Nucleotide sequencing and DNA polymorphism studies of beta-*lactoglobulin* gene in native Saudi goat breeds in relation to milk yield

A.A.M. El Hanafy^{1,2}, M.I. Qureshi¹, J. Sabir¹, M. Mutawakil¹, M.M.M. Ahmed^{1,2}, H. El Ashmaoui^{1,3}, H.A.M.I. Ramadan^{1,3}, M. Abou-Alsoud¹, M.A. Sadek¹

ABSTRACT: β-Lactoglobulin (β-LG) is the dominant non-casein whey protein found in milk of bovine and of most ruminants. The amino acid sequence of β-LG along with its 3-dimensional structure illustrates linkage with the lipocalin superfamily. Preliminary studies in goats indicated that milk yield can be influenced by polymorphism in genes coding for whey proteins. The aim of this study is to identify and evaluate the incidence of functional polymorphisms in the exonic and intronic portions of β -LG gene in native Saudi goat breeds (Ardi, Habsi and Harri). Blood samples were collected from 300 animals (100 for each breed) and genomic DNA was extracted using QIAamp DNA extraction kit. A fragment of the β -LG gene from exon 7 to 3' flanking region was amplified with pairs of specific primers. Subsequent digestion with Sac II restriction endonuclease revealed two alleles (A and B) and three different banding patterns or genotypes, i.e. AA, AB, and BB. The statistical analysis showed that β -LG AA genotype had higher milk yield than β -LG AB and β -LG BB genotypes. Nucleotide sequencing of the selected β -LG fragments was done and submitted to GenBank NCBI (Accession Nos. KJ544248, KJ588275, KJ588276, KJ783455, KJ783456, KJ874959, and KP269078). Two already established SNPs in exon 7 (+4601 and +4603) and one fresh SNP in the 3' UTR region were detected in the β-LG fragments with designated AA genotype. The exonic SNPs, i.e. +4601 (G/A) and +4603 (G/C), were found within the Sac II restriction site and accountable for generating the AA genotypic patterns. Hence, the allele characterized by the substitution G>A has been sub-designated as AA^A , while the one characterized by the substitution G>C as AA^{C} . The polymorphisms in exon 7 did not produce any amino acid substitution.

Keywords: β-LG; Saudi goats; functional polymorphism; PCR-RFLP; genotyping

INTRODUCTION

Domestic goat (*Capra hircus*) is among the first domesticated livestock species and hence integral to animal husbandry. Goats are extensively reared at a global scale particularly in the developing world and serve as a vital resource of meat, milk, fibre,

and pelts (Zeder and Hesse 2000; MacHugh and Bradley 2001; Qureshi et al. 2014). Variable milk yields have been recorded among different goat breeds and within breeds as well. Researchers have established that milk yield can be influenced by polymorphisms in genes coding for whey proteins (Kumar et al. 2006; El-Hanafy et al. 2010).

Supported by the NSTIP strategic technologies program, Kingdom of Saudi Arabia (Project No. 11-BIO518-03). The authors confirm that this article content has no conflicts of interest.

¹Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia

²Nucleic Acids Research Department, Genetic Engineering and Biotechnology Research Institute, City for Scientific Research and Technology Applications, Borg El-Arab, Egypt

³Cell Biology Department, Genetic Engineering and Biotechnology Division, National Research Centre, Cairo, Egypt

 β -Lactoglobulin (β -LG) is the dominant noncasein whey protein present in milk of bovine and of most ruminants. It has widely been accepted to be absent in humans, though there are indications of minor presence (Hambraeus and Lonnerdal 2003). The amino acid sequence of β-LG along with its 3-dimensional structure reveals association with the lipocalin superfamily (Kontopidis et al. 2004). Its ability to bind hydrophobic bioactive substances and amphiphilic molecules ranges from hexane to palmitic acid to retinol to vitamin D. Studies have revealed β-LG-pectin complexes as molecular nano-vehicles for delivering hydrophobic nutraceuticals such as ω-3 polyunsaturated fatty acids and vitamin D (Wang et al. 1997; Kontopidis et al. 2004; Zimet and Livney 2009; Ron et al. 2010; Cui et al. 2012). Other biological activities of β-LG include enzyme regulation, neonatal acquisition of passive immunity, a source of bioactive peptides, and antimicrobial activity against mastitis-causing bacteria. It is actively employed in the food industry for numerous characteristics (Kontopidis et al. 2004; Cui et al. 2012).

Milk proteins show genetic polymorphism due to nucleotide sequence substitution or deletion, various degrees of glycosylation and phosphorylation. To date, 14 β-LG variants in Bos genus (B. taurus, B. javanicus, and B. grunniens) have been identified at the protein and DNA levels (Caroli et al. 2009; Cui et al. 2012). Investigations have also revealed the polymorphic nature of β -LG in sheep with three genetic variants (A, B, and C). Possible relationship between β -LG polymorphism and yield, composition, and cheese-making ability of milk has been widely studied in different sheep varieties (Arora et al. 2010). Preliminary investigations have reported novel SNPs in the goat β-LG gene (Pena et al. 2000; Kumar et al. 2006; Jain et al. 2012).

Ardi, Habsi, and Harri are the three distinct Saudi Arabian goat breeds holding a special place in the regional agribusiness economy, though lack detailed genetic characterization. As the regional climatic conditions are quite harsh, livestock here has evolved to survive in extreme arid environment. Ardi and Harri goats share high genetic similarity (73.5%) and are wide-spread compared to Habsi animals. Harri breed is reared for its high milk productivity while Ardi lacks behind in terms of yield but is quite famous among the desert dwellers owing to its persistent milking

ability (Sabir et al. 2012, 2013). Limited information is available regarding Habsi goats. According to traditional breeders, Habsi is far more productive in terms of milk production and statistics from this study significantly support this notion. Earlier regional studies exploring β -LG polymorphism in goats have found novel genetic variants (Kumar et al. 2006; El-Hanafy et al. 2010). As no preliminary data is available regarding β -LG polymorphism in relation to milk production in Saudi breeds, the current investigation concentrated on analyzing β -LG polymorphism at the DNA level and analyzing potential associations with utility traits like milk yield.

MATERIAL AND METHODS

Experimental animals. Blood samples from Ardi and Harri goats (n = 100 for each breed) were collected from private farms located in the suburbs of Jeddah province and Riyadh city, while samples belonging to Habsi breed (n = 100) were obtained from two farms located near Al-Qunfudhah village, South Jeddah. All animals were female goats in their first to second season of lactation with an average age of 1-1.5 years.

Milk recording and statistical analysis. Animals were raised under extensive husbandry conditions and were fed on hay (ad libitum) and concentrated feed mixture (14% crude protein (CP)) according to their requirements. Daily recording of milk yield was initiated following 4 days of kidding. Milk yield was recorded 3 days per week during the first 16 weeks of lactation and average daily and weekly yield were then estimated. To document daily milk yield, female goats were separated from their kids for 9 h and afterwards hand milked and yield subsequently recorded. Total milk yield for each studied breed was statistically analyzed by one way ANOVA and the variations were tested by *F*-test at significance level (P < 0.05). Genotypes were determined by direct counting of restriction fragments observed in the gel. The frequency of different genotypes was calculated using the standard procedure given by Falconer and Mackay (1996):

Genotype frequency = n of individuals of a particular genotype/total n of individuals of all genotypes

Blood collection and DNA isolation. Around 10 ml of blood was collected from each goat using 0.5 ml ethylene diamine tetra acetate (EDTA; 0.5M, pH = 8) as an anticoagulant. The samples

Table 1. Primers (forward and reverse) utilized for the amplification of exon 7 to 3' UTR region of the β -LG gene

Primer	Primer sequence (5'-3')	Molecular weight	Tm (°C)	Length (mer)	GC content (%)	Size of amplicon (bp)
Primer (F)	CGGGAGCCTTGGCCCCTCTG	6126	69	20	75.0	427
Primer (R)	CCTTTGTCGAGTTTGGGTGT	6161	58	20	50.0	

Tm = primer melting temperature

were transported to the animal genetics laboratory and kept at -20°C until the isolation of genomic DNA. Genomic DNA was extracted using the QIAamp DNA extraction kit.

PCR-RFLP. A region of the β -LG gene spanning over exon 7' to 3' flanking area was amplified by employing a set of forward (5'CGGGAGCCTTG-GCCCCTCTG3') and reverse (5'CCTTTGTC-GAGTTTGGGTGT3') primers (Table 1) (Kumar et al. 2006). Amplification was achieved through a pre-programmed MultiGeneTM Thermal Cycler (Labnet International Inc., New Jersey, USA) in 25 µl reaction mixture (reaction recipe in Table 2). While formulating final recipe in 200 µl PCR tube, 23 µl of master mix and 2 µl of genomic DNA were blended in each animal specific reaction (amplification profile in Table 3). PCR amplicons were digested overnight at 37°C with 10 Units of Sac II restriction endonuclease. Restriction enzyme, 10× reaction-buffer, and autoclaved triple distilled water (TDW) were mixed for all the mandatory reactions in a 1.5 ml micro-centrifuge tube to design a master mix which was then dispensed into labelled 200 μl PCR tubes. Following digestion, the samples were resolved in 3% (w/v) agarose gel in $1\times$ TAE buffer stained with ethidium bromide for distinguishing genotypes.

Sequencing analysis. Selected β -LG fragments with designated genotypes were sequenced on both strands with the same set of primers used earlier

Table 2. PCR recipe

Reaction components	Quantity (1×)
dNTP Mix	0.5 μl (200μΜ)
Forward primer (10 pmol)	1.0 µl
Reverse primer (10 pmol)	1.0 µl
$10\times$ buffer Complete with MgCl_2	2.5 μl
Taq DNA polymerase	0.5 μl (1 U)
Autoclaved TDW	16.5 μl
Genomic DNA	2.0 μl (100 ng)
Total volume	25 μl

TDW = triple distilled water

for PCR amplification following Sanger's dideoxy chain termination method (Sanger et al. 1977). The sequences of different genotypes were analyzed using the ClustalW sub-programme of the BioEdit Sequence Alignment Editor (BioEdit Version 7.2.5, 2013) to generate sequence alignment report.

RESULTS AND DISCUSSION

PCR-RFLP analysis. The 427 bp β -LG amplicons were digested with Sac II to establish polymorphic sites within the coding region of the gene. Sac II with a recognition site (5'...CCGC \downarrow GG...3'/3'... $GG\uparrow CGCC...5'$) revealed two alleles (A and B) with three different restriction patterns or genotypes (Figure 1). The β -LG AB genotype had two restriction sites and generated three bands, i.e. 427 bp, 349 bp, and 78 bp. The β -LG BB genotype with only one restriction site revealed two bands of sizes 349 bp and 78 bp. An undigested product of size 427 bp termed as β-LG AA (AA^A and AA^C; see nucleotide sequence comparison) genotype was also obtained. These variable banding patterns also recognized the polymorphic site produced due to a single nucleotide substitution at position +4601 (Pena et al. 2000). The distribution of genotypic and allelic frequencies is presented in Table 4. The frequency of A allele was found to be lower compared to that of the B allele in all the

Table 3. Thermal Cycler parameters

- (1) Initial cycle at 95°C for 5 min Then 35 cycles of steps containing:
- (2) Denaturation at 95°C for 30 s Annealing at 64°C for 1 min Extension at 72°C for 90 s
- (3) Final extension at 72°C for 5 min

Amplified products were visualized through agarose gel electrophoresis (2%)

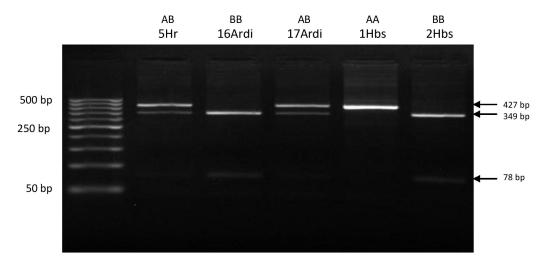


Figure 1. RFLP electrophoretic patterns of β -LG amplicons generated by Sac II digestion Enzymatic digestion resulted in different restriction patterns or genotypes across the studied goat breeds. Lane 1: 50 bp DNA ladder, lanes 2–6: 5Harri, 16Ardi, 17Ardi, 1Habsi and 2Habsi

studied breeds, and in close agreement to the data presented earlier by Kumar et al. (2006) in Indian goats. The distribution patterns of β -LG genotypes exhibited that 48% of all the investigated animals showed as homozygous BB, 38% as heterozygous AB, and only 13% as homozygous AA.

Preliminary data has been presented in goats regarding β -LG genotypic variants and associated disparities in milk production traits (Kumar et al. 2006; El-Hanafy et al. 2010), though, no conclusive evidence is available regarding similar affects in sheep and cattle (Botaro et al. 2008; Kawecka and Radko 2011). As obvious from Table 4, Habsi breed with the highest presence of A allele showed significantly higher milk yield (138.26 ± 1.26 l) than Ardi (132.11 ± 1.08 l) and Habsi (131.32 ± 1.08 l) breeds (P < 0.05). Table 5 depicts maximum recorded values for milk yield at 16 weeks of lactation in the β -LG AA genotype, i.e. 146.82 \pm 0.21 l in Ardi, 152.75 ± 0.25 l in Habsi, and 146.97 ± 0.25 l in Harri breed (P < 0.05) compared to β -LG AB genotype, where estimated yields stood at 141.01 ± 0.9 l in Ardi, 143.48 ± 0.65 l in Habsi, and 141.20 ± 0.92 l in Harri breed (P<0.05). Minimal figures were obtained in the β -LG BB genotype, i.e. 122.99 ± 0.65 l in Ardi, 123.06 ± 0.81 l in Habsi, and 123.01 ± 0.62 l in Harri breed (P<0.05). The results presented in this study are in close agreement with the data generated by Kumar et al. (2006) in Indian goats and El-Hanafy et al. (2010) in Egyptian goat breeds.

Nucleotide sequence comparison. Nucleotide sequencing of the selected β -LG fragments based on genotypic data was completed and following successful BLAST analysis (http://www.ncbi.nlm. nih.gov/BLAST/) submitted to GenBank NCBI (Accession Nos. KJ544248, KJ588275, KJ588276, KJ783455, KJ783456, KJ874959, and KP269078). Comparison with already published sequence of the goat β -LG gene (NCBI Accession No. Z33881) via ClustalW revealed complete similarity except SNPs at specific points within fragments designated as AA genotype (NCBI Accession Nos. KJ544248, KJ783456, KP269078, and KJ874959) (Figure 2).

Table 4. Genotypic and allelic frequencies following *Sac II* enzymatic digestion of the amplified β -*LG* fragments and estimated milk yields for the studied goat breeds

Duood		Ge	notypic freque	ency	Allelic fr	requency	Milk yield (kg/16 weeks)	
Breed	п -	AA	AB	BB	A	В	(mean ± SE)	
Ardi	100	0.08	0.4	0.52	0.28	0.72	132.11 ± 1.08 ^a	
Habsi	100	0.23	0.41	0.36	0.43	0.57	138.26 ± 1.26^{b}	
Harri	100	0.09	0.34	0.57	0.26	0.74	131.32 ± 1.08^{a}	

 $^{^{}a,b}P < 0.05$

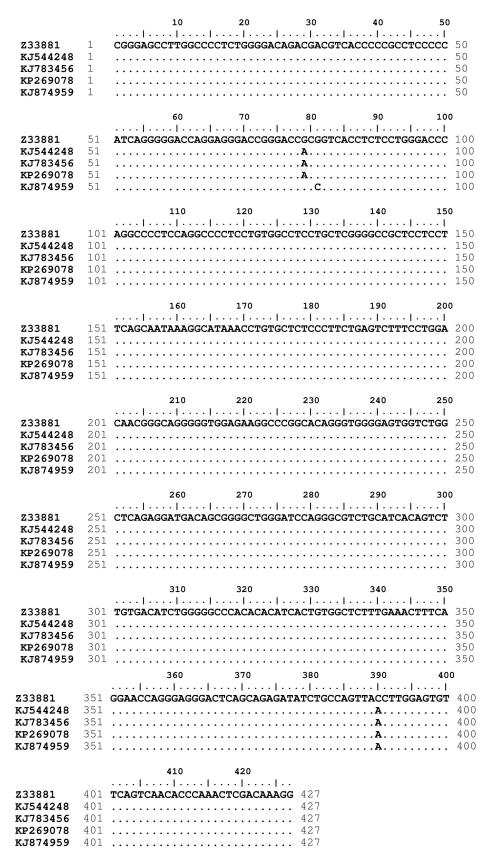


Figure 2. Nucleotide sequence comparison of the β -LG fragments (NCBI Accession Nos. KJ544248, KJ783456, KP269078, and KJ874959) with GenBank reference sequence of the goat β -LG gene (NCBI Accession No. Z33881)

Table 5. Milk yield (kg/16 weeks) for designated genotypes within the studied goat breeds

Construe		Milk yield (kg/16 weeks)	
Genotype	Ardi	Habsi	Harri
\overline{AA}	146.82 ± 0.21 ^a	152.75 ± 0.25^{a}	146.97 ± 0.25^{a}
AB	141.01 ± 0.9^{b}	143.48 ± 0.65^{b}	141.20 ± 0.92^{b}
BB	$122.99 \pm 0.65^{\circ}$	$123.06 \pm 0.81^{\circ}$	123.01 ± 0.62^{c}

a-cP < 0.05

SNPs at position 79 (+4601) (G/A) (NCBI Accession Nos. KJ544248, KJ783456, and KP269078) and 81 (+4603) (G/C) (NCBI Accession No. KJ874959) were found in exon 7 and one at position 390 (+4912) in the flanking 3' UTR region (C/A) (NCBI Accession No. KJ544248, KJ783456, KP269078, and KJ874959). The exonic SNPs, i.e. +4601 (G/A)and +4603 (G/C), were present within the Sac II restriction site and accountable for generating the AA genotypic patterns. As a consequence, the allele characterized by the substitution G>A has been sub-designated as AA^A , while the one characterized by the substitution G>C as AA^C . Frequency of the AAA allele in the analyzed sequences was 0.925 while that of AA^{C} allele was 0.075. No simultaneous existence was recorded for these exonic SNPs in any of the examined sequences. The +4601 (7E+83) nucleotide substitution has been established earlier in Spanish and French goats by Pena et al. (2000). The incidence of the tandemly repeated 10 bp (CCAGGCCCCT) sequence, specific to ruminants only, has also been recorded (NCBI Accession Nos. KJ544248, KJ783456, KP269078, and KJ874959). Pena et al. (2000) named this polymorphic site as S₂I₂ variant due to the presence of the I_2 allele. However, an additional duplication of the inserted 10 bp sequence (the S₂I₃ variant) has not been recognized in any of the studied animals. Kumar et al. (2006) examined these alterations at both DNA and protein levels in Indian breeds and reported that β -LG AA genotype had a higher milk yield than β -LG AB genotype. The +4603 (G/C) nucleotide substitution has been recognized by Jain et al. (2012) in Indian goat breeds, though they did not elucidate associations, if any, with the utility traits. Two similar polymorphic sites (+4601 and +4603) in the exonic region of Saudi goats imply a selection pressure within this coding region. The polymorphisms in the coding region did not produce any amino acid change in the protein. The +4912 nucleotide substitution within the 3' UTR region is a novel SNP. The 3' non-coding region has been established to control stabilizing factors on the mRNA (Bellasco and Brawerman 1993). Exon 7 comprises most of the 3' UTR region on the β -LG mRNA (Folch et al. 1994). Hence, intronic SNPs can influence regulatory switches controlling translation of the β -LG transcript.

CONCLUSION

The exon 7 to 3' flanking region of β -LG gene of C. hircus was analyzed to detect novel SNPs and to present a baseline data for Saudi goats. Two already annotated SNPs in exon 7 (Pena et al. 2000; Kumar et al. 2006; Jain et al. 2012) and one fresh SNP within the 3' UTR region have been detected. Based on this first hand data, further analyses are warranted as exploring such mutations for the phenotypic variation of milk yield could provide a means for improving this trait on commercial scale. SNPs within non-coding regions can also influence gene expression by controlling important regulatory switches. Thus, the identification of variations/variants across the entire gene in the existing gene pool is imperative.

Acknowledgement. The authors acknowledge the assistance from the Science and Technology Unit, Deanship of Scientific Research and Deanship of Graduate Studies, King Abdulaziz University, Jeddah, KSA.

REFERENCES

Arora R., Bhatia S., Mishra B.P., Sharma R., Pandey A.K., Prakash B., Jain A. (2010): Genetic polymorphism of the beta-*lactoglobulin* gene in native sheep from India. Biochemical Genetics, 48, 304–311.

Bellasco J., Brawerman G. (1993): Control of Messenger RNA Stability. Academic Press, San Diego, USA.

Botaro B.G., Lima Y.V., Aquino A.A., Fernandes R.H., Garcia J.F., Santos M.V. (2008): Effect of beta-lactoglobulin poly-

- morphism and seasonality on bovine milk composition. Journal of Dairy Research, 75, 176–181.
- Caroli A.M., Chessa S., Erhardt G.J. (2009): Invited review: Milk protein polymorphisms in cattle, effect on animal breeding and human nutrition. Journal of Dairy Science, 92, 5335–5352.
- Cui Y., Cao Y., Ma Y., Qu X., Dong A. (2012): Genetic variation in the β-*lactoglobulin* of Chinese yak (*Bos grunniens*). Journal of Genetics, 91, 44–48.
- El-Hanafy A.A., El-Saadani M.A., Eissa M., Maharem G.M., Khalifa Z.A. (2010): Polymorphism of β-lactoglobulin gene in Barki and Damascus and their crossbred goats in relation to milk yield. Biotechnology in Animal Husbandry, 26, 1–12.
- Falconer D.S., Mackay T.F.C. (1996): Introduction to Quantitative Genetics. Addison Wesley Longman, Harlow, UK.
- Folch J.M., Coll A., Sanchez A. (1994): Complete sequence of the caprine β -*lactoglobulin* gene. Journal of Dairy Science, 77, 3493–3497.
- Hambraeus L., Lonnerdal B. (2003): Nutritional aspects of milk proteins. In: Fox P.F. and McSweeney P.L.H. (eds): Advanced Dairy Chemistry. Volume I: Proteins. Kluwer Academic/Plenum Publishers, New York, USA, 605–645.
- Jain A., Gour D.S., Bisen P.S., Dubey P.P., Sharma D.K., Tiwari R.P., Gupta N., Kumar D. (2012): Allele mining in β-lactoglobulin gene of Capra hircus. African Journal of Biotechnology, 11, 11057–11064.
- Kawecka A., Radko A. (2011): Genetic polymorphism of β-*lactoglobulin* in sheep raised for milk production. Journal of Applied Animal Research, 39, 68–71.
- Kontopidis G., Holt C., Sawyer L. (2004): Invited review: β-*Lactoglobulin*: binding properties, structure, and function. Journal of Dairy Science, 87, 785–796.
- Kumar A., Rout P.K., Roy R. (2006): Polymorphism of β-*lacto globulin* gene in Indian goats and its effect on milk yield. Journal of Applied Genetics, 47, 49–53.
- MacHugh D.E., Bradley D.G. (2001): Livestock genetic origins: goats buck the trend. Proceedings of the National Academy of Sciences of the United States of America, 98, 5382–5384.

- Pena R.N., Sanchez A., Folch J.M. (2000): Characterization of genetic polymorphism in the goat beta-*lactoglobulin* gene. Journal of Dairy Research, 67, 217–224.
- Qureshi M.I., Sabir J.S.M., Mutawakil M.H.Z., El Hanafy A.A., Ashmaoui H.E., Ramadan H., Anwar Y., Sadek A.M., Alsoud M.A., Saini K.S., Ahmed M.M. (2014): Review of modern strategies to enhance livestock genetic performance: from molecular markers to next-generation sequencing technologies in goats. Journal of Food, Agriculture and Environment, 12, 752–761.
- Ron N., Zimet P., Bargarum J., Livney Y.D. (2010): Betalactoglobulin–polysaccharide complexes as nanovehicles for hydrophobic nutraceuticals in non fat foods and clear beverages. International Dairy Journal, 20, 686–693.
- Sabir J.S.M., Mutawakil M.H.Z., El-Hanafy A.A., Ahmed M.M.M. (2012): Genetic similarity among four breeds of goats in Saudi Arabia detected by random amplified polymorphic DNA marker. African Journal of Biotechnology, 11, 3958–3963.
- Sabir J.S.M., Sabry A.M., Awad N.S., Mohamed A.A., Mutawakil M.H.Z. (2013): Najdi, Harri and Aradi Saudi goat breeds possess genetic variation required for genetic improvement. World Applied Sciences Journal, 26, 867–872.
- Sanger F., Nicklen S., Coulson A.D. (1977): DNA sequencing with chain terminating inhibitor. Proceedings of the National Academy of Sciences of the United States of America, 74, 5436–5467.
- Wang Q.W., Allen J.C., Swaisgood H.E. (1997): Binding of vitamin D and cholesterol to β-lactoglobulin. Journal of Dairy Science, 80, 1054–1059.
- Zeder M.A., Hesse B. (2000): The initial domestication of goats (*Capra hircus*) in the Zagros Mountains 10,000 years ago. Science, 287, 2254–2257.
- Zimet P., Livney Y.D. (2009): Beta-lactoglobulin and its nano-complexes with pectin as vehicles for ω -3 polyunsaturated fatty acids. Food Hydrocolloids, 23, 1120–1126.

Received: 2014–07–03 Accepted after corrections: 2014–12–17

Corresponding Author

Muhammad Iqbal Qureshi, Research Assistant, King Abdulaziz University, Faculty of Science, Department of Biological Sciences, Section of Animal Biotechnology, P.O. Box 80203, Jeddah, 21589, Kingdom of Saudi Arabia Phone: +966 590 834 299, e-mail: miq.qureshi@gmail.com