

Article

Open Access

Genomic insights into the convergent evolution of desert adaptation in camels and antelopes

Jiong Zhou^{1,✉}, Xiao-Fang Zhou^{1,✉}, Hui-Shan Yue^{1,✉}, Wu Chen^{2,✉}, Bin Li¹, Bo-Tong Zhou¹, Zi-He Li¹, Ze-Cheng Du¹, Yi-Fan Mao¹, Wen Wang¹, Dong-Dong Wu³, Ge Han^{1,*}, Bao Wang^{1,*}, Lei Chen^{1,*}

¹ Shaanxi Key Laboratory of Qinling Ecological Intelligent Monitoring and Protection, School of Ecology and Environment, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China

² Guangzhou Zoo and Guangzhou Wildlife Research Center, Guangzhou, Guangdong 510070, China

³ State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China

ABSTRACT

Extreme heat and chronic water scarcity present formidable challenges to large desert-dwelling mammals. In addition to camels, antelopes within the Hippotraginae and Alcelaphinae subfamilies also exhibit remarkable physiological and genetic specializations for desert survival. Among them, the critically endangered addax (*Addax nasomaculatus*) represents the most desert-adapted antelope species. However, the evolutionary and molecular mechanisms underlying desert adaptations remain largely unexplored. Herein, a high-quality genome assembly of the addax was generated to investigate the molecular evolution of desert adaptation in camels and desert antelopes. Comparative genomic analyses identified 136 genes harboring convergent amino acid substitutions implicated in crucial biological processes, including water reabsorption, fat metabolism, and stress response. Notably, a convergent R146S amino acid mutation in the prostaglandin EP2 receptor gene *PTGER2* significantly reduced receptor activity, potentially facilitating large-mammal adaptation to arid environments. Lineage-specific innovations were also identified in desert antelopes, including previously uncharacterized conserved non-coding elements. Functional assays revealed that several of these elements exerted significant regulatory effects *in vitro*, suggesting potential roles in adaptive gene expression. Additionally, signals of introgression and variation in genetic load were observed, indicating their possible influence on desert adaptation. These findings provide insights into the sequential evolutionary processes that drive physiological resilience in arid environments and

highlight the importance of convergent evolution in shaping adaptive traits in large terrestrial mammals.

Keywords: Desert adaptation; Convergent evolution; *Addax nasomaculatus*

INTRODUCTION

Deserts cover over 30% of the Earth's terrestrial surface and represent some of the most physiologically demanding environments on the planet. Characterized by chronic water scarcity, extreme temperature fluctuations, intense solar radiation, and nutrient limitations, these ecosystems exert powerful selective pressures on resident fauna (Ward, 2016). Despite these challenges, a diverse range of organisms, including reptiles, birds, and mammals, have independently evolved strategies to persist in arid environments. Among mammals, camels and desert-dwelling antelopes exemplify large-bodied species that exhibit highly specialized adaptations for desert survival. Camels, long recognized for their ability to endure extreme heat and aridity, have become an emblem of desert endurance. Similarly, several African antelope species, including the addax and oryx in the subfamily Hippotraginae and blue wildebeest (*Connochaetes taurinus*) and topi (*Damaliscus lunatus*) in the subfamily Alcelaphinae, display extraordinary physiological and behavioral traits that permit survival under prolonged aridity. These species can survive many months without direct water intake, relying instead on dietary and metabolic water to meet hydration needs (Kingdon, 2014).

Desert antelopes and camels evolved similar morphological and physiological characteristics to adapting deserts. Although morphological convergence—such as the evolution of wide,

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright ©2025 Editorial Office of Zoological Research, Kunming Institute of Zoology, Chinese Academy of Sciences

Received: 10 December 2024; Accepted: 26 March 2025; Online: 27 March 2025

Foundation items: This work was supported by the National Key R&D Program of China (2022YFF1000100), Shaanxi Program for Support of Top-notch Young Professionals, and Fundamental Research Funds for the Central Universities

*Authors contributed equally to this work

*Corresponding authors, E-mail: gehan1992@163.com; wangbao0716@163.com; chen_lei@nwpu.edu.cn

flattened hooves for improved locomotion on sandy substrates—is evident, it is the physiological parallels between camels and desert antelopes that most strongly suggest adaptive convergences. Both groups demonstrate remarkable water-conserving capabilities, including the production of highly concentrated urine and desiccated feces. For instance, dromedaries can concentrate urine to 3 170 mOsmol/kg H₂O, and gemsbok up to 2 900 mOsmol/kg H₂O, significantly higher than most similarly sized animals (Beuchat, 1990; Rocha et al., 2021a). These species also exhibit exceptional thermal tolerance, enabling water conservation through reduced need for sweating. Camels can withstand temperatures exceeding 42°C and tolerate water losses exceeding 25% of their total body weight (Schmidt-Nielsen, 1959), while oryxes can allow core body temperatures to rise to 47°C, enabling survival in extreme heat. Recent comparative genomic and transcriptomic analyses have further revealed complex adaptive features related to desert environments, including regulatory networks involved in lipid and water homeostasis, mechanisms of osmoregulation and osmoprotection, and molecular responses to thermal stress, dehydration, and ultraviolet (UV) radiation (Alvira-Iraizoz et al., 2021; Wu et al., 2014). Despite these insights, whether camels and desert antelopes share common genetic architectures underlying their convergent adaptations remains unresolved. Elucidating the molecular basis of these traits would deepen understanding of adaptive evolution and may inform strategies for enhancing environmental resilience in domesticated species.

Among desert antelopes, the addax (*Addax nasomaculatus*), the sole representative of its genus, exhibits some of the most extreme adaptations to hostile desert conditions (Krausman & Casey, 2007). Native to the Sahara and Sahel, the addax is capable of surviving in extreme arid conditions where annual precipitation is negligible. It is also among the most critically endangered hoofed mammals, with an estimated wild population of only 30–90 individuals (IUCN, 2024), raising concerns about inbreeding depression and the accumulation of deleterious mutations that may impact its desert adaptability. Both the addax and scimitar oryx (*Oryx dammah*) possess distinctive white skin, potentially helping to counteract the detrimental effects of solar radiation. However, the genetic and physiological mechanisms underlying these traits remain poorly understood.

In this study, a high-quality genome of the addax was assembled and annotated to investigate the molecular basis of desert adaptation. Phylogenomic and comparative analyses, integrating genomes from gemsbok (*Oryx gazella*), scimitar oryx, sable antelope (*Hippotragus niger*), topi, and blue wildebeest, were performed to explore convergent genetic signatures shared with camels. In addition to parallel adaptations, lineage-specific molecular changes were detected in desert antelopes, alongside evidence suggesting potential roles of introgression and genetic load in shaping adaptive trajectories. Collectively, these results provide a genomic framework for understanding the molecular architecture of desert adaptation in large mammals and identify candidate genetic elements for further functional investigation.

MATERIALS AND METHODS

Sample collection and ethics statement

Whole blood was collected from a female addax housed at

Guangzhou Zoo. Ethical approval for the study was granted by the Ethics Committee of the Guangzhou Wildlife Research Center (permit number: GZZOO20230101A). All experimental procedures were approved by the Animal Care and Use Committee of Northwestern Polytechnical University following the guidelines outlined in the Guide for the Care and Use of Laboratory Animals in China.

Genome assembly and annotation

Oxford Nanopore (ONT) long reads exceeding 1 kb were self-corrected and assembled using NextDenovo v.2.2-beta.0 (Hu et al., 2024). Contigs were subsequently polished using NextPolish v.1.2.1 (Hu et al., 2020), with three rounds of alignment using long reads and three rounds of alignment using short reads. Genome assembly completeness was assessed using BUSCO v.5.1.2 (Manni et al., 2021) with the Laurasiatheria_odb10 dataset and Compleasm v.0.2.2 (Huang & Li, 2023). Primary contigs were scaffolded using Ragtag v.1.0.0 (Alonge et al., 2019) based on syntenic alignment to the cattle reference genome (ARS-UCD1.2).

Repeat elements were identified using RepeatMasker v.4.1.2 (<http://www.repeatmasker.org>) with the Repbase transposable element library (Jurka et al., 2005), RepeatModeler v.2.0.2a (<http://www.repeatmasker.org/RepeatModeler.html>), TRF v.4.09 (Benson, 1999), LTR finder v.1.07, and RepeatProteinMasker v.4.1.2 (<http://www.repeatmasker.org>). BEDTools v.2.30.0 (Quinlan & Hall, 2010) was used to merge results and mask repetitive regions. Protein-coding gene annotation integrated multiple approaches. First, protein sequences from humans, cattle, and goats (Ensembl annotation) were mapped to the genome using tBLASTn (Camacho et al., 2009) with an *E*-value threshold of 1×10^{-5} . GeneWise v.2.4.1 (Birney et al., 2004) was then used to predict gene models based on protein alignment. Second, whole-genome alignments to human, cattle, and goat genomes were performed using an in-house whole-genome alignment pipeline. Gene projection was inferred using TOGA v.1.0.1 (Jebb et al., 2020) with target genome annotation from ENSEMBL. In addition, *de novo* gene prediction was performed using AUGUSTUS v.3.3.3 (Stanke et al., 2006) with default parameters. All predicted gene models were integrated using EvidenceModeler v.2.1.0 (Haas et al., 2008).

Pairwise whole-genome alignments

Pairwise alignments were conducted using the UCSC whole-genome alignment pipeline (Kent et al., 2003). Soft-masked query and target genomes were aligned using lastz v.1.04.03 (Harris, 2007) with parameters “O=400 E=30 M=254”. Alignment blocks were chained using axChain and netted into whole-genome alignments using chainNet with parameters “-minSpace=25, -minFill=1, -minScore=2000”. Genomic synteny was visualized with Circos v.0.69-8 (Krzywinski et al., 2009).

Phylogeny reconstruction

To reconstruct phylogenetic relationships, genome data from the addax and 17 additional species were analyzed, including gemsbok, scimitar oryx, sable antelope, topi, blue wildebeest, goat (*Capra hircus*), cattle (*Bos taurus*), musk deer (*Moschus berezovskii*), Chinese muntjac (*Muntiacus reevesi*), giraffe (*Giraffe camelopardalis*), lesser mousedeer (*Tragulus kanchil*), hippo (*Hippopotamus amphibius*), pig (*Sus scrofa*), camel (*Camelus ferus*), horse (*Equus caballus*), dog (*Canis lupus familiaris*), and human (*Homo sapiens*) (Supplementary Table

S1). A total of 12 234 conserved genes from the BUSCO Laurasiatheria_odb10 dataset were used to identify single-copy orthologous genes. From these, 9 991 genes present in at least 80% of the species were retained for further analysis. Nucleotide sequences of each gene were aligned using MAFFT v.7.471 (Katoh & Standley, 2013). Multiple sequence alignment was used to construct a concatenated maximum-likelihood tree using IQ-TREE v.2.1.2 (Minh et al., 2020). Gene trees of each single-copy orthologous gene were also reconstructed using IQ-TREE v.2.1.2 with parameters “--runs 2 -B 1000 --boot-trees -m MFP” (Minh et al., 2020). A coalescent-based species tree was inferred from the gene trees using ASTRAL v.5.7.3 (Zhang et al., 2018). Topological frequencies among gene trees were estimated with DiscoVista v.1.0 (Sayyari et al., 2018). Divergence times were estimated with MCMCTree in PAML v.4.9 (Yang, 2007), calibrated with fossil-based constraints from TimeTree (www.timetree.org).

Single nucleotide polymorphism (SNP) calling

Re-sequencing data from three addax and 38 scimitar oryx individuals were downloaded from NCBI (Hempel et al., 2021; Humble et al., 2023). Raw reads were quality filtered using fastp v.0.21.0 (Chen et al., 2018). Clean reads were mapped to the addax genome assembly and scimitar oryx reference genome assembly (GCF_014754425.2) using BWA MEM v.0.7.17 (Li & Durbin, 2009). Variant calling was performed using HaplotypeCaller and GenotypeGVCFs from GATK v.3.8 (Van der Auwera et al., 2013). SNPs were filtered to retain only biallelic sites using BCFtools v.1.9 (Danecek et al., 2011). Subsequently, to obtain a high-quality set of variants, VCFtools v.0.1.16 (Danecek et al., 2011a) was employed to remove loci with a quality score less than 30, mean depth of coverage less than 5 or greater than 20, genotyping rate less than 95%, or minor allele count less than 1. Genome-wide heterozygosity was calculated as the proportion of polymorphic sites over the total number of SNP sites.

Demographic history

The demographic history of desert antelopes was inferred based on pairwise sequential Markovian coalescent (PSMC) analysis calibrated with substitution rates estimated by r8s v.1.8.1 (Sanderson, 2003). Illumina reads were aligned to their respective reference genomes using Bowtie2 (Langmead & Salzberg, 2012) with default parameters. Scaffolds, aligned to cattle sex chromosomes, were subsequently excluded from downstream analyses to avoid artifacts. Consensus diploid sequences were obtained using SAMtools v.1.9 (Li et al., 2009) and BCFtools v.1.9 (Danecek et al., 2011) with the parameters recommended by PSMC v.0.6.5 (Li & Durbin, 2011). Generation times were set at 7.6 years for addax, 7.1 years for gemsbok, and 7.9 years for scimitar oryx, based on IUCN data (IUCN, 2024). Population trajectories were visualized with 100 bootstraps. Additionally, effective population sizes (N_e) of addax and scimitar oryx were estimated using SMC++ v.1.15.4 (Terhorst et al., 2017) based on genome-wide SNP datasets.

Convergent evolution in large desert mammals

To investigate genome-wide signatures of convergence between desert antelopes and camels, 9 991 single-copy orthologous genes were examined for potential convergent amino acid (AA) substitutions. The convergence-at-conservative sites (CCS) method (Xu et al., 2017) was applied to detect convergent residues based on the inferred species

tree comprising 18 mammals. To account for stochastic convergence, 10 000 ancestral sequence simulations were performed using PAML v.4.9 (Yang, 2007). Gene Ontology (GO) enrichment analysis of positively selected genes (PSGs) was conducted using Enrichr (Kuleshov et al., 2016), and multiple sequence alignments were visualized using ggmsa (Zhou et al., 2022). To further identify convergent PSGs, the BUSTED model in HyPhy v.2.5.15 (Pond et al., 2005) was employed, consistent with methods used to detect species-specific selection. In parallel, the branch-site model in PAML (Yang, 2007) was used to detect signatures of positive selection in camels and desert antelopes. Likelihood ratio tests (LRT) were used to compare models, with $2\Delta\ln L$ values assumed to follow a χ^2 distribution and degrees of freedom corresponding to the parameter differential between models. Genes were considered candidates for positive selection if they met two criteria: False discovery rate (FDR)-adjusted $P < 0.05$ in the LRT and the presence of at least one codon site with a posterior probability > 0.8 identified by Bayesian Empirical Bayes (BEB).

Protein structure simulation

The three-dimensional structure of PARP3 was predicted using AlphaFold3 (Abramson et al., 2024). Molecular docking simulations were then performed to assess NAD⁺ binding interactions with both addax and human PARP3 proteins. Polar contact interactions between PARP3 and NAD⁺ were visualized to examine potential differences in ligand-binding dynamics.

Cyclic adenosine monophosphate (cAMP) accumulation assay

The cAMP accumulation assay was conducted by ICE Bioscience (China). The wild-type (WT) human *PTGER2* gene was cloned into the pcDNA3.1 vector with an N-terminal hemagglutinin signal peptide. Site-directed mutagenesis was performed using QuickChange polymerase chain reaction (PCR), with all constructs verified by Sanger sequencing. HEK293 cells were cultured in 1× Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% (v/v) fetal bovine serum and incubated in 5% CO₂ at 37°C. The constructed plasmid was transfected into HEK293 cells using Lipofectamine 3000 transfection reagent. PTGER2-HEK cells were treated with trypsin and resuspended in 1× stimulation buffer after removing the culture medium by centrifugation at 25°C and 1 000 r/min for 5 minutes, then seeded into 384-well plates. A 1 µL volume of diluted 10× test compound was added to each well, followed by centrifugation at 25°C and 1 000 r/min for 1 minute and incubation at 37°C for 30 minutes. Detection buffer was used to dilute Eu-cAMP to working concentrations, with 5 µL of this solution added to each well. ULight-anti-cAMP was also diluted with detection buffer to working concentrations, with 5 µL of this solution added to each well. The plate was centrifuged at 25°C and 1 000 r/min for 1 minute and incubated at room temperature for 1 h. After incubation, a BioTek multifunctional microplate reader was used to detect fluorescence signals at 665 nm and 620 nm. The ratio (665/620) was plotted against compound concentration, and half-maximal inhibitory concentrations (IC₅₀) were calculated using nonlinear regression in GraphPad Prism v.7.

Identification of conserved non-coding elements (CNEs)

Genomes from 18 mammalian species, including addax,

gemsbok, scimitar oryx, sable antelope, topi, blue wildebeest, goat, cattle, musk deer, Chinese muntjac, giraffe, lesser mousedeer, hippo, pig, camel, horse, dog, and human, were used to identify CNEs (Supplementary Table S1). Whole-genome alignments were performed by aligning each genome against the cattle genome (GCA_002263795.2) using the previously described pairwise whole-genome alignment pipeline. The resulting pairwise alignments in MAF format were merged into a multiple alignment using MULTIZ v.11.2 (Blanchette et al., 2004; Mirdita et al., 2022), with cattle as the reference. Conserved regions across the 18 species were identified and filtered using PHAST v.1.4 (Hubisz et al., 2011). Coding and non-coding elements were separated using ANNOVAR v.2015-12-14 (Wang et al., 2010) based on gene annotations from the cattle GFF file, with only non-coding elements retained for subsequent analysis. CNEs conserved across all 18 species were further screened to identify CNEs with fixed insertion/deletion substitutions (>5 bp in length) specific to desert antelope lineages. To examine regulatory potential, FIMO v.5.5.6 (Grant et al., 2011) was used to identify motifs using the JASPAR database (Castro-Mondragon et al., 2022). Bulk DNase-seq data were downloaded from ENCODE (ENCODE Project Consortium et al., 2020) and ATAC-seq (ERR9846793) data were downloaded from NCBI. Raw ATAC-seq reads were filtered using fastp v.0.23.1 (Chen et al., 2018) and clean reads were subsequently aligned to the reference genome using Bowtie2 (Langmead & Salzberg, 2012). Improperly paired reads and PCR duplicated reads were removed using Picard v.2.25.7 (<https://broadinstitute.github.io/picard/>). Only reads with a high mapping quality score (MAPQ>13) were retained for peak calling, which was performed with MACS2 v.2.1.0 (Zhang et al., 2008). Promoter and enhancer regions were retrieved from the human ENCODE database (ENCODE Project Consortium et al., 2020).

Dual-luciferase reporter assay

Chemically synthesized DNA fragments corresponding to Cattle-*AGTR1*-Promoter (Chr1:119 489 350–119 489 600), Cattle-*AGTR1*-Enhancer (Chr1:119 521 302–119 521 553), Human-*AQP1*-Promoter (chr7:30 911 588–30 911 927), and Human-*AQP1*-Enhancer (chr7:30 914 634–30 914 983) were cloned into the pGL4.23 vector (Promega, USA) and sequenced (Tsingke Biotech, China). HEK293T cells were seeded in 24-well plates, and after 24 h, 100 ng of each pGL4.23 construct was co-transfected with 5 ng of pRL-CMV (internal control) using Turbofect reagent (ThermoFisher Scientific, USA). Cells were collected 36 h post-transfection using passive lysis buffer. Luciferase activity was measured using the Dual-Luciferase Reporter Assay System (Promega) on a multimode microplate reader (Spark Tecan, Switzerland). Light output from transcriptional activity was divided by the output from *Renilla* luciferase activity to normalize the samples.

Gene flow analysis

To distinguish introgression from incomplete lineage sorting (ILS), Quantifying Introgression via Branch Length (QuIBL) analysis was performed (Edelman et al., 2019). A total of 3 582 gene trees covering all 18 species were selected for evaluation. To distinguish whether discordances between gene and species trees were more likely due to introgression or ILS, delta Bayesian Information Criterion (BIC) values were determined for inner branch lengths by subtracting the BIC

value of scenario 1 (ILS alone) from the BIC value of scenario 2 (ILS and introgression combined). Stringent cut-off values of delta BIC>10 (favoring the ILS-alone model) or BIC<-10 (favoring the combined ILS and introgression model) were set to accept the better model.

To further assess gene flow, *D*-statistics (*ABBA-BABA* test) were computed using D-suite v.0.5 (Malinsky et al., 2021), with goat designated as the outgroup. Variant call format (VCF) files containing SNP data were generated from the whole-genome alignment MAF files using snp-sites (Page et al., 2016). The “Dtrios” option was applied to calculate *D*-statistics for the triplet consisting of addax, gemsbok, and scimitar oryx, using goat as the outgroup. A model-based estimate of introgression was obtained by constructing a maximum-likelihood (ML) tree using TreeMix v.1.13 (Pickrell & Pritchard, 2012), accounting for linkage disequilibrium (LD) by grouping sites in blocks of 10 000 SNPs (-k 10 000). After constructing the ML tree, migration events were sequentially introduced (-m) and iterated 50 times for each value (m=1–2) to check for convergence based on likelihood scores and the proportion of variance explained. The inferred ML trees were visualized using the built-in TreeMix R script plotting functions.

To reconstruct the evolutionary history and introgression events of desert-adapted species, the multispecies coalescent with introgression (MSCi) framework was applied using BP&P v.4.7.0 (Flouri et al., 2020), which incorporates both incomplete lineage sorting and introgression into phylogenomic inference. Three independent replicates were run for 500 000 generations and sampled every two generations, with a burn-in of 32 000 generations. TRACER v.1.7.2 (Rambaut et al., 2018) was used to combine the results of the three independent runs and assess convergence (effective sample size>200).

Runs of homozygosity (ROH) and identification of deleterious mutations

ROH segments were called with a minimum length of 500 kb and a minimum of 50 SNPs using the --homozyg function in PLINK v.1.9 (Chang et al., 2015) and parameters: --homozyg-window-snp 50 --homozyg-snp 50 --homozyg-kb 500 --homozyg-gap 1000 --homozyg-density 50 --homozyg-window-missing 5 and --homozyg-window-het 3.

Missense variants were identified using snpEff v.4.30 (Cingolani et al., 2012). Ancestral states of missense variants were determined based on sequence comparisons with sable antelope, topi, and blue wildebeest. Heterozygous and homozygous mutation loads of missense variants were estimated in addax and scimitar oryx, respectively. Heterozygous mutation load was measured as the number of heterozygotes for each individual at missense sites, and homozygous mutation load was measured as the total number of homozygotes for each individual at missense sites. Deleterious mutations were identified using SIFT4G, which provided a score indicating the putative deleterious effect for each position of the protein using protein alignment and UniRef100 (Vaser et al., 2016). Positions with scores <0.05 were classified as deleterious.

RESULTS

Genome assembly and phylogeny of the addax

A highly contiguous reference genome of the addax was generated, with initial contigs assembled from ONT long reads

and polished using Illumina short reads. The resulting primary assembly achieved a contig N50 of 72.85 Mb, representing a three-order-of-magnitude improvement over the previous assembly (ASM1959352v1; contig N50: 17 kb) (Hempel et al., 2021) (Supplementary Table S2). Chromosomal synteny analysis between the addax and cattle (Figure 1A) enabled the construction of a chromosome-level assembly using the cattle genome (GCA_002263795.2) as a reference. The final assembly was anchored to 29 chromosomes ($2n=58$), with a total genome length of 2.85 Gb (Supplementary Tables S3, S4), consistent with the known addax karyotype (Claro et al., 1996). Moreover, BUSCO analysis using the Laurasiatheria_odb10 dataset (12 234 genes) indicated a high level of completeness, with 95.99% of genes identified. Repetitive elements accounted for 1.35 Gb, or 47.76% of the genome (Supplementary Table S5). A total of 22 211 protein-coding genes were annotated, with gene structure metrics, such as gene length, exon number, and codon usage, comparable to those of other mammals (Supplementary Table S6) (Hempel et al., 2021; Humble et al., 2020).

Phylogenetic reconstruction was performed using 9 991

single-copy orthologous genes across 18 species (Supplementary Table S1). Both the concatenated ML tree and coalescent-based species tree supported the placement of *Addax* as the sister lineage to *Oryx*, consistent with previous findings (Bibi, 2013; Fernández & Vrba, 2005) (Figure 1B).

Genome-wide heterozygosity in the addax was estimated at approximately 0.09%, one of the lowest levels reported among endangered mammals (Figure 1C), underscoring the urgency of conservation efforts. Demographic histories of the addax, scimitar oryx, and gemsbok were reconstructed using PSMC models. Results indicated that both the addax and scimitar oryx experienced pronounced reductions in N_e beginning around the mid-Pleistocene transition (1.25–0.7 million years ago) (Figure 1D), likely in response to global cooling trends (Herbert, 2023). The N_e trajectories of the addax and gemsbok exhibited the same trend, whereas the N_e of the addax and scimitar oryx showed the opposite trend prior to the Last Glacial Maximum (LGM, 20 000 years ago). This divergence may reflect differing ecological tolerances, with the addax potentially less competitive in humid environments

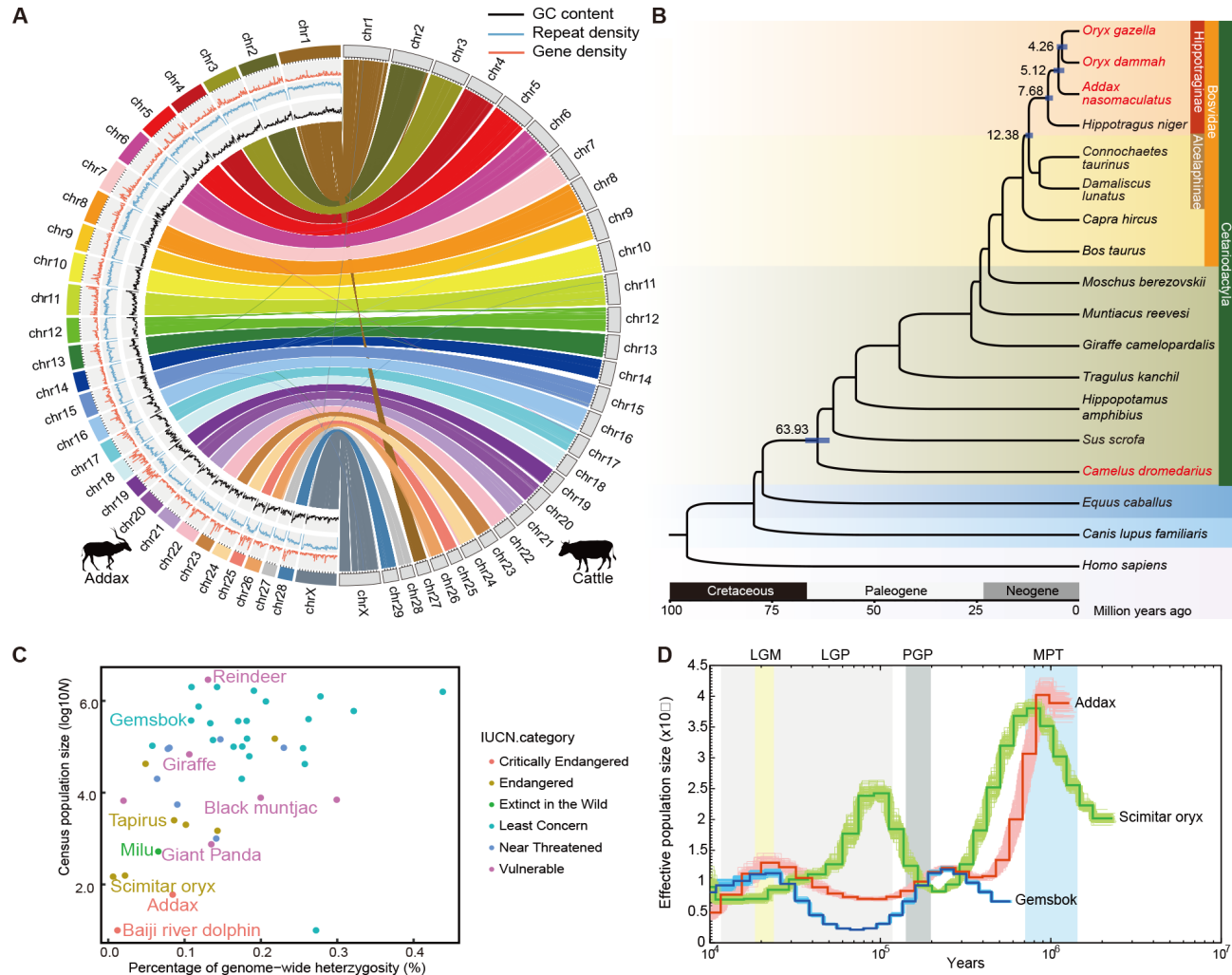


Figure 1 Genome assembly and evolutionary history of the addax

A: Synteny alignment of the addax and cattle. Densities of coding genes and repeat sequences were calculated with a window size of 1 Mb. B: Phylogenetic relationships of the addax and mammals. Numbers labeled on the tree refer to estimated divergence time, and blue rectangles on each node represent 95% confidence interval. C: Genome-wide heterozygosity distribution. D: Recent population demographic history of three desert antelopes (addax: $g=7.6$, $\mu=2.46\times10^{-8}$; scimitar oryx: $g=7.9$, $\mu=2.52\times10^{-8}$; gemsbok: $g=7.1$, $\mu=2.52\times10^{-8}$). (Blue: Mid-Pleistocene Transition (MPT), dark gray: Penultimate Glacial Period (PGP), light gray: Last Glacial Period (LGP), yellow: Last Glacial Maximum (LGM).

occupied by the scimitar oryx. Since the LGM, all three desert antelope species have exhibited a declining trend in N_e as inferred by SMC++ (Supplementary Figure S1), a pattern broadly consistent with other ruminants (Chen et al., 2019) and likely attributable to the impact of human activities.

Convergent evolution underlying desert adaptation in large mammals

Phylogenetic analysis confirmed that desert antelopes are distantly related to camels, with an estimated divergence time of approximately 64 million years ago (Figure 1B). To ascertain potential convergent adaptations to arid environments in these lineages, two distinct methodologies were applied: detection of convergent AA substitutions and identification of shared signals of natural selection. Convergent AA substitutions were identified using the CCS method (Xu et al., 2017) and verified against the ruminant genome database (RGD) (Fu et al., 2022). Raw sequencing reads were also analyzed to confirm the reliability of the convergent sites (Supplementary Figure S2). A total of 136 genes containing 149 convergent AA substitutions were identified in camels and three desert antelope lineages (L1: ancestral lineage of *Oryx* and *Addax*; L2: ancestral lineage of Hippotraginae; L3: ancestral lineage of Hippotraginae and Alcelaphinae; Figure 2A, B; Supplementary Table S7). Specifically, 36, 78, and 27 convergent genes were detected between camels and L1, L2, and L3, respectively (Supplementary Table S7). GO enrichment analysis revealed distinct functional profiles among lineages. Convergent genes in L3 were enriched in energy metabolism functions, including cellular response to ketone, glycerol-3-phosphate metabolism, and galactose metabolism. In contrast, genes in the L1 and L2 lineages were associated with functions related to heat stress, including interstrand cross-link repair and regulation of the

DNA damage checkpoint (Figure 2C; Supplementary Table S8). These lineage-specific functional enrichments suggest that desert adaptation in antelopes followed a stepwise evolutionary trajectory. Further analysis revealed that 21 of the 136 convergent genes were also under positive selection in both camels and desert antelopes, while 42 genes showed lineage-specific signals of positive selection in either group (Supplementary Table S9).

Genomic insights into water reabsorption: Efficient water reabsorption is critical for desert-dwelling animals to survive under conditions of scarce and unpredictable water availability. Comparative genomic analyses identified several genes related to osmotic regulation (*PTGER2*, *SIK3*, *SAT1*, and *SLC38A4*) that exhibited convergent AA substitutions in camels and desert antelopes. These genes participate in the maintenance of water balance via kidney osmotic regulation (Li et al., 2019; Mackenzie & Erickson, 2004; Markovich, 2011; Olesen & Fenton, 2013). Among them, *PTGER2* exhibited a convergent R146S AA substitution and evidence of positive selection in desert-adapted lineages (Figure 3A, B). *PTGER2* encodes a receptor for prostaglandin E2 (PGE2), which is activated by G (s) proteins, stimulating adenylyl cyclase and subsequently increasing intracellular cAMP levels (Qu et al., 2021) (Figure 3A). To evaluate the functional implications of the R146S mutation, a cAMP accumulation assay was conducted to compare the agonist sensitivity of the WT human *PTGER2* receptor and the R146S mutant. The WT receptor exhibited a half maximal effective concentration (EC_{50}) of 0.055 nmol/L, consistent with high agonistic potency (Qu et al., 2021). Conversely, the R146S mutant showed a markedly reduced sensitivity, with an EC_{50} of 0.689 nmol/L, indicating a 12.5-fold decrease in cAMP accumulation relative to WT ($P < 0.001$, one-way analysis of variance (ANOVA)). These results suggest that the R146S substitution significantly

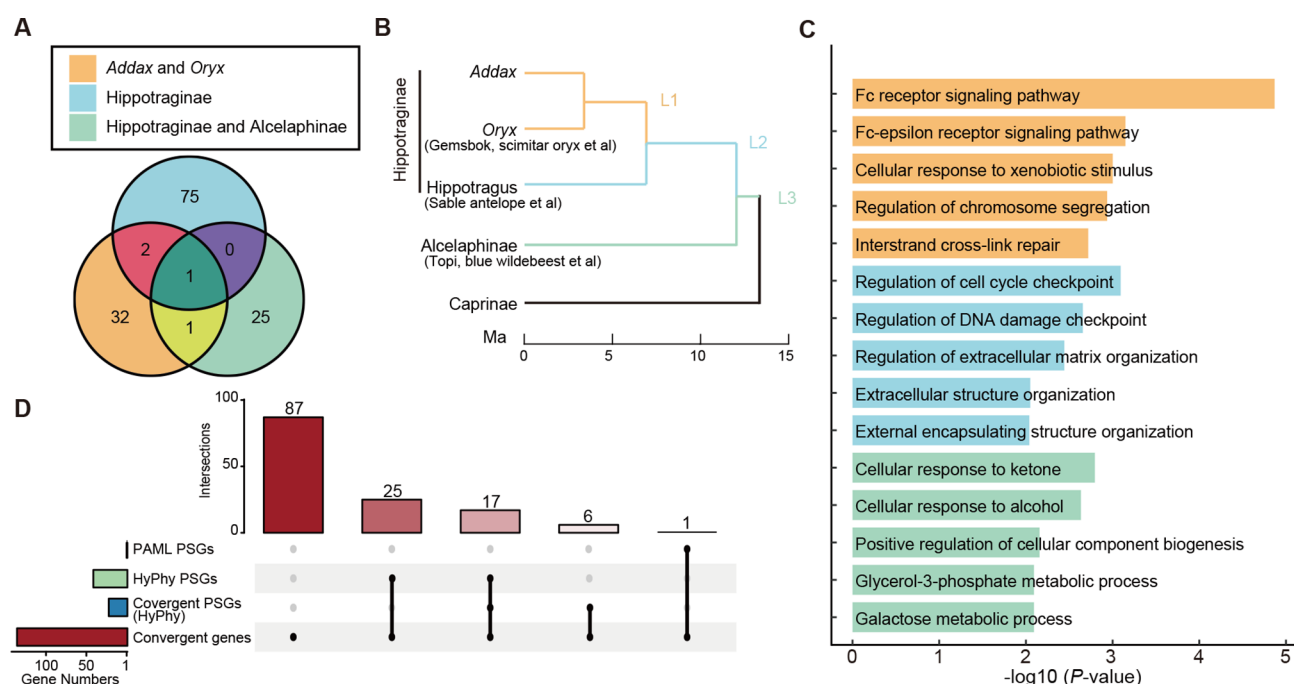


Figure 2 Convergent evolution between camels and desert antelopes

A: Convergent genes between camels and antelopes. B: Phylogenetic tree of desert antelopes (L1: ancestral lineage of *Addax* and *Oryx*; L2: ancestral lineage of Hippotraginae; L3: ancestral lineage of Hippotraginae and Alcelaphinae). C: GO enrichment analysis of convergent genes between camels and antelopes (orange: *Addax* and *Oryx*; blue: Hippotraginae; green: Hippotraginae and Alcelaphinae). D: Convergent positively selected genes (PSGs) and desert-mammal-specific PSGs in convergent genes.

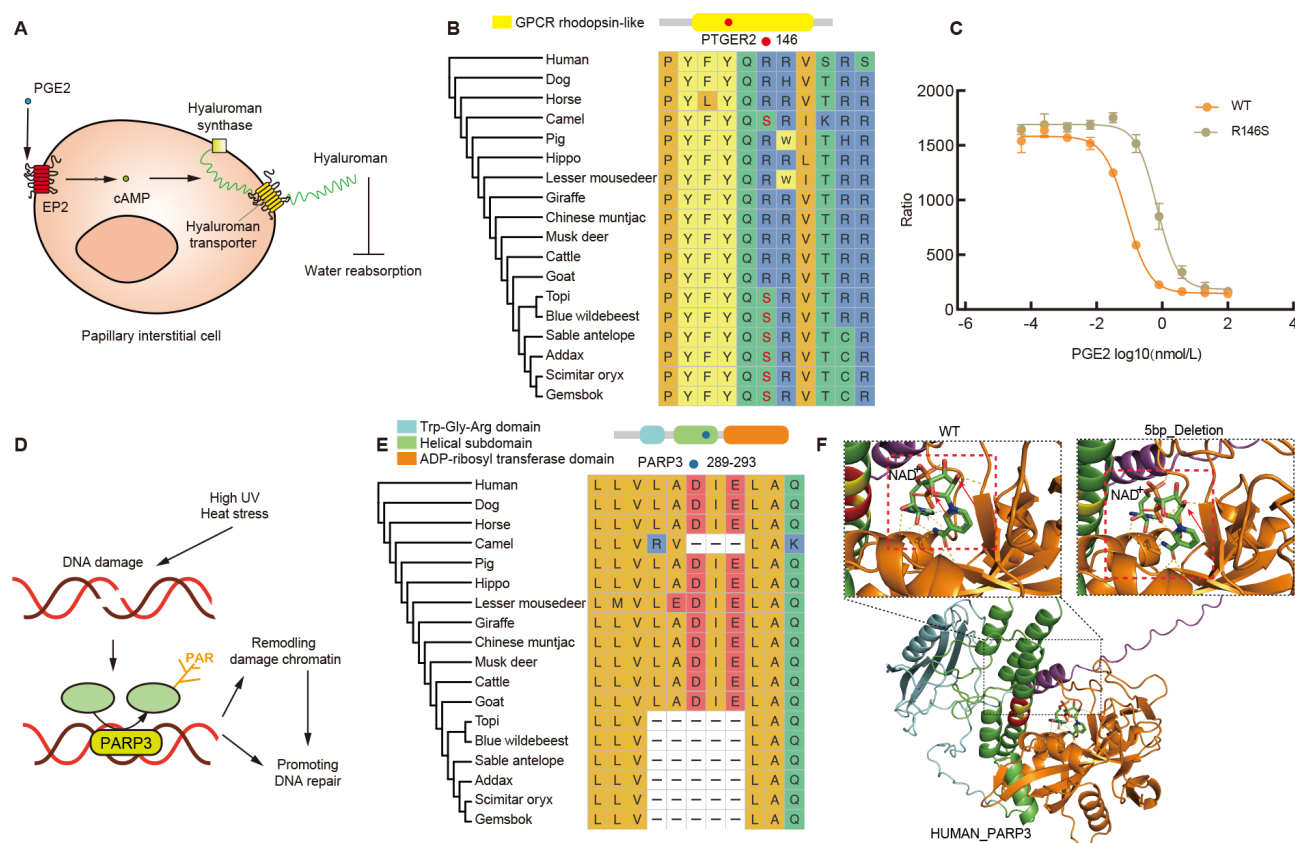


Figure 3 Convergent evolution in desert mammals

A: Schematic overview of molecular pathways involved in water reabsorption. B: Convergent amino acid (AA) substitution (R146S) in PTGER2 observed in desert-adapted mammals. C: Functional impact of the R146S mutation on PTGER2-mediated ligand binding and Gs signaling, with signaling values negatively correlated with cAMP concentration. D: Schematic overview of the mechanisms associated with DNA repair in desert mammals. E: PARP3 contains a 5 AA indel (residues 289–293) in desert antelopes and a 3 AA indel (residues 291–293) in camels. F: AlphaFold3-predicted structure of PARP3 in complex with NAD⁺. Domains are color-coded: NTR (purple), Trp-Gly-Arg (WGR) (blue), helical subdomain (HD) (green), and ADP-ribosyl transferase (ART) (orange). The 5 AA deletion is highlighted in red; NAD⁺ is shown as a green molecule within a red dashed box; hydrogen bonds are indicated in yellow; and significantly altered hydrogen bonds are denoted by arrows.

attenuates PTGER2 activity in desert animals (Figure 3C). Previous studies have shown that PGE2 stimulates hyaluronan synthesis in stromal mesenchymal cells. Hyaluronan, a glycosaminoglycan with high water-binding capacity, enhances water retention in the extracellular matrix (Rügheimer et al., 2008). PTGER2 is the only prostaglandin receptor expressed in mesenchymal cells (Zhang et al., 2009), and Gs-coupled signaling has been implicated in hyaluronan synthase activation in other tissues (Olesen & Fenton, 2013) (Figure 3A). In small desert rodents, reduced renal hyaluronan levels are associated with enhanced water conservation (Göransson et al., 2002). Thus, the R146S mutation may diminish PTGER2 activity, leading to a potential decrease in hyaluronan expression and increase in water reabsorption, thereby contributing to survival in desert-dwelling animals. Taken together, these findings suggest that the ancestral lineages of Hippotraginae and Alcelaphinae may have already acquired physiological mechanisms for coping with aridity early in their evolutionary history.

Genomic insights into DNA repair: In desert environments, large mammals face persistent exposure to extreme heat and UV radiation, both of which can cause extensive DNA damage (Sejian et al., 2018), including single-strand (SSBs) and double-strand breaks (DSBs). As such, enhanced DNA repair capacity may represent a critical component of desert adaptation. Several genes related to DNA repair pathways

(e.g., *PARP3*, *FANCA*, *FANCC*, and *BRCA2*) (Walden & Deans, 2014) were found to contain convergent AA substitutions at multiple ancestral nodes in the desert antelope lineage (Figure 2B; Supplementary Figure S3). Among them, *PARP3* emerged as a PSG, encoding poly (ADP-ribose) (PAR) polymerase 3, which catalyzes the synthesis of PAR. PAR production also directs the rapid local recruitment of chromatin remodeling factors and various repair proteins mediated via their PAR-binding domains (De Vos et al., 2012) (Figure 3D). PARP3-dependent mono ADP-ribosylation (MARylation) is essential for both DSB and SSB repair processes (Rulten et al., 2011). Notably, insertions and deletions (indels) were identified in PARP3 at conserved regions in desert-adapted species: a 5 AA indel (residues 289–293) in desert antelopes and a 3 AA indel (residues 291–293) in camels (Figure 3E). These indels were located in the helical subdomain (HD) in the catalytic (CAT) domain of PARP3. The HD plays a crucial role in function PARP3 by coordinating PARP1 multimeric interactions, enhancing DNA damage binding affinity, and regulating NAD⁺ binding site exposure and enzyme activity (Rouleau-Turcotte et al., 2022). Loss of the HD has been shown to increase DNA-independent catalytic activity of PARP3 (Dawicki-McKenna et al., 2015).

To evaluate the structural implications of these indels, AlphaFold3 (Abramson et al., 2024) was used to model the three-dimensional structure of addax PARP3. The predicted

model revealed alterations compared to the human PARP3 structure (Figure 3F). Furthermore, molecular docking simulations with NAD⁺ showed an increase in hydrogen bonds in the addax protein (13) compared to the human protein (12), suggesting enhanced binding affinity for the NAD⁺ ligand. Structural heatmaps indicated enhanced probability of contact between the WGR domain and DNA in the indel region (Supplementary Figure S4). These results suggest that indels in the HD may influence the regulation of NAD⁺ binding site exposure and DNA affinity. Given the association between UV-induced DNA damage and desert exposure, such modifications may represent adaptive enhancements in DNA repair capacity under extreme environmental stress. These findings support a role for the evolution of DNA repair mechanisms in facilitating mammalian survival in arid, high-radiation habitats.

Genomic insights into food uptake and energy metabolism: Food scarcity also represents a significant challenge faced by desert animals, necessitating efficient strategies for energy acquisition and storage. Triacylglycerol (TAG) serves as the primary form of energy storage in mammalian cells, supporting survival during prolonged nutritional deprivation (Csaki et al., 2013). Comparative genomic analysis revealed convergent AA substitutions in two key genes, *GPAT2* and *AGPAT2*, which belong to four core enzyme families responsible for TAG synthesis (Supplementary Figure S3). These genetic modifications may enhance lipid storage capacity, enabling desert mammals to survive long periods of food scarcity.

In addition to metabolic adaptations, alterations in dietary sensing mechanisms were also observed. The gene *TAS1R2*, which is associated with the perception of sweet and acid tastes (Liman & Kinnamon, 2021; Nelson et al., 2001), exhibited convergent signals of relaxed selective constraint across multiple desert mammal species, suggesting unique foraging strategies. By decreasing their sensitivity to sweet and acidic tastes, desert animals may more effectively exploit a wide spectrum of limited and sporadic food resources, thereby improving their survival under resource-limited conditions (Liu et al., 2016).

Lineage-specific adaptations in desert antelopes

To investigate antelope-specific genetic adaptations to desert environments, lineage-specific coding innovations were analyzed in desert-adapted taxa. Among orthologous genes, a total of 269 were identified as PSGs, including 100, 101, and 100 genes in L1, L2, and L3, respectively ($P < 0.05$), and 184 genes in the addax lineage (Supplementary Table S10). GO enrichment analysis of the addax genes revealed significant enrichment of the term “Microtubule Binding”, suggesting adaptive enhancements in microtubule-associated processes (Supplementary Table S11). These modifications may confer improved cellular integrity and stability, efficient resource transport, robust stress responses, and effective cell division, potentially contributing to physiological robustness under desert conditions.

CNEs, which often function as *cis*-regulatory elements (Pennacchio et al., 2006; Woolfe et al., 2004), were also examined as potential contributors to desert adaptation. Alterations within CNEs, such as indels, have the potential to impact genetic regulation. Using whole-genome alignments from 15 mammalian species in combination with ATAC-seq data from cattle and DNase-seq data from humans, 404

lineage-specific CNEs were identified (103 in L1, 209 in L2, and 92 in L3), with specific structural variants located within ATAC-seq or DNase-seq peaks in the ancestor branches of desert antelopes (Supplementary Table S12). Genes associated with these CNEs were enriched in “adherens junction”, “semaphorin receptor activity”, and “cell junction” terms (Supplementary Table S13), functions implicated in neural development and intercellular stability that may support physiological resilience to heat stress.

Notably, several CNE-associated genes are involved in pathways regulating fluid homeostasis, such as the renin secretion pathway (Figure 4A; Supplementary Figures S5, S6). Among these, *AGTR1* and *AQP1* are of particular functional relevance. *AGTR1* encodes the angiotensin II type 1 receptor, a key mediator of the renin-angiotensin-aldosterone system (RAAS), regulating thirst, blood pressure, and urine concentration. Mice lacking *Agtr1* exhibit polyuria and impaired urine concentration (Stegbauer et al., 2011). In addition to its roles in cardiovascular function (Catt et al., 1984), *AGTR1* is also involved in energy metabolism (Kouyama et al., 2005), underscoring its importance in arid-adapted physiology (Figure 4B). A 30 bp deletion, exclusive to desert antelopes, was identified within the intronic region of *AGTR1*. Notably, this deletion overlapped with a strong ATAC-seq peak in cattle kidney cells (Figure 4C) and contained a highly conserved motif matching the BATF3 transcription factor binding site ($P = 2.704 \times 10^{-4}$). Given the pivotal role of *AGTR1*, alterations in its expression may significantly affect RAAS activity and related traits (Stegbauer et al., 2011). Functional validation using dual-luciferase reporter assays demonstrated that the cattle-derived CNE_285 sequence displayed strong repressor activity ($P < 0.001$; Figure 4D), whereas the desert antelope-specific deletion substantially reduced this repressive effect ($P < 0.001$; Figure 4D). Therefore, this desert antelope-specific deletion in CNE_285 may increase the expression of *AGTR1*, potentially impacting blood pressure regulation, urine concentration, and adaptation to desert environments.

Similarly, a desert antelope-specific 9 bp insertion was identified within the intronic region of *AQP1*, coinciding with a notable DNase-seq peak in human kidney cells. This region contained a highly conserved motif matching the TCF12 transcription factor binding site (Figure 4E). *AQP1* encodes a widely expressed water channel, which has been thoroughly characterized in the kidney (Preston et al., 1992), especially the apical and basolateral cell membranes of the proximal tubules and the descending limb of the loop of Henle, accounting for over 70% of total water reabsorption (Su et al., 2020). *Aqp1* knockout in mice leads to impaired urine concentration and reduced osmolality (Fenton & Knepper, 2007). To assess the regulatory potential of this region, dual-luciferase reporter assays were conducted using constructs spanning the 9 bp insertion. Results demonstrated that desert antelope CNE_138 exhibited a significantly greater enhancing effect (Figure 4F), suggesting that the insertion increases *AQP1* expression. This up-regulation may, in turn, increase water reabsorption capacity and contribute to improved osmoregulation, supporting adaptation to desert environments.

Potential effects of introgression and genetic load

Substantial gene tree discordance was observed, with only 49.06% of gene trees aligning with the species tree topology,

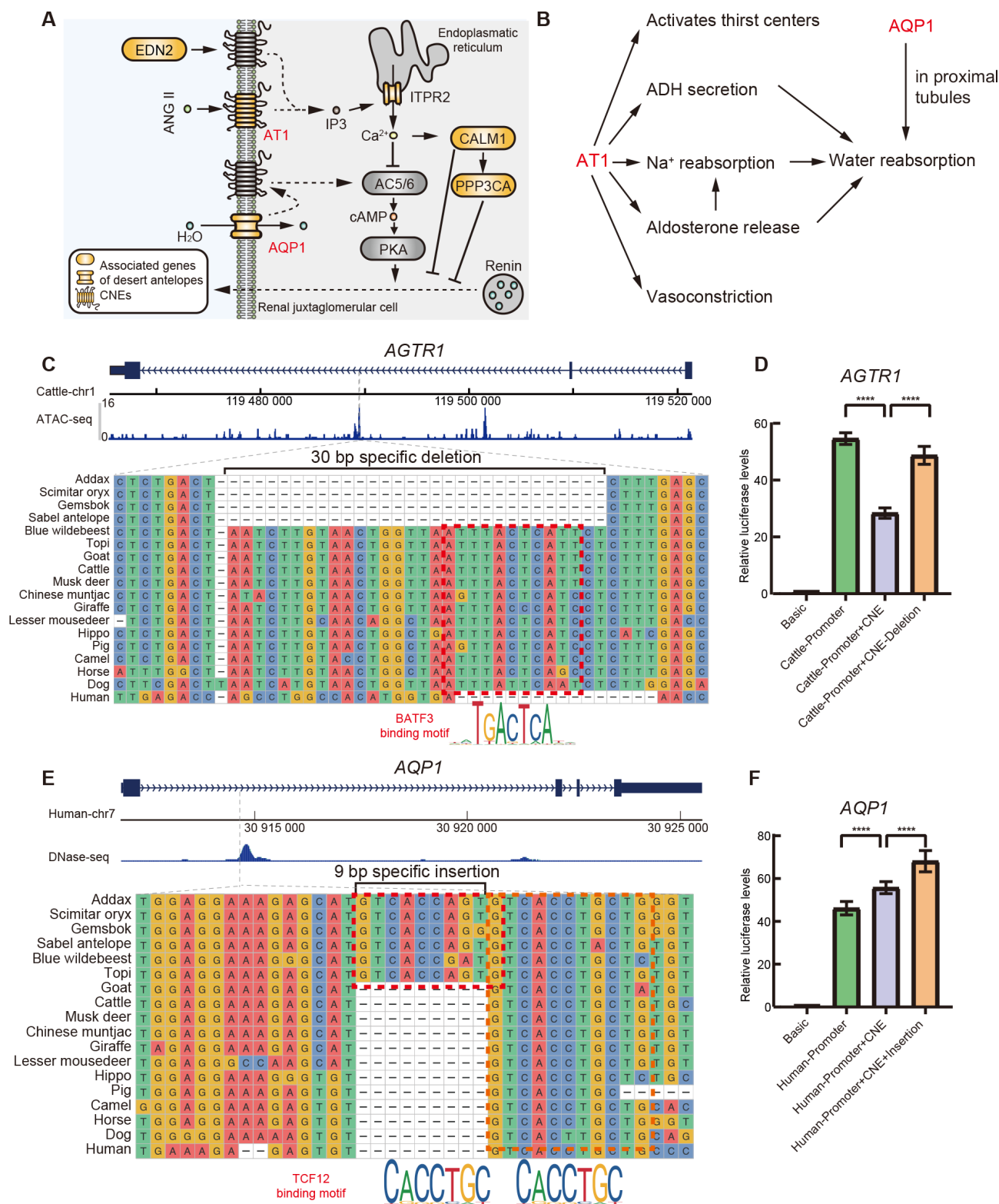


Figure 4 Innovation of conserved non-coding elements in desert antelopes

A: Schematic overview of renin secretion. B: Schematic overview of *AGTR1* and *AQP1* functions. C: Similarity of sequence alignments for CNE_285. ATAC-seq data of cattle kidney cells are presented. BATF3 motif is shown. D: *In vitro* reporter assays showing that desert-antelope-specific CNE_285 significantly increased *AGTR1* expression. Two-sided *t*-test was used (****: $P < 0.0001$; Exact *P*-values: Cattle promoter-cattle promoter+CNE, 3.80×10^{-21} ; Cattle promoter+CNE-cattle promoter+CNE-deletion, 5.36×10^{-15}). E: Similarity of sequence alignments for CNE_138. DNase-seq data of human kidney cells are presented. TCF12 binding motif is shown. F: *In vitro* reporter assays showing that desert-antelope-specific CNE_138 significantly increased *AQP1* expression. Two-sided *t*-test was used (****: $P < 0.0001$; Exact *P*-values: Human promoter-cattle promoter+CNE, 3.16×10^{-8} ; Human promoter+CNE-human promoter+CNE-deletion, 2.07×10^{-8}).

indicating a high level of ILS or introgression. In addition, unequal proportions of the remaining topologies (31.72% and 19.22%) further suggested the potential presence of introgression (Figure 5A). To distinguish introgression from ILS, three complementary approaches were employed. First,

QuIBL analysis (Edelman et al., 2019) provided strong support for introgression between addax and scimitar oryx ($\Delta\text{BIC} < -10$; Supplementary Table S14). This result was corroborated by TreeMix (Pickrell & Pritchard, 2012) and *D*-statistics (*ABBA-BABA* test) (Figure 5A–C), both of which supported

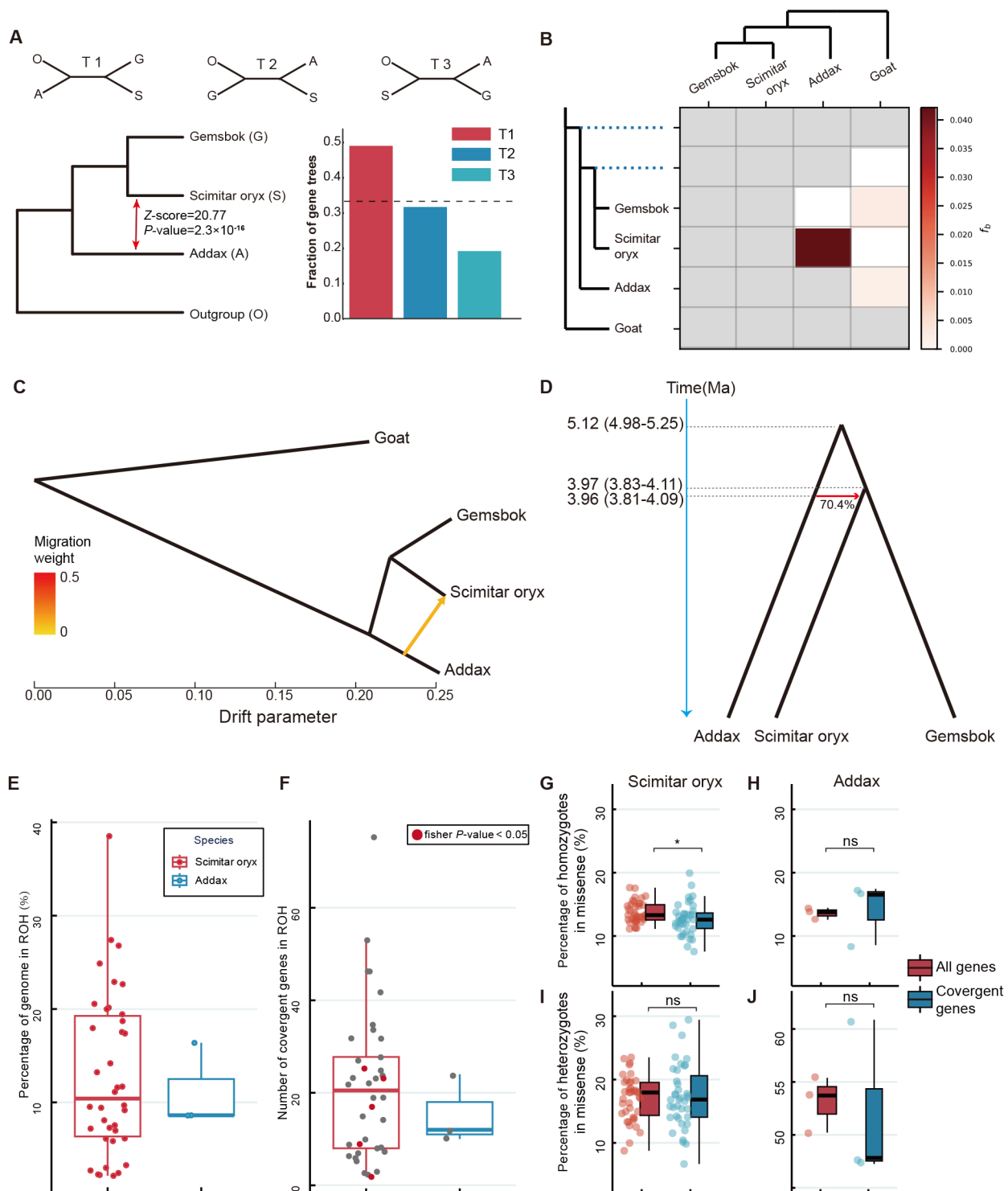


Figure 5 Effects of introgression, genetic load, and runs of homozygosity (ROH)

A: Significant introgression between addax and scimitar oryx inferred from DiscoVista and *ABBA-BABA* analyses. Node numbers are indicated in black boxes. B: F-branch (f_b) statistics showing excess allele sharing between addax and scimitar oryx. C: TreeMix analysis depicting ML tree with inferred migration edges. D: MSCi results from BPP analysis. E: Proportion of the autosomal genome in ROH for each individual. F: Number of convergent genes located within ROH across individuals. Gray points indicate individuals without significant enrichment based on Fisher's statistics. G–J: Distribution of heterozygous and homozygous missense mutations per individual in addax and scimitar oryx based on SNPeff annotations. Two-sided *t*-test was used. ns: Not significant; *, $P < 0.05$.

introgression from addax to scimitar oryx. Further evidence came from MSCi analysis, which yielded a high introgression probability ($\phi=0.705$) between these species. The high percentage of infiltration (70.5%; Figure 5D) implies extensive historical genetic exchange, consistent with the observed gene tree heterogeneity. Historical overlap in the distribution of addax and scimitar oryx (Kingdon, 2014) may have facilitated such gene flow. Notably, several genes within significantly introgressed genomic regions (top 1% of *ABBA-BABA* signals) are associated with desert adaptation (Supplementary Table S15). These include *SLCO3A1*, *SLCO4A1*, and *SLCO6A1*, which encode sodium-independent organic anion transporters (OATs) involved in metabolic function and detoxification (Sekine et al., 2000), and *SLC45A2*, *RXFP3*, *ADAMTS12*, and *AMACR*, which are associated with melanin metabolism and deposition (Le et al., 2020; Verhagen et al., 2012), potentially contributing to the pale skin pigmentation of both species.

To assess the impact of genetic load on desert adaptation, re-sequencing data from addax ($n=3$) and scimitar oryx ($n=39$), downloaded from NCBI (Humble et al., 2023), were analyzed to detect ROH region and identify possible deleterious mutations. Several convergent genes were found within ROH regions, although no significant enrichment signal was detected based on Fisher's statistics (Figure 5E, F; Supplementary Table S16). Heterozygous and homozygous mutation loads in addax and scimitar oryx were also estimated to investigate the effects of genetic load on desert adaptations (Figure 5G–J; Supplementary Table S17). In addax, no significant enrichment of deleterious homozygous or heterozygous mutations was observed in the 136 convergent genes compared to all genes across the genome. However, these results require further validation due to the limited number of addax individuals (3) used in the study. In contrast, scimitar oryx exhibited a significantly lower proportion of homozygous mutations in convergent genes ($P<0.05$; Figure 5G), suggesting potential purging of deleterious alleles as a contributing factor to desert adaptation (Supplementary Figure S7 and Table S18).

DISCUSSION

Although the desert adaptability of camels has been extensively studied, research on wild large mammals remains limited. The assembled addax genome complements existing genomic resources and provides novel insights into the molecular basis of desert adaptation in large-bodied mammals. Our findings underscore the critical need to conserve the remaining addax population and offer a genetic foundation for future conservation strategies.

Survival in arid environments requires specialized physiological and molecular adaptations. Large desert-dwelling mammals face intense selective pressures to minimize energy expenditure, enhance water conservation, and maintain homeostasis under extreme environmental stress. These challenges are compounded by the energetic and thermoregulatory demands associated with their large body size. Previous studies have revealed that desert adaptation involves coordinated selection of genes across diverse functional pathways, reflecting the complexity of adaptive phenotypes involved. Although functional convergence has been documented across various mammalian lineages (Rocha et al., 2021b), most analyses have focused on camels, with minimal representation of other

large-bodied taxa. In this study, comparative genomic analysis of camels and desert antelopes revealed convergent signatures in genes related to heat stress responses and energy metabolism (Figure 6), consistent with the greater metabolic and thermoregulatory demands experienced by large desert mammals. While few convergent amino acid substitutions were observed in pathways related to water reabsorption and thermoregulation, numerous CNEs in desert antelopes were identified in regulatory regions near relevant genes. For example, CNEs linked to *SLC22A11*, *CUBN*, *ATP6V0A1*, and *ATP6V1E2* are implicated in renal function and osmotic regulation (Supplementary Figures S4, S5), potentially influencing osmotic water reabsorption capacity. Similarly, CNEs associated with *CALM1*, *IPTR2*, *AGTR1*, and *MYL6* are involved in the regulation of vascular smooth muscle contraction and relaxation, which is essential for body temperature control. These findings underscore the likely importance of CNE-mediated regulatory changes in shaping physiological traits relevant to desert survival. Although no common variant genes were identified, our findings align with previously reported functions in related studies. The lack of prominent major-effect genes (genes with obvious phenotype effects) suggests that the mechanisms underlying convergent adaptation in animals to different environments vary considerably and constitute a complex process.

Notably, several convergent selection signals in desert antelopes appear to have originated from ancestral nodes shared by Hippotraginae and Alcelaphinae. Given that some Alcelaphinae species exhibit partial drought tolerance, this ancestral inheritance may have contributed to the emergence of key adaptive traits. In addition, evidence of gene flow between the addax and scimitar oryx, both of which inhabit the Sahara, suggests that introgression may have facilitated the transfer of adaptive alleles between lineages occupying similar ecological niches.

In conclusion, the high-quality genome sequence of the addax provides molecular evidence of convergent desert adaptation in large mammals. Genetic changes associated with water reabsorption, thermoregulation, heat stress response, and energy metabolism contribute to physiological resilience in arid environments. The adaptation of desert antelopes appears to represent a stepwise evolutionary trajectory shaped by both ancestral inheritance and selective pressures. These findings advance our understanding of adaptive evolution in extreme environments and inform both conservation efforts and the potential improvement of livestock resilience in arid regions.

DATA AVAILABILITY

The genome assembly and all sequencing data have been deposited in NCBI under BioProjectID PRJNA1009402, Genome Sequence Archive (<https://ngdc.cnca.ac.cn/gsa>) under accession PRJCA034978, and Science Data Bank (www.scidb.cn/en/anonymous/emVVtkJ6, doi: 10.57760/sciencedb.19938).

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

L.C. and B.W. designed and managed the project. W.C. collected and prepared the samples. J.Z. performed genome assembly and gene

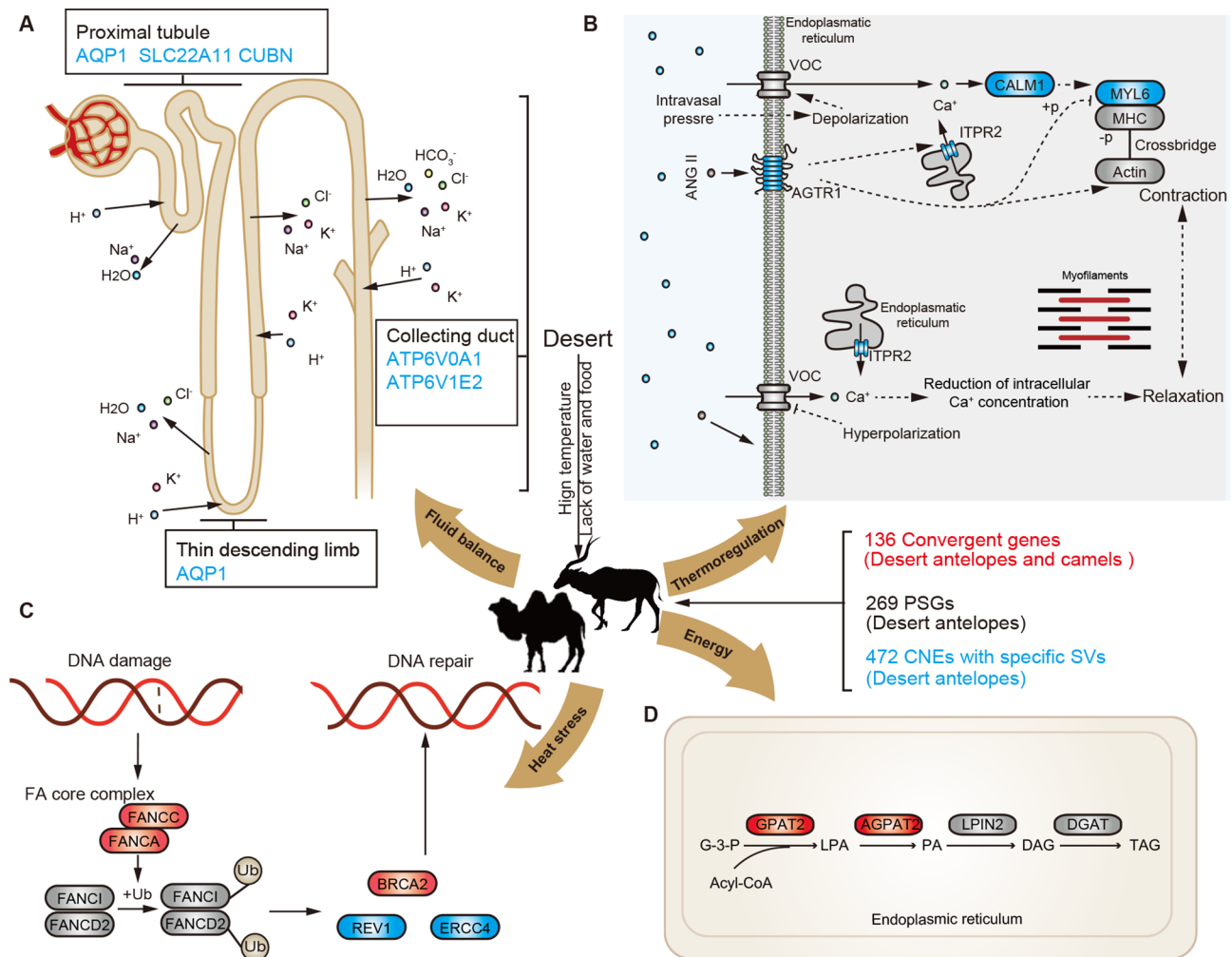


Figure 6 Relevant pathways for desert adaptation in mammals

A: Water reabsorption and fluid balance regulation mediated by renal tubules. B: Blood pressure regulation mediated by smooth muscle contraction and dilation of blood vessels. C: DNA repair mediated by the Fanconi anemia pathway. D: Synthesis of triglycerides.

annotation. J.Z., X.F.Z., and B.L. performed the evolutionary analyses. H.S.Y. conducted the experiments. H.S.Y., Z.C.D., and Y.F.M. contributed to figure design. J.Z., L.C., and B.T.Z. drafted the manuscript. J.Z., B.T.Z., Z.H.L., B.W., W.W., D.D.W., G.H., and L.C. revised the manuscript. All authors read and approved the final version of the manuscript.

REFERENCES

- Abramson J, Adler J, Dunger J, et al. 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, **630**(8016): 493–500.
- Alonge M, Soyk S, Ramakrishnan S, et al. 2019. RaGOO: fast and accurate reference-guided scaffolding of draft genomes. *Genome Biology*, **20**(1): 224.
- Alvira-Iraizoz F, Gillard BT, Lin PJ, et al. 2021. Multiomic analysis of the Arabian camel (*Camelus dromedarius*) kidney reveals a role for cholesterol in water conservation. *Communications Biology*, **4**(1): 779.
- Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research*, **27**(2): 573–580.
- Beuchat CA. 1990. Body size, medullary thickness, and urine concentrating ability in mammals. *American Journal of Physiology: Regulatory, Integrative and Comparative Physiology*, **258**(2): R298–R308.
- Bibi F. 2013. A multi-calibrated mitochondrial phylogeny of extant Bovidae (Artiodactyla, Ruminantia) and the importance of the fossil record to systematics. *BMC Evolutionary Biology*, **13**(1): 166.
- Birney E, Clamp M, Durbin R. 2004. GeneWise and genomewise. *Genome Research*, **14**(5): 988–995.

- Blanchette M, Kent WJ, Riemer C, et al. 2004. Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Research*, **14**(4): 708–715.
- Camacho C, Coulouris G, Avagyan V, et al. 2009. BLAST+: architecture and applications. *BMC Bioinformatics*, **10**(1): 421.
- Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I, et al. 2022. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Research*, **50**(D1): D165–D173.
- Catt KJ, Mendelsohn FAC, Millan MA, et al. 1984. The role of angiotensin II receptors in vascular regulation. *Journal of Cardiovascular Pharmacology*, **6**: S575–S586.
- Chang CC, Chow CC, Tellier LC, et al. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, **4**(1): 7.
- Chen L, Qiu Q, Jiang Y, et al. 2019. Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. *Science*, **364**(6446): eaav6202.
- Chen SF, Zhou YQ, Chen YR, et al. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, **34**(17): i884–i890.
- Cingolani P, Platts A, Wang LL, et al. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w¹¹¹⁸; iso-2; iso-3. *Fly*, **6**(2): 80–92.
- Claro F, Hayes H, Cribiu EP. 1996. The karyotype of the addax and its comparison with karyotypes of other species of Hippotraginae antelopes.

Hereditas, **124**(3): 223–227.

Csaki LS, Dwyer JR, Fong LG, et al. 2013. Lipins, lipinopathies, and the modulation of cellular lipid storage and signaling. *Progress in Lipid Research*, **52**(3): 305–316.

Danecek P, Auton A, Abecasis G, et al. 2011. The variant call format and VCFtools. *Bioinformatics*, **27**(15): 2156–2158.

Dawicki-McKenna JM, Langelier MF, DeNizio JE, et al. 2015. PARP-1 activation requires local unfolding of an autoinhibitory domain. *Molecular Cell*, **60**(5): 755–768.

De Vos M, Schreiber V, Dantzer F. 2012. The diverse roles and clinical relevance of PARPs in DNA damage repair: current state of the art. *Biochemical Pharmacology*, **84**(2): 137–146.

Edelman NB, Frandsen PB, Miyagi M, et al. 2019. Genomic architecture and introgression shape a butterfly radiation. *Science*, **366**(6465): 594–599.

Fenton RA, Knepper MA. 2007. Mouse models and the urinary concentrating mechanism in the new millennium. *Physiological Reviews*, **87**(4): 1083–1112.

Fernández MH, Vrba ES. 2005. A complete estimate of the phylogenetic relationships in Ruminantia: a dated species-level supertree of the extant ruminants. *Biological Reviews*, **80**(2): 269–302.

Flouri T, Jiao XY, Rannala B, et al. 2020. A bayesian implementation of the multispecies coalescent model with introgression for phylogenomic analysis. *Molecular Biology and Evolution*, **37**(4): 1211–1223.

Fu WW, Wang R, Nanaei HA, et al. 2022. RGD v2.0: a major update of the ruminant functional and evolutionary genomics database. *Nucleic Acids Research*, **50**(D1): D1091–D1099.

Göransson V, Johnsson C, Nylander O, et al. 2002. Renomedullary and intestinal hyaluronan content during body water excess: a study in rats and gerbils. *The Journal of Physiology*, **542**(1): 315–322.

Grant CE, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics*, **27**(7): 1017–1018.

Haas BJ, Salzberg SL, Zhu W, et al. 2008. Automated eukaryotic gene structure annotation using EVIDENCEModeler and the Program to Assemble Spliced Alignments. *Genome Biology*, **9**(1): R7.

Harris RS. 2007. Improved Pairwise Alignment of Genomic DNA. Ph. D. dissertation, The Pennsylvania State University, State College.

Hempel E, Westbury MV, Grau JH, et al. 2021. Diversity and paleodemography of the addax (*Addax nasomaculatus*), a saharan antelope on the verge of extinction. *Genes*, **12**(8): 1236.

Herbert TD. 2023. The mid-Pleistocene climate transition. *Annual Review of Earth and Planetary Sciences*, **51**: 389–418.

Hu J, Fan JP, Sun ZY, et al. 2020. NextPolish: a fast and efficient genome polishing tool for long-read assembly. *Bioinformatics*, **36**(7): 2253–2255.

Hu J, Wang Z, Sun ZY, et al. 2024. NextDenovo: an efficient error correction and accurate assembly tool for noisy long reads. *Genome Biology*, **25**(1): 107.

Huang N, Li H. 2023. Compleasm: a faster and more accurate reimplementation of BUSCO. *Bioinformatics*, **39**(10): btad595.

Hubisz MJ, Pollard KS, Siepel A. 2011. PHAST and RPHAST: phylogenetic analysis with space/time models. *Briefings in Bioinformatics*, **12**(1): 41–51.

Humble E, Dobrynin P, Senn H, et al. 2020. Chromosomal - level genome assembly of the scimitar - horned oryx: insights into diversity and demography of a species extinct in the wild. *Molecular Ecology Resources*, **20**(6): 1668–1681.

Humble E, Stoffel MA, Dicks K, et al. 2023. Conservation management strategy impacts inbreeding and mutation load in scimitar-horned oryx. *Proceedings of the National Academy of Sciences of the United States of America*, **120**(18): e2210756120.

IUCN. 2024. IUCN red list of threatened species. <https://www.iucnredlist.org/en>.

Jebb D, Huang ZX, Pippel M, et al. 2020. Six reference-quality genomes

reveal evolution of bat adaptations. *Nature*, **583**(7817): 578–584.

Jurka J, Kapitonov VV, Pavlicek A, et al. 2005. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research*, **110**(1–4): 462–467.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, **30**(4): 772–780.

Kent WJ, Baertsch R, Hinrichs A, et al. 2003. Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes. *Proceedings of the National Academy of Sciences of the United States of America*, **100**(20): 11484–11489.

Kingdon J. 2014. Mammals of Africa: Volume VI: Hippopotamuses, Pigs, Deer, Giraffe and Bovids. A&C Black.

Kouyama R, Suganami T, Nishida J, et al. 2005. Attenuation of diet-induced weight gain and adiposity through increased energy expenditure in mice lacking angiotensin II type 1a receptor. *Endocrinology*, **146**(8): 3481–3489.

Krausman PR, Casey AL. 2007. Addax nasomaculatus. *Mammalian Species*, (807): 1–4.

Krzywinski M, Schein J, Birol I, et al. 2009. Circos: an information aesthetic for comparative genomics. *Genome Research*, **19**(9): 1639–1645.

Kuleshov MV, Jones MR, Rouillard AD, et al. 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*, **44**(W1): W90–W97.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods*, **9**(4): 357–359.

Le L, Escobar IE, Ho T, et al. 2020. SLC45A2 protein stability and regulation of melanosome pH determine melanocyte pigmentation. *Molecular Biology of the Cell*, **31**(24): 2687–2702.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**(14): 1754–1760.

Li H, Durbin R. 2011. Inference of human population history from individual whole-genome sequences. *Nature*, **475**(7357): 493–496.

Li H, Handsaker B, Wysoker A, et al. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**(16): 2078–2079.

Li HL, Russo A, Diantonio A. 2019. SIK3 suppresses neuronal hyperexcitability by regulating the glial capacity to buffer K⁺ and water. *Journal of Cell Biology*, **218**(12): 4017–4029.

Liman ER, Kinnamon SC. 2021. Sour taste: receptors, cells and circuits. *Current Opinion in Physiology*, **20**: 8–15.

Liu ZJ, Liu GJ, Hailer F, et al. 2016. Dietary specialization drives multiple independent losses and gains in the bitter taste gene repertoire of Laurasiatherian Mammals. *Frontiers in Zoology*, **13**(1): 28.

Mackenzie B, Erickson JD. 2004. Sodium-coupled neutral amino acid (System N/A) transporters of the SLC38 gene family. *Pflügers Archiv*, **447**(5): 784–795.

Malinsky M, Matschiner M, Svardal H. 2021. Dsuite - Fast D - statistics and related admixture evidence from VCF files. *Molecular Ecology Resources*, **21**(2): 584–595.

Manni M, Berkeley MR, Seppey M, et al. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Molecular Biology and Evolution*, **38**(10): 4647–4654.

Markovich D. 2011. Physiological roles of renal anion transporters NaS1 and Sat1. *American Journal of Physiology: Renal Physiology*, **300**(6): F1267–F1270.

Minh BQ, Schmidt HA, Chernomor O, et al. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, **37**(5): 1530–1534.

Mirdita M, Schütze K, Moriawaki Y, et al. 2022. ColabFold: making protein folding accessible to all. *Nature Methods*, **19**(6): 679–682.

Nelson G, Hoon MA, Chandrashekar J, et al. 2001. Mammalian sweet taste

- receptors. *Cell*, **106**(3): 381–390.
- Olesen ETB, Fenton RA. 2013. Is there a role for PGE2 in urinary concentration?. *Journal of the American Society of Nephrology*, **24**(2): 169–178.
- Page AJ, Taylor B, Delaney AJ, et al. 2016. *SNP-sites*: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microbial Genomics*, **2**(4): e000056.
- Pennacchio LA, Ahituv N, Moses AM, et al. 2006. *In vivo* enhancer analysis of human conserved non-coding sequences. *Nature*, **444**(7118): 499–502.
- Pickrell JK, Pritchard JK. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, **8**(11): e1002967.
- Pond SLK, Frost SDW, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics*, **21**(5): 676–679.
- Preston GM, Carroll TP, Guggino WB, et al. 1992. Appearance of water channels in *Xenopus* oocytes expressing red cell CHIP28 protein. *Science*, **256**(5055): 385–387.
- Qu CX, Mao CY, Xiao P, et al. 2021. Ligand recognition, unconventional activation, and G protein coupling of the prostaglandin E₂ receptor EP2 subtype. *Science Advances*, **7**(14): eabf1268.
- Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**(6): 841–842.
- Rambaut A, Drummond AJ, Xie D, et al. 2018. Posterior summarization in Bayesian phylogenetics using tracer 1.7. *Systematic Biology*, **67**(5): 901–904.
- Rocha JL, Brito JC, Nielsen R, et al. 2021a. Convergent evolution of increased urine - concentrating ability in desert mammals. *Mammal Review*, **51**(4): 482–491.
- Rocha JL, Godinho R, Brito JC, et al. 2021b. Life in deserts: the genetic basis of mammalian desert adaptation. *Trends in Ecology & Evolution*, **36**(7): 637–650.
- Rouleau-Turcotte E, Krastev DB, Pettitt SJ, et al. 2022. Captured snapshots of PARP1 in the active state reveal the mechanics of PARP1 allostery. *Molecular Cell*, **82**(16): 2939–2951.
- Rügheimer L, Johnsson C, Maric C, et al. 2008. Hormonal regulation of renomedullary hyaluronan. *Acta Physiologica*, **193**(2): 191–198.
- Rulten SL, Fisher AEO, Robert I, et al. 2011. PARP-3 and APLF function together to accelerate nonhomologous end-joining. *Molecular Cell*, **41**(1): 33–45.
- Sanderson MJ. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics*, **19**(2): 301–302.
- Sayyari E, Whitfield JB, Mirarab S. 2018. DiscoVista: interpretable visualizations of gene tree discordance. *Molecular Phylogenetics and Evolution*, **122**: 110–115.
- Schmidt-Nielsen K. 1959. The physiology of the camel. *Scientific American*, **201**(6): 140–151.
- Sejian V, Bhatta R, Gaughan JB, et al. 2018. Review: adaptation of animals to heat stress. *Animal*, **12 Suppl 2**: s431–s444.
- Sekine T, Cha SH, Endou H. 2000. The multispecific organic anion transporter (OAT) family. *Pflügers Archiv*, **440**(3): 337–350.
- Stanke M, Keller O, Gunduz I, et al. 2006. AUGUSTUS: *ab initio* prediction of alternative transcripts. *Nucleic Acids Research*, **34**(S2): W435–W439.
- Stegbauer J, Gurley SB, Sparks MA, et al. 2011. AT₁ receptors in the collecting duct directly modulate the concentration of urine. *Journal of the American Society of Nephrology*, **22**(12): 2237–2246.
- Su W, Cao R, Zhang XY, et al. 2020. Aquaporins in the kidney: physiology and pathophysiology. *American Journal of Physiology: Renal Physiology*, **318**(1): F193–F203.
- Terhorst J, Kamm JA, Song YS. 2017. Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nature Genetics*, **49**(2): 303–309.
- The ENCODE Project Consortium, Moore JE, Purcaro MJ, et al. 2020. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature*, **583**(7818): 699–710.
- Van der Auwera GA, Carneiro MO, Hartl C, et al. 2013. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics*, **43**(1110): 11.10. 1–11.10. 33.
- Vaser R, Adusumalli S, Leng SN, et al. 2016. SIFT missense predictions for genomes. *Nature Protocols*, **11**(1): 1–9.
- Verhagen JMA, Huijman JG, Williams M, et al. 2012. Incidental finding of alpha - methylacyl - CoA racemase deficiency in a patient with oculocutaneous albinism type 4. *American Journal of Medical Genetics Part A*, **158A**(11): 2931–2934.
- Walden H, Deans AJ. 2014. The fanconi anemia DNA repair pathway: structural and functional insights into a complex disorder. *Annual Review of Biophysics*, **43**: 257–278.
- Wang K, Li MY, Hakonarson H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, **38**(16): e164.
- Ward D. 2016. *The Biology of Deserts*. 2nd ed. Oxford: Oxford University Press.
- Woolfe A, Goodson M, Goode DK, et al. 2004. Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biology*, **3**(1): e7.
- Wu HG, Guang XM, Al-Fageeh MB, et al. 2014. Camelid genomes reveal evolution and adaptation to desert environments. *Nature Communications*, **5**: 5188.
- Xu SH, He ZW, Guo ZX, et al. 2017. Genome-wide convergence during evolution of mangroves from woody plants. *Molecular Biology and Evolution*, **34**(4): 1008–1015.
- Yang ZH. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, **24**(8): 1586–1591.
- Zhang C, Rabiee M, Sayyari E, et al. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, **19**(S6): 153.
- Zhang L, Bowen T, Grennan-Jones F, et al. 2009. Thyrotropin receptor activation increases hyaluronan production in preadipocyte fibroblasts. *Journal of Biological Chemistry*, **284**(39): 26447–26455.
- Zhang Y, Liu T, Meyer CA, et al. 2008. Model-based analysis of ChIP-seq (MACS). *Genome Biology*, **9**(9): R137.
- Zhou L, Feng TZ, Xu SB, et al. 2022. ggmsa: a visual exploration tool for multiple sequence alignment and associated data. *Briefings in Bioinformatics*, **23**(4): bbac222.