J Genet Resour 2025; 11(2): 191-202 *Homepage: http://sc.journals.umz.ac.ir/*

RESEARCH ARTICLE *DOI: 10.22080/jgr.2025.28696.1424*

Development and Characterization SSR Makers Based on Next-generation Sequencing Technology in Iranian Turkmen Camel

Elham Rezvannejad[1], Karim Nobari[2*] and Mohammad Reza Bakhtiari Zadeh[3]

[1]Department of Biotechnology, Institute of Science and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman, Iran
[2]Animal Science Research Department, Tehran Agricultural and Natural Resources Research and Education Center, AREEO, Varamin, Iran
[3]Department of Animals and Poultry Science, Campus of Agriculture and Natural Resources, Aburaihan Campus, Tehran

ARTICLE INFO

Article history:
Received 17 May 2025
Accepted 06 July 2025
Available 24 July 2025

Keywords:
Primer
SSR marker
Turkmen camel
Whole genome sequencing

Supplementary information:
Supplementary information for this article is available at *http://sc.journals.umz.ac.ir/*

Corresponding authors:
✉ K. Nobari
k.nobari@areeo.ac.ir

p-ISSN 2423-4257
e-ISSN 2588-2589

ABSTRACT

The Iranian Turkmen one-humped camel (*Camelus dromedarius*) is an economically and culturally significant species valued for its milk and meat production. It is well-adapted to the arid and semi-arid climates of Middle East. Despite its importance, the genetic diversity and historical demography of this species remain poorly characterized compare to two-humped camel species, which exhibit limited genetic differentiation, the genetic diversity and historical demography of one-humped camels remain poorly understood. This study developed novel simple sequence repeat (SSR) markers for the Iranian Turkmen camel genome to facilitate genetic studies and marker-assisted breeding. We aimed to identify polymorphic loci to support analyses of genetic diversity, population structure, and evolutionary biogeography. Using Illumina HiSeq 2000 sequencing, we generated 589,326,158 clean 150 bp paired-end reads at 50x coverage. *De novo* assembly produced 235,978 contigs (N50= 8,526 bp) from which 151,556 SSR motifs were identified, primarily 2-mer (38.80%), 4-mer (21.85%), and 1-mer (16.20%) motifs. Primer pairs were designed for 144,184 loci (95%) (amplicon sizes of 100-180 bp). The Turkmen camel exhibited high SSR diversity (6,201 unique motifs), significantly exceeding than of Arabian and Bactrian camels (χ^2 = 7.14, p= 0.007, df= 1) with an estimated 78 to 510 additional unique motifs. These markers enable genetic diversity analysis, historical demography studies, and identification of breed-specific motifs (e.g., 1,179 novel motifs in Turkmen camels). Additionally, 19,425 codon-repeat loci, predominantly leucine-rich repeats (42%), were identified, potentially linked to stress response genes, offering insights into functional genomics. These SSR markers support conservation genomics and breeding programs for *Camelus* species, enhancing their long-term sustainability. Limitations include potential bias from using short contigs rather than scaffolds, and the need for functional validation of 19,425 codon-repeat loci, particularly leucine-rich repeats (42%).

Please cite this paper as: Rezvannejad, E., Nobari, K., & Bakhtiari Zadeh, M. R. (2025). Development and characterization SSR makers based on next-generation sequencing technology in Iranian Turkmen camel. *Journal of Genetic Resources,* 11(2), 191-202. doi: 10.22080/jgr.2025.28696.1424

Introduction

The Camelidae family, originating in North America during the Tortonian stage of the Miocene, later diverged into two tribs: Lamini and Camelini. Camelini, or the old-world camels, belong to the order Artiodactyla and migrated to the Eastern Hemisphere via the Bering land bridge (Wu *et al.*, 2014; Wu *et al.*, 2015). The genus *Camelus* comprises three extant species:

the one-humped dromedary (*Camelus dromedarius*), also known as Arabian camels; the two-humped domestic Bactrian (*Camelus bactrianus*); and the wild Bactrian camel (*Camelus ferus*) (Rao *et al.*, 1970). The extinction of wild dromedaries underscores the importance of conserving domestic popolations (Roth and Merz, 2013).

Dromedaries, constituting ~90% of the global camel population of 27.7 million, are primarily found in North Africa and parts of Asia (Kadim *et al.*, 2008). They valued for their milk, meat, fiber, transportation, and even dung as fuel (Burger, 2016). Their physiological traits enable them to survive in ability to survive in arid and semi-arid environments with low rainfall, extreme heat, and sparse vegetation, underscoring their role in food security. However, limited research and genetic improvement efforts, particularly in developing countries, have hindered their sustainable management (Fitak *et al.*, 2016).

In Iran, the Turkmen camel, *Camelus dromedaries* breed, holds significant cultural and economic value in Golestan Province. Prized for its milk and meat, this breed exhibits unique adaptations to local conditions, distinguishing it from other dromedary and Bactrian populations. Despite its importance, its genetic diversity and conservation status remain poorly characterized, exacerbated by threats like climate change, habitat loss, and genetic dilution through crossbreeding.

Microsatellites, or simple sequence repeats (SSRs), are powerful tools for studying genetic variation, population structure, and evolutionary relationships due to their high polymorphism, co-dominance, and multiallelic nature (Mahmoud *et al.*, 2012; Piro *et al.*, 2020; Yousefzadeh *et al*., 2024). Compared to single nucleotide polymorphisms (SNPs), SSRs offer higher discriminatory power per locus, requiring fewer markers for population-level analyses and are cost-effective for resource-limited settings (Nobari *et al.*, 2010; Guichoux *et al.*, 2011; Fatemi *et al*., 2025). However, fragmented draft genomes for *C. dromedarius*, *C. bactrianus*, and *C. ferus* limit comprehensive genetic studies, necessitating de novo SSR marker development using next-generation sequencing (NGS) (Fitak *et al.*, 2016; Ming *et al.*, 2020). NGS enables efficient identification of genome-wide SSRs, even without fully annotated reference genomes, facilitating the discovery of novel, breed-specific markers.

This study aims to address these gaps by developing novel SSR markers for the Iranian Turkmen camel using NGS. By comparing these markers with those from *Camelus dromedarius*, *Camelus bactrianus*, and *Camelus ferus*, we characterize genetic diversity, population structure, and breed-specific adaptations. These markers provide a foundation for targeted breeding and conservation strategies to ensure the long-term sustainability of this culturally and economically significant breed.

Material and Methods

Sampling, DNA extraction, and sequencing

Blood was collected from the jugular vein of a single female Iranian Turkmen camel (*C. dromedaries*), reared for hand milking, on a private farm in Golestan province, Iran. A 5ml syringe was used, and the sample was stored in a 5ml EDTA tube to prevent coagulation. Genomic DNA was extracted from whole blood using the salting-out method. DNA quality and integrity were assessed via 1% agarose gel electrophoresis strained with ethidium bromide, visualized under UV light. DNA concentration was quantified using a Qubit Fluorometer, yielding 304 ng/µl, confirming high molecular weight DNA suitable for sequencing. The DNA sample was sent for high-throughput whole genome sequencing using the Illumina HiSeq™2000 platform generating 150 bp paired-end reads at 50x coverage. *De novo* assembly was performed to generate contigs. Although based on a single individual, this approach is sufficient for de novo SSR discovery. However, it limits the assessment of allele frequencies and population-level polymorphism, which requires a larger sample sizes and validation across multiple individuals.

Quality filtering and de novo assembly

Raw sequence reads were quality-checked using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), confirming no adapter contamination. Reads shorter than 40 bp were removed, and a quality threshold of Q20 was applied, retaining sequences with at least 50% of bases meeting this score, using CLC Genomics Workbench (CLC Bio, Aarhus, Denmark). De

novo assembly was conducted with CLC Genomics Workbench using the following parameters: mismatch cost= 3, insertion/deletion cost= 3, length fraction= 0.5, and similarity fraction= 0.8. The assembled contings were compared to the dromedary reference genome (accession number: GCA_000803125.1) using MUMmer3 (nucmer).

SSR identification and primer design

Microsatellites (SSRs) were defined as repeat motifs of 1–10 nucleotides to include longer motifs (hepta- to deca-nucleotide repeats), which are less prone to replication slippage and stutter artifacts, enhancing specificity for population genetics and linkage mapping. SSRs were identified using SSR Locator software with the following criteria: mononucleotide motifs with repeats ≥ 20, dinucleotide motifs with repeats ≥ 10, trinucleotide motifs with repeats ≥ 7, tetranucleotide motifs with repeats ≥ 5, pentanucleotide and hexanucleotide motifs with repeats ≥ 4, and hepta-, octa-, nona-, and deca-nucleotide motifs with repeats ≥3. Primer pairs were designed in Primer3 for amplicons of 100-180 bp, with primer lengths of 15-25 bp (optimal: 20 bp), temperature range of 55-65°C (default 60°C), and the GC content was targeted to be between 40-60% (optimal: 50%) to ensure robust amplification and specificity (no experimental validation was performed in this study). SSR identification and primer design were also performed for Arabian dromedary (CamDro3), domestic Bactrian (Ca_bactrianus_MBC_1.0), and wild Bactrian (BCGSAC_Cfer_1.0) genomes, downloaded in FASTA format from the NCBI genome database. Consistent criteria were applied using SSR Locator to enable comparative analyses.

Results

DNA extraction and sequence assembly

Genomic DNA extracted from a female Iranian Turkmen camel (*Camelus dromedaries*) yielded a concentration of 304 ng/µl, quantified using a Qubit Fluorometer. Whole-genome sequencing on the Illumina HiSeq™ 2000 platform produced 589,326,158 clean 150 bp paired-end reads at 50x coverage, ensuring robust SSR detection. De novo assembly generated 235,978 contigs with an N50 of 8,526 bp, ranging from 200 bp to 683,514 bp (SD = 25,268). Contig size distribution (Fig. 1) showed a peak at ~200 bp, reflecting fragmentation typical of short-read sequencing. Approximately 92% of contigs aligned to the Arabian dromedary reference genome (GCA_000803125.1) using MUMmer3, confirming high sequence homology.

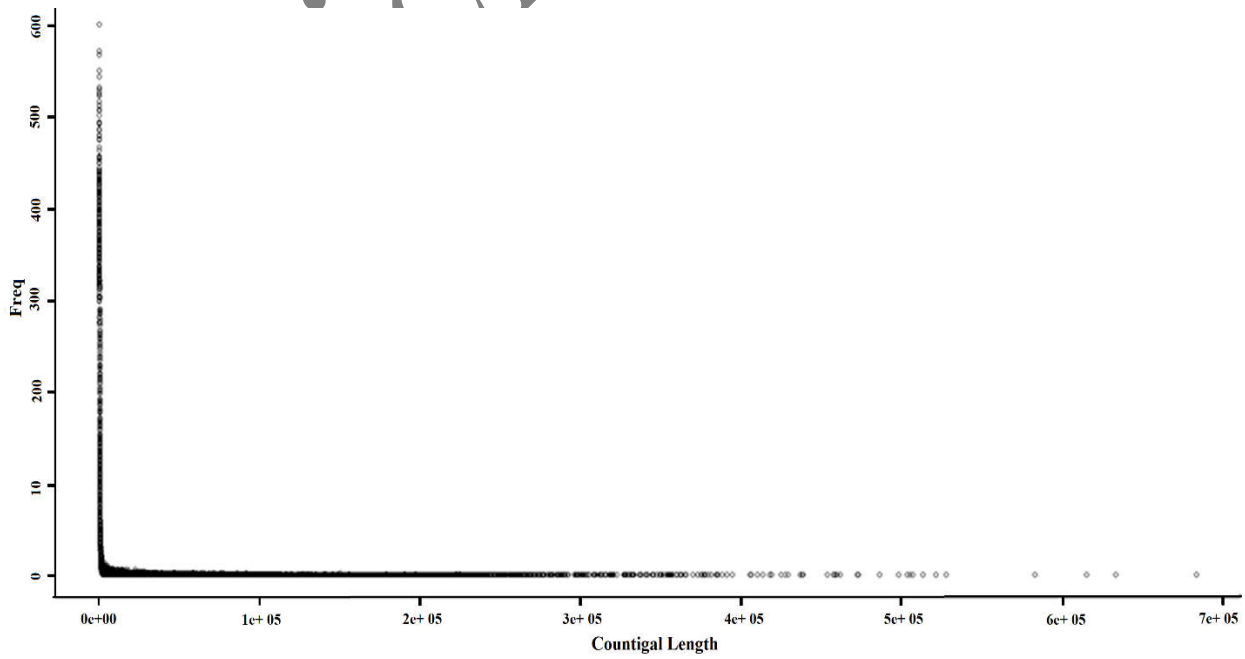

Fig. 1: Frequency of contigs of different sizes

SSR identification and comparative analysis

From the Turkmen camel genome, 152,381 microsatellites (SSRs) were identified, with 151,556 retained after filtering for ≥90% sequence identity and excluding low-complexity regions. Dinucleotide repeats were most abundant (38.80%), followed by tetranucleotide (21.85%), and mononucleotide (16.20%) repeats (Table 1).

Table 1. Number of different motifs for Turkmen dromedaries, Arabian dromedaries, domestic Bactrian camels, and wild Bactrian camels.

Ecotypes	Frequency	1-mer	2-mer	3-mer	4-mer	5-mer	6-mer	7-mer	8-mer	9-mer	10-mer	Total
Camelus dromedarius (Turkmen camel)	N	24550	58808	8414	33112	14182	4121	4583	2373	923	502	151556
	%	16.2	38.8	5.5	21.8	9.3	2.7	3.0	1.6	0.6	0.3	
Camelus dromedarius (Arabian camel)	N	29369	62208	8131	36188	13963	4042	4296	2496	918	497	162108
	%	18.1	38.4	5.0	22.3	8.6	2.5	2.7	1.5	0.6	0.3	
Camelus bacterianus (Bactrian camel)	N	28929	59913	8084	37820	14379	4311	4394	2660	907	484	161881
	%	17.9	37.0	5.0	23.4	8.9	2.7	2.7	1.6	0.6	0.3	
Camelus ferus (Wild Bactrian camel)	N	30182	60399	8142	36674	14208	4270	4382	2508	882	450	162097
	%	18.6	37.3	5.0	22.6	8.8	2.6	2.7	1.5	0.5	0.3	

Primer pairs were designed for 144,184 loci (95%) for genotyping. Repeat thresholds were: mononucleotide ≥20, dinucleotide ≥10, trinucleotide ≥7, and tetranucleotide ≥5, pentanucleotide and hexanucleotide ≥4, and hepta- to deca-nucleotide ≥3. Fig. 2 illustrates the number of repeats for each identified motif across the genome of the Turkmen camel and other camels for comparison purposes. The 10 motifs with the highest number of repeats for each camel are listed in Table 2.

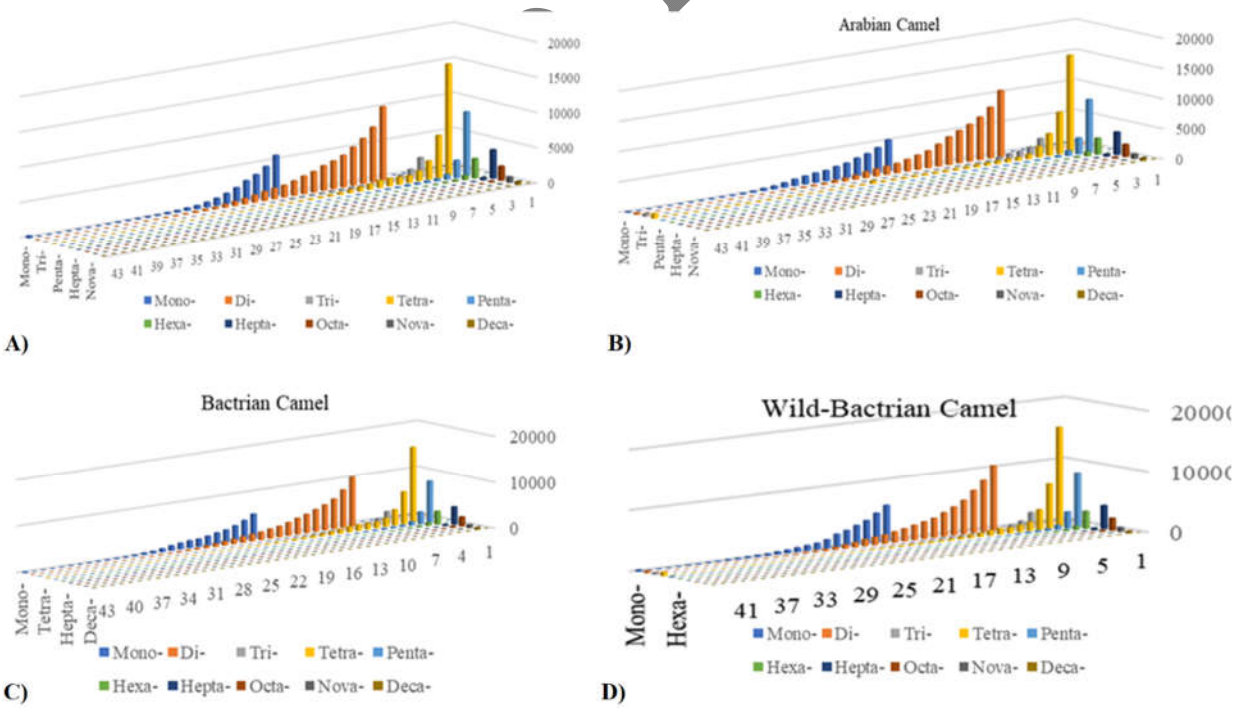

Fig. 2. Number of repeats for motifs ranging from Mono-nucleotide to Deca-nucleotide in the genomes: A) Turkmen camels; B) Arabian camels; C) *Camelus bacterianus*; D) Wild Bactrian camels.

Table 2. Frequency (percentage) of the 10 most repeated motifs in Turkmen camels, Arabian camels, Bactrian camels, and Wild Bactrian camels.

Strand1/Strand2	Turkmen camels	Arabian camels	Bactrian camels	Wild Bactrian camels
A/T	14.49	16.8	17.44	18.20
C/G	1.64	-	-	-
AC/GT	16.40	16.6	16.71	16.56
CA/TG	13.16	13.1	13.24	13.52
TA/TA	3.22	3.0	2.06	2.14
AT/AT	2.96	2.6	1.80	1.89
GA/TC	1.60	1.7	1.73	1.72
AG/CT	-	1.4	1.44	1.44
AAAC/GTTT	2.12	2.1	2.12	2.12
AAAAC/GTTTT	1.72	1.7	1.71	1.72
TAAA/TTTA	1.39	1.4	-	-
CAAA/TTTG	-	-	1.44	1.37

The Turkmen camel exhibited 6,201 unique motifs, significantly higher than Arabian dromedary (5,895), domestic Bactrian (5,979), and wild Bactrian (5,847) camels ($\chi^2 = 7.14$, p= 0.007, df= 1) (Table 3). Genomic site frequency for unique motifs (Fig. 3) showed larger motifs (hepta- to deca-nucleotide) typically occurred at single sites, while mono-, di-, and trinucleotide motifs appeared at multiple sites. Shared motifs across species are summarized in Table 4.

The Turkmen camel had 1,179 novel motifs unique to this breed, absent from NCBI databases, compared to 932 for Arabian dromedary, 971 for domestic Bactrian camel, and 945 for wild Bactrian camel. Dromedaries (Turkmen and Arabian) shared 2,870 novel motifs unique to their clade, while Bactrian camels shared 2,633. These novel motifs, defined as previously unreported in the NCBI databases, enhance breed-specific genetic characterization and conservation efforts. Of the 151,556 SSRs, primers were designed for 89-98% of motifs across categories (*e.g.*, 89% for dinucleotide, 97% for trinucleotide) (Table 5), with counts of 24550, 58808, 8414, 33112, 14182, 4121, 4583, 2373, 923, and 502 for di, tri, tetra, penta, hexa, hepta, octa, nona, and deca nucleotid repeats, respectively.

Table 3. Number of unique motifs for each group and their nucleotide lengths across the genome of different camels.

Ecotypes	Motif (Xmer)	1	2	3	4	5	6	7	8	9	10
Torkaman camel	Unique Motifs	4	12	60	225	573	871	1864	1440	716	436
	NoMG[1]	579835	863316	88329	234579	70704	19356	14336	7640	2907	1544
	MotNucCou[2]	579835	1726632	264987	938316	353520	116136	100352	61120	26163	15440
Arabian camel	Unique Motifs	4	12	60	215	566	821	1756	1385	663	413
	NoMG[1]	714047	913615	110567	343895	79592	21128	13834	10583	3058	1658
	MotNucCou[2]	714047	1827230	331701	1375580	397960	126768	96838	84664	27522	16580
Bactrian camel	Unique Motifs	4	12	59	223	572	855	1854	1361	630	409
	NoMG[1]	705731	884515	86045	282223	72380	20460	14048	10766	3297	1564
	MotNucCou[2]	705731	1769030	258135	1128892	361900	122760	98336	86128	29673	15640
Wild Bactrian camel	Unique Motifs	4	12	60	217	544	839	1802	1342	634	393
	NoMG[1]	741827	933429	129738	338148	87538	23934	14277	11101	3293	1562
	MotNucCou[2]	741827	1866858	389214	1352592	437690	143604	99939	88808	29637	15620

NoMG[1]: The number of motifs in the genome; MotNucCou[2]: The motif's nucleotide counts on the genome

Table 4. Number of shared unique motifs among different camels.

Camels	Turkmen camel	Wild Bactrian camel	Arabian camel	Bactrian camel
Torkaman camel				
Wild Bactrian camel	3327			
Arabian camel	3777	3384		
Bactrian camel	3508	3745	3376	

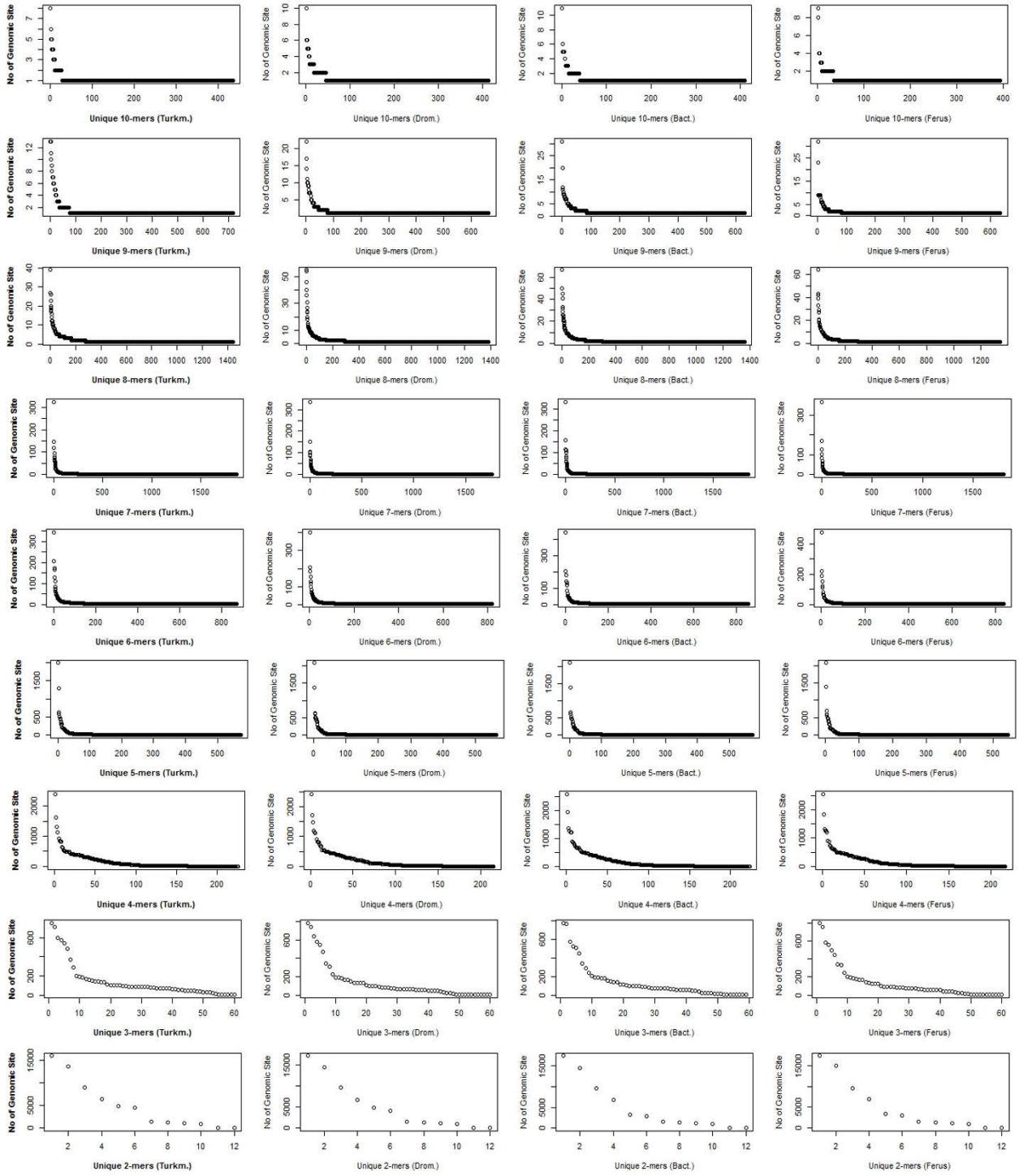

Fig. 3. The number of genomic sites for unique motifs (sorted by frequency) identified in the genomes of Turkmen, Arabian, Bactrian, and wild Bactrian camels.

Additionally, 19,425 codon-repeat loci were identified, predominantly leucine-rich (42%), with 68% in coding regions, suggesting possible roles in stress response and immune function genes (e.g., HSP70 and IL10), pending functional validation (Table 6). Repeat count distributions are shown in Fig. 4, and motif frequencies in the Fig. 5. A supplementary table details unique motifs, their genomic positions, and nucleotide counts across species (Supplement 1).

Table 5. Number and percentage of primers developed for each SSR motif

	1-mer	2-mer	3-mer	4-mer	5-mer	6-mer	7-mer	8-mer	9-mer	10-mer
Number	21928	57205	8001	31186	13617	4056	4450	2332	907	490
Percent	0.89	0.97	0.95	0.94	0.95	0.97	0.96	0.94	0.94	0.98
Total	144184 (0.95%)									

Table 6. Amino acid codon SSR loci, their frequencies, and repeat counts.

Amino acid	Number loci	Total repeats	Amino acid	Number loci	Total repeats
Leu	2526	20490	Pro	885	4470
Asn	1812	16021	Arg	793	4324
Ile	1335	12901	Gly	858	4253
Term	1070	10589	His	651	4013
Tyr	959	9687	Gln	585	3753
Lys	1309	6531	Ala	610	3598
Phe	1362	6424	Met	302	2513
Ser	905	5681	Asp	306	2276
Cys	786	4926	Glu	368	2039
Val	870	4852	Trn	349	1922
Thr	784	4499	total	19425	135762

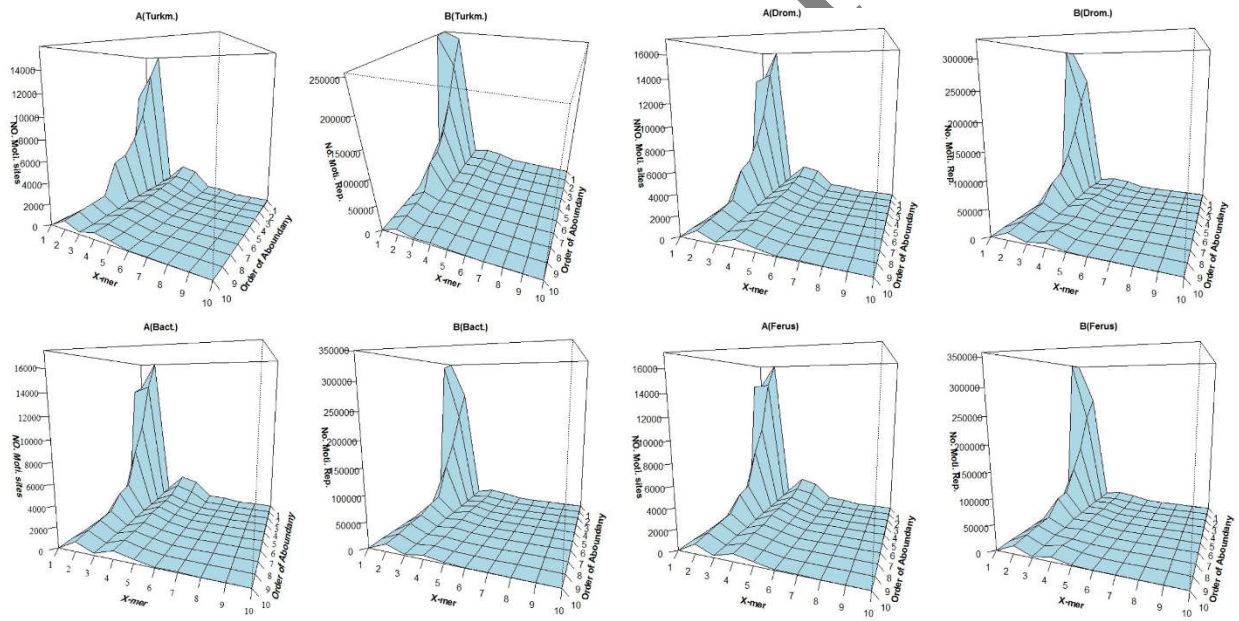

Fig. 4. Motifs identified in the genomes of Turkmen, Arabian, Bactrian, and wild Bactrian camels: A) The number of genomic sites; B) Repeat counts; Arabian showed with Drom.

Discussion

Contigs length is a critical factor in *de novo* whole-genome sequencing, influencing assembly quality for downstream analyses like SSR detection. Longer contigs reduce gaps, improving continuity and accuracy (Chen *et al*. 2015). The N50 metric, representing the contig length at which 50% of the genome is covered by contigs of that length or longer, assembly contiguity (Desai *et al*., 2013). Here, de novo assembly yielded 235,978 contigs with an N50 of 8,526 bp, sufficient to support robust SSR detection, despite a peak in short ~ 200 bp (Fig. 1). This fragmentation, typical of Illumina HiSeq 2000 short-read sequencing, may underrepresent SSR-rich regions, introducing bias in motif identification (Rhie *et al.*, 2021). Further studies

using long-read sequencing (e.g., PacBIO, Oxford Nanopore) could enhance contiguity and resolve repetitive regions (Srivastava *et al.*, 2019). Sequencing generated 589,326,158 clean reads at 50x coverage, calculated using SAMtools depth (v1.9, minimum mapping quality of 20) ensuring reliable SSR detection, including low-frequency motifs (Tóth *et al.*, 2000; Jeanjean *et al.*, 2025). Alignment of 92% of Turkmen camel contigs to the Arabian dromedary reference genome (GCA_000803125.1) using MUMmer3 confirmed high sequence homology and assembly quality. However, reliance on the Arabian dromedary reference may introduce mapping biases for Bacterian camels, where genomic divergence could lead to underrepresentation of lineage-specific SSRs.

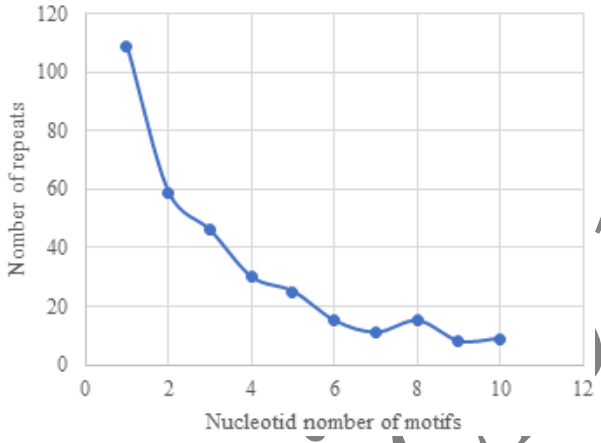

Fig. 5. The highest repeat counts for each SSR motif

The Turkemen camel genome contained 151,556 SSRs, with dinucleotide repeats (38.80%) most prevalent, followed by tetranucleotide (21.85) and mononucleotide (16.20%) repeats (Table 1). This distribution aligns with patterns in other camels (Hedayat-Evrigh *et al.* 2018; Barazandeh, *et al.* 2020; Shaltenbay *et al.*, 2025) and animals (Yan *et al.*, 2008; Bakhtiarizadeh *et al.*, 2012, Abe & Gemmell 2014, Sadder *et al.*, 2015), where shorter repeats predominate due to higher mutation rates and replication slippage (Tóth *et al.*, 2000). In contrast, some plants show higher trinucleotide repeat frequencies (Varshney *et al.*, 2002; Durand *et al.*, 2010; Joshi *et al.*, 2010; Zhou *et al.*, 2016; Wang *et al.* 2017).

Table 2 demonstrated that the frequent dinucleotide across all camels was AC/GT, with an approximate frequency of 16% for all camels. This was followed by the mononocleotide A/T, which exhibited variable frequencies among different camels. While breed-specific motifs (*e,g,* C/G in Turkmen, CAAA/TTTG in Bacterian camels) suggest evolutionary divergence, valuable for breed identification and conservation efforts.

Compared to Barazandeh *et al.* (2020), who reported AC/TG (54%) and AT/TA (32.8%), with GC/CG having the lowest frequency (1.2%). Their most frequent 3-mer motif was GCC/GGC (19.2%), followed by AGC/GCT (10.3%). These differences likely stem from methodological distinctions: our study used DNA-seq, capturing both coding and non-coding regions, whereas Barazandeh *et al.* (2020) used RNA-seq, which focuses on transcribed regions. Coding regions, subject to purifying selection, may limit SSR expansion to maintain protein function, leading to lower 2-mer frequencies in RNA-seq data. In contrast, non-coding regions, under relaxed selection, allow greater SSR variability, explaining the higher 2-mer frequency in our DNA-seq data. This highlights the complementary nature of DNA-seq and RNA-seq approaches in camel genomics, as they capture different compartments and selective pressures (Tóth *et al*., 2000).

The Turkmen camel exhibited 6,201 unique motifs, surpassing Arabian dromedary (5,895), domestic Bactrian (5,979) and wild Bactrian (5,847) camels (Table 3), indicating greater SSR diversity potentially linked to environmental adaptations (Pandey *et al*., 2018). Of these, 1,179 novel motifs were unique to the Turkmen camel, enhancing its genetic characterization. Dromedaries shared 2,870 novel motifs, and Bactrian camels 2,633, reflecting clade-specific variation. Although our study analyzed a single individual, Turkmen camels belong to a broader dromedary population shaped by geographic isolation, migration, and crossbreeding. Future multi-individual studies will be required to assess population substructure, genetic drift, and gene flow, which may significantly influence SSR diversity.

The results presented in Table 5 showed that out of a total of 151,556 SSRs identified in the Turkmen camel, primers could be designed for 144,184 (95%) of them. Therefore, the SSR

markers developed and their corresponding primers in this study provide a valuable resource for future research on genetic diversity, historical and geographical demography, evolution, and conservation genomics of various camel species. SSR markers, along with their corresponding primers, can also aid in enhancing genetic studies of the species with limited genetic characterization and registration.

Among 19,425 codon-repeat loci, 68% were in coding regions, with leucine-rich repeats (42%) (Table 6) suggesting roles in stress response and immune function (*e.g.*, HSP70, IL10) genes (Kumar *et al*., 2016; Glazko *et al*., 2023). These findings warrant functional validation.

Limitations include the reliance on a single-individual sample, restricting allele frequency and polymorphism assessments. Short-read sequencing technologies (*e.g.*, Illumina) are prone to systematic biases in repetitive and GC-rich regions, which may result in underrepresentation of certain SSR motifs. This limitation highlights the need for integrating long-read sequencing platforms in future camel genomics to overcome repetitive region biases. In the end, the SSR primers were designed *in silico* and were not experimentally validated; while computational primer design using Primer3 provides a strong foundation, laboratory-based testing is necessary to confirm amplification efficiency, specificity, and reproducibility.

Conclusion

This study identified 151,556 SSR motifs in the Turkmen camel genome, with dinucleotide (38%), tetranucleotide (21%), and mononucleotide (16%) repeats most prevalent. The Turkmen camel exhibited the highest SSR diversity (6,201 unique motifs) significantly exceeding other camels ($\chi^2 = 7.14$, p= 0.007, df= 1), with 78 to 510 additional unique motifs. Primer pairs for 144,184 loci (95%) support population genetic studies, linkage mapping, and conservation. Additionally, 19,425 codon-repeat loci, predominantly leucine-rich (42%), with 68% in coding regions, suggest roles in stress response and immune function. Despite robust 50x coverage and N50 of 8,526 bp, limitations include the single-individual sample, restricting population-level analyses, and short contings potentially missing SSR-rich regions. Although 144,184 primer pairs were computationally designed using Primer3, reliance on the Arabian dromedary reference genome may also introduce alignment bias for Bactrian camels. Further studies should use longer-read sequencing (e.g., PacBio, Oxford Nanopore), larger sample sizes, and functional asseys to validate codon-repeat roles and marker polymorphism across Camelus species, enhancing breeding and conservation strategies for the Turkmen.

Acknowledgment

The authors gratefully acknowledge for the research funding provided for this project (No. 02/3365) by Institute of Science and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman, Iran.

Data availability statement

The sequencing data generated in this study are available from the corresponding author upon request.

Conflict of interest

There is no conflict of interest between authors.

References

Abe, H., & Gemmell, N. J. (2014). Abundance, arrangement, & function of sequence motifs in the chicken promoters. *BMC Genomics*, 15(1), 900. https://doi.org/10.1186/1471-2164-15-900

Bakhtiarizadeh, M., Arefnejad, B., Ebrahimie, E., & Ebrahimi, M. (2012). Application of functional genomic information to develop efficient EST-SSRs for the chicken (*Gallus gallus*). *Genetics and Molecular Research*, 11(2), 1558-1574. https://doi.org/10.4238/2012.may.21.12

Barazandeh, A., Mokhtari, M., Moghbeli Damaneh, M., & Roudbari, Z. (2020). Analyzing simple sequence repeats derived from expressed sequence tags in dromedary camels. *Iranian Journal of Applied Animal Science*, 10(3), 555-565. http://www.ijas.ir/

Burger, P. A. (2016). The history of old world camelids in the light of molecular genetics.

Tropical Animal Health and Production, 48(5), 905-913.

Chen, T. W., Gan, R. C., Chang, Y. F., Liao, W. C., Wu, T. H., Lee, C. C., ... & Tang, P. (2015). Is the whole greater than the sum of its parts? De novo assembly strategies for bacterial genomes based on paired-end sequencing. *BMC Genomics*, 16(1), 648. https://doi.org/10.1186/s12864-015-1859-8

Cuadrado, A., & Schwarzacher, T. (1998). The chromosomal organization of simple sequence repeats in wheat and rye genomes. *Chromosoma*, 107(8), 587-594. https://doi.org/10.1007/s004120050345

Desai, A., Marwah, V. S., Yadav, A., Jha, V., Dhaygude, K., Bangar, U., Kulkarni, V., & Jere, A. (2013). Identification of optimum sequencing depth especially for *de novo* genome assembly of small genomes using next generation sequencing data. *PloS One*, 8(4), e60204. https://doi.org/10.1371/journal.pone.0060204

Durand, J., Bodénès, C., Chancerel, E., Frigerio, J.-M., Vendramin, G., Sebastiani, F., Buonamici, A., Gailing, O., Koelewijn, H.-P., & Villani, F. (2010). A fast & cost-effective approach to develop and map EST-SSR markers: oak as a case study. *BMC Genomics*, 11, 570. https://doi.org/10.1186/1471-2164-11-570

Fatemi, M., Estaji, A., Ghanbari, A., Torabi Giglou, M., & Jamali, M. (2025). Evaluation of genetic diversity of wild raspberry genotypes in the western areas of the Caspian Sea by ISSR molecular markers. *Journal of Genetic Resources*, 11(1), 98-104. https://doi.org/10.22080/jgr.2025.28406.1420

Fitak, R. R., Mohandesan, E., Corander, J., & Burger, P. A. (2016). The *de novo* genome assembly and annotation of a female domestic dromedary of North African origin. *Molecular Ecology Resources*, 16(1), 314-324. https://doi.org/10.1111/1755-0998.12443

Glazko, V. I., Kosovsky, G. Y., Glazko, T. T., & Fedorova, L. M. (2023). DNA markers and microsatellite code. *Sel'skokhozyaistvennaya Biologiya* ,58(2), 223-248. https://doi.org/10.15389/agrobiology.2023.2.223

Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O., ... & Petit, R. J. (2011). Current trends in microsatellite genotyping. *Molecular Ecology Resources*, 11(4), 591-611. https://doi.org/10.1111/j.1755-0998.2011.03014.x

Hedayat-Evrigh, N., Miraei-Ashtiani, S. R., Khalkhali Ivrigh, R., & Pourasad, K. (2018). Molecular assessment of genetic diversity in dromedaries and Bactrian camel using microsatellite markers. *Journal of Agricultural Science and Technology*, 20(6), 1137-1148. https://doi.org/10.1007/s11250-016-1032-7

Jeanjean, S. I., Shen, Y., Hardy, L. M., Daunay, A., Delépine, M., Gerber, Z., ... & How-Kit, A. (2025). A detailed analysis of second and third-generation sequencing approaches for accurate length determination of short tandem repeats and homopolymers. *Nucleic Acids Research*, 53(5), gkaf131. https://doi.org/10.1093/nar/gkaf131

Joshi, R. K., Kuanar, A., Mohanty, S., Subudhi, E., & Nayak, S. (2010). Mining and characterization of EST derived microsatellites in *Curcuma longa* L. *Bioinformation*, *5*(3), 128- 131. https://doi.org/10.6026/97320630005128

Kadim, I., Mahgoub, O., & Purchas, R. (2008). A review of the growth, and of the carcass and meat quality characteristics of the one-humped camel (*Camelus dromedaries*). *Meat Sci*ence, 80(3), 555-569. https://doi.org/10.1016/j.meatsci.2008.02.010

Kelkar, Y. D., Tyekucheva, S., Chiaromonte, F., & Makova, K. D. (2008). The genome-wide determinants of human and chimpanzee microsatellite evolution. *Genome Research*, 18(1), 30-38. http://www.genome.org/cgi/doi/10.1101/gr.7113408

Kumar, A. S., Sowpati, D. T., & Mishra, R. K. (2016). Single amino acid repeats in the

proteome world: structural, functional, and evolutionary insights. *PloS One*, 11(11), e0166854. https://doi.org/10.1371/journal.pone.0166854

Mahmoud, A., Alshaikh, M., Aljumaah, R., & Mohammed, O. (2012). Genetic variability of camel (*Camelus dromedarius*) populations in Saudi Arabia based on microsatellites analysis. *African Journal of Biotechnology*, 11(51), 11173-11180. https://doi.org/10.5897/AJB12.1081

Ming, L., Yuan, L., Yi, L., Ding, G., Hasi, S., Chen, G., Jambl, T., Hedayat-Evright, N., Batmunkh, M., & Badmaevna, G. K. (2020). Whole-genome sequencing of 128 camels across Asia reveals origin and migration of domestic Bactrian camels. *Communications Biology*, 3(1), 1. https://doi.org/10.1038/s42003-019-0734-6

Nobari, K., Ghazanfari, S., Nassiry, M. R., Tahmoorespur, M., & Jorjani, E. (2010). Relationship between leptin gene polymorphism with economical traits in Iranian Sistani and Brown Swiss cows. https://doi.org/10.3923/javaa.2010.2807.2810

Pandey, M., Kumar, R., Srivastava, P., Agarwal, S., Srivastava, S., Nagpure, N. S., Jena, J. K., & Kushwaha, B. (2018). WGSSAT: a high-throughput computational pipeline for mining and annotation of SSR markers from whole genomes. *Journal of Heredity*, 109(3), 339-343. https://doi.org/10.1093/jhered/esx075

Piro, M., Mabsoute, F., El Khattaby, N., Laghouaouta, H., & Boujenane, I. (2020). Genetic variability of dromedary camel populations based on microsatellite markers. *Animal*, 14(12), 2452-2462. https://doi.org/10.1017/S1751731120001573

Rao, M., Gupta, R., & Dastur, N. (1970). Camels' milk and milk products. *Indian Journal of Dairy Science*, 23, 71-78. https://epubs.icar.org.in/index.php/IJDS

Rhie, A., McCarthy, S. A., Fedrigo, O., et al. (2021). Towards complete and error-free genome assemblies of all vertebrate species. *Nature*, 592(7856), 737-746. https://doi.org/10.1038/s41586-021-03451-0

Roth, H. H., & Merz, G. (2013). Wildlife resources: a global account of economic use. Springer Science and Business Media.

Sadder, M., Migdadi, H., Al-Haidary, A., & Okab, A. (2015). Identification of simple sequence repeat markers in the dromedary (*Camelus dromedarius*) genome by next-generation sequencing. *Turkish Journal of Veterinary and Animal Sciences*, 39(2), 218-228. https://doi.org/10.3906/vet-1402-72

Shaltenbay, G., Ualiyeva, D., Kapassuly, T., Kozhakhmet, A., Orazymbetova, Z., Kulboldin, T., ... & Dossybayev, K. (2025). Genetic variability and population structure of *Camelus* from Kazakhstan Inferred from 17 STR markers. *Diversity*, 17(7), 459. https://doi.org/10.3390/d17070459

Srivastava, S., Avvaru, A.K., Sowpati, D.T., & Mishra, R.K. (2019). Patterns of microsatellite distribution across eukaryotic genomes. *BMC Genomics*, 20(1), 153. https://doi.org/10.1186/s12864-019-5516-5

Teneva, A., Dimitrov, K., Petrović, C. V., Petrović, M. P., Dimitrova, I., Tyufekchiev, N., & Petrov, N. (2013). Molecular genetics and SSR markers as a new practice in farm animal genomic analysis for breeding and control of disease disorders. *Biotechnology in Animal Husbandry*, 29(3), 405-429. https://doi.org/10.2298/BAH1303405T

Tóth, G., Gáspári, Z., & Jurka, J. (2000). Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research*, 10(7), 967-981. https://doi.org/10.1101/gr.10.7.967

Varshney, R. K., Thiel, T., Stein, N., Langridge, P., & Graner, A. (2002). In silico analysis on frequency & distribution of microsatellites in ESTs of some cereal species. *Cellular and Molecular Biology Letters*, 7(2A), 537-546. http://www.cmbl.org.pl

Wang, P., Yang, L., Zhang, E., Qin, Z., Wang, H., Liao, Y., Wang, X., & Gao, L. (2017). Characterization and development of EST-SSR markers from a cold-stressed transcriptome of centipedegrass by illumina paired-end sequencing. *Plant Molecular Biology Reporter*, 35, 215-223. https://doi.org/10.1007/s11105-016-1017-8

Wu, H., Guang, X., Al-Fageeh, M. B., Cao, J., Pan, S., Zhou, H., ... & Wang, J. (2014). Camelid genomes reveal evolution and adaptation to desert environments. *Nature Communications*, 5(1), 5188. https://doi.org/10.1038/ncomms6188

Yan, Q., Zhang, Y., Li, H., Wei, C., Niu, L., Guan, S., Li, S., & Du, L. (2008). Identification of microsatellites in cattle unigenes. *Journal of Genetics and Genomics*, 35(5), 261-266. https://doi.org/10.1016/S1673-8527(08)60037-5

Young, E. T., Sloan, J. S., & Van Riper, K. (2000). Trinucleotide repeats are clustered in regulatory genes in *Saccharomyces cerevisiae*. *Genetics*, 154(3), 1053-1068. https://doi.org/10.1093/genetics/154.3.1053

Yousefzadeh, H., Nasiri, M., Amirchakhmaghi, N., Sabbagh, S. A., Walas, Ł., & Kozlowski, G. (2024). Transmission of genetic variation from the adult generation to naturally established seedlings of Fagus orientalis in the Hyrcanian forest. *Journal of Genetic Resources*, 10(2), 152-163. https://doi.org/10.22080/jgr.2024.26853.1387

Zhou, Q., Luo, D., Ma, L., Xie, W., Wang, Y., Wang, Y., & Liu, Z. (2016). Development and cross-species transferability of EST-SSR markers in Siberian wildrye (*Elymus sibiricus* L.) using Illumina sequencing. *Scientific Reports*, 6(1), 20549. https://doi.org/10.1038/srep20549