accinelli
717 G., Van Belleghem S.M., Patterson N., Neafsey D.E., Challis R., Kumar S., Moreira
718 G.R.P., Salazar C., Chouteau M., Counterman B.A., Papa R., Blaxter M., Reed R.D.,

719 Dasmahapatra K.K., Kronforst M., Joron M., Jiggins C.D., Owen McMillan W., Palma
720 F. Di, Blumberg A.J., Wakeley J., Jaffe D., Mallet J. 2019. Genomic architecture and
721 introgression shape a butterfly radiation. Science. 366:594–599.
722 Feist D.D., Rosenmann M. 1975. Seasonal sympatho-adrenal and metabolic responses to cold
723 in the Alaskan snowshoe hare (*Lepus americanus macfarlani*). Comp. Biochem. Physiol.
724 -- Part A Physiol. 51:449–455.
725 Ferreira M.S., Alves P.C., Callahan C.M., Giska I., Farelo L., Jenny H., Mills L.S.,
726 Hackländer K., Good J.M., Melo-Ferreira J. 2020. Transcriptomic regulation of seasonal
727 coat color change in hares. Ecol. Evol. 10:1180–1192.
728 Ferreira M.S., Alves P.C., Callahan C.M., Marques J.P., Mills L.S., Good J.M., Melo-Ferreira
729 J. 2017. The transcriptional landscape of seasonal coat color molt in the snowshoe hare.
730 Mol. Ecol. 26:4173–4185.
731 Fontaine M.C., Pease J.B., Steele A., Waterhouse R.M., Neafsey D.E., Sharakhov I. V, Jiang
732 X., Hall A.B., Catteruccia F., Kakani E., Mitchell S.N., Wu Y.-C., Smith H.A., Love
733 R.R., Lawniczak M.K., Slotman M.A., Emrich S.J., Hahn M.W., Besansky N.J. 2015.
734 Extensive introgression in a malaria vector species complex revealed by phylogenomics.
735 Science. 347:1258522–1258522.
736 Furumura M., Potterf S.B., Toyofuku K., Matsunaga J., Muller J., Hearing V.J. 2001.
737 Involvement of ITF2 in the transcriptional regulation of melanogenic genes. J. Biol.
738 Chem. 276:28147–54.
739 Ge D., Wen Z., Xia L., Zhang Z., Erbajeva M., Huang C., Yang Q. 2013. Evolutionary
740 History of Lagomorphs in Response to Global Environmental Change. PLoS One.
741 8:e59668.
742 Giska I., Farelo L., Pimenta J., Seixas F.A., Ferreira M.S., Marques J.P., Miranda I., Letty J.,
743 Jenny H., Hackländer K., Magnussen E., Melo-Ferreira J. 2019. Introgression drives

repeated evolution of winter coat color polymorphism in hares. Proc. Natl. Acad. Sci. 116:24150–24156.

Gittelman R.M., Schraiber J.G., Vernot B., Mikacenic C., Wurfel M.M., Akey J.M. 2016. Archaic Hominin Admixture Facilitated Adaptation to Out-of-Africa Environments. Curr. Biol. 26:1–8.

Grange W.B. 1932. The Pelages and Color Changes of the Snowshoe Hare, *Lepus americanus phaeonotus*, Allen. J. Mammal. 13:99.

Grant P.R., Grant B.R. 2019. Hybridization increases population variation during adaptive radiation. Proc. Natl. Acad. Sci. 116:23216–23224.

Green R.E., Krause J., Briggs A.W., Maricic T., Stenzel U., Kircher M., Patterson N., Li H., Zhai W., Fritz M.H.-Y., Hansen N.F., Durand E.Y., Malaspinas A.-S., Jensen J.D., Marques-Bonet T., Alkan C., Prüfer K., Meyer M., Burbano H.A., Good J.M., Schultz R., Aximu-Petri A., Butthof A., Höber B., Höffner B., Siegemund M., Weihmann A., Nusbaum C., Lander E.S., Russ C., Novod N., Affourtit J., Egholm M., Verna C., Rudan P., Brajkovic D., Kucan Ž., Gušic I., Doronichev V.B., Golovanova L. V, Lalueza-Fox C., de la Rasilla M., Fortea J., Rosas A., Schmitz R.W., Johnson P.L.F., Eichler E.E., Falush D., Birney E., Mullikin J.C., Slatkin M., Nielsen R., Kelso J., Lachmann M., Reich D., Pääbo S. 2010. A draft sequence of the Neandertal genome. Science. 328:710–722.

Guerrero R.F., Hahn M.W. 2017. Speciation as a sieve for ancestral polymorphism. Mol. Ecol. 26:5362–5368.

Halanych K.M., Demboski J.R., van Vuuren B.J., Klein D.R., Cook J. a. 1999. Cytochrome b phylogeny of North American hares and jackrabbits (*Lepus*, Lagomorpha) and the effects of saturation in outgroup taxa. Mol. Phylogenet. Evol. 11:213–221.

Hansen R.M., Bear G.D. 1963. Winter coats of white-tailed jackrabbits in southwestern

769 Colorado. J. Mammology. 44:420–422.
770 Hart J.S., Pohl H., Tener J.S. 1965. Seasonal acclimatization in varying hare (*Lepus*
771 *americanus*). Can. J. Zool. 43:731–744.
772 Hewson R. 1958. Moults and winter whitening in the mountain hare *Lepus timidus scoticus*,
773 Hilzheimer. Proc. Zool. Soc. London. 131:99–108.
774 Hibbard C. 1963. The Origin of the P 3 Pattern of *Sylvilagus*, *Caprolagus*, *Oryctolagus* and
775 *Lepus*. J. Mammal. 44:1–15.
776 Hoberg E.P., Galbreath K.E., Cook J.A., Kutz S.J., Polley L. 2012. Northern Host-Parasite
777 Assemblages. History and Biogeography on the Borderlands of Episodic Climate and
778 Environmental Transition. Advances in Parasitology. Academic Press. p. 1–97.
779 Hopkins D.M. 1959. Cenozoic history of the Bering land bridge. Science. 129:1519–1528.
780 Huson D.H., Bryant D. 2006. Application of phylogenetic networks in evolutionary studies.
781 Mol. Biol. Evol. 23:254–267.
782 Inoue J., Donoghue P.C.J., Yang Z. 2010. The Impact of the Representation of Fossil
783 Calibrations on Bayesian Estimation of Species Divergence Times. Syst. Biol. 59:74–89.
784 Jamie G.A., Meier J.I. 2020. The Persistence of Polymorphisms across Species Radiations.
785 Trends Ecol. Evol. 35:795–808.
786 Jiao X., Flouri T., Rannala B., Yang A.Z. 2020. The Impact of Cross-Species Gene Flow on
787 Species Tree Estimation. Syst. Biol. 0:1–18.
788 Jones M.R., Mills L.S., Alves P.C., Callahan C.M., Alves J.M., Lafferty D.J.R., Jiggins F.M.,
789 Jensen J.D., Melo-Ferreira J., Good J.M. 2018. Adaptive introgression underlies
790 polymorphic seasonal camouflage in snowshoe hares. Science. 360:1355–1358.
791 Jones M.R., Mills L.S., Jensen J.D., Good J.M. 2020a. The origin and spread of locally
792 adaptive seasonal camouflage in snowshoe hares. Am. Nat. 196:316–332.
793 Jones M.R., Mills L.S., Jensen J.D., Good J.M. 2020b. Convergent evolution of seasonal

794 camouflage in response to reduced snow cover across the snowshoe hare range.
795 Evolution. 74:2033–2045.
796 Kinoshita G., Nunome M., Kryukov A.P., Kartavtseva I. V., Han S.-H., Yamada F., Suzuki
797 H. 2019. Contrasting phylogeographic histories between the continent and islands of
798 East Asia: Massive mitochondrial introgression and long-term isolation of hares
799 (Lagomorpha: *Lepus*). Mol. Phylogenet. Evol. 136:65–75.
800 Kobert K., Salichos L., Rokas A., Stamatakis A. 2016. Computing the Internode Certainty
801 and Related Measures from Partial Gene Trees. Mol. Biol. Evol. 33:1606–1617.
802 Kohl S., Coppieters F., Meire F., Schaich S., Roosing S., Brennenstuhl C., Bolz S., van
803 Genderen M.M., Riemslag F.C.C., Lukowski R., den Hollander A.I., Cremers F.P.M.,
804 De Baere E., Hoyng C.B., Wissinger B. 2012. A Nonsense Mutation in *PDE6H* Causes
805 Autosomal-Recessive Incomplete Achromatopsia. Am. J. Hum. Genet. 91:527–532.
806 Kozak K.M., McMillan O., Joron M., Jiggins C.D. 2018. Genome-wide admixture is common
807 across the *Heliconius* radiation. bioRxiv. 414201.
808 Lado S., Alves P.C., Islam M.Z., Brito J.C., Melo-Ferreira J. 2019. The evolutionary history
809 of the Cape hare (*Lepus capensis sensu lato*): insights for systematics and biogeography.
810 Heredity. 123:634–646.
811 Lamichhaney S., Berglund J., Almén M.S., Maqbool K., Grabherr M., Martinez-Barrio A.,
812 Promerová M., Rubin C.-J., Wang C., Zamani N., Grant B.R., Grant P.R., Webster M.T.,
813 Andersson L. 2015. Evolution of Darwin's finches and their beaks revealed by genome
814 sequencing. Nature. 518:371–375.
815 Landis M.J., Matzke N.J., Moore B.R., Huelsenbeck J.P. 2013. Bayesian analysis of
816 biogeography when the number of areas is large. Syst. Biol. 62:789–804.
817 Leaché A.D., Harris R.B., Rannala B., Yang Z. 2014. The Influence of Gene Flow on Species
818 Tree Estimation: A Simulation Study. Syst. Biol. 63:17–30.

819 Levänen R., Thulin C.-G., Spong G., Pohjoismäki J.L.O. 2018. Widespread introgression of
820 mountain hare genes into Fennoscandian brown hare populations. PLoS One.
821 13:e0191790.
822 Li G., Figueiró H. V, Eizirik E., Murphy W.J., Yoder A. 2019. Recombination-Aware
823 Phylogenomics Reveals the Structured Genomic Landscape of Hybridizing Cat Species.
824 Mol. Biol. Evol. 36:2111–2126.
825 Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping
826 and population genetical parameter estimation from sequencing data. Bioinformatics.
827 27:2987–2993.
828 Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-
829 MEM. arXiv Prepr. arXiv:1303.
830 Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G.,
831 Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics.
832 25:2078–9.
833 Liu J., Yu L., Arnold M.L., Wu C.-H., Wu S.-F., Lu X., Zhang Y.-P. 2011. Reticulate
834 evolution: frequent introgressive hybridization among chinese hares (genus *Lepus*)
835 revealed by analyses of multiple mitochondrial and nuclear DNA loci. BMC Evol. Biol.
836 11:223.
837 Liu K.J., Steinberg E., Yozzo A., Song Y., Kohn M.H., Nakhleh L. 2015. Interspecific
838 introgressive origin of genomic diversity in the house mouse. 112:196–201.
839 Long C., Kubatko L. 2018. The Effect of Gene Flow on Coalescent-based Species-Tree
840 Inference. Syst. Biol. 67:770–785.
841 Lopez-Martinez N. 2008. The lagomorph fossil record and the origin of the European rabbit.
842 In: Alves P.C., Ferrand N., Häcklander K., editors. Lagomorph Biology: Evolution,
843 Ecology and Conservation. Verlag Berlin Heidelberg: Springer. p. 27–46.

844 Malinsky M., Svardal H., Tyers A.M., Miska E.A., Genner M.J., Turner G.F., Durbin R.
845 2018. Whole-genome sequences of Malawi cichlids reveal multiple radiations
846 interconnected by gene flow. Nat. Ecol. Evol. 2:1940–1955.
847 Marques D.A., Meier J.I., Seehausen O. 2019. A Combinatorial View on Speciation and
848 Adaptive Radiation. Trends Ecol. Evol. 34:531–544.
849 Marques J.P., Seixas F.A., Farelo L., Callahan C.M., Good J.M., Montgomery W.I., Reid N.,
850 Alves P.C., Boursot P., Melo-Ferreira J. 2020. An Annotated Draft Genome of the
851 Mountain Hare (*Lepus timidus*). Genome Biol. Evol. 12:3656–3662.
852 Martin S.H., Davey J.W., Jiggins C.D. 2015. Evaluating the use of ABBA-BABA statistics to
853 locate introgressed loci. Mol. Biol. Evol. 32:244–257.
854 Matthee C., Van Vuuren B., Bell D., Robinson T. 2004. A Molecular Supermatrix of the
855 Rabbits and Hares (Leporidae) Allows for the Identification of Five Intercontinental
856 Exchanges During the Miocene. Syst. Biol. 53:433–447.
857 Matzke N.J. 2013. Probabilistic historical biogeography: new models for founder-event
858 speciation, imperfect detection, and fossils allow improved accuracy and model-testing.
859 Front. Biogeogr. 5.
860 Matzke N.J. 2014a. BioGeoBEARS: BioGeography with Bayesian (and likelihood)
861 Evolutionary Analysis with R Scripts. Version 1.1.2. .
862 Matzke N.J. 2014b. Model selection in historical biogeography reveals that founder-event
863 speciation is a crucial process in island clades. Syst. Biol. 63:951–970.
864 Mayr E. 1963. Animal species and evolution. Eugen. Rev. 55:226–228.
865 Meier J.I., Marques D.A., Mwaiko S., Wagner C.E., Excoffier L., Seehausen O. 2017.
866 Ancient hybridization fuels rapid cichlid fish adaptive radiations. Nat. Commun.
867 8:14363.
868 Melo-Ferreira J., Alves P.C. 2018. Systematics of lagomorphs. In: Smith A., Alves P.C.,

869 Häcklander K., editors. Lagomorphs: pikas, rabbits, and hares of the World. John
870 Hopkins University Press. p. 9–12.
871 Melo-Ferreira J., Boursot P., Carneiro M., Esteves P.J., Farelo L., Alves P.C. 2012. Recurrent
872 introgression of mitochondrial DNA among hares (*Lepus* spp.) revealed by species-tree
873 inference and coalescent simulations. Syst. Biol. 61:367–381.
874 Melo-Ferreira J., Seixas F.A., Cheng E., Mills L.S., Alves P.C. 2014. The hidden history of
875 the snowshoe hare, *Lepus americanus*: extensive mitochondrial DNA introgression
876 inferred from multilocus genetic variation. Mol. Ecol.:4617–4630.
877 Meyer M., Kircher M. 2010. Illumina Sequencing Library Preparation for Highly Multiplexed
878 Target Capture and Sequencing. Cold Spring Harb. Protoc. 2010.
879 Mills L.S., Bragina E., Kumar A. V., Zimova M., Feltner J., Davis B., Lafferty D.,
880 Häcklander K., Alves P.C., Good J.M., Melo-Ferreira J., Abramov A., Lopatina N., Fay
881 K. 2018. Winter coat color polymorphisms identify global hotspots for evolutionary
882 rescue from climate change. Science. 359:1033–1036.
883 Mills L.S., Zimova M., Oyler J., Running S., Abatzoglou J.T., Lukacs P.M. 2013.
884 Camouflage mismatch in seasonal coat color due to decreased snow duration. Proc. Natl.
885 Acad. Sci. U. S. A. 110:7360–7365.
886 Molloy E.K., Warnow T. 2018. To Include or Not to Include: The Impact of Gene Filtering
887 on Species Tree Estimation Methods. Syst. Biol. 67:285–303.
888 Osborne C.P., Beerling D.J. 2006. Nature's green revolution: the remarkable evolutionary rise
889 of C 4 plants. Philos. Trans. R. Soc. B Biol. Sci. 361:173–194.
890 Page A.J., Taylor B., Delaney A.J., Soares J., Seemann T., Keane J.A., Harris S.R. 2016.
891 SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. Microb.
892 genomics. 2:e000056.
893 Le Pape E., Passeron T., Giubellino A., Valencia J.C., Wolber R., Hearing V.J. 2009.

894 Microarray analysis sheds light on the dedifferentiating role of agouti signal protein in
895 murine melanocytes via the Mc1r. Proc. Natl. Acad. Sci. U. S. A. 106:1802–1807.
896 Paradis E., Claude J., Strimmer K. 2004. APE: Analyses of Phylogenetics and Evolution in R
897 language. Bioinformatics. 20:289–290.
898 Pease J.B., Hahn M.W. 2015. Detection and Polarization of Introgression in a Five-Taxon
899 Phylogeny. Syst. Biol. 64:651–662.
900 Pickrell J.K., Pritchard J.K. 2012. Inference of Population Splits and Mixtures from Genome-
901 Wide Allele Frequency Data. PLoS Genet. 8:e1002967.
902 Pyörnilä A., Putaala A., Hissa R., Sulkava S. 2008. Adaptations to environment in the
903 mountain hare (*Lepus timidus*): thermal physiology and histochemical properties of
904 locomotory muscles. Can. J. Zool. 70:1325–1330.
905 Quinlan A.R., Hall I.M. 2010. BEDTools: a flexible suite of utilities for comparing genomic
906 features. Bioinformatics. 26:841–842.
907 Rajakumari S., Wu J., Ishibashi J., Lim H.-W., Giang A.-H., Won K.-J., Reed R.R., Seale P.
908 2013. EBF2 Determines and Maintains Brown Adipocyte Identity. Cell Metab. 17:562–
909 574.
910 Raudvere U., Kolberg L., Kuzmin I., Arak T., Adler P., Peterson H., Vilo J. 2019. g:Profiler:
911 a web server for functional enrichment analysis and conversions of gene lists (2019
912 update). Nucleic Acids Res. 47:191–198.
913 Ree R.H., Smith S.A. 2008. Maximum Likelihood Inference of Geographic Range Evolution
914 by Dispersal, Local Extinction, and Cladogenesis. Syst. Biol. 57:4–14.
915 Reich D., Thangaraj K., Patterson N., Price A.L., Singh L. 2009. Reconstructing Indian
916 population history. Nature. 461:489–494.
917 dos Reis M., Yang Z. 2011. Approximate Likelihood Calculation on a Phylogeny for
918 Bayesian Estimation of Divergence Times. Mol. Biol. Evol. 28:2161–2172.

dos Reis M., Yang Z. 2019. Bayesian Molecular Clock Dating Using Genome-Scale Datasets. In: Anisimova M., editor. Evolutionary Genomics. New York, NY: Springer New York. p. 309–330.

Robinson T.J., Yang F., Harrison W.R. 2002. Chromosome painting refines the history of genome evolution in hares and rabbits (order Lagomorpha). Cytogenet. Genome Res. 96:223–227.

Rogowitz G.L. 1990. Seasonal Energetics of the White-Tailed Jackrabbit (*Lepus townsendii*). J. Mammal. 71:277–285.

Ronquist F. 1997. Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. Syst. Biol. 46:195–203.

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

Salichos L., Stamatakis A., Rokas A. 2014. Novel information theory-based measures for quantifying incongruence among phylogenetic trees. Mol. Biol. Evol. 31:1261–1271.

Sambrook E., Fritsch F., Maniatis T. 1989. Molecular cloning. Cold Spring Harbour, NY: Cold Spring Harbour Press.

Sarver B.A.J., Keeble S., Cosart T., Tucker P.K., Dean M.D., Good J.M. 2017. Phylogenomic Insights into Mouse Evolution Using a Pseudoreference Approach. Genome Biol. Evol. 9:726–739.

Sasaki M., Yoshitane H., Du N.-H., Okano T., Fukada Y. 2009. Preferential Inhibition of BMAL2-CLOCK Activity by PER2 Reemphasizes Its Negative Role and a Positive Role of BMAL2 in the Circadian Transcription. J. Biol. Chem. 284:25149–25159.

Sayyari E., Whitfield J.B., Mirarab S. 2018. DiscoVista: Interpretable visualizations of gene tree discordance. Mol. Phylogenet. Evol. 122:110–115.

Schliep K.P. 2011. phangorn: phylogenetic analysis in R. Bioinform. 27:592–593.

944 Schumer M., Xu C., Powell D.L., Durvasula A., Skov L., Holland C., Blazier J.C.,
945 Sankararaman S., Andolfatto P., Rosenthal G.G., Przeworski M. 2018. Natural selection
946 interacts with recombination to shape the evolution of hybrid genomes. Science.
947 360:656–660.
948 Seixas F.A., Boursot P., Melo-Ferreira J. 2018. The genomic impact of historical
949 hybridization with massive mitochondrial DNA introgression. Genome Biol. 19:91.
950 Sheriff M.J., Kuchel L., Boutin S., Humphries M.M. 2009. Seasonal Metabolic
951 Acclimatization in a Northern Population of Free-Ranging Snowshoe Hares, *Lepus*
952 *americanus*. J. Mammal. 90:761–767.
953 Simpson G.G. 1947. Holarctic mammalian faunas and continental relationships during the
954 cenozoic. Bull. Geol. Soc. Am. 58:613–688.
955 Smith A.T., Johnston C.H., Alves P.C., Häcklander K. 2018. Lagomorphs: Pikas, Rabbits and
956 Hares of the World. Baltimore: Johns Hopkins University Press.
957 Smith J., Kronforst M.R. 2013. Do *Heliconius* butterfly species exchange mimicry alleles?
958 Biol. Lett. 9:20130503–20130503.
959 Solís-Lemus C., Bastide P., Ané C. 2017. PhyloNetworks: A Package for Phylogenetic
960 Networks. Mol. Biol. Evol. 34:3292–3298.
961 Solís-Lemus C., Yang M., Ané C. 2016. Inconsistency of Species Tree Methods under Gene
962 Flow. Syst. Biol. 65:843–851.
963 Stamatakis A. 2018. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
964 large phylogenies. Bioinformatics. 30:1312–1313.
965 Stokkan K.-A., Folkow L., Dukes J., Neveu M., Hogg C., Siefken S., Dakin S.C., Jeffery G.
966 2013. Shifting mirrors: adaptive changes in retinal reflections to winter darkness in
967 Arctic reindeer. Proc. R. Soc. B Biol. Sci. 280:20132451.
968 Svardal H., Quah F.X., Malinsky M., Ngatunga B.P., Miska E.A., Salzburger W., Genner

969 M.J., Turner G.F., Durbin R. 2020. Ancestral Hybridization Facilitated Species
970 Diversification in the Lake Malawi Cichlid Fish Adaptive Radiation. Mol. Biol. Evol.
971 37:1100–1113.
972 Swofford D.L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (* and Other
973 Methods). .
974 Taylor S.A., Larson E.L. 2019. Insights from genomes into the evolutionary importance and
975 prevalence of hybridization in nature. Nat. Ecol. Evol. 3:170–177.
976 Than C., Degnan J.H., Nakhleh L. 2011. Coalescent Histories on Phylogenetic Networks and
977 Detection of Hybridization Despite Incomplete Lineage Sorting. Syst. Biol. 60:138–149.
978 Tolesa Z., Bekele E., Tesfaye K., Ben Slimen H., Valqui J., Getahun A., Hartl G.B.,
979 Suchentrunk F. 2017. Mitochondrial and nuclear DNA reveals reticulate evolution in
980 hares (*Lepus* spp., Lagomorpha, Mammalia) from Ethiopia. PLoS One. 12:e0180137.
981 Vanderpool D., Minh B.Q., Lanfear R., Hughes D., Murali S., Harris R.A., Raveendran M.,
982 Muzny D.M., Gibbs R.A., Worley K.C., Rogers J., Hahn M.W. 2020. Primate
983 phylogenomics uncovers multiple rapid radiatio. bioRxiv. 043786.
984 Wang R.J., Hahn M.W. 2018. Speciation genes are more likely to have discordant gene trees.
985 Evol. Lett. 2:281–296.
986 Wen D., Yu Y., Hahn M.W., Nakhleh L. 2016. Reticulate evolutionary history and extensive
987 introgression in mosquito species revealed by phylogenetic network analysis. Mol. Ecol.
988 25:2361–2372.
989 White J.A. 1991. North American Leporinae (Mammalia: Lagomorpha) from late Miocene
990 (Clarendonian) to latest Pliocene (Blancan). J. Vertebr. Paleontol. 11:67–89.
991 Yamada F., Takaki M., Suzuki H. 2002. Molecular phylogeny of Japanese Leporidae, the
992 Amami rabbit *Pentalagus furnessi*, the Japanese hare *Lepus brachyurus*, and the
993 mountain hare *Lepus timidus*, inferred from mitochondrial DNA sequences. Genes

994 Genet. Syst. 77:107–16.
995 Yang Z. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. Mol. Biol. Evol.
996 24:1586–1591.
997 Yu Y., Degnan J.H., Nakhleh L. 2012. The Probability of a Gene Tree Topology within a
998 Phylogenetic Network with Applications to Hybridization Detection. PLoS Genet.
999 8:1002660.
1000 Yu Y., Dong J., Liu K.J., Nakhleh L. 2014. Maximum likelihood inference of reticulate
1001 evolutionary histories. Proc. Natl. Acad. Sci. 111:16448–16453.
1002 Yu Y., Nakhleh L. 2015. A maximum pseudo-likelihood approach for phylogenetic networks.
1003 BMC Genomics. 16:S10.
1004 Zhang C., Rabiee M., Sayyari E., Mirarab S. 2018. ASTRAL-III: Polynomial time species
1005 tree reconstruction from partially resolved gene trees. BMC Bioinformatics. 19:153.
1006 Zimova M., Hackländer K., Good J.M., Melo-Ferreira J., Alves P.C., Mills L.S. 2018.
1007 Function and underlying mechanisms of seasonal colour moulting in mammals and
1008 birds: what keeps them changing in a warming world? Biol. Rev. 93:1478–1498.
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018

1019 **FIGURE LEGENDS**

1020

1021 **Fig. 1 - Hare (*Lepus* spp.) evolutionary history and biogeography. (a)** The distribution of
1022 the 15 hare species studied in this work obtained from the IUCN database
1023 (http://www.iucnredlist.org); **(b)** Coalescent species tree estimated with *ASTRAL* assigning
1024 individuals to species groups (n > 1 except for the white-sided jackrabbit); **(c)** Ancestral range
1025 reconstruction under the DIVALIKE+J model implemented in *BioGeoBEARS* on the
1026 divergence time tree estimated with *MCMCtree* and calibrated with deep fossil calibration
1027 points. Pie charts represent the probability of each potential range, and squares represent the
1028 current range of extant taxa. Node labels represent estimated divergence times in millions of
1029 years. Confident intervals, an alternative biogeographic reconstruction based on *Lepus* fossil
1030 calibrations and ancestral ranges for the shoulders of the tree are provided in Supplementary
1031 Material.

1032

1033 **Fig. 2 - The hare (*Lepus* spp.) species tree is underlined by extensive gene tree**
1034 **incongruence.** Split network constructed from 8889 gene trees (5% threshold) shows
1035 discordance among the gene tree topologies (cuboid structures represent alternative
1036 topologies) in deeper nodes of the species tree. Species are marked in accordance with the
1037 continents where they are distributed.

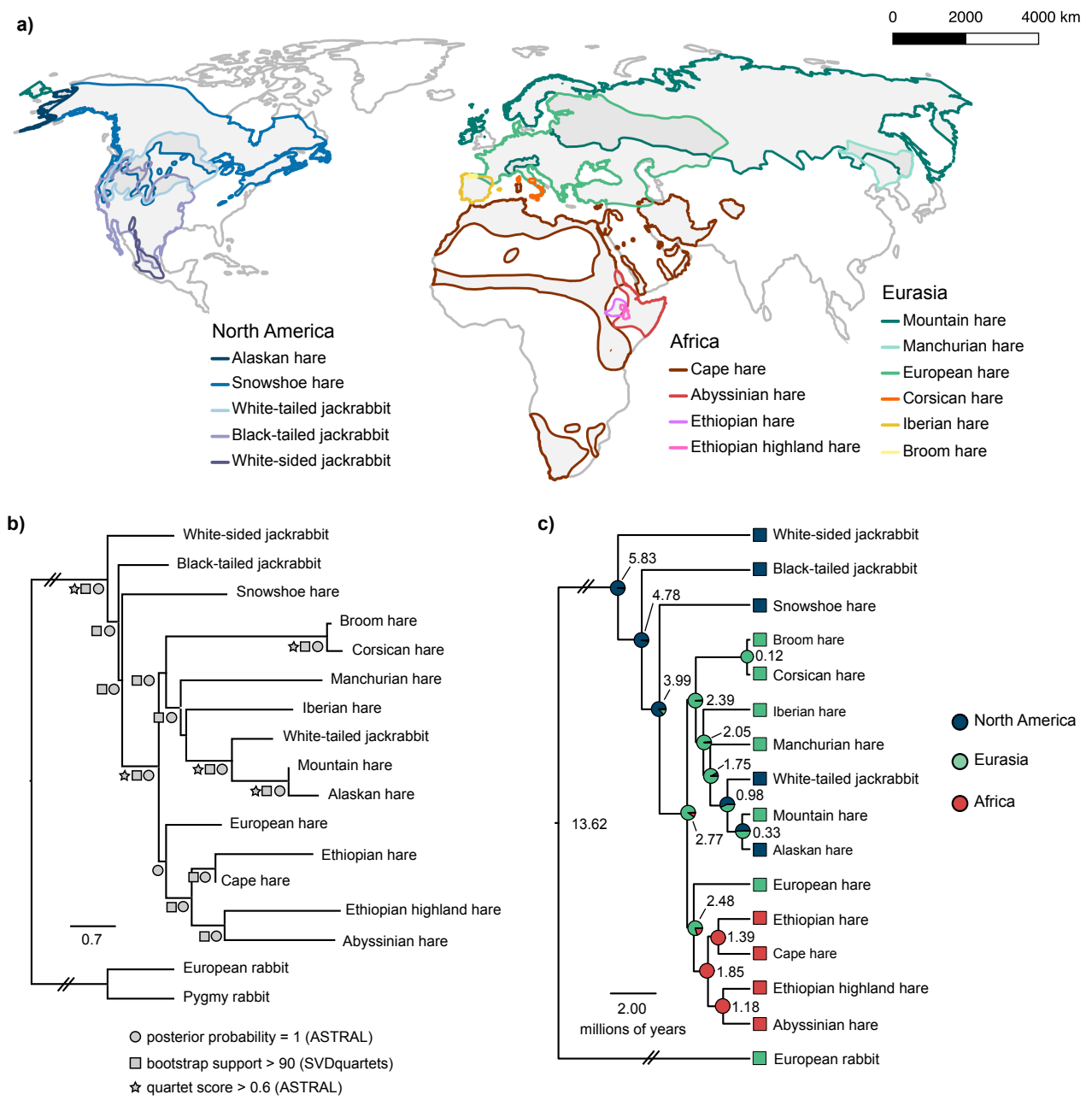
1038

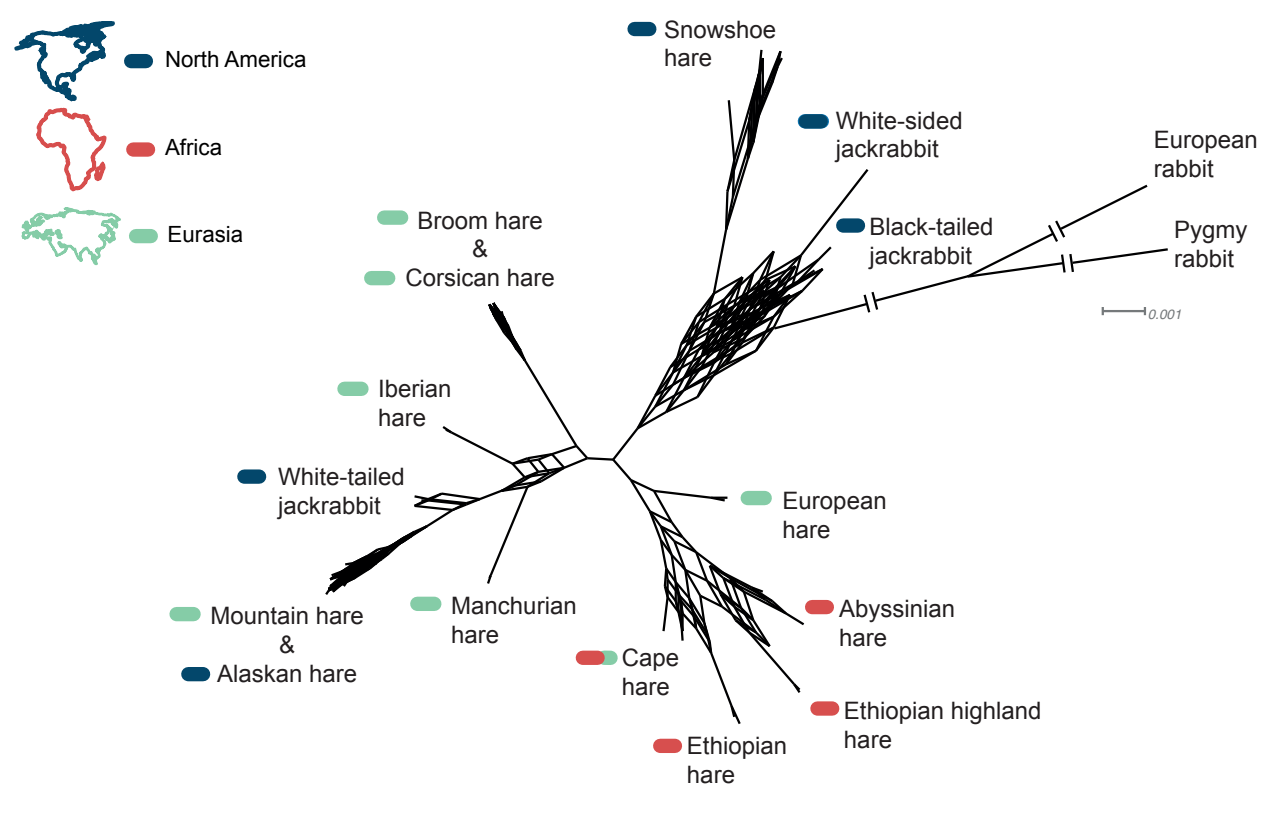
1039 **Fig. 3 – Admixture events are distributed across the hare (*Lepus* spp.) species tree.** The
1040 'f-branch' statistic $f_b(C)$ represents excess allele sharing between branches *b* (y-axis) and *C*
1041 (x-axis) of the species tree in Figure 1b. The gradient represents the $f_b(C)$ score, dark gray
1042 represents tests not consistent with the species tree (for each branch b, having itself or a sister
1043 taxon as donor C) and asterisks denote block jackknifing significance at $P \leq 0.05$ (after

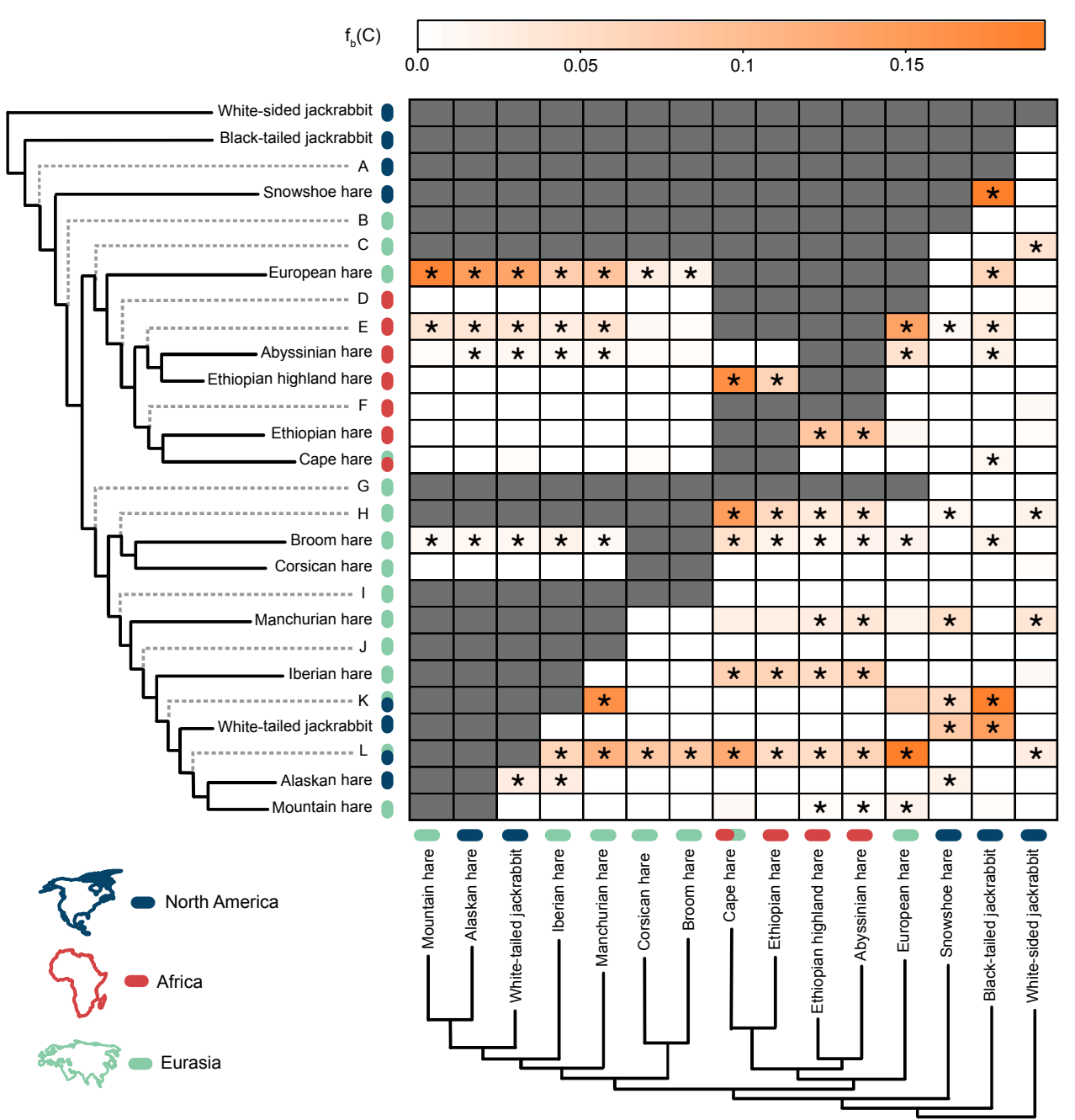
Bonferroni correction). Tips of the tree are marked according to their current distribution, and ancestral tips (dashed lines and labeled with letters) are marked according to the ancestral range reconstruction in Figure 1c (two colors represent the inference of two ranges with equivalent (~0.5) probability). Ancestral tips are labeled from A to L corresponding to labels in Supplementary Table S11 available on Dryad.

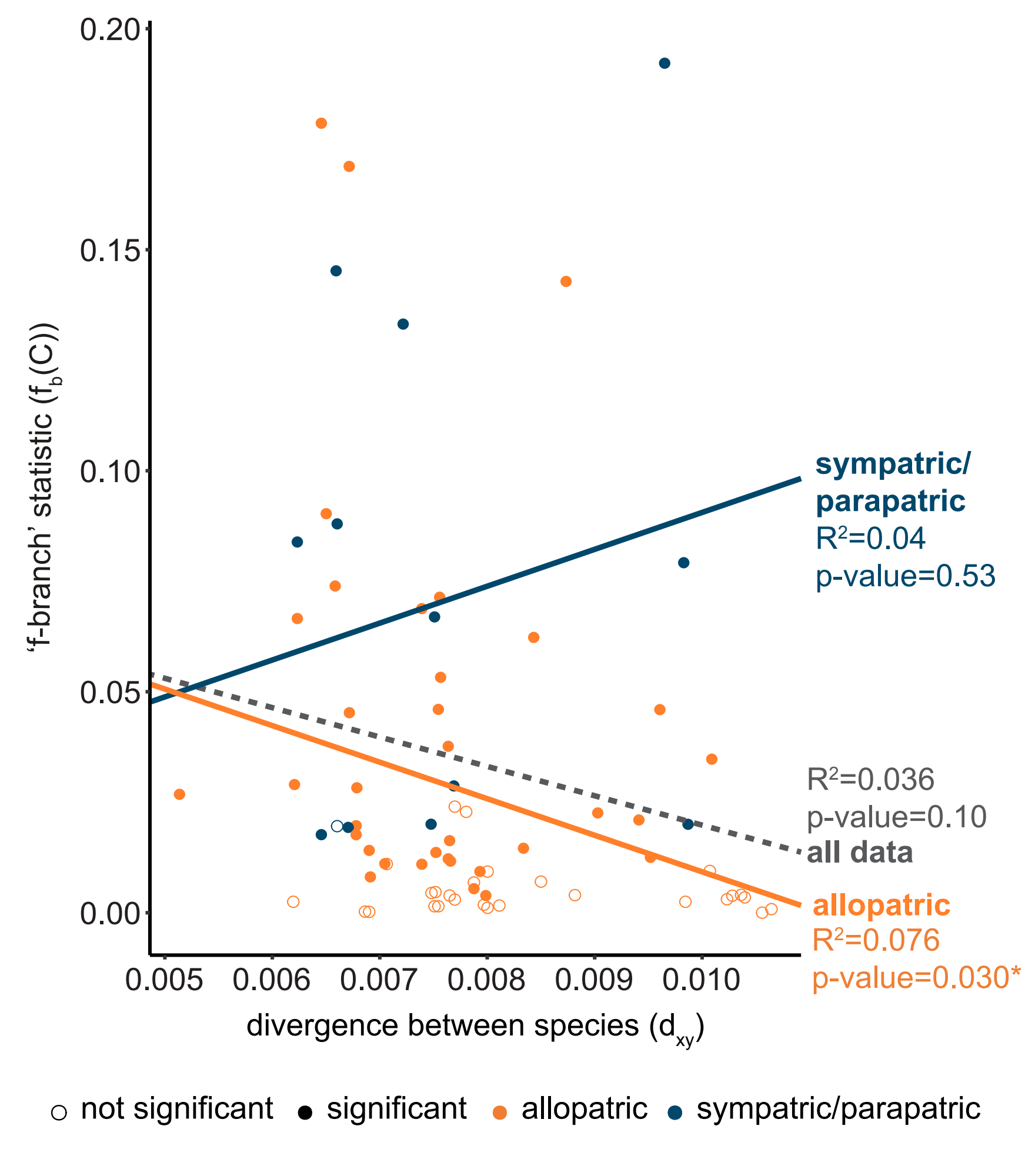
Fig. 4 – Admixture proportions decrease with genetic divergence between allopatric species. We plot the 'f-branch' ($f_b(C)$) values against exome-wide divergence (d_{xy}) for extant species pairs, differentiating species with overlapping (sympatric/parapatric) or non-overlapping (allopatric) distributions. The tendency line represents a linear regression relating d_{xy} and all (significant and non-significant) $f_b(C)$ values calculated with function *lm()* in R.

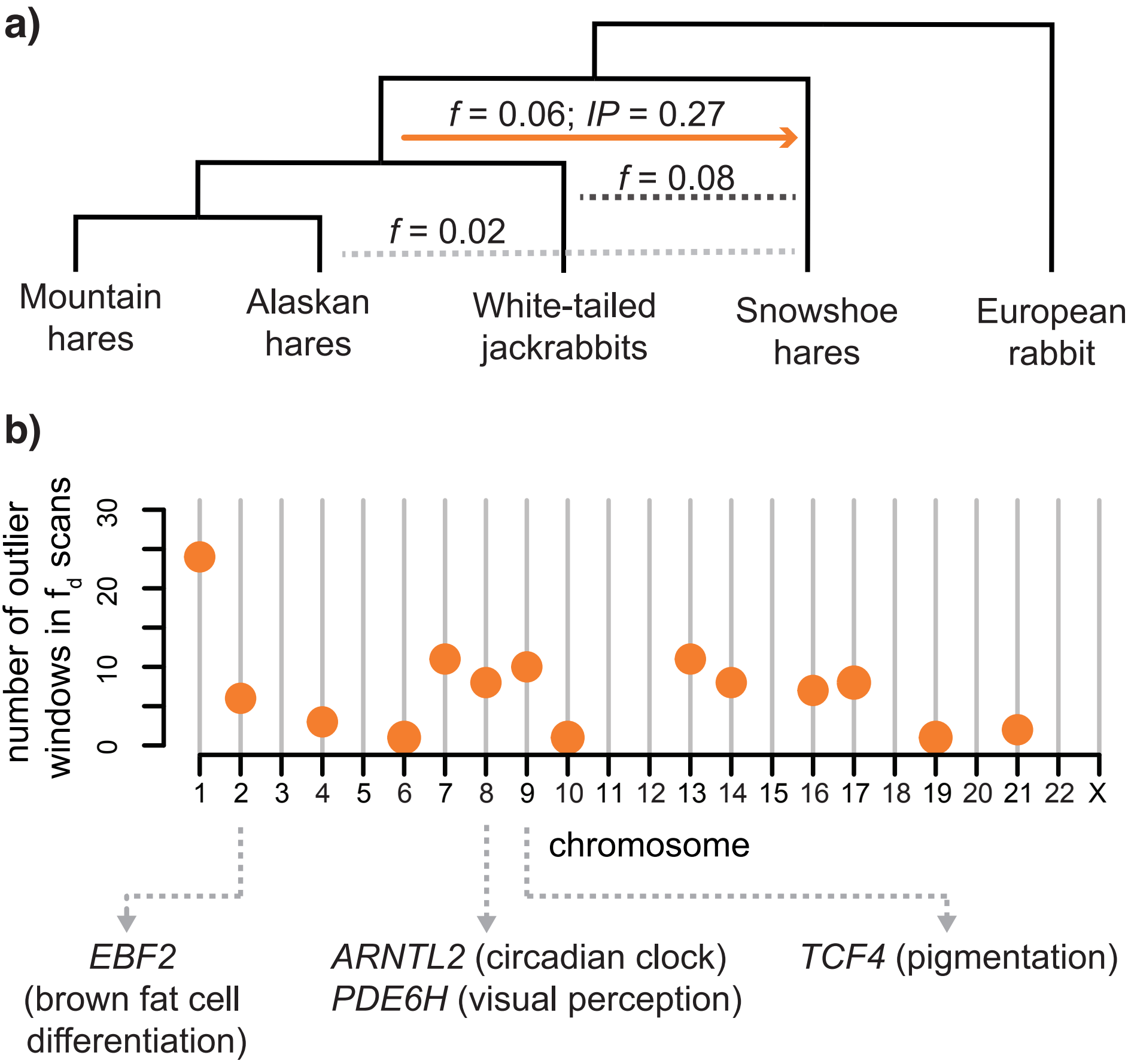
Fig. 5 – The impact of ancestral introgression on extant northern latitude species. (a) Events of past and recent admixture inferred in this study involving snowshoe, Alaskan hares, mountain hares and white-tailed jackrabbits. The arrow indicates the direction of ancestral introgression inferred by *PhyloNet* and 'f-branch' and dashed lines indicate recent introgression inferred with 'f-branch'. Values above the lines represent admixture proportion (f) estimated with 'f-branch' or inheritance probability (IP) estimated with *PhyloNet;* **(b)** Genomic distribution of ancestral blocks of introgression (circles) inferred as shared outlier windows among fraction of admixture (f_d) analysis testing for admixture among snowshoe hares and each one of the three other northern latitude species. Tests were run in 50 kb genomic sliding windows, and outlier windows are in the top 0.5% of the f_d distribution.


a)
0
2000
4000 km
North America
Alaskan hare
Snowshoe hare
White-tailed jackrabbit
Black-tailed jackrabbit
White-sided jackrabbit
Africa
Cape hare
Abyssinian hare
Ethiopian hare
Ethiopian highland hare
Eurasia
Mountain hare
Manchurian hare
European hare
Corsican hare
Iberian hare
Broom hare
b)
White-sided jackrabbit
Black-tailed jackrabbit
Snowshoe hare
Broom hare
Corsican hare
Manchurian hare
Iberian hare
White-tailed jackrabbit
Mountain hare
Alaskan hare
European hare
Ethiopian hare
Cape hare
Ethiopian highland hare
Abyssinian hare
European rabbit
Pygmy rabbit
0.7
posterior probability = 1 (ASTRAL)
bootstrap support > 90 (SVDquartets)
quartet score > 0.6 (ASTRAL)
c)
White-sided jackrabbit
Black-tailed jackrabbit
Snowshoe hare
Broom hare
Corsican hare
Iberian hare
Manchurian hare
White-tailed jackrabbit
Mountain hare
Alaskan hare
European hare
Ethiopian hare
Cape hare
Ethiopian highland hare
Abyssinian hare
European rabbit
5.83
4.78
0.12
3.99
2.39
2.05
1.75
0.98
13.62
2.77
0.33
2.48
1.39
1.85
1.18
2.00
millions of years
North America
Eurasia
Africa


North America
Africa
Eurasia
Snowshoe hare
White-sided jackrabbit
European rabbit
Black-tailed jackrabbit
Pygmy rabbit
0.001
Broom hare & Corsican hare
Iberian hare
White-tailed jackrabbit
European hare
Mountain hare & Alaskan hare
Manchurian hare
Abyssinian hare
Cape hare
Ethiopian highland hare
Ethiopian hare


$f_b(C)$
0.0
0.05
0.1
0.15
White-sided jackrabbit
Black-tailed jackrabbit
A
Snowshoe hare
B
C
European hare
D
E
Abyssinian hare
Ethiopian highland hare
F
Ethiopian hare
Cape hare
G
H
Broom hare
Corsican hare
I
Manchurian hare
J
Iberian hare
K
White-tailed jackrabbit
L
Alaskan hare
Mountain hare
North America
Africa
Eurasia
Mountain hare
Alaskan hare
White-tailed jackrabbit
Iberian hare
Manchurian hare
Corsican hare
Broom hare
Cape hare
Ethiopian hare
Ethiopian highland hare
Abyssinian hare
European hare
Snowshoe hare
Black-tailed jackrabbit
White-sided jackrabbit


0.20
0.15
0.10
0.05
0.00
'f-branch' statistic ($f_b(C)$)
0.005
0.006
0.007
0.008
0.009
0.010
divergence between species (d_{xy})
sympatric/
parapatric
R^2=0.04
p-value=0.53
R^2=0.036
p-value=0.10
all data
allopatric
R^2=0.076
p-value=0.030*
not significant
significant
allopatric
sympatric/parapatric


a)
f = 0.06; IP = 0.27
f = 0.08
f = 0.02
Mountain hares
Alaskan hares
White-tailed jackrabbits
Snowshoe hares
European rabbit
b)
number of outlier windows in f_d scans
0
10
20
30
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X
chromosome
EBF2 (brown fat cell differentiation)
ARNTL2 (circadian clock)
PDE6H (visual perception)
TCF4 (pigmentation)