2026 SCIENCELINE

Online Journal of Animal and Feed Research

Volume 16, Issue 4: 231-239; July 30, 2026 ISSN 2228-7701

DOI: https://dx.doi.org/10.51227/ojafr.2026.28

IN SILICO ANALYSIS OF THE PROMOTER REGION OF PROLACTIN GENE IN CATTLE TO UNDERSTAND ITS GENE REGULATION

Bakhtiyor ADILOV, Boburbek AKHMADALIEV, Anvar SHERIMBETOV, Bobur ABDIKARIMOV, Bobir ABDUVALIEV, and Dilshodbek RUZMETOV

Animal genetics laboratory, Institute of Genetics and Plant Experimental Biology Academy of Sciences of the Republic of Uzbekistan, Kibray District, Tashkent Region 111226, Uzbekistan

Email: ahmadaliyev_bobur@mail.ru

Supporting Information

RESEARCH ARTICLE

PII: S2228770126000028-16

Received: April 20, 2026

Revised: July 19, 2026

Accepted: July 22, 2026

ABSTRACT: *Prolactin* (*PRL*) is a key endocrine regulator of lactation, mammary gland growth, and systemic metabolic homeostasis in cattle. The transcriptional architecture of bovine *PRL* regulation remains incompletely characterized, despite its physiological importance. In this study we used an *in silico* TSS, CpG island, and TFBS profiling approach to analyze the ~5 kb upstream regulatory region of the *Bos taurus PRL* gene. By employing Neural Network Promoter Prediction under stringent criteria, we identified three highly reliable TSSs located at -4478, -2123, and -517 base pairs relative to the translational start codon, indicating a multi-TSS promoter complex that facilitates alternative transcript initiation. Canonical CpG islands were absent, consistent with tissue-specific, fine-tuned transcriptional control. TFBS analysis revealed a densely populated, hierarchically organized promoter enriched for pituitary-specific POU1F1 motifs, hormone-responsive ESRRA and RARA elements, signal-dependent STAT5 and AP-1 sites, and developmental transcription factors such as NANOG, HOXD9, and FOXI3. High-affinity TFBS clusters, including cis-regulatory modules of activators and repressors (like PRDM1) suggest combinatorial control and dynamic transcriptional responsiveness. These findings give mechanistic insights into the transcriptional complexity of bovine *PRL* locus and built an essential resource base on functional genomics studies to enhance efficiency of lactation and endocrine resilience in cattle.

Keywords: *Bos taurus*, Lactation, Mammary gland growth, Prolactin gene, Promoter region, Transcription start sites.

INTRODUCTION

In cattle, the genetic and regulatory basis of economically important traits such as milk production has attracted considerable attention. The bovine *prolactin* (*PRL*) gene has been extensively investigated because of its associations with milk yield, fertility, and lactation performance. Several studies have reported polymorphisms within the *PRL* gene and their associations with milk production traits (Motmain et al., 2022; Ilie et al., 2023). In addition, previous studies have characterized specific regulatory variants and promoter polymorphisms of the bovine *PRL* gene. However, most of these studies have focused primarily on genetic associations and sequence polymorphisms rather than on the comprehensive regulatory architecture of the *PRL* promoter (Brym et al., 2007).

To date, information regarding putative transcription start sites, the distribution of transcription factor binding motifs, and the presence of CpG islands within the bovine *PRL* promoter remains limited. Therefore, a comprehensive *in silico* characterization of the bovine *PRL* promoter is still lacking. Recent advances in genomics and bioinformatics have made it possible to predict promoter regions, transcription factor binding sites (TFBS), CpG islands, and conserved regulatory motifs using computational approaches without requiring immediate experimental validation (Matys et al., 2006; Fornes et al., 2020).

Such analyses are useful for generating hypotheses regarding regulatory mechanisms prior to experimental validation. Therefore, the present study aimed to perform a comprehensive *in silico* characterization of the bovine *PRL* promoter region and to identify putative regulatory elements that warrant further experimental validation and functional investigation.

MATERIALS AND METHODS

Identification of transcription start sites

The genomic sequence of the bovine *PRL* gene was retrieved from the NCBI Gene database (Gene ID: 280901; RefSeq transcript accession: NM_173953.2) based on the *Bos taurus* reference genome assembly ARS-UCD1.2. According to the NCBI reference genome, the *PRL* gene is located on chromosome 23 (NC_037350.1:35332693–

Citation: Adilov B, Akhmadaliev B, Sherimbetov A, Abdikarimov B, Abduvaliev B, and Ruzmetov D (2026). *In silico* analysis of the promoter region of prolactin gene in cattle to understand its gene regulation. *Online J. Anim. Feed Res.*, 16(4): 231-239. DOI: https://dx.doi.org/10.51227/ojafr.2026.28

35341308) on the forward (+) strand. A 5-kb genomic region upstream of the translation start codon (ATG) of the reference transcript was extracted (NC_037350.1:35327693–35332692) and used for promoter analysis. The 5-kb upstream region was selected to include both proximal and potential distal regulatory elements that may contribute to transcriptional regulation. Putative transcription start sites (TSSs) were predicted using the Neural Network Promoter Prediction (NNPP) server (https://www.fruitfly.org/seq_tools/promoter.html) (Reese, 2001). NNPP employs a time-delay artificial neural network trained on experimentally validated eukaryotic promoter sequences to identify potential promoter regions and putative TSSs based on sequence characteristics and promoter-associated motifs. Predictions were performed using the default parameters with a minimum promoter prediction score threshold of 0.80.

Transcription factor binding site prediction

The transcription factor binding site (TFBS) prediction was done from the AnimalTFDB v4.0 platform (https://guolab.wchscu.cn/AnimalTFDB4/#/TFBS_Predict), a broad animal transcription factors and regulatory annotations repository (Shen et al., 2023). AnimalTFDB integrates genome-wide TF repertoires from Ensembl (release 105) and provides a TFBS prediction module based on curated motif libraries across multiple species. The TFBS identification was based on a position weight matrix (PWM)-based scanning implemented in AnimalTFDB. This prediction module includes quality and computationally inferred TF binding motifs from several public databases like HOCOMOCO, TRANSFAC, JASPAR, CIS-BP, and hTFtarget (Kulakovskiy et al., 2018; Fornes et al., 2020). These motifs provide high-confidence transcription factor binding profiles. The input nucleotide sequence is scanned against the integrated motif datasets using PWM-based scoring algorithms which assess the potential of TF–DNA interactions through sequence similarity to established binding motifs (Stormo, 2013). AnimalTFDB provides a motif score and p-value for each predicted binding site. The score reflects the log-likelihood ratio of the observed sequence against a background model, and the p-value describes the probability of obtaining a match with the same or better strength by chance. For TFBS prediction, motifs with p-values ≤ 1×10^{-4} and motif match scores exceeding the cut-off values recommended by AnimalTFDB were retained for analysis, whereas motifs with $P \leq 1 \times 10^{-5}$ were considered high-confidence predictions. These criteria are commonly used in genome-wide motif scanning to reduce false-positive predictions (Fornes et al., 2020). Unless otherwise stated, the default background nucleotide frequencies were used for TFBS prediction by AnimalTFDB. The predicted TFBSs were annotated with the names of transcription factors, genomic coordinates, orientation of strands, and metrics of statistical significance.

Identification of CpG islands in promoter regions

CpG islands were determined using a strict computational scheme within the promoter regions. The criteria included a CpG island with a GC content of at least 50%, an observed to expected CpG ratio of ≥ 0.60, and an island minimum length of 200bp (Takai and Jones, 2002; Long et al., 2013). CpG islands were identified using CpG Island Searcher (CpGi130) program at http://www.bioinformatics.org/sms/cpg_island.html (Takai and Jones, 2002).

RESULTS

Identification of TSSs

Neural Network Promoter Prediction (NNPP) analysis predicted three putative transcription start sites (TSSs) within the *Bos taurus PRL* promoter region (Table 1). These putative TSSs received NNPP scores of 0.98, 0.94, and 0.84, suggesting the presence of promoter-like sequence features. Positional mapping relative to the translational start codon (ATG) indicated that the predicted putative TSSs were located approximately at −4478 bp, −2123 bp, and −517 bp upstream. However, these predictions should be considered putative and require experimental validation or support from public transcriptomic datasets. Multiple TSSs may be alternative transcription starts, with a possible implication for complex regulation of *PRL* expression in cattle. The proximal TSS with -517 bp and a higher prediction score is presumably representing the core promoter region while the more distal sites may function as auxiliary or tissue-specific regulatory elements. These findings, therefore, implied a multiple-TSS architecture of the *PRL* promoter in cattle, which could integrate differential transcriptional control mechanisms with potential functional and physiological implications in lactation and endocrine regulations in cattle.

Table 1. Number and predictive score value for cattle *PRL* gene TSSs.

Promoter region name	No. of TSS identified	Predictive score at cutoff value of 0.8	Distance from start codon (ATG)
Bos_taurus_pro-*PRL*	3	0.98, 0.94, 0.84	–4478, –2123, –517

Identification of CpG islands

Analysis of the *PRL* promoter revealed no detectable CpG islands. Despite comprehensive screening using established prediction algorithms and criteria for CpG island identification (length ≥ 200 bp, GC content ≥ 50%, observed-to-expected CpG ratio ≥ 0.6), this region did not exhibit characteristics consistent with canonical CpG islands.

Common motifs and transcription factors

Transcription Factor Binding Site Landscape of the *Bos taurus Prolactin* 5 kb Promoter. The promoter region of the *Bos taurus PRL* gene (5kb) contains a high density of transcription factor binding sites (TFBS) that are functionally diverse and hierarchically organized based on an *in silico* analysis. Motif prediction was conducted via several highly confident databases (TRANSFAC, JASPAR, hTFtarget, CIS-BP, HOCOMOCO), and all predicted TFBS were mapped with respect to the start codon (ATG) to allow for spatial characterisation of regulatory elements. Overall, an estimated 76 high-confidence TFBS were identified in the promoter region (Table 2), and the average density of TFBS (~16 sites per kb) indicates that the *PRL* promoter is highly compact and rich in regulatory elements. Additionally, the TFBS distribution in the *PRL* promoter is not evenly distributed; two regions of enrichment were identified: a distal regulatory region (−5000 to −3000 bp) and a mid-promoter region (−3000 to −1000 bp), both of which contain most of the predicted binding sites. In contrast, the proximal region of the promoter (−1000 to −1 bp) has a lower density of TFBS but does contain high concentrations of core promoter elements. This non-random pattern of organization indicates that the *PRL* promoter has a modular regulatory architecture, with the distal and mid-promoter regions acting as enhancer-like domains with multiple TFBS clustered together, whereas the proximal region is involved in supporting the initiation of transcription. Distribution of transcription factor binding sites (TFBS) by functional category revealed a non-uniform pattern. Hormone receptor-associated factors, including ESRRA, RARA, and NR4A1, constituted the largest proportion (31.6%). This was followed by POU1F1 (17.1%) and the STAT/JUN group (11.8%). SP1 accounted for 6.6%, and developmental transcription factors comprised 10.5% of the predicted TFBS. The remaining 22.4% of predicted TFBS were classified as “Other TFs”, representing transcription factors with diverse or unclassified biological functions (Figure 1).

Table 2 - Identified Transcription Factor Binding Sites (TFBs) within the *PRL* promoter region.

TF	Source	location in the promoter	Strand	Score	P value	Mached sequence
POU1F1*	hTFtarget	-107	-	16.2022	1.47E-07	TTCATGAATATAT
POU1F1	CISBP	-2521	+	9.77692	4.91E-06	CTAATTAGG
POU1F1	hTFtarget	-393	-	13.427	8.96E-06	TTCTTGAATATAG
POU1F1	hTFtarget	-1160	-	13.2809	1.06E-05	GTAATGAATATAA
POU1F1	CISBP	-208	-	8.61062	1.58E-05	CTTATTCATA
POU1F1	hTFtarget	-1866	+	12.7245	1.87E-05	TTGCATTATTTAG
POU1F1	CISBP	-2754	+	8.48673	3.03E-05	ATAATTCATT
POU1F1	hTFtarget	-1739	+	11.618	5.07E-05	CCATTGAATAAAC
POU1F1	hTFtarget	-4042	-	7.29592	5.89E-05	TTATGTACATTAAC
POU1F1	hTFtarget	-2832	-	7.0102	6.37E-05	ATATAGTAATTAGA
POU1F1	CISBP	-265	-	8.82308	6.64E-05	TTAATTAGA
POU1F1	hTFtarget	-2615	+	10.6458	7.73E-05	TTTCTGCATCATTAAGG
POU1F1	hTFtarget	-4225	-	11.1011	7.74E-05	GAAATGAAAATTT
NR4A1	hTFtarget	-1225	+	15.0	2.21E-06	TAAAGGTCAC
NR4A1	hTFtarget	-736	-	12.7	2.13E-05	TTAAATGTCAAG
NR4A1	hTFtarget	-2165	-	12.4364	2.56E-05	TAAAATGTCATA
NR4A1	HOCOMOCO	-2339	+	10.824	3.21E-05	AAAATGGCA
NR4A1	hTFtarget	-365	+	10.9817	4.86E-05	AGGCAAATGTTA
NR4A1	HOCOMOCO	-1991	+	10.488	5.84E-05	GAAATGGCA
NR4A1	hTFtarget	-1572	+	10.6301	7.44E-05	AAATGGCAC
NR4A1	HOCOMOCO	-1182	+	10.232	8.96E-05	AGAAAGTCA
ESRRA	hTFtarget	-1226	+	15.5	3.45E-06	CTAAAGGTCACAGGC
ESRRA	hTFtarget	-926	-	14.6	6.86E-06	TTCAAGGGCACACAC
ESRRA	hTFtarget	-1512	-	13.3939	1.17E-05	GGGTCAAAATGCCAT
ESRRA	hTFtarget	-1081	-	12.187	2.42E-05	GGCCAAGGTGACA
ESRRA	hTFtarget	-1917	+	7.87755	3.57E-05	GGGGTCGCCCAGAGTCA
ESRRA	hTFtarget	-2979	+	10.2653	6.82E-05	TCAGGTGAAGGGGACCTGG
ESRRA	hTFtarget	-843	+	5.06122	9.76E-05	GAGGGCAACGTATGTTG
RARA	hTFtarget	-2415	-	13.3776	1.94E-06	AGGTCAAATAACGTTCA
RARA	TRANSFAC	-1228	+	15.6053	2.95E-06	AACTAAAGGTCACAGG
RARA	hTFtarget	-1066	+	13.5606	1.03E-05	AGAGTGGCTTGAGGTCAGAG
RARA	hTFtarget	-2068	+	13.1712	1.30E-05	CAGTCCATGGGGTCAC
RARA	CISBP	-3159	-	5.82653	1.57E-05	AATTTCAAAAGATCA

RARA	hTFtarget	-1511	-	6.44898	2.71E-05	TGGGTCAAAATGCCA
RARA	hTFtarget	-1918	+	5.52041	2.73E-05	TGGGGTCGCCCAGAGTCA
RARA	hTFtarget	-737	-	11.8333	3.14E-05	GGGGAGTTTAAATGTCAAGG
RARA	hTFtarget	-241	+	5.47959	6.56E-05	GATGTTCATTTCTGGTCA
Sp1	hTFtarget	-720	+	11.6966	4.29E-05	CCCATCCCAC
Sp1	TRANSFAC	-134	+	11.0843	5.46E-05	GCAAAAGGGAAGGG
Sp1	TRANSFAC	-814	-	10.6386	7.28E-05	TGAAGAGGGCTGGA
Sp1	hTFtarget	-1609	-	10.6923	7.47E-05	TCCCCTCCTTT
Sp1	hTFtarget	-3334	-	9.94231	9.95E-05	CTCCCACCCTT
STAT5A	hTFtarget	-3290	+	14.4242	1.15E-05	TTCTAGGAAC
STAT5A	HOCOMOCO	-503	-	9.84848	9.69E-05	TACTAAGAAA
STAT6	HOCOMOCO	-790	-	13.973	7.86E-06	CTTCAAGAGAAATC
JUN	hTFtarget	-4794	-	11.4343	4.87E-05	CATTCCATTGTGGTGAGAGA
JUND	hTFtarget	-524	+	11.1327	5.25E-05	CTTACTCATCCTT
JUND	hTFtarget	-3485	+	11.0	5.73E-05	GCTGATTCATG
JUNB	hTFtarget	-420	-	10.4184	7.69E-05	GGTGTCTCATG
JUNB	hTFtarget	-1818	-	10.6364	7.74E-05	GCTGATTCACA
JUND	hTFtarget	-2063	+	7.93878	9.48E-05	CATGGGGTCACTAG
Blimp-1	TRANSFAC	-2026	-	19.303	9.36E-08	AAAAGGGAAAGTGA
Irx6	TRANSFAC	-677	+	15.8204	8.12E-07	TACATACATGTTAAAAT
Hoxd9	TRANSFAC	-1301	+	13.4306	1.64E-06	GCAATAAAAA
Zfp445	TRANSFAC	-550	+	10.2787	2.87E-06	GAGTATAACAGGAACTGA
Hic1	TRANSFAC	-3233	-	12.3103	3.00E-06	TTGCCCACCAATGCCATG
Blimp-1	TRANSFAC	-1548	-	15.3333	3.52E-06	GACAGTGAAAGTCT
Foxj3	TRANSFAC	-4009	-	14.8621	3.70E-06	ATAAACAAAAGCA
Lhx4	TRANSFAC	-2525	+	14.1503	3.73E-06	ATTTCTAATTAGGGATG
Evx2	TRANSFAC	-1304	-	14.597	3.89E-06	TTTTATTGCATT
Irx2	TRANSFAC	-677	+	14.3133	4.38E-06	TACATACATGTTAAAAT
Zfp90	TRANSFAC	-891	-	15.4754	4.96E-06	AGGGAGTTCATA
Foxj3	TRANSFAC	-3056	+	14.0862	5.44E-06	CAAAACAAAAACA
Sall3	TRANSFAC	-1024	+	14.0161	5.98E-06	TAACAGTTT
Nanog	TRANSFAC	-3915	+	12,8358	1,47E-05	TATCAAAAACAATACAGGCA
Nanog	TRANSFAC	-3060	+	13.6866	5.99E-06	ATAGCAAAACAAAAACAAAT
BACH2	hTFtarget	-1817	-	14.3818	6.61E-06	TGCTGATTCAC
Blimp-1	TRANSFAC	-3004	-	14.0788	7.34E-06	CACTTTCCC
Blimp-1	TRANSFAC	-2025	+	14.0788	7.34E-06	CACTTTCCC
Alx1	TRANSFAC	-2522	-	12.3152	7.98E-06	CCTAATTAGA
Nanog	TRANSFAC	-2565	+	13.3881	8.31E-06	ATAGATGAACAAAAGAAAAT
Zfp37	TRANSFAC	-1676	-	12.7869	8.35E-06	ATTTAAAGGTGC
Arid5a	TRANSFAC	-4207	-	13.1786	9.34E-06	GCAATATTGGAAGC
Blimp-1	TRANSFAC	-909	-	13.0642	1.11E-05	GGGGAAGTGACAGTGG
Foxj3	TRANSFAC	-3732	+	12.5862	1.15E-05	ACAAACAAAAGCA
Blimp-1	TRANSFAC	-3870	+	12.8624	1.20E-05	CTAAATGGGAAAGGAA

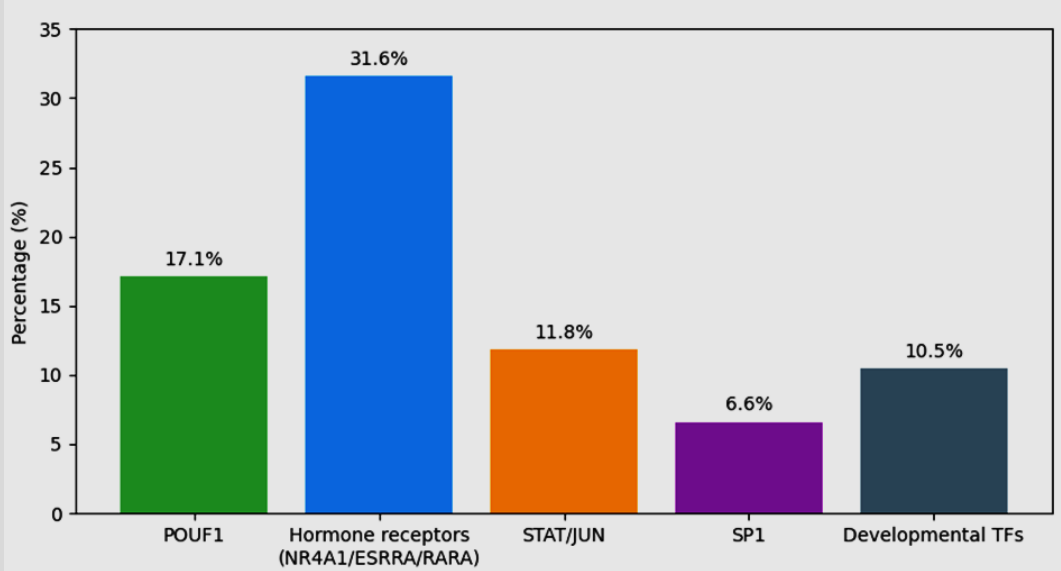

Figure 1 - Distribution of transcription factor binding sites (TFBS) by functional category. Bar chart showing the percentage of TFBS associated with different transcription factor families. Values indicate the relative contribution of each functional category to the total TFBS.

Hormone-responsive regulatory elements

One of the notable features of the *PRL* promoter was the presence of numerous predicted hormone-responsive elements, particularly those associated with nuclear receptors. Multiple putative binding sites for estrogen-related receptor alpha (ESRRA) were identified across a substantial portion of the *PRL* promoter (~−3084 to −841 bp). These predicted motifs exhibited relatively high scores (up to 15.5) and low P-values (as low as 3.45×10^{-6}), suggesting sequence similarity to known ESRRA binding motifs. The occurrence of predicted ESRRA motifs in the mid and proximal promoter regions may indicate potential regulatory regions; however, their functional significance and direct involvement in *PRL* transcription require experimental validation. In addition, several putative binding sites for retinoic acid receptor alpha (RARA) were identified throughout the promoter (~−2586 to −841 bp). A cluster of predicted RARA motifs was observed near −2586 bp (score = 13.38; $P = 1.94 \times 10^{-6}$). The co-occurrence of predicted ESRRA and RARA motifs may suggest potential interactions between regulatory pathways, although no functional relationship can be inferred from computational predictions alone.

Pituitary-specific and prolactin-regulatory factors

Multiple predicted POU1F1 (Pit-1) binding motifs were identified at several positions throughout the *PRL* promoter, spanning approximately −4894 to −776 bp. These putative motifs represent the largest group of predicted TFBS identified in the promoter region. The distribution of predicted POU1F1 motifs was not uniform, with apparent clusters observed in the distal (−4800 to −3800 bp) and mid-promoter (−2400 to −2200 bp) regions. The occurrence of these motif clusters may suggest potential cis-regulatory regions; however, their functional significance cannot be inferred from sequence analysis alone and requires experimental validation. In addition, several predicted NR4A1 binding motifs were identified within the −4600 to −2600 bp region, some of which were located near predicted POU1F1 motifs. The spatial proximity of these predicted motifs may indicate potential regulatory interactions; however, no functional relationship or coordinated regulatory network can be concluded based solely on computational predictions.

General transcription factors and signal-responsive elements

Numerous general transcription factor binding sites have been found in the *PRL* promoter, including SP1, STAT5A, STAT5B and JUN. Many of the SP1 binding sites occur predominately within a proximal promoter (approximately ~−720 to −130 bp) and correspond with the canonical function of SP1 in regulating regions of core promoters to promote basal levels of transcription. The presence of multiple SP1 motifs in close proximity suggests a strong basal transcriptional framework necessary for *PRL* expression. The presence of STAT5A and STAT5B binding sites at both distal (~−3290 bp) and proximal (~−503 bp) positions strongly supports the established role of STAT5 in prolactin signaling and cytokine-induced responses. Therefore, these binding sites play a role in the feedback mechanism regulating transcriptional activity and signal-dependent transcriptional activation. STAT5 binding sites are located throughout the promoter, indicating that STAT5 may regulate transcription through both direct binding as well as through long-distance interactions. JUN binding sites, which are part of the AP-1 complex, were identified in both distal (~−4794 bp) and proximal (~−524 bp) areas. The pattern of binding sites therefore indicates that *PRL* transcriptional activity may be regulated by external signals, including stress and inflammation and growth factors, and demonstrates that this gene is involved in broader physiological processes.

Developmental and pluripotency-associated factors

Various developmental and lineage-associated transcription factors (e.g., NANOG, FOXI3, IRX2, HOXD9) could putatively bind to the distal promoter region (~−4300 to −1700 bp), in addition to the classical prolactin regulators. Though these transcription factors are not traditionally linked to prolactin gene regulation, they provide evidence that the *PRL* promoter may contain evolutionarily-conserved regulatory elements that may be functional during certain developmental or physiological stages. Additionally, these motifs may also reflect latent regulatory potential that can be activated in non-canonical frameworks such as during early embryonic development or cellular reprogramming.

High-affinity and clustered binding regions

One of the main features of the *PRL* promoter was the presence of TFBS clusters located in two major regions (−4500 to −3000 bp and −2500 to −700 bp). Several predicted motifs for POU1F1, ESRRA, RARA, SP1, and NR4A1 were found in close proximity within these regions, suggesting potential cis-regulatory modules. The highest-scoring predicted PRDM1 (Blimp-1) binding motif was identified at approximately −2026 bp (score = 19.303, $P = 9.36 \times 10^{-8}$). This high-scoring motif may represent a potential regulatory element; however, its functional significance remains to be experimentally validated. The proximity of predicted activator (e.g., POU1F1) and repressor (e.g., PRDM1) motifs may indicate potential regulatory complexity, although no conclusions regarding promoter dynamics or regulatory interactions can be drawn from computational predictions alone.

DISCUSSION

Promoter analysis of the *Bos taurus PRL* gene predicted the presence of three putative transcription start sites (TSSs). The occurrence of multiple putative TSSs is consistent with the multi-TSS promoter architecture reported for many tissue-specific and developmentally regulated genes; however, the predicted TSSs identified in the present study require experimental validation or support from public transcriptomic datasets before being considered active transcription initiation sites. The proximal putative TSS located at −517 bp received the highest prediction score and may correspond to a potential core-promoter region responsible for basal transcriptional activity. In contrast, the more distal putative TSSs may represent alternative regulatory regions that could contribute to context-dependent regulation of *PRL* expression. The multi-TSS organization has been associated with dynamic transcriptional regulation and alternative promoter usage in other endocrine genes (Carninci et al., 2006; Haberle and Stark, 2018).

Similarly, the presence of multiple putative TSSs in the bovine *PRL* promoter may suggest the possibility of alternative 5′ UTR isoforms; however, this hypothesis remains to be experimentally verified. Collectively, these observations suggest that the bovine *PRL* promoter may possess a multi-layered regulatory architecture involving multiple putative transcription initiation regions and diverse cis-regulatory elements. Using the classical criteria of CpG island length, GC content, and observed/expected CpG ratio (Takai and Jones, 2002; Long et al., 2013), this study did not identify a canonical CpG island within the analyzed *PRL* promoter sequence. Previous studies have shown that canonical CpG islands are commonly associated with housekeeping genes and permissive chromatin environments (Deaton and Bird, 2011), whereas tissue-specific promoters frequently lack classical CpG island features (Saxonov et al., 2006). Therefore, the absence of a canonical CpG island in the bovine *PRL* promoter may suggest that epigenetic regulation of this gene could involve mechanisms other than broad CpG island-associated methylation. However, no direct evidence regarding DNA methylation or chromatin organization was generated in the present study, and the potential contribution of local CpG methylation and chromatin modifications to *PRL* regulation remains speculative and requires further investigation. *In silico* analysis of the ~5 kb *Bos taurus PRL* promoter suggested a potentially complex regulatory architecture containing numerous predicted transcription factor binding motifs. Prolactin is a key regulator of mammary gland development, metabolism, immune function, and reproduction in cattle (Bole-Feysot et al., 1998; Zhou et al., 2025). The high density of predicted transcription factor binding sites (TFBS) observed in the present study may indicate the presence of multiple regulatory regions capable of integrating diverse physiological signals. However, the functional significance of these predicted motifs cannot be inferred from sequence analysis alone and requires experimental validation. Previous experimental studies have demonstrated that POU1F1 (Pit-1) plays an essential role in lactotroph differentiation and prolactin production (Li et al., 1990; Sönmez and Ünal, 2023).

In the present study, multiple predicted POU1F1 binding motifs were identified throughout the bovine *PRL* promoter, both proximally and distally. The apparent clustering of these predicted motifs may suggest potential cis-regulatory regions; however, their occupancy and regulatory function *in vivo* remain unknown. Similarly, although cooperative interactions among adjacent POU1F1 binding sites have been reported in other experimental systems (Dasen and Rosenfeld, 2001; Kawauchi et al., 2025), the present study did not evaluate transcription factor binding or promoter activity. Therefore, any relationship between the predicted POU1F1 motifs and *PRL* transcription should be regarded as a hypothesis requiring future experimental validation. Single nucleotide polymorphisms (SNPs) within regulatory regions have been reported to alter the binding affinity of transcription factors, including POU1F1, and may influence gene expression in certain biological contexts (Lü et al., 2010). Because prolactin plays an important role in mammary gland development and lactation, variation within predicted POU1F1 binding motifs could potentially affect *PRL* regulation. However, the present study did not evaluate genetic polymorphisms, gene expression, or production traits, and therefore no conclusions can be drawn regarding the functional effects of these predicted motifs on milk production or lactational performance. Future studies integrating polymorphism analyses, gene expression data, and phenotypic measurements will be necessary to determine the biological significance of these predicted regulatory elements. The presence of multiple predicted POU1F1 motifs in the *PRL* promoter may indicate potential regulatory regions; however, their functional roles remain to be experimentally validated.

The presence of multiple motifs known to have high-affinity for estrogen receptor α (ESR1 and ESRRA), as well as retinoic acid receptor-α (RARA), in the *PRL* promoter for underscores the fundamental importance of hormone action in the expression of prolactin gene. Prolactin is known to be synthesized in response to the action of estrogen, especially during pregnancy and early lactation, when increasing levels of estradiol sensitize both the lactotrophs and the mammary epithelium to become maximally prepared for their functions (Freeman et al., 2000). There is evidence that, at the regulatory regions of the *PRL* gene, ESR1-binding is synergistic with Pit-1, thereby increasing the activity of the *PRL* promoter in the presence of estrogen. Furthermore, there is considerable evidence to support the conclusion that retinoic acid signaling via the activation of RAR/RXR heterodimers plays an important role in cell fate decision making for mammary epithelial cells and their differentiation into functional lactocytes (Hannan et al., 2023). The position of the putative binding sites for ESRRA and RARA in relation to the POU1F1 motifs suggests that hormones may also alter the basal transcriptional program mediated by pituitary-specific factors. In addition to being hormone responsive, the *PRL*

promoter contains putative binding sites for the transcription factors STAT5A/B and AP-1 (JUN/FOS), which may increase the responsiveness of the *PRL* promoter to extra-cellular signals such as cytokines, growth factors, and metabolic stress.

STAT5 is activated by the signal from the prolactin receptor and is an established downstream effector of this signaling pathway in the mammary gland, resulting in the activation of milk protein genes and enhancing the lactogenic pathways via a positive feedback mechanism (Liu et al., 1997; Rädler et al., 2017; Zhou et al., 2025). The presence of motifs that can bind to STAT5 at the *PRL* locus indicates that autoregulatory loops may be present at this locus and that prolactin itself could modulate its own expression in an indirect manner using receptor-mediated signaling pathways. In addition, stress signals, the availability of nutrients and growth factor signaling pathways can all modulate the activity of the AP-1 complex (also known as JUN, FOS, and all related family members) indicating that transcriptional activity of prolactin can potentially be affected by metabolic and environmental stressors (Shaulian and Karin, 2002).

Several predicted motifs corresponding to developmental transcription factors, including NANOG, IRX2, FOXI3, and HOXD9, were identified in the distal region of the bovine *PRL* promoter. However, these transcription factors are not currently known to participate in *PRL* regulation, and their predicted occurrence may reflect the limitations of motif-scanning approaches and the potential for false-positive predictions. Although members of the HOX and FOX transcription factor families have been implicated in developmental and tissue-specific regulatory processes in mammals (Golson and Kaestner, 2016; Hubert and Wellik 2023), there is currently no direct evidence supporting a role for NANOG, IRX2, FOXI3, or HOXD9 in the regulation of bovine *PRL* expression. Therefore, the biological significance of these predicted motifs remains uncertain and requires further investigation through analyses of sequence conservation, tissue-specific expression, and experimental validation in bovine pituitary and mammary tissues.

The occurrence of transcription factor-binding sites (TFBS) near each other, especially those that contain POU1F1, ESRRA, RARA, SP1 and STAT motifs, provides evidence for the presence of cis-regulatory modules (CRM) that serve as integrative hubs for grouping together different transcription factors (TFs). CRMs are being increasingly recognized as playing an essential role for context-dependent gene regulation, providing a means for multiple TFs to combine their efforts to influence levels of gene transcription (Spitz and Furlong, 2012). In the bovine *PRL* promoter, these CRMs may enable dynamic switching between activation and repression states in response to developmental stage, endocrine milieu, or physiological stress Additionally, the finding of a high affinity PRDM1 (also known as Blimp-1) binding site suggests that there are repressive mechanisms in place in addition to activating platforms that provide another layer of regulation to prevent excessive *PRL* production (Minnich et al., 2016). These types of repressive modules are important to maintaining normal cellular homeostasis and preventing abnormal hormone production. Collectively, the architecture of the *PRL* promoter provides evidence that there is a multi-layered regulatory system that enables prolactin expression to be dynamically adjusted throughout pregnancy, parturition, and lactation. This ensures that mammary gland development, the production of milk and maintenance of overall metabolic homeostasis of the whole organism are coordinated in response to both intrinsic cues and external environmental conditions (Bauman and Currie, 1980). The distal clusters and CRMs that were found in this *in-silico* study may be good candidates to target for genetic selection or epigenetic modulation to enhance the efficiency of lactation and the resilience of dairy cattle. Functional validation of these elements through reporter assays, CRISPR gene editing, and chromatin accessibility profiling will be essential for elucidating their roles *in vivo*.

CONCLUSION

The present study provides an *in silico* characterization of the bovine *PRL* promoter and identifies several putative regulatory elements, including predicted transcription start sites and transcription factor binding motifs. These findings should be regarded as hypotheses for future research and require experimental validation to determine their functional significance in bovine *PRL* gene regulation.

DECLARATIONS

Corresponding author

Correspondence and requests for materials should be addressed to Boburbek Akhmadaliev; E-mail: ahmadaliyev_bobur@mail.ru ORCID: 0009-0007-9938-150X

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Authors' contribution

Author's contribution should be explained in the title page as in the example. Initials of the authors must be in compliance with the following templates:

Conceptualization, study design, supervision, and critical revision of the manuscript: Bakhtiyor ADILOV;

Draft manuscript preparation, integration of all sections, and manuscript formatting: Boburbek AKHMADALIEV;
Data curation, genomic sequence retrieval from genomic databases, and transcription start site (TSS) analysis: Anvar SHERIMBETOV;
Transcription factor binding site (TFBS) prediction, bioinformatics analysis, and data visualization: Bobur ABDIKARIMOV;
CpG island analysis, statistical evaluation, and interpretation of epigenetic features: Bobir ABDUVALIEV;
Results interpretation, literature review, and drafting of the Discussion section: Dilshodbek RUZMETOV.

Funding sources
The authors gratefully acknowledge the financial support provided by the Academy of Sciences of the Republic of Uzbekistan under the budget-funded research program entitled "Genotyping of Genes Responsible for Milk Productivity and Quality in Cattle Breeds".

Ethical consideration
This study was conducted using publicly available nucleotide sequence data and bioinformatic analyses. Since no live animals, human participants, or biological samples were involved, ethical approval and informed consent were not required.

Competing interests
The authors have not declared any competing interests.

Funding
The authors declare that no funds, grants, or other support were received during the preparation or publication of this manuscript.

REFERENCES

Bauman DE and Currie WB (1980). Partitioning of nutrients during pregnancy and lactation: A review of mechanisms involving homeostasis and homeorhesis. Journal of Dairy Science, 63(9): 1514–1529. https://doi.org/10.3168/jds.S0022-0302(80)83111-0

Bole-Feysot C, Goffin V, Edery M, Binart N and Kelly PA (1998). Prolactin (PRL) and its receptor: Actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. Endocrine Reviews, 19(3): 225–268. https://doi.org/10.1210/edrv.19.3.0334

Brym P, Malewski T, Starzyński R, Flisikowski K, Wójcik E, Ruść A, Zwierzchowski L and Kamiński S (2007). Effect of new SNP within bovine prolactin gene enhancer region on expression in the pituitary gland. Biochemical Genetics, 45(9–10): 743–754. https://doi.org/10.1007/s10528-007-9115-9

Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J, et al. (2006). Genome-wide analysis of mammalian promoter architecture and evolution. Nature Genetics, 38(6): 626–635. https://doi.org/10.1038/ng1789

Dasen JS and Rosenfeld MG (2001). Signaling and transcriptional mechanisms in pituitary development. Annual Review of Neuroscience, 24: 327–355. https://doi.org/10.1146/annurev.neuro.24.1.327

Deaton AM and Bird A (2011). CpG islands and the regulation of transcription. Genes and Development, 25(10): 1010–1022. https://doi.org/10.1101/gad.2037511

Fornes O, Castro-Mondragon JA, Khan A, van der Lee R, Zhang X, Richmond PA, et al. (2020). JASPAR 2020: update of the open-access database of transcription factor binding profiles. Nucleic Acids Research, 48(D1): D87-D92. https://doi.org/10.1093/nar/gkz1001

Freeman ME, Kanyicska B, Lerant A and Nagy G (2000). Prolactin: Structure, function, and regulation of secretion. Physiological Reviews, 80(4): 1523–1631. https://doi.org/10.1152/physrev.2000.80.4.1523

Golson ML and Kaestner KH (2016). Fox transcription factors: From development to disease. Development, 143(24): 4558–4570. https://doi.org/10.1242/dev.112672

Haberle V and Stark A (2018). Eukaryotic core promoters and the functional basis of transcription initiation. Nature Reviews Molecular Cell Biology, 19(10): 621–637. https://doi.org/10.1038/s41580-018-0028-8

Hannan FM, Elajnaf T, Vandenberg LN, Kennedy SH and Thakker RV (2023). Hormonal regulation of mammary gland development and lactation. Nature Reviews Endocrinology, 19(1): 46–61. https://doi.org/10.1038/s41574-022-00742-y

Hubert KA and Wellik DM (2023). Hox genes in development and beyond. Development, 150(1): dev192476. https://doi.org/10.1242/dev.192476

Ilie DE, Mizeranschi AE, Mihali CV, Neamț RI, Cziszter LT, Carabaș M and Grădinaru AC (2023). Polymorphism of the prolactin (PRL) gene and its effect on milk production traits in Romanian cattle breeds. Veterinary Sciences, 10(4): 275. https://doi.org/10.3390/vetsci10040275

Kawauchi Y, Sasaki S, Matsushita A, Misawa Nakamura H, Yamashita M, Kakizawa K, Ohba K, Tsuriya D, Tanaka T and Suda T (2025). Functional properties of POU1F1 mutants in the transcriptional regulation of the thyrotropin β gene compared with the prolactin gene. International Journal of Molecular Sciences, 27(1): 119. https://doi.org/10.3390/ijms27010119

Kulakovskiy IV, Vorontsov IE, Yevshin IS, Sharipov RN, Fedorova AD, Rumynskiy EI, Medvedeva YA, Magana-Mora A, et al. (2018). HOCOMOCO: Towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. Nucleic Acids Research, 46(D1): D252–D259. https://doi.org/10.1093/nar/gkx1106

Li S, Crenshaw EB, Rawson EJ, Simmons DM, Swanson LW and Rosenfeld MG (1990). Dwarf locus mutants lacking three pituitary cell types result from mutations in the POU-domain gene Pit-1. Nature, 347(6293): 528–533. https://doi.org/10.1038/347528a0

Liu X, Robinson GW, Wagner KU, Garrett L, Wynshaw-Boris A and Hennighausen L (1997). Stat5a is mandatory for adult mammary gland development and lactogenesis. Genes and Development, 11(2): 179–186. https://doi.org/10.1101/gad.11.2.179

Long HK, Blackledge NP and Klose RJ (2013). ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection. Biochemical Society Transactions, 41(3): 727–740. https://doi.org/10.1042/BST20130028

Lü A, Hu X, Chen H, Jiang J, Zhang C, Xu H and Gao X (2010). Single nucleotide polymorphisms in bovine PRL gene and their associations with milk production traits in Chinese Holsteins. Molecular Biology Reports, 37(1): 547–551. https://doi.org/10.1007/s11033-009-9762-5

Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chekmenev D, et al. (2006). TRANSFAC and its module TRANSCompel: Transcriptional gene regulation in eukaryotes. Nucleic Acids Research, 34(Database issue): D108–D110. https://doi.org/10.1093/nar/gkj143

Minnich M, Tagoh H, Bönelt P, Axelsson E, Fischer M, Cebolla B, Tarakhovsky A, Nutt SL, et al. (2016). Multifunctional role of the transcription factor Blimp-1 in coordinating plasma cell differentiation. Nature Immunology, 17(3): 331–343. https://doi.org/10.1038/ni.3349

Motmain Z, Özdemir M, Ekinci K, Saygili E and Bilgin E (2022). A meta-analysis of the associations between prolactin (PRL) gene polymorphism and milk production traits in cattle. Kafkas Üniversitesi Veteriner Fakültesi Dergisi, 28(5): 627–631. https://doi.org/10.9775/kvfd.2022.27857

Rädler PD, Wehde BL and Wagner KU (2017). Crosstalk between STAT5 activation and PI3K/AKT functions in normal and transformed mammary epithelial cells. Molecular and Cellular Endocrinology, 451: 31–39. https://doi.org/10.1016/j.mce.2017.04.025

Reese MG (2001). Application of a time-delay neural network to promoter annotation in the Drosophila melanogaster genome. Computers & Chemistry, 26(1): 51–56. https://doi.org/10.1016/S0097-8485(01)00099-7

Saxonov S, Berg P and Brutlag DL (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proceedings of the National Academy of Sciences of the United States of America, 103(5): 1412–1417. https://doi.org/10.1073/pnas.0510310103

Shaulian E and Karin M (2002). AP-1 as a regulator of cell life and death. Nature Cell Biology, 4(5):E131–E136. https://doi.org/10.1038/ncb0502-e131

Shen WK, Chen SY, Gan ZQ, Zhang YZ, Yue T, Chen MM, et al. (2023). AnimalTFDB 4.0: A comprehensive animal transcription factor database updated with variation and expression annotations. Nucleic Acids Research, 51(D1): D39–D45. https://doi.org/10.1093/nar/gkac907

Sönmez Z and Ünal H (2023). The relationship of POU1F1-HinfI gene polymorphisms on milk yields in Simmental cattle. Yuzuncu Yil University Journal of Agricultural Sciences, 33(4): 513–521. https://doi.org/10.29133/yyutbd.1271873

Spitz F and Furlong EEM (2012). Transcription factors: From enhancer binding to developmental control. Nature Reviews Genetics, 13(9): 613–626. https://doi.org/10.1038/nrg3207

Stormo GD (2013). Modeling the specificity of protein-DNA interactions. Quantitative Biology, 1(2):115–130. https://doi.org/10.1007/s40484-013-0012-4

Takai D and Jones PA (2002). Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proceedings of the National Academy of Sciences of the United States of America, 99(6): 3740–3745. https://doi.org/10.1073/pnas.052410099

Zhou X, Ullah A, Shi L, Dou M, Wang C, Khan MZ, Wang C and Zhang X (2025). Molecular regulatory mechanisms of mammary gland development: A review. Animals, 15(23): 3480. https://doi.org/10.3390/ani15233480