Arch. Anim. Breed., 59, 363–372, 2016
www.arch-anim-breed.net/59/363/2016/
doi:10.5194/aab-59-363-2016

Open Access
Archives Animal Breeding

Molecular cloning, sequence characterization, and tissue expression analysis of three water buffalo (*Bubalus bubalis*) genes – *ST6GAL1*, *ST8SIA4,* and *SLC35C1*

Shen Song[1,*], Yina Ou-Yang[1,*], Jinlong Huo[1], Yongyun Zhang[2], Changlin Yu[1], Minhui Liu[3], Xiaohong Teng[1], and Yongwang Miao[1]

[1]Faculty of Animal Science and Technology, Yunnan Agricultural University, Kunming 650201, Yunnan, China
[2]Teaching Demonstration Center of the Basic Experiments of Agricultural Majors, Yunnan Agricultural University, Kunming 650201, Yunnan, China
[3]Yunnan Agricultural University Library, Kunming 650201, Yunnan, China
[*]These authors contributed equally to this work.

Correspondence to: Yongwang Miao (yongwangmiao1@126.com)

Received: 9 March 2016 – Revised: 26 July 2016 – Accepted: 1 August 2016 – Published: 16 August 2016

Abstract. Recent studies have shown that ST6 beta-galactosamide alpha-2,6-sialyltransferase 1 (ST6GAL1), ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 4 (ST8SIA4), and solute carrier family 35, member C1 (SLC35C1) play essential roles in the metabolism of milk glycoconjugates in mammals. However, studies on their coding genes in water buffalo have not been reported. In the present study, cloning and sequencing showed that the coding sequences (CDSs) of buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* were 1218, 1080, and 1095 bp in length, which encoded a precursor protein composed of 405, 359, and 364 amino acids, respectively. The deduced sequences of these three proteins in turn showed 97.6–98.5, 98.6–99.7, and 97.8–99.2 % similarities with other bovine species. Both buffalo ST6GAL1 and ST8SIA4 were predicted to be a member of glycosyltransferase family 29 and were all hydrophilicity proteins functioning in the Golgi apparatus. Buffalo SLC35C1 was a hydrophobic membrane protein located in the Golgi membrane, containing a TPT domain that is found in a number of sugar phosphate transporters. In addition, semi-quantitative RT-PCR analysis in 13 lactating buffalo tissues revealed that the *ST6GAL1* and *ST8SIA4* were expressed in 9 tissues, while *SLC35C1* was expressed in 11 tissues. The expression levels of these three genes in the mammary gland were significantly higher in lactating than in non-lactating stage. Collectively, our data indicate that *ST6GAL1*, *ST8SIA4*, and *SLC35C1* are potentially involved in the process of buffalo lactation.

1 Introduction

Glycoconjugates are some simple molecules, but they are the most important bioactive components in milk (Gopal and Gill, 2000). They protect against pathogens by acting not only as competitive inhibitors through binding on the epithelial surface of the intestine but also as growth promoters for colonic bacterial flora (Newburg, 1999; Gopal and Gill, 2000; Nakano et al., 2001). A number of studies on major components of milk have been performed; however, less attention has been paid to minor components in milk such as glycoconjugates (Martín-Sosa et al., 2009).

Glycosyltransferases and sugar transporters are glycosylation-related enzymes/proteins which are important to milk oligosaccharide metabolism (Wickramasinghe et al., 2011). Sialyltransferases are Golgi type II transmembrane glycosyltransferases (Harduin-Lepers et al., 2005). As one member of the sialyltransferases family, ST6 beta-galactosamide alpha-2,6-sialyltransferase 1 (ST6GAL1) transfers sialic acid from CMP-sialic acid to either type II (Galβ1, 4GlcNAc) free disaccharides or the N- or O-linked oligosaccharides using an α2,6-linkage (Maksimovic et al., 2011). Both human and bovine *ST6GAL1* genes have a coding region of 1218 nucleotides, which consists of

five exons (Wang et al., 1993; Mercier et al., 1999). Six single nucleotide polymorphisms (− 106C>T, −399A>G, −736G>A, −753A>C, −751C>T, and −736G>A) in human *ST6GAL1* gene were shown to affect transcription factor binding, of which three (−753A>C, −751C>T, and −736G>A) significantly affect promoter activity (Lee, 2008). Studies in human, mouse, and bovine showed that there is a trend of increasing *ST6GAL1* gene expression during lactation (Wang et al., 1993; Mercier et al., 1999; Maksimovic et al., 2011). ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 4 (ST8SIA4) is a key enzyme that catalyses the polycondensation of alpha-2,8-linked sialic acids during polysialic acid synthesis. Two residues which can affect the function of ST8SIA4 were identified in human. Arg82 significantly affects the ability of ST8SIA4 to polysialylate neuropilin-2 and SynCAM 1, and Arg93 plays a key role in substrate recognition (Harduin-Lepers et al., 1995, 2005). A study on transcriptome profiling of bovine milk oligosaccharides metabolism genes showed a higher level of expression of the *ST8SIA4* gene in late lactation than in transition, suggesting that the ST8SIA4 may play a vital role in sialylated oligosaccharides metabolism during cow lactation (Wickramasinghe et al., 2011). Solute carrier 35 (SLC35) is a family of nucleotide sugar transporters (Ishida and Kawakita, 2004). Solute carrier family 35, member C1 (SLC35C1) plays an important role in regulation fucosylation of glycans (Zhang et al., 2012). It transports GDP-L-fucose into the Golgi apparatus for further modification and conjugation by fucosyltransferases (Ishida and Kawakita, 2004). The increased expression of *SLC35C1* gene will increase the synthesis of GDP-L-fucose in milk (Wickramasinghe et al., 2011). In addition, studies in human demonstrated that single amino acid changed at some positions in the SLC35C1 will lead to the reduction of transporting activity. Then, the fucosylated glycoconjugates on the leukocyte surface will be lost, causing immunodeficiency (Gerardy-Schahn et al., 2001; Hirschberg, 2001).

Water buffalo contributes significantly to the animal production and the dairy industry in the tropical and subtropical countries (Singh et al., 2000; Khan et al., 2011; Perera, 2011). It has become the second largest source of milk supply in the world in recent years. Buffalo milk contains less water and more fat, lactose, protein, and minerals than that of Holstein cow milk (Vijh et al., 2008; Mahmood and Usman, 2010; Yindee et al., 2010). Compared with the milk of other domestic animals, the levels of oligosaccharides are much higher in buffalo milk (Bhanu et al., 2015). However, studies on the encoding genes of the synthesis and metabolism of glycoconjugates in water buffalo have not been reported. Since both water buffalo and cattle belong to the family Bovidae, a large amount of genetic/genomic resources from cattle research could serve as shortcuts for the water buffalo community to initiate genomic science in this species (Michelizzi et al., 2010). The key genes of bovine milk oligosaccharide metabolism identified by RNA sequencing in a previous study will provide some candidate genes for studying oligosaccharides in water buffalo milk (Wickramasinghe et al., 2011). Here, we focus on three milk oligosaccharide metabolism genes, *ST6GAL1*, *ST8SIA4*, and *SLC35C1*, which were identified in a previous study (Wickramasinghe et al., 2011). The aim of the present study is to clone the full-length coding sequences (CDSs) of buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes and further perform necessary bioinformatics analysis and tissue expression analysis. The results will establish a primary foundation for further understanding the biochemical functions of these three genes in water buffalo.

2 Material and methods

2.1 Sample collection, RNA extraction, and cDNA synthesis

All procedures for sample collection were performed in accordance with the Guide for Animal Care and Use of Experimental Animals and approved by the Institutional Animal Care and Use Committee of Yunnan Agricultural University. Six adult female Binglangjiang water buffalo, of which three were in non-lactating stage (the dry period) and another three were in the peak of lactation, respectively, were selected and slaughtered for sample collecting. The heart, pituitary gland, small intestine (the duodenum), longissimus dorsi muscle, spleen, liver, mammary gland, dorsal skin, lung, brain, kidney, rumen, and back adipose tissue were collected and preserved in liquid nitrogen immediately and then stored at −80 °C until further processing. Total RNA was extracted using the RNAiso Plus kit (TaKaRa, Dalian, China) following the manufacturer's instructions. To avoid potential DNA contamination, the RNA was treated with DNase I (TaKaRa, Dalian, China). The RNA quality from the different types of tissues was first assessed on a 1.5 % agarose gel electrophoresis containing ethidium bromine; then its concentration and purity were determined using the NanoDrop 2000 UV–Vis spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). The cDNA was synthesized from 3 µg RNA for each sample using an oligo(dT) 18 primer and M-MLV reverse transcriptase (Invitrogen, USA).

2.2 Primer design and gene isolation

Based on the mRNA sequence of cattle *ST6GAL1* (accession no. NM_001035373) and its highly homologous expressed sequence tags (accession no. DT809938, EE365129, EV620792, EE980740, GO784097, and EE903357), which were downloaded from the National Center for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/), a pair of primers were designed using Oligo 6.0 software to amplify the full-length CDS of buffalo *ST6GAL1* gene. Similarly, the primers for isolating buffalo *ST8SIA4* and *SLC35C1* genes were designed

Table 1. Primers used for isolation of water buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes.

Gene	Primers (5′ to 3′)	Amplicon length (bp)	Annealing temperature (°C)
ST6GAL1	F: CTCAGAACAGCGTGGTTTCC R: CACACACTCCCGTGACAACA	1451	55
ST8SIA4	F: CTATGAAGAGGAGGTGAGGGAG R: TTTCAGGTAAGTGGTGGATGCT	1255	57
SLC35C1	F: CAGGAAATCAGTCTGGCTGTGA R: TCAAAGGTCGTGTGGAGGTAGA	1494	56

Table 2. Primers used for semi-quantitative RT-PCR.

Gene (accession no.)	Primers (5′ to 3′)	Annealing temperature (°C)	References
ST6GAL1	F: GGTGTGCTGTGGTCTCTTCA R: CCCACGTCTTGTTGGAATTT	60	Maksimovic et al. (2011)
ST8SIA4	F: ACGCAACTCATCGGAGATGGTGA R: GTGTCCGTCGTCTGTCCAGC	60	Desanti et al. (2011)
SLC35C1 (NM_001101210)	F: GCATCTGGCGTCTGACCTT R: CGTCTGGGCACAGGCTTT	61	this study
18S rRNA (JN412502)	F: GGACATCTAAGGGCATCACAG R: AATTCCGATAACGAACGAGACT	55	this study

based on the nucleotide sequences of cattle *ST8SIA4* and *SLC35C1* genes and their respective homologous sequences (NM001001163, EE900497, EE965073, EE828093, EH146320, and NM_001101210), respectively. Detailed primer information is described in Table 1.

PCR reactions for cloning buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes were performed in a final volume of 25 µL containing 2.0 µL of 50 ng µL^{-1} cDNA, 2.0 µL of 2.5 mM dNTPs mixed (TaKaRa, Dalian, China), 2.5 µL of 10× Taq DNA polymerase buffer (Mg^{2+} Plus), 0.5 µL of 10 µM forward primer, 0.5 µL of 10 µM reverse primer, 0.25 µL of 5 U µL^{-1} Ex Taq DNA polymerase (TaKaRa, Dalian, China), and 17.25 µL of sterile water. The PCR mixtures underwent 5 min at 95 °C followed by 35 cycles of denaturing at 95 °C for 30 s, annealing with different temperature according to different primers and extension for 2 min at 72 °C and with a final extension cycle of 72 °C for 10 min. The annealing temperatures for RT-PCR are shown in Table 1. The amplified fragments were sub-cloned into the pMD18-T vector (TaKaRa, Dalian, China) and then sequenced bi-directionally using an automated DNA sequencer (ABI3730). At least eight independent clones were sequenced.

2.3 Semi-quantitative RT-PCR

Semi-quantitative RT-PCR (reverse transcription polymerase chain reaction) was conducted to evaluate the gene expression level of 13 tissues. Also, it was carried out to determine the differential expression in the mammary gland of both lactating and non-lactating stage. To eliminate the effect of cDNAs concentration, we repeated the RT-PCR five times using 1, 2, 3, 4, and 5 µL cDNAs as templates, respectively. The housekeeping gene *18S ribosomal RNA* was selected as an endogenous control for determination of targeted mRNA relative quantity because of its stable expression in most tissues of the body. The primers designed for the semi-quantitative RT-PCR analysis were presented in Table 2. The PCR conditions were as follows: 95 °C for 4 min followed by 30 cycles of denaturation at 95 °C for 30 s, annealing at a suitable temperature (Table 2) for 30 s, and extension at 72 °C for 30 s and a final extension cycle of 72 °C for 5 min.

2.4 Bioinformatics analysis

The amino acid sequences of the ST6GAL1, ST8SIA4, and SLC35C1 were deduced from the coding region via DNAStar version 7.0 (DNAStar, Inc., USA). Physicochemical characteristics, including theoretical molecular weight and isoelectric point, hydropathy, transmembrane region, signal peptide, and subcellular localization, were predicted using the Compute pI/Mw tool (http://web.expasy.org/compute_pi/; Walker, 2005), ProtParam tool (http://web.expasy.org/protparam/), ProtScale (http://web.expasy.org/protscale/), TMHMM version 2.0 (http://www.cbs.dtu.dk/services/TMHMM/; Krogh et al., 2001), SignalP 4.1 Server (http://www.cbs.dtu.dk/services/SignalP/; Pe-

tersen et al., 2011), and ProtComp 9.0 (http://linux1.softberry.com/berry.phtml), respectively. The conserved domains and functional sites were analysed using the Conserved Domain Architecture Retrieval Tool in BLAST at the NCBI server (http://www.ncbi.nlm.nih.gov/BLAST) and SMART (http://smart.emblheidelberg.de/). Secondary structures of deduced AA sequences were predicted by SOPMA (http://npsa-pbil.ibcp.fr/). The three-dimensional structure homology models of the ST6GAL1, ST8SIA4, and SLC35C1 were constructed using Swiss-Model (http://swissmodel.expasy.org/; Biasini et al., 2014). Multiple sequence alignment was performed using Clustal X 2.0 (Larkin et al., 2007). Neighbour-joining phylogenetic trees were generated based on the *ST6GAL1*, *ST8SIA4*, and *SLC35C1* sequences by applying MEGA version 6.0 (Tamura et al., 2013), which were subsequently subjected to be edited manually. Statistical reliability of the groups within phylogenetic trees was assessed using the bootstrap method with 10 000 replications.

3 Results

3.1 RT-PCR results for buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes

RT-PCR was adopted to amplify the full-length CDSs of buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes using the cDNAs as templates. The resulting PCR products were 1451, 1255, and 1494 bp, respectively (Fig. S1 in the Supplement).

3.2 Sequence analysis and genetic relationships

Sequence prediction showed that the cDNA sequences from this study in turn contained an open-reading frame of 1218, 1080, and 1095 bp, respectively. The nucleotide sequence analysis using the BLAST programme at the NCBI server (http://www.ncbi.nlm.nih.gov/) displayed that the gene sequences acquired were not homologous to any of the known water buffalo genes. Then the cDNA sequences of this study were deposited into the GenBank database under accession number KF360006, JX891653, and JX888717. The CDS of buffalo *ST6GAL1* has an overall base composition of 26.11 % A, 27.5 % C, 24.3 % G, and 22.09 % T, and that of buffalo *ST8SIA4* has a composition of 29.81 % A, 19.81 % C, 23.52 % G, and 26.85 % T. The *SLC35C1* CDS has a composition of 17.63 % A, 32.6 % C, 27.85 % G, and 21.92 % T.

Sequence alignment analysis between water buffalo and other common vertebrates was performed based on the ST6GAL1, ST8SIA4, and SLC35C1 sequences of this study and that published in the GenBank (http://www.ncbi.nlm.nih.gov/) (Fig. S2). Multiple alignment revealed that the amino acid sequences of buffalo ST6GAL1, ST8SIA4, and SLC35C1 had high homology with previously reported gene sequences in other species. Buffalo ST6GAL1 sequence has 97.6–98.5 % similarity with its homologous sequences of the species in the family Bovidae and has 74.3–82.9 % similarity with other mammalian species (Fig. S2a). Buffalo ST8SIA4 sequence shows 98.6–99.7 % homology with the species in the Bovidae and shows 93.3–99.7 % homology with other vertebrate species (Fig. S2b). However, the sequence of buffalo SLC35C1 shares 97.8–99.2 % similarity with its homologous sequences of the species in the Bovidae, and shares 86.3–99.2 % similarity with other mammalian species (Fig. S2c). In order to gain a better understanding of the genetic relationships of buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes to that of other species, phylogenetic trees are constructed on the basis of the CDSs of these three genes by the neighbour-joining method (Fig. 1). The *ST6GAL1*, *ST8SIA4*, and *SLC35C1* CDSs of water buffalo and other bovine species formed a sub-group in the corresponding trees with high support, which indicates that the three buffalo genes had a higher sequence identity with cattle, yak, and bison than with other species (Fig. 1).

3.3 Characteristics and structures of ST6GAL1, ST8SIA4, and SLC35C1 proteins

Buffalo ST6GAL1 includes 405 amino acids with a predicted molecular weight (MW) of ∼ 46.27 kDa and a theoretical pI of 9.04. Buffalo ST8SIA4 is 359 amino acids long and has a MW of about 41.29 kDa and a theoretical pI of 9.73. The SLC35C1 contains 364 amino acids with a predicted MW of about 40.11 kDa and a theoretical pI of 8.78. Both the ST6GAL1 and ST8SIA4 were classified as an unstable protein, with an instability index (II) of 41.12 for the ST6GAL1 and that of 47.36 for the ST8SIA4. The SLC35C1 was classified as a stable protein (instability index (II) = 33.75). Comparison of buffalo ST6GAL1, ST8SIA4, and SLC35C1 amino acid sequences with the sequences previously published in some representative species showed that the buffalo proteins all contained one conserved functional domain. The conserved domains of the ST6GAL1, ST8SIA4, and SLC35C1 are displayed in Fig. S3. Both buffalo ST6GAL1 and ST8SIA4 have a Glyco_transf_29 domain (aa 149–383 for the ST6GAL1, aa 94–354 for the ST8SIA4), which indicates these two proteins belong to glycosyltransferase family 29. However, the SLC35C1 contains a TPT domain (aa 39–338) that is found in a number of sugar phosphate transporters, including those with a specificity for triose phosphate. One transmembrane region (aa 9–27) was predicted in the ST6GAL1 (Fig. 2a). No transmembrane region was predicted in the ST8SIA4 (Fig. 2b), and eight transmembrane regions were predicted in the SLC35C1 (aa 36–58, aa 73–95, aa 116–135, aa 139–161, aa 168–185, aa 195–214, aa 227–249, and aa 264–286) (Fig. 2c). There were five, five and three kinds of functional sites predicted in buffalo ST6GAL1, ST8SIA4, and SLC35C1 proteins, respectively (Table S1). Cytplasmic–nuclear discrimination sug-

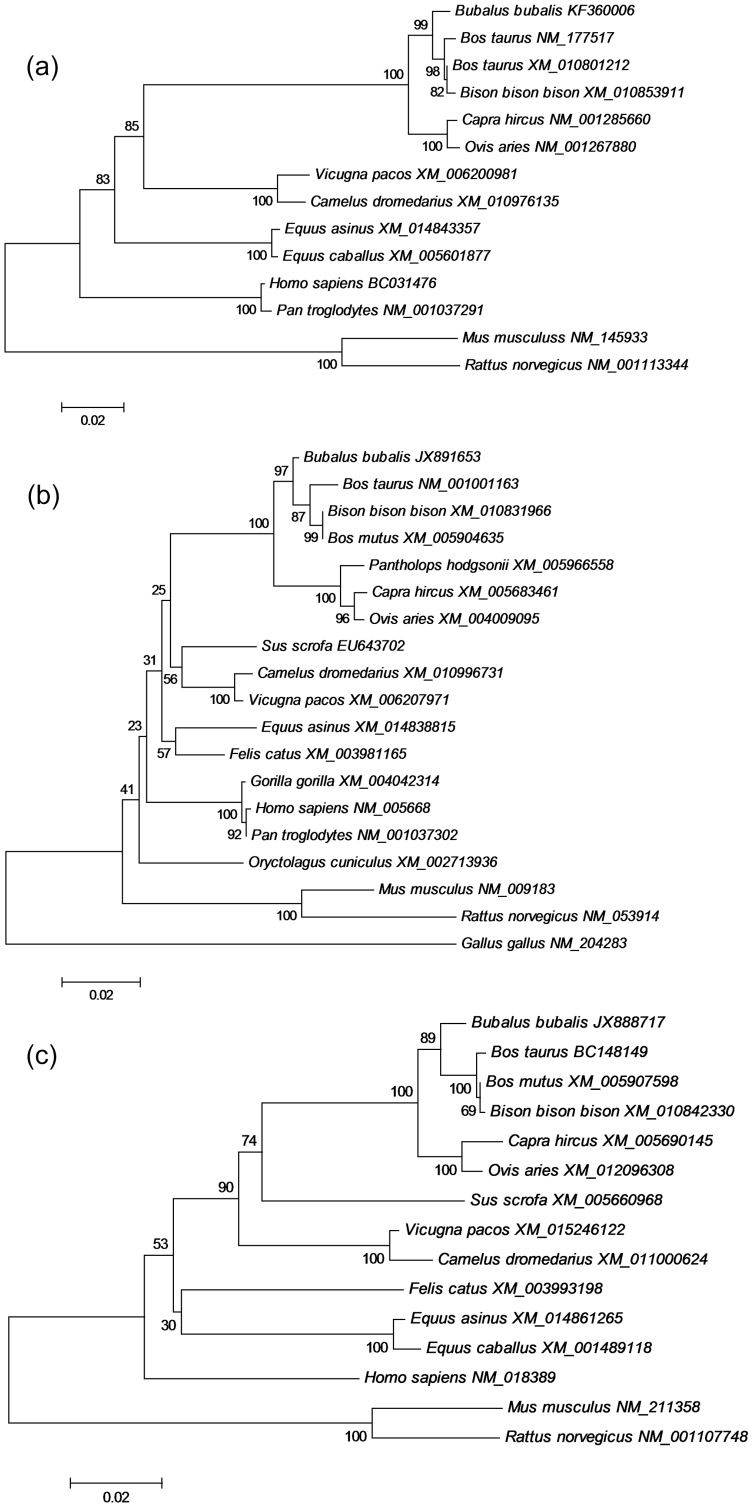

Figure 1. Phylogenetic analyses based on the CDSs of the three genes between buffalo and other species. **(a)** *ST6GAL1*, **(b)** *ST8SIA4*, and **(c)** *SLC35C1*.

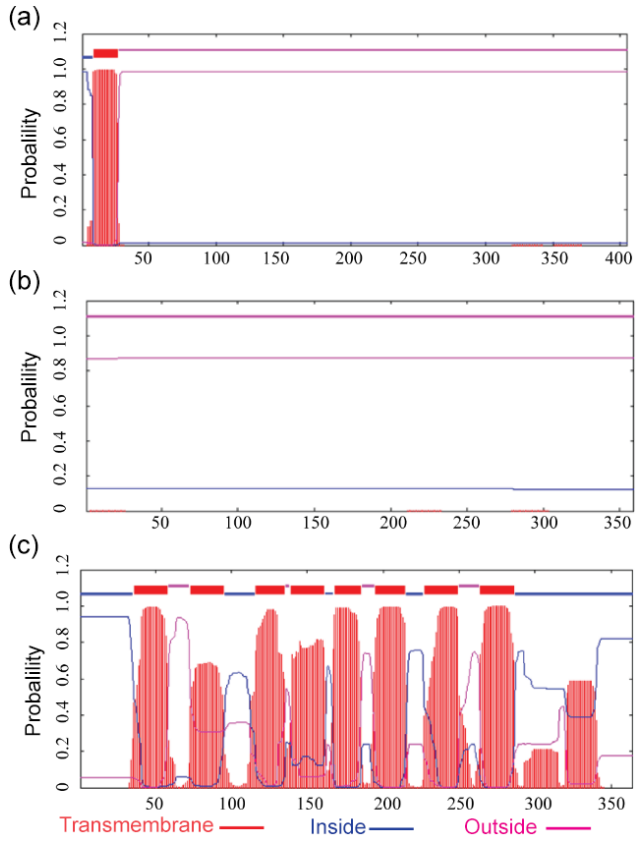

Figure 2. Transmembrane regions of buffalo ST6GAL1 **(a)**, ST8SIA4 **(b)**, and SLC35C1 **(c)** predicted by TMHMM programme.

gested that both buffalo ST6GAL1 and ST8SIA4 are possibly located in the Golgi apparatus with more than 98 % reliability, and buffalo SLC35C1 functions in membrane bound the Golgi. A putative N-terminal signal peptide was predicted in buffalo ST8SIA4, and its most likely peptide cleavage sites are between the 26th and 27th amino acids. However, no N-terminal signal peptide has been predicted in the amino acid sequences of both buffalo ST6GAL1 and SLC35C1. Hydropathy analysis showed that the grand averages of hydropathicity (GRAVY) for buffalo ST6GAL1, ST8SIA4, and SLC35C1 were −0.362, −0.247, and 0.493, respectively, which suggested that both buffalo ST6GAL1 and ST8SIA4 were hydrophilicity proteins, whereas buffalo SLC35C1 was a hydrophobin.

Prediction of secondary structure indicated that the deduced buffalo ST6GAL1 contains 32.35 % α-helix, 15.56 % extended strand, 3.70 % β-turn ,and 48.39 % random coil. The ST8SIA4 is composed of 37.88 % α-helix, 18.38 % extended strand, 5.57 % β-turn, and 38.17 % random coil. The SLC35C1 consists of 44.23 % α-helix, 24.45 % extended strand, and 2.20 % β-turn, and 29.12 % random coil (Fig. 3). The three-dimensional structure homology models of both buffalo ST6GAL1 and ST8SIA4 were built based on the target-template alignments (Fig. 3). The results showed that buffalo ST6GAL1 was similar to human beta-galactoside alpha-2,6-sialyltransferase 1 with 87.70 % identity and 78 % coverage, and buffalo ST8SIA4 was similar to the structure of human alpha 2,8-sialyltransferase with 38.14 % identity and 87 % coverage (Fig. 3). Buffalo SLC35C1 was unable to build a model since raw model contained fewer than three amino acid residues.

3.4 Tissue expression analysis

Tissue expression profiles of the *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes were assayed via semi RT-PCR in 13 tissues of lactating buffalo. The results demonstrated that the relative expression levels of the buffalo *ST6GAL1* gene were high in the pituitary gland; moderate in the spleen, liver, mammary gland, brain, and kidney; and weak in the small intestine, lung, and rumen, and there was no expression in the heart, longissimus dorsi muscle, dorsal skin, and back adipose tissues. Buffalo *ST8SIA4* gene was predominantly expressed in the longissimus dorsi muscle, moderately expressed in the pituitary gland, spleen, liver, mammary gland and lung, weakly expressed in brain, kidney and rumen, and minimally expressed in the heart, small intestine, dorsal skin and back adipose tissue. The *SLC35C1* gene exhibited highest expression level in the pituitary gland; moderate in the longissimus dorsi muscle, spleen, liver, mammary gland, lung, brain, kidney, and rumen; weak in the heart and small intestine; and minimal expression in the dorsal skin and back adipose tissue (Fig. 4a and c).

The expression pattern of buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes in the mammary tissues of lactating and non-lactating buffalo was analysed in the present study. Remarkable differences of the *ST6GAL1*, *ST8SIA4*, and *SLC35C1* mRNA expressions between the lactating and non-lactating period were observed in the mammary gland of water buffalo. These three genes showed higher expression levels in lactating than in non-lactating period. The expression levels of water buffalo *STGAL1*, *ST8SIA4*, and *SLC35C1* genes showed a significant difference between lactating stage and non-lactating stage (Fig. 4b). Their expression levels in lactating stage were 2.28, 1.30, and 1.69 times that of non-lactating stage, respectively (Fig. 4b).

4 Discussion

As more and more genomes are sequenced, comparative genomics offers new insights into the structural and functional characteristics of genes through comparison of individual gene sequences between different species (Ellegren, 2008). Comparative genomic-based strategies have begun to aid in the identification of functional sequences based on their high level of evolutionary conservation (Nobrega and Pennacchio, 2004). Sequence similarity information among the genomes of different species has become a major resource for cloning

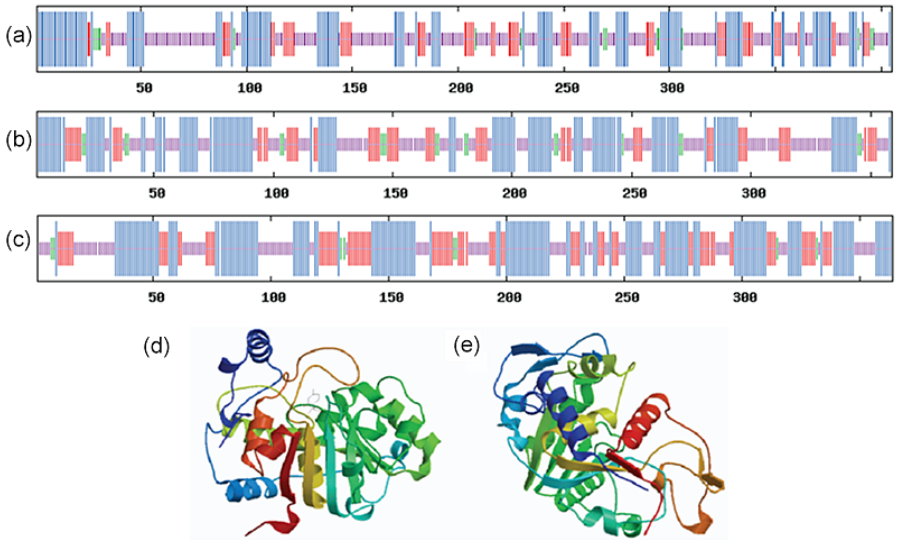

Figure 3. Predicted secondary structures of buffalo ST6GAL1 **(a)**, ST8SIA4 **(b)**, and SLC35C1 **(c)**, and tertiary structures of buffalo ST6GAL1 **(d)** and ST8SIA4 **(e)**. Alpha helix, extended strand, beta turn, and random coil are indicated with the longest, second longest, third longest, and shortest vertical lines, respectively.

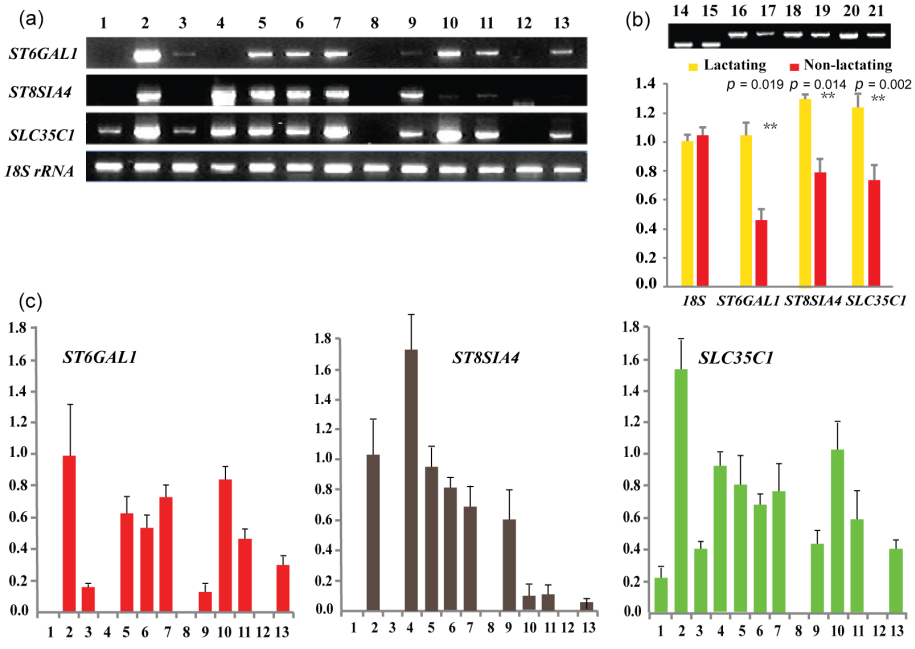

Figure 4. Expression analysis of buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes. Results are expressed as the ratio between the intensity of bands corresponding to the gene studied vs. the intensity of bands corresponding to *18S rRNA* internal control. The vertical axis represents gene relative quantification (mean ± SE) estimated from three biological replicates, and the horizontal axis indicates different tissues. Error bars represent the standard deviation of three samples. **(a)** Tissue expression profiles of buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes in the 13 tissues of lactating buffalo. **(b)** Differential expression of the reference gene *18S rRNA*, and buffalo *ST6GAL1*, *ST8SIA4*, *SLC35C1* genes in the mammary gland of lactating and non-lactating stage. Each lane in the gel image is corresponding to the bar of the bar chart. **(c)** Relative expression levels of buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes in the 13 tissues of lactating buffalo. Note: 1 – heart, 2 – pituitary gland, 3 – small intestine, 4 – longissimus dorsi muscle, 5 – spleen, 6 – liver, 7 – mammary gland, 8 – dorsal skin, 9 – lung, 10 – brain, 11 – kidney, 12 – back adipose tissue, 13 – rumen. 14, 16, 18, and 20 respectively represent the expressions of buffalo *18S rRNA*, *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes in lactating mammary gland, while 15, 17, 19, and 21 denote the corresponding expressions of these four buffalo genes in non-lactating mammary gland.

new genes, finding functional regions, and predicting functions, especially for improvement in identification of protein-coding genes in closely related species (Hardison, 2003). This extensive conservation in protein-coding regions between diverse species in mammals will provide us with a powerful approach to identify the functional regions of different genes for water buffalo. Water buffalo and cattle are the members of the family Bovidae, they have been shown to be closely related, and their chromosomes can be matched arm to arm at the cytogenetic level (Amaral et al., 2008). Compared with other common domestic animals, at present the studies on buffalo genome have lagged behind relatively. Taking advantage of the extensive resources and data now being accessible from the cattle-related studies, we can easily study water buffalo protein-coding genes through comparision with that of cattle.

Glycoconjugates are some of the most important bioactive components in milk, but little or no attention has been paid to these minor components in milk (Martín-Sosa et al., 2009). Previous studies have shown that *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes play important roles in the synthesis of glycoconjugates (Hamamoto and Tsuji, 2002; Hellbusch, 2007; Yang et al., 2012). In the present study, the complete CDSs of buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes were firstly isolated from buffalo cDNAs based on their counterpart sequence information of cattle and other species, and the CDSs of these three genes and their encoding proteins were characterized by adopting the method of comparative genomics. The CDSs of buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* are 1218, 1080, and 1095 nucleotides in length, which encode a protein containing 405, 359, and 364 amino acids, respectively. Sequence analysis showed that buffalo ST6GAL1, ST8SIA4, and SLC35C1 had high identity at the amino acid level with other animals, especially with bovine species, confirming that the genes were isolated from the correction. In addition, genetic relationships based on buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes revealed that water buffalo had a closer relationship with bovine species than with other species. These results indicate that buffalo ST6GAL1, ST8SIA4, and SLC35C1 have a similar function as that of bovine species. Our results also showed that the method of comparative genomics is extremely effectual to identify novel genes and to predict probable functions for the novel genes.

Previous studies in some vertebrates revealed that both ST6GAL1 and ST8SIA4 predominantly reside in the Golgi compartment where ST6GAL1 serves as a sialyltransferase transferring sialic acid from CMP-sialic acid to type II free disaccharides or the termini of N- or O-linked oligosaccharides (Maksimovic et al., 2011), and ST8SIA4 as a polysialyltransferase responsible for the polysialylation of the neural cell adhesion molecule (Close et al., 2000). They all are a member of glycosyltransferase family 29 (Harduin-Lepers et al., 2005; Foley et al., 2009). The SLC35C1 protein is a type III membrane protein which transports nucleotide sugars pooled in the cytosol into the lumen of Golgi apparatus, where most glycoconjugate synthesis occurs (Ishida and Kawakita, 2004). In the present study, a same function domain, Glyco_transf_29 domain, was predicted in both buffalo ST6GAL1 and ST8SIA4, suggesting they belong to the glycosyltransferase family 29. Both buffalo ST6GAL1 and ST8SIA4 are classified as hydrophilicity proteins and predicted possibly to function in the Golgi apparatus. A transmembrane domain of 19 amino acids (aa 9–27) was predicted in buffalo ST6GAL1, which is consistent with the observation in bovine (Mercier et al., 1999). Buffalo SLC35C1 was presumed as a hydrophobic protein existing in the Golgi membrane. A prior study showed that SLC35C1 was a protein with 10 transmembrane domains (Ishida and Kawakita, 2004). However, there were only eight transmembrane regions predicted in the present study. These results confirm the functional similarity for buffalo ST6GAL1, ST8SIA4, and SLC35C1 to that of other vertebrates. Protein analysis showed that all of these three proteins contain numerous putative sites, such as protein kinase C phosphorylation site, N-myristoylation site, casein kinase II phosphorylation site, tyrosine kinase phosphorylation site, and cAMP- and cGMP-dependent protein kinase phosphorylation site. Most protein functions are regulated by the modification of some amino acids in the polypeptide chain. Whether these putative functional sites play crucial roles in their corresponding protein remains a further investigation.

The tissue expression profile showed that these three genes were obviously differentially expressed in different tissues of the lactating buffalo. A possible explanation for the results is that at the same stage of development those biological activities associated with the function of the ST6GAL1, ST8SIA4, and SLC35C1 were presented in a wide range between different tissues within an individual organism. Tissue expression analysis in a dairy cow which was slaughtered 34 days following birth indicated that the expression level of the *ST6GAL1* mRNA in the liver was significantly higher than in any other tissues examined (Maksimovic et al., 2011). According to the EST profiles of cattle in UniGene (*ST8SIA4*, http://www.ncbi.nlm.nih.gov/UniGene/ESTProfileViewer.cgi?uglist=Bt.17789 and *SLC35C1*, http://www.ncbi.nlm.nih.gov/UniGene/ESTProfileViewer.cgi?uglist=Bt.27334), the *ST8SIA4* and *SLC35C1* showed the highest expression level in the kidney and intestine, respectively. However, our results showed that buffalo *ST6GAL1* and *SLC35C1* mRNAs were expressed more abundantly in the pituitary gland, whereas the *ST8SIA4* mRNA was expressed more abundantly in the longissimus dorsi muscle than in other tissues, suggesting that the genes play particularly key roles in water buffalo pituitary gland or longissimus dorsi muscle. The expressions of these three genes seem to display different patterns in water buffalo with cattle, and the reasons for these differences still need further investigation.

The expression of some lipogenic genes was detected in lactating and non-lactating mammary tissue of buffalo by semi-quantitative RT-PCR method (Yadav et al., 2012). In the current study, the expressions of buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes were compared in the mammary gland of both lactating and non-lactating stages using the same strategy. The results showed that buffalo *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes were moderately expressed in lactating stage. Furthermore, the three buffalo genes manifested higher expression in the lactating than in the non-lactating stage, which is consistent with the observation in cattle (Wickramasinghe et al., 2011). This implies that these buffalo genes may participate in some biological processes during lactation.

Progress in lactation biology of the bovine mammary has advanced substantially in recent years (Bauman et al., 2006). The knowledge of lactation biology and biosynthesis of milk will be used further for exploring the genes controlling the milk yield and composition in dairy animals for the improvement in quality and quantity of dairy products. Key roles have been attributed to water buffalo in providing milk in many developing countries in Asia, especially in tropical and subtropical countries (Perera, 2011). However, due to limited prior genomic characterization, the genetic improvement of water buffalo has lagged behind other bovine species (Amaral et al., 2008). With the promotion of consumer awareness of the links between diet and health, nutrition quality of milk is becoming increasingly important in food choice. Identification of the key genes involved in milk glycoconjugate synthesis was essential for improving milk quality in water buffalo.

5 Conclusions

In summary, three water buffalo novel genes, *ST6GAL1*, *ST8SIA4*, and *SLC35C1*, were firstly isolated and characterized. The results indicated that they encode three functional proteins which have the similar function to their counterpart proteins of other vertebrates, especially that of bovine species. These three buffalo genes manifested differential expression in 13 tissues during lactation, and compared with the non-lactating stage, the relative mRNA levels of these three buffalo genes remarkably increased in the mammary gland during the lactating stage. Our results suggested they may play critical roles in glycoconjugate synthesis and metabolism during lactation in water buffalo. However, there have been no genetic association studies between the polymorphisms of the *ST6GAL1*, *ST8SIA4*, and *SLC35C1* genes and milk traits in water buffalo. This study will provide a primary foundation for further insights into the association between their polymorphisms and milk composition traits, and functions of these three buffalo genes.

The Supplement related to this article is available online at doi:10.5194/aab-2-363-2016-supplement.

Acknowledgements. This study was financially supported by the National Natural Science Foundation of China (no. 31460582 and no. 30660024), the Natural Science Foundation Key Project of Yunnan Province, China (no. 2014FA032 and no. 2007C0003Z), the Applied and Basic Research Foundation of Yunnan Province, China (no. 2006C0034M), and the Foundation of Yunnan Department of Finance, China (study on the germ-plasm characteristics of Binglangjiang buffalo).

Edited by: S. Maak
Reviewed by: W. Deng and three anonymous referees

References

Amaral, M. E. J., Grant, J. R., Riggs, P. K., Stafuzza, N. B., Filho, E. A. R., Goldammer, T., Weikard, R., Brunner, R. M., Kochan, K. J., Greco, A. J., Jeong, J., Cai, Z., Lin, G., Prasad, A., Kumar, S., Saradhi, G. P., Mathew, B., Kumar, M. A., Miziara, M. N., Mariani, P., Caetano, A. R., Galvão, S. R., Tantia, M. S., Vijh, R. K., Mishra, B., Kumar, ST B., Pelai, V. A., Santana, A. M., Fornitano, L. C., Jones, B. C., Tonhati, H., Moore, S., Stothard, P., and Womack, J. E.: A first generation whole genome RH map of the river buffalo with comparison to domestic cattle, BMC Genomics, 9, 1–11, 2008.

Bauman, D. E., Mather, I. H., Wall, R. J., and Lock, A. L.: Major advances associated with the biosynthesis of milk, J. Dairy. Sci., 89, 1235–1243, 2006.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T. G., Bertoni, M., and Bordoli, L.: SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, Nucleic. Acids. Res., 42, 252–258, 2014.

Bhanu, L. S., Amano, M., Nishimura, S. I., and Aparna, H. S.: Glycome characterization of immunoglobulin G from buffalo (*Bubalus bubalis*) colostrum, Glycoconj J., 32, 625–634, 2015.

Close, B. E., Tao, K., and Colley, K. J.: Polysialyltransferase-1 autopolysialylation is not requisite for polysialylation of neural cell adhesion molecule, J. Biol. Chem., 275, 4484–4491, 2000.

Desanti, G. E., Jenkinson, W. E., Parnell, S. M., Boudil, A., Gautreau-Rolland, L., Eksteen, B., Ezine, S., Lane, P. J., Jenkinson, E. J., Anderson, G.: Clonal analysis reveals uniformity in the molecular profile and lineage potential of CCR9(+) and CCR9(−) thymus-settling progenitors, J. Immunol., 186, 5227–5235, 2011.

Ellegren, H.: Comparative genomics and the study of evolution by natural selection, Mol. Ecol., 17, 4586–4596, 2008.

Foley, D. A., Swartzentruber, K. G., and Colley, K. J.: Identification of sequences in the polysialyltransferases ST8Sia II and ST8Sia IV that are required for the protein-specific polysialylation of the neural cell adhesion molecule, NCAM, J. Biol. Chem., 284, 15505–15516, 2009.

Gerardy-Schahn, R., Oelmann, S., and Bakker, H.: Nucleotide sugar transporters: biological and functional aspects, Biochimie, 83, 775–782, 2001.

Gopal, P. K. and Gill, H. S.: Oligosaccharides and glycoconjugates in bovine milk and colostrum, Br. J. Nutr., 84, 69–74, 2000.

Hamamoto, T. and Tsuji, S.: ST6Gal-I, Handbook of Glycosyltransferases and Related Genes, Springer, 295–300, 2002.

Hardison, R. C.: Comparative genomics, PLoS. Biol., 1, 156–160, 2003.

Harduin-Lepers, A., Recchi, M. A., and Delannoy P.: 1994, the year of sialyltransferases, Glycobiology, 5, 741–758, 1995.

Harduin-Lepers, A., Mollicone, R., Delannoy, P., and Oriol, R.: The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach, Glycobiology, 15, 805–817, 2005.

Hellbusch, C.: Ein Knockout-Mausmodell für Congenital Disorder of Glycosylation-IIc: Defizienz des Golgi-GDP-Fucose-Transporters, eDiss, 2007.

Hirschberg, C. B.: Golgi nucleotide sugar transport and leukocyte adhesion deficiency II, J. Clin. Invest., 108, 3–6, 2001.

Ishida, N. and Kawakita, M.: Molecular physiology and pathology of the nucleotide sugar transporter family (SLC35), Pflugers. Arch., 447, 768–775, 2004.

Khan, F. A., Nabi, S. U., Pande, M., Das, G. K., and Sarkar, M.: Bilateral follicular cysts in a water buffalo, Trop. Anim. Health. Prod., 43, 539–541, 2011.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L.: Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J. Mol. Biol., 305, 567–580, 2001.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G.: Clustal W and Clustal X version 2.0, Bioinformatics, 23, 2947–2948, 2007.

Lee, H. L.: Genetic and Epigenetic Factors Controlling the Expression of Sialyltransferase Gene *ST6GAL1*: University of Hong Kong, 2008.

Mahmood, A. and Usman, S.: A comparative study on the physicochemical parameters of milk samples collected from buffalo, cow, goat and sheep of Gujrat, Pakistan, J. Nutr., 9, 1192–1197, 2010.

Maksimovic, J., Sharp, J. A., Nicholas, K. R., Cocks, B. G., and Savin, K.: Conservation of the ST6Gal I gene and its expression in the mammary gland, Glycobiology, 21, 467–481, 2011.

Martín-Sosa, S., Alonso, J., Sánchez-Juanes, F., Zancada, L., García-Pardo, L. A., and Hueso, P.: Sialoglycoconjugate content of milk replacers for neonatal calves, Span. J. Agric. Res., 7, 322–329, 2009.

Mercier, D., Wierinckx, A., Oulmouden, A., Gallet, P. F., Palcic, M. M., Harduin-Lepers, A., Delannoy, P., Petit, J. M., Leveziel, H., and Julien, R.: Molecular cloning, expression and exon/intron organization of the bovine beta-galactoside alpha2,6-sialyltransferase gene, Glycobiology, 9, 851–863, 1999.

Michelizzi, V. N., Dodson, M. V., Pan, Z., Amaral, M. E., Michal, J. J., McLean, D. J., Womack, J. E., and Jiang, Z.: Water buffalo genome science comes of age, Int. J. Biol. Sci., 6, 333–349, 2010.

Nakano, T., Sugawara, M., and Kawakami, H.: Sialic acid in human milk: composition and functions, Acta. Paediatr. Taiwan, 42, 11–17, 2001.

Newburg, D. S.: Human milk glycoconjugates that inhibit pathogens, Curr. Med. Chem., 6, 117–127, 1999.

Nobrega, M. A. and Pennacchio, L. A.: Comparative genomic analysis as a tool for biological discovery, J. Physiol., 554, 31–39, 2004.

Perera, B. M.: Reproductive cycles of buffalo, Anim. Reprod. Sci., 124, 194–199, 2011.

Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H.: SignalP 4.0: discriminating signal peptides from transmembrane regions, Nat. Methods., 8, 785–786, 2011.

Singh, J., Nanda, A. S., and Adams, G. P.: The reproductive pattern and efficiency of female buffaloes, Anim. Reprod. Sci., 60–61, 593–604, 2000.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S.: MEGA6: Molecular Evolutionary Genetics Analysis version 6.0, Mol. Biol. Evol., 30, 2725–2729, 2013.

Vijh, R. K., Tantia, M. S., Mishra, B., and Bharani Kumar, S. T.: Genetic relationship and diversity analysis of Indian water buffalo (*Bubalus bubalis*), J. Anim. Sci., 86, 1495–1502, 2008.

Walker, J. M.: The proteomics protocols handbook: Springer, 2005.

Wang, X., Vertino, A., Eddy, R. L., Byers, M. G., Jani-Sait, S. N., Shows, T. B., and Lau, J. T.: Chromosome mapping and organization of the human beta-galactoside alpha 2,6-sialyltransferase gene, Differential and cell-type specific usage of upstream exon sequences in B-lymphoblastoid cells, J. Biol. Chem., 268, 4355–4361, 1993.

Wickramasinghe, S., Hua, S., Rincon, G., Islas-Trejo, A., German, J. B., Lebrilla, C. B., and Medrano, J. F.: Transcriptome Profiling of Bovine Milk Oligosaccharide Metabolism Genes Using RNA-Sequencing, PLoS One, 6, e18895, doi:10.1371/journal.pone.0018895, 2011.

Yadav, P., Mukesh, M., Kataria, R. S., Yadav, A., Mohanty, A. K., and Mishra, B. P.: Semi-quantitative RT-PCR analysis of fat metabolism genes in mammary tissue of lactating and non-lactating water buffalo (*Bubalus bubalis*), Trop. Anim. Health. Prod., 44, 693–696, 2012.

Yang, Z. M., Chen, W. W., and Wang, Y. F.: Study on gene differential expressions of substance and energy metabolism in chronic superficial gastritis patients of Pi deficiency syndrome and of piwei hygropyrexia syndrome, Chinese Journal of Integrated Traditional and Western Medicine, 32, 1180–1187, 2012.

Yindee, M., Vlamings, B. H., Wajjwalku, W., Techakumphu, M., Lohachit, C., Sirivaidyapong, S., Thitaram, C., Amarasinghe, A. A. A. W. K., Alexander, P. A. B. D. A., Colenbrander, B., and Lenstra, J. A.: Y-chromosomal variation confirms independent domestications of swamp and river buffalo, Anim. Genet., 41, 433–435, 2010.

Zhang, P., Haryadi, R., Chan, K. F., Teo, G., Goh, J., Pereira, N. A., Feng, H., and Song, Z.: Identification of functional elements of the GDP-fucose transporter *SLC35C1* using a novel Chinese hamster ovary mutant, Glycobiology, 22, 897–911, 2012.