Iraqi Journal of Veterinary Sciences

www.vetmedmosul.com

Novel polymorphisms in the Alpha-Lactalbumin (Lalba) gene and analysis of gene expression in dairy buffaloes (*Bubalus bubalis*) of the Amazon

C.L. Santos[1], E.C. Sousa[1], R.C. Guimaraes[1], J.V. Figueiredo[1], R.A. Silva[1], E.M. Barbosa[2], P.P. Santana[3] and E.S. Filho[1]

[1]Animal Health and Production Institute, Molecular Biology Laboratory, Federal Rural University of the Amazon, Belem, [2]Department of the Collegiate Degree in Education, Federal University of Amapa, Mazagao, Brazil, [3]Social-Environmental and Water Resources Institute, Federal Rural University of the Amazon, Belem, Brazil

Article information

Article history:
Received 18 January, 2025
Accepted 18 March, 2025
Published 01 April, 2025

Keywords:
Milk production
Molecular markers
Ruminants
Selection

Correspondence:
E.S. Filho
silva.filho@ufra.edu.br

Abstract

This research aimed to find new SNPs in the Lalba gene to associate them with milk production and analyze their expression profile. A total of 104 females from a property located in Bujaru/PA, were used. DNA extractions were performed using the phenolic method, followed by PCR, purification, and sequencing to investigate SNPs in the promoter region and 5' UTR. RNA was extracted from milk epithelial cells to determine the expression profile of Lalba in 7 haplotypes selected for average milk production. Ten new SNPs were found in the studied region, with emphasis on the indel (-555>TAAA) observed in 7% of the population. None of the SNPs showed a significant association with average milk production (ML). Seven SNPs showed wild-type nucleotide allele frequencies greater than 0.5, with 90% of them detected as HWE. The SNP -42 (A>G) showed a strong binding site for the transcription factor HNF3. Haplotypes 4 and 2 were significant when compared to the others and had higher expressions. Although no SNP influenced milk production, haplotypes 4 and 2 showed the opposite, as they presented higher levels of mRNA expression in somatic milk cells, which is probably due to the additive effect of the combined alleles and, therefore, are desirable to be fixed in the population. However, studies focusing on the relationship between SNPs and expression levels in these cells, together with other genes that may be influencing milk production levels, should be carried out to investigate whether these variants are confirmed by the phenotypes of the animals.

DOI: 10.33899/ijvs.2025.156753.4092,

Introduction

Domestic buffaloes (*Bubalus bubalis*) are of Asian origin and are known for their docility, longevity, hardiness, and adaptability in different regions of the world with distinct soil and climate characteristics. Thus, they are found inhabiting areas where cattle would be unable to be raised (flooded areas). In addition, they have high feed conversion rates, generating products of high nutritional quality through a diet with low dietary attributes (1,2). Therefore, raising dairy buffaloes is considered an agricultural activity of great financial importance, capable of providing milk with excellent nutritional properties. Initially, buffalo farming in the country was focused on beef, but from the 1980s and 1990s onwards, the scenario began to change, mainly due to the enthusiasm of producers in investigating milk production, a fact that was noticed by the growth of industrial units dedicated to the production of buffalo milk derivatives (3). This commitment was also due to the composition that buffalo milk presents compared to that of other ruminants, mainly about its composition and characteristics, especially when it comes to the levels of protein, fat, lactose, mineral residue (calcium and phosphorus), and total solids since it allows high industrial yield in the manufacture of

derivatives, which can exceed 40-50% of the industrial yield when compared to bovine milk (4,5). In addition, buffalo milk is more digestible for people with lactose intolerance (6). However, even though the species displays its full potential for the activity, its production remains behind that of bovine milk. Because of this, the search for improving buffalo milk production has emerged among livestock farmers and researchers in the field, shown mainly in recent years by the growth of studies aimed at detecting variations at the level of deoxyribonucleic acid (DNA) in the buffalo genome that is capable of explaining the differences in their milk production. Among the technologies used in the search for candidate genes, chips with a high density of SNP markers and different GWAS methodologies have revealed the most genes for milk production in buffaloes (7-9). One of these is the alpha-lactalbumin gene (Lalba). It encodes whey protein (α-LA), which is essential for the synthesis of lactose in the mammary gland, which in turn is an important osmotic factor present in milk, playing a critical role in regulating the volume produced. It is also classified as an acute phase protein (APP). Since the concentration of α-LA during infections was reduced, it was then classified as APP negative (10). In a study with Polish Holstein cows, it was discovered that the SNP g.-1001T>C in the promoter region of the gene influenced milk production characteristics. Among them, high daily milk and dry matter production, in addition to high lactose production and concentration, are associated with the TT genotype ($P \leq 0.05$) (11). In Bhadawari buffaloes, the AB and BC genotypes were found to be significantly associated ($P \leq 0.05$) with total milk production and daily milk production (12). In Nilli Ravi buffaloes, 5 SNPs were detected; however, only Lalba2 at position 34310940 was associated with milk protein ($P \leq 0.05$) (5).

Although few studies have demonstrated the presence of SNPs in the Lalba gene of buffaloes, it is considered polymorphic. Because of this, researchers are increasingly using sequencing tools in different populations of the same species to prove the existence of these genes associated with zootechnical indices for milk production, which are responsible for ensuring an improvement in herd productivity. Therefore, the present research aimed to investigate and detect new possible SNPs in the Lalba gene to associate them with zootechnical data on milk production in a herd of crossbred buffaloes from the Amazon region.

Materials and methods

Ethical aspects

The study was approved by the Animal Use Ethics Committee (CEUA) of the Federal Rural University of the Amazon (2820150222/2022/CEUA/UFRA).

Animals

A total of 104 crossbred Murrah and Mediterranean buffaloes aged between 1 and 5 years weighing between 450 and 500 kg were used. All were healthy and free of mastitis and belonged to a property located in the municipality of Bujaru, Belém, Pará, Brazil (1°37'40''S, 48°13'21''W). Five ml of blood were collected from these animals in tubes containing anticoagulant (EDTA).

For the evaluation of gene expression, 42 buffaloes with average milk production ranging from 2 to 10 L/day were selected. From these, 10 mL of milk were collected in sterile 15 mL Falcon tubes, free of DNase, RNase, and non-pyrogenic, with 2 mL of RNAlatter added to stabilize and protect the RNA in the samples. Both samples were stored in a thermal box at a maximum temperature of 4°C and taken to the Serology and Molecular Biology Laboratory of the Federal Rural University of the Amazon-UFRA for further analysis.

DNA Extraction and Polymerase Chain Reaction (PCR)

It was performed using the phenol-chloroform method (13) and stored at 20 °C. The investigation of genetic variation based on SNPs was performed through PCR of the Lalba gene (ID: 102410146) in its promoter region and 5' UTR. The following primer pair F 5' ACC TTT CCC AAA AAG GTT GG 3' and R 5' ACC ACA CTG CTC ACC TCC TT 3' used in the research was designed through Primer3plus based on the reference sequence deposited in Genbank, amplifies a fragment of 833 bp. The reactions were performed for a final volume of 25 µL, being 2 µL of DNA, 12.5 µL Taq DNA Polymerase 2x Master Mix (Ampliqon, Odense, Denmark), 1.0 µL of each primer - forward and reverse - (10 pmol/µL) and 8.5 µL of ultra-pure water. All reactions were performed in a Thermal Cycler 2720 thermocycler (Applied Biosystems, Foster City, CA, USA), with the following amplification protocol: initial denaturation at 95 ºC for 5 min, followed by 35 cycles with denaturation at 95 ºC for 30 s, annealing at 55 ºC for 30 s, extension at 72 ºC for 50 s and final extension at 72 ºC for 5 min. The PCR products were analyzed using 0.8% agarose gel.

DNA Purification and Sequencing

The PCR products from 83 samples were purified using the PCR Product Purification Kit on a Column (Ludwig Biotec LTDA, Alvorada, RS, BR) and visualized on 0.8% agarose gels. These purified products were sequenced using the Sanger method using the BigDye Terminator v3.1 Cycle Sequencing Kit (Invitrogen California/USA) in a final volume of 10 µL in an ABI 3500 Genetic Analyzer automatic sequencer (Applied Biosystems). The sequences obtained were edited using the Finch TV software version 1.4.0 (Geospiza Research Team, USA), compared with the reference sequence for buffaloes entered in GenBank in the BLAST system, and finally aligned in BioEdit 7.2 using the ClustalW tool. Transcription factor domains in the promoter region were identified using Nsite v. 5 software (14).

RNA Extraction

For RNA isolation, milk samples were processed to separate somatic cells. Then, 1000 µL of milk from 15 mL Falcon tubes added with RNAlatter were centrifuged at 12,000 rpm at 4°C for 8 minutes. The supernatant was carefully removed with a spatula. The remaining material was transferred to sterile 2.0 mL Eppendorf tubes, where 700 µL of Trizol® (Invitrogen, USA) was added. The extraction protocol was followed according to the manufacturer's recommendations. The samples were quantified in a BioDrop µLITE spectrophotometer (BioDrop) using the A260nm/A280nm ratio.

Gene expression (RT-qPCR)

The gene expression profile was investigated from the 42 buffaloes in exon 2 of the Lalba gene. For this, the primer pairs described in table 1 were designed using Primer3Plus based on the reference sequences deposited in GenBank. The RT-qPCR reactions were performed in triplicate for Lalba and the endogenous Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene. The amplification reaction protocol was performed for a final volume of 10 µL according to the recommendations of the manufacturer of the RealQ Plus and RealQ Fast kit (Ampliqon, Odense, Denmark). Relative expression values were determined by the equation 2-ΔCt, where ΔCt is the ratio between the Ct of the target gene and the Ct of the endogenous gene (15).

Statistical analyses

Genetic diversity data such as allele and genotypic frequencies, Hardy Weinberg Equilibrium (HWE), and F_{IS} were obtained using PopGene (16). Expression patterns were analyzed using SAS OnDemand, assuming a normal distribution. Mean milk production data were analyzed using simple descriptive statistics. The Tukey test was used to assess the association with polymorphism data and the gene expression profile of haplotypes, with a significance level of 0.05.

Table 1: Primer pairs designed in Primer3plus according to the reference genes deposited in GenBank (ID) for gene expression analysis in dairy buffaloes

Genes	Primers (5' - 3')	Exon	ID Genbank
LALBA	F: CCA GTG GTT ATG ACA CAC AAGC R: GAG GGT TCT GGT CGT CTT TG	2	102410146
GAPDH	F: ACC CAG AAG ACG GTG GATG R: CCG TTG AGC TCA GGG ATGA	7	102404028

Results

The alignment performed on all the sequences obtained revealed the presence of 10 SNPs, seven of which were transitional and two of which were transversional. In addition, the presence of an insertion block with the nucleotide sequence TAAA between positions -552 and -555 was also detected in 92.8% of the samples (Table 2).

Table 2: Summary of single nucleotide alterations (SNPs) in the LALBA gene in buffaloes from the Bujaru-Pa region compared to Bubalus bubalis (GenBank ID: 102410146)

PCR fragment mutation	Genomic position based on reference	*Bubalus bubalis*	Bujaru-Pa buffaloes	Variation
-42 A>G	89119342	A	G	Transition
-60 T>C	89119324	T	C	Transition
-292 G> A	89119092	G	A	Transition
-339 C>T	89119045	C	T	Transition
-444 T>G	89118940	T	G	Transversion
-445 C>T	89118941	C	T	Transition
-501 T >G	89118883	T	G	Transversion
-555>TAAA	89118828- 89118847	-	TAAA	InDel
-591 C>T	89118797	C	T	Transition
-720 C>T	89118668	C	T	Transition

UTR = untranslated region; bp: base pairs.

The analysis of population genetic diversity detected that only eight SNPs presented the frequency of the wild-type allele greater than 0.5, equivalent to 90% of the total. These alleles of the SNPs in positions -339 (97.6%), -60 (92.5%), and -720 (91.5%) were those that exhibited high frequencies in the population that exceeded 90%. Regarding the frequency of the mutant allele, SNPs -42, -444, -501, and -591 were the exceptions, as they were the only ones that exhibited the frequency of the mutant allele (G) greater than 0.5, being 68.1%, 95.1%, 95.1%, and 97.5%, respectively.

Furthermore, the insertion mutation was detected in 77 buffaloes evaluated between positions -555 to -552, which contributed 4 new nucleotides (TAAA) to the population, also at a high frequency of 96.3%. Six of the 83 buffaloes are heterozygous for this SNP, that is, characterized as InDel, and no buffalo in the present study presented the deletion, as verified in the reference sequence for the species deposited in the GenBank database. All values for the effective number of alleles in the population for all SNPs evaluated were above 1.

The SNPs -42, -292, -445, and -720 presented lower observed heterozygosity rates (Ho) than expected (He), according to the HWE proportions, which is evidenced by the positive F_{IS} values, suggesting an excess of homozygotes, with the inbreeding process occurring for these. The only SNPs that exhibited the presence of homozygotes for the mutant allele were -42, -292, -445, and -720. On the other hand, for loci -60, -339, -444, -501, -591 and the insertion block together, the Ho rates were higher than those of He, data confirmed by negative F_{IS} values, thus suggesting an excess of heterozygotes and possible selection of them in the population. Locus -591 presented the lowest value of 0.037. This analysis of the SNPs found suggests the presence of low diversity. However, loci -42, -292, and -445 presented the highest values of 0.626, 0.601, and 0.601, respectively, which may suggest the presence of greater diversity for these loci within the population. Regarding the proportions of HWE, of all the SNPs found in the present study, nine are under equilibrium conditions (P>0.05), which are in the following positions -42 (A>G), -60 (T>C), -292 (G>A), -339 (C>T), -444 (T>G), -445 (C>T), -501 (T>G), -555 (TAAA) and -591 (C>T). The SNP in position -720 (C>T) is deviating from the HWE (P<0.05). All the information mentioned above is described in table 3.

From the analysis of the binding to transcription factors performed by the software above, it was possible to see that of the ten SNPs found, only one, at position -42 (TCATAAATA), has a strong binding site for the transcription factor hepatocyte nuclear factor 3 (HNF3) in its promoter region. However, the same analysis also found five other transcription factors, namely, Myocyte enhancer factor 2 (MEF2), Specificity protein 1 and specificity protein 3 (SP1; SP3), Sterol regulatory element-binding protein (SREBP), and Zinc finger proteins and Myc-associated zinc finger protein (MAZ/Sp1). However, these are not associated with any of the nine SNPs but are binding sites for the regulatory region of the Lalba gene and are, therefore, important to mention.

Table 3: SNP positions and respective descriptions of genetic diversity in the promoter region and 5' UTR of the Lalba gene in dairy buffaloes

SNP	Genotype			Alleles		Ho	He	Ne*	I*	F_{IS}	HWE
-42 A>G	AA 0.13(11)	AG 0.37(31)	GG 0.49(41)	A (0.319)	G (0.681)	0.373	0.437	1.768	0.626	0.140	0.180
-60 T>C	TT 0.90(75)	TC 0.09 (8)	CC 0 (0)	T (0.952)	C (0.048)	0.096	0.092	1.101	0.193	-0.050	0.667
-292 G>A	GG 0.53(44)	GA 0.36(30)	AA 0.11 (9)	G (0.711)	A (0.289)	0.361	0.413	1.698	0.601	0.120	0.246
-339 C>T	CC 0.95(79)	CT 0.04 (4)	TT 0 (0)	C (0.976)	T (0,024)	0.048	0.047	1.049	0.113	-0.024	0.845
-444 T>G	GG 0.90(75)	TG 0.09 (8)	TT 0 (0)	T (0.048)	G (0.951)	0.096	0.092	1.101	0.193	-0.050	0.667
-445 C>T	CC 0.53(44)	CT 0.36(30)	TT 0.11 (9)	C (0.711)	T (0.289)	0.361	0.413	1.698	0.601	0.120	0.246
-501 T>G	TT 0 (0)	TG 0.09 (8)	GG 0.90(75)	T (0.048)	G (0.951)	0.096	0.092	1.101	0.193	-0.050	0.667
-555>TAAA	IN 0,92(77)	INDEL 0,07 (6)	DEL 0 (0)	IN (0.963)	DEL (0.036)	0.072	0.070	1.074	0.155	-0.037	0.755
-591 C>T	CC 0 (0)	CT 0.04 (4)	TT 0.95(79)	C (0.024)	T (0.975)	0.048	0.047	1.049	0.113	-0.024	0.845
-720 C>T	CC 0.89(74)	CT 0.04 (2)	TT 0.08 (7)	C (0.936)	T (0.096)	0.024	0.155	1.182	0.289	0.844	0.000

Ho: Observed heterozygosity, He: Expected heterozygosity, Ne*: Effective number of alleles, I*: Shannon diversity index, F_{IS}: Inbreeding coefficient, HWE: Hardy Weinberg equilibrium (P>0.05).

The association of the SNPs found with the average milk production (L/day) of the buffaloes evaluated is described in table 4. The genotypic comparisons of all the SNPs found in the study with the average milk production did not show any statistical difference (P>0.05). However, even though it was not possible to observe a significant effect in the association,

it was possible to notice that some SNPs stood out because they were responsible for increasing milk production in certain buffaloes.

Therefore, for SNPs at positions -60, -339, -501, -720 and for Indel (-555), it was possible to observe that all heterozygotes of the same exhibited greater milk production compared to their wild homozygotes. Highlighting the heterozygote (CT) of SNP -720, which presented the highest milk production observed among all other SNPs. In addition, when compared to the production exhibited by its wild homozygote (CC), there was an increase in production of more than 2 liters of milk for the heterozygote, being 7.44 L/day; the homozygote (TT) also followed this increase, as it exhibited 7.01 L/day. Based on these findings, we can infer that the presence of mutant alleles in these SNPs may be contributing positively to the increase in milk production, having the so-called additive effect.

This deduction can be better observed and explained in SNPs -444, -501, and -591, for which it was not possible to find any buffalo belonging to the respective wild homozygous genotypes. However, the mutant homozygotes all demonstrated an increase of more than 0.5 liters in milk production compared to their heterozygotes. SNPs -42, -292, and -445 were the only ones that presented the three genotypes in the population, with the buffaloes of the wild homozygous genotypes exhibiting higher milk production.

Table 4: Association of novel polymorphisms detected in the regulatory region of the Lalba gene with buffaloes' average milk production (L/day)

SNPs	Milk production (L/day) (means±SD) Genotypes			Probability
-42	AA 5.70±1.77	AG 5.55±1.81	GG 5.4±1.83	0.983
-60	TT 5.46±1.80	TC 5.84±1.73	CC 0	0.623
-292	GG 5.80±1.71	GA 5.09±1.74	AA 5.39±2.28	0.490
-339	CC 5.48±1.79	CT 5.86±0.88	TT 0±0	0.772
-444	TT 0±0	TG 4.87±1.63	GG 5.57±1.79	0.461
-445	CC 5.80±1.71	CT 5.09±1.74	TT 5.39±2.28	0.490
-501	TT 0±0	TG 4.87±1.63	GG 5.57±1.79	0.461
-555	IN5.48±1.80	INDEL 5.80±1.63	DEL 0±0	0.809
-591	CC 0±0	CT3.81±1.18	TT 5.58±1.76	0.170
-720	CC 5.37±1.78	CT 7.44±0	TT 7.01±0.05	0.249

Regarding the analysis of the gene expression profile performed in the present study, all haplotypes evaluated exhibited expression of the Lalba gene, as well as the selected endogenous gene, GAPDH, in the somatic cells of milk. Haplotypes three (ATGCGCGINTC), four (GTCGACGCTGINTC), and two (AGTGCGCGINTC) presented the lowest (0.632±0.263) and highest (18.986±16.264; 15.386±12.956) expression levels, respectively. The last two mentioned were statistically significant ($P<0.05$) when compared to haplotypes one, three, five, and seven. The only one that was similar to the significant ones was haplotype six, which exhibited an expression profile of 9.706±5.199, a value close to those with the highest expression (Figure 1).

The gene expression profiles obtained from the seven haplotypes were compared with the average production data of the buffaloes evaluated in the study. However, as with the ten new SNPs found, there was no statistically significant difference in the average milk production ($P>0.05$). However, based on the analysis performed, it was possible to perceive that haplotype four, which presented the highest expression, could be considered a candidate for greater synthesis of mRNA of the Lalba gene in milk and, consequently, greater production of the α-lactalbumin protein in buffaloes.

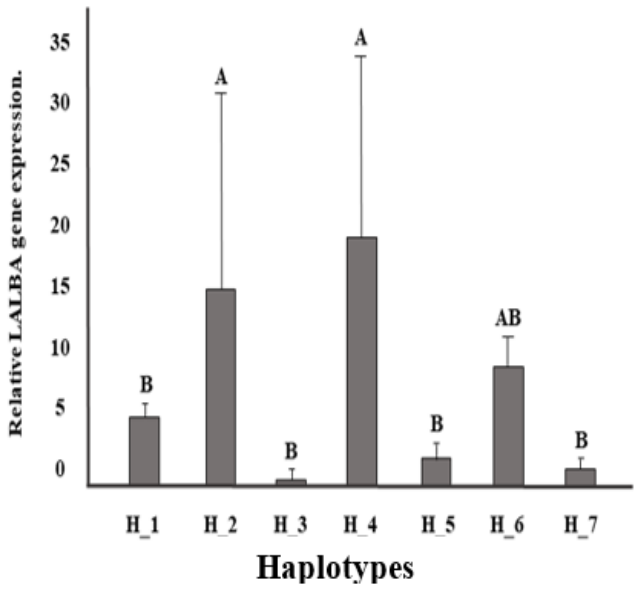

Figure 1: Relative gene expression from mRNA of the Lalba gene obtained from somatic cells of milk of the seven different haplotypes. According to the equation 2-ΔCt. Different letters above the bars indicate significant pairwise comparisons according to the Tukey test ($P<0.05$).

Discussion

The characteristics related to milk production in ruminants are controlled by quantitative loci traits (QTLs). These are spread throughout the genome and are regions in which genotypes are associated with phenotypes demonstrating differences in milk production and composition, for example. Therefore, the wide variety of cows selected for greater and better production and composition, both in quantity and quality, depends on identifying several candidate genes and their polymorphic variants (17,18). The search for so-called promising genes is commonly used due to the ease of understanding certain metabolic pathways and physiological processes and because they are close to QLTs (5).

The Lalba gene has been reported as polymorphic in many production animals, such as goats, sheep, cattle, and buffaloes. Therefore, there is evidence that nucleotide substitution mutations in regulatory and non-coding regions of genes influence the definition of complex characteristics (5,11,19,20).

In the present study, it was possible to perceive the variety of SNPs that can be found in buffaloes since ten new SNPs were detected in the regulatory region of the Lalba gene, with the one at position -42 (A>G) presenting a binding site for the transcription factor hepatocyte nuclear factor 3 (HNF-3). HNF-3 is a factor that plays a crucial role in the regulation of metabolism and tissue differentiation, mainly in the pancreas and liver (21). In another study with dairy cattle, it was observed that the SNP at chromosomal position g.146704699 G>A of the AGPAT3 gene, especially the A allele, created a transcription binding site (TFBS) for FOXA1 (Forkhead box protein A1, also known as hepatocyte nuclear factor 3-alpha (HNF-3-alpha). In addition, it showed a significant effect ($P<0.05$) on three short-chain fatty acids (C6:0: caproic fatty acid, C8:0: caprylic fatty acid, and C10:0: capric fatty acid) present in the milk of Holstein cows (22).

Regarding the genetic diversity parameters evaluated, the values found for allele and genotypic frequencies of wild-type alleles in most SNPs were high, which supports the degree of consanguinity observed in some SNPs through positive F_{IS} values. However, even with these SNPs suggesting an inbreeding process, 90% of those evaluated in the study are under HWE conditions due to the proximity in He and Ho values. Furthermore, this fact can also be explained by the assumption that inbreeding does not necessarily eliminate all alleles from a population but rather reorganizes them when present. Significant differences in He and Ho were observed only for the SNP (-720 C>T), which presented a deviation in HWE ($P>0.05$), so that the population may be subdivided by the presence of significant inbreeding or by the displacement of genes even from another population (23,24).

Many studies aimed at verifying genetic variability in water buffaloes are being carried out in many regions of the world; however, most are focused on microsatellite markers. In a study carried out with seventeen populations of Turkish water buffaloes, it was found that all twenty microsatellite loci tested were highly informative and polymorphic. However, the average F_{IS} value obtained in the study was 0.091, with all loci in HWE deviation (25). In the present work, the average F_{IS} obtained was 0.098, similar to what occurred in the previous study, but with divergence regarding HWE where of the ten SNPs found, nine are under HWE and only one in deviation. However, regarding the F_{IS} values isolated from each of the SNPs found, it can be suggested that the population studied presents only four in the process of inbreeding (positive values), characterizing a reduction in diversity based on many causes, such as the Wahlund effect, which consists of an increase in homozygotes and a decrease in heterozygotes in samples of individuals belonging to subdivided populations (26).

Regarding genetic diversity studies carried out with buffaloes in the Amazon region, the F_{IS} values found in the present work were higher than those found by other researchers in the literature (27-29), 0.0294 to 0.0329, 0.0741, and 0.0730 respectively. To investigate the genetic variability of the SNP (LEP-1620) in the leptin gene, a study was carried out with two buffalo breeds (Mediterranean and Murrah) and their crossbreds in two distinct breeding systems characteristics of the Amazon region (dry land and floodplain). In these populations, distinct FIS values were found, being negative (-0.125, -0.171, and -0.142) for Mediterranean and Murrah from dry land Mediterranean from floodplain and positive (0.057) for crossbreds raised in a floodplain system with all racial groups in HWE ($P>0.05$) (30). And the leptin gene is strongly related to milk production and characteristics such as fat and protein in buffaloes (31).

Research based on SNP detection in the regulatory region of the Lalba gene has been studied since the 1990s, where many SNPs have been reported since then in several ruminant species such as cattle, goats, sheep, and buffaloes, most of which are significantly associated with increased milk production, in addition to being associated with the main components of milk such as proteins and fats. And because they are located in the regulatory region, they can also influence the binding to many transcription factors and the levels of gene expression profiles (11,32-36).

Association studies with milk production traits for the Lalba gene are still scarce; however, one of the first studies aimed at elucidating the effect of an SNP in exon 1 of the gene in two distinct buffalo breeds, Bhadawari and Murrah, was able to demonstrate that the AB and BC genotypes present in Bhadawari individuals were significantly associated ($P<0.05$) with total milk production and daily milk production. Still, the AB, BB, BC, CC, and CD genotypes present in Murrah did not show significant effects

on these traits (12). Another non-significant SNP, the α-LA2516 SNP, located in exon 4 of the gene, was found to have no relationship with protein percentage, fat percentage, milk production, and number of somatic cells in milk (P>0.05) in Chinese Holstein cows (37).

Other authors have also found SNPs in the Lalba regulatory region, both in its 5' flanking region and in intron regions, which also had no significant effect on performance and milk production traits in cattle (38,39). This corroborates the findings of the present study, in which no significant association was observed regarding average milk production (L/day). The lack of significance in the associations with the animals' dairy performance may be due to the environmental, epigenetic effect, or intrinsic differences that the studied breeds have, which, in the case of buffaloes, are still little investigated (5).

However, more recent studies have demonstrated that there are significant effects between associations of certain types of mutations such as SNPs, copy number variants (CNVs), and even microsatellite studies in both river and swamp buffaloes (5,25,40). Regarding environmental and epigenetic effects, variations based on CNVs contribute to better environmental adaptation in buffaloes across the continent where they are raised. Thus, it was reported in the literature that a CNV was observed in the promoter region 1647 bp upstream of Lalba, and this overlaps with two different transposons (RTE-BovB and Bov-Ta3). But the interesting thing about this finding is that this CNV was common (CN = 2) in 95.08% of river buffaloes but was presented as a deletion (CN = 0) in 82.50% of swamp buffaloes, which may have contributed to the higher milk production of river buffaloes compared to swamp buffaloes.

SNP mutations can be found throughout the genome of a wide variety of species. However, when they are located along the regulatory region, they can interfere with the binding of existing transcription factors to the target DNA sequence analyzed, thus modulating the gene expression process through transcription inhibition or activation (41). Concerning animal production, this phenomenon can result in divergent reproductive and productive efficiencies in milk production.

The isolation of epithelial cells derived from somatic milk cells (MECs) is being used as a suggestion for non-invasive collection methods to obtain expression analysis of genes involved in milk production as well as its components. High amounts of mRNA from protein and fat genes were observed in MECs obtained from the Murrah breed compared to the Sahiwal cattle breed. The expressions for casein and serum protein (Lalba) genes were significantly (P<0.05) higher in MSCs purified during the early stage of lactation compared to those purified from other stages (peak, middle, and late) (42).

Another alternative to obtain mRNA of milk production genes is the somatic cells present in milk (MSCs), which is done by separating these cells by centrifugation. This was the method used in the present study to obtain mRNA from MSCs in the milk of buffaloes selected for evaluation of the expression of the Lalba gene. In the study carried out in buffaloes at the end of lactation, the expression of Lalba was low in MSCs; however, it was higher and significant (P<0.05) in buffalo milk fat globules (MFG), indicating that MFG can be used as another alternative for collecting these transcripts since the total RNA present in MFG derives from mammary epithelial cells (43).

From transcriptome analyses of four dairy buffaloes, it was possible to demonstrate that genes involved in production were highly expressed during all phases of lactation (initial, middle, and final). One of them was the Lalba gene, which exhibited high levels of expression in early lactation, with a subtle decline in the middle stage, but maintained its increased expression at the end (44). The increased expression of this gene is relevant to the synthesis of milk proteins since α-lactalbumin generally makes up about 20% of whey proteins (45). In addition, α-lactalbumin encoded by the Lalba gene is responsible for regulating lactose synthesis in many mammals, so lactose is one of the most valuable carbohydrates present in milk and plays an important role in the regulation of osmolarity. Therefore, the Lalba gene is considered relevant to the lactose production pathway in milk (46,47).

The data shown in the literature corroborate the gene expression values of the Lalba transcripts obtained in the present study, which were derived from milk MSCs, and all the haplotypes evaluated showed expression of the same, also showing that MSCs are excellent sources that provide transcripts to be evaluated in gene expression analyses. In which of the seven haplotypes evaluated, sets four and two were highly expressed and significant (P<0.05) in the MSCs of the crossbred buffaloes evaluated with levels much higher than those demonstrated by the authors cited above.

Conclusion

This study was able to demonstrate that the Lalba gene is polymorphic in crossbred dairy buffaloes raised in the Amazon region. Although no significant association with milk production was found, these haplotypes may have contributed to the significant expression of serum protein mRNA in them. Lalba gene transcripts in MSCs are expressed in buffalo milk. Of all the SNPs found, InDel stands out as an important piece of information for the genetic characterization of the population. This is present in 7% of the population and, in addition, approximately 93% of the animals are characterized by the presence of insertion of four nucleotide bases (TAAA) in their regulatory region, which can contribute with great importance to the evolutionary processes of living beings, as it can guarantee genetic variability in addition to the emergence of new adaptive characteristics. However, more associative studies should be carried out on buffalo breeds, especially those

raised in the Amazon region, to further promote and improve the dairy production chain.

Conflict of interest

This manuscript's authors declare no conflict of interest regarding the writing process or data analysis.

Acknowledgments

Gratitude to the team of the Molecular Biology laboratory of the Federal Rural University of Amazon and the financial support of this research by the Coordination for the Improvement of Higher Education Personnel Foundation (CAPES).

References

1. Cavali J, Pereira RA. Buffalo milk production. In: Salman AD, Pfeifer LM, editors. Dairy farming in the Amazon. Brasília: Embrapa; 2020. 391-399 p. [available at]
2. Filho WL, Mota AV, Mendonça RA, Domingues FN, Rêgo AC, Faturi C. Buffalo milk production system in the Arari region of the Marajó archipelago, Pará. Amazon J Agric Environ Sci. 2022;65:1-9. [available at]
3. Bernardes O. Buffaloes breeding in Brasil: position and economic relevancy. Rev Bras Reprod Anim. 2007;31:293-298. [available at]
4. Andrade KD, Rangel, NA, Araújo VM, Lima Júnior DM, Oliveira NA. Effect of season on the quality of buffalo milk. Rev Verd. 2011.6(3), 33-37. [available at]
5. Manzoor S, Nadeem A, Javed M. Polymorphism association and expression analysis of alpha-lactalbumin (LALBA) gene during lactation in Nili Ravi buffalo. Trop Anim Health Prod. 2020;52(1):265-271. DOI: 10.1007/s11250-019-02010-0
6. Hedkar CD, Kalyankar SD, Deosarkar SS. Buffalo milk. In: Caballero B, Finglas PM, Toldrá F, editors. Encyclopedia of Food and Health. UK: Academic Press; 2016. 522–528 p. [available at]
7. Pauciullo A, Cosenza G, Steri R, Coletta A, Jemma L, Feligini M, Di Berardino D, Macciotta NP, Ramunno L. An association analysis between OXT genotype and milk yield and flow in Italian Mediterranean river buffalo. J Dairy Res. 2012;79:150-156. DOI: 10.1017/s0022029911000914
8. Deng TX, Pang CY, Lu XR, Zhu P, Duan AQ, Liang XW. Associations between polymorphisms of the STAT1 gene and milk production traits in water buffaloes. J Anim Sci. 2016;94:927-935. DOI: 10.2527/jas.2015-0139
9. Li J, Liang A, Li Z, Du C, Hua G, Salzano A, Campanile G, Gasparrini B, Yang L. An association analysis between PRL genotype and milk production traits in Italian Mediterranean river buffalo. J Dairy Res. 2017;84:430-433. DOI: 10.1017/S0022029917000693
10. Reczyńska D, Zalewska M, Czopowicz M, Kaba J, Zwierzchowski L, Bagnicka E. Acute phase protein levels as an auxiliary tool in diagnosing viral diseases in ruminants – a review. Viruses 2018;10:E502. DOI: 10.3390/v10090502
11. Ostrowska M, Zwierzchowski L, Brzozowska P, Kawecka-Grochocka E, Żelazowska B, Bagnicka E. The effect of single-nucleotide polymorphism in the promoter region of bovine alpha-lactalbumin (LALBA) gene on LALBA expression in milk cells and milk traits of cows. J Anim Sci. 2021;99(7):skab169. DOI: 10.1093/jas/skab169
12. Dayal S, Bhattacharya TK, Vohra V, Kumar P, Sharma A. Genetic polymorphism of alpha-lactalbumin gene in riverine buffalo. DNA Seq. 2005;16(3):173-9. DOI: 10.1080/10425170500088205
13. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: A laboratory manual. 2nd ed. USA: Cold Spring Harbor Laboratory Press; 1989. 1659 p. [available at]
14. Solovyev VV, Shahmuradov IA, Salamov AA. Identification of promoter regions and regulatory sites. Methods Mol Biol. 2010;674:57-83. DOI: 10.1007/978-1-60761-854-6_5
15. Livak KL, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2 –ΔCt method. Methods. 2001;25:402-408. DOI: 10.1006/meth.2001.1262
16. Yeh FC, Boyle TB. Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belg J Bot. 1997;129:157-163. [available at]
17. Shannon CE. A mathematical theory of communication. Bell Sys Tech J. 1948;27:379-423. [available at]
18. Lopdell TJ. Using QTL to Identify Genes and Pathways Underlying the Regulation and Production of Milk Components in Cattle. Animals. 2023;2;13(5):911. DOI: 10.3390/ani13050911
19. Gárcia-gámez E, Gutiérrez Gil B, Sahana G, Sánchez JP, Bayón Y, Arranz JJ. GWAS analysis for milk production traits in dairy sheep and genetic support for a QTN influencing milk protein percentage in the LALBA gene PLoS One. 2012;7(10): e47782. DOI: 10.1371/journal.pone.0047782
20. Rahmatalla SA, Arends D, Ahmed AS, Reissmann M, Brockmann GA. Whey protein polymorphisms in Sudanese goat breeds. Trop Anim Health Prod. 2020;52(3):1211-1222. DOI: 10.1007/s11250-019-02119-2
21. Kaestner KH. The hepatocyte nuclear factor 3 (HNF3 or FOXA) family in metabolism. Trends Endocrinol Metab. 2000;11(7):281-5. DOI: 10.1016/s1043-2760(00)00271-x
22. Shi L, Wu X, Yang Y, Ma Z, Lv X, Liu L, Li Y, Zhao F, Han B, Sun D. A post-GWAS confirming the genetic effects and functional polymorphisms of AGPAT3 gene on milk fatty acids in dairy cattle. J Anim Sci Biotechnol. 2021;1;12(1):24. DOI: 10.1186/s40104-020-00540-4
23. Hartl DL, Clark AG. Principles of Population Genetics. 4th ed. USA: Sinauer Associates, Inc. Sunderland; 2006. 672 p. [available at]
24. Mcmanus C, Paiva S, Corrêa PS, Seixas L, Melo CB. Statistics for describing population genetics. Brasília: Série Técnica – Genética; 2011. 50 p. [available at]
25. Ünal EÖ, Isık R, Sen A, Geyik Kus E, Soysal MI. Evaluation of Genetic Diversity and Structure of Turkish Water Buffalo Population by Using 20 Microsatellite Markers. Animals. 2021;11:1067. DOI: 10.3390/ani11041067
26. Garnier-géré P, Chikhi L. Population Subdivision, Hardy–Weinberg Equilibrium and the Wahlund Effect. USA: John Wiley & Sons, Ltd; 2013. DOI: 10.1002/9780470015902.a0005446.pub3
27. Silveira KR. Population structure of dairy buffaloes using genomic data [master's thesis]. Brazil: Faculty of Agricultural Sciences, Federal University of Grande Dourados; 2019. [available at]
28. Marques JF, Martínez AM, Costa MR, Albuquerque MM, Quiroz J, Vega pla JL, Delgado JV. Genetic Diversity of Brazilian Buffaloes Bubalus Bubalis Using DNA Microsatellites Arch Zootec. 2011;60:1213-1221. [available at]
29. Ángel-Marín PA, Cardona H, Moreno-Ochoa M, Cerón-Muñoz M. Analysis of genetic diversity in Colombian buffalo herds. Rev Colomb Cienc Pecu 2010;23:411-421. [available at]
30. Silva LK, Barbosa EM, Silva SB, Campelo JG, Gonçalves EC, Silva CS, Marques JF, Silva Filho E. Polymorphism of the leptin gene in buffalo breed groups from eastern Amazon. Vet Ital. 2021;31;57(4):257-263. DOI: 10.12834/VetIt.1739.9163.2
31. Zetouni L, de Camargo GM, da Silva Fonseca PD, Gil FM, Lugo NA, Aspilcueta-Borquis RR, Cervini M, Tonhati H. Effects of a single nucleotide polymorphism in the leptin gene on the productive traits of dairy buffaloes (*Bubalus bubalis*). Mol Biol Rep. 2013;40(8):5159-63. DOI: 10.1007/s11033-013-2618-z
32. Bleck GT, Bremel RD. Sequence and single base polymorphisms of the bovine lactalbumin 5 flanking region. Gene. 1993;126:213–218. DOI: 10.1016/0378-1119(93)90369-E

33. Bleck GT, Bremel RD. Correlation of the alpha-lactalbumin (+15) polymorphism to milk production and milk composition of Holsteins. J Dairy Sci. 1993b;76:2292–2298. DOI: 10.3168/jds.S0022-0302(93)77566-9
34. Voelker GR, Bleck GT, Wheeler MB. Single-base polymorphisms within the 5' flanking region of the bovine a-lactoalbumin gene. J Dairy Sci. 1997;80:194–197. DOI: 10.3168/jds.s0022-0302(97)75927-7
35. Kaminski S, Rymkiewicz-Schymczyk J, Wojcik1 E, Rusé A. Associations between bovine milk protein genotypes and haplotypes and the breeding value of Polish Black-and-White bulls. J Anim Feed Sci. 2002;11:205–221. [available at]
36. Visker MW, Heck JL, Van Valenberg HF, Van Arendonk JM, Bovenhuis H. Short communication: A new bovine milk-protein variant: α-lactalbumin variant D. J Dairy Sci. 2012;95:2165–2169. DOI: 10.3168/jds.2011-4794
37. Zhou JP, Dong CH. Association between a polymorphism of the α-lactalbumin gene and milk production traits in Chinese Holstein cows. Genet Mol Res. 2013;12(3):3375–3382. DOI: 10.4238/2013.September.4.3
38. Zhang JD, Sun JE, Womack Y, Zhang Y, Wang Y, Zhang Y. Polymorphism Identification, RH Mapping and Association of α-Lactalbumin Gene with Milk Performance Traits in Chinese Holstein. Asian-Aust J Anim Sci. 2007;20(9):1327–1333. [available at]
39. Wickramasinghe S, Rincon G, Islas-Trejo A, Medrano JF. Transcriptional profiling of bovine milk using RNA sequencing. BMC Genom. 2012;13:45–56. DOI: 10.1186/1471-2164-13-45
40. Yang L, Han J, Deng T, Li F, Han X, Xia H, Quan F, Hua G, Yang L, Zhou Y. Comparative analyses of copy number variations between swamp buffaloes and river buffaloes. Anim Genet. 2023; 54(2):199-206. DOI: 10.1111/age.13288
41. Castellanos M, Mothi N, Muñoz V. Eukaryotic transcription factors can track and control their target genes using DNA antennas. Nat Commun. 2020;28;11(1):540. DOI: 10.1038/s41467-019-14217-8
42. Sharma A, Shandilya UK, Sodhi M, Jatav P, Mohanty A, Jain P, Verma P, Kataria RS, Kumari P, Mukesh M. Milk-derived mammary epithelial cells as non-invasive source to define stage-specific abundance of milk protein and fat synthesis transcripts in native Sahiwal cows and Murrah buffaloes. 3 Biotech. 2019;9(3):106. DOI: 10.1007/s13205-019-1642-7
43. Chen Q, Wu Y, Zhang M, Xu W, Guo X, Yan X, Deng H, Jiang Q, Yang X, Lan G, Guo Y, Qin G, Jiang H. Milk fat globule is an alternative to mammary epithelial cells for gene expression analysis in buffalo. J Dairy Res. 2016;83(2):202-8. DOI: 10.1017/S0022029916000133
44. Arora R, Sharma A, Sharma U, Girdhar Y, Kaur M, Kapoor P, Ahlawat S, Vijh RK. Buffalo milk transcriptome: A comparative analysis of early, mid and late lactation. Sci Rep. 2019;12;9(1):5993. DOI: 10.1038/s41598-019-42513-2
45. Linn JG. Factors affecting the composition of milk from dairy cows. Designing Foods: Animal Product Options in the Marketplace. USA: National Academy Press; 1988. [available at]
46. Miglior F, Sewalem A, Jamrozik J, Lefebvre DM, Moore RK. Analysis of milk urea nitrogen and lactose and their effect on longevity in Canadian dairy cattle. J Dairy Sci. 2006;89:4886–4894. DOI: 10.3168/jds.S0022-0302(06)72537-1
47. Qasba PK, Kumar S. Molecular divergence of lysozymes and alpha-lactalbumin. Crit Rev Biochem Mol Biol. 1997;32(4):255–306. DOI: 10.3109/10409239709082574

تعدد الأشكال الجديدة في جين ألفا لاكتالبومين وتحليل التعبير الجيني في جاموس الألبان في الأمازون

سينتيا لوانا بينهيرو سانتوس[١]، إليم كريستينا ماسيدو بارا دي سوزا[١]، رافاييل كاسيب غيماريش[١]، جوليانا فاسكونسيلوس فيغيريدو[١]، روبرتا دي أروجو سيلفا[١]، إليزابيث ماتشادو باربوسا[٢]، بريسيلا دي باولا بيسا سانتانا[٣] و إدنالدو دا سيلفا فيلهو[١]

[١]معهد الصحة والإنتاج الحيواني، مختبر البيولوجيا الجزيئية، الجامعة الفيدرالية الريفية في منطقة الأمازون، بيليم، [٢]قسم الشهادة الجامعية في التربية، جامعة أمابا الفيدرالية، مازاغاو، [٣]معهد الموارد الاجتماعية والبيئية والمياه، الجامعة الفيدرالية الريفية في منطقة الأمازون، بيليم، البرازيل

الخلاصة

يهدف هذا البحث إلى العثور على تعدد الأشكال الجديدة في جين Lalba لربطها بإنتاج الحليب وتحليل ملف تعريف التعبير الخاص بها. تم استخدام ١٠٤ إناث الجاموس في بوجارو / بنسلفانيا. تم إجراء استخراج الحمض النووي باستخدام الطريقة الفينولية، متبوعة بتفاعل البوليميراز المتسلسل والتنقية والتسلسل للتحقيق في تعدد الأشكال في منطقة المحفز و 5' UTR. تم استخراج الحمض النووي الريبي من الخلايا الظهارية للحليب لتحديد ملف تعريف التعبير عن لالبا في ٧ أنماط فردانية تم اختيارها لمتوسط إنتاج الحليب. تم العثور على عشرة تعدد أشكال جديدة في المنطقة المدروسة، مع التركيز على (- indel 555>TAAA) التي لوحظت في ٧٪ من المجموعة. لم يظهر أي من تعدد الأشكال ارتباطا كبيرا بمتوسط إنتاج الحليب. أظهرت سبعة تعدد الأشكال ترددات أليل النوكليوتيدات من النوع البري أكبر من ٠,٥، مع اكتشاف ٩٠٪ منها على أنها HWE. أظهر SNP -42 (A>G) موقع ربط قوي لعامل النسخ HNF3. كانت الأنماط الفردانية ٤ و ٢ مهمة عند مقارنتها بالآخرين وكان لهما تعبيرات أعلى. على الرغم من عدم تأثير SNP على إنتاج الحليب ، إلا أن الأنماط الفردانية ٤ و ٢ أظهرت عكس ذلك ، حيث قدموا مستويات أعلى من تعبير mRNA في خلايا الحليب الجسدية، والذي ربما يرجع إلى التأثير الإضافي للأليلات المدمجة، وبالتالي، من المستحسن أن يتم تثبيتها في السكان. ومع ذلك، يجب إجراء دراسات تركز على العلاقة بين تعدد الأشكال ومستويات التعبير في هذه الخلايا، جنبا إلى جنب مع الجينات الأخرى التي قد تؤثر على مستويات إنتاج الحليب، للتحقيق فيما إذا كانت هذه المتغيرات مؤكدة من خلال الأنماط الظاهرية للحيوانات.