DOI : 10.5958/2277-8934.2024.00061.4 Vol 31 No 3, p 341-354

GENOME-WIDE COMPARATIVE ANALYSES REVEAL SELECTION SIGNATURES UNDERLYING ADAPTATION IN DOMESTIC BACTRIAN AND WILD TWO-HUMPED CAMEL

Siriguleng[1#], Aodungaowa[2#], Wutaodao[1], Zhenji Li[3], Chaijin[4], Duhao[5], Shana Huang[6], Hangfei Gao[6], Naqin[6], Lixian Li[7], Wei Cao[8], Liang Ming[9], Jirimutu[5*], Alateng baolige[5*] and Yi Heng[5*]

[1]Ured Rear Banner Livestock and Aquatic Workstation, Inner Mongolia 015543, China
[2]Animal Disease Prevention and Control Center of Ured Rear Banner, Inner Mongolia 015543, China
[3]Bayannur City Animal Husbandry Service Center, Inner Mongolia 015000, China
[4]Agriculture, Animal Husbandry and Science and Technology Bureau of Ured Rear Banner, Inner Mongolia 015543, China
[5]Ured Rear Banner People's Government, Inner Mongolia 015543, China
[6]Inner Mongolia Agricultural and Animal Husbandry Technology Promotion Center, Inner Mongolia 010013, China
[7]Urad Rear Banner Green Industry Coordination and Development Center, Inner Mongolia 015543, China
[8]Bayannur Agricultural and Animal Husbandry Science Researc h Institute, Inner Mongolia 015199, China
[9]Key Laboratory of Dairy Biotechnology and Bioengineering, Ministry of Education, College of Food Science and Engineering, Inner Mongolia Agricultural University, Hohhot, Inner Mongolia, 010018, China
#First authors, *correspondence authors

ABSTRACT

Bactrian camels are vital large mammals that adapt well to the harsh environment of the desert. Genome-wide selection signatures can provide insights into natural and/or artificial selection and reveal functional genes related to biological characteristics and/or phenotypes. Here, we investigated genomic diversity and structure, and identified selection signatures of domestic Bactrians from China and Mongolia (IMG_D and MG_D) and wild two-humped camels (MG_W). The average sequencing depth reached 12.40× for each population, and more than 4.01, 3.58, and 2.70 million single-nucleotide polymorphisms (SNPs) were detected covering all autosomes and the X chromosome in the IMG_D, MG_D, and MG_W. The population structure suggested gene flows between IMG_D and MG_D, but no strong signal migration between the domestic and wild two-humped camels. Following the F_{ST} and θ_{π} approaches, candidate evolving genes in the camel lineage were significantly enriched in insulin secretion, insulin signaling pathway, lipid metabolism, immune system, and adaptation for desert, which may be the target of selection in domestic Bactrians and wild two-humped camels during the breeding and survival process. Furthermore, screened candidate genes, including ABCC8, KCNJ11, FFAR1, PRKACB, CREB1, PRKACB, ACACA, and SLC2A4, were associated with insulin pathways and putatively related to insulin resistance. We also identified candidate genes and KEGG pathways associated with olfactory transduction and environmental adaptation, implying a greater desert adaptation capacity in Bactrian camels. In conclusion, the present study provides a greater understanding of genome diversity and variations associated with adaptive and biological characteristics in Bactrian camels.

Key words: Domestic Bactrian, selection signature, genome-wide, wild two-humped camel

The Bactrian and dromedary camels are economically important livestock in desert and semi-desert areas, providing meat, milk, and wool and serving as a vital mode of transportation for local residents (Saipolda, 2004). According to current Food and Agriculture Organisation (FAO) statistics only 5% are Bactrian (two-humped) camels (Sikkema *et al,* 2019). Bactrian camels (*Camelus bactrianus*) are the last domesticated large mammals whose ancestors were primarily in Northeast and Central Asia approximately 4450 years ago (Ming *et al,* 2020). Today, more than 90% of Bactrian camels are distributed across China, Mongolia, and Kazakhstan, numbering about 952,000 (Ming *et al,* 2022), and formed different breeds or populations from different geographical locations, which have constituted an extensive genetic resource pool. In addition, wild two-humped camels (*Camelus ferus*), with morphological similarities to its domestic counterpart, is a critically endangered ungulate that inhabits Central Asia's

SEND REPRINT REQUEST TO JIRIMUTU email: yeluotuo1999@vip.163.com

desert ecosystems (Yadamsuren *et al,* 2019) and stores genetic resources differently from domestic Bactrian camels. In recent years, genetic diversity in domestic Bactrians and wild two-humped camels have been recognised as an important attribute (Ming *et al,* 2020; Ming *et al,* 2017; Yi *et al,* 2017; Ming *et al,* 2021), and the conservation of camels' genetic diversity is urgent for improving production and balancing the ecological environment (Burger *et al,* 2019).

Through long-term natural and artificial selection, such as climate change, environmental pressure, human migration and socioeconomic practices, Bactrian camels shaped their unique adaptability to harsh environmental conditions in desert and semi-desert areas—cold, hot, arid, poor grazing, and even food shortages. They can withstand extreme thirst and hunger for prolonged periods (Ali *et al,* 2019). Long-term selection leads to different phenotypic structures in the Bactrian camel population across world regions (Ming *et al,* 2022). These changes have left their footprint on specific regions of the organism genome as the selection signature associated with adaptive and productive phenotypes (Bigham *et al,* 2010; Lv *et al,* 2014). Detecting selection signatures can shed light on the processes involved in genome evolution and determine functional gene or genomic regions associated with unique biological characteristics and economic traits (Hayes *et al,* 2009; Nielsen, 2005).

Domestic Bactrians and wild two-humped camels belong to the genus Camelus. However, they originated from different ancestors and underwent different evolutionary processes (Ji *et al,* 2009; Jirimutu *et al,* 2012). To explore their genetic variance and identify candidate regions and genes related to important traits, we resequenced whole genomes belonging to 39 domestic Bactrians from China and combined whole genome sequencing from NCBI (National Centre for Biotechnology Information) database of domestic Bactrians and wild two-humped camels in Mongolia to detect a within and between domestic and wild comparative selection signature analysis. We aimed to provide a theoretical basis for improving economically important traits in Bactrian camels and further insight into the mechanisms underlying adaptation to extreme desert environments.

Materials and Methods

Ethics statement

The procedures and protocols were approved by the animal care committee of the Camel Protection Association of Inner Mongolia. The research was supported by the Review Committee for the Use of Human or Animal Subjects of the Food Science and Engineering College of Inner Mongolia Agricultural University (Hohhot, China). All experimental procedures used in this study were conducted in compliance with the ARRIVE guidelines (https://arriveguidelines.org).

Sample collection

Thirty nine Gobi Red Bactrian camel samples (IMG_D) were sampled from Bayan Nur city, Inner Mongolia, China. The Gobi Red Bactrian is an ancient population raised for milk, wool, and meat production. Notably, Gobi Red Bactrian has undergone a long-term natural and artificial selection and is characterised by high milk production (3.6-4.2 kg/day), red coats, and excellent wool quality (long fibre, fine velvet, and high yield). We collected 5 mL of blood in EDTA anticoagulant tubes from each camel's jugular vein after disinfection treatment, which was then stored at -80°C until further processing.

DNA extraction and sequencing

Blood sample DNA was isolated using the QIAamp DNA Blood Mini Kit (Qiagen) according to the manufacturer's instructions. The quality and integrity of DNA were detected by the OD260/280 ratio and a 2% agarose gel electrophoresis. The DNA concentration was controlled using the Quant-iT PicoGreen dsDNA Reagent Kit (Thermo Fisher Scientific, USA) according to the manufacturer's instructions. Libraries were prepared using Illumina kits and then sequenced on the Illumina HiSeq platform (Illumina; CA, USA) with the standard paired-end mode. By combining the raw data by sequencing with the genome data from the database, filtration was performed as follows: (1) Reads containing adapter sequences were removed; (2) discarded bases with consecutive quality <20 for both ends of the sequencing read; (3) removed the final length of the sequencing read, which was <50 bp; (4) removed reads with >5% of unknown bases. We based all the bioinformatic analyses on the filtration data (clean reads) from the Illumina quality control filter.

Data Availability

The datasets generated and analysed during the persent study are available with the National Centre for Biotechnology Information (NCBI) Sequence Read Archive (SRA) GenBank repository, data-base under accession number PRJNA890442 (https://www.

ncbi.nlm.nih.gov/sra/PRJNA890442). Additionally, whole genome resequencing data from 16 Mongolian domestic Bactrians (MG_D) and 13 Mongolian wild two-humped camels (MG_W) were obtained from Sequence Read Archive at the NCBI (SAMN06759127-132, 134, 136, 141-145, 162-163, 169-178, 182-184, 187).

Read alignment and variant calling and annotation

We used the Burrows–Wheeler Aligner (v0.7.9a, MEM) (Li and Durbin, 2009) to map clean reads to the reference genome assembly of the wild two-humped camel (*Camelus ferus*, VSZR00000000; Assembly: GCA_009834535.1) (Ming *et al,* 2020). Duplicates reads were discarded using the Picard MarkDuplicates tool (v1.115). Statistics of alignment rate, coverage, and sequencing depth were performed on the deduplicated data. The sample alignment rate reflects the similarity between the sample sequencing data and the reference genome. The coverage depth directly reflects the homogeneity of the sequencing data and homology of the reference sequence. We followed the Genome Analysis Toolkit (GATK) (DePristo *et al,* 2011) pipeline for variant calling and ANNOVAR (Wang *et al,* 2010) was performed to assign SNPs.

Phylogenetic and population structure

In order to ensure the reliability of subsequent population analyses, raw variants were filtered according to the following criteria: minor allele frequency >5%, the proportion of the sample covered by SNP to the total sample >90%, and variants with <20% of individuals missing genotypes. We calculated the distance matrix with Treebest software (Vilella *et al,* 2009) and constructed a phylogenetic tree by the neighbor-joining method. We identified the top four principal components accounting for variation in the dataset. The population ancestry was inferred by Admixture (v1.3.0)v (Alexander *et al,* 2009) based on Bayesian mathematical models. Migration events among camel populations were inferred using TreeMix (v1.12) (Pickrell and Pritchard, 2012). We estimated the optimal number of ancestral clusters K with the cross-validation error. The filtered SNP set was also used to estimate genome-wide linkage disequilibrium (LD). The LD decay was calculated with PopLDdecay (Zhang *et al,* 2019) using default parameters.

Calculation of θπ and Fst

We used a sliding-window approach (100 kb windows sliding in 50 kb steps) quantify polymorphism levels (θπ, pairwise nucleotide variation as a measure of variability), and genetic differentiation (FST) between domestic Bactrian and wild two-humped camels.

Identification of selected regions

To detect regions associated with selective sweep, we calculated the distribution of the θπ ratios (θπ, wild/ θπ, domestic) and FST values according to Li *et al* (2013). We simultaneously used empirical procedures and selected windows with significantly low and high θπ ratios (the 5% left and right tails) and significantly high FST values (the 5% right tail) of the empirical distribution as regions with strong selective sweep signals along the genome, which could harbour genes that underwent a selective sweep.

Functional enrichment analysis

Functional classification of GO (Genen Ontology) categories was performed using Blast2GO (Conesa and Götz, 2008). We also tested each set of putative positively selected genes for overrepresentation of the KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways. The P values were adjusted by FDR (False Discovery Rate) and the adjusted P value cut-off was 0.05.

Results

Genome data analysis and SNP identification

The present study reports on the whole genome sequencing of the Gobi Red Bactrian camel originating from Inner Mongolia, China. We characterised the genetic variations of combined Mongolian domestic and wild two-humped camel whole genome data from the National Centre for Biotechnology Information (NCBI). All methods in this study were performed in accordance with the relevant guidelines and regulations. We integrated and aligned 2×125 bp of paired-end reads, more than 1200 Gb, using Burrows–Wheeler Aligner (BWA) (Li and Durbin, 2009) software for the wild two-humped camel reference genome assembly. The average sequencing coverage rate reached 96.84%, and we thus inferred that the sequencing data covered most of the genome.

We achieved an average sequencing depth of 12.40× for each population and more than 97.77% of the sequence reads were mapped to the reference genome, indicating that the sequencing data covered most of the genome, and high-quality sequences were obtained (Table S1). The number of detected single-nucleotide polymorphisms (SNPs) for the IMG_D, MG_D, and MG_W camels were 4.01, 3.58, and 2.70

Table S1. Whole genome resequencing for 39 domestic Bactrians from China, 16 domestic Bactrians from Mongolia and 13 wild two-humped camels from Mongolia.

NO.	Sample	Clean reads	Clean Bases (bp)	Sequencing depth (×)	Coverage Rate (%)	Mapping Rate (%)	Singletons mapped(%)
1	IMG_D_1	233,377,308	34,163,102,124	13.25	98.5	98.28	0.36
2	IMG_D_2	235,673,208	34,590,568,017	12.54	98.59	98.13	0.33
3	IMG_D_3	230,639,824	33,926,116,838	12.00	98.28	98.05	0.33
4	IMG_D_4	230,481,092	33,914,522,791	12.17	98.48	98.27	0.31
5	IMG_D_5	259,457,164	37,977,083,433	14.48	98.56	98.07	0.35
6	IMG_D_6	233,882,908	34,060,358,933	13.47	98.53	98.05	0.40
7	IMG_D_7	272,247,348	40,076,051,109	13.85	98.42	98.37	0.30
8	IMG_D_8	240,660,096	35,427,587,595	12.72	98.38	98.34	0.35
9	IMG_D_9	229,351,106	33,535,939,131	13.01	98.44	97.97	0.37
10	IMG_D_10	226,821,290	33,183,779,029	13.03	98.48	98.11	0.36
11	IMG_D_11	258,208,284	37,749,957,795	14.35	98.45	98.23	0.37
12	IMG_D_12	235,661,296	34,400,249,280	13.50	98.60	98.15	0.39
13	IMG_D_13	264,040,292	38,610,372,980	14.92	98.51	97.96	0.39
14	IMG_D_14	248,890,494	36,440,198,986	14.15	98.58	98.17	0.35
15	IMG_D_15	250,601,884	36,896,180,928	12.92	98.33	98.12	0.32
16	IMG_D_16	252,689,076	37,150,798,850	13.00	98.48	98.25	0.33
17	IMG_D_17	275,330,276	406,12,701,892	14.85	98.52	98.15	0.32
18	IMG_D_18	262,921,260	38,802,491,675	14.22	98.6	98.04	0.32
19	IMG_D_19	247,285,228	36,320,848,848	14.02	98.49	98.22	0.33
20	IMG_D_20	250,795,352	36,841,958,919	13.35	98.46	98.08	0.37
21	IMG_D_21	235,153,794	34,517,978,197	13.25	98.43	98.17	0.36
22	IMG_D_22	256,840,936	37,753,221,093	13.90	98.60	98.00	0.37
23	IMG_D_23	264,183,004	38,901,619,901	13.52	98.67	98.17	0.34
24	IMG_D_24	256,203,234	37,783,618,394	13.57	98.45	98.13	0.31
25	IMG_D_25	253,721,376	37,403,517,863	13.63	98.49	98.17	0.33
26	IMG_D_26	255,661,860	37,518,440,629	14.26	98.53	98.08	0.34
27	IMG_D_27	234,668,174	34,419,215,228	13.25	98.44	97.97	0.33
28	IMG_D_28	227,115,944	33,337,568,585	12.98	98.33	98.28	0.30
29	IMG_D_29	250,766,240	36,816,697,912	14.00	98.47	97.97	0.33
30	IMG_D_30	230,007,272	33,729,098,117	13.09	98.29	98.26	0.32
31	IMG_D_31	245,052,612	35,970,636,344	13.76	98.42	97.92	0.35
32	IMG_D_32	270,646,976	39,673,054,313	10.05	98.16	98.16	0.33
33	IMG_D_33	238,430,504	34,959,568,806	13.36	98.39	97.83	0.37
34	IMG_D_34	261,378,634	38,477,622,495	13.28	98.25	98.16	0.32
35	IMG_D_35	227,745,594	33,402,148,164	12.97	98.45	98.24	0.33
36	IMG_D_36	225,840,872	33,059,604,073	13.07	98.33	98.30	0.33
37	IMG_D_37	241,568,264	35,400,380,850	13.65	98.31	98.18	0.36
38	IMG_D_38	228,566,354	33,490,388,190	12.97	98.54	98.06	0.34
39	IMG_D_39	245,314,832	35,830,593,946	13.82	98.42	98.08	0.34
40	IMG_D_40	234,122,566	34,479,066,182	12.59	98.41	98.31	0.30
Average		245,550,096	36,040,122,711	13.37	98.45	98.14	0.34
1	MG_D_126	206,602,240	24,317,444,074	11.09	92.82	97.28	0.73
2	MG_D_127	186,533,826	21,566,060,531	9.96	95.38	97.66	0.61

3	MG_D_135	218,518,034	26,477,212,053	12.27	98.27	97.75	0.57
4	MG_D_136	244,234,498	29,608,862,005	13.61	98.23	97.12	0.67
5	MG_D_137	229,412,608	27,216,482,487	12.60	94.11	97.8	0.63
6	MG_D_138	210,662,068	24,344,621,671	11.12	89.67	97.56	0.78
7	MG_D_139	189,265,650	21,954,426,844	10.12	89.00	97.84	0.72
8	MG_D_140	230,032,266	28,051,883,744	12.99	98.21	97.70	0.50
9	MG_D_141	214,667,082	26,189,095,619	12.11	89.71	97.62	0.57
10	MG_D_142	228,685,732	27,870,865,144	12.85	89.5	97.59	0.59
11	MG_D_143	210,397,592	25,808,492,572	11.89	98.34	97.55	0.51
12	MG_D_144	232,321,890	28,331,227,058	13.00	98.28	97.31	0.64
13	MG_D_148	213,008,706	25,962,184,933	12.00	98.66	97.89	0.60
14	MG_D_149	192,585,844	22,324,278,455	10.22	89.14	97.41	0.95
15	MG_D_150	229,141,296	27,907,655,650	12.98	88.40	97.97	0.55
16	MG_D_153	211,826,506	25,881,058,295	11.96	98.29	97.71	0.56
Average		215,493,490	25,863,240,696	11.92	94.13	97.61	0.64
1	MG_W_088	291,402,732	28,755,444,143	13.32	98.48	97.82	0.61
2	MG_W_089	353,854,564	35,214,965,866	16.29	98.94	97.73	0.58
3	MG_W_090	298,721,994	29,131,818,611	13.46	98.55	97.29	0.59
4	MG_W_091	281,696,446	27,426,967,391	12.69	97.49	97.61	0.63
5	MG_W_092	248,797,732	23,578,570,077	10.85	93.77	97.54	0.59
6	MG_W_093	215,445,682	21,161,872,442	9.52	97.89	97.84	0.54
7	MG_W_095	187,680,432	18,403,139,067	8.18	97.63	97.14	0.58
8	MG_W_097	325,823,712	32,357,429,084	14.70	98.80	97.64	0.51
9	MG_W_102	236,433,494	23,506,551,016	10.62	98.53	97.67	0.52
10	MG_W_103	231,188,056	22,934,262,953	10.25	98.4	97.43	0.58
11	MG_W_104	299,329,996	29,557,118,508	13.51	98.91	97.73	0.52
12	MG_W_105	282,630,858	28,023,934,404	12.82	98.77	97.78	0.51
13	MG_W_106	193,435,104	19,247,570,617	8.71	97.08	97.26	0.58
Average		265,110,831	26,099,972,629	11.92	97.94	97.58	0.59

IMG_D: Gobi Red Bactrian from Inner Mongolia, China. MG_D: Mongolian domestic Bactrian camel, from Mongolia. MG_W: Mongolian wild two-humped camel, from Mongolia.

million variants using the Genome Analysis Toolkit. The Ts/Tv ratio of the SNPs was 2.16, 2.21, and 2.17, respectively (Table 1), which agreed with previous research on the domestic Bactrian camel (2.18) (Ming *et al,* 2020), but were slightly lower than wild two-humped camels (2.26) and dromedaries (2.31 and 2.34) (Fitak *et al,* 2016; Khalkhali-Evrigh *et al,* 2018).

Table 1. Summary of identified variants for domestic Bactrians from China and Mongolia, and wild two-humped camels from Mongolia.

Populations	IMG_D	MG_D	MG_W
Number of SNPs	4,011,826	3,584,401	2,703,238
Ts/Tv	2.16	2.21	2.17
Heterozygous SNPs	2,326,859	2,043,109	1,811,169
Heterozygous ratio (%)	0.58	0.57	0.67
Homozygous SNPs	1,684,967	1,541,292	892,069
Homozygous ratio (%)	0.42	0.43	0.33

IMG_D: Gobi Red Bactrian from Inner Mongolia, China. MG_D: Mongolian domestic Bactrian camel from Mongolia. MG_W: Mongolian wild two-humped camel from Mongolia.

We also identified heterozygous 2,326,859, 2,043,109, and 1,811,169 SNPs in the IMG_D, MG_D, and MG_W populations. The heterozygosity ratio was similar in domestic Bactrian camels from China and Mongolia, and slightly higher in wild two-humped camels across whole genomes. Furthermore, the SNPs were classified at the chromosome level (Table 2).

SNP annotation and functional classification.

In the high-quality SNPs from three Bactrian camel populations, most of which were located in intergenic regions (57.79-56.87%), only 0.11% were located in exonic regions (Table S3). This intergenic

region was the most mutated in the Bactrian camel genome, whereas the exonic regions had fewer mutations. Compared to domestic Bactrian camel populations, fewer non-synonymous (12,555) and synonymous (14,774) SNPs in wild two-humped camel were localised within exons, resulting in a non-synonymous/synonymous ratio of 0.8498 (Table 2).

Table 2. The distribution of SNPs in the camel whole genome.

Populations		IMG_D	MG_D	MG_W
Intergenic[a]		2,318,481	2,038,605	1,548,855
ncRNA[b]		1,757	1,629	840
UTR[c]		59,785	56,829	40,340
Intronic		1,588,212	1,444,238	1,083,026
Splicing		188	182	137
exonicg	Synonymous	26,424	21,520	14,774
	Non-synonymous	17,963	17,687	12,555
	Stop altering	251	238	187

[a] Including "intergenic", "upstream", and "downstream" given by ANNOVAR.
[b] Including "ncRNA_exonic", "ncRNA_intronic", "ncRNA_splicing", and "ncRNA_UTR".
[c] Including "UTR5" and "UTR3".

In addition, the non-synonymous SNPs were classified and mapped to the KEGG pathway. Interestingly, 524 genes from IMG_D and 492 genes from MG_D were significantly enriched in the olfactory transduction (ko04740) pathway, whereas only 441 genes from MG_W were enriched in this pathway (Table S4). These results indicated that olfactory transduction was vital to artificial selection during the domestication of Bactrian camels.

Phylogenetic analysis

Principle component analysis (PCA)

To examine the genetic relationship among and within the 3 camel populations, we first completed a PCA analysis. The first and second eigenvectors clearly distinguished the domestic Bactrians and wild two-humped camels (Fig 1a). Unexpectedly, IMG_D and MG_D clustered together, indicating a close genetic relationship due to either close geographic proximity or deliberate intercrossing. Although, IMG_D and MG_D belong to the domestic population, they have distinct morphological characteristics and breeding histories. MG_D is an ancient breed bred by Mongolians and adapted to cold environments. IMG_D is known for its red coat colour and outstanding wool quality.

Phylogenetic tree

Furthermore, a phylogenetic tree was constructed with the filtered SNP using the neighbour-joining algorithm. The NJ (Neighbouring Tree) tree clustered the 3 studied populations into separate genetic groups confirming their genetic distinction (Fig 1b). Consistent with the PCA results, two individuals of MG_D clustered with the IMG_D, confirming a close genetic relationship or gene flow between IMG_D and MG_D Bactrian camels due to close geographic proximity.

Population genetic structure

To estimate the proportion of shared genetic ancestry and/or levels of admixture, we performed a population structure analysis with admixture (Alexander *et al,* 2009) (Fig 1c). The cross-validation procedure supported that K = 2 was optimal (Fig S1), showing a clear division between the wild two-humped camels (MG_W), and domestic Bactrian camels (MG_D and IMG_D). By contrast, domestic Bactrian camels from China and Mongolia were grouped from K = 2 to K = 3 (Fig 1c). Evident introgression of IMG_D camels into MG_D was observed. Consistent with the results of PCA and NJ, the admixture analysis further confirmed, with higher resolution, the intermixed genetic makeup of the two domestic Bactrian camel populations. Another method to infer the camels' population tree was TreeMix (Pickrell and Pritchard, 2012). It is worth mentioning that there was no strong migration signal between the domestic and wild two-humped camels (Fig 2a); however, gene flows from dromedaries to domestic Bactrian camels were identified.

Linkage disequilibrium (LD) analysis

The PopLDdecay software was used to explore genome-wide patterns of LD in each camel population by the default parameters. The IMG_D and MG_D populations had similar and lower LD values, suggesting a relatively early origin of the domestic Bactrian camels. The MG_W population had higher LD values, indicating that they were derived from a relatively small ancestral population (<1000) (Fig 2b).

Genome-wide selection signature analysis

The differentially selected genes and genomic regions by selection signatures have been identified, which are vital to revealing the genetics of economic and adaptive traits. To better understand the underlying genetics of their unique biological properties, the adaptation among domestic and wild two-humped camel populations, the fixation index

Table S2. Genome-wide summary of SNPs from IMG_D, MG_D and MG_W.

Chromosome	SNP Count			Chromosome	SNP Count		
	IMG_D	MG_D	MG_W		IMG_D	MG_D	MG_W
1	660985	555214	349038	20	261828	220856	128018
2	661441	534063	333531	21	166683	139387	91001
3	568323	431077	331244	22	168384	141202	81472
4	360613	298968	231269	23	209926	174802	105974
5	467046	380748	251238	24	199467	157019	108360
6	495467	390327	265460	25	241030	173302	113146
7	392311	327206	222000	26	202939	164138	99326
8	429930	320694	231109	27	159647	137692	92239
9	450362	389531	256474	28	57163	46799	28755
10	371975	315930	202625	29	176777	138295	96807
11	508255	434003	253039	30	181285	136771	94817
12	328280	259162	231019	31	150705	140119	91803
13	421560	303863	189248	32	184404	144969	80215
14	441653	342608	210909	33	130749	110245	77526
15	265101	227771	160315	34	140414	99646	73064
16	349410	268003	167139	35	214226	179007	106212
17	259301	225785	154052	36	320949	262025	194201
18	183787	160376	124721	X	351632	289319	263155
19	318083	250401	166088				

IMG_D: Gobi Red Bactrian from Inner Mongolia, China. MG_D: Mongolian domestic Bactrian camel, from Mongolia. MG_W: Mongolian wild two-humped camel, from Mongolia.

Table S3. Numbers and distribution of SNPs in different camel populations.

Population	IMG_D		MG_D		MG_W	
	Number	Per cent(%)	Number	Per cent(%)	Number	Per cent(%)
Total	4011826	100	3584401	100	2703238	100
UTR5	15765	0.39	15894	0.44	10640	0.39
UTR3	44020	1.10	40935	1.14	29700	1.10
UTR5;UTR3	65	0	59	0	47	0
exonic	42008	1.05	41562	1.16	29084	1.08
splicing	188	0	182	0.10	137	0.01
exonic;splicing	24	0	24	0	18	0
upstream	30087	0.75	29197	0.81	20233	0.75
downstream	31088	0.77	28945	0.81	20653	0.76
upstream;downstream	1306	0.03	1273	0.04	892	0.03
intronic	1588212	39.59	1444238	40.29	1083026	40.06
intergenic	2257306	56.27	1980463	55.25	1507969	55.78
ncRNA_exonic	346	0.01	330	0.01	238	0.01
ncRNA_splicing	1	0	1	0	1	0
ncRNA_intronic	1410	0.04	1298	0.04	601	0.02

IMG_D: Gobi Red Bactrian from Inner Mongolia, China. MG_D: Mongolian domestic Bactrian camel, from Mongolia. MG_W: Mongolian wild two-humped camel, from Mongolia.

(FST), and quantifying polymorphism level (θπ) tests involving IMG_D vs. MG_W, MG_D vs. MG_W, and IMG_D vs. MG_D were performed in 100 kb windows with a 50 kb sliding step. We selected the top 5% of candidate genes for further analysis. Up to 233 genes were positively selected for FST (IMG_D vs MG_W) comparison, 665 genes for FST (MG_D vs MG_W) comparison, and 576 genes for the FST (IMG_D vs MG_D) comparison (Fig 3); these candidates were distributed across different chromosomes.

Functional analysis of candidate genes

Furthermore, candidate genes from domestic Bactrians (IM_D and MG_D) and wild two-humped camels were mapped to the KEGG pathways (Table S8). A few genes were significantly enriched in the adipocytokine signaling pathway (ko04920), insulin signaling pathway (ko04910), insulin secretion (ko04911), B cell receptor signaling pathway (ko04662), IL-17 signaling pathway (ko04657), Th17 cell differentiation (ko04659), fatty acid biosynthesis and degradation (ko00071), glycerophospholipid metabolism (ko00564), glycerolipid metabolism (ko00561), circadian entrainment (ko04713), plant–pathogen interaction (ko04626), and carbohydrate metabolism (ko00030) (Fig S2). These enriched signaling pathways had significant enrichment ($P<0.05$). The findings suggested that insulin signaling, lipid metabolism and the immune system relate to desert adaptation could be the target of

Table 3. Significant KEGG pathway enrichment for candidate genes in domestic Bactrians and wild two-humped camels.

Pathway Hierarchy	KEGG Pathway	Accession Code	Tota Genes	P-value	Gene Name
Endocrine system	Adipocytokine signaling pathway	ko04920	7	4.80×10^{-3}	AKT2, NFKBIE, ACSBG2, NFKBIA, NFKB1, PPARGC1A, MAPK10
	Insulin signaling pathway	ko04910	9	4.69×10^{-2}	EIF4EBP1, PPARGC1A, CBLC, PRKACB, SLC2A4, PDE3B, AKT3, ACACA, PIK3CD
	Insulin secretion	ko04911	16	8.32×10^{-4}	ABCC8, PRKCB, KCNJ11, FFAR1, PCLO, CREB1, PRKCA, RAPGEF4, PRKACB, KCNMB2, ADCY8, GNAQ, CAMK2G, CREB3L2, ADCY1, CACNA1D
Immune system	B cell receptor signaling pathway	ko04662	8	1.10×10^{-3}	PPP3CA, PIK3CD, NFKBIE, AKT2, PLCG2, LOC102517806, NFKB1, NFKBIA
	Th1 and Th2 cell differentiation	ko04658	8	3.76×10^{-3}	NFKBIE, JAK1, NFKB1, NFKBIA, IL12A, PPP3CA, MAML3, MAPK10
	Fc epsilon RI signaling pathway	ko04664	5	2.80×10^{-2}	PLCG2,PIK3CD,PRKCA,MAPK10, AKT2
	IL-17 signaling pathway	ko04657	6	2.95×10^{-2}	MUC5AC, USP25, TNFAIP3, NFKBIA, NFKB1, MAPK10
	T cell receptor signaling pathway	ko04660	7	3.14×10^{-2}	PIK3CD, PPP3CA, PAK4, NFKBIA, NFKB1, AKT2, NFKBIE
	Th17 cell differentiation	ko04659	7	3.32×10^{-2}	RORA, MAPK10, PPP3CA, NFKBIA, NFKB1, JAK1, NFKBIE
	Toll-like receptor signaling pathway	ko04620	6	4.29×10^{-3}	MAPK10, PIK3CD, AKT2, NFKBIA, NFKB1, IL12A
Lipid metabolism	Fatty acid degradation	ko00071	3	3.07×10^{-2}	ECI1, ACADL, ACSBG2
	Glycerophospholipid metabolism	ko00564	6	2.57×10^{-2}	DGKB,DGKH, PHOSPHO1, DGKD, DGKI, CDS1
	Glycerolipid metabolism	ko00561	4	3.48×10^{-3}	DGKB, DGKI, DGKH, DGKD
Environmental adaptation	Circadian entrainment	ko04713	14	3.73×10^{-2}	PRKCA, PRKACB, ADCY8, GNAQ, CREB1, CACNA1H, ADCY1, CACNA1D, CAMK2G, ITPR1, GNB1, PRKG2, GNG7, PRKCB
	Plant–pathogen interaction	ko04626	1	1.49×10^{-2}	HSP90AA1
	Circadian rhythm - fly	Ko04711	1	4.24×10^{-2}	GSK3B
Carbohydrate metabolism	Pentose phosphate pathway	ko00030	2	4.21×10^{-2}	TALDO1,DERA

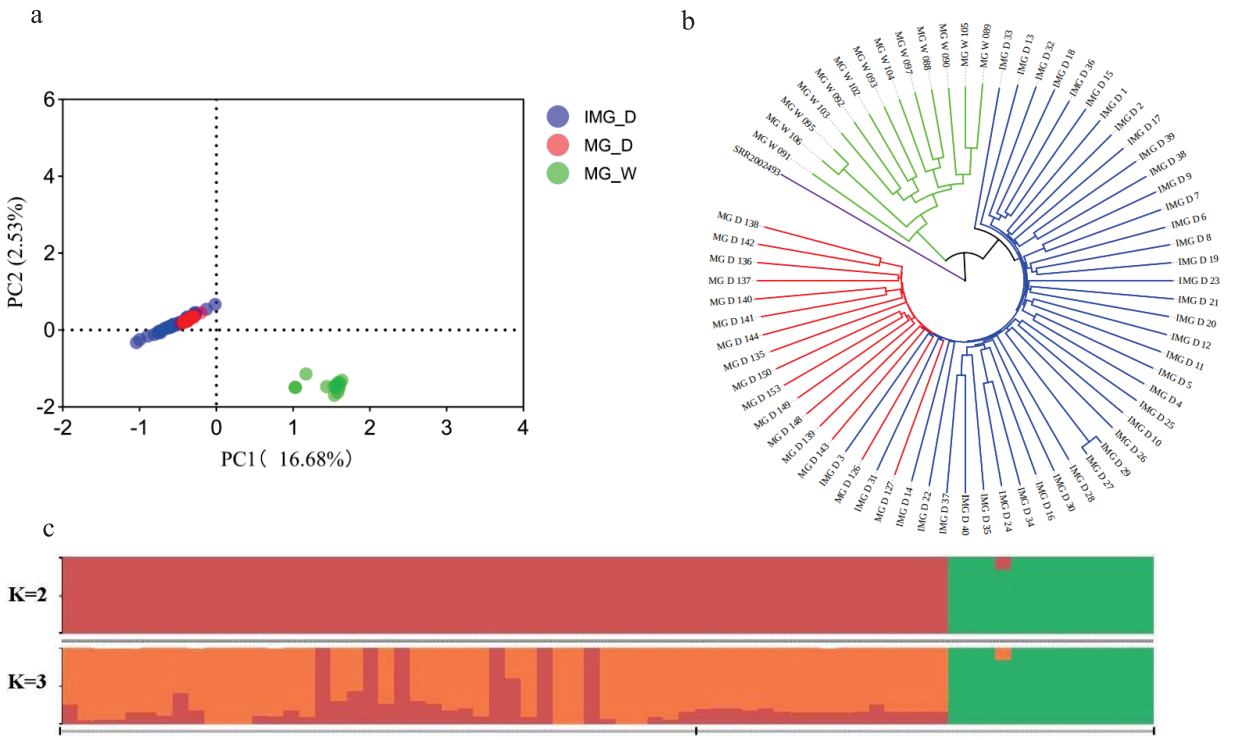

Fig 1. Population genetics analyses of Bactrian camels on genome-wide SNPs. (a) Principal component analysis (PCA) results of 3 Bactrian camel populations. (b) NJ tree constructed using p-distances between individuals. (c) Admixture analysis assuming different numbers of ancestry K. The proportion of an individual's genome assigned to each ancestry is represented by different colours.

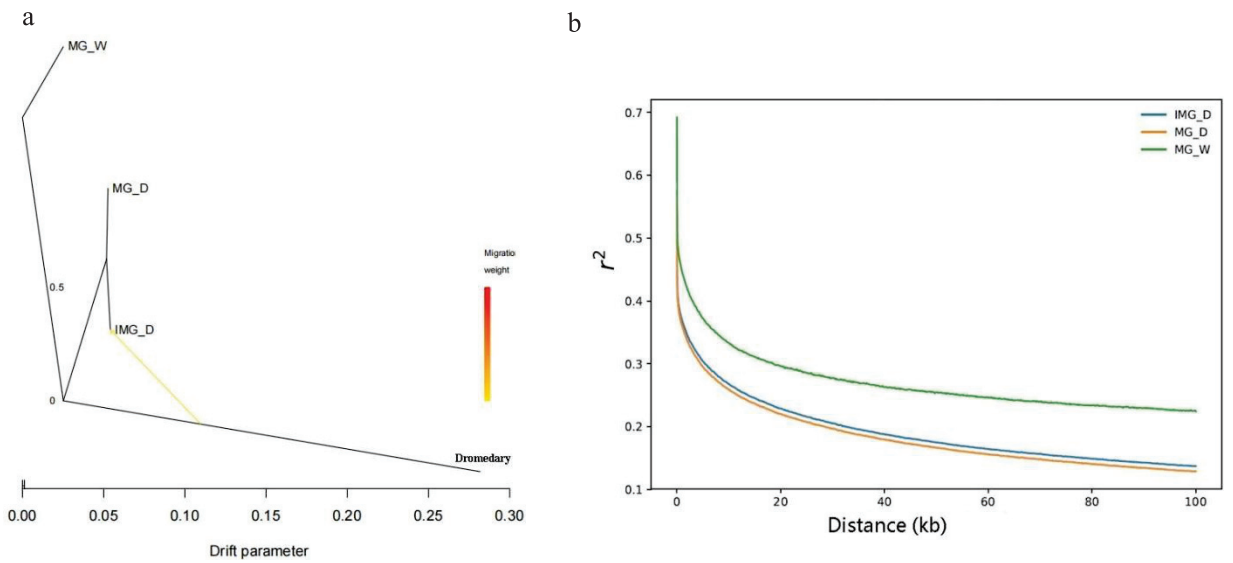

Fig 2. TreeMix analysis of migration events m (a) and decay of linkage disequilibrium (LD) in the Bactrian camel populations with one line per population (b).

selection in domestic Bactrians and wild two-humped camels during the breeding and survival process (Table 3).

Candidate genes related to insulin secretion and insulin signaling pathways

Bactrian camels exhibit insulin resistance, which maintains high blood sugar levels in their body. There are 16 functional candidate genes involved in the insulin secretion pathway and 9 in the insulin signaling pathway (Table 3, Fig 4). The changes in the ABCC8 and KCNJ11 genes disrupt the KATP channel's potentiality and regulate the secretion of insulin, thereby maintaining glucose homeostasis

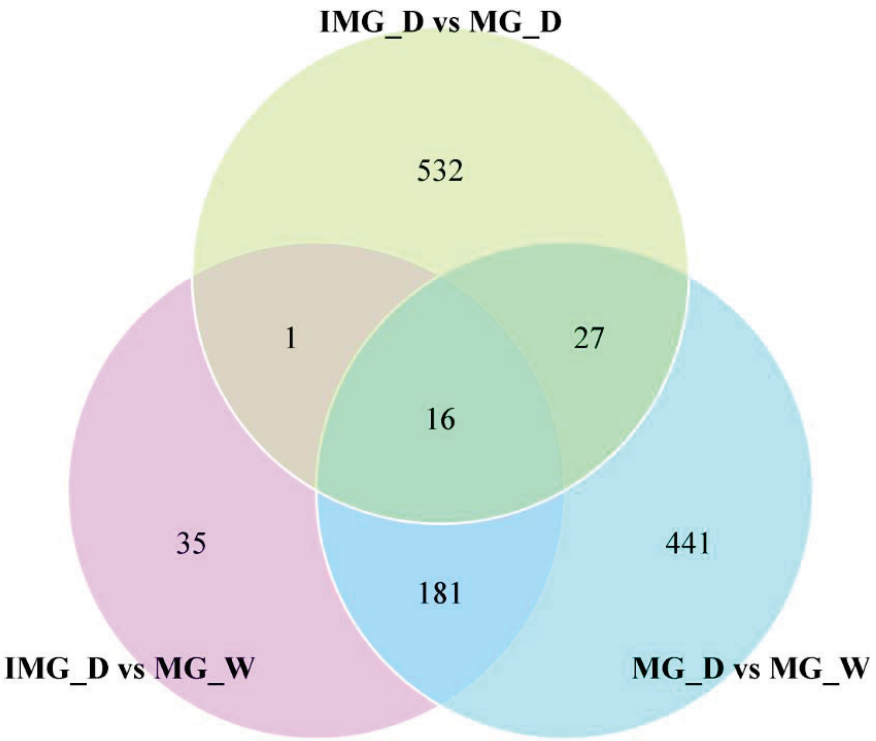

Fig 3. Venn diagram showing comparative analysis of candidate genes among different Bactrian camel populations.

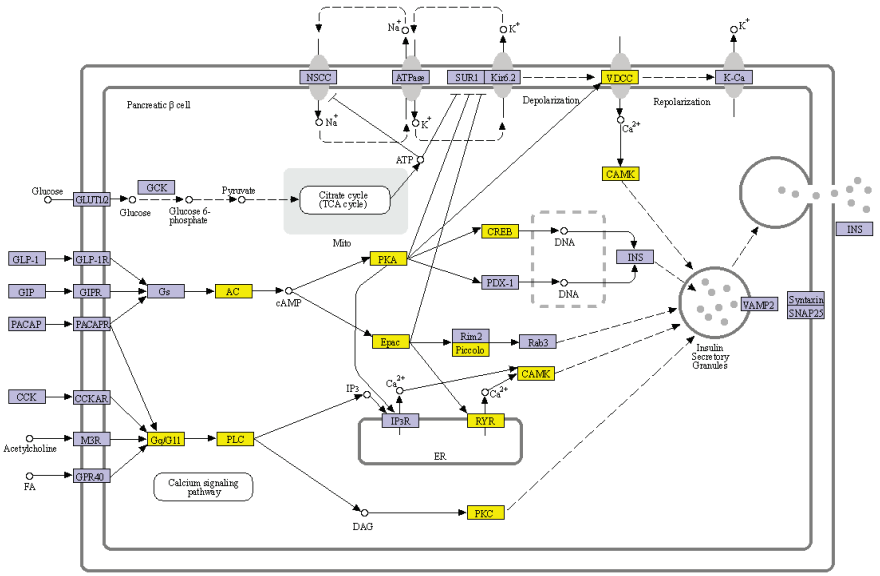

Fig 4. Insulin secretion pathway. Candidate genes are marked in yellow.

(Edghill *et al,* 2010; Reddy *et al,* 2021). The FFAR1 receptor is a long-chain fatty acid G-protein coupled receptor widely distributed in the pancreas and central nervous system. It can act on islet β cells to promote insulin secretion and activate islet α cells to secrete glucagon, regulating gastrointestinal endocrine cells and adjusting glycolipid levels (Walker *et al,* 2011). The CREB1 gene promotes insulin synthesis and secretion. The central gene PRKACB is involved in the insulin secretion pathway. The ACACA gene controls the secretion of insulin, and the GNAS gene is an important regulator of insulin secretion's capacity for pancreatic beta cells (Dalle *et al,* 2011; Taneera *et al,* 2019; Beale, 2013). Insulin resistance is also associated with SLC2A4 genes. During the Bactrian camels' evolution, these insulin-relevant pathways and the candidate genes under selection are similar to the previous study, explaining insulin resistance in Bactrian camels (Jirimutu *et al,* 2012).

Candidate genes related to environmental adaptation

The desert environment where Bactrian camels live is harsh, and a lack of food and water resources is common. In order to survive in this environment, Bactrian camels have developed a unique environmental adaptation mechanism. A few genes subject to selection were associated with circadian entrainment or rhythm—fly pathways (Fig 5), such as the PRKACB gene, which indirectly affects cell proliferation and differentiation; CACNA1H may regulate intracellular processes such as contraction, secretion, neurotransmission, and gene expression. ADCY1 is involved in the regulatory processes of the central nervous system (Chen *et al,* 2013; Santos-Cortez *et al,* 2014). In addition, limited protein-coding genes are involved in plant–pathogen interaction pathways, such as HSP90AA1 aids protein folding and quality control for many 'client' proteins (Zuehlke *et al,* 2015).

Discussion

Domestic animal characteristics result from a high-intensity artificial selection over a short period for wild ancestral species. In the fields of evolutionary

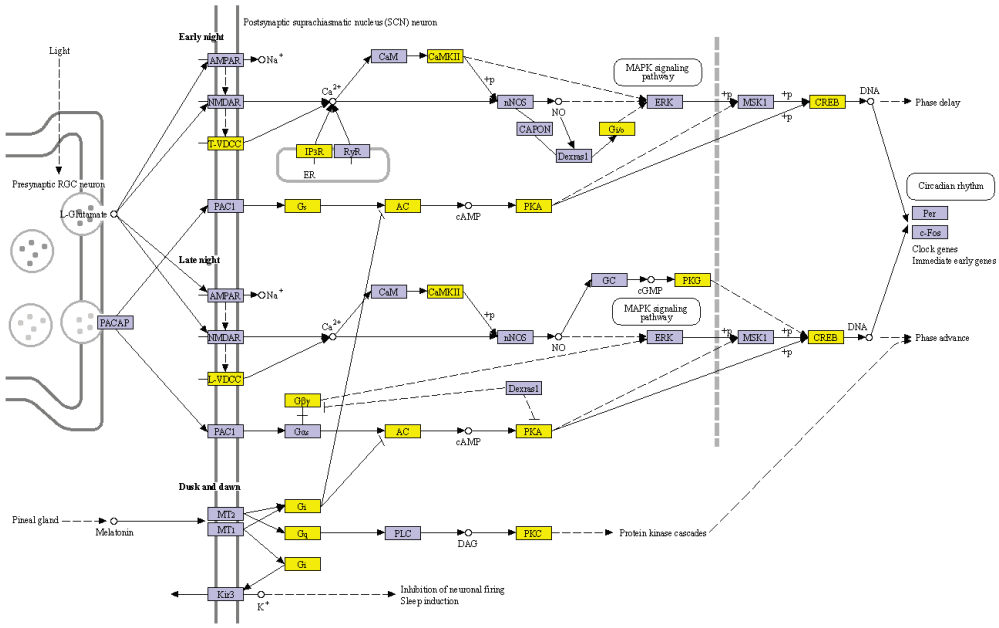

Fig 5. Circadian entrainment pathway. Candidate genes are marked in yellow.

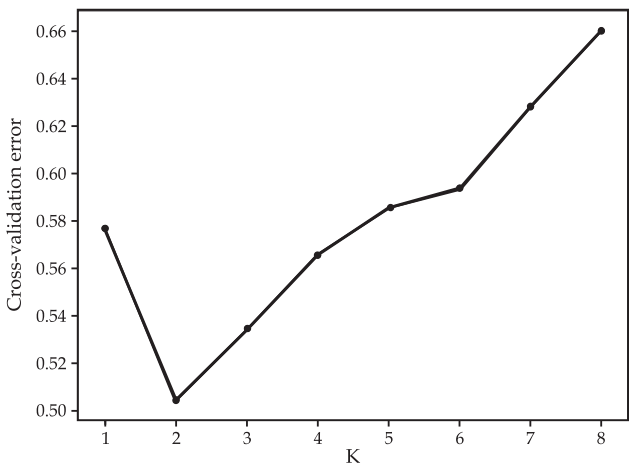

Fig S1. Cross-validation errors in the ADMIXTURE analysis. The number of ancestry K was assumed from 1 to 3 and K = 2 is the optimum number.

genetics and genomics, many theories and methods for testing natural selection at the gene and genome levels have been developed. Artificial selection is much stronger than natural selection; therefore, evolutionary genomic detection methods detect artificial selection signals more effectively (Jensen *et al,* 2007). In the present study, the whole genome variations of domestic Bactrians and wild two-humped camels were characterised. Genome-wide selection signatures were also performed between domestic and wild two-humped camels, providing vital genomic information under the influence of natural and artificial selection. From the genomic structure, due to their close geographic location and genetic relationship, gene flow exists between domestic Bactrian camel populations from China's Inner Mongolia and Mongolia (Fig 1a,b). However, gene flow signaling is not strong in the domestic and wild two-humped camel populations mainly due to their independent maternal origin (Ji *et al,* 2009) and the limited survival environment (Taklamakan Desert, Arjin Mountains in the Lop Nur Lake region, and the Great Gobi Strictly Protected Area 'A') (Yadamsuren *et al,* 2019) of wild two-humped camels, which may also lead to the relatively pure preservation of their genomes.

Bactrian camels are an important animal species in the Gobi Desert of China and Mongolia. Long-term evolution and natural selection have resulted in unique biological characteristics for adapting to harsh desert environments, including cold and hot resistance, anti-starvation, and a strong immune system. The analysis showed that a large number of positive candidate genes in Bactrian camels were involved in circadian entrainment pathways (ko 04713), plant–pathogen interactions (ko04626), and circadian rhythm—fly pathways (ko04711). These enrichment pathways involved in the environmental adaptation hierarchy may explain Bactrian camels' unique desert environment adaptability. It is worth mentioning that HSP90AA1, a member of the HSP protein (heat-shock proteins) family, completes positive selection in domestic Bactrian camels from IMG_D and MG_D, which relates to the temperature of living camels. The function of HSP is to protect cells from heat shock by resisting the denaturation of cellular proteins (Feder and Hofmann, 1999). In

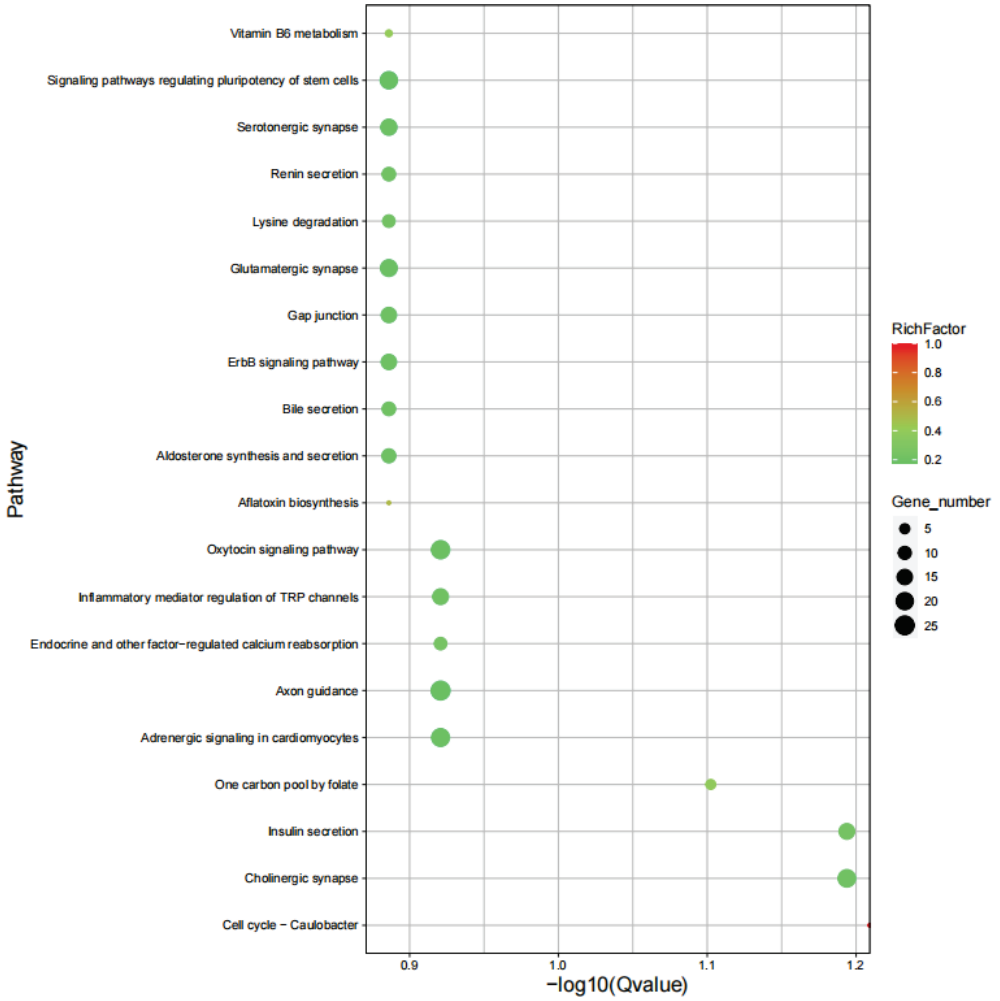

Fig S2. Significantly enrichment KEGG pathways canditate genes from domestic Bactrians and wild two-humped camels.

China's Inner Mongolia, the desert's temperature in summer is as high as 40 degrees, whereas in Mongolia, even in summer, the desert temperature is not as high. Therefore, this gene has been differentiated in Bactrian camel populations in Inner Mongolia and Mongolia.

Blood glucose levels in camels are higher than those of other animals (Al-Ali *et al,* 1988). In general, the blood glucose level of camels in a normal state (7.1 ± 0.3mmol/L) was higher than ruminants (2.5-3.5 mmol/L) and monogastric animals (3.5-5.0 mmol/L). The insulin content in camel blood (5 ± 1μU/mL) was lower than in sheep (12 ± 2μU/mL) and horses (7 ± 1μU/mL) (Elmahdi *et al,* 1997). Research has shown that the high level of blood glucose and low level of insulin in camels may be caused by their insulin resistance (Kaske *et al,* 2001; Guo *et al,* 2021). Insulin binding to receptors results in the tyrosine phosphorylation of insulin receptor substrates (IRS) by the insulin receptor tyrosine kinase (INSR). This process allows IRS association with the regulatory subunit of phosphoinositide 3-kinase (PI3K), which activates Akt. Insulin is mainly passed through the PI3K/Akt pathway to maintain the balance of glucose and lipid metabolism (Jensen *et al,* 2007; Taniguchi *et al,* 2006). In the present study, some candidate genes such as ABCC8, KCNJ11, FFAR1, PRKACB, CREB1, PRKACB, ACACA, SLC2A4, and AKT were enriched in the insulin signaling (ko04910) and secretion pathways (ko04911) between domestic Bactrian and wild two-humped camels, which may result in insulin responsiveness. According to reports, hibernating brown bears (Nelson *et al,* 2014) and reindeer (Elmahdi *et al,* 1997) that normally live during the snowy months or when food supplies are low also exhibit insulin resistance, which may benefit their survival in harsh conditions, similar to camels.

Bactrian camels have also well-developed, extraordinarily strong, and sensitive olfactory senses that can detect odours over distances up to 3 km

away. If they are downwind, the distance will be longer, at dozens of kilometers away. The KEGG pathway analysis of non-synonymous variations showed that more functional genes were enriched in the olfactory transduction (ko04740) pathway of domestic Bactrians (IMG_D and MG_D) than in wild two-humped camels (MG_W) (Fig 3). In wild two-humped camels, only 3 distribution areas have relatively simple vegetation; therefore, they do not need to distinguish many odours, and their olfactory receptors may be relatively small compared with domestic Bactrians. For these reasons, olfactory receptors are strengthened in wild two-humped camels and were an artificial selection during the domestication of domestic Bactrian camels, as confirmed by previous findings (Jirimutu *et al,* 2012).

Furthermore, several significant pathways associated with lipid metabolism and the immune system have been analysed. Candidate genes such as ECI1, ACADL, ACSBG2, DGK, and CDS1 were enriched in fatty acid degradation and glycerophospholipid metabolism. The protein encoded by the ECI1 gene is a key mitochondrial enzyme involved in the β-oxidation of unsaturated fatty acids. The ACADL gene is responsible for the beta-oxidation of fatty acids within the mitochondria. The ACSBG2 gene indirectly acts upstream or within the fatty acid metabolic process. Diacylglycerol kinases (DGKs) are regulators of the intracellular concentration of the second messenger diacylglycerol (DAG) and thus play a key role in cellular processes (Riese *et al,* 2016). The CDS1 gene encodes an enzyme that regulates the amount of phosphatidylinositol available for signaling by catalysing the conversion of phosphatidic acid to CDP-diacylglycerol (Huang and Freter, 2015). These genes may enhance a camel's energy storage and production capacity in the desert. In addition, seven significant immune systems related to KEGG pathways were identified, and 19 candidate genes were enriched in these pathways. These genes may be related to Bactrian camels' unique immune system, which allows them to adapt to the changing desert environment.

Conclusions

The present study provided comprehensive insights into the candidate regions for signatures of positive selection in the genome of domestic Bactrians from China and Mongolia and wild two-humped camels from Mongolia. Several candidate genes in have been identified 3 camel populations, which have essential roles in metabolism, insulin resistance, olfactory transduction, environmental adaptation, and other characteristics. These results provide evidence of selection in camels for adapting to the harsh arid conditions of desert environments and may provide some perspective on disease-resistance research.

Competing Interests

The authors declare that they have no competing interests.

References

Al-Ali A, Husayni H and Power D. A comprehensive biochemical analysis of the blood of the camel (*Camelus dromedarius*). Comparative Biochemistry and Physiology Part B: Comparative Biochemistry. 1988; 89:35-37.

Alexander D H, Novembre J and Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Research. 2009; 19:1655-1664.

Ali A, Baby B and Vijayan R. From desert to medicine: a review of camel genomics and therapeutic products. Frontiers in Genetics. 2019; 10:17-26.

Beale E G. Insulin signaling and insulin resistance. Journal of Investigative Medicine. 2013; 61:11–14.

Bigham A, *et al*. Identifying signatures of natural selection in Tibetan and Andean populations using dense genome scan data. PLoS Genetics. 2010; 6.

Burger P A, Ciani E and Faye B. Old World camels in a modern world - a balancing act between conservation and genetic improvement. Animal Genetics. 2019; 50:598-612.

Chen Y, Gao Y, Tian Y, *et al*. PRKACB is downregulated in non-small cell lung cancer and exogenous PRKACB inhibits proliferation and invasion of LTEP-A2 cells. Oncology Letters. 2013; 5:1803-1808.

Conesa A and Götz S. Blast2GO: a comprehensive suite for functional analysis in plant genomics. International Journal of Plant Genomics. 2008; 619832.

Dalle S, Quoyer J, Varin E, *et al*. Roles and regulation of the transcription factor CREB in pancreatic β-cells. Current Molecular Pharmacology. 2011; 4:187-195.

DePristo M A, *et al*. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature Genetics. 2011; 43:491-498.

Edghill E L, Flanagan S E and Ellard S. Permanent neonatal diabetes due to activating mutations in ABCC8 and KCNJ11. Reviews in Endocrine and Metabolic Disorders. 2010; 11:193-198.

Elmahdi B, Sallmann H P, Fuhrmann H, *et al*. Comparative aspects of glucose tolerance in camels, sheep, and ponies. Comparative Biochemistry and Physiology Part A: Physiology. 1997; 118:147-151.

Feder M E and Hofmann G E. Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. Annual Review of Physiology 1999; 61:243-282.

Fitak R R, Mohandesan E, Corander J, *et al*. The de novo genome assembly and annotation of a female domestic dromedary of North African origin. Molecular Ecology Resources. 2016; 16:314-324.

Guo F *et al*. Reversible insulin resistance helps Bactrian camels survive fasting. Scientific Reports. 2021; 11:1-12.

Hayes B J, Bowman P J, Chamberlain A J, *et al*. Invited review: Genomic selection in dairy cattle: Progress and challenges. Journal of Dairy Science. 2009; 92:433-443.

Huang C and Freter C. Lipid metabolism, apoptosis and cancer therapy. International Journal of Molecular Science. 2015; 16:924-949.

Jensen J D, Wong A and Aquadro C F. Approaches for identifying targets of positive selection. TRENDS in Genetics. 2007; 23:568–577.

J R, *et al*. Monophyletic origin of domestic bactrian camel (*Camelus bactrianus*) and its evolutionary relationship with the extant wild camel (*Camelus bactrianus* ferus). Animal Genetics. 2009; 40:377-382.

Jirimutu, *et al*. Genome sequences of wild and domestic bactrian camels. Nature Communications. 2012; 3:1202.

Kaske M, Elmahdi B, Engelhardt W, *et al*. Insulin responsiveness of sheep, ponies, miniature pigs and camels: results of hyperinsulinemic clamps using porcine insulin. Journal of Comparative Physiology B. 2001; 171:549-556.

Khalkhali-Evrigh R, Hafezian S H, Hedayat-Evrigh N, *et al*. Genetic variants analysis of three dromedary camels using whole genome sequencing data. PloS One. 2018; 13.

Li H and Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25:1754-1760.

Li M, *et al*. Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars. Nature Genetics. 2013; 45:1431-1438.

Lv F H, *et al*. Adaptations to climate-mediated selective pressures in sheep. Molecular Biology and Evolution. 2014; 31:3324-3343.

Ming L, *et al*. Genetic diversity and phylogeographic structure of Bactrian camels shown by mitochondrial sequence variations. Animal Genetics. 2017; 48:217-220.

Ming L, *et al*. Chromosome-level assembly of wild Bactrian camel genome reveals organization of immune gene loci. Molecular Ecology Resources. 2020; 20:770-780.

Ming L, *et al*. Whole-genome sequencing of 128 camels across Asia reveals origin and migration of domestic Bactrian camels. Communications Biology. 2020; 3:1.

Ming L, *et al*. Mitochondrial DNA variation and phylogeography of Old World camels. Animal Bioscience. 2021; 34:525.

Ming L, Siren D, Hasi S, *et al*. Review of genetic diversity in Bactrian camel (*Camelus bactrianus*). Animal Frontiers. 2022; 12:20-29.

Nelson O L, *et al*. Grizzly bears exhibit augmented insulin sensitivity while obese prior to a reversible insulin resistance during hibernation. Cell Metabolism. 2014; 20:376-382.

Nielsen R. Molecular signatures of natural selection. Annual Review of Genetics. 2005; 39:197-218.

Pickrell J and Pritchard J. Inference of population splits and mixtures from genome-wide allele frequency data. Nature Precedings. 2012; 8:1002967.

Reddy S, *et al*. Association of ABCC8 and KCNJ11 gene variants with type 1 diabetes in south Indians. Egyptian Journal of Medical Human Genetics. 2021; 22:1–11.

Riese M J, *et al*. Diacylglycerol kinases (DGKs): novel targets for improving T cell activity in cancer. Frontiers in Cell and Development Biology. 2016; 4:108.

Saipolda T. Mongolian camels. In: Cardellino R, Rosati A, Mosconi C, editors. Proceedings of the current status of genetic resources, recording and production systems in African, Asian and American camelids; Sousse, Tunisia. Rome, Italy: ICAR. 2004; pp 73-79.

Santos-Cortez R L P, *et al*. Adenylate cyclase 1 (ADCY1) mutations cause recessive hearing impairment in humans and defects in hair cell function and hearing in zebrafish. Human Molecular Genetics. 2014; 23:3289-3298.

Sikkema R S, *et al*. Global status of Middle East respiratory syndrome coronavirus in dromedary camels: a systematic review. Epidemiology and Infection. 2019; 147:84.

Taneera J, *et al*. GNAS gene is an important regulator of insulin secretory capacity in pancreatic β-cells. Gene. 2019; 715:144028.

Taniguchi C M, Emanuelli B and Kahn C R. Critical nodes in signalling pathways: insights into insulin action. Natture Revieve Molecular Cell Biology. 2006; 7:85-96.

Vilella A J, *et al*. Ensembl Compara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. Genome Research. 2009; 19:327-335.

Walker C G, *et al*. Variation in the FFAR1 gene modifies BMI, body composition and beta-cell function in overweight subjects: an exploratory analysis. PloS One. 2011; 6:19146.

Wang K, Li M and Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research. 2010; 38:164.

Yadamsuren A, Daria O and Liu S. The seasonal distribution of wild camels (*Camelus ferus*) in relation to changes of the environmental conditions in Mongolia. Open Journal of Ecology. 2019; 9:293–314.

Yi L, *et al*. Molecular diversity and phylogenetic analysis of domestic and wild Bactrian camel populations based on the mitochondrial ATP8 and ATP6 genes. Livestock Science. 2017; 199:95-100.

Zhang C, Dong S S, Xu J Y, *et al*. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. Bioinformatics. 2019; 35:1786-1788.

Zuehlke A D, Beebe K, Neckers L, *et al*. Regulation and function of the human HSP90AA1 gene. Gene. 2015; 570:8-16.