Eurasia Specialized Veterinary Publication

**International Journal of Veterinary Research and Allied Science**

**ISSN:3062-357X**

2022, Volume 2, Issue 2, Page No: 125-139

Available online at: www.esvpub.com/

# Genomic Insights into the Geographical Structuring, Genetic Diversity, and Environmental Adaptations of Saudi Arabian Dromedary Camel (Camelus dromedarius) Populations

**Sarah Wilhelm[1]*, Leon Krüger[1]**

[1]Department of Pathobiology, Faculty of Veterinary Medicine, University of Vienna, Vienna, Austria.

***E-mail** ✉ s.wilhelm.vet@gmail.com

**ABSTRACT**

Dromedary camels (Camelus dromedarius) in Saudi Arabia show remarkable genetic variability, reflecting their adaptation to a range of habitats, from arid deserts to mountainous regions and coastal zones. In this study, we examined the population genetics of these camels, focusing on geographic influences rather than strictly ecological classifications. Analysis of whole-genome sequences from 63 individuals revealed pronounced differences in heterozygosity and inbreeding levels among populations, with mountainous camels displaying the highest diversity and coastal populations the lowest. Furthermore, we identified significant enrichment of genes linked to environmental resilience, including the HECT domain in desert-adapted camels, which contributes to protein stability under harsh conditions. Multivariate and admixture analyses highlighted the genetic uniqueness of certain breeds, particularly the Awarik (beach ecotype), which appears to have experienced prolonged genetic isolation.

**Keywords:** Dromedary camels, SNPs, genetic variation, Majaheem, Awarik,, Environmental adaptation

**Received:** 29 May 2022
**Revised:** 27 September 2022
**Accepted:** 01 October 2022

**How to Cite This Article:** Wilhelm S, Krüger L. Genomic Insights into the Geographical Structuring, Genetic Diversity, and Environmental Adaptations of Saudi Arabian Dromedary Camel (Camelus dromedarius) Populations. Int J Vet Res Allied Sci. 2022;2(2):125-39. https://doi.org/10.51847/2kdhV8WWBx

## Introduction

The Arabian Peninsula is home to the dromedary camel (Camelus dromedarius), a species that has successfully adapted to a range of habitats, including deserts, mountains, and coastal areas, resulting in diverse camel populations [1]. Saudi Arabia, with its varied ecological landscapes, offers an ideal setting for investigating the genetic diversity and adaptive mechanisms of these camels [2]. Previous genetic studies indicate that geographic regions better reflect the genetic structure of Saudi Arabian camels than ecological classifications, suggesting a complex interplay between genetic variation and environmental adaptation [2]. Consequently, camel populations from the northern, central, and western regions exhibit genetic distinctions from those in the southwest and southeast [3].

Whole-genome analyses of dromedaries across the Arabian Peninsula have revealed an overall homogeneous gene pool, with only minor geographic differentiation [4]. This genetic similarity is largely due to historical trade routes and movement of camels, which facilitated interbreeding between distant populations. Furthermore, the absence of structured breeding programs has led to unmanaged mating, further homogenizing genetic traits. These

observations align with Al Abri and Faye [5], who emphasized that historical and ongoing camel movements for trade and other purposes have significantly shaped their genetic landscape. Despite widespread genetic mixing, some geographic-associated structuring persists, leading to three primary population groups: northern, central, and western, and southwestern and southeastern camels [6].

Camels have played a central role in the socio-economic development of human societies throughout the Arabian Peninsula and beyond [7]. Domesticated approximately 3,000–4,000 years ago in southern Arabia, dromedaries facilitated transportation and trade across arid regions, enhancing cultural and economic interactions among distant communities [8]. Their provision of milk, meat, and wool, combined with remarkable adaptability to harsh desert conditions, has made camels invaluable to nomadic and semi-nomadic populations [9]. Additionally, they hold enduring cultural significance, featuring prominently in traditional festivals and competitions across the Arabian Peninsula [10].

Globally, the dromedary camel population is estimated at around 35 million, predominantly in arid and semi-arid regions of Africa, Asia, Australia, and the Arabian Peninsula [11]. Saudi Arabia alone hosts approximately 1.6 million camels, reflecting its longstanding traditions of camel husbandry and breeding [2]. The country recognizes fourteen distinct camel breeds, each adapted to specific environmental conditions. Desert breeds such as Magaheem, Wodeh, Sofor, and Shual inhabit the northern and central regions [12], while coastal breeds like Sahlia and Awarik are found in western and southwestern coastal zones, and mountain breeds such as Hadana and Awadi occupy western and southwestern highlands [13].

The primary objective of this study is to investigate the genetic relationships among Saudi camel populations, with particular emphasis on groupings defined by geographic and environmental factors, including desert, mountainous, and coastal habitats. Understanding the genetic diversity and population structure of these camels is essential for their conservation and improvement. Insights gained from such studies can inform breeding strategies aimed at maintaining resilience and sustainability in the face of environmental and climatic challenges [14].

**Sample Collection**

Camels included in this study were collected from three distinct ecological zones: coastal, desert, and mountainous regions **(Table 1, Figure 1)**. Animals were classified according to their breed and documented origin, based on historical records and official listings provided by the Ministry of Agriculture and Environment. The coastal cohort comprised the Sahliah and Awarik breeds, primarily distributed along the Jeddah and southern Jazan coastlines. The desert group included the Sofor, Shul, and Majaheem breeds, which are adapted to the extreme arid conditions of the Najad and Riyadh regions. The mountainous group consisted of the Awadi and Haddana breeds, native to the rugged landscapes of the Hijaz Mountains in southwestern Saudi Arabia. Each breed represented a proportional subset of the subpopulation within its ecological zone. Sampling was conducted across multiple farms, with individual camels selected based on farmer-provided pedigree information to minimize the inclusion of closely related animals.

**Table 1.** Sampling distribution and characteristics across different habitat groups

| Group | Breed | Sample count | Sample location | Sex (M/F) | Age range |
|---|---|---|---|---|---|
| Beach | Sahliah | 12 | Coastal regions near Jeddah and South Jazan | 8 M / 4F | Juvenile to Adult |
| | Awarik | 9 | Coastal regions near Jeddah and South Jazan | 5 M / 4F | Juvenile to Adult |
| Deserts | Sofor | 8 | Najad and its eastern areas | 4 M / 4F | Juvenile to Subadult |
| | Shul | 10 | Najad and its eastern areas | 5 M / 5F | Subadult to Adult |
| | Majaheem | 11 | Najad and its eastern areas | 6 M / 5F | Subadult to Adult |
| Mountain | Awadi | 4 | Hijaz mountains and southwestern KSA | 2 M / 2F | Juvenile to Subadult |
| | Haddana | 9 | Hijaz mountains and southwestern KSA | 5 M / 4F | Juvenile to Adult |

The table summarizes the distribution of sampled camel breeds across three ecological groups (beach, deserts, and mountain), detailing sample count, geographic location, sex distribution (M/F), and age range.

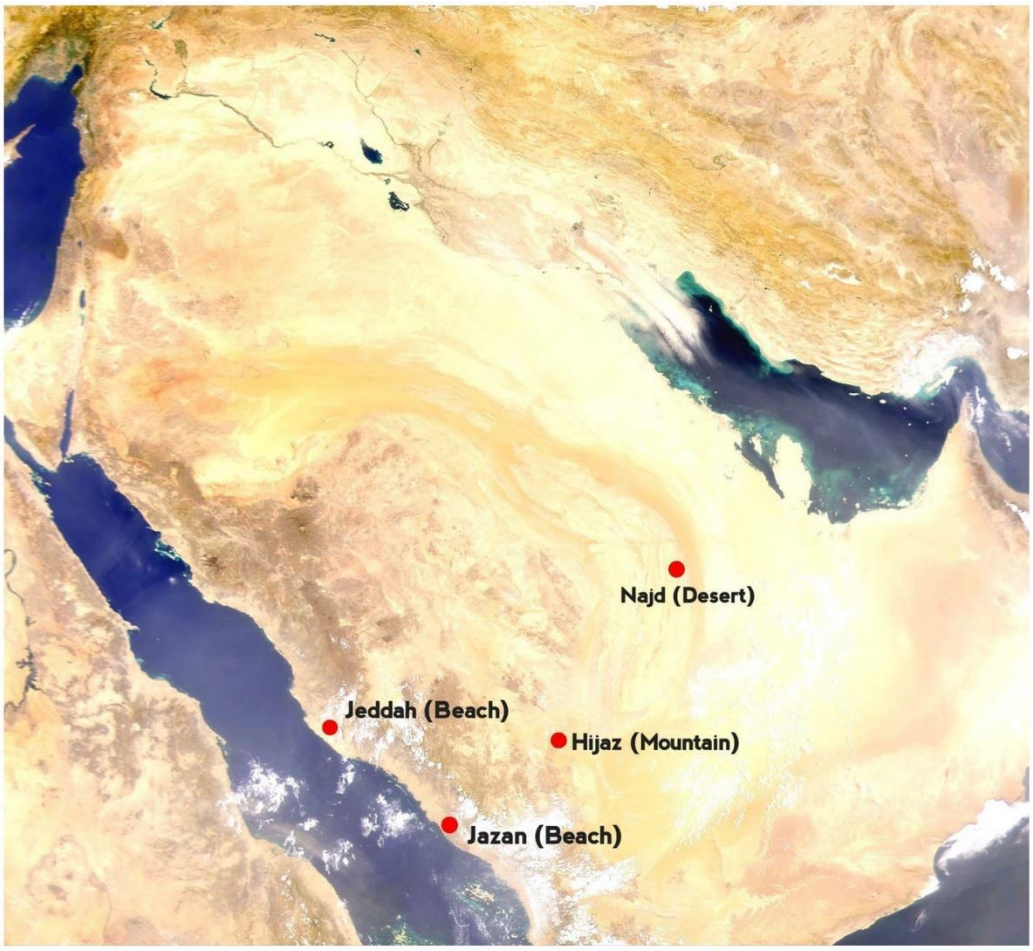


**Figure 1.** Geographic distribution of sampled camels in Saudi Arabia. Red markers indicate collection sites, with key regions labeled as Jeddah (Sahliah), Jazan (Awarik), Hijaz Mountains (Awadi, Haddana), and Najd (Majaheem, Shul, and Sofor)

*DNA extraction and sequencing*

Blood samples were collected from each camel, and genomic DNA was isolated using the Maxwell RSC system (Promega, USA) following the protocol of Wu *et al.* [15]. DNA concentration was measured with the QuantiFluor dsDNA system on a Quantus Fluorometer, and the integrity and purity of DNA were confirmed via agarose gel electrophoresis. Animals were selected from multiple farms based on pedigree and farmer-provided information to avoid sampling closely related individuals. Whole-genome sequencing was carried out by the Beijing Genomics Institute (BGI). Genomic DNA was fragmented to approximately 300 bp using Covaris shearing, end-repaired, adenylated at the 3′ ends, and ligated to sequencing adapters. Libraries were amplified into DNA nanoballs (DNBs) using rolling circle amplification and loaded onto patterned nanoarrays for sequencing on the Illumina NovaSeq 6000 platform. Sequencing generated paired-end reads with an average coverage of 30× per genome. Quality metrics were assessed with FastQC v0.12.0 [16].

*Data processing and alignment*

The workflow for sequence analysis began with quality inspection using FastQC to detect low-quality reads and adapter contamination. Reads were cleaned using SOAPnuke [17], discarding any read with more than 50% adapter content, over 50% of bases with a Phred score below 20, or at least 2% ambiguous nucleotides (“N”). Cleaned reads were aligned to the CamDro3 reference genome (NCBI Assembly) using BWA v0.7.18 [18], with indexing performed to optimize mapping efficiency. Alignment outputs were initially stored in SAM format, then converted to coordinate-sorted BAM files with SAMtools [19, 20]. Additional processing steps included marking

secondary alignments (-M), enabling YML output (-Y), and specifying read group information (-R) to ensure compatibility with subsequent analyses.

*Variant detection and filtering*

Genetic variants were identified using GATK HaplotypeCaller v4.5.0.0 [21], applying parameters recommended by the GATK Best Practices workflow [22]. Only variants with a base quality score of at least 20, a mapping quality of 20 or higher, and up to six alternative alleles were considered. To ensure reliability, low-confidence variants were removed based on stringent criteria: quality by depth (QD) below 2.0, total quality score (QUAL) under 40.0, Fisher Strand bias (FS) above 60.0, mapping quality (MQ) less than 40.0, mapping quality rank sum (MQRankSum) below −12.5, and read position rank sum (ReadPosRankSum) below −8.0. These filters help eliminate sequencing artifacts and retain only high-confidence SNPs [23]. Individual GVCF files were combined into a single dataset using GATK CombineGVCFs, and the consolidated file was converted to standard VCF format with GenotypeGVCFs to enable further analyses.

*SNP quality control and pruning*

Further refinement of SNPs was conducted using PLINK v1.9 [24, 25]. Variants with excessive missing data (>5%) were discarded (--geno), and alleles with a minor allele frequency (MAF) below 0.05 were removed (--maf) to focus on informative, common SNPs. Linkage disequilibrium (LD) pruning was performed with --indep-pairwise, using a 2000 kb window, one-SNP step, and $r^2$ threshold of 0.5. These settings reflect the extended LD observed in camel genomes, similar to patterns seen in other livestock species such as cattle [26]. Only autosomal markers were retained (--chr), and closely related individuals were filtered out using --rel-cutoff, following established relatedness estimation methods [25, 27].

*Assessment of genetic diversity*

Observed heterozygosity (Ho) for each camel was calculated using PLINK (--het), comparing the number of observed homozygous genotypes to the expected value. This measure provides insight into within-individual genetic diversity and allows estimation of inbreeding coefficients (F) [24, 25].

*Analysis of runs of homozygosity (ROH)*

ROHs, which indicate segments of the genome where both alleles are identical, were identified with the --homozyg command in PLINK. Detection relied on criteria for minimum SNP count, marker density, and permissible gaps. Identified ROHs were classified as small (<0.1 Mbp), medium (0.1–5 Mbp), or large (>5 Mbp), consistent with thresholds commonly used in livestock studies (28, 29). These classifications provide information on historical and recent inbreeding patterns in camels, taking into account their genome size and extended LD structure [26].

*Effective population size estimation*

The effective population size (Ne) for each camel population was inferred using SNeP v1.11 [30]. Analyses incorporated corrections for unphased genotype data, adjustments for mutation rates, and the Sved and Feldman [31] mutation modifier to improve the accuracy of Ne estimates in populations with complex breeding histories.

*Assessment of population structure*

Genetic relationships among populations were explored using a dual approach. Principal Component Analysis (PCA) in PLINK v1.9 provided a visualization of the primary axes of genetic variation, allowing the identification of clusters and patterns of differentiation among populations. In parallel, ADMIXTURE v1.3.0 [32] was employed to infer the number of ancestral gene pools (K) and estimate the proportional ancestry of each individual. Cross-validation was applied across K values from 2 to 5 to determine the most reliable number of ancestral populations based on predictive accuracy.

*Identification of selection signatures*

Regions of the genome under positive selection were detected using the XP-nSL statistic in Selscan v2.0.2 [33, 34]. Pairwise comparisons of phased data were conducted between populations, with parameters adapted to the camel genome: a scaling factor of 20,000 for normalization of EHH decay curves, a maximum inter-marker gap of 200,000 bp to account for extended LD, and an EHH threshold of 0.05 to reduce false positives. Standardized

XP-nSL scores were used to identify SNPs exceeding the significance threshold (two-sided $\alpha = 0.05$), and neighboring regions were expanded by half the average SNP spacing to define extended selection windows accurately.

*Annotation and functional enrichment*

Candidate regions were annotated using a reference GFF file to map overlapping genes. Populations were organized into three ecological categories, and genes were selected based on consistency and uniqueness: only genes showing selection signals across all populations within an ecological group were retained, while genes exclusive to a single ecological group were highlighted as potentially involved in environment-specific adaptation. Functional enrichment analyses, including Gene Ontology (GO) terms and pathway identification, were performed with DAVID [35] to elucidate the biological processes associated with adaptive loci. This approach ensured that the detected selection signals corresponded to biologically meaningful and ecologically relevant genes.

**Results**

*Sequencing output and data processing*

Whole-genome sequencing of 63 dromedary camels generated roughly 385 million raw reads, corresponding to approximately 57 billion base pairs. After applying stringent quality filters—removing reads with over half of their bases below a Phred score of 20—the dataset retained 373 million reads, representing around 56 billion bases. This filtering resulted in only a 2.95% reduction in data, indicating minimal loss of information while maintaining high quality for downstream analyses. The filtered reads showed an average G.C. content of 42.4% and a Q30 rate of 97.07%, reflecting strong sequencing accuracy .

Reads were aligned to the CamDro3 reference genome using BWA v0.7.18. Alignment performance was excellent, with a mean mapping rate of 99.81% and 99.69% of reads properly paired. Sequencing depth averaged 21.5×, covering more than five reads for 92.87% of the genome and at least one read for 93.74%, ensuring robust genome-wide representation.

*Variant identification and filtering*

Initial variant calling detected approximately 2.96 million SNPs, dominated by homozygous variants (2.87 million; 97.04%) and a smaller fraction of heterozygous variants (88,927; 2.96%). The transition-to-transversion ratio was 1.70, indicating typical patterns of nucleotide substitution within these populations.

Starting with 63 samples and nearly 13 million variants, quality control reduced the dataset in several stages. Filtering for missing genotype data lowered the count to 4.5 million variants while retaining all 63 samples, achieving a genotyping rate of 0.63. Applying a minor allele frequency (MAF) cutoff removed rare alleles, leaving 4,077,109 variants. Linkage disequilibrium pruning further reduced the dataset to 392,139 loci, and restricting analysis to autosomes yielded 367,871 SNPs. Finally, after excluding closely related individuals, the dataset comprised 34 camels and 367,871 high-quality SNPs, providing a reliable foundation for downstream population genetic analyses **(Figure 2)**.

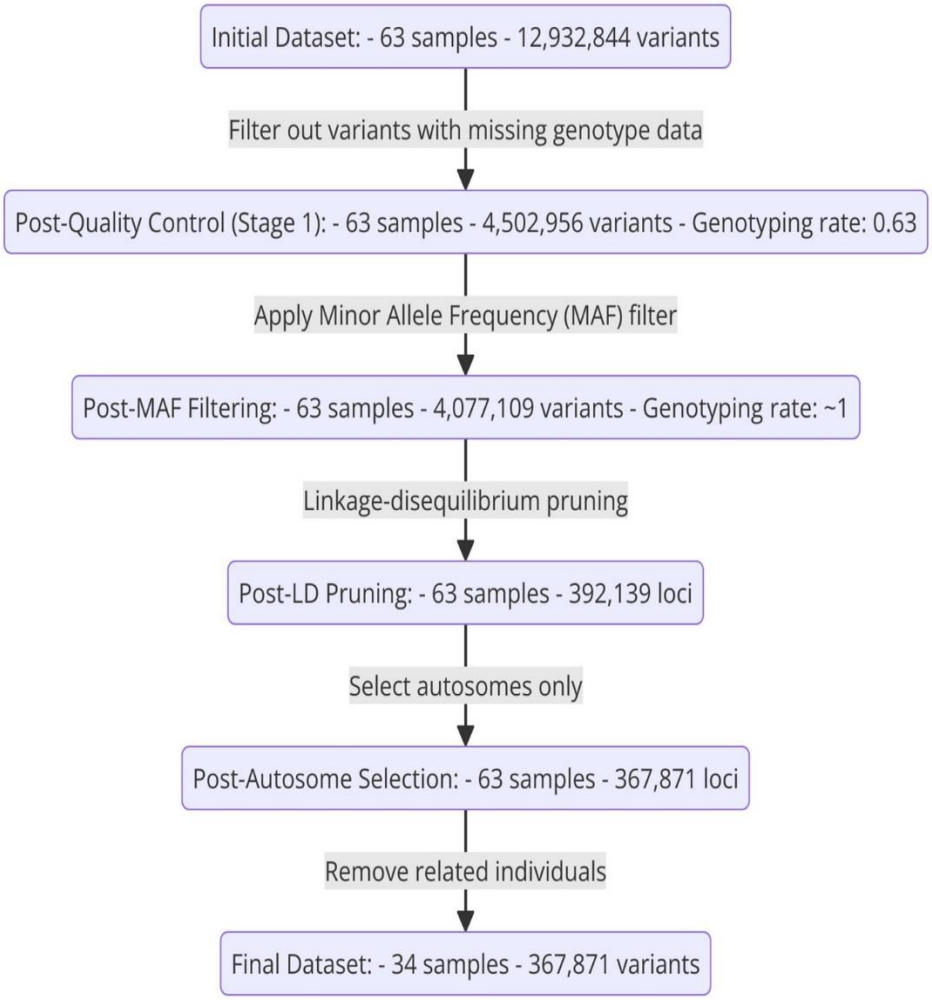


**Figure 2**. A diagram illustrating the sequential steps in the SNP filtering process, from the initial to the final refined dataset

*Genetic variation across populations*

Analysis of observed heterozygosity (Ho) revealed notable differences in genetic diversity among the six camel populations. The HAD population exhibited the highest level of heterozygosity (Ho = 0.414), suggesting a relatively rich genetic makeup. In contrast, the AWA population showed the lowest Ho at 0.374, indicating a more limited genetic variation. Intermediate Ho values were observed in the MAJ, SAH, SHU, and SOF populations, ranging from 0.396 to 0.405, reflecting moderate diversity within these groups **(Table 2)**. These findings highlight how genetic variability differs between populations, likely influenced by ecological habitat and breeding practices.

**Table 2.** Overview of the six studied camel populations, including their abbreviations, ecological classification (beach, mountain, desert), sample size (N), observed heterozygosity (Ho ± SD), inbreeding coefficient based on runs of homozygosity (FROH ± SD), and estimated effective population size (Ne)

| Population | Acronym | Ecotype | *N* | $H_o$ (S.D.) | $F_{ROH}$ (S.D.) | $N_e$ |
|---|---|---|---|---|---|---|
| Awarik | AWA | Beach | 5 | 0.374 (0.03) | 0.015 (0.01) | 15 |
| Haddana | HAD | Mountain | 4 | 0.414 (0.00) | 0.003 (0.00) | 11 |
| Majaheem | MAJ | Deserts | 9 | 0.400 (0.01) | 0.002 (0.00) | 37 |
| Sahliah | SAH | Beach | 7 | 0.405 (0.01) | 0.000 (0.00) | 24 |
| Shul | SHU | Deserts | 4 | 0.396 (0.01) | 0.004 (0.01) | 17 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Sofor | SOF | Deserts | 5 | 0.401 (0.01) | 0.000 (0.00) | 23 |

*Population size, inbreeding, and genetic differentiation*

Effective population size (Ne) showed substantial variation among the camel groups. The MAJ population had the largest Ne [36], indicating a relatively high number of breeding individuals, while the HAD population exhibited the smallest Ne [11], suggesting limited genetic contributors. Intermediate values were observed in the AWA [15], SHU [17], SAH [24], and SOF [23] populations, reflecting differing degrees of genetic representation across ecological zones.

Inbreeding levels, measured as the proportion of the genome in runs of homozygosity (FROH), were generally low, pointing to minimal recent inbreeding. The AWA population displayed the highest FROH (0.015), whereas SAH and SOF populations showed no detectable homozygosity (0.000). Other populations fell within the 0.002–0.004 range, consistent with a history of outcrossing and gene flow.

Principal Component Analysis (PCA) revealed clear patterns of genetic structuring. PC1 and PC2 accounted for 7.7% and 5.6% of the total variance, respectively. Desert populations clustered tightly along PC1, forming a linear alignment indicative of shared ancestry. Mountain populations grouped closely along PC2, reflecting their genetic cohesion. Beach populations were divided into two distinct clusters: one aligned with the mountain populations, suggesting admixture, and the other separated along the negative axis of PC1, demonstrating pronounced differentiation within this ecotype. These results illustrate how geography and ecological adaptation shape the genetic landscape of Saudi Arabian camels **(Figure 3)**.

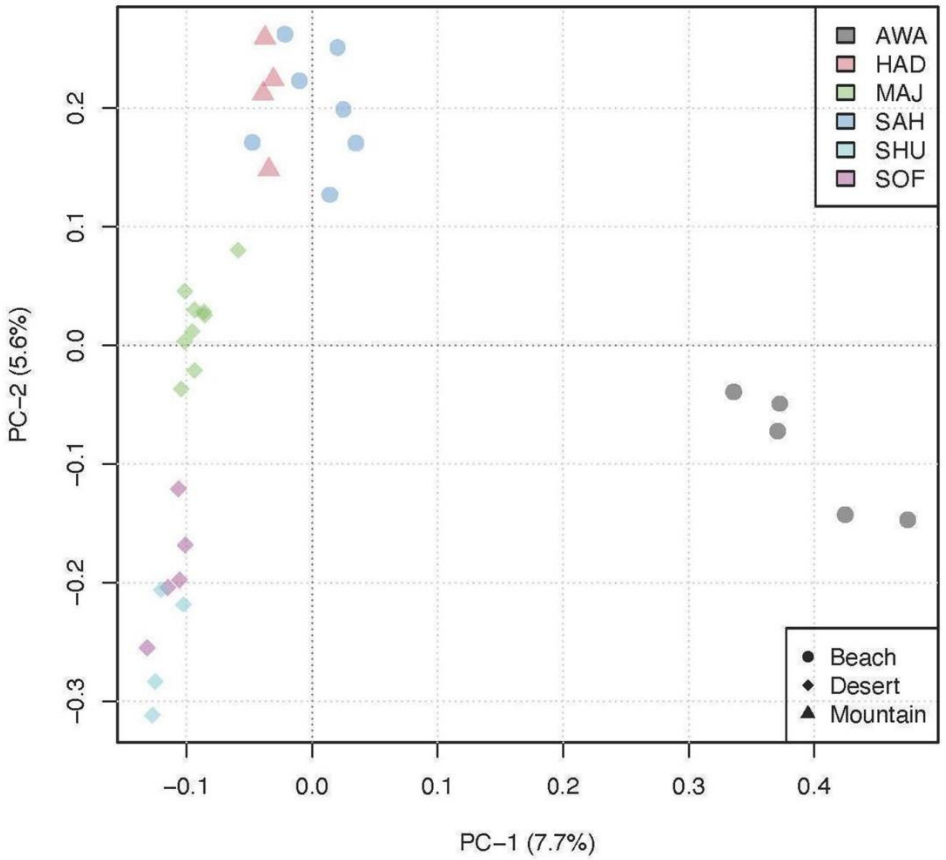


**Figure 3.** Principal Component Analysis (PCA) depicting the genetic relationships among six camel breeds across three ecological environments

Admixture analysis, conducted for K values ranging from 2 to 5, revealed the underlying population structure **(Figure 4)**. The AWA population consistently formed a distinct genetic cluster, contributing minimally (<10%) to the ancestral components observed in SAH and HAD, and showing almost no overlap with MAJ, SHU, or SOF. At K = 3, SHU exhibited a unique cluster that was partially shared with SOF and, to a lesser extent, with MAJ. Increasing K to 4 revealed a fourth ancestral component, further differentiating HAD and MAJ. At higher K

values, the population structure became increasingly intricate, with multiple ancestral components distributed across all populations, highlighting admixture and complex genetic relationships among the camel breeds.

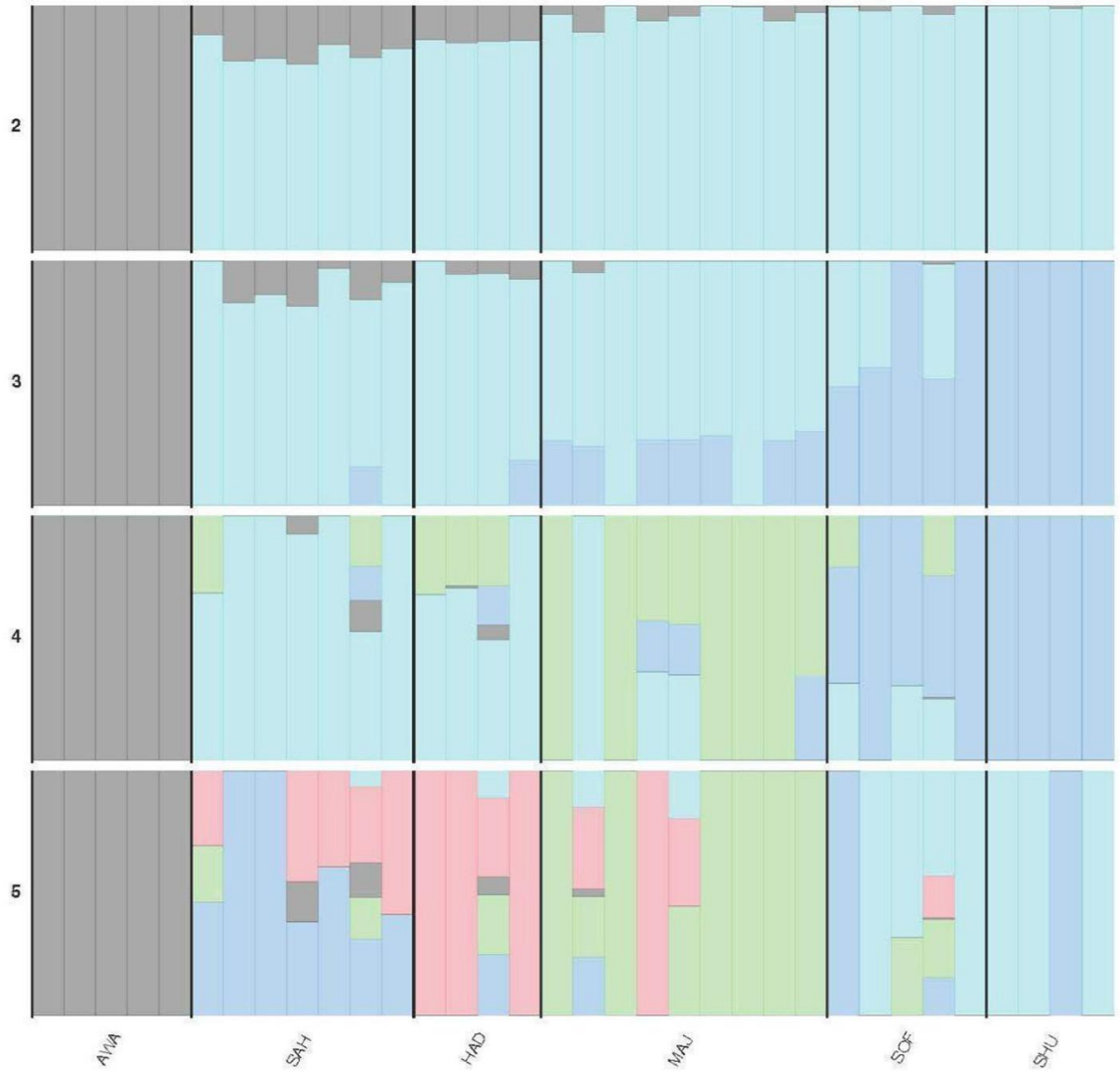


**Figure 4**. Admixture plot showing the genetic composition of six camel populations across K-values (K = 2 to 5). See Table 2 for population acronyms

Examination of runs of homozygosity (ROH) highlighted notable differences in genomic homozygosity among camel ecotypes. ROHs were classified into three size categories: small (1–2 Mb), medium (2–5 Mb), and large (>5 Mb) to reflect varying levels of historical inbreeding and genetic drift **(Table 3; Figure 5)**.

Camels from coastal regions exhibited three extensive homozygous stretches totaling 15.76 Mb. Additionally, they possessed 113 intermediate-length ROHs summing to 295.96 Mb, and 468 short ROHs with a combined length of 646.34 Mb. Desert-adapted camels showed a reduced number and size of homozygous segments, with two long ROHs (12.94 Mb), 52 medium ROHs (134.97 Mb), and 271 small ROHs (371.00 Mb). Mountain populations displayed a pattern comparable to coastal camels, including three long ROHs (15.76 Mb), 115 medium ROHs totaling 299.59 Mb, and 431 short ROHs covering 595.75 Mb.

These observations suggest that homozygosity patterns are shaped by ecological adaptation, historical population size, and breeding practices, with desert camels showing lower overall homozygosity, likely due to greater gene flow and larger effective population sizes, whereas coastal and mountain populations retain longer stretches of homozygosity indicative of more restricted mating pools.

**Table 3.** Comparison of runs of homozygosity (ROH) categories across camel ecotypes

| ROH category | Beach (Count) | Deserts (Count) | Mountain (Count) |
|---|---|---|---|

| | | | |
|---|---|---|---|
| Large | 3 | 2 | 3 |
| Medium | 113 | 52 | 115 |
| Small | 468 | 271 | 431 |

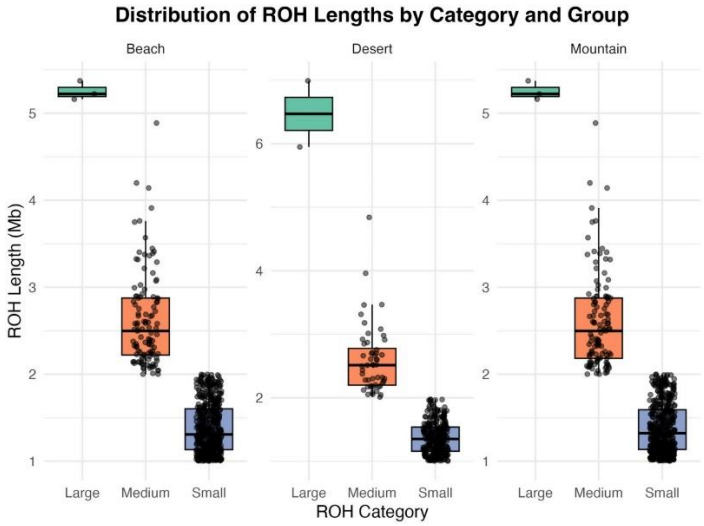


**Figure 5.** Box plot showing the distribution of ROH segment lengths across three categories—small, medium, and large—among camel ecotypes: beach, desert, and mountain

Analysis of selection signatures using the XP-nSL statistic, which identifies regions of extended haplotype homozygosity from phased genomic data, uncovered distinct patterns of adaptive variation across the camel populations. Pairwise comparisons among ecotypes revealed that the beach population contained 105 genes under unique selection pressure, the desert population exhibited 30 uniquely selected genes, and the mountain population displayed the most pronounced selection landscape, encompassing 322 unique genes **(Figure 6)**. These findings highlight ecotype-specific genomic regions potentially associated with environmental adaptation and functional traits.

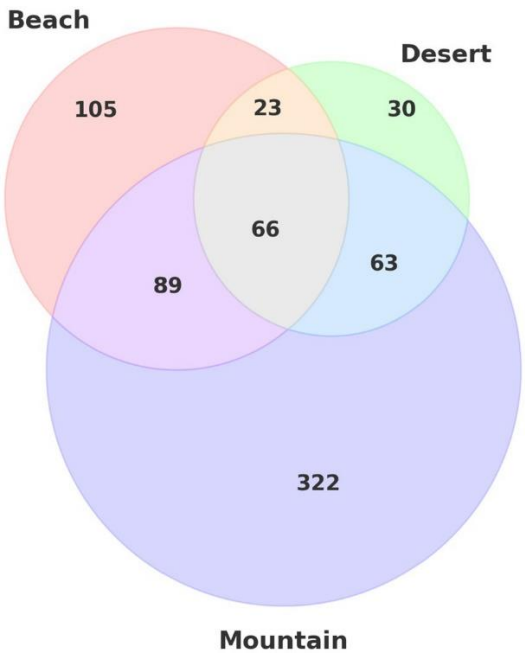


**Figure 6.** Venn diagram showing the overlap and distribution of gene sets under selection across beach, desert, and mountain camel populations

In addition to genes uniquely selected within each ecological group, several loci were shared between populations. Eighty-nine genes were common to both beach and mountain camels, 23 genes overlapped between beach and

desert populations, 63 genes were shared between desert and mountain groups, and 66 genes were detected across all three ecotypes.

Enrichment analysis revealed that the desert-adapted camels contained 21 genes with notable functional features, including HECT domains, regions rich in charged amino acids, and intrinsically disordered segments. These features were associated with an enrichment score of 0.11 and statistically significant p-values of 0.024 and 0.026, suggesting meaningful biological roles. Pathway-level examination highlighted Ubiquitin-ligase activity as the most prominent functional pathway in desert camels, with key genes such as HECTD4 and TMED5 contributing to this enrichment **(Tables 5–6)**.

**Table 4.** Overview of genes identified in the three camel ecotypes, including functional annotations, enrichment metrics, statistical significance, and representative genes

| Group | Recognized genes | Enrichment score (key cluster) | Significant *p*-values | Key functional annotations | Example genes |
|---|---|---|---|---|---|
| Deserts | 21 | 0.11 | 0.024 (HECT domain), 0.026 (HECT domain) | HECT domain, basic and acidic residues, disordered | HECTD4, TMED5 |
| Beach | 90 | 1.864 | 0.009 (P.H. domain), 0.030 (P.H. domain), 0.024 (HECT domain) | P.H. domain, HECT domain, thyroid hormone synthesis | DAGLB, TMEM70 |
| Mountain | 274 | 2.655 | 0.000345 (B30.2/SPRY domain), 0.000410 (SPRY_dom), 0.024 (HECT domain) | B30.2/SPRY domain, HECT domain, transferase activity | ITGB1, KIF3B |

**Table 5.** The top enriched pathways for the deserts, beach, and mountain groups and the corresponding *p*-values

| Group | Top enriched pathways | *p*-values |
|---|---|---|
| Deserts | Ubiquitin-ligase activity | 0.024 (HECT domain), 0.026 (HECT domain) |
| Beach | Phosphatidylinositol signaling, thyroid hormone synthesis | 0.009 (P.H. domain), 0.033 (Thyroid hormone synthesis), 0.024 (HECT domain) |
| Mountain | Immune response (B30.2/SPRY domains), metabolic adaptations | 0.000345 (B30.2/SPRY domain), 0.000410 (SPRY domain), 0.024 (HECT domain) |

**Table 6.** The table highlights specific genes, their functions, and their significance within the deserts, beach, and mountain ecotypes

| Group | Gene | Function | Significance |
|---|---|---|---|
| Deserts | *HECTD4* | Ubiquitin-ligase activity | Essential for protein maintenance in extreme temperatures |
| Beach | *TMED5* | Protein trafficking within the cell | Vital for maintaining cellular functions under stress. |
| | *DAGLB* | Endocannabinoid biosynthesis | Manages stress from varying salinity and nutrient availability |
| Mountain | *TMEM70* | Mitochondrial function and energy metabolism | Critical for adapting to the dynamic energy demands of coastal environments |
| | *ITGB1* | Cell adhesion and signal transduction | Adaptations to the physical demands of mountainous terrain |
| | *KIF3B* | Intracellular transport along microtubules | Supports cellular functions under high-altitude stress |

*Functional gene enrichment in beach and mountain camels*

Analysis of the beach-adapted camels identified 90 genes with notable functional features, including the P.H. domain, HECT domain, and genes involved in thyroid hormone synthesis **(Table 4)**. These annotations produced an enrichment score of 1.864, with significant p-values of 0.009 and 0.030 for the P.H. domain and 0.024 for the HECT domain. Further pathway analysis **(Table 5)** revealed that Phosphatidylinositol signaling and thyroid hormone synthesis were the most enriched processes in this group, with DAGLB and TMEM70 among the key contributing genes **(Table 6)**.

In the mountain ecotype, 274 genes exhibited enrichment in domains such as B30.2/SPRY, HECT, and transferase activity, achieving the highest overall enrichment score of 2.655 **(Table 4)**. Highly significant p-values were observed for the B30.2/SPRY domain (0.000345), SPRY domain (0.000410), and HECT domain (0.024). Enrichment analysis highlighted pathways associated with immune response and metabolic adaptation as the most prominent, with critical genes including ITGB1, KIF3B, and TMEM70 contributing to these processes **(Tables 5–6)**.

**Discussion**

Genetic variation among Saudi Arabian dromedary camels is shaped by a combination of environmental pressures, traditional breeding practices, and historical trade routes. These influences have contributed to the development of resilient populations capable of thriving in extreme desert conditions. Although the overall genetic diversity is relatively uniform, clear geographic and ecological structuring is evident [6, 23, 36].

In this study, high-quality genomic data were obtained, with 99.81% of sequencing reads successfully aligned to the reference genome. After thorough filtering, 34 samples containing 367,871 high-confidence variants were retained for subsequent analyses.

Population-level analyses revealed notable differences in heterozygosity, effective population size, and genomic inbreeding. PCA **(Figure 3)** illustrated a distinct separation between desert breeds (MAJ, SHU, SOF), which clustered tightly, and beach populations (SAH, AWA), which showed more dispersed clustering. The AWA breed, in particular, appeared as an independent cluster along PC1, consistent with its low heterozygosity (Ho = 0.374) and elevated FROH (0.015). These patterns likely reflect the combined effects of geographic isolation and selective breeding.

Admixture analyses corroborated these findings, highlighting historical gene flow among certain populations, such as SAH (beach) and HAD (mountain). At K = 3, desert populations displayed a distinct ancestral component, emphasizing their unique genetic makeup, while the AWA breed remained largely separate, suggesting limited admixture and retention of population-specific alleles. These observations align with previous studies on Omani dromedaries, which similarly showed maintenance of distinct local genetic lineages [23].

*Genomic insights into camel adaptation and population structure*

Investigating runs of homozygosity (ROH; **Table 3, Figure 5**) offered a window into the recent and historical demographic events across camel populations. In the AWA population, short ROHs (1–2 Mb), often indicative of past population bottlenecks or genetic drift, were the most frequent. HAD camels displayed a predominance of medium ROHs (2–5 Mb), reflecting moderate inbreeding or subtle population substructure. Extended ROHs exceeding 5 Mb, signaling very recent inbreeding, were more evenly observed in beach and mountain ecotypes but were comparatively scarce in desert-adapted breeds. These distributions appear to reflect both ecological limitations and historical breeding patterns.

Estimates of effective population size (Ne) were generally low across the groups, highlighting their susceptibility to the erosion of genetic variation. The MAJ population stood out with the largest Ne [36], consistent with its broader geographic spread and increased opportunities for gene flow. These findings underscore the roles of historical admixture and environmental pressures in shaping genetic diversity patterns [6].

Functional analyses revealed ecotype-specific adaptive signatures. Desert camels exhibited enrichment of HECT family E3 ubiquitin ligases, which are crucial for preserving protein stability under extreme oxidative stress, a common feature of arid environments [37–42]. In beach populations, DAGLB and TMEM70 were prominent, implicating lipid signaling and mitochondrial efficiency in coping with fluctuating salinity and nutrient conditions [43–45]. Genes involved in immune defense, including CX3CR1, IL6R, and CCR8, were also notable in desert populations, reflecting adaptation to pathogen-rich environments [46, 47]. Fertility-related genes such as ESR1 and SPACA5 point to mechanisms supporting reproductive success under harsh climatic pressures, consistent with findings in African zebu cattle [48, 49].

Other genes highlighted energy regulation and structural adaptation. ESRRG and CRTC1 play key roles in energy homeostasis, essential for survival in desert habitats and for high endurance in racing camels [50, 51]. Chondrogenesis-related genes like CHSY1 and CRLF1 indicate skeletal adaptations necessary for carrying loads and long-distance travel [52, 53]. Selective pressures from human use were evident in genes related to milk

production (PICALM) and locomotor performance (NAA16), mirroring domestication-driven selection observed in other livestock species [54, 55].

Collectively, these results reveal a complex interplay between natural environmental pressures and human-mediated breeding that has shaped the genetic landscape of Saudi Arabian dromedaries. The observed patterns of ROH, effective population size, and functional gene enrichment illustrate how adaptation, demography, and management practices converge to influence resilience, productivity, and performance. Integrating these genomic insights with ecological and physiological data will be crucial for designing informed conservation and breeding strategies that maintain genetic diversity and enhance the adaptive potential of camel populations [15, 23].

**Acknowledgments:** None

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

**Financial Support:** The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by Ministry of Environment, Water and Agriculture and the Genalive Medical Laboratory, Riyadh, Saudi Arabia.

**Ethics Statement:** The animal study protocol was approved by the Research Ethics Committee of King Saud University, Riyadh, Saudi Arabia (Ethics Reference No. KSU-SE-23-93, 19/10/2023). The study was conducted in accordance with the local legislation and institutional requirements.

**References**

1. Burger PA, Ciani E, Faye B. Old world camels in a modern world – a balancing act between conservation and genetic improvement. Anim Genet. 2019;50(6):598-612. doi:10.1111/age.12858
2. Hossam Mahmoud A, Abu-Tarbush FM, Alshaik M, Aljumaah R, Saleh A. Genetic diversity and population genetic structure of six dromedary camel (Camelus dromedarius) populations in Saudi Arabia. Saudi J Biol Sci. 2020;27(5):1384-89. doi:10.1016/j.sjbs.2019.11.041
3. Al-Swailem AM, Al-Busadah KA, Shehata MM, Al-Anazi IO, Askari E. Classification of Saudi Arabian camel (Camelus dromedarius) subtypes based on RAPD technique. J Food Agric Environ. 2007;5. doi:10.1234/4.2007.749
4. Khalkhali-Evrigh R, Hafezian SH, Hedayat-Evrigh N, Farhadi A, Bakhtiarizadeh MR. Genetic variants analysis of three dromedary camels using whole genome sequencing data. PLoS One. 2018;13(8):e0204028. doi:10.1371/journal.pone.0204028
5. Al Abri MA, Faye B. Genetic improvement in dromedary camels: challenges and opportunities. Front Genet. 2019;10:167. doi:10.3389/fgene.2019.00167
6. Almathen F, Bahbahani H, Elbir H, Alfattah M, Sheikh A, Hanotte O. Genetic structure of Arabian Peninsula dromedary camels revealed three geographic groups. Saudi J Biol Sci. 2022;29(1). doi:10.1016/j.sjbs.2021.11.032
7. Burger PA. The history of Old world camelids in the light of molecular genetics. Trop Anim Health Prod. 2016;48(5):905-13. doi:10.1007/s11250-016-1032-7
8. Legesse YW, Dunn CD, Mauldin MR, Ordonez-Garza N, Rowden GR, Gebre YM, et al. Morphometric and genetic variation in 8 breeds of Ethiopian camels (Camelus dromedarius). J Anim Sci. 2018;96(10). doi:10.1093/jas/sky351
9. Ahmad S, Yaqoob M, Hashmi N, Ahmad S, Zaman MA, Tariq M. Economic importance of camel: a unique alternative under crisis. Pak Vet J. 2010;30.
10. Khalaf S. Camel racing in the Gulf notes on the evolution of a traditional cultural sport. Anthropos. 1999;94.
11. Food and Agriculture Organization. FAOSTAT. 2018.

12. Mahmoud AH, Abou-tarboush FM, Alshaikh MA, Aljumaah RS. Genetic characterization and bottleneck analysis of Maghateer camel population in Saudi Arabia using microsatellite markers. J King Saud Univ Sci. 2020;32(1). doi:10.1016/j.jksus.2019.11.027
13. Harek D, Ikhlrf H, Bouhadad R, Sahel H, Cherifi Y, Djellout NE, et al. Genetic diversity status of camel's resources (Camelus dromedarius Linnaeus, 1758) in Algeria. GABJ. 2017;1. doi:10.46325/gabj.v1i1.87
14. Obšteter J, Jenko J, Pocrnic I, Gorjanc G. Investigating the benefits and perils of importing genetic material in small cattle breeding programs via simulation. J Dairy Sci. 2023;106(4). doi:10.3168/jds.2022-23132
15. Wu H, Guang X, Al-Fageeh MB, Cao J, Pan S, Zhou H, et al. Camelid genomes reveal evolution and adaptation to deserts environments. Nat Commun. 2014;5:5188. doi:10.1038/ncomms6188
16. Andrews S. FastQC. Babraham Bioinformatics. 2010.
17. Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, Li S, et al. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. GigaScience. 2018;7(6):gix120. doi:10.1093/gigascience/gix120
18. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-60. doi:10.1093/bioinformatics/btp324
19. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. GigaScience. 2021;10(2):giab008. doi:10.1093/gigascience/giab008
20. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078-9. doi:10.1093/bioinformatics/btp352
21. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297-303. doi:10.1101/gr.107524.110
22. Broad Institute. HaplotypeCaller. GATK best practices. Available at: https://gatk.broadinstitute.org/hc/en-us/articles/360037225632-HaplotypeCaller (accessed December 12, 2024).
23. Bahbahani H, Musa HH, Wragg D, Shuiep ES, Almathen F, Hanotte O. Genome diversity and signatures of selection for production and performance traits in dromedary camels. Front Genet. 2019;10:893. doi:10.3389/fgene.2019.00893
24. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience. 2015;4(1):s13742-015-0047-8. doi:10.1186/s13742-015-0047-8
25. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559-75. doi:10.1086/519795
26. El Hou A, Rocha D, Venot E, Blanquet V, Philippe R. Long-range linkage disequilibrium in French beef cattle breeds. Genet Sel Evol. 2021;53:63. doi:10.1186/s12711-021-00657-8
27. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. Bioinformatics. 2010;26(22):2867-73. doi:10.1093/bioinformatics/btq559
28. Ferenčaković M, Sölkner J, Curik I. Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. Genet Sel Evol. 2013;45:42. doi:10.1186/1297-9686-45-42
29. Purfield DC, Berry DP, McParland S, Bradley DG. Runs of homozygosity and population history in cattle. BMC Genet. 2012;13:70. doi:10.1186/1471-2156-13-70
30. Barbato M, Orozco-terWengel P, Tapio M, Bruford MW. SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. Front Genet. 2015;6:109. doi:10.3389/fgene.2015.00109
31. Sved JA, Feldman MW. Correlation and probability methods for one and two loci. In: ___*; 1973. p.129-32.
32. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 2009;19(9):1655-64. doi:10.1101/gr.094052.109
33. Ferrer-Admetlla A, Liang M, Korneliussen T, Nielsen R. On detecting incomplete soft or hard selective sweeps using haplotype structure. Mol Biol Evol. 2014;31(9):248-62. doi:10.1093/molbev/msu077
34. Szpiech ZA, Hernandez RD. Selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. Mol Biol Evol. 2014;31(10):2824-7. doi:10.1093/molbev/msu211

35. Garduño López VI, Martínez-Rocha R, Núñez Domínguez R, Ramírez Valverde R, Domínguez Viveros J, Reyes Ceron A, et al. Genome-wide scan for selection signatures in Mexican Sardo negro zebu cattle. PLoS One. 2024;19(2):e0312453. doi:10.1371/journal.pone.0312453
36. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res. 2022;50(10):gkac194. doi:10.1093/nar/gkac194
37. Al Abri M, Al-Sumry HS, Kashmar M, Al-Ward H, Al-Azri SS. Assessing genetic diversity and defining signatures of positive selection on the genome of dromedary camels from the southeast of the Arabian Peninsula. Front Vet Sci. 2023;10:1296610. doi:10.3389/fvets.2023.1296610
38. Abdelnour SA, Swelum AA, Abd El-Hack ME, Khafaga AF, Taha AE, Abdo M. Cellular and functional adaptation to thermal stress in ovarian granulosa cells in mammals. J Therm Biol. 2020;92:102688. doi:10.1016/j.jtherbio.2020.102688
39. Hoter A, Rizk S, Naim HY. Cellular and molecular adaptation of Arabian camel to heat stress. Front Genet. 2019;10:588. doi:10.3389/fgene.2019.00588
40. Marín I. Animal HECT ubiquitin ligases: evolution and functional implications. BMC Evol Biol. 2010;10:56. doi:10.1186/1471-2148-10-56
41. Metzger MB, Hristova VA, Weissman AM. HECT and RING finger families of E3 ubiquitin ligases at a glance. J Cell Sci. 2012;125(Pt 6):1181-7. doi:10.1242/jcs.091777
42. Qian H, Zhang Y, Wu B, Wu S, You S, Sun Y. Structure and function of HECT E3 ubiquitin ligases and their role in oxidative stress. J Transl Intern Med. 2020;8(2):71-9. doi:10.2478/jtim-2020-0012
43. Qian J, Chen W, Ye Y. Regulation of proteasome degradation by ubiquitination and deubiquitination. Front Mol Neurosci. 2020;13:140. doi:10.3389/fnmol.2020.00140
44. Bahri H, Buratto J, Rojo M, Dompierre JP, Salin B, Blancard C, et al. TMEM70 forms oligomeric scaffolds within mitochondrial cristae promoting in situ assembly of mammalian ATP synthase proton channel. Biochim Biophys Acta Mol Cell Res. 2021;1868(3):118942. doi:10.1016/j.bbamcr.2020.118942
45. Brar NK, Waggoner C, Reyes JA, Fairey R, Kelley KM. Evidence for thyroid endocrine disruption in wild fish in San Francisco Bay, California, USA. Aquat Toxicol. 2010;96(2). doi:10.1016/j.aquatox.2009.10.023
46. Ruiz-Jarabo I, Martos-Sitcha JA, Barragán-Méndez C, Martínez-Rodríguez G, Mancera JM, Arjona FJ. Gene expression of thyrotropin- and corticotrophin-releasing hormones is regulated by environmental salinity in the euryhaline teleost Sparus aurata. Fish Physiol Biochem. 2018;44(3). doi:10.1007/s10695-017-0457-x
47. Coghill JM, Fowler DH, West ML, Morice WG, Blazar BR. CCR8 expression identifies CD4+ T cells enriched for regulatory T-cell activity. Transplantation. 2013;85(4):620-5. doi:10.1097/TP.0b013e3181638625
48. Nesargikar PN, Spiller B, Chavez R. The complement system: history, pathways, cascade, and inhibitors. Eur J Microbiol Immunol. 2012;2(2):103-11. doi:10.1556/EuJMI.2.2012.2.2
49. Matthews J, Gustafsson JA. Estrogen receptor and the future of estrogen therapy. Nat Rev Cancer. 2003;3(10):802-11. doi:10.1038/nrc1206
50. Bahbahani H, Salim B, Almathen F, Al Enezi F, Mwacharo JM, Hanotte O, et al. Signatures of positive selection in African Butana and Kenana dairy zebu cattle. PLoS One. 2018;13(7):e0190446. doi:10.1371/journal.pone.0190446
51. Altarejos JY, Montminy M. CREB and the CRTC co-activators: sensors for hormonal and metabolic signals. Nat Rev Mol Cell Biol. 2008;12(3):141-51. doi:10.1038/nrm2069
52. Eichner LJ, Giguère V. Estrogen-related receptors (ERRs): a new dawn in transcriptional control of mitochondrial gene networks. Mol Endocrinol. 2011;25(4):585-93. doi:10.1210/me.2010-0028
53. Stefanovic B, Stefanovic L. TGF-β induction of cytokine receptor-like factor 1 in chondrocytes regulates cartilage metabolism. J Cell Physiol. 2012;227(3):901-9. doi:10.1002/jcp.22892
54. Wilson DG, Phamluong K, Lin WY, Barck K, Carano RA, Diehl L, et al. Chondroitin sulfate synthase 1 (Chsy1) is critical for cartilage and limb development in mice. Dev Dyn. 2012;241(11):1949-57. doi:10.1002/dvdy.23895
55. Kelly SA, Bell TA, Selitsky SR, Buus RJ, Hua K, Weinstock GM, et al. A novel N(alpha)-acetyltransferase gene (NAA16) is associated with increased running endurance in mice. Physiol Genomics. 2014;46(5):473-82. doi:10.1152/physiolgenomics.00120.2013

56. Sanchez MP, Ramayo-Caldas Y, Wolf V, Laithier C, Martin P, Palhière I, et al. Genetic control of milk production and composition in dairy sheep through variants in PICALM and other candidate genes. J Dairy Sci. 2017;100(12):9381-93. doi:10.3168/jds.2017-12988