Guo et al. *Zool. Res.* 2024, 45(6): 1209−1220
https://doi.org/10.24272/j.issn.2095-8137.2024.029
https://cstr.cn/32248.14.zr.2024.029

ZR | Zoological Research

Article Open Access

Convergent evolution in high-altitude and marine mammals: Molecular adaptations to pulmonary fibrosis and hypoxia

Bo-Xiong Guo[1], Ya Zhang[1], Xiao-Yu Sun[1], Yi-Xuan Sun[1], Wen-Jun Lv[1], Shi-Xia Xu[1], Guang Yang[1], Wen-Hua Ren[1,*]

[1] *Jiangsu Key Laboratory for Biodiversity and Biotechnology*, *College of Life Sciences*, *Nanjing Normal University*, *Nanjing*, *Jiangsu* 210046, *China*

ABSTRACT

High-altitude and marine mammals inhabit distinct ecosystems but share a common challenge: hypoxia. To survive in low-oxygen environments, these species have evolved similar phenotypic pulmonary adaptations, characterized by a high density of elastic fibers. In this study, we explored the molecular mechanisms underlying these adaptations, focusing on pulmonary fibrosis and hypoxia tolerance through comparative genomics and convergent evolution analyses. We observed significant expansions and contractions in certain gene families across both high-altitude and marine mammals, closely associated with processes involved in pulmonary fibrosis. Notably, members of the keratin gene family, such as *KRT17* and *KRT14*, appear to be associated with the development of the dense elastic fiber phenotype observed in the lungs of hypoxia-tolerant mammals. Through selection pressure and amino acid substitution analyses, we identified multiple genes exhibiting convergent accelerated evolution, positive selection, and amino acid substitution in these species, associated with adaptation to hypoxic environments. Specifically, the convergent evolution of *ZFP36L1*, *FN1*, and *NEDD9* was found to contribute to the high density of elastic fibers in the lungs of both high-altitude and marine mammals, facilitating their hypoxia tolerance. Additionally, we identified convergent amino acid substitutions and gene loss events associated with sperm development, differentiation, and spermatogenesis, such as amino acid substitutions in *SLC26A3* and pseudogenization of *CFAP47*, as confirmed by PCR. These genetic alterations may be linked to changes in the reproductive capabilities of these animals. Overall, this study offers novel perspectives on the genetic and molecular adaptations of high-altitude and marine mammals to hypoxic environments, with a particular emphasis on pulmonary fibrosis.

Keywords: Convergent evolution; Pulmonary fibrosis; High-altitude mammals; Marine mammals; Hypoxia

INTRODUCTION

Mammals have adapted to various extreme environments, encompassing both high-altitude plateaus and deep oceans (Li et al., 2021a; Withers et al., 2016). Despite the distinct nature of these habitats, both high-altitude and marine environments pose a shared survival challenge: hypoxia (Li et al., 2021a). To overcome the reduced availability of oxygen, species in these regions have evolved specialized physiological traits. In high-altitude environments, mammals have evolved mechanisms such as increased red blood cell counts and hemoglobin with higher oxygen-binding affinity (Ma et al., 2023; Storz, 2007; Storz & Bautista, 2022). These adaptations improve oxygen transport in the blood, crucial for survival in low-oxygen conditions (Storz, 2007; Storz & Bautista, 2022). Additionally, many high-altitude mammals exhibit lower basal metabolic rates to conserve energy under hypoxic stress and display morphological traits such as shorter limbs and thicker fur, optimizing heat retention in cold climates (Pu et al., 2019; Wu et al., 2020). Marine mammals, such as whales and dolphins, exhibit distinct adaptations for underwater life. Notably, these species possess elevated levels of myoglobin in their muscles, enabling enhanced oxygen storage, with their hemoglobin-rich blood facilitating efficient oxygen transport during prolonged underwater activities (Isogai et al., 2021). Moreover, they regulate body temperature and metabolic rates to conserve oxygen and adapt to high-pressure environments during deep dives (Favilla & Costa, 2020). Their streamlined body shapes also minimize drag to optimize underwater movement (Zhou et al., 2018a).

Lung function is central to the adaptation of mammals to hypoxic conditions, serving as a vital interface with the

Received: 16 April 2024; Accepted: 17 June 2024; Online: 18 June 2024
Foundation items: This work was supported by the National Natural Science Foundation of China (32270442, 31872219, 31370401, 32030011, 31630071, 31772448), National Key Research and Development Program of China (2022YFF1301602), and Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX23_1747, KYCX23_1740)
*Corresponding author, E-mail: 08162@njnu.edu.cn

external environment (Lee et al., 2020). Comparative analyses have demonstrated convergent adaptations in lung structure between high-altitude and marine mammals, with both groups possessing lungs with a notably increased surface area. This anatomical characteristic significantly increases oxygen absorption and the oxygen-carrying capacity of blood, distinguishing them from other species (Kooyman, 1973). Additionally, these mammals exhibit a specialized hemoglobin variant with heightened oxygen affinity, facilitating oxygen transport in low-oxygen environments (Storz, 2007; Storz & Bautista, 2022). The expansion of alveolar surfaces and increased vascular contact areas further optimize the transfer of oxygen into the bloodstream (Storz, 2007; Storz & Bautista, 2022). Recent investigations have revealed that wild yaks possess a comprehensive elastic system in their lungs, comprising pulmonary pleura, interlobular connective tissue, and alveolar septa, which aids lung dilation, contraction, and improved oxygen absorption at high altitudes (Chen et al., 2006; Gao et al., 2022; Li et al., 2021b). Similarly, research on Xizang antelopes has shown that the pulmonary arteriole walls thicken at higher altitudes, with increased elastic fibers enhancing lung tension, preventing pulmonary edema, and ensuring efficient gas exchange under hypoxic conditions, crucial for survival in hypoxic environments (Zhao et al., 2022). These adaptations are not exclusive to these species but are also observed in other high-altitude species, such as the plateau pika and Xizang pig (Yang et al., 2021; Zhang et al., 2022). In marine mammals like whales, thickened alveolar walls rich in elastic fibers enable lung collapse during prolonged dives (Fahlman, 2024), reducing the risk of decompression sickness and facilitating breath-holding (Henk & Haldiman, 1990; Piscitelli et al., 2013). Studies suggest that the presence of elastic fibers in marine mammal lungs may also play a role in their ability to endure anoxic environments (Guo et al., 2024). Pinnipeds, including seals and sea lions, similarly have lungs rich in elastic fibers, which aid in both hypoxic tolerance and decompression sickness risk (Boyd, 1975; Fahlman et al., 2017; Falke et al., 1985). The unique pulmonary phenotypes of these mammals may also provide insights into fibrotic lung diseases in humans (Guo et al., 2024). At the molecular level, numerous studies have investigated the genetic adaptations of high-altitude and marine mammals to hypoxia (Li et al., 2021, 2021c; Lyu et al., 2022; Tian et al., 2016, 2017; Wu et al., 2020; Yépez et al., 2023). For example, Lyu et al. (2022) explored the evolution of critical genes, including *CXCL10* and *SCAMP1*, in six high-altitude mammalian species, highlighting their importance in adapting to low oxygen levels. Similarly, Li et al. (2021c) analyzed how terrestrial species adjust to hypoxia through both physiological changes and gene regulation from physiological and molecular viewpoints. Furthermore, Wu et al. (2020) focused on the adaptive evolution of domestic high-altitude mammals, such as yaks and sheep, emphasizing the pivotal role of *EPAS1* within the context of convergent evolution. In marine mammals, comparative genomic analysis and other techniques have examined gene evolution related to energy metabolism, a key factor in surviving hypoxic conditions (Tian et al., 2017). The hypoxia-inducible factor (HIF) signaling pathway and selective pressure on hypoxia-specific genes, such as *EDN1* and *AGTR1*, are vital for the adaptation of marine mammals to low-oxygen environments (Tian et al., 2016; Yépez et al., 2023). Despite substantial research on hypoxic adaptability in high-altitude and marine mammals, studies on their common phenotypes, such as widespread pulmonary elastin fibrosis, and shared genetic mechanisms remain limited. This gap suggests significant opportunities for further research into the mechanisms underlying mammalian adaptation to extreme environments.

Advancements in sequencing technologies and the expansion of mammalian genomic data have significantly enhanced the ability to explore the molecular mechanisms by which high-altitude and marine mammals adapt to pulmonary fibrosis and hypoxic environments. In this study, we employed genomic approaches to analyze 23 mammalian species, including high-quality genome-sequenced high-altitude and marine mammals and closely related species. We focused on the convergent evolution of single-copy orthologous genes, examining selective pressures, amino acid substitutions, gene loss, and the expansion and contraction of gene families. Overall, our findings reveal key molecular evolutionary processes that allow high-altitude and marine mammals to thrive under hypoxic conditions, providing deeper insight into their unique adaptative strategies.

MATERIALS AND METHODS

Identification of orthologous genes

In this study, we curated a dataset from the genomic annotations of 23 species, excluding those with low-quality assemblies. The dataset included five cetacean species (*Tursiops truncatus*, *Physeter catodon*, *Balaenoptera musculus*, *Phocoena sinus*, and *Balaenoptera acutorostrata*), four other marine mammals (*Ursus maritimus*, *Trichechus manatus*, *Neomonachus schauinslandi*, and *Zalophus californianus*), three high-altitude mammals (*Bos mutus*, *Panthera uncia*, and *Rhinopithecus roxellana*), and their close relatives (*Bos taurus*, *Bubalus bubalis*, *Capra hircus*, *Hippopotamus amphibius*, *Ailuropoda melanoleuca*, *Canis familiaris*, *Felis catus*, *Mus musculus*, *Choloepus didactylus*, *Ornithorhynchus anatinus*, and *Homo sapiens*) (Supplementary Table S1). Due to its high-quality genome, *C. didactylus* was selected as a close relative of *T. manatus* for subsequent analysis. Genomic data were downloaded from the National Center for Biotechnology Information (https://www.ncbi.nlm.nih.gov).

Transcripts shorter than 150 base pairs or not a multiple of three base pairs in length were excluded using our custom Perl script. Single-copy orthologous genes were identified using OrthoFinder v.2.5.5 (Emms & Kelly, 2019) employing an all-against-all BLASTP approach.

Phylogenetic analysis and divergence time estimation

Molecular phylogenetic analysis was conducted using shared single-copy genes based on orthologous gene sets identified by OrthoFinder v.2.5.5 (Emms & Kelly, 2019). This involved aligning 5 247 single-copy orthologous genes using the “alignSequences” and “exportAlignment” modules in MACSE v.2 (Ranwez et al., 2011) and PRANK (Löytynoja, 2014), with default parameters. For alignment optimization, the Gblocks server (Castresana, 2000) was used to eliminate poorly aligned regions, gaps, and non-homologous fragments. The optimized sequences were then concatenated to form single reads in Fasta format. A custom Perl script was used to identify fourfold degenerate sites for phylogenetic tree construction. These alignments were then amalgamated into a super alignment matrix, which was used to infer the

phylogenetic relationships of the 23 species using maximum-likelihood (ML Tree), implemented in RAxML v.8 (Stamatakis, 2014).

Divergence time estimation was performed using the MCMCTree package in PAML v.4 (Yang, 2007). The resulting phylogenetic tree was displayed using FigTree (http://tree.bio.ed.ac.uk/software/figtree/) and iTOL (https://itol.embl.de/). Species silhouettes were obtained from PhyloPic (http://phylopic.org/). Fossil-based constraints were applied to seven nodes (*R. roxellana–H. sapiens*: 26.8–30.6 million years ago (Ma), *B. acutorostrata–B. musculus*: 12.1–15.7 Ma, *P. sinus–T. truncatus*: 16.0–19.2 Ma, *F. catus–P. uncia*: 11.6–14.7 Ma, *N. schauinslandi–Z. californianus*: 22.1–30.0 Ma, *A. melanoleuca–U. maritimus*: 17.3–28.0 Ma, and *C. didactylus–T. manatus*: 83.7–97.9 Ma) based on Timetree (Kumar et al., 2017) (http://www.timetree.org). The posterior distribution of parameters was estimated using Markov chain Monte Carlo (MCMC) sampling with 1 000 000 iterations, a burn-in of 2 000 000, and a sampling frequency of 100. Analysis results were separately examined using Tracer v.1.7.2 (Rambaut et al., 2018) to determine the convergence among parameters, node ages, and likelihood values. The estimated sample size for each parameter was above 200.

Gene family expansion and contraction analysis

Detection of homologous gene family expansions and contractions was performed using CAFE v.5.0.0 (Mendes et al., 2020), based on the OrthoFinder v.2.5.5 results (Emms & Kelly, 2019). OrthoFinder models gene gains and losses along phylogenies by employing birth and death processes. This analysis specifically focused on gene families exhibiting significant expansion or contraction (P<0.05), with at least one-third of the species adapted to hypoxic environments (Lyu et al., 2022). The goal was to identify gene families displaying evidence of convergent evolution in these species.

Gene annotation

For gene families showing significant expansion or contraction, the corresponding protein sequences were retrieved from the OrthoFinder v.2.5.5 analysis. These sequences were subsequently annotated for gene functions using eggNOG-mapper v.2 (Cantalapiedra et al., 2021) (http://eggnog-mapper.embl.de/).

Selection pressure analysis

The neutral theory of molecular evolution posits that the ratio of nonsynonymous (*K*a) and synonymous (*K*s) substitution rates in protein-coding genes serves as an indicator of natural selection. Accordingly, we calculated the average *K*a/*K*s ratios and applied the two-ratio model using Codeml from the PAML package. This approach was employed to discern rapidly evolving genes (REGs) and positively selected genes (PSGs) in species inhabiting hypoxic environments.

In the branch model, the null hypothesis posits uniform evolutionary rates across all tree branches, achieved by setting the “model” parameter to zero. Conversely, the alternative hypothesis permits differing rates for hypoxic tolerant mammals (foreground) and other branches (background) by setting the “model” parameter to two. Significant differences between these models were evaluated using the likelihood ratio test (LRT) (Lewis et al., 2011), where a significant LRT value suggests adopting the alternative hypothesis. To control for multiple comparisons, we applied the false discovery rate (FDR) method (Korthauer et al., 2019) in R software (Benjamini & Hochberg, 1995). Genes with high ω (*K*a/*K*s) values and significant *P*-adjusted values in the hypoxic tolerant branches were classified as REGs.

For the branch-site model, the null hypothesis assumes neutral evolution of all codons and branches, with parameters set to “model=2, NSsites=2, fix_omega=1, omega=1”. The alternative hypothesis, enabling positive selection at specific codons in hypoxic tolerant mammals branch, was tested using the settings “model=2, NSsites=2, fix_omega=0”. LRT-identified PSGs exhibited a Bayes Empirical Bayes posterior probability greater than 0.8.

To assess the reliability of the identified PSGs, we employed the Fixed Effects Likelihood (FEL) model in Datamonkey (https://www.datamonkey.org/). This model utilizes fixed effects likelihood for statistical testing, with a *P*-value threshold of less than 0.1 to ascertain significant selective differences in gene sequences.

Convergent amino acid substitution analysis

The FasParser v.2 tool (Sun, 2018) was employed to identify convergent amino acids in species inhabiting hypoxic environments. Human canonical sequences from UniProt (https://www.uniprot.org/) were used as the reference for amino acid positions. To assess the potential impact of convergent amino acid replacement on corresponding protein function, predictions were made using PolyPhen-2 (Adzhubei et al., 2013) (http://genetics.bwh.harvard.edu/pph2).

Convergent gene loss analysis

A detailed analysis of gene losses in various species was conducted using the Tool to Infer Orthologs from Genome Alignments (TOGA) (Kirilenko et al., 2023), entailing comparative genomic analysis between target species and the human genome to ascertain gene losses, including insertions, deletions, large fragment deletions, premature termination codons, and exons. The utilized dataset has been authorized and recommended by the developers of TOGA (https://genome.senckenberg.de/download/TOGA/) (Kirilenko et al., 2023).

Evaluation of gene loss and its evolution

To investigate the presence of inactivating mutations in hypoxia-tolerant species, multiple sequence alignments of *CFAP47* were performed across 75 mammalian species. The occurrence of stop codons and frameshift insertions was examined using human genes as a reference. Analysis suggested a potential loss of the *CFAP47* gene. This gene loss was validated by polymerase chain reaction (PCR) on the genomic DNA of bottlenose dolphins, Yangtze River dolphins, and pygmy sperm whales. Details of PCR primers are provided in Supplementary Table S2 The validation site was amplified using 2×Taq Plus Master Mix II enzyme (Vazyme Biotech, China), and the amplification products were detected through Sanger sequencing.

To assess whether *CFAP47* shows signs of relaxed selection in hypoxia-tolerant species compared to other mammals, the RELAX tool (Wertheim et al., 2015) implemented in Datamonkey (http://www.datamonkey.org) (Weaver et al., 2018) was employed. RELAX evaluates whether a test branch (hypoxia-tolerant species) undergoes relaxed or intensified selection compared to reference branches (all other lineages), allowing for variation in evolutionary rates across sites and branches (Wertheim et al.,

2015) and enabling assessment of relaxed selection in pseudogene branches. The “codeml” program from PAML was utilized to determine whether *CFAP47* has undergone neutral evolution. In this analysis, Model A assumes a uniform ω value across all branches, while Model B permits ω=1 across all branches. Purifying selection was identified by comparing Models A and B. Model C was used to calculate separate ω2 values for *CFAP47* in pseudogenized hypoxia-tolerant species and ω1 values for other mammals. Comparison between Models A and C suggested a relaxation in selective pressure. To investigate the complete absence of selective pressure in branches with a pseudogenized *CFAP47*, Model D was tested, fixing ω2=1 for pseudogenized species and assigning a distinct ω1 to all other branches. Finally, Model E was evaluated, permitting variation in ω across branches.

Enrichment analysis

To determine the biological processes associated with REGs, PSGs, gene losses, and gene families, enrichment analyses were conducted using predefined gene sets. Gene Ontology (GO, including biological processes) analysis was performed using Metascape (Zhou et al., 2019) (http://metascape.org), with *H. sapiens* set as the “Input as species” and “Analysis as species”, with *P*-value cutoff of less than 0.05.

RESULTS

Identification of orthologous gene set and phylogenetic analysis

Based on genomic analysis of the 23 mammalian species, 5 247 one-to-one orthologous genes were identified using OrthoFinder v.2.5.5 (https://github.com/davidemms/OrthoFinder/tree/master). After alignment with MACSE v.2-PRANK-Gblocks, 5 243 genes remained for further analysis.

The ML phylogenetic tree was constructed using 37 500 bp fourfold degenerate sites, with all nodes receiving 100% bootstrap support. As shown in Figure 1, the species were distributed across various taxonomic orders, including three plateau mammals, five cetaceans, and four other marine mammals. The cetaceans were classified within Cetartiodactyla, while the other marine mammals were placed within Carnivora and Sirenia and the plateau mammals were distributed across Carnivora, Cetartiodactyla, and Primates, consistent with the branch relationships reported in Foley et al. (2016). The ML tree and sequence data were used as MCMCTree inputs to estimate divergence times for all nodes in the phylogenetic tree (Figure 1). The divergence between cetaceans and *H. amphibius* was estimated at 65.28 Ma, with a 95% highest posterior density (HPD) of 46.15–85.49 Ma, older than the 54 Ma estimate by Nikaido et al. (1999) based on short and long interspersed elements. The divergence time between *B. mutus* and *B. taurus* was estimated at 5.19 Ma (95% HPD: 2.13–8.73 Ma), similar to the 4.83 Ma estimation reported in Lyu et al. (2022). Similarly, the split between *R. roxellana* and *H. sapiens* was estimated at 28.76 Ma (95% HPD: 26.82–30.61 Ma), consistent with Glazko & Nei (2003). The divergence between *T. manatus* and *C. didactylus* was estimated at 92.47 Ma (95% HPD: 84.49–98.37 Ma). These results generally align with those reported in the Timetree database. Overall, these species demonstrate independently evolved hypoxia-adaptation mechanisms, essential for

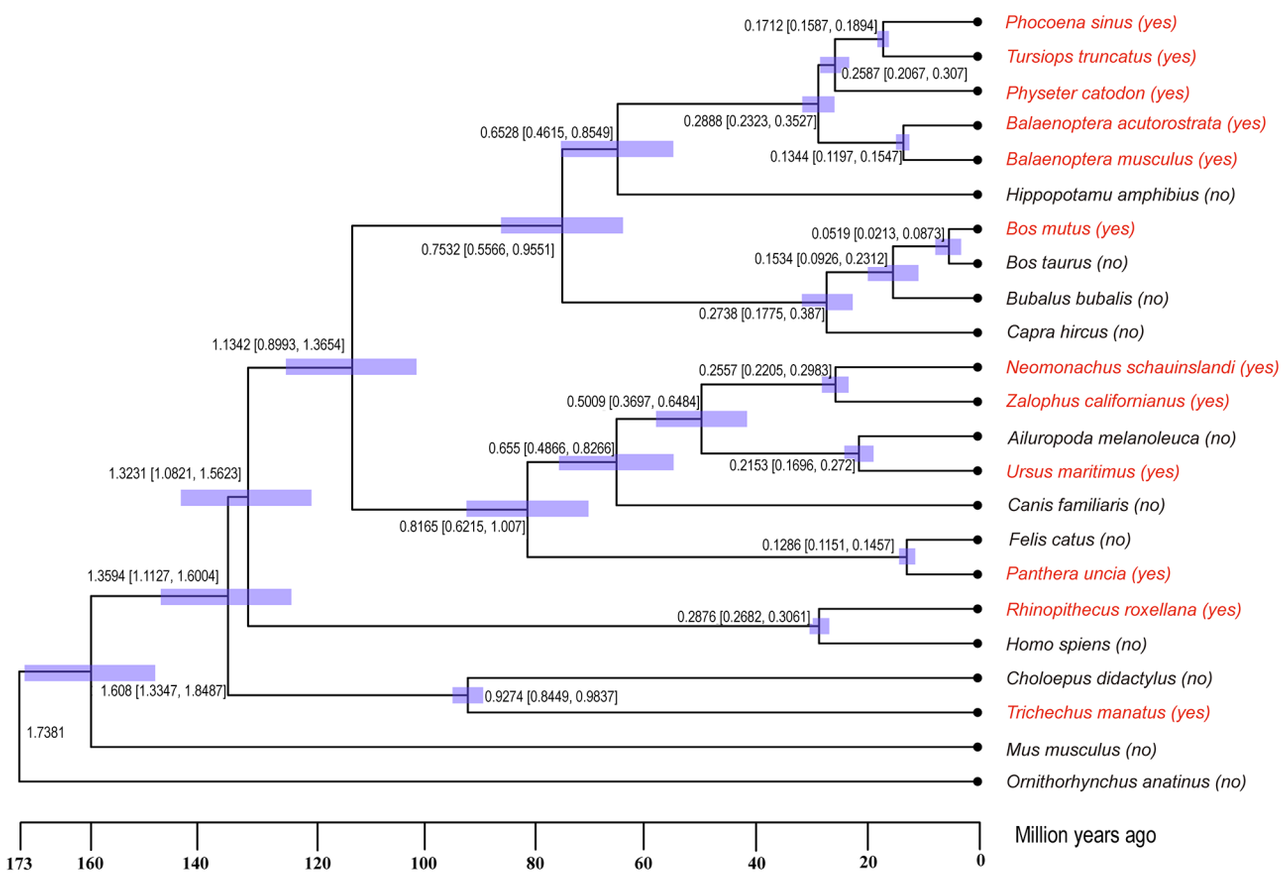

Figure 1 Phylogenetic tree and divergence times based on fourfold degenerate sites of 5 243 one-to-one orthologous genes

Red font and “yes” represent species living in low-oxygen environments, black font and “no” represent closely related species.

convergent evolution studies.

Gene family expansion and contraction analysis

Our analysis of gene family expansion and contraction revealed that nine marine mammals, comprising five cetaceans and four pinnipeds, showed less variation in gene families relative to their close relatives. Notably, *P. catodon* and *T. manatus* experienced a higher number of gene family contractions (479 and 460, respectively) compared to *B. musculus* and *Z. californianus* (115 and 79, respectively). Conversely, *H. amphibius*, *A. melanoleuca*, and *C. didactylus*, close relatives of the aforementioned species, demonstrated 209, 471, and 243 gene family contractions, respectively. Furthermore, three high-altitude mammals exhibited more significant gene family changes compared to their relatives (Figure 2A). Notably, *B. mutus* demonstrated fewer gene family expansions and significantly more contractions (+72/−1 022) relative to *B. taurus* (+160/−115). Similarly, *P. uncia* and *R. roxellana* exhibited more pronounced gene family alterations, both in expansions and contractions (+206/−345 and +223/−117, respectively), compared to their close relatives *F. catus* (+85/−85) and *H. sapiens* (+123/−94).

Regarding convergent evolutionary patterns among gene families in hypoxia-adapted environments, we identified those with significant expansion or contraction (P<0.05) in at least four hypoxia-tolerant mammals. Results indicated a higher prevalence of gene family contractions than expansion across the 12 hypoxia-tolerant mammals (Figure 2B; Supplementary Figure S1). Further annotation and enrichment analysis of these gene families highlighted their involvement in key biological processes (Figure 2C; Supplementary Figure S2). Notably, gene families with significant convergent expansion were associated with cell morphogenesis (GO:0000902, three genes, P=0.0498), protein folding (GO:0006457, six genes, P=2.56998e-07), cell-cell adhesion (GO:0098609, five genes, P=0.0005), and regulation of cellular response to stress (GO:0080135, three genes, P=0.0465) (Figure 2C). These processes likely play a critical role in the development and maintenance of lung tissue elasticity, influencing cellular responses in hypoxic environments and regulating signal transduction pathways. Moreover, we observed a significant contraction in certain members of the Keratin gene family in at least four hypoxia-tolerant species. This gene family is implicated in pulmonary fibrosis (Ficial et al., 2014; Habermann et al., 2020; Jaeger et al., 2022), and its alteration may contribute to the production of abundant elastic fibers in the lungs of these species, facilitating adaptation to low-oxygen environments.

REGs and PSGs

To investigate the genetic adaptations of species to hypoxic environments, we analyzed genes from 12 hypoxia-adapted species using PAML and Datamonkey, identifying 2 683 REGs and 167 PSGs. Enrichment analysis of GO biological processes associated with these REGs revealed several terms linked to adaptive evolution (Figure 3A, B), pulmonary fibrosis phenotype, and hypoxic adaptation, including cellular response to reduced oxygen levels (GO:0036294, 23 genes, P=0.0048), lung morphogenesis (GO:0060425, 11 genes, P=0.0027), enhancement of stress fiber assembly (GO:0051496, 10 genes, P=0.0129), stimulation of fibroblast proliferation (GO:0048146, 10 genes, P=0.0189), stress

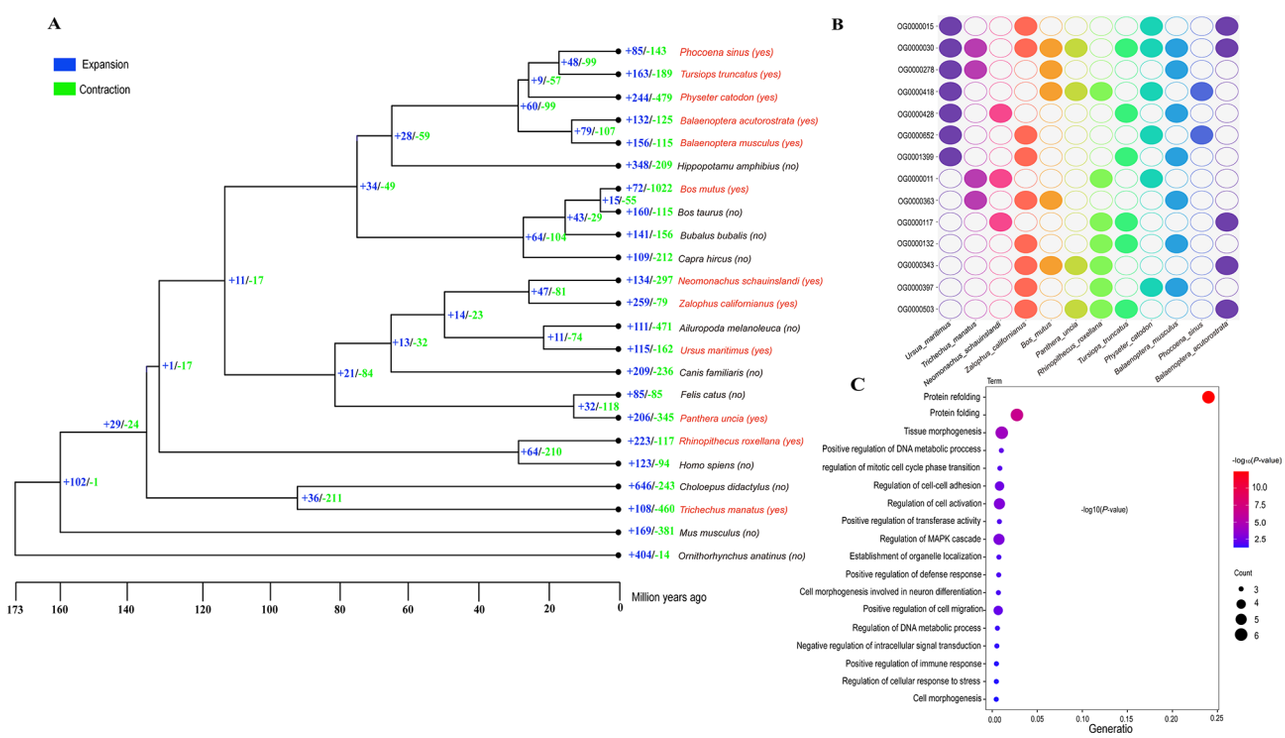

Figure 2 Gene family expansion and contraction in 23 mammalian species

A: Phylogenetic tree of 23 mammalian species displaying gene family expansions (blue) and contractions (green). Hypoxia-tolerant species are indicated in red and labeled “yes”, while closely related species are shown in black and labeled “no”. B: Significant gene family expansions (P<0.05) in hypoxia-tolerant mammals. X-axis represents hypoxia-tolerant species, Y-axis represents expanded gene families. Filled circles indicate significant expansions, unfilled circles denote no significant expansion. C: Functional annotation of significantly expanded gene families was performed using eggNOG-mapper v.2 (Cantalapiedra et al., 2021) and gene enrichment analysis of GO biological processes was performed using Metascape.

response in multicellular organisms (GO:0033555, 16 genes, *P*=0.0030), collagen biosynthesis (GO:0032965, eight genes, *P*=0.0165), development of lung vasculature (GO:0060426, three genes, *P*=0.0277), and cellular response to ultraviolet (UV) radiation (GO:0034644, 17 genes, *P*=0.0022) (Figure 3A). These findings provide key insights into the molecular mechanisms that drive adaptation in high-altitude and marine mammals.

In contrast to the large number of REGs, only 167 PSGs were identified across the 12 hypoxia-adapted species. Functional enrichment analysis of these PSGs identified several GO terms pertinent to pulmonary fibrosis formation and hypoxic adaptation, including organization of supramolecular fibers (GO:0097435, 13 genes, *P*=2.54074e-05), actin filament organization (GO:0007015, six genes, *P*=0.0044), microtubule cytoskeletal organization (GO:0000226, nine genes, *P*=0.0035), GTPase activity regulation (GO:0043087, eight genes, *P*=0.0003), actin filament-based process regulation (GO:0032970, nine genes, *P*=0.0003), supramolecular fiber organization regulation (GO:1902903, eight genes, *P*=0.0015), cell shape regulation (GO:0008360, four genes, *P*=0.0075), cellular response to UV (GO:0034644, 17 genes, *P*=0.0022), multicellular organismal response to stress (GO:0033555, 16 genes, *P*=0.0030), and cellular response to osmotic stress (GO:0071470, nine genes, *P*=0.0136) (Figure 3B). These enriched terms suggest convergent adaptive mechanisms that enable these species to adapt to challenging environments characterized by low oxygen and high UV exposure, which contribute to pulmonary fibrosis.

The genes *ZFP36L1*, *FN1*, and *NEDD9* exhibited convergent accelerated evolution and positive selection in both high-altitude and marine mammals (Supplementary Tables S3, S4). These genes play critical roles in lung adaptation to hypoxic environments, pulmonary fibrosis, and pulmonary hypertension (Loh et al., 2020; Samokhin et al., 2018; Sun et al., 2022; Xu et al., 2020). This implies that, in the evolution of hypoxia-adapted mammals, these genes have significantly contributed to producing abundant elastic fibers in the lungs, thereby facilitating adaptation to low-oxygen environments.

Convergent amino acid substitution

Fixed amino acid substitutions within specific mammalian groups can provide insights into the development of particular phenotypes (Chai et al., 2021; Guo et al., 2024). We identified 5 243 one-to-one orthologous genes using OrthoFinder v.2.5.5 and employed FasParser v.2 to designate hypoxia-tolerant species as the foreground branch in our analysis of genes that may have undergone convergent amino acid substitutions. This process yielded 50 genes showing such substitutions in hypoxia-tolerant species (Supplementary Table S5). GO enrichment analysis of these genes (Figure 4A) revealed several key terms associated with adaptive evolution, including cell-cell adhesion (GO:0098609, five genes, *P*=0.0017), supramolecular fiber organization (GO:0097435, four genes, *P*=0.0143), cilium or flagellum-dependent cell motility (GO:0001539, three genes, *P*=0.0015), and blood vessel development (GO:0001568, four genes, *P*=0.0101), which may be associated with pulmonary fibrosis phenotypes in hypoxia-tolerant species. Additional GO terms, such as cellular response to UV radiation (GO:0034644, three genes, *P*=0.0004), DNA damage-triggered signal transduction (GO:0042770, three genes, *P*=0.0016), and response to environmental stimuli (GO:0042770, three genes, *P*=0.0016), suggest adaptations to high-pressure and UV-intense environments. Furthermore, processes such as spermatid development (GO:0007286, four genes, *P*=0.0003) and differentiation (GO:0048515, four genes, *P*=0.0004) indicate convergent evolutionary trends related to reproductive functions in high-altitude and marine mammals, potentially linked to their hypoxia tolerance. Figure 4B displays convergent amino acid substitutions (leucine to isoleucine) in the *SLC26A3* gene across marine and high-altitude mammals, with assessment of the functional effects of amino acid replacement suggesting that mutations at this site could be damaging (Supplementary Table S5).

Convergent gene loss

To explore patterns of gene loss across multiple mammalian species in comparison to humans, we conducted a comprehensive analysis of gene deletions in 22 non-human mammalian species (Kirilenko et al., 2023). As shown in

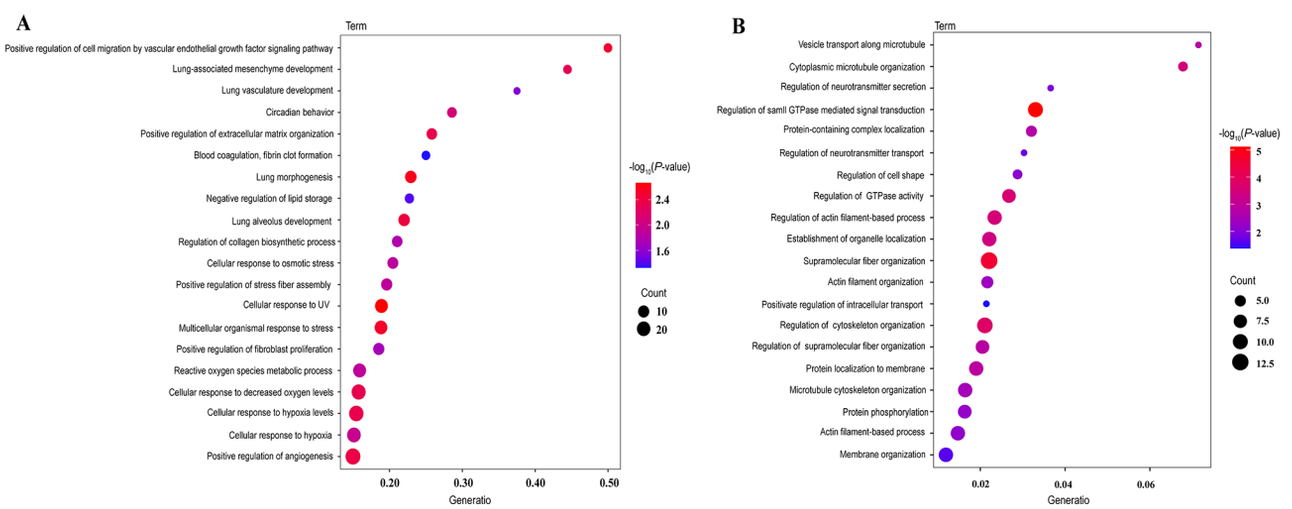

Figure 3 Functional enrichment of convergent accelerated evolution genes and convergent positive selection genes in GO biological processes

A: GO term enrichment analysis of convergent REGs. B: GO term enrichment analysis of convergent PSGs. Circles represent quantities of different genes, size of each circle is proportional to number of genes, color transition of circles from blue to red indicates an increase in $-\log_{10}$ (*P*-value).

Figure 5A, cetaceans lost 814 genes, other marine mammals lost 1 149 genes, and high-altitude mammals lost 1 201 genes. Among these, 304 genes were convergently lost in hypoxia-tolerant species (Figure 5B). GO enrichment analysis of these genes indicated significant enrichment in various biological processes, notably spermatogenesis (GO:0007283, 12 genes, *P*=0.0172), male gamete generation (GO:0048232, 12 genes, *P*=0.0207), down-regulation of adaptive immune response (GO:0002823, three genes, *P*=0.0188), protein deubiquitination (GO:0016579, five genes, *P*=0.0035), vesicle cargo loading (GO:0035459, three genes, *P*=0.0038), and ER-to-Golgi vesicle transport (GO:0006888, four genes, *P*=0.0320) (Figure 5C). We performed an in-depth analysis of the *CFAP47* gene, crucial for sperm morphogenesis and male fertility. Notably, variations in *CFAP47* are closely associated with male infertility. Pseudogenization of *CFAP47* was observed in several hypoxia-tolerant species, including the bottlenose dolphin (*T. truncatus*), Yangtze River dolphin (*L. vexillifer*), and pygmy sperm whale (*K. breviceps*), characterized by premature stop codons and base insertions, respectively (Supplementary Figure S3; Supplementary Table S6. PCR verification confirmed pseudogenization of *CFAP47* in the hypoxia-tolerant species (Supplementary Figure S3).

Evolutionary analyses of *CFAP47* showed an average ω of 0.131 across the tree (model A), significantly lower than 1.0 (model B; Supplementary Table S7), indicative of strong purifying selection acting on *CFAP47* during mammalian evolution. In model C (Supplementary Table S7), which allows two different ω values for pseudogenized (ω2=0.532) and functional (ω1=0.092) lineages, the data fit significantly better than the null model A, which assumes a single ω for all branches (Supplementary Table S7). This suggests that selective purifying pressure on *CFAP47* is markedly relaxed (increased ω) in hypoxia-tolerant species. Furthermore, the ω value in hypoxia-tolerant species (ω2=0.532) (model D vs. model C, *P*>0.05; Supplementary Table S7) indicates a shift toward neutral evolution (ω=1.0) in *CFAP47*. Model E, which allows ω variation in all branches, fit significantly better than model C, suggesting that ω varies significantly among non-hypoxia-tolerant species. Similarly, the RELAX model demonstrated that *CFAP47* evolved under relaxed selection (*K*<1 at *P*<0.001) in all hypoxia-tolerant species (Supplementary Table S8). This finding suggests potential alterations in the reproductive capabilities of these species under hypoxic conditions.

DISCUSSION

Throughout evolution, mammals residing in high altitudes and marine environments have needed to adapt to hypoxic conditions. This adaptation encompasses not only the functional optimization of hemoglobin but also significant physiological and anatomical modifications (Li et al., 2021a; Tian et al., 2017; Wu et al., 2020), particularly alterations in lung structure (Chen et al., 2006; Gao et al., 2022; Guo et al., 2024). The lungs, acting as a vital interface between the environment and organisms, exhibit extensive adaptive transformations in both high-altitude and marine mammals (Tian et al., 2016), including an abundance of loosely arranged elastic fibers within a distinctive lung architecture, crucial for efficient gas exchange in hypoxic conditions (Chen et al., 2006; Gao et al., 2022; Guo et al., 2024; Piscitelli et al., 2013). From a molecular standpoint, we employed a range of

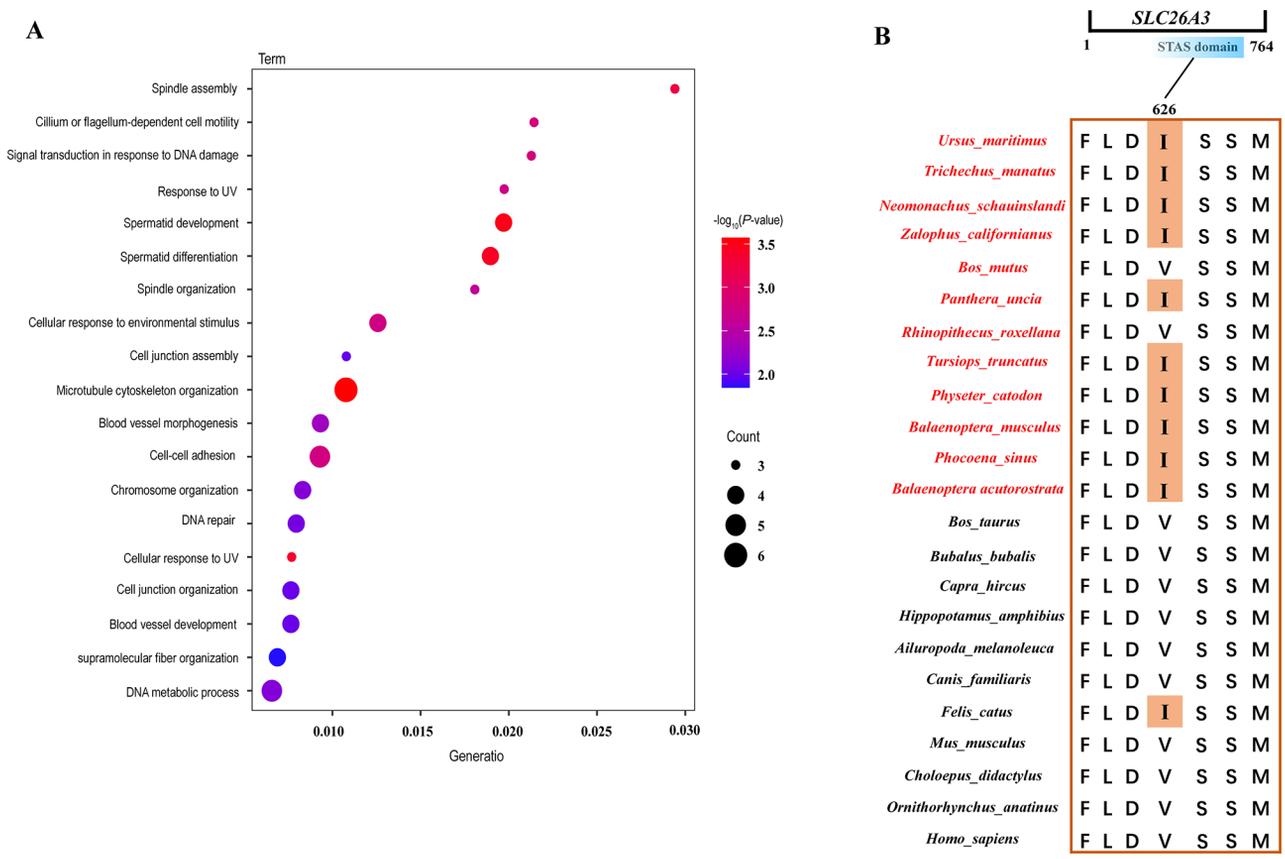

Figure 4 Convergent amino acid substitution and GO enrichment analysis

A: GO term enrichment analysis of convergent amino acid substitution genes. B: Convergent amino acid substitution sites in *SLC26A3* gene.

convergent evolutionary analyses, including gene family expansion and contraction, to identify various genes and gene families that have undergone significant adaptive evolution in hypoxia-tolerant mammals. These genetic adaptations provide critical insights into the molecular frameworks underlying respiratory efficiency in low-oxygen environments. These instances of convergent evolution underscore the shared molecular pathways that organisms utilize to overcome analogous physiological challenges in different settings, highlighting the complexity and diversity of biological adaptation strategies in response to environmental extremes.

Convergent patterns of gene family expansion and contraction

Prior research on convergent evolution has shown that gene family expansion and contraction primarily result from directional changes in response to selective pressures, including environmental factors and dietary habits (Freitas & Nery, 2020; Lyu et al., 2022; Rogers et al., 2022). This study examined gene family expansion and contraction in three plateau-dwelling mammals and their closely related species (Figure 2A). Results revealed that the extent of gene family changes in these high-altitude mammals was significantly more pronounced than in their close relatives, suggesting that the intense selective pressure exerted by the plateau environment has driven extensive gene family modifications in order to adapt to high-altitude conditions. Conversely, among marine mammals, only the sperm whale (*P. catodon*) and manatee (*T. manatus*) exhibited notable gene family contractions compared to their close relatives, with no significant gene family changes observed in other species (Figure 2A).

To further investigate whether the observed gene family expansions and contractions are crucial for adaptation to high-altitude and marine environments, we conducted a convergent analysis of gene family changes. Given the considerable phylogenetic distance between plateau and marine mammals

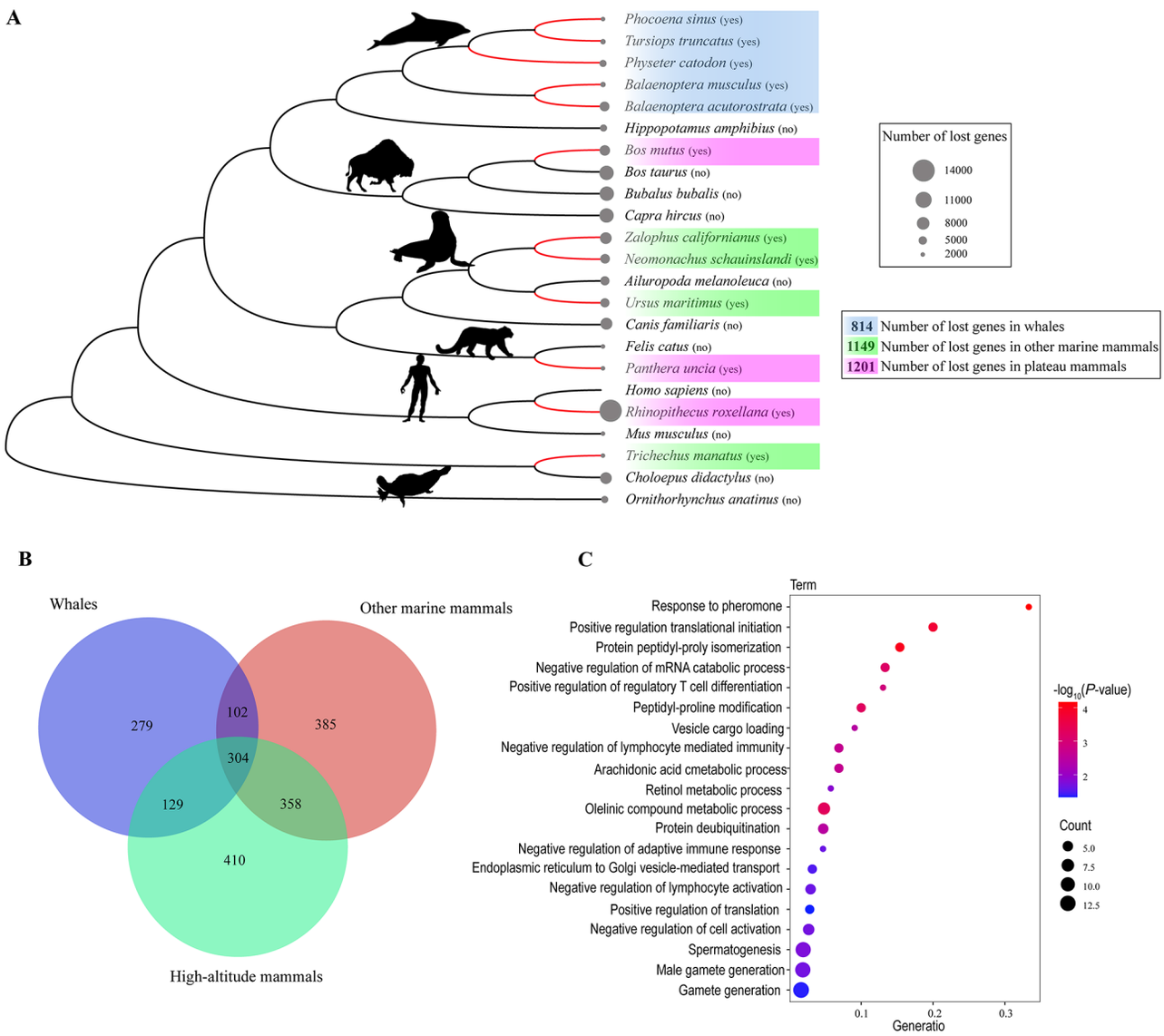

Figure 5 Convergent gene loss and GO term enrichment analysis

A: Number of gene losses in whales, other marine mammals, high-altitude mammals, and their close relatives. Circle sizes represent quantity of gene loss, with blue circles for whales, green for other marine mammals, pink for high-altitude mammals, and colorless for close relatives. Animal silhouettes were obtained from PhyloPic (https://www.phylopic.org/). B: Venn diagram of number of genes commonly lost among cetaceans, other marine mammals, and high-altitude mammals. C: GO term enrichment analysis of convergent loss genes.

(e.g., cetaceans and pinnipeds), we initially identified gene families with significant expansion or contraction in each species, using a significance threshold of $P<0.05$. Subsequently, we focused on gene families with common expansion or contraction in at least one-third of the studied plateau and marine mammals (Figure 2B; Supplementary Figure S1), which were then subjected to annotation enrichment analysis (Figure 2C; Supplementary Figure S2).

Analysis of the 12 hypoxia-tolerant mammals revealed a higher prevalence of gene family contractions compared to expansions (Figure 2B; Supplementary Figure S1). These findings align with the "*less-is-more hypothesis*" (Albalat & Cañestro, 2016; Olson, 1999), which posits that gene loss is a significant source of genetic variation, fostering adaptive phenotypic diversity. Further enrichment analysis revealed that both the expansion and contraction of various gene families were closely linked to cellular responses to hypoxia, signal transduction regulation, cell structure stability, and formation and maintenance of elastic fibers in lung tissue. These gene family modifications were enriched in several important GO terms, including cell adhesion (GO:0098609, five genes, $P=0.0005$), cellular stress response regulation (GO:0080135, three genes, $P=0.0465$), cell morphogenesis (GO:0000902, three genes, $P=0.0498$), intermediate filament organization (GO: 0045109, seven genes, $P=2.61473e-09$), and epithelial cell differentiation (GO: 0030855, 11 genes, $P=3.37805e-06$) (Figure 2C; Supplementary Figure S2). Notably, we identified that certain members of the Keratin gene family (e.g., *KRT14*, *KRT17*) have undergone convergent evolution in hypoxia-tolerant species (Ficial et al., 2014; Habermann et al., 2020; Jaeger et al., 2022). This is consistent with the observed up-regulation of these genes in lung cells under pathological conditions such as idiopathic pulmonary fibrosis (Ficial et al., 2014; Habermann et al., 2020; Jaeger et al., 2022). Experimental studies on mice have demonstrated that differential regulation of *KRT14* is associated with lung injury and pulmonary fibrosis (Ievlev et al., 2023). Similarly, research on airway basal cells has indicated that *KRT17* plays a pivotal role in the development of idiopathic pulmonary fibrosis (Jaeger et al., 2022). Collectively, these findings suggest that the Keratin gene family has undergone adaptive evolution in both marine and high-altitude mammals, potentially contributing to the emergence of a distinct fibrotic phenotype in the lungs. This molecular adaptation may be crucial for enhancing survival in low-oxygen environments.

Convergent adaptation patterns of REGs and PSGs

Positive selection and rapid gene evolution are key aspects of biological adaptation, driven by the accumulation of genetic variation in response to environmental changes (Aguileta et al., 2009). In this study, we explored genes that have undergone convergent rapid evolution and positive selection in both plateau and marine mammals, which face similar anoxic environments. In total, 2 683 genes showing convergent accelerated evolution in both plateau and marine mammals were identified. Subsequent analysis indicated that these genes were predominantly involved in pathways related to hypoxic adaptation, lung morphology, vascular development, pulmonary elastic fiber formation, and cellular responses to UV radiation and stress (Figure 3A, B). For example, the *ZFP36L1* gene, a post-transcriptional regulator, plays a vital role in hypoxic signaling and the regulation of abnormal cell cycle progression (Loh et al., 2020) (Supplementary Table S3). The *FN1* gene, implicated in pulmonary fibrosis, has been shown to decelerate fibrotic progression when its major functional domains are inhibited in fibrosis models (Sun et al., 2022; Xu et al., 2020) (Supplementary Table S3). In summary, these genes associated with hypoxic environments, lung morphology, vascular development, and pulmonary elastic fiber formation may reflect convergent adaptive mechanisms to hypoxic conditions in plateau and marine mammals.

In addition to REGs, we identified 167 PSGs exhibiting convergent positive selection signatures in hypoxia-tolerant mammals. Functional enrichment analysis of these PSGs revealed several GO terms relevant to pulmonary fibrosis formation and hypoxic adaptation, including actin filament organization (GO:0007015, six genes, $P=0.0045$), microtubule cytoskeletal organization (GO:0000226, nine genes, $P=0.0035$), and GTPase activity regulation (GO:0043087, eight genes, $P=0.0003$). Additionally, *NEDD9* plays a crucial role in the pathogenesis of pulmonary arterial hypertension (PAH) (Hassoun, 2021), particularly in promoting vascular fibrosis (Samokhin et al., 2018) (Supplementary Table S4). Elevated *NEDD9* expression in the lung endothelium has also been observed in human lung diseases characterized by hypoxia and pulmonary thrombosis (Alba et al., 2021, 2022; Samokhin et al., 2018) (Supplementary Table S4). Thus, the convergent positive selection of *NEDD9* in hypoxia-tolerant species likely represents an important adaptive mechanism to hypoxic environments through the development of pulmonary fibrosis. These genes exhibit signs of convergent positive selection and rapid evolution in both marine and high-altitude mammals, suggesting their involvement in pulmonary fibrosis development and adaptation to low-oxygen environments.

Convergent adaptation patterns of amino acid substitutions and gene loss

Fixed amino acid substitutions and gene losses in certain mammalian groups are associated with distinct phenotypes that arise in response to unique environmental challenges (Albalat & Cañestro, 2016; Huang et al., 2021). In mammals inhabiting high-altitude and marine hypoxic environments, habitat specificity has precipitated the evolution of specialized reproductive strategies. For example, compared to low-altitude mammals, high-altitude mammals exhibit adaptations in reproductive cycles and decreased litter sizes, producing fewer, yet more mature and resilient offspring (Robertson & Wilsterman, 2020). Similarly, marine mammals living in hypoxic environments have evolved comparable reproductive adaptations (Pomeroy, 2011; Rendell et al., 2019; Zhou et al., 2018b), essential for survival of their offspring in harsh environments. Notably, marine mammals, such as whales, exhibit cryptorchidism, an evolutionary trait that enhances fertilization and reproduction efficiency in aquatic settings (Chai et al., 2021; Ding et al., 2021). This adaptation not only facilitates underwater reproduction but also serves as a physiological response to cold seawater environments. Our study identified 50 genes with convergent amino acid substitutions and 304 genes with evidence of gene loss across hypoxia-tolerant species (Figure 5A; Supplementary Table S5). Enrichment analysis revealed significant involvement of these genes in lung development, lung fiber formation, and blood vessel development, as well as in sperm differentiation, development, and male gametogenesis (Figures 4A, 5C). These findings suggest a potential link between genetic

adaptations and spermatogenesis in hypoxic environments.

Based on convergent evolution analysis of high-altitude and marine mammals in hypoxic environments, our study revealed significant convergent evolution in spermatogenesis-related genes. These genetic adaptations play a pivotal role in ensuring successful reproduction under conditions of oxygen scarcity. We identified convergent amino acid substitutions in genes integral to sperm development, differentiation, and ciliary motion (Figure 4A). For example, the *SLC26A3* gene, previously linked to impaired epididymal development and reduced sperm fertilization capacity in mouse models (El Khouri et al., 2018), was found to have a substitution at position 626, replacing valine with isoleucine, in hypoxia-tolerant species. Functional assessment of this amino acid replacement indicated potential detrimental effects (Figure 4B; Supplementary Table S5). Located in the STAS domain, this mutation may influence spermatogenesis in both plateau and marine mammals (Wedenoja et al., 2017). Furthermore, we identified the loss of several genes related to sperm formation and vesicle transport in hypoxia-tolerant mammals (Figure 5C), including the *CFAP47* gene, crucial for sperm morphology and male fertility (Liu et al., 2021, 2023). Using bioinformatics and PCR validation, we detected pseudogenization of *CFAP47* among high-altitude and marine mammals. Specifically, a 1 bp insertion at position 224 in *T. truncatus*, premature stop codons at position 61 in *L. vexillifer*, and a 1 bp insertion at position 433 in *K. breviceps* were observed (Supplementary Figure S3). These genetic alterations in *CFAP47* may impact spermatogenesis in hypoxia-tolerant mammals. Research suggests that captive male *P. uncia* exhibit significantly reduced fecundity (Herrick et al., 2020). This decline in fertility is not only due to the adverse effects of inbreeding but also to the pseudogenization of *CFAP47*, which plays a crucial role in spermatogenesis. The loss of function in this gene likely contributes to the lower reproductive success observed in these animals (Herrick et al., 2020). Furthermore, mutations in *CFAP47* have also been implicated in respiratory anomalies observed in primary ciliary dyskinesia patients (Ge et al., 2024), highlighting its dual role in reproductive and respiratory adaptation. These findings suggest that *CFAP47* pseudogenization is integral to the adaptive responses related to hypoxic environments, affecting both respiration and reproduction in marine and high-altitude mammals (Ge et al., 2024; Herrick et al., 2020). Overall, our findings establish a robust link between various key genes and spermatogenesis, suggesting that reproductive adaptations are essential for their survival in low-oxygen habitats. Moreover, our results support the previously proposed "*training hypothesis*", which suggests that testicular descent can reduce the blood supply to mature sperm cells (Freeman, 1990), thereby inducing adaptations in sperm mitochondria to augment their oxidative metabolism, akin to how muscle cells adapt under oxygen stress. This improved aerobic efficiency may confer a competitive advantage by optimizing sperm performance in hypoxic conditions (Freeman, 1990).

In conclusion, this study explored the genetic mechanisms underlying the adaptations of marine and high-altitude mammals to hypoxic conditions through convergent evolution of a pulmonary fibrosis phenotype. Through multiple analytical methods, we demonstrated that gene family dynamics —characterized by expansions, contractions, and shifts in selective pressure, including positive selection and accelerated evolution —played a critical role in shaping the convergent evolution of the pulmonary fibrosis phenotype, enabling these mammals to survive in low-oxygen environments. Additionally, we identified several evolutionary events linked to spermatogenesis and reproductive capacity, informed by convergent amino acid substitutions and gene losses. These genetic changes suggest a possible concurrent evolution of reproductive systems, further enhancing the environmental adaptability of these species. Overall, our findings provide a deeper understanding of the adaptive evolution of lung morphology in these species, offering valuable insights for developing interventions for diseases impacting human reproductive capacity.

DATA AVAILABILITY

Orthologous sequences for 5 243 genes across 23 species, along with data on REGs, PSGs, and the *CFAP47* gene, are available in the Figshare dataset (https://doi.org/10.6084/m9.figshare.25966309.v1).

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

W.H.R., G.Y., and S.X.X. conceived the study and assisted with manuscript revision. B.X.G., Y.Z., X.Y.S., and W.J.L. collected the data, B.X.G. performed the molecular evolutionary analysis and wrote the paper. Y.X.S. conducted PCR experiments. All authors read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

We thank members of Jiangsu Key Laboratory for Biodiversity and Biotechnology for suggestions and support during the project. We thank Dr. Si-Min Chai, Dr. Tian-Zhen Wu, and Dr. Xu Zhou for their helpful suggestions. We are particularly grateful to Professor Michael Hiller for his assistance in gene loss data.

REFERENCES

Adzhubei I, Jordan DM, Sunyaev SR. 2013. Predicting functional effect of human missense mutations using PolyPhen-2. *Current Protocols in Human Genetics*, **76**(1): 7.20. 1–7.20. 41.

Aguileta G, Refrégier G, Yockteng R, et al. 2009. Rapidly evolving genes in pathogens: methods for detecting positive selection and examples among fungi, bacteria, viruses and protists. *Infection, Genetics and Evolution*, **9**(4): 656–670.

Alba GA, Samokhin AO, Wang RS, et al. 2021. NEDD9 is a novel and modifiable mediator of platelet–endothelial adhesion in the pulmonary circulation. *American Journal of Respiratory and Critical Care Medicine*, **203**(12): 1533–1545.

Alba GA, Samokhin AO, Wang RS, et al. 2022. Pulmonary endothelial NEDD9 and the prothrombotic pathophenotype of acute respiratory distress syndrome due to SARS-CoV-2 infection. *Pulmonary Circulation*, **12**(2): e12071.

Albalat R, Cañestro C. 2016. Evolution by gene loss. *Nature Reviews Genetics*, **17**(7): 379–391.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, **57**(1): 289–300.

Boyd RB. 1975. A gross and microscopic study of the respiratory anatomy of the Antarctic Weddell seal. *Leptonychotes weddelli. Journal of Morphology*, **147**(3): 309–335.

Cantalapiedra CP, Hernández-Plaza A, Letunic I, et al. 2021. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Molecular Biology and Evolution*, **38**(12): 5825–5829.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, **17**(4): 540–552.

Chai SM, Huang X, Wu TZ, et al. 2021. Comparative genomics reveals molecular mechanisms underlying health and reproduction in cryptorchid mammals. *BMC Genomics*, **22**(1): 763.

Chen QS, Feng X, Jiang SC. 2006. Structural study on plateau adaptability of yak lung. *Scientia Agricultura Sinica*, **39**(10): 2107–2113. (in Chinese)

Ding XY, Cao L, Zheng Y, et al. 2021. Insights into the evolution of spermatogenesis-related ubiquitin–proteasome system genes in abdominal testicular Laurasiatherians. *Genes*, **12**(11): 1780.

El Khouri E, Whitfield M, Stouvenel L, et al. 2018. Slc26a3 deficiency is associated with epididymis dysplasia and impaired sperm fertilization potential in the mouse. *Molecular Reproduction and Development*, **85**(8-9): 682–695.

Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, **20**(1): 238.

Fahlman A. 2024. Cardiorespiratory adaptations in small cetaceans and marine mammals. *Experimental Physiology*, **109**(3): 324–334.

Fahlman A, Moore MJ, Garcia-Parraga D. 2017. Respiratory function and mechanics in pinnipeds and cetaceans. *Journal of Experimental Biology*, **220**(10): 1761–1773.

Falke KJ, Hill RD, Qvist J, et al. 1985. Seal lungs collapse during free diving: Evidence from arterial nitrogen tensions. *Science*, **229**(4713): 556–558.

Favilla AB, Costa DP. 2020. Thermoregulatory strategies of diving air-breathing marine vertebrates: a review. *Frontiers in Ecology and Evolution*, **8**: 555509.

Ficial M, Antonaglia C, Chilosi M, et al. 2014. Keratin-14 expression in pneumocytes as a marker of lung regeneration/repair during diffuse alveolar damage. *American Journal of Respiratory and Critical Care Medicine*, **189**(9): 1142–1145.

Foley NM, Springer MS, Teeling EC. 2016. Mammal madness: is the mammal tree of life not yet resolved?. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **371**(1699): 20150140.

Freeman S. 1990. The evolution of the scrotum: a new hypothesis. *Journal of Theoretical Biology*, **145**(4): 429–445.

Freitas L, Nery MF. 2020. Expansions and contractions in gene families of independently-evolved blood-feeding insects. *BMC Evolutionary Biology*, **20**(1): 87.

Gao X, Wang S, Wang YF, et al. 2022. Long read genome assemblies complemented by single cell RNA-sequencing reveal genetic and cellular mechanisms underlying the adaptive evolution of yak. *Nature Communications*, **13**(1): 4887.

Ge HJ, Zhou WJ, He M, et al. 2024. Mutations in *CFAP47*, a previously reported MMAF causative gene, also contribute to the respiratory defects in patients with PCD. *Molecular Genetics & Genomic Medicine*, **12**(1): e2278.

Glazko GV, Nei M. 2003. Estimation of divergence times for major lineages of primate species. *Molecular Biology and Evolution*, **20**(3): 424–434.

Guo BX, Sun YX, Wang YH, et al. 2024. Evolutionary genetics of pulmonary anatomical adaptations in deep-diving cetaceans. *BMC Genomics*, **25**(1): 339.

Habermann AC, Gutierrez AJ, Bui LT, et al. 2020. Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Science Advances*, **6**(28): eaba1972.

Hassoun PM. 2021. Pulmonary arterial hypertension. *New England Journal of Medicine*, **385**(25): 2361–2376.

Henk WG, Haldiman JT. 1990. Microanatomy of the lung of the bowhead whale *Balaena mysticetus*. *The Anatomical Record*, **226**(2): 187–197.

Herrick JR, Iske CJ, Santymire RM, et al. 2020. Factors affecting reproductive traits in male snow leopards (*Unciauncia*). *Reproduction and Fertility*, **1**(1): 35–49.

Huang X, Sun D, Wu TZ, et al. 2021. Genomic insights into body size evolution in Carnivora support Peto's paradox. *BMC Genomics*, **22**(1): 429.

Ievlev V, Lynch TJ, Freischlag KW, et al. 2023. Krt14 and Krt15 differentially regulate regenerative properties and differentiation potential of airway basal cells. *JCI Insight*, **8**(2): e162041.

Isogai Y, Imamura H, Nakae S, et al. 2021. Common and unique strategies of myoglobin evolution for deep-sea adaptation of diving mammals. *iScience*, **24**(8): 102920.

Jaeger B, Schupp JC, Plappert L, et al. 2022. Airway basal cells show a dedifferentiated KRT17^{high}Phenotype and promote fibrosis in idiopathic pulmonary fibrosis. *Nature Communications*, **13**(1): 5637.

Kirilenko BM, Munegowda C, Osipova E, et al. 2023. Integrating gene annotation with orthology inference at scale. *Science*, **380**(6643): eabn3107.

Kooyman GL. 1973. Respiratory adaptations in marine mammals. *American Zoologist*, **13**(2): 457–468.

Korthauer K, Kimes PK, Duvallet C, et al. 2019. A practical guide to methods controlling false discoveries in computational biology. *Genome Biology*, **20**(1): 118.

Kumar S, Stecher G, Suleski M, et al. 2017. TimeTree: a resource for timelines, timetrees, and divergence times. *Molecular Biology and Evolution*, **34**(7): 1812–1819.

Lee P, Chandel NS, Simon MC. 2020. Cellular adaptation to hypoxia through hypoxia inducible factors and beyond. *Nature Reviews Molecular Cell Biology*, **21**(5): 268–283.

Lewis F, Butler A, Gilbert L. 2011. A unified approach to model selection using the likelihood ratio test. *Methods in Ecology and Evolution*, **2**(2): 155–162.

Li F, Qiao ZL, Duan QJ, et al. 2021a. Adaptation of mammals to hypoxia. *Animal Models and Experimental Medicine*, **4**(4): 311–318.

Li JY, Meng XQ, Wang LH, et al. 2021b. Changes in the expression levels of elastic fibres in yak lungs at different growth stages. *BMC Developmental Biology*, **21**(1): 9.

Li MK, Pan D, Sun H, et al. 2021c. The hypoxia adaptation of small mammals to plateau and underground burrow conditions. *Animal Models and Experimental Medicine*, **4**(4): 319–328.

Liu CY, Tu CF, Wang LB, et al. 2021. Deleterious variants in X-linked *CFAP47* induce asthenoteratozoospermia and primary male infertility. *The American Journal of Human Genetics*, **108**(2): 309–323.

Liu MH, Dai SY, Zhang JY, et al. 2023. A novel mutation in *CFAP47* causes male infertility due to multiple morphological abnormalities of the sperm flagella. *Frontiers in Endocrinology*, **14**: 1155639.

Loh XY, Sun QY, Ding LW, et al. 2020. RNA-binding protein *ZFP36L1* suppresses hypoxia and cell-cycle signaling. *Cancer Research*, **80**(2): 219–233.

Löytynoja A. 2014. Phylogeny-aware alignment with PRANK. *In*: Russell DJ. Multiple Sequence Alignment Methods. Totowa: Humana, 155–170.

Lyu T, Zhou SY, Fang JH, et al. 2022. Convergent genomic signatures of high-altitude adaptation among six independently evolved mammals. *Animals*, **12**(24): 3572.

Ma M, Pu P, Niu ZY, et al. 2023. A novel mechanism for high-altitude adaptation in hemoglobin of black-spotted frog (*Pelophylax nigromaculatus*). *Frontiers in Ecology and Evolution*, **11**: 1103406.

Mendes FK, Vanderpool D, Fulton B, et al. 2020. CAFE 5 models variation in evolutionary rates among gene families. *Bioinformatics*, **36**(22-23): 5516–5518.

Nikaido M, Rooney AP, Okada N. 1999. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interpersed elements: hippopotamuses are the closest extant relatives of whales. *Proceedings of the National Academy of Sciences of the United States of America*, **96**(18): 10261–10266.

Olson MV. 1999. When less is more: gene loss as an engine of evolutionary change. *The American Journal of Human Genetics*, **64**(1): 18–23.

Piscitelli MA, Raverty SA, Lillie MA, et al. 2013. A review of cetacean lung morphology and mechanics. *Journal of Morphology*, **274**(12): 1425–1440.

Pomeroy P. 2011. Reproductive cycles of marine mammals. *Animal Reproduction Science*, **124**(3-4): 184–193.

Pu P, Lu SS, Niu ZY, et al. 2019. Oxygenation properties and underlying molecular mechanisms of hemoglobins in plateau zokor (*Eospalax baileyi*). *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, **317**(5): R696–R708.

Rambaut A, Drummond AJ, Xie D, et al. 2018. Posterior summarization in Bayesian phylogenetics using tracer 1.7. *Systematic Biology*, **67**(5): 901–904.

Ranwez V, Harispe S, Delsuc F, et al. 2011. MACSE: multiple alignment of coding sequences accounting for frameshifts and stop codons. *PLoS One*, **6**(9): e22594.

Rendell L, Cantor M, Gero S, et al. 2019. Causes and consequences of female centrality in cetacean societies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **374**(1780): 20180066.

Robertson CE, Wilsterman K. 2020. Developmental and reproductive physiology of small mammals at high altitude: challenges and evolutionary innovations. *Journal of Experimental Biology*, **223**(24): jeb215350.

Rogers LW, Koehler AM, Crouch JA, et al. 2022. Comparative genomic analysis reveals contraction of gene families with putative roles in pathogenesis in the fungal boxwood pathogens *Calonectria henricotiae* and *C. Pseudonaviculata. BMC Ecology and Evolution*, **22**(1): 79.

Samokhin AO, Stephens T, Wertheim BM, et al. 2018. NEDD9 targets *COL3A1* to promote endothelial fibrosis and pulmonary arterial hypertension. *Science Translational Medicine*, **10**(445): eaap7294.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**(9): 1312–1313.

Storz JF. 2007. Hemoglobin function and physiological adaptation to hypoxia in high-altitude mammals. *Journal of Mammalogy*, **88**(1): 24–31.

Storz JF, Bautista NM. 2022. Altitude acclimatization, hemoglobin-oxygen affinity, and circulatory oxygen transport in hypoxia. *Molecular Aspects of Medicine*, **84**: 101052.

Sun J, Guo YY, Chen TT, et al. 2022. Systematic analyses identify the anti-fibrotic role of lncRNA TP53TG1 in IPF. *Cell Death & Disease*, **13**(6): 525.

Sun YB. 2018. FasParser2: a graphical platform for batch manipulation of tremendous amount of sequence data. *Bioinformatics*, **34**(14): 2493–2495.

Tian R, Wang ZF, Niu X, et al. 2016. Evolutionary genetics of hypoxia tolerance in cetaceans during diving. *Genome Biology and Evolution*, **8**(3): 827–839.

Tian R, Yin DQ, Liu YZ, et al. 2017. Adaptive evolution of energy metabolism-related genes in hypoxia-tolerant mammals. *Frontiers in Genetics*, **8**: 205.

Weaver S, Shank SD, Spielman SJ, et al. 2018. Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. *Molecular Biology and Evolution*, **35**(3): 773–777.

Wedenoja S, Khamaysi A, Shimshilashvili L, et al. 2017. A missense mutation in *SLC26A3* is associated with human male subfertility and impaired activation of CFTR. *Scientific Reports*, **7**(1): 14208.

Wertheim JO, Murrell B, Smith MD, et al. 2015. RELAX: detecting relaxed selection in a phylogenetic framework. *Molecular Biology and Evolution*, **32**(3): 820–832.

Withers PC, Cooper CE, Maloney SK, et al. 2016. Physiological adaptations to extreme environments. *In*: Withers PC, Cooper CE, Maloney S K, et al. Ecological and Environmental Physiology of Mammals. Oxford: Oxford University Press, 290–392.

Wu DD, Yang CP, Wang MS, et al. 2020. Convergent genomic signatures of high-altitude adaptation among domestic mammals. *National Science Review*, **7**(6): 952–963.

Xu JC, Xu XY, Jiang LN, et al. 2020. SARS-CoV-2 induces transcriptional signatures in human lung epithelial cells that promote lung fibrosis. *Respiratory Research*, **21**(1): 182.

Yang YN, Yuan HN, Yang TL, et al. 2021. The expression regulatory network in the lung tissue of Tibetan pigs provides insight into hypoxia-sensitive pathways in high-altitude hypoxia. *Frontiers in Genetics*, **12**: 691592.

Yang ZH. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, **24**(8): 1586–1591.

Yépez Y, Marcano-Ruiz M, Bortolini MC. 2023. Adaptive strategies of aquatic mammals: exploring the role of the HIF pathway and hypoxia tolerance. *Genetics and Molecular Biology*, **46**(3 Suppl 1): e20230140.

Zhang XZ, Fu L, Zou XY, et al. 2022. Lung transcriptome analysis for the identification of genes involved in the hypoxic adaptation of plateau pika (*Ochotona curzoniae*). *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, **41**: 100943.

Zhao PF, Zhao FF, Hu J, et al. 2022. Physiology and transcriptomics analysis reveal the contribution of lungs on high-altitude hypoxia adaptation in Tibetan sheep. *Frontiers in Physiology*, **13**: 885444.

Zhou XM, Guang XM, Sun D, et al. 2018a. Population genomics of finless porpoises reveal an incipient cetacean species adapted to freshwater. *Nature Communications*, **9**(1): 1276.

Zhou XM, Sun D, Guang XM, et al. 2018b. Molecular footprints of aquatic adaptation including bone mass changes in cetaceans. *Genome Biology and Evolution*, **10**(3): 967–975.

Zhou YY, Zhou B, Pache L, et al. 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications*, **10**(1): 1523.