Acta Scientiarum
http://periodicos.uem.br/ojs/acta
ISSN on-line: 1807-8672
Doi: 10.4025/actascianimsci.v47i1.70103

Single nucleotide polymorphism in *Osteopontin* gene and its association with milk traits in Azikheli buffalo

Arshia Rauf[1], Sohail Ahmad[1], Sundus Altaf[1], Shehryar Khattak[2], Bibi Sabiha[1,3] and Muhammad Ibrahim[1*]

[1]Institute of Biotechnology and Genetic Engineering, Faculty of Crop Production Sciences, The University of Agriculture, 25000, Peshawar, Khyber-Pakhtunkhwa, Pakistan. [2]Pakistan Red Crescent Society Bannu Branch, 28101, Bannu, Khyber-Pakhtunkhwa, Pakistan. [3]Institute of Pathology and Diagnostic Medicine, Khyber Medical University, Peshawar, Khyber-Pakhtunkhwa, Pakistan. *Author for Correspondence. E-mail: ibrahimfaqir@aup.edu.pk

ABSTRACT. *Osteopontin* (*OPN*) is known to effect milk composition traits. This study aimed to associate *OPN* gene polymorphism with milk traits in Azikheli buffaloes. Data were collected for milk yield and milk composition from 30 buffaloes. DNA samples of these specimen were used to amplify exon 4, intron 4 and exon 6 of the *OPN* gene using predesigned primers. The PCR products were sequenced through Sanger sequencing. The results showed that the milk yield varied significantly ($p < 0.001$) among Azikheli buffaloes. Sanger sequencing revealed 24 SNPs in the targeted regions of *OPN*, among which 2 were found in the high yielding buffaloes, while 23 were in the low yielding buffaloes of which one SNP was shared. One novel SNP g.5096T>C in the intron 5 of the *OPN* gene showed significant association with milk yield and milk protein. non-synonymous substitutions were observed at different loci i.e., g.5521C>T (Asp108Glu), g.5505C>T (Ala128Val), g.5446T>A (Thr149Ala), and 5395CGA>DEL (Asp92Del). Among the non-synonymous mutations only Ala128Val was found to have effect on protein stability (DDG = – 0.92 kcal mol^{-1}) due to its presence in the conserved region of the protein. In conclusion, our results suggest SNP g.5096T>C as a potential genetic marker for high milk yield in Azikheli buffalo.

Keywords: protein stability; genetic markers; in-silico analysis; gene sequencing; genetic association.

Received on October 25, 2023.
Accepted on November 14, 2024.

Introduction

Buffalo milk is considered healthier than cow milk as it has higher calcium and protein levels, has more saturated fatty acids and low cholesterol levels (Medhammar et al., 2012). There are three buffalo breeds well recognized in Pakistan namely Kundi, Nili-Ravi and Azikheli. Azikheli is found in the Swat region of Pakistan (Khan et al., 2019). Azikheli have exceptional ability to produce high quality milk and meat from low-grade or even insufficient feed given to them (Nadeem et al., 2013). Quantitative Trait Loci have been detected in chromosome 6 of bovine which harbors genes like *ABCG2* (ATP-binding cassette super-family G member 2), *IFG1* (Insulin-like growth factor 1), *POU2F1* (POU Class 2 Homeobox 1) and *OPN* which are responsible for milk composition traits (Abdel-Shafy et al., 2020; Moravčikova et al., 2019).

OPN or secreted phosphoprotein 1 (*SPP1*) is a glycoprotein that is positioned on chromosome 7 in buffalo which contains 6-8 exons. Its length is 1383-1935 bp and has 280 Amino acids (http://www.ensembl.org). It is extremely phosphorylated and expressed in numerous cells and tissues including blood (Frank et al., 2021), milk (Suchit, 2018), seminal plasma (Bustamante-Filho et al., 2021), and macrophages (Murthy et al., 2022) etc. OPN becomes functional when it binds to several cell surface receptors including multiple integrins and CD44 (Messex et al., 2022; Wang et al., 2020) which activates many cellular pathways involved in processes like cell-mediated immune responses (Aasmul-Olsen et al., 2021), bone mineralization and hematopoiesis (Al-Bari & Al Mamun, 2020), cancer metastasis (Khongsti & Das, 2021), pregnancy (Wang et al., 2018), and neural development (Joung et al., 2020). The suppression of *OPN* adversely effects the expression of *CSN2* and *CSN3* (Sheehy et al., 2009) which are responsible for producing almost 80% of milk protein (Farrell Jr et al., 2004).

Various studies have been done on the associations of *OPN* with various milk production traits like milk composition (Raza et al., 2023), milk yield (Dettori et al., 2020), and lactation persistency (Bissonnette, 2018). SNPs g.58899C>A and g.2916G>A have been detected in *OPN* gene of the Chinese Holstein cattle that showed significant association with milk composition traits (Raza et al., 2023). One SNP (OPN3907) showed high association with protein percentage of milk (Schnabel et al., 2005). Another SNP A>C detected in *OPN* gene had significant associations with protein and fat percentages in milk of North America Holstein population

(Leonard et al., 2005). In riverine buffaloes, an SNP (g.38329758 T>C) in *OPN* causing a substitution of valine to alanine (V127A) showed association with higher milk protein (Manzoor et al., 2018).

Selective breeding of individuals with superior phenotypes has been the main focus for the improvement of livestock (Ramón et al., 2021). But Little to almost no work is done on *OPN*'s effect on milk production trait associations in the local livestock breeds of Pakistan even though it is an important gene involved in many processes. Hence the purpose of our study was to investigate polymorphism in the targeted regions of OPN gene of high and low yielding Azikheli buffaloes and to determine association between gene polymorphisms and milk yield/composition traits.

Material and methods

Animal selection

Azikheli buffaloes from Government farm of Charbagh, Swat, Khyber Pakhtunkhwa were selected and categorized based on their average milk yield per day. A total of 30 Azikheli buffalo specimens were categorized into low yielding buffaloes (8 liters milk day^{-1}) and high yielding buffaloes (14 liters milk day^{-1}).

Milk samples and analysis

Milk samples were collected from the selected Azikheli buffaloes. The samples were immediately transported to the laboratory in ice box. Analysis of total protein content in milk was performed at Azikheli research station using Speedy Lab Auto Milk Analyzer (Astori Tecnica, Model 110V).

Blood sample collection

Following the standard biosafety protocol, blood samples were collected from 6 buffaloes (three high yielding and three low yielding) at their concurrent lactation stages. A total volume of 3ml of blood was drawn out using disposable syringes and stored in EDTA tubes immediately to avoid blood clotting. Then the tubes were brought to the laboratory in icebox for DNA extraction using GeneJET™ Genomic DNA Purification Kit (Thermo Fisher, Catalogue no. 01263698). The quality of the extracted genomic DNA was checked by running the DNA on 1% agarose gel for intactness, and using nanodrop for DNA purity and concentration.

Primer designing

Primers for the targeted regions of *OPN* gene were designed using NCBI Primer3 software, Insilco PCR, and BLAST tools. The finally working primers are given in Table 1. Two sets of primers were designed for the targeted regions of *OPN* gene including exon 4, intron 4 and exon 6 using the FASTA sequences obtained from NCBI (NC_059163.1).

Table 1. Primers designed to amplify different regions of OPN gene in Azikheli buffalo

Direction	Primer Sequence	Target	GC %	Tm	Product size
F	GCTTATCACTTAGAGACCCCTG	Exon 4 & intron 4	50	57	804
R	CTGGAGGGTTTGCTAGTGGA		55		
F	TCCACTAGCAAACCCTCCAG	Exon 6	55	59	1049
R	CACGCAGACTCTAGTTTCCTA		50		

Polymerase chain reaction

For the 15 µL final reaction volume of PCR, reagents added were 1.5 µL DNA sample, 5.2 µL water, 7.5 µL high fidelity Phusion® PCR master mix and 0.8 µL primers. They were added to the PCR tubes. The thermal profile for amplification is given in Table 2. Gel electrophoresis was performed for visualization of PCR products on 1% agarose gel. For visualization, a 1.5 µL of PCR product was loaded into the wells along with 1 µL loading dye (catalogue no. R1161) and the reaction was run for 30 m at 90 volts. Lastly, the gel was placed under Gel Doc UV-Transilluminator (MGIS-21-C2-9M, company Taiwan) to visualize the bands.

Table 2. Thermal profile for PCR

Cycles	PCR steps	Temperature (°C)	Duration
1	Initial Denaturation	98	30 sec
	Denaturation	98	5 sec
30	Annealing	63	3 sec
	Extension	72	20 sec
1	Final Extension	72	3min

Sequencing

The Amplified products were further sent for Sanger sequencing to Beijing Genomic Institute (BGI) for identification of the polymorphisms in the targeted regions of *OPN* gene.

Data analysis

Means, standard errors, and ANOVA were calculated for different milk trait parameters among different buffalo categories using SPSS Software (IBM SPSS Statistics 23.0). The sequencing data were analyzed using BioEdit software (version 7.2.5). POPGENE 32 (version 1.3) was used for checking allelic frequencies. The associations analysis was performed using principal component analysis (PCA) test in SPSS software. For the exonic variants, in-silico analysis was performed using EXPASY (https://web.expasy.org/translate/) for translation, FALCON2 (http://falcon.ictbda.cn:89/serve/) for protein modeling, DYNAMUT (https://biosig.lab.uq.edu.au/dynamut2/) for protein stability and CONSURF (https://consurf.tau.ac.il/consurf_index.php) for finding the conserved regions in the protein.

Results

Milk yield and composition

The high yielding animals were reported to produce 13.7 ± 0.6 liters of milk per day, while the low yielding animals produced 8.0 liters of milk per day on average. The difference in milk yield between high yielding and low yielding animals was significantly different ($p < 0.001$). In high yielding animals the milk protein concentration was slightly higher (3.17 ± 0.23%) compared to the milk protein concentration of low yielding animals (2.88 ± 0.02%); however, the difference was not significant as shown in Figure 1.

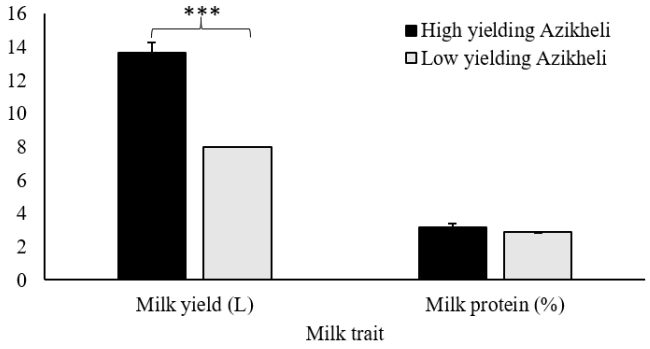

Figure 1. Milk traits comparison among the two categories of Azikheli buffalo.

The bars represent the mean value of milk yield in liters and milk protein content in percentage respectively. The milk yield shows significant difference between the two categories (*** = $p < 0.001$) and milk protein content shows no significant effect.

PCR Results

The agarose gel pictures of the PCR products of the targeted regions of *OPN* gene through primer 1 and primer 2 are shown in Figure 2. Primer 1 amplified the intron 4 while primer 2 amplified the regions of intron 5 and exon 6 region for the samples. To determine the product size, a 1 kb DNA ladder was used. Primer 1 yielded a product of 804 bp whereas primer 2 yielded a product of 1049 bp. The remaining PCR products were sent for sanger sequencing for the identification of *OPN* gene variants after confirmation of PCR on gel.

Total variants

BioEdit tool was used for analyzing the sequences received back from BGI. A total number of 22 SNPs and 2 deletions were identified in the targeted regions. High yielding buffaloes had only two SNPs in intron 5, while low yielding buffaloes showed a total of 23 variants including two deletions and 21 SNPs. In intron 4, 7 SNPs were observed in low yielding buffaloes. In Intron 5, a total of 10 variants were observed in which one was shared among high and low yielding buffaloes, one was a deletion and remaining variants were unique to low yielding buffaloes. In exon 6, one deletion and 7 SNPs were identified in low yielding buffalo. After determining total variants, most of SNPs were in low yielding buffaloes. The results showed that 22 out of 24 were unique to

the low yielding Azikheli buffalo whereas only one variant found on intron 5 was unique in the high yielding buffalo. Another variant g.4948G>A on intron 5 was shared among both types of Azikheli buffaloes.

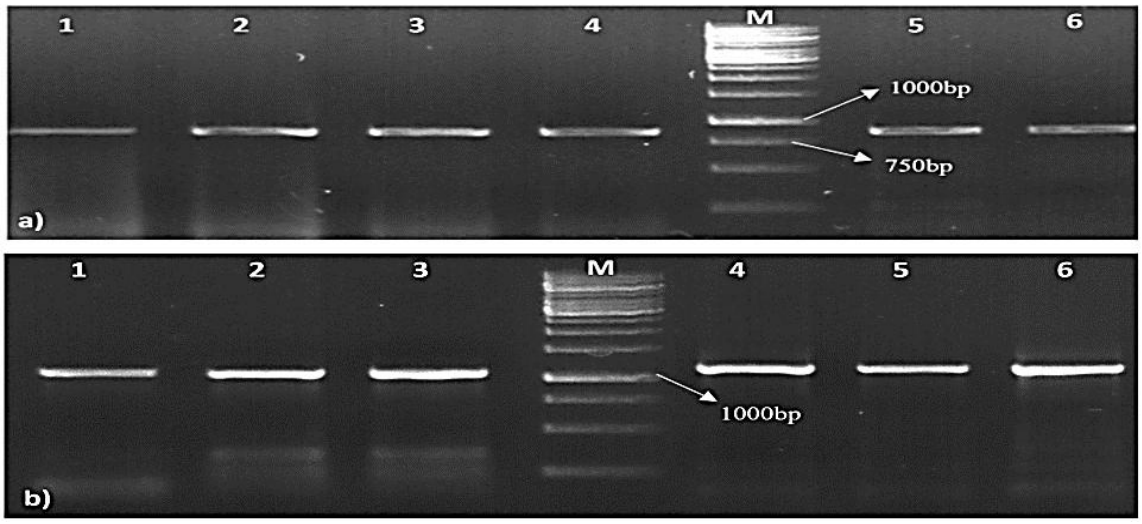

Figure 2. Gel pictures showing PCR amplification. (a) Primer 1 which yielded product size of 804bp, and (b) primer 2 yielding product size of 1049bp. Key: 1,2, and 3 = samples from low yielding buffalo; 4,5, and 6 = samples from high yielding buffalo; M = 1 kb DNA ladder.

Genetic diversity

No observed numbers of alleles were found in the intron 4 of the high yielding buffaloes, whereas in the low yielding buffaloes one allele was observed at each position of intron 4. The average effective number of alleles in the low yielding buffaloes was observed to be 0.94 which is lower than the average observed number of alleles. Only one variant in low yielding buffaloes g.4234T>C was within Hardy-Weinberg equilibrium (HWE), whereas other variants significantly deviated from HWE ($p > 0.05$). Similarly in the region of intron 5, in high yielding buffaloes, two alleles were observed at positions g.4948G>A and g.5096T>C, whereas different alleles for eight SNPs and one deletion were observed in low yielding category.

Table 3. Allelic frequency in different regions of *OPN* gene.

		Na		Ne		HWE P value	
Regions	Position	High Yielding	Low Yielding	High Yielding	Low Yielding	High Yielding	Low Yielding
Intron 4	g.4055T>C	0	1	0	1		0.045
	g.4148A>G	0	1	0	1		0.045
	g.4202G>A	0	1	0	1		0.045
	g.4226T>C	0	1	0	1		0.045
	g.4234T>C	0	1	0	0.6		1
	g.4235A>T	0	1	0	1		0.045
	g.4236T>C	0	1	0	1		0.045
Intron 5	g.4187A>G	0	1	0	1		0.045
	g.4833G>A	0	1	0	1		0.045
	g.4866A>T	0	1	0	1		0.045
	g.4890C>G	0	1	0	1		0.045
	g.4948G>A	1	0	0.8	0	0.02	
	g.4973A>G	0	1	0	1		0.045
	g.5032A>T	0	1	0	1		0.045
	g.5096T>C	1	1	0.8	1	0.02	0.045
	g.5101A>G	0	1	0	1		0.045
	g.5279DEL	0	1	0	1		0.045
Exon 6	5395ACGA>DEL	0	1	0	1		0.045
	g.5446T>A	0	1	0	1		0.045
	g.5455C>T	0	1	0	1		0.045
	g.5473T>C	0	1	0	1		0.045
	g.5505C>T	0	1	0	1		0.045
	g.5521C>T	0	1	0	1		0.045
	g.5567A>G	0	1	0	1		0.045

Key: Na= number of observed alleles, Ne=Number of effective alleles, HWE= Hardy Weinberg Equilibrium.

The average observed number of alleles for high yielding buffaloes was 0.2 which was lower than in low yielding buffaloes (0.9). The average effective number of alleles in high yielding buffaloes (0.16) was also lower than the value observed for the low yielding buffaloes. One variant g.5096T>C was showing shared frequency in both low and high yielding buffaloes with a high number of effective alleles in low yielding buffaloes compared to the high yielding. All of the variants deviated from HWE in high yielding buffaloes as well as in low yielding buffaloes ($p > 0.05$) except g.4234T>C. The genetic diversity was evaluated in the exon 6 of *OPN* gene, which resulted in 7 variants, including six SNPs and one deletion. All variants were identified in low yielding buffaloes. In high yielding buffaloes, no observed or effective number of alleles were found in the exon 6. In low yielding buffaloes, the observed and effective number of alleles were found to be 1 at each position. All the variants identified in the exon 6 of low yielding buffalo significantly deviated from HWE ($p > 0.05$) as shown in Table 3.

Heterozygosity

The observed heterozygosity in the region of intron 4 was observed only for one variant (g.4234T>C) in low yielding buffaloes as shown in Table 4, while the remaining variants showed no heterozygosity in this region. The average observed heterozygosity in intron 4 was lower than the expected heterozygosity. The observed heterozygosity at intron 5 in *OPN* gene was not identified in both categories of buffaloes whereas the expected heterozygosity was observed in the positions g.4948G>A and g.5096T>C of high yielding buffaloes with a value of 0.53.

Table 4. Heterozygosity observed in different regions of the *OPN* gene of Azikheli buffalo

		Observed He		Expected He		Average He	
Regions	Position	High Yielding	Low Yielding	High Yielding	Low Yielding	High Yielding	Low Yielding
Intron 4	g.4055T>C	0	0	0	0.666	0.25	0.25
	g.4148A>G	0	0	0	0.666	0.25	0.25
	g.4202G>A	0	0	0	0.666	0.25	
	g.4226T>C	0	0	0	0.666	0.25	0.25
	g.4234T>C	0	0.5	0	0.5	0.187	0.187
	g.4235A>T	0	0	0	0.666	0.25	0.25
	g.4236T>C	0	0	0	0.666	0.25	0.25
Intron 5	g.4187A>G	0	0	0	0.666	0.25	0.25
	g.4833G>A	0	0	0	0.666	0.25	0.25
	g.4866A>T	0	0	0	0.666	0.25	0.25
	g.4890C>G	0	0	0	0.666	0.25	0.25
	g.4948G>A	0	0	0.533	0	0.222	0.222
	g.4973A>G	0	0	0	0.666	0.25	0.25
	g.5032A>T	0	0	0	0.666	0.25	0.25
	g.5096T>C	0	0	0.53	0	0.22	0.22
	g.5101A>G	0	0	0	0.666	0.25	0.25
	g.5279DEL	0	0	0	0.666	0.25	0.25
Exon 6	5395ACGA>DEL	0	0	0	0.666	0.25	0.25
	g.5446T>A	0	0	0	0.666	0.25	0.25
	g.5455C>T	0	0	0	0.666	0.25	0.25
	g.5473T>C	0	0	0	0.666	0.25	0.25
	g.5505C>T	0	0	0	0.666	0.25	0.25
	g.5521C>T	0	0	0	0.666	0.25	0.25
	g.5567A>G	0	0	0	0.666	0.25	0.25

Key: He: Heterozygosity.

The expected heterozygosity in low yielding buffaloes at all positions were 0.66. The expected heterozygosities of both high and low yielding buffaloes were higher than their observed values. Seven variations were observed in the exon 6 of OPN including a deletion. All of the variants were observed in the low yielding buffaloes whereas no variations were observed in the high yielding buffaloes. All the variants observed in exon 6 of low yielding buffaloes were homozygous. The expected heterozygosity in the low yielding category had a value of 0.66 at all the positions.

Mutation frequency

Variation in the mutation frequency in the high yielding buffaloes was observed only at positions g.4948G>A and g.5096T>C at the region of intron 5 with values 0.667 and 0.333 respectively. The mutation

frequency at position g.4948G>A was also observed in low yielding buffaloes with a value of 0.8 which was highest among all the mutation frequency in other positions. The lowest mutation frequency was observed at position g.4234T>C in region of intron 4 in low yielding buffaloes with a value of 0.25.

Correlation matrix

Of the total 24 variants, only one SNP showed positive correlation with both milk yield and milk protein content with values of 0.4 and 0.2, respectively. The rest of the variants showed negative correlation with milk yield and protein content with correlation values between – 0.2 and – 0.6 as shown in Figure 3

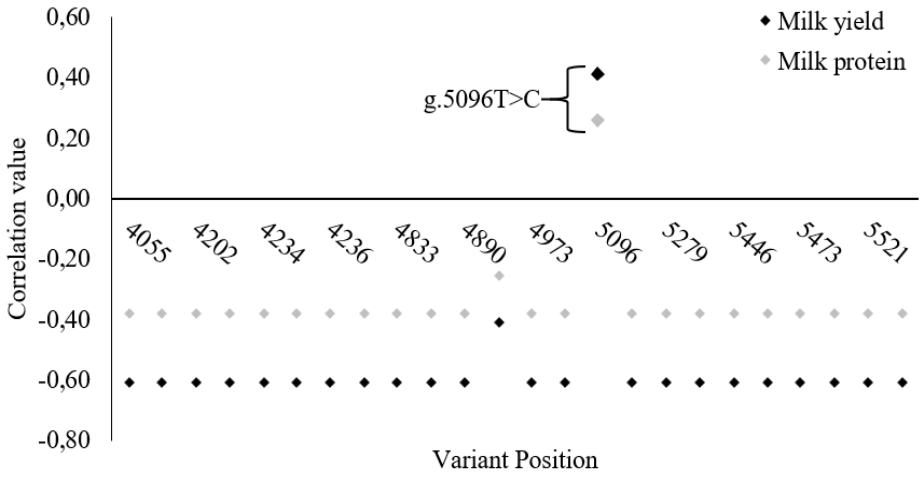

Figure 3. Correlation matrix of different SNPs with milk yield and milk protein content.

The value above 0 indicates positive correlation and below zero shows negative correlation.

In-silico Analysis

Translation of the exon variants to amino acid change

Firstly, the mutated region of exon 6 was used to translate into its corresponding amino acid sequence through Translate Expasy Tool. The results were compared with the original protein sequence of *OPN* (XP_006042334.3) obtained from NCBI. The alignment results identified one deletion (Asp92DEL), three synonymous mutations (no change in the amino acid) and three non-synonymous mutations changing the amino acids at different positions (Asp108Glu, Ala128Val, and Thr149Ala) as shown in. Table 5.

Table 5. Exonic variants and their corresponding amino acid change.

TYPE OF MUTATION	SNP POSITION	AA CHANGE
Deletion	5395CGA>DEL	Asp92DEL
	g.5455C>T	Asp111Asp
Synonymous	g.5473T>C	Asp117Asp
	g.5567A>G	Asp133Asp
	g.5521C>T	Asp108Glu
Non-Synonymous	g.5505C>T	Ala128Val
	g.5446T>A	Thr149Ala

Protein modeling

The wild type and mutant type OPN protein models were generated using Falcon2. The results were compared and a visible change in structure was observed as shown in Figure 4.

Arrows indicating change of amino acids at several positions in both wild and mutant type protein after superimpositions.

Both of the wild type and mutant type protein models were superimposed to see the effect of the mutations on the structure of the protein and it was clearly seen that the mutations in the mutant protein model had somewhat altered the original protein structure but all of the mutations were present on the loops of the protein so their presence did not damage the protein structure.

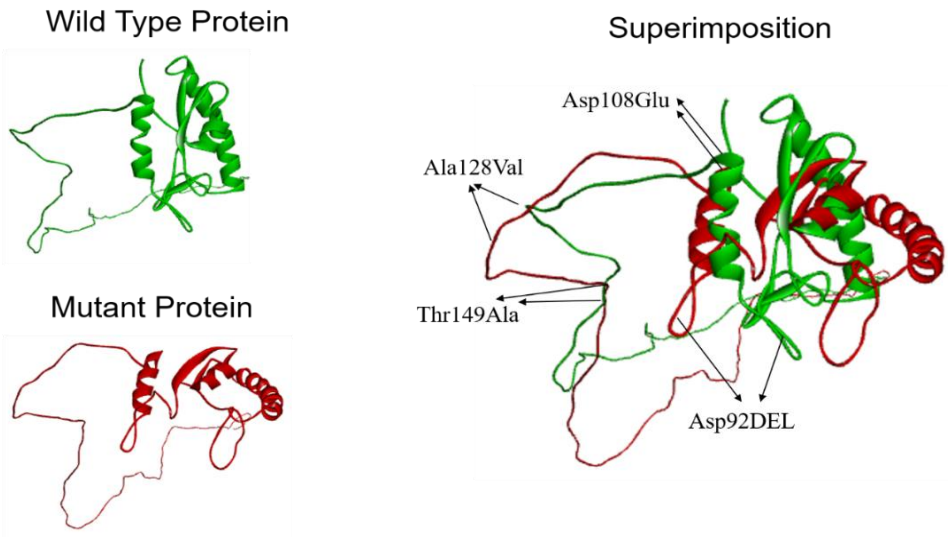

Figure 4. Structure of the Wild-type and mutant OPN protein and their superimposition.

Effect of non-synonymous mutations on protein structure

The Amino acid change in the mutated protein sequence and its effect on the protein structure was checked through missense3D software, individually. The substitution (Asp108Glu, Ala128Val, Thr149Ala) showed no damage to the protein structure as described in Table 6.

Table 6. The effect of the altered amino acid in protein sequence of OPN on its structure.

SNPs	Altered amino acids	Result
g.5521C>T	Asp108Glu	No Structural Damage detected
g.5505C>T	Ala128Val	No Structural Damage detected
g.5446T>A	Thr149Ala	No structural damage detected

Effect of mutations on protein stability

Protein Stability is critical for a protein to function properly which was determined individually using DynaMut2 as shown in Table 7. All of the non-synonymous mutations decreased the stability of the protein with Ala128Val having the highest energy of decreasing protein stability i.e., – 0.92 kcal mol^{-1}.

Table 7. Stability prediction of the OPN protein as a result of the altered amino acid sequence.

Amino Acid Position	Altered amino acid	Stability	Predicted Stability Change
g.5521C>T	Asp108Glu	Decrease	– 0.32 kcal mol^{-1}
g.5505C>T	Ala128Val	Decrease	– 0.92 kcal mol^{-1}
g.5446T>A	Thr149Ala	Decrease	– 0.64 kcal mol^{-1}

Effect of mutations on the conserved regions of the protein

The identified non-synonymous mutations were evaluated for their presence in the conserved regions of the protein using ConSurf tool. The ConSurf tool showed the conserved regions of a protein in a colorimetric scale from 1-9 (1 being most variable and 9 being most conserved). According to the scale, Asp108Glu was found to be reasonably conserved with a value of 5. Both Asp and Glu are negatively charged and polar amino acids, their substitution does not effect the net charge of the protein since both play similar role in protein active or binding sites. Similarly, the mutation Thr149Ala was found in the slighlty conserved region with a value of 6. Since theronine side chain is really reactive, it can form hydrogen bonds with polar substrates but Alanine has a very non-reactive side chain which is why it has rare direct involvment in protein functions, hence their substitution can have adverse effects on the protein. Whereas, Ala128Val was found in a highly conserved region with a value of 8 as shown in Figure 5. Valine has an unreactive side chain like alanine and is not involved in protein functions but plays a role in substrate recognition and involved in binding of hydrophobic ligands like lipids so their substitution can have lethal effect on the protein.

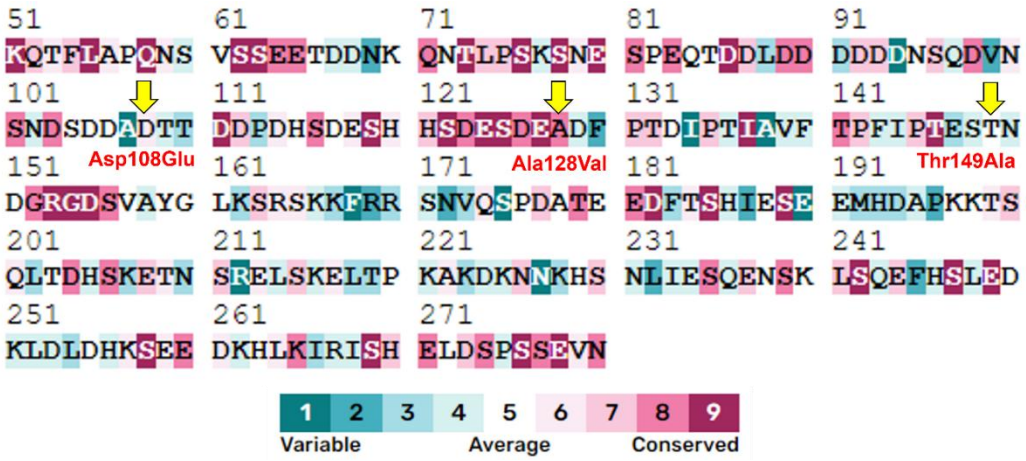

Figure 5. The effect of non-synonymous mutations on the conserved regions of OPN protein sequence.

Discussion

Milk production traits have been of economic importance in dairy industry (Ma et al., 2021) especially the buffalo milk because of its higher calcium and protein levels, more saturated fatty acids and low cholesterol levels which is more beneficial for human health (Medhammar et al., 2012). Since milk production is a phylogenetic trait, it is controlled by many genes, Osteopontin being one of them. It is a multifunctional protein that is involved not only in local immunity but mammary gland development and milk production traits as well (Dettori et al., 2020). Many researchers have studied *OPN* and found polymorphisms in several regions of the gene that show associations with milk producing traits like milk yield and milk protein content in different animal breeds like Polish Holstein Friesian (Kułaj et al., 2022), Iranian Holstein (Pasandideh et al., 2015) and Nili-Ravi (Manzoor et al., 2018) etc. In our study, we analyzed the OPN gene for SNPs in Azikheli buffalo to check its association with milk traits as data regarding milk production in Azikheli is sparsely available but it has a potential role in milk production.

In the current study, six Azikheli buffaloes were selected from government dairy farm of Charbagh, Swat, KP, Pakistan. They were categorized into high yielding buffaloes and low yielding buffaloes. High significant difference was recorded between the high yielding buffaloes (13.7 ± 0.6 liters/day) and low yielding buffaloes (8.0 liters/day). In previous studies, Nili-Ravi buffalo's average milk yield was recorded to be 7 liters/day (Khan et al., 2022). This variation in the milk yield can be due to many factors including environmental stresses as well as genetic capacity of an animal (Marumo et al., 2022). The average milk protein concentration of Azikheli buffalo was recorded to be 3.06 ± 0.23%. Where the milk protein concentration in high yielding buffaloes was slightly higher (3.17 ± 0.23%) compared to the milk protein concentration of low yielding buffaloes (2.88 ± 0.02%) representing no significant difference in current study which was contradictory to study conducted by others (Schnabel et al., 2005). Previous authors have reported higher protein content of about 3.2-3.5% in milk of Holstein Friesian which is higher than milk protein content of Azikheli buffalo (Curone et al., 2018; Kułaj et al., 2022). However, different factors can bring variation in protein levels in the milk of an animal e.g. health status, different stages of lactation, different breeds, and nutritional status of the animal (Heck et al., 2009).

The target regions of current study were exon 4, intron 4 and exon 6 of OPN gene in Azikheli buffalo where a total of 24 SNPs were identified. Previous literature reported 10 SNPs in riverine buffalo in both exonic as well as intronic regions (Manzoor et al., 2018) and 3 SNPs in Chinese Holstein cows each in intronic, exonic and 3'UTR regions of OPN gene (Raza et al., 2023). Many SNPs in OPN gene have been reported by researchers but the SNP c.8514C>T in intron 4 has been studied frequently as it shows significant association with milk protein content in cattle (Kułaj et al., 2019; Pasandideh et al., 2015). Some contradictory results have also been found by scientists (De Mello et al., 2012; Lali et al., 2020) showing no significant association of c.8514C>T with milk traits. The reason for the contrasting results may be due to epigenetic or effect of the environmental, breed or sampling difference (Heck et al., 2009).

The SNP c.8514C>T has also been associated with beef traits (White et al., 2007) as well as with growth traits in Holstein-Friesian (Pareek et al., 2008). Our target region included intron 4 but no such SNP was identified in this region, however, seven other unique SNPs were identified in in low yielding Azikheli buffaloes, all of them showing no significant association with the milk traits. Another novel SNP (g.5096T>C)

was identified in the region of intron 5 in the current study, that showed significant association with milk traits i.e., with milk yield (0.40) and milk protein (0.20) suggesting that it might be responsible for the higher milk yield in Azikheli buffalo. The SNP g.5096T>C could be used as a molecular marker for milk traits in Azikheli buffalo, however, it should be further evaluated using a larger population. In past decades, intronic SNPs were considered useless as scientists believed that introns lose their function during splicing but recently they have been gathering attention in the scientific world as polymorphisms in intronic region of a gene can result in a truncated protein by changing its stability and mRNA splicing sites (Manzoor et al., 2018).

The exonic mutations are the basis of genetic diversity and are able to shape the structure or expression level of protein (Manzoor et al., 2018). Exonic SNPs in OPN gene have also been associated with growth and production traits. A missense mutation on exon 7 causing Gln235Arg has been significantly associated with somatic cell count (Dettori et al., 2020). Similarly, a non-synonymous mutation from threonine to methionine at locus g.58675C > T has been associated with carcass weight in sheep (Matsumoto et al., 2019). In the current study, 3 non-synonymous mutations were identified in exon 6, meaning it altered the amino acid in the protein sequence i.e., Asp108Glu, Ala128Val, Thr149Ala. None of the mutations showed any effect on the stability of protein structure but had somewhat altered the protein structure and stability.

The missense mutation Ala128Val effected the protein structure the most because of its presence in the highly conserved region of OPN protein sequence. Such a substitution in conserved regions can lead to impaired or enhanced misfolding of a protein, poor stability, translocation problems and deprivation of function. A reported non-synonymous mutation (Ala128Val) showed association with milk protein (Tantia et al., 2008), whereas the same mutation (Ala128Val) in our study showed no association with milk traits. However, it needs to be further validated by protein studies as the mutation is present in a conserved region of OPN protein meaning its presence can cause altered or malfunctioned protein. Similarly, another study have shown significant association of SNPs in exon 6 with milk production traits (Raza et al., 2023). The other two non-synonymous mutations identified in the current study were novel and have not been reported in previous studies. All the non-synonymous mutations in the current study were unique to the low yielding Azikheli buffalo, suggesting their negative effect on milk yield.

Conclusion

Variation in milk yield and quality parameters among Azikheli buffaloes may be utilized to identify genetic basis for these traits. The novel SNP g.5096T>C in the intron 5 of OPN gene may have significant association with milk yield and protein content. The non-synonymous mutation g.5505C>T (Ala128Val) in the exon 6 of OPN gene affecting protein stability, may be responsible for low milk yield in buffaloes. The associated SNPs may further be validated through expression analysis and proteomics. These SNPs may be developed as potential selection markers for milk production and quality traits in buffaloes.

Acknowledgments

The Financial support for this study was provided by Higher Education Commission of Pakistan under the PSDP project titled Establishment of National Center for Livestock Breeding Genetics and Genomics (NCLBG&G), and the HEC-NRPU project No. 6306/KPK/NRPU//R&D/HEC/2016.

References

Aasmul-Olsen, K., Henriksen, N. L., Nguyen, D. N., Heckmann, A. B., Thymann, T., Sangild, P. T., & Bering, S. B. (2021) Milk osteopontin for gut, immunity and brain development in preterm pigs. *Nutrients*, *13*(8), 2675. https://doi.org/10.3390/nu13082675

Abdel-Shafy, H., Awad, M. A., El-Regalaty, H., Ismael, A., El-Assal, S. E.-D., & Abou-Bakr, S. (2020) A single-step genomic evaluation for milk production in Egyptian buffalo. *Livestock Science*, *234*, 103977. https://doi.org/10.1016/j.livsci.2020.103977

Al-Bari, A. A., & Al Mamun, A. (2020) Current advances in regulation of bone homeostasis. *FASEB BioAdvances*, *2*(11), 668-679. https://doi.org/10.1096/fba.2020-00058

Bissonnette, N. (2018) Short communication: Genetic association of variations in the osteopontin gene (SPP1) with lactation persistency in dairy cattle. *Journal of Dairy Science*, *101*(1), 456-461. https://doi.org/10.3168/jds.2017-13129

Bustamante-Filho, C. I., Menegassi, R. S., Pereira, R. G., Salton, D. G., Munari, M. F., Schneider, R. M., Mattos, R. C., Barcellos, J. O. J., Laurino, J. P., Cirne-Lima, E. O., & Jobim, M. I. M. (2021) Bovine seminal plasma osteopontin: Structural modelling, recombinant expression and its relationship with semen quality. *Andrologia*, *53*(1), e13905. https://doi.org/10.1111/and.13905

Curone, G., Filipe, J., Cremonesi, P., Trevisi, E., Amadori, M., Pollera, C., Bianca Castiglioni, B., Turin, L., Tedde, V., Vigo, D., Moroni, P., Minuti, A., Bronzo, V., Filippa Addis, M., & Riva, F. Vigo, D. (2018) What we have lost: Mastitis resistance in Holstein Friesians and in a local cattle breed. *Research in Veterinary Science*, *116*, 88-98. https://doi.org/10.1016/j.rvsc.2017.11.020

De Mello, F., Cobuci, J., Martins, M., Silva, M., & Neto, J. (2012) Association of the polymorphism g. 8514C> T in the osteopontin gene (SPP1) with milk yield in the dairy cattle breed Girolando. *Animal Genetics*, *43*(5), 647-648.

Dettori, M. L., Pazzola, M., Petretto, E., & Vacca, G. M. (2020) Association analysis between SPP1, POFUT1 and PRLR gene variation and milk yield, composition and coagulation traits in Sarda sheep. *Animals*, *10*(7), 1216. https://doi.org/10.3390/ani10071216

Farrell Jr, H., Jimenez-Flores, R., Bleck, G., Brown, E., Butler, J., Creamer, L., Hicks, C. L., Hollar, C. M., Ng-Kwai-Hang, K. F., & Swaisgood, H. (2004) Nomenclature of the proteins of cows' milk—Sixth revision. *Journal of Dairy Science*, *87*(6), 1641-1674. https://doi.org/10.3168/jds.S0022-0302(04)73319-6

Frank, J. W., Steinhauser, C. B., Wang, X., Burghardt, R. C., Bazer, F. W., & Johnson, G. A. (2021) Loss of ITGB3 in ovine conceptuses decreases conceptus expression of NOS3 and SPP1: implications for the developing placental vasculature. *Biology of Reproduction*, *104*(3), 657-668. https://doi.org/10.1093/biolre/ioaa212

Heck, J., Van Valenberg, H., Dijkstra, J., & Van Hooijdonk, A. (2009) Seasonal variation in the Dutch bovine raw milk composition. J*ournal of Dairy Science*, *92*(10), 4745-4755. https://doi.org/10.3168/jds.2009-2146

Joung, S., Fil, J. E., Heckmann, A. B., Kvistgaard, A. S., & Dilger, R. N. (2020) Early-life supplementation of bovine milk osteopontin supports neurodevelopment and influences exploratory behavior. *Nutrients*, *12*(8), 2206. https://doi.org/10.3390/nu12082206

Khan, H., Ali, S. M., & Yousafzai, A. A. (2019) *Azikheli Buffalo Breed of Swat-Khyber Pukhtun Khwa-Pakistan*. Engormix / Dairy Cattle. https://en.engormix.com/dairy-cattle/articles/azikheli-buffalo-breed-swat-t43992.htm

Khan, N., Khan, S., Israr, M., Hilal, M. G., Nawaz, M., Ibrahim, M., & Khan, N. A. (2022) Comparing Azikheli buffaloes with Nili Ravi buffaloes for morphometry, productivity and reproductivity. Pak*istan Journal of Agriculture Sciences*, *59*(4), 693-701. https://doi.org/10.21162/PAKJAS/22

Khongsti, K., & Das, B. (2021) Osteopontin and breast cancer metastasis: Possible role of genistein on the regulation of osteopontin. *Phytomedicine Plus*, *1*(4), 100138. https://doi.org/10.1016/j.phyplu.2021.100138

Kułaj, D., Pokorska, J., & Bauer, E. A. (2022) Associations between osteopontin gene polymorphism (OPN), milk yield, somatic cell score and age of cattle on milk antioxidant capacity. *International Dairy Journal*, *127,* 105295. https://doi.org/10.1016/j.idairyj.2021.105295

Kułaj, D., Pokorska, J., Ochrem, A., Dusza, M. & Makulska, J. (2019) Effects of the c. 8514C> T polymorphism in the osteopontin gene (OPN) on milk production, milk composition and disease susceptibility in Holstein-Friesian cattle. *Italian Journal of Animal Science*, *18*(1), 546-553. https://doi.org/10.1080/1828051X.2018.1547129

Lali, F. A., Anilkumar, K. & Aravindakshan, T. (2020) Effect of Osteopontin gene variants on milk production traits in Holstein Friesian crossbred cattle of Kerala. *Turkish Journal of Veterinary and Animal Sciences*, *44*(3), 695-701. https://doi.org/10.3906/vet-1812-4

Leonard, S., Khatib, H., Schutzkus, V., Chang, Y. & Maltecca, C. (2005) Effects of the osteopontin gene variants on milk production traits in dairy cattle. *Journal of Dairy Science*, *88*(11), 4083-4086. https://doi.org/10.3168/jds.S0022-0302(05)73092-7

Ma, Y., Khan, M. Z., Xiao, J., Alugongo, G. M., Chen, X., Chen, T., ... Shah, M. K. (2021) Genetic markers associated with milk production traits in dairy cattle. *Agriculture*, *11*(10), 1018. https://doi.org/10.3390/agriculture11101018

Manzoor, S., Nadeem, A., Maryam, J., Hashmi, A. S., Imran, M., & Babar, M. E. (2018) Osteopontin gene polymorphism association with milk traits and its expression analysis in milk of riverine buffalo. *Tropical Animal Health and Production*, *50*(2), 275-281. https://doi.org/10.1007/s11250-017-1426-1

Marumo, J., Lusseau, D., Speakman, J., Mackie, M., & Hambly, C. (2022) Influence of environmental factors and parity on milk yield dynamics in barn-housed dairy cattle. *Journal of Dairy Science*, *105*(2), 1225-1241. https://doi.org/10.3168/jds.2021-20698

Matsumoto, H., Kohara, R., Sugi, M., Usui, A., Oyama, K., Mannen, H., & Sasazaki, S. (2019) The non-synonymous mutation in bovine SPP1 gene influences carcass weight. *Heliyon*, *5*(12), e03006. https://doi.org/10.1016/j.heliyon.2019.e03006

Medhammar, E., Wijesinha-Bettoni, R., Stadlmayr, B., Nilsson, E., Charrondiere, U. R., & Burlingame, B. (2012) Composition of milk from minor dairy animals and buffalo breeds: a biodiversity perspective. *Journal of the Science of Food and Agriculture*, *92*(3), 445-474. https://doi.org/10.1002/jsfa.4690

Messex, J. K., Byrd, C. J., Thomas, M. U., & Liou, G.-Y. (2022) Macrophages cytokine spp1 increases growth of prostate intraepithelial neoplasia to promote prostate tumor progression. *International Journal Of Molecular Sciences*, *23*(8), 4247. https://doi.org/10.3390/ijms23084247

Moravčikova, N., Trakovicka, A., Kadlečik, O., & Kasarda, R. (2019). Genomic signatures of selection in cattle through variation of allele frequencies and linkage disequilibrium. *Journal of Central European Agriculture*, *20*(2), 576-580. https://doi.org/10.5513/JCEA01/20.2.2552

Murthy, S., Karkossa, I., Schmidt, C., Hoffmann, A., Hagemann, T., Rothe, K., Seifert, O., Ulf Anderegg, U., von Bergen, M., Schubert, K., & Rossol, M. (2022) Danger signal extracellular calcium initiates differentiation of monocytes into SPP1/osteopontin-producing macrophages. *Cell Death and Disease*, *13*(1), 1-15. https://doi.org/10.1038/s41419-022-04507-3

Nadeem, A., Majeed, R., Babar, M. E., Yaqub, T., Javed, M., Hussain, T., & Khosa, A. N. (2013) Genetic Variants of POU1F1 gene in Azakheli buffalo breed of Pakistan. *Buffalo Bulletin*, *32*(SI 2), 692-696.

Pareek, C., Zięba, M., Michno, J., Czarnik, U., & Zwierzchowski, L. (2008) Study of SNP C> T polymorphism within the candidate genes for dairy and beef traits in a panel. *Journal of Agrobiology*, *25*, 121-124.

Pasandideh, M., Mohammadabadi, M., Esmailizadeh, A., & Tarang, A. (2015) Association of bovine PPARGC1A and OPN genes with milk production and composition in Holstein cattle. *Czech Journal of Animal Science*, *60*(3), 97-104. https://doi.org/10.17221/8074-CJAS

Ramón, M., Carabaño, M. J., Díaz, C., Kapsona, V. V., Banos, G., & Sánchez-Molano, E. (2021) Breeding strategies for weather resilience in small ruminants in Atlantic and Mediterranean climates. *Frontiers in Genetics*, *12*, 1588.

Raza, S. H. A., Khan, R., Pant, S. D., Shah, M. A., Quan, G., Feng, L., Cheng, G., Gui, L.-S., & Zan, L. (2023) Genetic variation in the OPN gene affects milk composition in Chinese Holstein cows. *Animal Biotechnology*, *34*(4), 893-899. https://doi.org/10.1080/10495398.2021.2001343

Schnabel, R. D., Kim, J.-J., Ashwell, M. S., Sonstegard, T. S., Van Tassell, C. P., Connor, E. E., & Taylor, J. F. (2005) Fine-mapping milk production quantitative trait loci on BTA6: analysis of the bovine osteopontin gene. *Proceedings of the National Academy of Sciences*, *102*(19), 6896-6901. https://doi.org/10.1073/pnas.0502398102

Sheehy, P., Riley, L., Raadsma, H., Williamson, P., & Wynn, P. (2009) A functional genomics approach to evaluate candidate genes located in a QTL interval for milk production traits on BTA6. *Animal Genetics*, *40*(4), 492-498. https://doi.org/10.1111/j.1365-2052.2009.01862.x

Suchit, K. (2018) Role of osteopontin in bovines and their association with different traits: A review. *International Journal of Genetics*, *10*(3), 370-372. https://doi.org/10.9735/0975-2862.10.3.370-372

Tantia, M. S., Mishra, B., Kumar, S. B., Mishra, B., Kataria, R., Mukesh, M., & Vijh, R. (2008) Characterization of osteopontin gene of Bubalus bubalis. *Animal*, *2*(7), 987-990. https://doi.org/10.1017/S1751731108002073

Wang, J.-B., Zhang, Z., Li, J.-N., Yang, T., Du, S., Cao, R.-J., & Cui, S.-S. (2020) SPP1 promotes Schwann cell proliferation and survival through PKCα by binding with CD44 and αvβ3 after peripheral nerve injury. *Cell and Bioscience*, *10*(98), 1-14. https://doi.org/10.1186/s13578-020-00458-4

Wang, X.-B., Qi, Q.-R., Wu, K.-L. & Xie, Q.-Z. (2018) Role of osteopontin in decidualization and pregnancy success. Reproduction, *155*(5), 423-432. https://doi.org/10.1530/REP-17-0782

White, S., Casas, E., Allan, M., Keele, J., Snelling, W., Wheeler, T., Shackelford, S. D., & Koohmaraie, M.,Smith, T. (2007) Evaluation in beef cattle of six deoxyribonucleic acid markers developed for dairy traits reveals an osteopontin polymorphism associated with postweaning growth. *Journal of Animal Scienc*e, *85*(1), 1-10. https://doi.org/10.2527/jas.2006-314