RESEARCH ARTICLE

Transcriptome analysis provides new insights into cold adaptation of corsac fox (*Vulpes Corsac*)

Xiufeng Yang[1] | Guolei Sun[1] | Tian Xia[1] | Muha Cha[2] | Lei Zhang[1] | Bo Pang[2] |
Qingming Tang[3] | Huashan Dou[2] | Honghai Zhang[1]

[1]College of Life Science, Qufu Normal University, Qufu, China

[2]Hulunbuir Academy of Inland Lakes in Northern Cold & Arid Areas, Hulunbuir, China

[3]Hulun Buir Forestry and Grassland Business Development Center, Hulunbuir, China

Correspondence
Honghai Zhang, College of Life Science, Qufu Normal University, Qufu, China.
Email: zhanghonghai67@126.com

Abstract

Vulpes are widely distributed throughout the world and have undergone drastic physiological and phenotypic changes in response to their environment. However, little is known about the underlying genetic causes of these traits, especially *Vulpes corsac*. In this study, RNA-Seq was used to obtain a comprehensive dataset for multiple pooled tissues of corsac fox, and selection analysis of orthologous genes was performed to identify the genes that may be influenced by the low-temperature environment. More than 6.32 Gb clean reads were obtained and assembled into a total of 173,353 unigenes with an average length of 557 bp for corsac fox. Selective pressure analysis showed that 16 positively selected genes (PSGs) were identified in corsac fox, red fox, and arctic fox. Enrichment analysis of PSGs showed that the *LRP11* gene was enriched in several pathways related to the low-temperature response and might play a key role in response to environmental stimuli of foxes. In addition, several positively selected genes were related to DNA damage repair (*ELP2* and *CHAF1A*), innate immunity (*ARRDC4* and *S100A12*), and the respiratory chain (*NDUFA5*), and these positively selected genes might play a role in adaptation to harsh wild fox environments. The results of common orthologous gene analysis showed that gene flow or convergent evolution might be an important factor in promoting regional differentiation of foxes. Our study provides a valuable transcriptomic resource for the evolutionary history of the corsac fox and the adaptations to the extreme environments.

KEYWORDS
cold adaptation, gene tree discordance, selective pressure analysis, transcriptome, *Vulpes corsac*

TAXONOMY CLASSIFICATION
Genetics

1 | INTRODUCTION

Adaptation is a process in which organisms adapt to their environment to survive and reproduce in terms of morphological structures and physiological functions (Brandon, 1978). In particular, creatures in extreme environments have undergone adaptive evolution in phenotype, physiology, and behavior, and the study of its adaptive mechanism has always been a hot topic in field of evolutionary biology. Numerous studies have explored in depth the mechanisms of adaptation in different species at the genomic, metagenomic, and transcriptome levels, such as sheep (Ji et al., 2016), yak (Lan et al., 2021), goat (Martchenko et al., 2020; Zhang et al., 2021), and dog (Gou et al., 2014; Miao et al., 2017; Wu et al., 2016).

As a typical example of adaptation, the fox distributed widely and the population remains stable around the world (IUCN, 2021, Figure 1). The members of this genus play important roles in maintaining ecosystems, including tundra, forests, deserts, and grasslands (Statham et al., 2014). Accordingly, they have undergone adaptive evolution in physiology, morphology, and other aspects. However, little is currently known about their underlying genetic causes, something we aimed to explore in this study. Research on the molecular ecology of the *Vulpes* has mainly focused on red fox and arctic fox, while corsac fox currently exhibits only in behavior, diet, and zoonoses (Abdybekova et al., 2015; Guilherme et al., 1938; Lkhagvasuren et al., 2016). In the study of adaptive evolution, genome-wide extension of unique gene families and positively selected genes potentially play roles in the adaptation of extreme environments in arctic fox (Peng et al., 2021). Transcriptome analysis of arctics and the red fox has shown that fatty acid metabolism genes are under positive selection in the arctic fox and seem generally to be crucial for mammalian survival under arctic conditions (Kumar et al., 2015). In our previous study, several positively selected genes related to natural immunity (*CFI* and *LRRFIP1*), protein synthesis (*GOLGA4*, *CEP19*, and *SLC35A2*), and DNA damage repair (*MDC1*) were screened in farmed fox, which might play a role in adaptation to breeding environments (Yang et al., 2021).

FIGURE 1 Corsac fox in HulunBuir grassland, Inner Mongolia

Transcriptome analysis is an economic strategy of functional genomic studies that enables the collection of gene sequences for nonmodel animals. These advances greatly facilitate comparative evolutionary studies (Emrich et al., 2007; Garber et al., 2011; Zhang et al., 2010). Despite not representing an entire genome, this transcriptome analysis identified numerous genes that are relevant to adaptation in species. Transcriptome analysis has been used for the assembly and annotation of transcriptomes in many Canidae, including wolves (Liu et al., 2017, 2019), dogs (Yang et al., 2018), foxes (Kukekova et al., 2011; Kumar et al., 2015; Yang et al., 2021), and Raccoon dogs (Du et al., 2017). These studies have identified many genes involved in adaptive evolution and have contributed to the development of extensive genomic and transcriptomic resources for Canidae. To date, few studies have investigated wild environment adaptation in corsac fox, and the molecular mechanisms have not yet been illuminated. In this study, the pooled transcriptome of multiple tissues of corsac fox was sequenced and compared with red fox and arctic fox to explore the molecular mechanism and corresponding genes of their adaptation to cold environments. The findings provide insights into the molecular mechanisms of adaptation to extreme environments in *Vulpes*.

2 | MATERIALS AND METHODS

2.1 | Sample collection

Kidney, brain, and liver tissue of one *Vulpes corsac* (corsac fox, named CF) was collected from Hulun Lake National Nature Reserve, Inner Mongolia, China. The tissue sample was stored in RNAlater RNA stabilization reagent (QIAGEN, USA) and frozen at −80℃ in ultradeepfreeze equipment. The data of two *vulpes lagopus* (arctic fox, named AF1 and AF2) and one *vulpes vulpes* (red fox, named RF) were downloaded from the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database with accession numbers ERR687853, ERR687854, and ERR687855, respectively, and this sample was collected from Kronoby, Finland (Kumar et al., 2015).

2.2 | RNA extraction, library construction, and sequencing

Total RNA of CF was extracted using RNeasy Mini Kit (QIAGEN, USA) and checked for purity using 1% agarose gels and a NanoPhotometer® spectrophotometer (IMPLEN, USA). The integrity and concentration of RNA were measured and assessed using the RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, USA) and the Qubit® RNA Assay Kit (Invitrogen, USA) in a Qubit® 2.0 Flurometer (Life Technologies, USA), respectively. Total RNA (at least 3 mg per tissue) was used to construct sequencing libraries with the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA). Once the library was built, an Agilent Bioanalyzer 2100 system was used to assess the

library quality and then sequenced on an Illumina HiSeq 2500 platform.

2.3 | Transcriptome assembly, gene functional annotation, and CDS identification

Clean reads were obtained from raw reads by removing reads that contained adapter or poly-N and low-quality reads (Q-value ≤ 20) (Cock et al., 2010). Transcriptome assembly was carried out with Trinity (Grabherr et al., 2011) to obtain reference transcriptome for subsequent analysis. Annotation of gene functional included pathway annotation, protein functions annotation, Gene Ontology (GO) annotation, and COG/KOG functional annotation. The website and parameters of those databases are listed in Table S1. The NR and Swiss-Prot databases were used to blast the CDS of each putative unigene with a cut-off of 1e-5, and Estscan software was used to determine the unigene that did not have aligned results (Iseli et al., 1999).

2.4 | Identification of gene orthologous groups and calculation of Ka/Ks

OrthoMCL (Li et al., 2003) software was used to identify orthologous groups with *Ailuropoda melanoleuca* (aml) and *Canis lupus familiaris* (cfa) as the outside and internal reference genomes, respectively. Sequence and optimized protein alignment results were constructed by Muscle 3.8.3313 and Gblocks_0.91b (Edgar, 2004), respectively. The online software KOBAS (http://kobas.cbi.pku.edu.cn/), DAVID (https://david.ncifcrf.gov/tools.jsp), and the R package ClusterProfiler (Yu et al., 2012) were used for GO and KEGG enrichment analysis. The results of sequence alignment were used to construct evolutionary trees by Phyml (version: 20131016) (Guindon et al., 2005).

The ratio (ω) of the number of nonsynonymous substitutions per nonsynonymous site (Ka) to the number of synonymous substitutions per synonymous site (Ks) could be used as an indicator of selective pressure acting on a protein-coding gene in genetics. When $\omega > 1$ was usually said to be evolving under positive selection, $\omega < 1$ was usually said to be under purifying selection of homologous

genes. In this study, the value of ω was calculated for the pairwise alignments by the Paml-codeml 4.7 (Yang, 2007) package with default settings.

2.5 | Homologous gene analysis of three species of fox

The orthologous genes obtained in this study were compared with those screened in previous studies on silver fox and blue fox to acquire the orthologous genes shared by five samples (corsac fox, red fox, arctic fox, blue fox, and silver fox) (Yang et al., 2021). Then, phylogenetic tree was constructed based on orthologous genes and 12 protein-coding genes (except *ND6*) of mitochondria downloaded from the NCBI database by three methods (PAUP (Swofford, 1998), Bayes (Fredrik & Huelsenbeck, 2003) and RAxML (Alexandros, 2014)), and the reasons for the differences were discussed based on previous studies (Degnan & Rosenberg, 2009; Nakhleh, 2013).

3 | RESULTS

3.1 | Overview of transcriptome sequencing data and *de novo* assembly

One cDNA library was constructed from pooled RNA extracts from different tissues (liver, brain, and kidney) of *Vulpes corsac* (CF). The summary statistics of the RNA-Seq data are shown in Table 1. After filtering the raw data, a total of 73.37 million clean reads were obtained with an average read length of 150 bp base pairs for further analysis. The Q30% and Q20% of the CF sample were 94.13% and 97.69%, respectively. A total of 234.9 million clean reads remained from RF, AF1, and AF2, which were downloaded from the NCBI database.

3.2 | Function annotation and CDS identification

After assembly, we obtained 555,949 transcripts with a mean length range from 738 to 836 bp and 463,335 unigenes with a

TABLE 1 Summary of sequencing results

Sample	Raw reads	Clean reads	Clean bases (Gb)	Error (%)	Q20 (%)	Q30 (%)	GC (%)
CF	76,627,518	73,370,220	11.01	0.01	97.69	94.13	50.99
RF[a]	87,641,852	81,623,324	7.67	0.08	93.82	84.17	50.21
AF1[a]	72,243,730	67,227,516	6.32	0.09	93.45	83.75	50.39
AF2[a]	92,135,172	86,128,572	8.1	0.08	93.95	84.33	49.84

Abbreviations: AF, arctic fox; CF, corsac fox; RF, red fox.

[a]The data which download from NCBI database. Error: sequencing error rate. Q20/Q30: Percentage of bases with a Phred value of at least 20/30. GC: The content of G and C.

mean length range from 557 to 690 bp for the four samples. The N50 values for transcripts and unigenes ranged from 1497 to 1765 and 758 to 1352, respectively. Most of the transcripts and unigenes were distributed within the range of 200–500 bp, which comprised 60.3%–69.3% of the transcripts and 66.3%–76.6% of the unigenes of the whole dataset. For the rest, 8,022–12,876 unigenes had lengths of 1–2 kbp and 4984–10,548 were within >2 kbp. The length distributions of the transcripts and unigenes are shown in Table 2. This large number of reads with paired-end information produced much longer unigenes (mean: 618 bp) than those in previous studies (Novaes et al., 2008; Parchman et al., 2010).

To better understand the functional characterization of all assembled unigenes, BLASTX and BLASTN were employed to against five public databases (Nr, Nt, KO, Swiss-Prot, PFAM, GO, and KOG). A total of 92,697 (40.94%), 54,455 (68.15%), 52,204 (73.33%), and 54,888 (63.94%) unigenes were annotated in at least one database in CF, RF, AF1, and AF1, respectively (Table 3 and Figure 2). Specifically, the Nt database contributed the largest number of matches (CF: 30,901, RF: 27,843, AF1: 28,043, and AF2: 27,298). Furthermore, 5926 (2.62%), 5365 (6.71%), 5328 (7.48%), and 5803 (6.76%) annotated in all databases in CF, RF, AF1, and AF1, respectively.

A total of 265,355 (CF: 112,481, RF: 51,661, AF1: 47,323, and AF2: 53,890) CDSs were identified by BLASTX and prediction of the transcripts, and most of them were predicted by Estscan (72.6%–83.6%). To avoid spurious hits, only CDSs longer than 300 nucleotides, CF: 34,974, RF: 23,312, AF1: 22,486, and AF2: 23,490, were annotated for subsequent analysis (Table 4).

3.3 | Identification of gene orthologous groups

A total of 21,985 homologous genes were identified in the three species. After filtering, only 393 orthologous genes were selected for the following analysis. Two online software programs (DAVID and KOBAS) and an R package ClusterProfiler were used for enrichment analysis of orthologous genes. Different enrichment methods have been applied in research on the function of orthologous genes. Ultimately, 393 single-copy genes were annotated in 355 (ClusterProfiler: 149, DAVID: 17, KOBAS: 189, p-value > .05) GO terms and 23 (ClusterProfiler: 13, DAVID: 3, KOBAS: 7, p-value > .05, Figure 3) KEGG pathways.

Of these, several protein-related terms were identified, such as negative regulation of protein catabolic process (GO:0042177, p = .0006), negative regulation of cellular protein catabolic process (GO:1903363, p = .0006), protein catabolic process (GO:0030163, p = .004), fat digestion and absorption (cfa04975, p = .026). Three genes, _AQP3_, _EGR1_, and _FABP1_, were annotated with pathways associated with hypoxic response (cellular response to decreased oxygen levels, GO:0036294, p = .044 and cellular response to hypoxia, GO:0071456, p = .044). In addition, a large number of genes were also enriched in immune-related pathways.

**TABLE 2** Length interval and distribution of transcripts and unigenes

Length interval & distribution	Transcripts				Unigenes			
	CF	RF	AF1	AF2	CF	RF	AF1	AF2
200–500 bp	184,053	59,822	52,954	64,584	173,353	53,894	47,236	58,544
500–1 kbp	37,021	16,303	15,419	16,436	29,632	12,367	11,671	12,416
1k-2 kbp	21,320	12,073	11,054	12,436	12,876	8,022	7,296	8,120
>2 kbp	23,138	9,391	8,347	11,598	10,548	5,617	4,984	6,759
Total Number	265,532	97,589	87,774	105,054	226,409	79,900	71,187	85,839
Mean length	738	786	836	787	557	667	672	690
N50	1,596	1,532	1,765	1,497	758	1,196	1,185	1,352
N90	262	277	280	281	243	252	255	251
Total Nucleotides	195,990,946	76,748,930	87,804,175	69,105,503	126,079,486	53,321,702	47,842,732	59,207,490

Note: N50/N90: The transcript obtained by splicing was arranged from long to short and then accumulated. When the cumulative length>=50%/90% of the total length, then the transcript length is considered N50/N90.

TABLE 3 Statistics of blast results for Unigene against databases

Database	CF	RF	AF1	AF2
Annotated in Nr	30,901 (13.65%)	27,843 (34.85%)	28,043 (39.39%)	27,298 (31.80%)
Annotated in Nt	85,294 (37.67%)	53,036 (66.38%)	51,205 (71.93%)	53,220 (62.00%)
Annotated in KO	16,636 (7.35%)	15,288 (19.13%)	15,348 (21.56%)	16,116 (18.77%)
Annotated in Swiss-Prot	26,122 (11.54%)	25,661 (32.12%)	26,000 (36.52%)	24,866 (28.97%)
Annotated in PFAM	27,231 (12.03%)	18,866 (23.61%)	18,210 (25.58%)	19,069 (22.21%)
Annotated in GO	27,450 (12.12%)	19,047 (23.84%)	18,377 (25.82%)	19,246 (22.42%)
Annotated in KOG	9,040 (3.99%)	9,303 (11.64%)	9,330 (13.11%)	9,014 (10.50%)
Annotated in all Databases	5,926 (2.62%)	5,365 (6.71%)	5,328 (7.48%)	5,803 (6.76%)
Annotated in at least one Database	92,697 (40.94%)	54,455 (68.15%)	52,204 (73.33%)	54,888 (63.94%)
Total Unigenes	226,409	79,900	71,187	85,839

FIGURE 2 Venn diagram of annotation results from top five databases. Number of annotated unigenes were marked in each circle
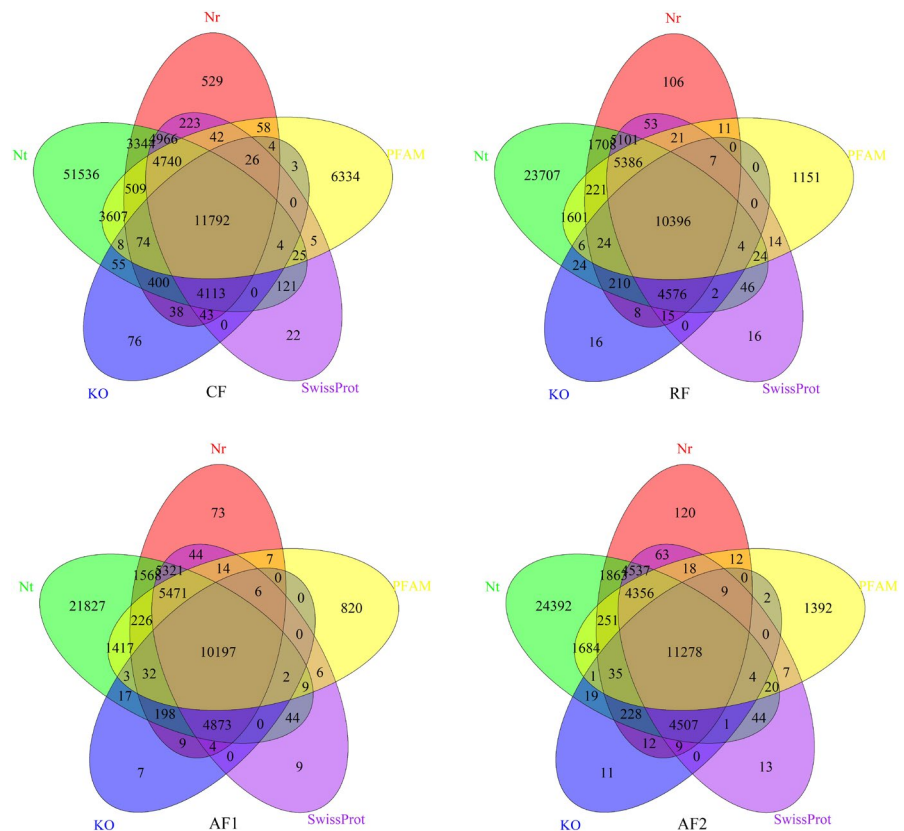
3.4 | Calculation of Ka/Ks to identify positively selected genes

To reveal the genetic mechanism of adaptation in response to the environment, Paml-codeml software was used to perform selective stress analysis on the identified single-copy orthologous genes. The 393 identified single-copy orthologous genes were used to calculate the Ka/Ks ratio. The results showed that 16 orthologous genes were identified to be under positive selection ($\omega > 1$, Table S2, Figure 4).

Enrichment analysis of positive genes showed that the *LRP11* gene was enriched in several pathways related to the

TABLE 4 Statistics of CDS

	Number of Blast to Protein database		Number of prediction by Estscan	
Sample	Total (Percent)	>300	Total (Percent)	>300
CF	18,476 (16.4%)	13,800	94,005 (83.6%)	21,074
RF	13,333 (25.8%)	10,256	38,328 (74.2%)	13,056
AF1	12,967 (27.4%)	10,013	34,356 (72.6%)	12,473
AF2	14,066 (26.1%)	11,000	39,824 (73.9%)	12,490

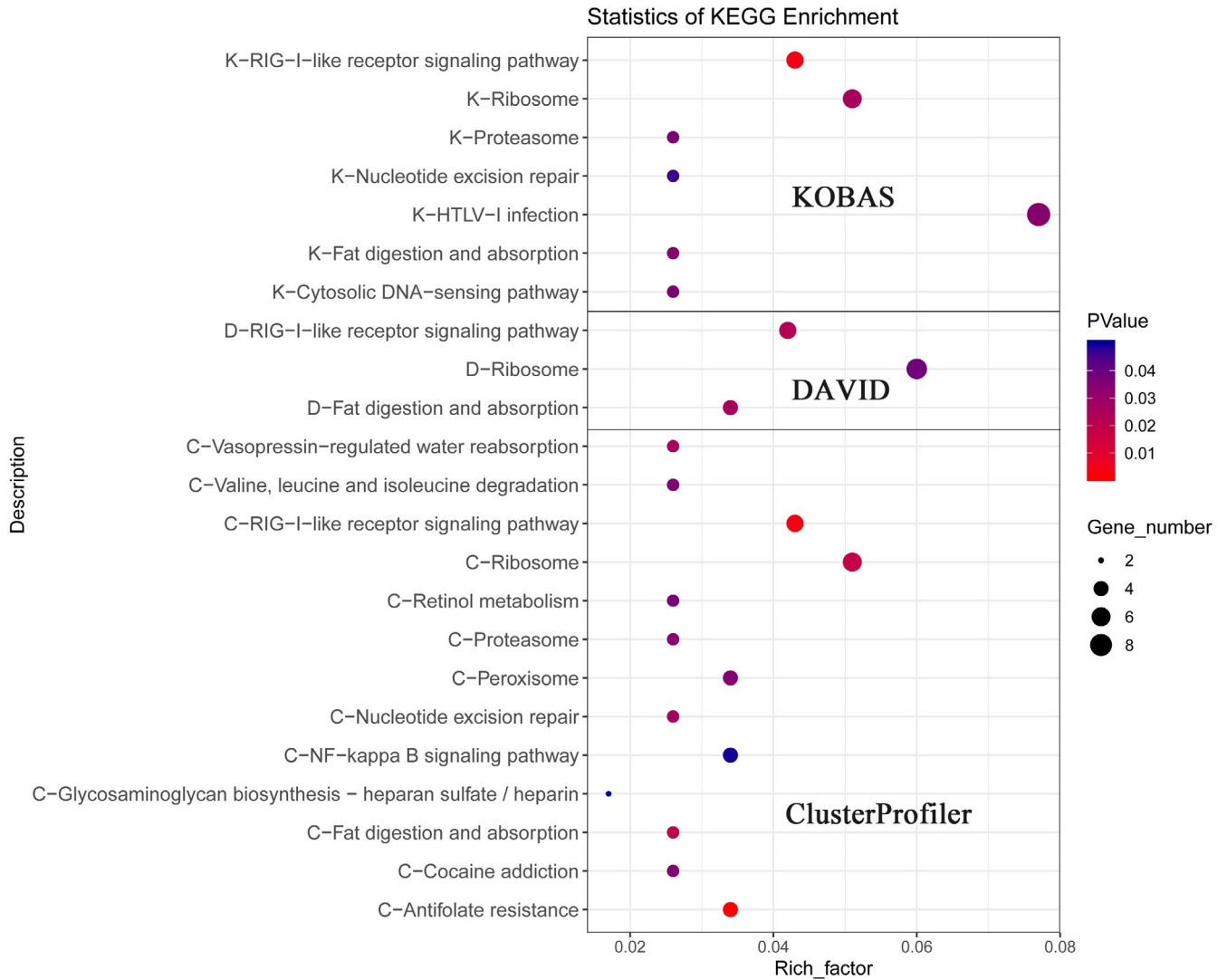*Note:* >300: Number of CDS longer than 300 nucleotides.

FIGURE 3 Scatter diagram of KEGG functional classifications for orthologous genes using the KOBAS (K), DAVID (D), and ClusterProfiler (C) methods. The x-axis shows the percentage of gene number/background of orthologous genes, and the y-axis shows the enriched KEGG pathway
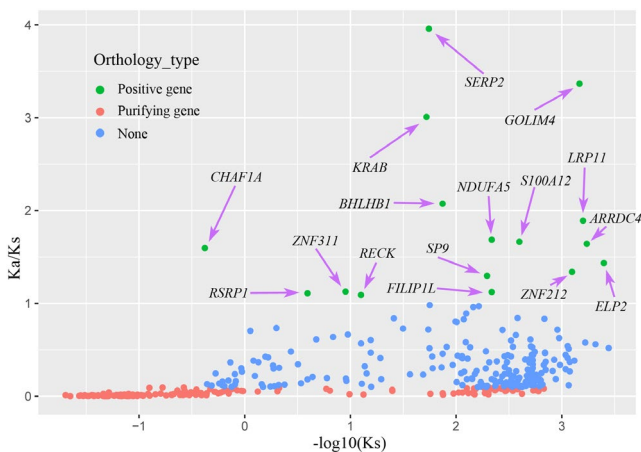

FIGURE 4 Ka/Ks distribution diagram of genes under positive and purifying selection. The x-axis shows the value of Ka/Ks. We set the y-axis to -log10(Ks) to highlight the positively selected genes

temperature response, such as response to cold (GO:0009409, $p = .087$), response to heat (GO:0009408, $p = .019$), and response to temperature stimulus (GO:0009266, $p = .029$). In addition, the *ARRDC4*, *NDUFA5*, and *OLIG2* genes were enriched in pathways related to ubiquitin proteins, respiration, and neuroregulation, respectively.

3.5 | Homologous gene analysis of three species of fox

Combined with our previously published study, a total of 27 orthologous genes were obtained in five samples. Among them, the *N4BP1* and *TWSG1* genes have sequence deletions in some individuals. It can be concluded that such deletions are caused by sequencing, since the targets of transcriptome sequencing were coding RNA. Therefore, the remaining 25 genes were used for subsequent

analysis. Three methods were used to construct ML and BI trees based on 25 genes (Figure 5a) and mitochondrial protein-coding genes (Table S3, Figure 5b).

As shown in Figure 5, the mitochondrial phylogenetic tree was the same as the traditional classification, but it was different from the results of orthologous gene construction. That is, the classification status of the corsac fox was different. To address this, gene trees of the 25 orthologous genes were constructed using RAxML and analyzed according to previous studies.

A total of 15 gene trees were constructed from 25 orthologous genes (Figure S1). Among them, the gene trees of *ARL3*, *CCDC151*, *GPC3*, *RDH10*, and *TP53I3* were the same as the tree constructed by mitochondria, and of the rest of the evolutionary tree, there were two main types: (1) the gene tree of 9 genes (*ARF4*, *F13B*, *OTUD5*, *RGS4*, *SCNN1B*, *ANAPC5*, *PXN*, *FAM122A*, and *CAMLG*) was the same as the tree constructed by orthologous genes. That is, corsac fox and arctic fox were the first cluster, with rad fox being the second cluster (Figure 5c); (2) the gene tree of 6 genes (*FAM171B*, *MGST1*, *GPLPH3*, *RARRES2*, *RND1*, and *LONRF1*) showed that rad fox and arctic fox were the first cluster, with corsac fox being the second cluster (Figure 5d).

4 | DISCUSSION

Vulpes corsac, *Vulpes vulpes*, and *Vulpes lagopus* have evolved different physiological, morphological, and behavior adaptations to cope with harsh environments. In the present study, comparative transcriptome analysis was used to identify the selected genes and pathways between them.

The results of orthologous gene enrichment showed that three genes (*AQP3*, *EGR1*, and *FABP1*) were annotated with pathways associated with the hypoxic response. The *AQP3* gene encodes aquaporin 3, which is located in the basolateral membrane of renal collecting duct cells and could promote the transport of small non-ionic solutes such as urea and glycerol (Hou et al., 2016). The *FABP1* gene encodes fatty acid-binding proteins, which are highly conserved hydrophobic proteins that are responsible for the uptake, transport, and metabolism of fatty acids (Linssen et al., 1990). In addition, there were also a large number of pathways related to nerve cells, immunity, and protein and lipid metabolism, including a large number of genes. These results suggest that these important homologous genes might have been preserved in response to harsh environments.

Selective pressure analysis results showed that a total of 16 genes were screened under positive selection from 393 single-copy orthologous genes. Enrichment analysis of positive genes showed that the *LRP11* gene which encodes low-density lipoprotein receptor-associated protein was enriched in several pathways related to the low-temperature response. In addition to our enrichment results, previous studies have shown that it is involved in animal anxiety responses (Cai, 2015). We speculated that this gene might play a key role in the response to environmental stimuli of animals, and the

positive selection of this gene might be the evolution of foxes to adapt to various environmental changes, such as cold and hunger.

In addition, several genes were related to DNA damage repair (*ELP2* and *CHAF1A*), innate immunity (*ARRDC4* and *S100A12*) and the respiratory chain (*NDUFA5*). Among these genes, the *ARRDC4* gene was involved in a pathway involving ubiquitin proteins. This gene has also been shown to promote *K63* polyubiquitination of *MDA5* through *TRIM65* and regulate the innate immune response induced by enterovirus (Meng et al., 2017). The protein encoded by the *S100A12* gene acts as an inflammatory mediator molecule that activates *RAGE* and TLR-4 mediated autoimmune responses. It could also affect the transport of some lipoproteins and increase the calcification of blood vessels (Li & Dewey, 2011). Among the remaining positive selection genes, there were three zinc finger protein genes (*ZNF212*, *KRAB*, and *ZNF311*). The main role of *KRAB* is to induce gene silencing and maintain genome stability, which plays a huge role in determining the evolution of species (Ivanov et al., 2007; Rowe & Trono, 2011). Studies on orthologous transcription factors in Drosophila have found that there were widespread differences in the binding of this protein to DNA, and such differences occur in a progressive and evolutionarily feasible way (Siggers et al., 2014). The adaptation of zinc finger proteins combined with the characteristics of new sites enhanced the adaptive evolution of species (Wang et al., 2016). These positively selected genes involved in the stress response, DNA damage repair, and innate immunity might play a role in adaptation to harsh wild fox environments.

The results of common orthologous gene analysis showed that the gene tree constructed by orthologous genes was quite different from the species tree. A literature review shows that the main causes of this situation were horizontal gene transfer (HGT), gene duplication and loss, hybridization, and convergent evolution (Degnan & Rosenberg, 2009). Among the possible reasons, HGT is generally believed to occur in microorganisms, such as bacteria, and does not affect phylogenetic relationships in higher animals (Yang, 2017). Sequence alignment results showed that gene duplication or loss did not occur in these orthologous genes. Thus, we speculate that gene flow or convergent evolution might have contributed to this result. (1) If this difference was due to hybridization, then we hypothesize that there was no reproductive isolation between the three species prior to a certain point in time, so gene flow may have existed between them. This gene flow may be an important factor in promoting regional differentiation of species (Yang, 2017). (2) If these differences were due to convergent evolution, it might mean that these genes evolved convergent in different foxes to adapt to the same environment. In addition, a large number of CDSs discovered in the fox transcriptome data will provide basic transcriptome data for future research.

5 | CONCLUSIONS

Despite not representing an entire genome, comparative transcriptome analysis was used to identify numerous genes and pathways
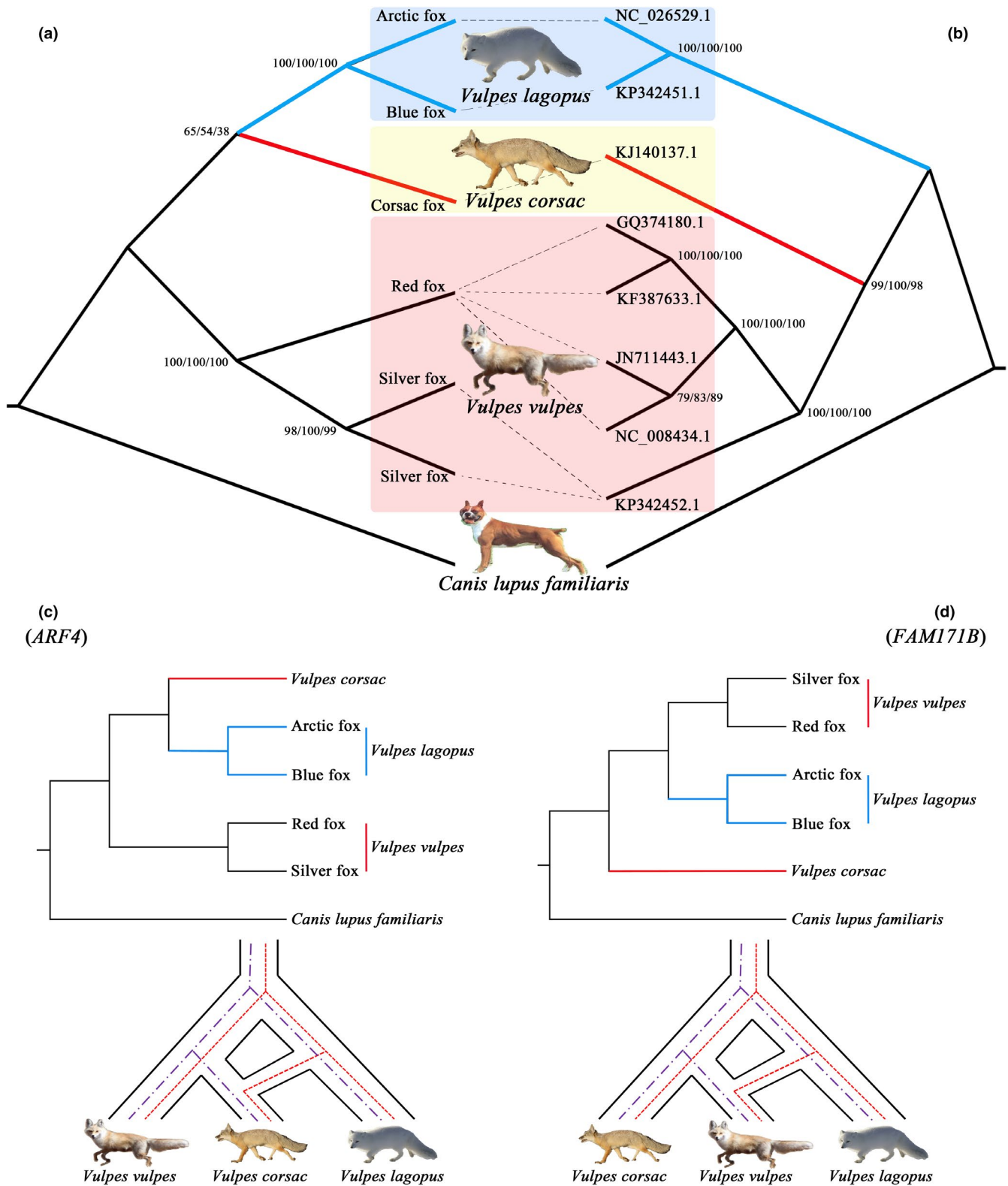
FIGURE 5 Homologous gene analysis of three species of fox. (a) Phylogenetic tree based on orthologous genes; (b) phylogenetic tree based on mitochondrial DNA. The values at each branch are the posterior probabilities of PAUP, Bayes, and RAxML. (c) and (d) Schematic diagram of the differences between species trees and gene trees. The purple dotted line represents the results of the species tree, and the red dotted line represents the results of the gene tree

that are relevant to adaptation in nonmodel organisms. Our study assembled and characterized the first transcriptome data for corsac fox using RNA-Seq technology and revealed candidate *LRP11* genes

that might be involved in cold environment adaptation in foxes. The respiratory chain, innate immunity, and DNA damage repair genes under positive selection might play a role in adaptation to

temperature fluctuations and energy expenditure in foxes. In addition, gene flow and convergent evolution might also improve the fitness of foxes. These transcriptomic data will not only promote future studies of evolution and adaptation in foxes and other canids but also provide important resources for the conservation of foxes.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interests.

AUTHOR CONTRIBUTIONS

Xiufeng Yang: Conceptualization (equal); Formal analysis (lead); Software (equal); Writing – original draft (lead). **Guolei Sun:** Funding acquisition (equal); Software (equal); Writing – review & editing (lead). **Tian Xia:** Formal analysis (equal); Software (equal). **Muha Cha:** Resources (equal). **Lei Zhang:** Software (equal). **Bo Pang:** Resources (equal). **Qingming Tang:** Resources (equal). **Huashan Dou:** Resources (lead). **Honghai Zhang:** Conceptualization (equal); Funding acquisition (lead).

ETHICAL APPROVAL

All sample procedures and experimental methods were approved by the Qufu Normal University Institutional Animal Care and Use Committee (No. QFNU2017-024), Qufu, China.

DATA AVAILABILITY STATEMENT

All data analyses during this study are deposited in the Sequence Read Archive database with accession number of SRR16229812. Appendices in the study are available in Dryad (https://doi.org/10.5061/dryad.h9w0vt4kx).

ORCID

Xiufeng Yang https://orcid.org/0000-0001-7521-954X
Honghai Zhang https://orcid.org/0000-0002-5010-2502

REFERENCES

Abdybekova, A., Sultanov, A., Zhumabayeva, A., Shapiyeva, Z., Toksanbayev, D., Budke, C. M., & Torgerson, P. R. (2015). Zoonotic cetodes of wild and domestic canids in southern kazakhstan. *American Journal of Tropical Medicine and Hygiene*, *93*(4), 313–314.

Alexandros, S. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Brandon, R. N. (1978). Adaptation and evolutionary theory. *Studies in History & Philosophy of Science*, *9*(3), 181–206. https://doi.org/10.1016/0039-3681(78)90005-5

Cai, R. (2015). *Building the regulatory pathway of Nr3c2 in anxious animal model by using system biology approach*. Nantong University.

Cock, P. J., Fields, C. J., Goto, N., Heuer, M. L., & Rice, P. M. (2010). The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research*, *38*(6), 1767–1771. https://doi.org/10.1093/nar/gkp1137

Degnan, J. H., & Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, *24*(6), 332–340. https://doi.org/10.1016/j.tree.2009.01.009

Du, Z., Huang, K., Zhao, J., Song, X., Xing, X., Wu, Q., Zhang, L., & Xu, C. (2017). Comparative transcriptome analysis of raccoon dog skin to determine melanin content in hair and melanin distribution in skin. *Scientific Reports*, *7*, 40903. https://doi.org/10.1038/srep40903

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.2460/ajvr.69.1.82

Emrich, S. J., Barbazuk, W. B., Li, L., & Schnable, P. S. (2007). Gene discovery and annotation using LCM-454 transcriptome sequencing. *Genome Research*, *17*(1), 69–73. https://doi.org/10.1101/gr.5145806

Fredrik, R., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, *19*(12), 1572–1574. https://doi.org/10.1093/bioinformatics/btg180

Garber, M., Grabherr, M. G., Guttman, M., & Trapnell, C. (2011). Computational methods for transcriptome annotation and quantification using RNA-seq. *Nature Methods*, *8*(6), 469–477. https://doi.org/10.1038/nmeth.1613

Gou, X., Wang, Z., Li, N., Qiu, F., Xu, Z. E., Yan, D., Yang, S., Jia, J., Kong, X., Wei, Z., Lu, S., Lian, L., Wu, C., Wang, X., Li, G., Ma, T., Jiang, Q., Zhao, X., Yang, J., ... Li, Y. (2014). Whole-genome sequencing of six dog breeds from continuous altitudes reveals adaptation to high-altitude hypoxia. *Genome Research*, *24*(8), 1308–1315. https://doi.org/10.1101/gr.171876.113

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., ... Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Guilherme, J. L., Burnside, R. J., Collar, N. J., & Dolman, P. M. (1938). Consistent nest-site selection across habitats increases fitness in Asian Houbara. *The Auk*, *135*(2), 192–205. https://doi.org/10.1642/AUK-17-156.1

Guindon, S., Lethiec, F., Duroux, P., & Gascuel, O. (2005). PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Research*, *33*, 557–559. https://doi.org/10.1093/nar/gki352

Hou, S. Y., Li, Y. P., Wang, J. H., Yang, S. L., Wang, Y., Wang, Y., & Kuang, Y. (2016). Aquaporin-3 inhibition reduces the growth of NSCLC cells induced by Hypoxia. *Cell Physiology and Biochemistry*, *38*(1), 129–140. https://doi.org/10.1159/000438615

Iseli, C., Jongeneel, C. V., & Bucher, P. (1999). ESTScan: A program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. In *International Conference on Intelligent Systems for Molecular Biology* (pp. 138–148). https://www.ncbi.nlm.nih.gov/pubmed/10786296

IUCN. (2021). *The IUCN Red List of Threatened Species*. Version 2021-2. https://www.iucnredlist.org

Ivanov, A. V., Peng, H., Yurchenko, V., Yap, K. L., Negorev, D. G., Schultz, D. C., Psulkowski, E., Fredericks, W. J., White, D. E., Maul, G. G., Sadofsky, M. J., Zhou, M.-M., & Rauscher, F. J. (2007). PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. *Molecular Cell*, *28*(5), 823–837. https://doi.org/10.1016/j.molcel.2007.11.012

Kukekova, A. V., Johnson, J. L., Teiling, C., Li, L., Oskina, I. N., Kharlamova, A. V., Gulevich, R. G., Padte, R., Dubreuil, M. M., Vladimirova, A. V., Shepeleva, D. V., Shikhevich, S. G., Sun, Q. I., Ponnala, L., Temnykh, S. V., Trut, L. N., & Acland, G. M. (2011). Sequence comparison of prefrontal cortical brain transcriptome from a tame and an aggressive silver fox (*Vulpes vulpes*). *BMC Genomics*, 12, 482. https://doi.org/10.1186/1471-2164-12-482

Kumar, V., Kutschera, V. E., Nilsson, M. A., & Janke, A. (2015). Genetic signatures of adaptation revealed from transcriptome sequencing of Arctic and red foxes. *BMC Genomics*, 16(1), 585. https://doi.org/10.1186/s12864-015-1724-9

Lan, D., Ji, W., Xiong, X., Liang, Q., Yao, W., Mipam, T.-D., Zhong, J., & Li, J. (2021). Population genome of the newly discovered Jinchuan yak to understand its adaptive evolution in extreme environments and generation mechanism of the multirib trait. *Integrative Zoology*, 16(5), 685–695. https://doi.org/10.1111/1749-4877.12484

Li, B., & Dewey, C. N. (2011). RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, 12(1), 323. https://doi.org/10.1186/1471-2105-12-323

Li, L., Stoeckert, C. J., Jr., & Roos, D. S. (2003). OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research*, 13(9), 2178–2189. https://doi.org/10.1101/gr.1224503

Linssen, M. C., Vork, M. M., deJong, Y. F., Glatz, J. F., & van derVusse, G. J. (1990). Fatty acid oxidation capacity and fatty acid-binding protein content of different cell types isolated from rat heart. *Molecular and Cellular Biochemistry*, 98(1–2), 19–25. https://doi.org/10.1007/BF00231363

Liu, G., Zhang, H., Sun, G., Zhao, C., Shang, S., Gao, X., Xia, T., & Yang, X. (2017). Characterization of the peripheral blood transcriptome and adaptive evolution of the MHC I and TLR gene families in the wolf (*Canis lupus*). *BMC Genomics*, 18(1), 584. https://doi.org/10.1186/s12864-017-3983-0

Liu, G., Zhao, C., Yang, X., Shang, J., Gao, X., Sun, G., Dou, H., & Zhang, H. (2019). Comparative analysis of peripheral blood reveals transcriptomic adaptations to extreme environments on the Qinghai-Tibetan Plateau in the gray wolf (*Canis lupus chanco*). *Organisms Diversity & Evolution*, 19(3), 543–556. https://doi.org/10.1007/s13127-019-00405-3

Lkhagvasuren, M., Murdoch, J. D., Munkhzul, T., & Strong, A. M. (2016). Predicting the effects of habitat loss on corsac fox occupancy in Mongolia. *Journal of Mammalogy*, 97(4), 1153–1163. https://doi.org/10.1093/jmammal/gyw067

Martchenko, D., Chikhi, R., & Shafer, A. B. A. (2020). Genome assembly and analysis of the North American Mountain Goat (*Oreamnos americanus*) reveals species-level responses to extreme environments. *G3-Genes Genomes Genetics*, 10(2), 437–442. https://doi.org/10.1534/g3.119.400747

Meng, J., Yao, Z., He, Y., Zhang, R., Zhang, Y., Yao, X., Yang, H., Chen, L., Zhang, Z., Zhang, H., Bao, X., Hu, G., Wu, T., & Cheng, J. (2017). *ARRDC4* regulates enterovirus 71-induced innate immune response by promoting K63 polyubiquitination of MDA5 through TRIM65. *Cell Death & Disease*, 8(6), e2866. https://doi.org/10.1038/cddis.2017.257

Miao, B., Wang, Z., & Li, Y. (2017). Genomic analysis reveals hypoxia adaptation in the Tibetan mastiff by introgression of the gray wolf from the Tibetan Plateau. *Molecular Biology and Evolution*, 34(3), 734–743. https://doi.org/10.1093/molbev/msw274

Nakhleh, L. (2013). Computational approaches to species phylogeny inference and gene tree reconciliation. *Trends in Ecology & Evolution*, 28(12), 719–728. https://doi.org/10.1016/j.tree.2013.09.004

Novaes, E., Drost, D. R., Farmerie, W. G., Pappas, G. J.Jr, Grattapaglia, D., Sederoff, R. R., & Kirst, M. (2008). High-throughput gene and SNP discovery in Eucalyptus grandis, an uncharacterized genome. *BMC Genomics*, 9(1), 312. https://doi.org/10.1186/1471-2164-9-312

Parchman, T. L., Geist, K. S., Grahnen, J. A., Benkman, C. W., & Buerkle, C. A. (2010). Transcriptome sequencing in an ecologically important tree species: Assembly, annotation, and marker discovery. *BMC Genomics*, 11(1), 180. https://doi.org/10.1186/1471-2164-11-180

Peng, Y., Li, H., Liu, Z., Zhang, C., Li, K., Gong, Y., & Li, X. (2021). Chromosome-level genome assembly of the Arctic fox (*Vulpes lagopus*) using PacBio sequencing and Hi-C technology. *Molecular Ecology Resources*, 21(6), 2093–2108. https://doi.org/10.1111/1755-0998.13397

Rowe, H. M., & Trono, D. (2011). Dynamic control of endogenous retroviruses during development. *Virology*, 411(2), 273–287. https://doi.org/10.1016/j.virol.2010.12.007

Siggers, T., Reddy, J., Barron, B., & Bulyk, M. L. (2014). Diversification of transcription factor paralogs via noncanonical modularity in C2H2 zinc finger DNA binding. *Molecular Cell*, 55(4), 640–648. https://doi.org/10.1016/j.molcel.2014.06.019

Statham, M. J., Murdoch, J., Janecka, J., Aubry, K. B., Edwards, C. J., Soulsbury, C. D., Berry, O., Wang, Z., Harrison, D., Pearch, M., Tomsett, L., Chupasko, J., & Sacks, B. N. (2014). Range-wide multilocus phylogeography of the red fox reveals ancient continental divergence, minimal genomic exchange and distinct demographic histories. *Molecular Ecology*, 23(19), 4813–4830. https://doi.org/10.1111/mec.12898

Swofford, D. (1998). *PAUP: Phylogenetic Analysis Using Parsimony*.

Wang, J., Wang, J., & Tian, C. Y. (2016). Evolution of KRAB-containing zinc finger proteins and their roles in species evolution. *Hereditas (Beijing)*, 38(11), 971–978. https://doi.org/10.16288/j.yczz.16-056

Wu, H., Liu, Y.-H., Wang, G.-D., Yang, C.-T., Otecko, N. O., Liu, F., Wu, S.-F., Wang, L. U., Yu, L. I., & Zhang, Y.-P. (2016). Identifying molecular signatures of hypoxia adaptation from sex chromosomes: A case for Tibetan Mastiff based on analyses of X chromosome. *Scientific Reports*, 6, 35004. https://doi.org/10.1038/srep35004

Yang, J. I., Li, W.-R., Lv, F.-H., He, S.-G., Tian, S.-L., Peng, W.-F., Sun, Y.-W., Zhao, Y.-X., Tu, X.-L., Zhang, M., Xie, X.-L., Wang, Y.-T., Li, J.-Q., Liu, Y.-G., Shen, Z.-Q., Wang, F., Liu, G.-J., Lu, H.-F., Kantanen, J., ... Liu, M.-J. (2016). Whole-genome sequencing of native sheep provides insights into rapid adaptations to extreme environments. *Molecular Biology and Evolution*, 33(10), 2576–2592. https://doi.org/10.1093/molbev/msw129

Yang, X. (2017). *The Phylogenomic of Aurantioideae and Phylogeography of Wild Citron (Citrus medica) and Ichang Papeda (Citrus ichangensis)*. Huazhong Agricultural University.

Yang, X., Liu, G., Wang, Q. I., Gao, X., Xia, T., Zhao, C., Dou, H., & Zhang, H. (2021). Comparative transcriptome provides insights into the selection adaptation between wild and farmed foxes. *Ecology and Evolution*, 11(19), 13475–13486. https://doi.org/10.1002/ece3.8071

Yang, X., Zhang, H., Shang, J., Liu, G., Xia, T., Zhao, C., Sun, G., & Dou, H. (2018). Comparative analysis of the blood transcriptomes between wolves and dogs. *Animal Genetics*, 49(4), 291–302. https://doi.org/10.1111/age.12675

Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, 24(8), 1586–1591. https://doi.org/10.1093/molbev/msm088

Yu, G., Wang, L. G., Han, Y., & He, Q. Y. (2012). clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS: A Journal of Integrative Biology*, 16(5), 284–287. https://doi.org/10.1089/omi.2011.0118

Zhang, G., Guo, G., Hu, X., Zhang, Y., Li, Q., Li, R., Zhuang, R., Lu, Z., He, Z., Fang, X., Chen, L. I., Tian, W., Tao, Y., Kristiansen, K., Zhang, X., Li, S., Yang, H., Wang, J., & Wang, J. (2010). Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome. *Genome Research*, 20(5), 646–654. https://doi.org/10.1101/gr.100677.109

Zhang, K., He, C., Xu, Y., Zhang, C., Li, C., Jing, X., & Chen, Y. (2021). Taxonomic and functional adaption of the gastrointestinal microbiome of goats kept at high altitude (4800 m) under intensive or extensive rearing conditions. *FEMS Microbiology Ecology*, *97*(3). https://doi.org/10.1093/femsec/fiab009

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.