Journal of Dairy Research (2018) **85** 133–137. © *Hannah Dairy Research Foundation 2018*
doi:10.1017/S0022029918000195

Molecular characterisation of the buffalo *SCAP* gene and its association with milk production traits in water buffaloes

Tingxian Deng, Xiaoya Ma, Chunying Pang, Shasha Liang, Xingrong Lu, Anqin Duan and Xianwei Liang*

Guangxi provincial Key Laboratory of Buffalo Genetics, Breeding and Reproduction technology, Buffalo Research Institute, Chinese Academy of Agricultural Sciences, Nanning 530001, China

Received 5 April 2017; accepted for publication 2 April 2018

The study reported in this Research Communication was conducted to investigate the molecular characterisation of buffalo *SCAP* gene, expression analysis, and the association between single nucleotide polymorphisms and milk production traits in 384 buffaloes. Sequence analysis revealed the *SCAP* gene had an open reading frame of 3837 bp encoding 1279 amino acids. A ubiquitous expression profile of *SCAP* gene was detected in various tissues with extreme predominance in the mammary gland during early lactation. Moreover, eleven SNPs in buffalo *SCAP* gene were identified, six of them (g.1717600A>G, g.1757922C>T, g.1758953G>A, g.1759142C>T, g.1760740G>A, and g.1766036T>C) were found to be significantly associated with 305-day milk yield. Thus, buffalo *SCAP* could sever as a candidate gene affecting milk production traits in buffalo and the identified SNPs might potentially be genetic markers.

Keywords: Buffalo, milk production traits, *SCAP*, single nucleotide polymorphism..

SREBP cleavage-activating protein (*SCAP*) is a family of endoplasmic reticulum (ER) membrane-bound transcription factors required for cleavage and activation of sterol regulatory element-binding proteins (SREBPs), which play a critical role in the regulation of cholesterol homeostasis (Weber et al. 2004). To date, *SCAP* gene has been cloned and characterised from several vertebrate (Qiu et al. 2006) and non-vertebrate species (Seegmiller et al. 2002). Several studies have confirmed that the single nucleotide polymorphisms (SNPs) in *SCAP* might affect intracellular lipid metabolism (Iwaki et al. 1999), metabolic phenotypes (Cao et al. 2002), and obesity (Liu et al. 2014). Dixit et al. (2015) reported that nine SNPs in *SCAP* gene had a significant effect on one or a group of fatty acids including conjugated linoleic acid in goat milk. However, there have not been any reports concerning the characterisation of buffalo *SCAP* gene and the identification of SNPs associated with milk production traits. Thus, we report the full coding sequences of buffalo *SCAP* gene, the tissue expression profiles, and the association between putative SNPs and milk production traits in Chinese buffaloes.

*For correspondence; e-mail: liangbri@126.com

Materials and methods

All animal procedures and the study design were approved by the Animal Ethics Committee of Buffalo Research Institute, Chinese Academy of Agricultural Sciences (BRI-CAAS).

Animal resource

Genomic DNA was isolated from blood samples taken from 384 crossbred buffaloes. These animals were randomly chosen from a farm belonging to the BRI-CAAS, and were a contemporaneous group. Records of milk traits including 305-day milk yield (MY), protein percentage (PP) and fat percentage (FP) were collected at the BRI-CAAS during a 6-year period (January 2010 to December 2015). Genomic DNA was extracted from the blood samples using TIANamp Blood DNA Kit (Tiangen Biotech (Beijing) Co., Ltd., China). The quality and purity of isolated DNA was evaluated using NanoDrop 2000 Spectrophotometer (Thermo Scientific, USA).

Ten healthy crossbred buffaloes were selected and sampled in the experiment. One buffalo was slaughtered by exsanguination and then fresh tissue samples were collected of heart, lung, kidney, epithelium, spleen, ovary, mammary gland, brain, large intestine, small intestine, rumen, muscle, uterus, lymph and liver. Mammary gland

https://doi.org/10.1017/S0022029918000195 Published online by Cambridge University Press

samples from nine individuals at different stages of lactation (Early lactation, $n = 3$; Peak lactation, $n = 3$; Late lactation, $n = 3$) were sampled by biopsy method. All the tissues were immediately frozen in liquid nitrogen and stored at −80 °C before processing for RNA isolation.

Isolation of buffalo SCAP cDNA sequence

Total RNA was isolated from the collected tissues using TRIzol® Reagent (Invitrogen, USA). Seven pairs of PCR primers (Supplementary Table S1) were designed and used to amplify the buffalo *SCAP* cDNA sequence based on bovine *SCAP* (GenBank accession number: NM_001101889.1). PCR products were amplified by RT-PCR using PrimeScript™ One Step RT-PCR Kit Ver2.0 (Takara Biotechnology (Dalian) Co., Ltd., China) under the procedure: 5 min at 94 °C, followed by 35 cycles of 30 s at 94 °C, 30 s at 60 °C, and 1 min at 72 °C and a final extension of 10 min at 72 °C. All PCR products were cloned into the pMD-19 T vectors (Takara Biotechnology (Dalian) Co., Ltd., China) and used for the sequence analysis (BGI, China).

Bioinformatics analysis of buffalo SCAP gene

Similarity searches of the sequenced cDNA of *SCAP* were conducted using the National Center for Biotechnology Information (NCBI) BLAST tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Multiple-sequence alignments based on the amino acid sequences were performed using ClustalW on MEGA v7.0. Phylogenetic tree was constructed using the neighbor-joining (NJ) method based on the Poisson-corrected distance with 1000 bootstraps.

Physical and chemical properties of the putative *SCAP* protein were predicted using ExPASy server (http://www.expasy.org/). Secondary structure of buffalo *SCAP* was predicted and analysed by using the Protter1·0 servers (http://wlab.ethz.ch/protter/#) and SOPMA software (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_seccons.html). 3D structures and precise substrate binding sites were predicted using the Phyre2 server (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index).

Gene expression assays of buffalo SCAP gene

Gene expression assays were conducted to further confirm the potential function of buffalo *SCAP* within different tissues and different stages of lactation, respectively. One pair of primers (Supplementary Table S1) was used for gene expression analysis using real-time quantitative PCR (qRT-PCR). The qRT-PCR used SYBR green fluorescence and was run in Light Cycler 480 II Instrument (Roche, Germany). PCR condition was 95 °C for 5 min; 45 cycles of 95 °C for 10 s, 60 °C for 20 s, 72 °C for 20 s; 95 °C for 5 s, 65 °C for 1 min, 97 °C continuous, 1 cycles; 40 °C for 10 s. All samples were amplified in triplicate and the relative gene expression was normalised by the *GAPDH* with the $2^{-\triangle\triangle Ct}$ method.

SNP screening and genotyping

Fourteen pairs of primers (Supplementary Table S1) were used for screening the putative SNP loci based on the predicted buffalo *SCAP* gene sequence (GenBank accession number: NW_005785163.1, region: 1673538 to1766641). Five buffalo samples were randomly chosen to detect the polymorphism of *SCAP* gene by pooled DNA-sequencing. PCR reaction, PCR procedure, and mutation search was performed as described previously Deng et al. (2016). The identified SNPs were genotyped in 384 individuals using iPLEX MassARRAY system (Sequenom's MassARRAY) by Guangzhou Life Technologies Company.

Statistical analysis

R language with genetics package was used to estimate the genetic parameters for each locus. The following statistical model was used to calculate the marker-trait association analysis by SAS version 9·4 software (SAS Institute Inc., Cary, NC) with the PROC MIXED procedure: $Y_{ijkl} = \mu + GC_i + S_j + M_k + \beta_1(I_{ijkl} - \bar{I}) + \beta_2(I_{ijkl} - \bar{I}) + e_{ijkl}$, where Y_{ijkl} = trait observation; μ = overall mean, GC_i = fixed effect of contemporary group consist of the combination of year and calving season; S_j = the random effect of sire; M_k = the fixed effect of genotype for SNP; β_1 = linear regression coefficient of trait Y_{ijkl} in relation to the testing age of buffalo; β_2 = quadratic regression coefficient of Y_{ijkl} in relation to the testing age of buffalo; I_{ijkl} = the testing age of buffalo; $\bar{I}$ = the testing mean age of the buffalo; and e_{ijklm} = the random residual with mean 0 and variance σ_e^2. The least square means with standard error for multiple comparison between different genotypes were performed by Bonferroni correction for multiple *F*-testing.

Results and discussion

Characterisation of buffalo SCAP gene

Buffalo *SCAP* cDNA sequence was 4135 bp in length and has been submitted to the NCBI (GenBank accession number: KX951421.1), including a 5'-untranslated region (UTR) of 50 bp, partial 3'-UTR of 95 bp, and an open reading frame (ORF) of 3837 bp encoded 1279 amino acids (Supplementary Fig. S1). Similar to the known *SCAP* protein, the buffalo *SCAP* has three structure domains: SSD (residues 308–457) with seven transmembrane helices, a 7WD40 motif (residues 308–457), and a WD40 domain (residues 1078–1235) (Supplementary Fig. S1). Previously, several studies have confirmed that the SSD was a cluster, around 180 residues long, of five membrane-spanning segments that are involved in sterol homeostasis and cell signalling (Luskey & Stevens, 1985), together with a WD40 domain consisting of more than 500 residues that could interact with the C-terminal regulatory domain of SREBPs to achieve the escort function (Stirnimann et al. 2010). These results suggested that the buffalo *SCAP* had

https://doi.org/10.1017/S0022029918000195 Published online by Cambridge University Press

similar functions to mammalian *SCAP* involved in the SREBP pathway contributing to lipid and cholesterol homeostasis.

Buffalo *SCAP* displayed high homology with other mammals, and their similarity ranged between 75 and 99% (Supplementary Table S2), implying that *SCAP* protein was greatly conserved within these species. The buffalo *SCAP* was shown to belong to the *Bovidae* with high bootstrap support (Supplementary Fig. S2). Buffalo *SCAP* shared the closest genetic relationship with cow, followed by sheep, goat, and pig that formed a separate group, while the non-artiodactyla species formed an even more distant group. These results suggest that *SCAP* protein is greatly conserved within these species.

The predicted analysis of secondary structure revealed that buffalo *SCAP* contains seven transmembrane helices and five putative N-glycosylation sites shaded in red (Supplementary Fig. S3a). The 3D structure analysis indicated buffalo *SCAP* protein contains 317·95 (25%) α-helices and 422·07(33%) β-strands (Supplementary Fig. S3c), which is similar to that of SOPMA analysis (Supplementary Fig. S3b). Eleven binding site residues were found to be located in the three α-helices of buffalo *SCAP*, that is, Ile-356, Glu-359, Phe-430, Leu-431, Gln-432, Phe-435/436, Thr-437, Val-439, Leu-440, and Asp-443 (Supplementary Fig. S3d).

Expression analysis of the buffalo SCAP gene

To further confirm the predicted function of buffalo *SCAP*, the relative mRNA expression levels were assessed by qRT-PCR. As shown in Fig. 1a, *SCAP* was highly expressed in the mammary gland, followed by the spleen and lung, but expression in other tissues was low. This result showed that the expression of buffalo *SCAP* may be related to milk fat synthesis and regulation of the SREBP pathway to increase milk production. Moreover, the buffalo *SCAP* gene had higher expression levels at early lactation compared with both peak and late lactations (Fig. 1b), meaning that the trend of *SCAP* expression was descendent in the lactation period. Aspilcueta-Borquis et al. (2010) reported the trend was similar to the typical lactation curve in Murrah buffalo, with higher means in early lactation. These trends are similar to those reported in Holstein cows (Swalve, 1995).

SNP screening and genotyping

Eleven SNP loci (g.1717600A>G, g.1718168G>A, g.1742850G>A, g.1753880A>G, g.1757922C>T, g.1758953G>A, g.1759116G>A, g.1759142C>T, g.1760740G>A, g.1762368A>G, and g.1766036T>C) were identified in buffalo *SCAP* using pooled DNA-sequencing, which has been submitted to the Single Nucleotide Polymorphism (dbSNP) database (Accession No. ss# 2137493882, 2137493883, 2137493884, 2137493888, 2137493889, 2137493890, 2137493891, 2137493892, 2137493893, 2137493894, and 2137493897). Here, only the SNP g.1766036T>C was found within the exon that belonged to the synonymous substitution, and the remaining SNPs were located in the intron region (Supplementary Table S3). Genotypic and allelic frequencies were also listed in Supplementary Table S3. All the identified SNPs were in accordance with Hardy-Weinberg Equilibrium (HWE) (χ^2 test, $P > 0{\cdot}01$) that could be used for further association analysis in the population.

Analysis of buffalo SCAP SNPs with milk traits

Mixed model was used to estimate the association between *SCAP* polymorphisms and milk traits in the studied population. Overall, a total of six loci in buffalo *SCAP* (g.1717600A>G, g.1757922C>T, g.1758953G>A, g.1759142C>T, g.1760740G>A, and g.1766036T>C) were found to be significantly ($P < 0{\cdot}05$) associated with MY (Table 1). Interestingly, the homozygous buffaloes with the above six loci had significantly higher milk yield

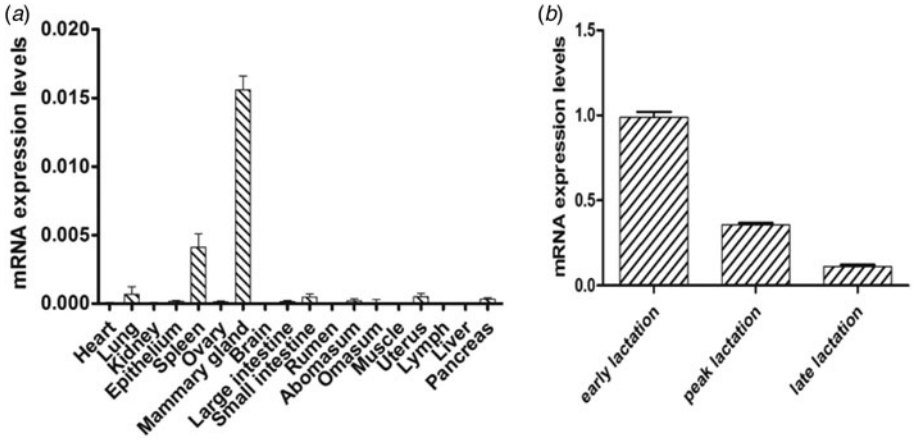

Fig. 1. Relative quantification of the buffalo SCAP gene in different tissues. (a) gene expression profile of *SCAP* in the collected tissues. (b) mRNA analysis of the buffalo *SCAP* gene in mammary gland at three lactations.

https://doi.org/10.1017/S0022029918000195 Published online by Cambridge University Press

Table 1. Association analysis between genotypes of 11 SNPs and three traits values in buffaloes

Loci	Genotype(n)	MY (kg)	PP (%)	FP (%)
g.1717600A>G	AA(125)	1393·65 ± 108·87a,b	4·44 ± 0·08	7·72 ± 0·31
	GA(59)	1088·08 ± 122·30a	4·52 ± 0·09	8·01 ± 0·35
	GG(9)	1754·67 ± 333·62a	4·57 ± 0·25	7·17 ± 0·96
	P value	0·0267	0·6293	0·5933
g.1718168G>A	AA(122)	1377·78 ± 110·44	4·45 ± 0·08	7·78 ± 0·28
	AG(62)	1148·27 ± 122·04	4·50 ± 0·09	8·00 ± 0·34
	GG(8)	1775·48 ± 336·84	4·57 ± 0·25	7·73 ± 0·77
	P value	0·0827	0·8148	0·4820
g.1742850G>A	AA(6)	1324·39 ± 365·91	4·38 ± 0·24	8·64 ± 1·00
	AG(53)	1372·74 ± 140·72	4·41 ± 0·10	7·32 ± 0·38
	GG(134)	1298·68 ± 105·67	4·48 ± 0·07	7·95 ± 0·29
	P value	0·8568	0·7528	0·1401
g.1753880A>G	AA(81)	1429·03 ± 124·54	4·48 ± 0·09	7·96 ± 0·34
	AG(71)	1202·92 ± 131·46	4·44 ± 0·09	7·73 ± 0·36
	GG(48)	1286·36 ± 136·30	4·49 ± 0·10	7·71 ± 0·37
	P value	0·2789	0·8941	0·7885
g.1757922C>T	CC(123)	1385·53 ± 110·59a	4·44 ± 0·08	7·74 ± 0·31
	TC(61)	1115·44 ± 122·07a	4·51 ± 0·09	8·06 ± 0·35
	TT(9)	1766·59 ± 338·22a,b	4·56 ± 0·25	7·21 ± 0·96
	P value	0·0472	0·7017	0·5622
g.1758953G>A	AA(13)	1394·87 ± 216·19a	4·50 ± 0·16	7·25 ± 0·60
	GA(53)	1107·61 ± 123·23a	4·46 ± 0·09	8·03 ± 0·34
	GG(132)	1454·42 ± 112·49a,b	4·47 ± 0·09	7·75 ± 0·31
	P value	0·0261	0·9691	0·4267
g.1759116G>A	AA(9)	1747·97 ± 330·26	4·45 ± 0·09	7·22 ± 0·96
	AG(62)	1130·36 ± 119·05	4·50 ± 0·10	7·98 ± 0·35
	GG(120)	1338·03 ± 107·80	4·57 ± 0·24	7·80 ± 0·31
	P value	0·0902	0·8409	0·7111
g.1759142C>T	CC(123)	1367·85 ± 110·43b	4·45 ± 0·09	7·72 ± 0·31
	CT(62)	1173·50 ± 122·77a	4·50 ± 0·10	8·05 ± 0·35
	TT(9)	2032·75 ± 310·60a,b	4·57 ± 0·24	6·99 ± 0·88
	P value	0·0244	0·8332	0·4244
g.1760740G>A	AA(84)	1204·98 ± 141·49a	4·49 ± 0·07	7·87 ± 0·39
	GA(74)	1203·79 ± 119·93a	4·44 ± 0·09	7·84 ± 0·33
	GG(84)	1540·80 ± 126·07a,b	4·48 ± 0·10	7·74 ± 0·35
	P value	0·0218	0·8741	0·9457
g.1762368A>G	AA(72)	1460·05 ± 129·18	4·48 ± 0·10	7·99 ± 0·36
	GA(85)	1322·37 ± 115·94	4·44 ± 0·09	7·56 ± 0·32
	GG(41)	1142·49 ± 142·12	4·52 ± 0·11	8·08 ± 0·40
	P value	0·1261	0·8000	0·3663
g.1766036T > C	CC(42)	1251·95 ± 139·52b	4·52 ± 0·11	7·88 ± 0·39
	CT(74)	1164·77 ± 120·35a	4·43 ± 0·09	7·78 ± 0·34
	TT(84)	1561·58 ± 125·93a,b	4·47 ± 0·10	7·81 ± 0·35
	P value	0·0103	0·7656	0·9752

Notes: Values with different superscripts within the same column in a particular population differ significantly at $P < 0·05$ or $P < 0·01$.

compared with the heterozygous individuals. However, no significant association was discovered between these SNPs and fat or protein percentages in the population ($P > 0·05$). Moreover, no significant association between the haplotype combinations and milk production traits was found in the studied population (data not shown).

SCAP gene involved in the SREBP pathway is pivotal in the regulation of milk fat synthesis in animals (Bionaz & Loor, 2008). Rincon et al. (2012) reported two quantitative trait loci (QTLs) in bovine *SCAP* associated with the monounsaturated fatty acid content and saturated fatty acid content. In the present study, six SNPs in buffalo *SCAP* gene had significant genetic effects on milk traits ($P < 0·05$). For all identified SNPs, the buffaloes with homozygous genotypes showed higher milk yield than the ones with the heterozygotes after Bonferroni test analysis, implying that these individuals showed no heterotic effect in population. In this case, we assumed that the reason for no heterotic effect of the identified SNPs in the selected crossbred buffaloes is that the current Chinese crossbred buffalo has undergone half a century of the blood purification, and could be considered to be close to purebred individuals.

https://doi.org/10.1017/S0022029918000195 Published online by Cambridge University Press

Moreover, no significant association was discovered between these SNPs and fat or protein percentages, which is inconsistent with the results described by Dixit et al. (2015). Although the expression analysis of *SCAP* was demonstrated to be related to the milk fat synthesis, no significant SNPs have been discovered in the population. The reason for that may be caused by the sample size and sample differences. Therefore, further study of the association analysis is needed to be performed in a larger population.

In conclusion, we reported the characterisation and expression analysis of buffalo *SCAP*. Buffalo *SCAP* showed close homology to mammalian *SCAP* sequences and displayed high expression levels in mammary gland, especially in early lactation. Moreover, eleven SNPs in *SCAP* were identified, six of which had significant genetic effects on milk yield. These findings provide important information for understanding the genetic effects of *SCAP* on milk production traits that could be used for marker-assisted selective breeding in buffaloes.

Supplementary material

The supplementary material for this article can be found at https://doi.org/10.1017/S0022029918000195

This work was supported by grants from the National Natural Science Foundation (No. 31660649), Natural Science Foundation of Guangxi Province, China (No.2015GXNSFBA139103) and Guangxi Program of Science and Technology Development (1504001-3).

References

Aspilcueta-Borquis RR, Sesana RC, Berrocal MH, de Oliveira Seno L, Bignardi AB, Faro LE, de Albuquerque LG, de Camargo GM & Tonhati H 2010 Genetic parameters for milk, fat and protein yields in Murrah buffaloes (*Bubalus bubalis* Artiodactyla, Bovidae). *Genetics and Molecular Biology* **33** 71–77

Bionaz M & Loor JJ 2008 Gene networks driving bovine milk fat synthesis during the lactation cycle. *BMC Genomics* **9** 366

Cao H, Miskie BA & Hegele RA 2002 Functional promoter polymorphism in SREBP cleavage-activating protein (SCAP). *Journal of Human Genetics* **47** 492–496

Deng T, Pang C, Ma X, Lu X, Duan A, Zhu P & Liang X 2016 Four novel polymorphisms of buffalo INSIG2 gene are associated with milk production traits in Chinese buffaloes. *Molecular and Cellular Probes* **30** 294–299

Dixit SP, Sivalingam J, Tyagi AK, Saroha V, Sharma A & Nagda RK 2015 Association of novel SNPs in the candidate genes affecting caprine milk fatty acids related to human health. *Meta Gene* **4** 45–56

Iwaki K, Nakajima T, Ota N & Emi M 1999 A common Ile796Val polymorphism of the human SREBP cleavage-activating protein (SCAP) gene. *Journal of Human Genetics* **44** 421–422

Liu FH, Song JY, Shang XR, Meng XR, Ma J & Wang HJ 2014 The gene-gene interaction of INSIG-SCAP-SREBP pathway on the risk of obesity in Chinese children. *BioMed Resarch International* **2014** 538564

Luskey KL & Stevens B 1985 Human 3-hydroxy-3-methylglutaryl coenzyme A reductase. Conserved domains responsible for catalytic activity and sterol-regulated degradation. *Journal of Biological Chemistry* **260** 10271–10277

Qiu H, Xia T, Chen X, Zhao X, Gan L, Feng S, Lei T & Yang Z 2006 Cloning, comparative characterization of porcine SCAP gene, and identification of its two splice variants. *Molecular Genetics and Genomics* **276** 187–196

Rincon G, Islas-Trejo A, Castillo AR, Bauman DE, German BJ & Medrano JF 2012 Polymorphisms in genes in the SREBP1 signalling pathway and SCD are associated with milk fatty acid composition in Holstein cattle. *Journal of Dairy Research* **79** 66–75

Seegmiller AC, Dobrosotskaya I, Goldstein JL, Ho YK, Brown MS & Rawson RB 2002 The SREBP pathway in drosophila: regulation by palmitate, not sterols. *Developmental Cell* **2** 229–238

Stirnimann CU, Petsalaki E, Russell RB & Muller CW 2010 WD40 proteins propel cellular networks. *Trends in Biochemical Science* **35** 565–574

Swalve HH 1995 The effect of test day models on the estimation of genetic parameters and breeding values for dairy yield traits. *Journal of Dairy Science* **78** 929–938

Weber LW, Boll M & Stampfl A 2004 Maintaining cholesterol homeostasis: sterol regulatory element-binding proteins. *World Journal of Gastroenterol* **10** 3081–3087

https://doi.org/10.1017/S0022029918000195 Published online by Cambridge University Press