TÜBİTAK

Research Article

Turk. J. Vet. Anim. Sci.
2012; 36(2): 107-113
© TÜBİTAK
doi:10.3906/vet-1009-475

Full length cDNA synthesis of differentially displayed ESTs during lactation in the Indian buffalo (*Bubalus bubalis*)

Ajai K. TRIPATHI[1,*], Krishna M. SINGH[1], Makrand N. VAZE[1], Umed V. RAMANI[1], Dharamsi N. RANK[2], Chaitanya G. JOSHI[1]

[1]Department of Animal Biotechnology, College of Veterinary Science & Animal Husbandry, Anand Agricultural University, Anand 388001, Gujarat - INDIA

[2]Department of Animal Genetics and Breeding, College of Veterinary Science & Animal Husbandry, Anand Agricultural University, Anand 388001, Gujarat - INDIA

Received: 17.09.2010

Abstract: In the present study, antisense primers were designed from 7 differentially displayed expressed sequence tags (ESTs) previously identified at our laboratory. Total RNA was isolated from the mammary gland tissue of lactating buffalo to carry out 5' rapid amplification cDNA ends (RACE) PCR. 5' RACE PCR amplicons, ranging from 161 bp to 336 bp, were cloned and sequenced. These 5' RACE sequences were analyzed using the blastn (Megablast) bioinformatics tool available from NCBI. Of these 7, only one 5' mRNA sequence did not show a match in the available database and another 1 failed to show a match with any specific gene due to the unavailability of buffalo genome information. Two other 5' mRNA sequences showed 82% to 100% homology with the 18S rRNA gene and the remaining 3 showed homology to the Talin-1, β-casein, and 28S rRNA genes. Full length cDNA were synthesized using primers designed from 5' mRNA sequences and ranged from 345 bp to 1030 bp. Finally, we conclude that the ESTs identified may act as markers for lactogenesis in other animals.

Key words: ESTs, RACE, full length, cDNA, rRNA

Introduction

Although the world's 188.3 million water buffalo comprise only 11.1% of the world's bovid population, more people depend on the water buffalo than on any other domesticated species in the world. The water buffalo is vital to the lives of small farmers and to the economy of many countries worldwide, including India (1). The water buffalo also provides more than 5% of the world's milk supply; this milk contains less water and more fat, lactose, protein, and minerals than cow's milk, and is used to make butter, butter oil, cheese, and other high quality dairy products. In addition, the animals have leaner meat that contains less fat and cholesterol than beef while offering a comparable taste (2). Genome analysis of the water buffalo has advanced significantly in recent years. However, in spite of the considerable information available on genome resources in the water buffalo in terms of cytogenetic characterization, whole genome mapping, and next generation sequencing, further work is still required.

The complete genome sequences of various organisms, including mammals, have recently become available as a consequence of rapid

* E-mail: drajais@gmail.com

advances in DNA sequencing technology (3). The principal route to understanding the biological significance of the genome sequence comes from the discovery and characterization of that portion of the genome that is transcribed into RNA products, the transcriptome, which represents the transacting fraction of the genetic information. Despite this, eukaryotic cDNA libraries are typically made from only the polyadenylated fraction. In terms of quantity, this polyadenylated fraction constitutes only 3% to 6% of the total RNA population. For these reasons, experimental representation of eukaryotic transcriptomes was usually done by constructing cDNA libraries from the polyA+ fraction of the RNA population (4).

Rapid amplification of cDNA ends (RACE) is a polymerase chain reaction (PCR)-based technique which was developed to facilitate the cloning of full length cDNA after a partial cDNA sequence has been obtained by other methods (5). The RACE RT-PCR strategy was first described by Frohman et al. (6). The missing sequence (cDNA ends) can be cloned by PCR using a technique variously called rapid amplification of cDNA ends (RACE), anchored PCR, or one-sided PCR (7).

A full length cDNA library is a powerful tool for functional genomics and is widely used as a physical resource for identifying genes. Genome-scale collections of full-length cDNA have become important for analyses of the structures and functions of expressed genes and their products. Since full-length cDNA carries complete protein coding sequences and untranslated regions, it is indispensable for the identification of genes and for the determination of protein primary structure (8). The generation of full-length cDNA from an mRNA template is a major challenge in biotechnology research. It constitutes the ultimate prerequisite in the construction of cDNA libraries and expression profiling and is a new area in DNA chip technology.

Milk production is a complex phenomenon involving many physiological systems and gene expression patterns during different stages of lactation add to its complexity. The complexity also lies in the involvement of many proteins yet to be evaluated completely for their specific role in milk synthesis. The physiology of milk production is so complex that we need a multidimensional approach in order to find answers to some of the most intriguing questions. Considering the aforementioned points, the present study was conducted to synthesize full length cDNA from differentially displayed expressed sequence tags (ESTs) during lactogenesis; the ESTs used had all been earlier identified at our laboratory and submitted to NCBI GenBank.

Materials and methods

Experimental material and total RNA extraction

For the present study, a total of 7 ESTs were chosen (NCBI Acc. No. FC456555, FC456563, FC456569, FC456579, FC456583, FC456589, and FC456593), all of which had been previously identified at our laboratory (9) and which were novel (showing no match in any database), or showed matches in a few but not in all databases or longer ESTs, respectively. Following all due aseptic precautions, the mammary gland tissue samples for the respective ESTs were collected from 3 Surti buffaloes (mean live weight of 300 ± 18 kg) reared at the Livestock Research Station of the College of Veterinary Science and Animal Husbandry at Anand Agricultural University, Anand. Biopsy samples were collected just after a period of peak yield (i.e. after 100 days), which is recorded as 90 days after parturition in buffaloes. Permission from the Committee for the Purpose of Control and Supervision of Experiments on Animals (CPCSEA) was obtained prior to initiation of the study.

Total RNA was extracted from each sample (100 mg) using Trizol reagent (Sigma) according to the manufacturer's instructions. The quality and quantity were checked with 1% formaldehyde agarose gel electrophoresis and an ND1000 spectrophotometer (NanoDrop, USA). The RNA samples were then treated with DNaseI (Fermentas) for the removal of possible genomic DNA contamination.

5' RACE PCR

5' RACE PCR was done using a FirstChoice RLM-RACE kit (Ambion) according to the manufacturer's instructions in combination with 7 antisense primers (Table 1). In brief, total RNA was dephosphorylated by incubation with calf intestinal phosphatase that removed the 5' phosphates from the truncated mRNA species and non-mRNA species. The mRNA

Table 1. The primers used for 5' RACE in the present study.

Sr. no.	ESTs	Accession no.	Primer sequences (5`-3`)	Annealing temperature
1	A1A11	FC456555	GTGATGTGGGGAAAGCAGTT	48 °C
2	H3C22.1	FC456562	GCTTACTTTGCCTTCTTGCATT	55 °C
3	H3G11.1	FC456569	CTGCTCAAGAACCCCAAGAA	54 °C
4	H25C21.7	FC456579	AGCTAGAGGCAGGCAAAGTC	53 °C
5	OA21.1	FC456583	ATCTGCACATGGACACCAGA	56 °C
6	OA32.7.10	FC456589	CAAAAAGCCCACCATCTCAC	54 °C
7	OC21.3.4	FC456593	TGGTCAGACAAGGGGATCTG	55 °C

was then decapped by incubation with a tobacco acid pyrophosphatase that exposes the 5' phosphates of the full-length mRNAs. An RNA oligo was ligated to the 5' end of the mRNA. The sample was reverse transcribed with random decamer and amplified by using a gene-specific primer (Table 1) and a primer targeting the RNA oligo, ligated to the 5' end of mRNA, (5'-GCTGATGGCGATGAATGAACACTG-3'), as per the manufacturer's instructions.

The 5' RACE PCR products were cloned using an InsT/Aclone PCR product cloning kit (Fermentas) and the vector containing the insert was propagated in *E. coli* host DH5-α as per the manufacturer's protocol. These clones were subjected to automated DNA sequencing on an ABI PRISM310 Genetic Analyzer (Applied Biosystems) using BigDye Terminator v3.1 Cycle sequencing chemistry.

Sequence analysis

Analysis of the 5' RACE amplified products of the respective ESTs was carried out using Megablast in databases like GenBank+EMBL+DDBJ+PDB sequences (nr/nt), GenBank non-mouse and non-human EST entries, human genomic+transcript, mouse genomic+transcript, and cow build 4 genome database.

Full length cDNA synthesis

Seven sense primers were designed and commercially synthesized (OCIMUM BIOSOLUTIONS, Hyderabad) from the 5' RACE sequences (Table 2) using PRIMER3 software (http://frodo.wi.mit.edu/).

For full length cDNA synthesis, cDNA was synthesized using primer ACAGAATTAATACGACTCACTATAGGT12VN-3' provided with an Ambion FirstChoice RLM-RACE kit as per the manufacturer's instructions. To amplify the full length cDNA, PCR was done for each of the 7 primers using 1.0 µL of cDNA, 0.25 µL (2.5 U) of Taq polymerase enzyme (Qiagen), 2.5 µL of 10 × Taq buffer, 2.0 µL of dNTPs, and 17.25 µL of nuclease free water with 5'-GCGAGCACAGAATTAATACGACT-3' and

Table 2. The primers used for full length cDNA synthesis.

Sr. no.	ESTs	Accession no.	Primer sequences (5`-3`)	Annealing temperature
1	A1A11	FC456555	ATCCCCTGAGAGACCCCTTG	58 °C
2	H3C22.1	FC456562	ATGATTATACGGCAGAGGAGAAG	56 °C
3	H3G11.1	FC456569	ATGCTTATGACCCGCACTTAC	54 °C
4	H25C21.7	FC456579	AGGTCTTGGCAGGTGGCTCT	60 °C
5	OA21.1	FC456583	ACGAACGTCTGCCCTATCAAC	54 °C
6	OA32.7.10	FC456589	AGCTCCTCCTTCACTTCTTGTCC	62 °C
7	OC21.3.4	FC456593	AACTTTGAAGGCCGAAGTGG	56 °C

5' RACE specific primer (Table 2) at 1.0 µL (10 pmol) each. The PCR condition was 94 °C for 5 min for the initial denaturation, then 35 cycles of 94 °C for 45 s for denaturation, the annealing temperature (Table 2) for respective ESTs, and 2 min for the extension at 72 °C with a final extension at 72 °C for 10 min; this was documented by a gel documentation system (SynGene, UK).

Results and discussion

To synthesize full length cDNA from differentially displayed ESTs previously identified at our lab, total RNA was extracted from the mammary gland tissue and 5' RACE PCR was done with 7 antisense primers. The 5' RACE PCR products, which varied in length from 180 bp to 350 bp (Figure 1), were cloned and sequenced. After a screening of the vector sequence through the VecScreen program, a sequence homology search (Megablast) was carried out for the respective ESTs in the GenBank database using the blastn algorithm. The nucleotide sequence (180-350 bp) data reported in this paper were submitted to GenBank with the accession numbers GH158820-GH158826.

5' RACE PCR for the first EST (FC456555) gave an amplification of 350 bp (Figure 1). From the cloning and sequencing of this product, a 320 bp nucleotide was found which showed no significant match with any database, except GenBank non-human and non-mouse EST, where it showed some similarity to a *Bos taurus* cDNA, but not much homology. In the cow build 4 genome database, it was located on chromosome 28 genomic contig (Table 3). The second differentially displayed EST (FC456562) through 5' RACE-PCR gave an amplification of 255 bp (Figure 1), which yielded 201 through sequencing. In the EST databases, it showed homology with reported 5' ends mRNA. In the cow build 4 genome, it showed high similarity (81% to 94%) to chromosomes 3, 15, and 5; however, no specific gene was identified (Table 3). The 5' RACE PCR of the third differentially displayed EST (FC4565569) was amplified to 345 bp (Figure 1), which gave 331 bp through 135 sequencing and showed homology (95%) with a 18S ribosomal RNA gene from the GenBank human and mouse database. In GenBank's EST database, it showed homology (95%) with the 5' end of reported mRNA and was located on *B. taurus* chromosomes 27, 3, and 11 (Table 3). The 5' RACE PCR of the fourth EST (FC456579) gave a product amplification of 290 bp (Figure 1) and 223 bp was obtained by cloning and sequencing, which showed homology with a Talin-1 protein of a different species. In the EST database, this EST showed 96% homology with the 5' ends of reported mRNA. Results showed that there was only 1 match

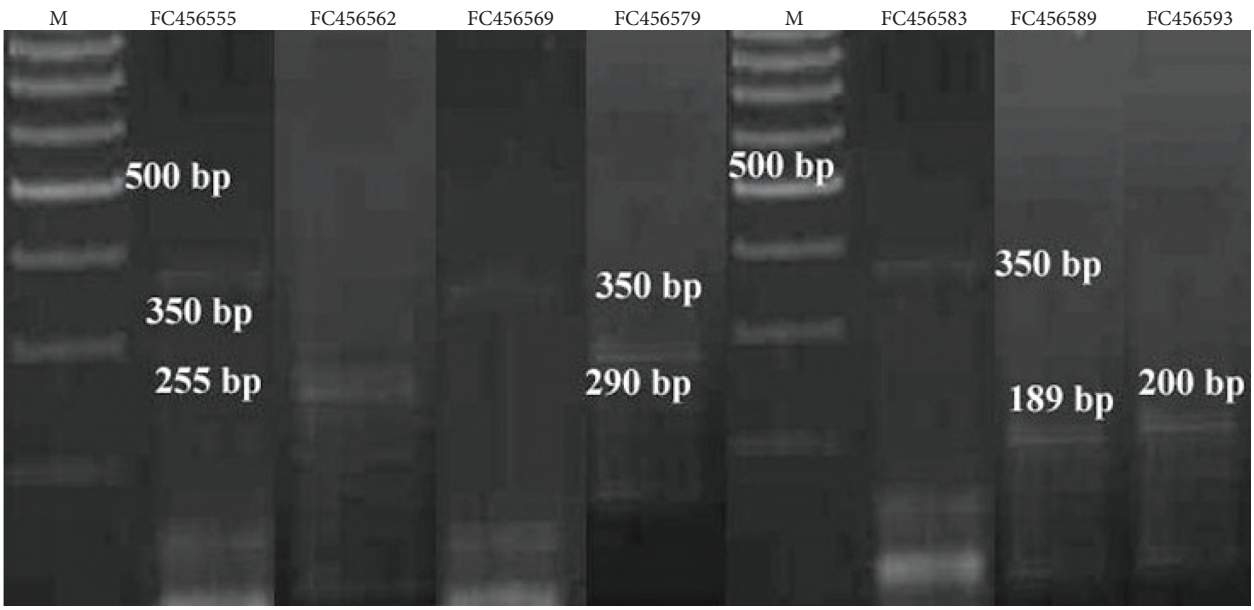

Figure 1. Agarose gel electrophoresis (2%) of the 5' RACE PCR products.

Table 3. A summary of the results of 5' RACE PCR in the present study.

Sr. no.	5' RACE product name	Size (bp)	Blastn (Megablast)					
			Nucleotide collection (nr/nt)	No. of nucleotides matched with the nucleotide collection (nr/nt)	ESTs	Human genomic + transcript	Mouse genomic + transcript	Cow build 4 genome (Ref only) (Chrm. No.)
1	5R FC55.2	320	NSSF	-	*Bos taurus* Corpus luteum	NSSF	NSSF	28
2	5R FC62.1	201	*Bos taurus* BAC CH240-6M21	158/189 (83%)	*Bos taurus* Skin	NSSF	NSSF	3, 15, 5
3	5R FC69.20	331	*Bos taurus* 18S ribosomal RNA gene, internal transcribed spacer 1, 5.8S ribosomal RNA gene	214/224 (95%)	*Ovis aries* Hind limb	*Homo sapiens* 18S ribosomal RNA, Genomic contig, 16, 2	misc RNA 17, 4, 8, 1	27, 3, 11
4	5R FC79.13	223	*Bos taurus* Talin 1 (TLN1), mRNA	208/216 (96%)	*Bos taurus* Ascending colon	*Homo sapiens* Talin 1 (TLN1), mRNA 9	Talin 1 (Tln1), mRNA 4	8
5	5R FC83.1	336	*Bos taurus* mitochondrial RNA, similar to 18S rRNA, clone: ORCS13221	319/327 (97%)	*Papaver somniferum* Root	*Homo sapiens* 18S ribosomal RNA 21, Y, Genomic contig	17, 6, Genomic contig	Genomic contig, 21
6	5R FC89.15	161	*Bubalus bubalis* beta-casein mRNA	140/141 (99%)	*Bubalus bubalis* Mammary gland	*Homo sapiens* casein beta (CSN2), mRNA	NSSF	6
7	5R FC93.4	167	*Bubalus bubalis* partial 28S ribosomal RNA	155/156 (99%)	000406BAGA004188HT BAGA *Bos taurus* cDNA 5', mRNA sequence.	*Homo sapiens* 28S ribosomal RNA 1, 2, 17, 6, 19	11, 13	1, 8, X

on chromosome 8, with 96% sequence homology (Table 3). It was observed that for upregulated EST FC456583, 5' RACE-PCR amplified a 350 bp product (Figure 1), which resulted in 336 bp after cloning and sequencing. A BLAST search showed homology (97%) with the 18s ribosomal RNA of different species (Table 3). 5' RACE PCR generated a 189 bp PCR product (Figure 1) for the sixth differentially displayed EST (FC456589). From the cloning and sequencing of this product, 161 bp nucleotides were found which showed high similarity with casein protein of different species including buffalo. In the EST database, it showed homology with the 5' ends of reported mRNA (100% homology). It was also found that there was only 1 match in the cow build 4 genome database (on chromosome 6) with 100% sequence homology (557458-557397) (Table 3). A 200 bp PCR product was amplified with the 5' RACE primer (Figure 1), which was designed with the upregulated EST FC456593. Cloning and sequencing of this product yielded a 167 bp product, which showed locations on chromosomes 1, 8, and X with 94%, 85%, and 89% sequence homology, respectively. It also showed sequence similarity with 28s ribosomal RNA of many species and in the EST database it showed homology with reported 5' ends mRNA (99%) (Table 3).

Full length cDNA was synthesized using cDNA, which was synthesized using 3' RACE adapter primer from an Ambion First Choice RLM RACE kit. The primers were designed from 5' RACE products for respective ESTs (Table 2) and 3' RACE specific adapter primer. In all 7 ESTs, there was amplification ranging from 345 bp to 1030 bp (Figure 2). In some ESTs, there was more than 1 band, which may be due to the premature termination of the transcription. In this study, 2 ESTs showed homology with 2 protein coding genes, Talin-1 and β-casein. Talin is responsible for myoblast fusion and sarcomere assembly and regulates the stability of myotendinous junctions (MTJs) (10) whereas β-casein, an important constituent of milk, shares 37% of its total caseins with those found in bovine milk (11).

There are many reports of full length cDNA synthesis of different genes in bovines, humans, mice, and other species but not for buffalo. Johnsen et al. (12) synthesized a full length proteose-peptone component 3 (PP3) cDNA of 679 bp from a bovine mammary gland cDNA library that encodes a

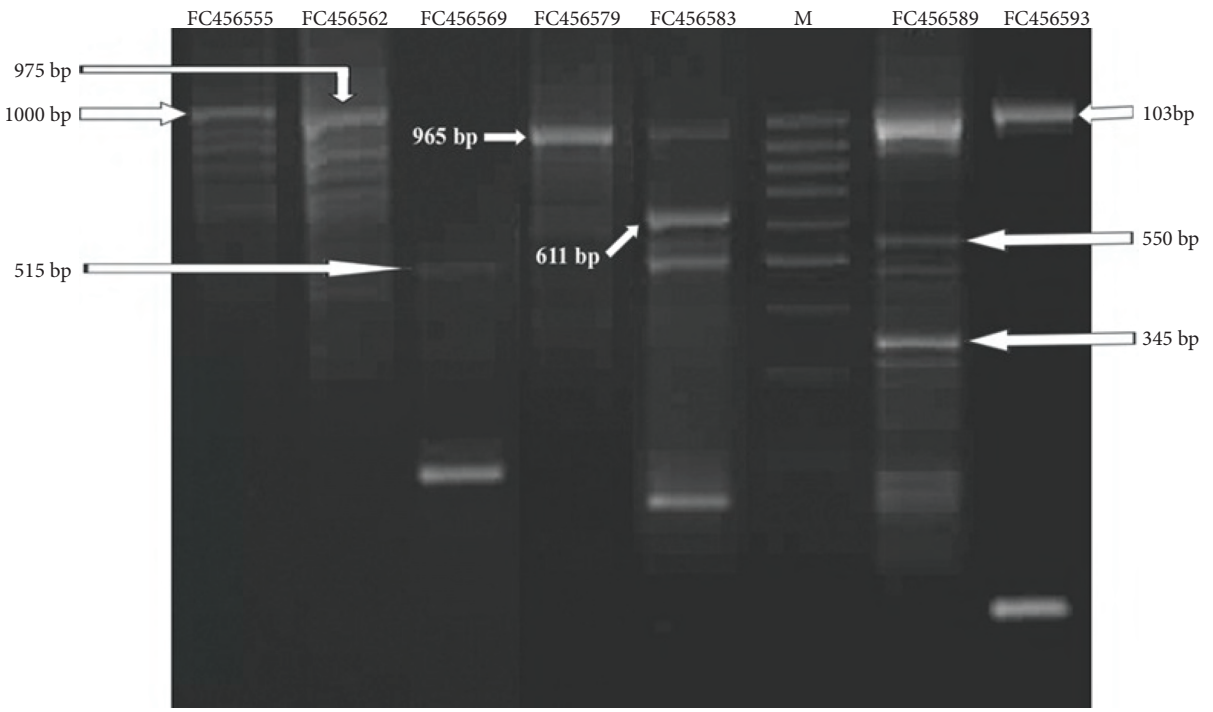

Figure 2. Agarose gel electrophoresis (2%) of the full length cDNA products.

signal peptide of 18 amino acids followed by the mature PP3 sequence of 135 amino acids, which showed homology with mouse and rat glycosylation dependent cell adhesion molecule 1 (GlyCAM-1). Kawachi et al. (13) characterized Bos taurus leptin receptor (Ob-R) isoform mRNA based on the GenBank database sequences of the partial sequence and cloned Ob-R isoforms full-length cDNA. Ma et al. (14) determined complete CDS sequence (994 bp) of the bovine FABGL gene by using a homology cloning approach combined with RT-PCR and RACE PCR. Similarly, Andreoni et al. (15) obtained a full-length cDNA sequence of bovine hexokinase type I from bovine brains with a sequence available in the database (GenBank acc. no. M65140) and analyzed a portion of the bovine hexokinase type I gene.

The differentially displayed ESTs used in this study, which had been previously identified gave amplification in 5' RACE PCR. Upon cloning and sequencing, the sequences revealed their similarities, mainly 18s and 28s rRNA and casein genes. During lactogenesis, as the transcription rate increases, there is also an increase in the level of ribosomal RNA to help in transcription. The synthesis of full length cDNA yielded sizes ranging from 345 bp to 1030 bp, which would be their full length for respective genes. We conclude that these identified genes may act as markers for lactogenesis in other animals. In systems biology, further investigation is required for the expression of full length cDNA in appropriate hosts in order to understand the interaction between these differentially expressed transcripts.

References

1. Tantia, M.S., Vijh, R.K., Bhasin, V., Sikka, P., Vij, P.K., Kataria, R.S., Mishra, B.P., Yadav, S.P., Pandey, A.K., Sethi, R.K., Joshi, B.K., Gupta, S.C., Pathak, K.M.: Whole-genome sequence assembly of the water buffalo (Bubalus bubalis). Indian J. of Animal Sci., 2011; 81: 38-46.
2. Food and Agriculture Organization of the United Nations (FAO): Water Buffalo: an asset undervalued. FAO Office for Asia and the Pacific, Bangkok, Thailand. 2000; 1-6.
3. Ohara, R., Kikuno, R.F., Kitamura, H., Ohara, O.: cDNA library construction from a small amount of RNA: adaptor-ligation approach for two-round cRNA amplification using T7 and SP6 RNA polymerases. BioTech., 2005; 38: 451-458.
4. Davis, C., Barvish, Z., Gitelman, I.: A method for the construction of equalized directional cDNA libraries from hydrolyzed total RNA. BMC Genomics, 2007; 8: 363.
5. Schaefer, B.C.: Revolutions in rapid amplification of cDNA ends: New strategies for polymerase chain reaction cloning of full-length cDNA ends. Analytical Biochem., 1995; 227: 255-273.
6. Frohman, M.A., Dush, M.K., Martin, G.R.: Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci., 1998; 85: 8998-9002.
7. Frohman, M.A.: On beyond classic RACE (rapid amplification of cDNA ends). PCR Methods Application, 1994; 4: 40-58.
8. Ling, P., Wang, M., Chen, X., Campbell, K.G.: Construction and characterization of a full-length cDNA library for the wheat stripe rust pathogen (Puccinia striiformis f. sp. tritici). BMC Genomics, 2007; 8: 145.
9. Ramani, U.V.: Study of differential gene expression in response to exogenous administration of bovine somatotropin hormone in Indian buffalo (Bubalus bubalis). MA Dissertation. Anand Agricultural University, Anand, Gujarat, India. 2007.
10. Conti, F.J., Monkley, S.J., Wood, M.R., Critchley, D.R., Müller, U.: Talin 1 and 2 are required for myoblast fusion, sarcomere assembly and the maintenance of myotendinous junctions. Development, 2009; 136: 3597-3606.
11. Choi, Y.J., Keller, W.L., Berg, J.E., Park, C.S., Mackinlay, A.G.: Casein gene expression in bovine mammary gland. J. Dairy. Sci., 1988; 71: 2898-2903.
12. Johnsen, L.B., Sorensen, E.S., Petersen, T.E., Berglund, L.: Characterization of a bovine mammary gland PP3 cDNA reveals homology with mouse and rat adhesion molecule GlyCAM-1. Biochim. Biophys. Acta., 1995; 1260: 116-8.
13. Kawachi, H., Yang, S.H., Hamano, A., Matsui, T., Smith, S.B., Yano-Comp, H.: Molecular cloning and expression of bovine (Bos taurus) leptin receptor isoform mRNAs. Biochem. Physiol. B. Biochem. Mol. Biol., 2007; 148(2): 167-73.
14. Ma, Y., Xu, S.Z., Gao, X., Ren, H.Y., Xin, Y.P., Gao, S.X., Zhang, Y.H.: Molecular cloning of bovine FABGL gene and its effects on bovine bioeconomic traits. Acta Genetica Sinica, 2006; 33: 1096-1104.
15. Andreoni, F., Serafini, G., Laguardia, M.E., Magnani, M.: Bovine hexokinase type I: Full-length cDNA sequence and characterisation of the recombinant enzyme. Mol. Cell Biochem., 2005; 268: 9-18.