Check for updates

Genome-wide analysis of the acyl-coenzyme A synthetase family and their association with the formation of goat milk flavour

Fuhong Zhang, Jun Luo*, Chenbo Shi, Lu Zhu, Qiuya He, Huibin Tian, Jiao Wu, Jianqing Zhao and Cong Li*

Key Laboratory of Animal Genetics, Breeding and Reproduction of Shaanxi Province, College of Animal Science and Technology, Northwest A&F University, Yangling, China

Goat milk is rich in fat and protein, thus, has high nutritional values and benefits human health. However, goaty flavour is a major concern that interferes with consumer acceptability of goat milk and the 4-alkyl-branched-chain fatty acids (vBCFAs) are the major substances relevant to the goaty flavour in goat milk. Previous research reported that the acyl-coenzyme A synthetases (ACSs) play a key role in the activation of fatty acids, which is a prerequisite for fatty acids entering anabolic and catabolic processes and highly involved in the regulation of vBCFAs metabolism. Although ACS genes have been identified in humans and mice, they have not been systematically characterized in goats. In this research, we performed genome-wide characterization of the ACS genes in goats, identifying that a total of 25 ACS genes (without *ACSM2A*) were obtained in the *Capra hircus* and each ACS protein contained the conserved AMP-binding domain. Phylogenetic analysis showed that out of the 25 genes, 21 belonged to the ACSS, ACSM, ACSL, ACSVL, and ACSBG subfamilies. However, *AACS*, *AASDH*, *ACSF*, and *ACSF3* genes were not classified in the common evolutionary branch and belonged to the ACS superfamily. The genes in the same clade had similar conserved structures, motifs and protein domains. The expression analysis showed that the majority of ACS genes were expressed in multi tissues. The comparative analysis of expression patterns in non-lactation and lactation mammary glands of goat, sheep and cow indicated that *ACSS2* and *ACSF3* genes may participate in the formation mechanisms of goaty flavour in goat milk. In conclusion, current research provides important genomic

resources and expression information for ACSs in goats, which will support further research on investigating the formation mechanisms of the goaty flavour in goat milk.

Introduction

The ACS family comprises a large and diverse group of enzymes. Each member of the ACS family contains a highly conserved amino acid sequence motif, an ATP/AMP binding domain (Black et al., 1997; Mashek et al., 2004). This motif locates at 200–300 amino acids from the N-terminus and is the marker of adenylate-forming enzymes (Watkins et al., 2007; Watkins and Ellis, 2012). In humans, the ACS gene family contains 26 members. Of these members, 22 are subdivided into five subfamilies based on their discrepancies in AMP/ATP and fatty acid-binding motifs (Steinberg et al., 2000). The five subfamilies are the short-chain acyl-CoA synthetase (ACSS), the medium-chain acyl-CoA synthetase (ACSM), the long-chain acyl-CoA synthase (ACSL), the very long-chain synthetase (ACSVL) and the bubblegum ACS synthetase (ACSBG) subfamilies (Grevengoed et al., 2014). Due to their structural features, the acetoacetyl-CoA synthetase (AACS), acyl-CoA synthetase family member 2 (ACSF2), acyl-CoA synthetase family member 3 (ACSF3), and 2-Aminoadipic 6-semialdehyde dehydrogenase (AASDH) genes are not classified into any subfamilies and are independent members of the ACS superfamily (Watkins et al., 2007). HUGO nomenclature advisors suggested name these four genes using the interim designation ACSF (ACSF1–4) family (Watkins et al., 2007). In mammals, the ACS family has been characterized in humans and mice, but not in goats.

Different ACS subfamilies exhibit their own preferences for different length of fatty acids (Rossi Sebastiano and Konstantinidou, 2019), with members of each subfamily showing tissue different expression profiles and subcellular locations (Grevengoed et al., 2014). Fatty acids with less than 6 carbons are typically catalyzed by ACSSs, C6—C10 fatty acids are catalyzed by ACSMs, while C12—C20 fatty acids and very long chain fatty acids (>20 carbons) are preferred by ACSLs and ACSVLs, respectively (Grevengoed et al., 2014; Tang et al., 2018). The long chain fatty acids are predominant fats and fulfil essential physiological functions in living organisms (O'Brien et al., 2020). ACSL family has been widely researched in the past (Ansari et al., 2017). ACSL family members, including ACSL1, ACSL3, ACSL4, ACSL5, and ACSL6, exhibit distinct substrate preferences (Rossi Sebastiano and Konstantinidou, 2019). ACSL1, highly expressed in multi tissues such as liver, kidney heart and muscle (Grevengoed et al., 2014), typically prefers oleate and linoleate (Kanter et al., 2012). The absence of ACSL1

inhibits the sensitivity of macrophages to oleic- and linoleic-mediated degradation of ABCA1 (ATP binding cassette transporter A1), and increase cholesterol spillage (Kanter et al., 2012). ACSL3 has been found in the endoplasmic reticulum and lipid droplets (Grevengoed et al., 2014), which preferentially catalyzes palmitic and arachidonic fatty acids (Ndiaye et al., 2020). ACSL4 has also prominent expressed in multi tissues, mainly adrenal gland, brain, ovary, and testis (Grevengoed et al., 2014), and prefers arachidonic acid (Ndiaye et al., 2020). Moreover, previous research suggests the dysregulated expression of both ACSL3 and ACSL4 is linked to several diseases, especially cancer (Tang et al., 2018). ACSL5 has the highest expression in intestinal mucosa relative to other tissues (Meller et al., 2013). The splice variants in ACSL5 is associated with several types of cancer (Perez-Nunez et al., 2019), and a deletion of ACSL5 can lead to intestinal lipid malabsorption (O'Brien et al., 2020). The expression of ACSL6 is large in brain (Takahiro and Tokuo, 1992; Grevengoed et al., 2014), and the absence of ACLS6 is the most likely cause of the omega-3 docosahexaenoic acid (DHA) deficiency in the brain and spine (Fernandez et al., 2018). Taken altogether, ACSs fulfil distinct roles in fatty acid metabolism.

Goat milk is rich in milk fat and protein, which has high nutritional value and is beneficial to human health (Teng et al., 2018). However, the consumer's acceptance of goat milk and dairy products is restricted because of the perceived characteristic goaty flavour (Kaffarnik et al., 2014). Previous studies have confirmed that 4-alkyl-branched-chain fatty acids (vBCFAs) are the main substances relevant to the goaty flavour in goat milk (Brennand et al., 2010; Salles et al., 2010; Teng et al., 2018). In addition, the concentration of vBCFAs, including 4-ethyloctanoic acid (4-Et-8:0), 4-methyloctanoic acid (4-Me-8:0) and 4-methylnonanoic acid (4-Me-9:0) in goat milk is much higher than in cow milk (Ha and Lindsay, 1993; Kaffarnik et al., 2014). Branched-chain fatty acids (BCFAs) are more likely to be synthesized in goat tissues rather than by rumen microbes (Berthelot et al., 2001). Fatty acids, including vBCFAs, are catalyzed to fatty acyl-CoAs by ACSs before involving in both anabolic and catabolic processes (Steinberg et al., 2000; Watkins et al., 2007; Grevengoed et al., 2014). However, research on the potential role of ACSs in the regulation of vBCFAs metabolism was limited.

Thus, the present study performed a genome-wide characterization of the ACS genes in goat and analyzed their expression profiles in dairy goat, and compared their expression

patterns in the non-lactation and lactation mammary glands of dairy goat, sheep, and cow. The major objective of this study was to investigate the function of ACSs in the formation of goaty flavour in goat milk, thus, to provide gene resources for the genetic improvement of goaty flavour by regulating the ACS-mediated vBCFAs metabolism.

Materials and methods

Genome-wide identification of acyl-coenzyme A synthetase genes

The goat (*Capra hircus*) reference genome (Accession NO. GCA_001704415.1) was downloaded from the Ensemble database (http://ftp.ensembl.org/pub/release-104/fasta/capra_hircus/dna). The ACS protein sequences of *Homo sapiens* were downloaded from the GenBank (https://www.ncbi.nlm.nih.gov/). To identify ACS genes of goat, we used human ACS protein sequences as queries to carry out local BLASTP ($p = 0.001$) searches against the goat genome database. To further improve the accuracy, we acquired the HMM (hidden Markov model) profile of the AMP-binding domain (PF00501.31) from the Pfam database (http://pfam.xfam.org) (Kochan et al., 2009), then searched the sequence of candidate genes identified by blastp, using HMMER 3.3 (http://eddylab.org/software/hmmer/hmmer-3.3.tar.gz) (Finn et al., 2011). Using the same criterion, ACS genes protein sequences of sheep and cow were obtained from *Ovis aries* (Accession NO, GCA_002742125.1) and *Bos taurus* reference genomes (Accession NO. GCA_002263795.2).

Each potential ACS genes of goat was further analyzed by the programs Pfam to identify the location of domains. The "Compute pI/Mw" tool was used to obtain the pI (theoretical isoelectric point) and MW (molecular weight) of identified ACS proteins (https://web.expasy.org/compute_pi/). Subcellular localizations of identified ACS genes were predicted, using WoLF PSORT website (https://wolfpsort.hgc.jp/). The location of ACS genes on goat genome was mapped using an online website (http://mg2c.iask.in/mg2c_v2.1/). Motif prediction of identified ACS proteins was analyzed using the MEME suite with a maximum number of 15 motifs (http://meme-suite.org/tools/meme). The gene structures and motif sequences were drawn using the EvolView online tool (https://www.evolgenius.info/evolview/#login) (Balakrishnan et al., 2019).

Multiple sequence alignment and phylogenetic tree construction

The ACS protein sequences from goat, human, sheep, and cow were analyzed together. Multiple sequence alignment was conducted using the MAFFT-7.429 (Katoh and Standley, 2013).

The results of multiple sequence alignments were used to predict conserved domains and motifs, and to construct phylogenetic trees. The phylogenetic tree was constructed using the IQ-TREE with a ultrafast bootstrap value of 1,000 (Lam-Tung et al., 2015). The phylogenetic tree and locations of conserved domains of ACSs were drawn using the EvolView online tool.

Expression analysis of acyl-coenzyme A synthetase genes

The expression levels of the ACS genes of dairy goat at non-lactation and lactation stages were analyzed based on the transcriptome data that were obtained by our laboratory (NCBI SRA accession: PRJNA637690) (Li et al., 2020). To understand the expression profiles of distinct tissues in dairy goat, transcriptome datasets of three organs (heart, kidney, and liver) and skeletal muscle tissue were downloaded, with 3 biological replicates (NCBI SRA accession: PRJNA309284 and NCBI SRA accession: PRJNA309345) from SRA database (https://www.ncbi.nlm.nih.gov/sra/). To further investigate the distinction in transcriptional responses of ACSs in the mammary glands of goat, sheep and cow, the sheep (NCBI SRA accession: PRJNA309284) and cow (NCBI SRA accession: PRJNA482783) transcriptome datasets of mammary gland in the non-lactation and lactation period were downloaded from SRA database (https://www.ncbi.nlm.nih.gov/sra/). The analysis procedures were described as bellow. The transcriptome datasets were converted to fastq format using fastq-dump. All clean reads were mapped to the respective reference genome sequence using Hisat2-2.1.0 (https://github.com/infphilo/hisat2/), and the transcription-level expression was calculated using StringTie (https://ccb.jhu.edu/software/stringtie/). The FPKM values were $\log_2$ transformed, and the heat map of gene expression levels was plotted using the EvolView online tool.

Results

Characterization of acyl-coenzyme A synthetase genes

To identify ACS genes in goat, ACS protein sequences of human were used as queries to search the goat genome. A total of 25 ACS genes (*ChACSs*) were identified after an analysis of conserved domains in goat. The full-length ChACS protein sequences were given in the Supplementary Sheet S1. We found that the *ACSM2A* gene was absent in goat reference genome, nor in sheep and cows' reference genome. The *ChACSs* shared homology to human ACSs (*HsACSs*), with the amino acid identity of 71.2%–96.4% (Supplementary Sheet S2).

TABLE 1 Information of ACS family genes in goat.

Subfamily	Gene	CDS (bp)	Exon	Intron	Amino acid (aa)	MW (Da)	pI	Localization
ACSBG	ChACSBG1	2,298	17	16	766	85,738.59	6.54	Mitochondrion
	ChACSBG2	1866	14	13	622	69,142.06	7.87	Plasma membrane
ACSL	ChACSL1	2,100	21	20	700	78,128.85	8.00	Cytoplasm
	ChACSL3	2,163	15	14	721	80,127.67	8.65	Cytoplasm
	ChACSL4	2013	16	15	671	74,477.15	8.41	Mitochondrion
	ChACSL5	2052	21	20	684	75,532.52	7.50	Endoplasmic reticulum
	ChACSL6	2,169	21	20	723	80,846.28	6.14	Endoplasmic reticulum
ACSM	ChACSM1	1734	14	13	578	64,819.43	7.84	Mitochondrion
	ChACSM2B	1734	15	14	578	64,636.62	7.21	Mitochondrion
	ChACSM3	1743	14	13	581	65,712.44	8.86	Mitochondrion
	ChACSM4	1743	13	12	581	64,965.59	8.64	Mitochondrion
	ChACSM5	1773	14	13	591	66,162.68	7.56	Mitochondrion
ACSS	ChACSS1	2028	14	13	676	74,352.33	6.59	Mitochondrion
	ChACSS2	2,145	19	18	715	80,245.85	6.34	Cytoplasm
	ChACSS3	2061	16	15	687	74,761.3	8.86	Mitochondrion
ACSVL	ChSLC27A1	1941	13	12	647	71,003.95	8.83	Peroxisome
	ChSLC27A2	1863	10	9	621	70,319.22	8.72	Endoplasmic reticulum
	ChSLC27A3	2,211	11	10	737	79,697.36	8.73	Plasma membrane
	ChSLC27A4	1932	13	12	644	72,242.34	8.92	Peroxisome
	ChSLC27A5	2073	10	9	691	75,683.17	8.70	Plasma membrane
	ChSLC27A6	1905	10	9	635	71,710.21	8.71	Endoplasmic reticulum

TABLE 2 Information of the other ACS genes.

Gene	CDS (bp)	Exon	Intron	Amino acid (aa)	MW (Da)	pI	Localization
ChAACS	2019	18	17	673	74,924.77	6.14	cytoplasm
ChACSF2	1848	16	15	616	68,226.71	8.13	mitochondrion
ChACSF3	1761	10	9	587	65,045.76	6.75	mitochondrion
ChAASDH	3,321	15	14	1,107	123,626.02	6.24	plasma membrane

Among 25 identified genes, 22 were divided into five groups based on the sequence similarity and the principle of the human ACS family nomenclature. There were three genes in ACSS subfamily, five genes in the ACSM subfamily, five genes in the ACSL subfamily, six genes in the ACSVL subfamily, and two genes in the ACSBG subfamily (Table 1). However, ChAACS, ChACSF2, ChACSF3, and ChAASDH were not classified in any subfamilies and they were independent members of the ACS superfamily (Table 2).

Subcellular localizations of ChACSs were predicted using wolfpsort software. The results revealed that the ChACSMs, ChACSF2, ChACSF3, ChACSBG1, ChACSL4, ChACSS1, and ChACSS3 genes were located in the mitochondrion, the ChAACS, ChACSL1, ChACSL3, and ChACSS2 genes in the cytoplasm, the ChACSL5, ChACSL6, ChSLC27A2, and ChSLC27A6 genes in endoplasmic reticulum, the ChAASDH, ChACSBG2, ChSLC27A3, and ChSLC27A5 genes in plasma membrance and others in peroxisome. The lengths, molecular weight and theoretical isoelectric point of ACS proteins exhibited substantial variation. The lengths of 26 ACS proteins ranged from 578 to 1,107 amino acids, the MW varied from 64,636.62 to 123,626.02 Da and the pI value changed from 6.14 to 8.92 (Tables 1, 2).

FIGURE 1
Phylogenetic tree of ACS proteins from *Capra hircus*, *Homo sapiens*, *Ovis aries*, and *Bos taurus*. Legends: The tree was generated using the IQ-TREE with an ultrafast bootstrap value of 1,000. Each ACS subfamiliy is marked by a different colour. Note: Ch, *Capra hircus*; *HS, Homo sapiens*; Oa, *Ovis aries*; Bt, *Bos taurus*.
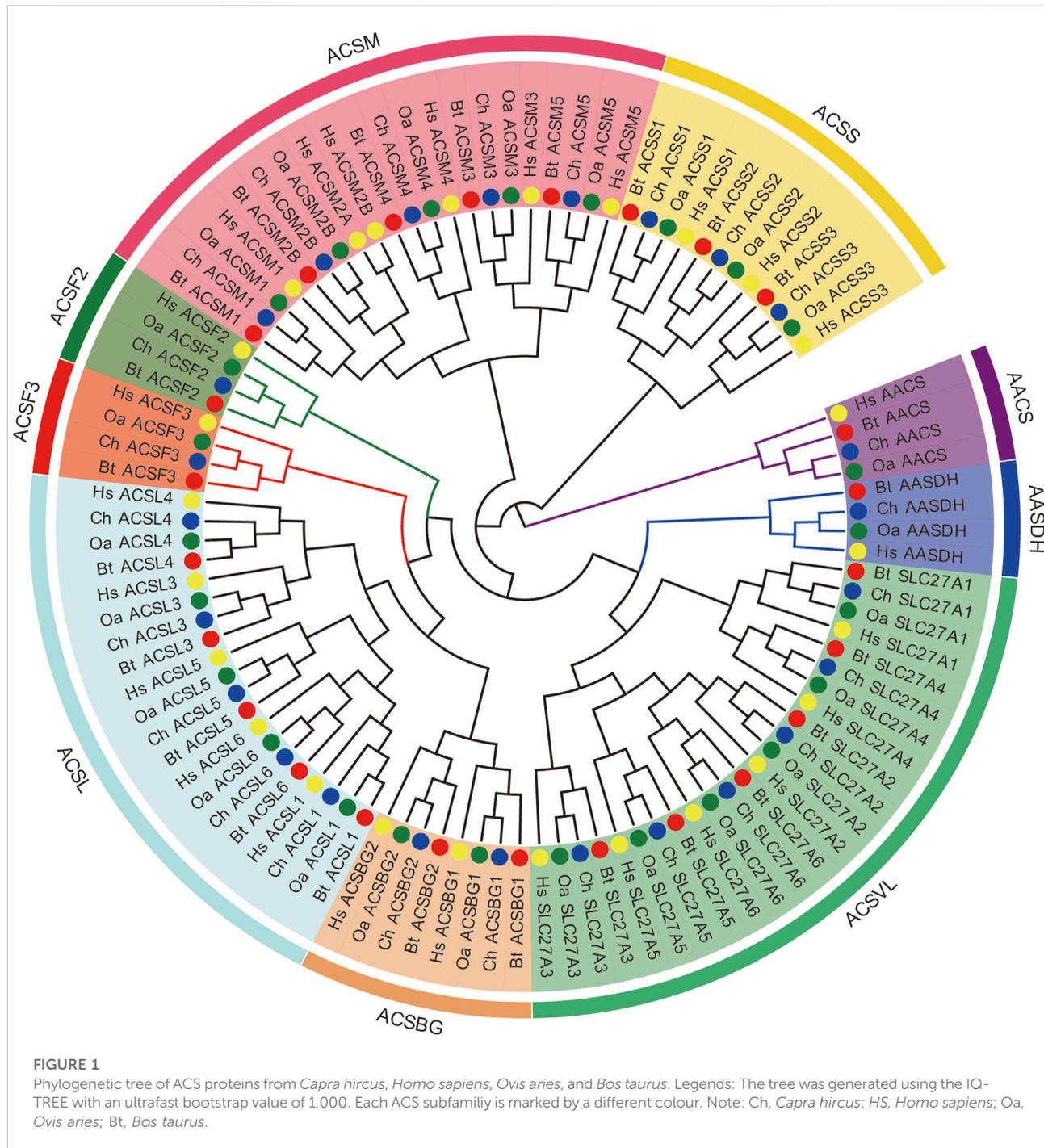
Phylogenetic analysis and multiple alignments

In order to investigate the evolutionary relationships of the ACS proteins among human, goat, sheep, and cow, a phylogenetic tree was constructed based on their full-length protein sequences, using IQ-TREE software. According to the topological structure of the phylogenetic tree, all the ACS proteins are clustered into nine distinct clades, including the ACSS, ACSM, ACSL, ACSVL, and ACSBG families (Figure 1). The other four clades (ACSS, ACSF1, ACSF2, and AASDH) were not classified in the common evolutionary branch, and belonged to the greater ACS family (Watkins et al., 2007). Hereafter, we characterized these four genes using the ACSF group for facilitating description. Three species (goat, sheep, and cow) had the same amount of
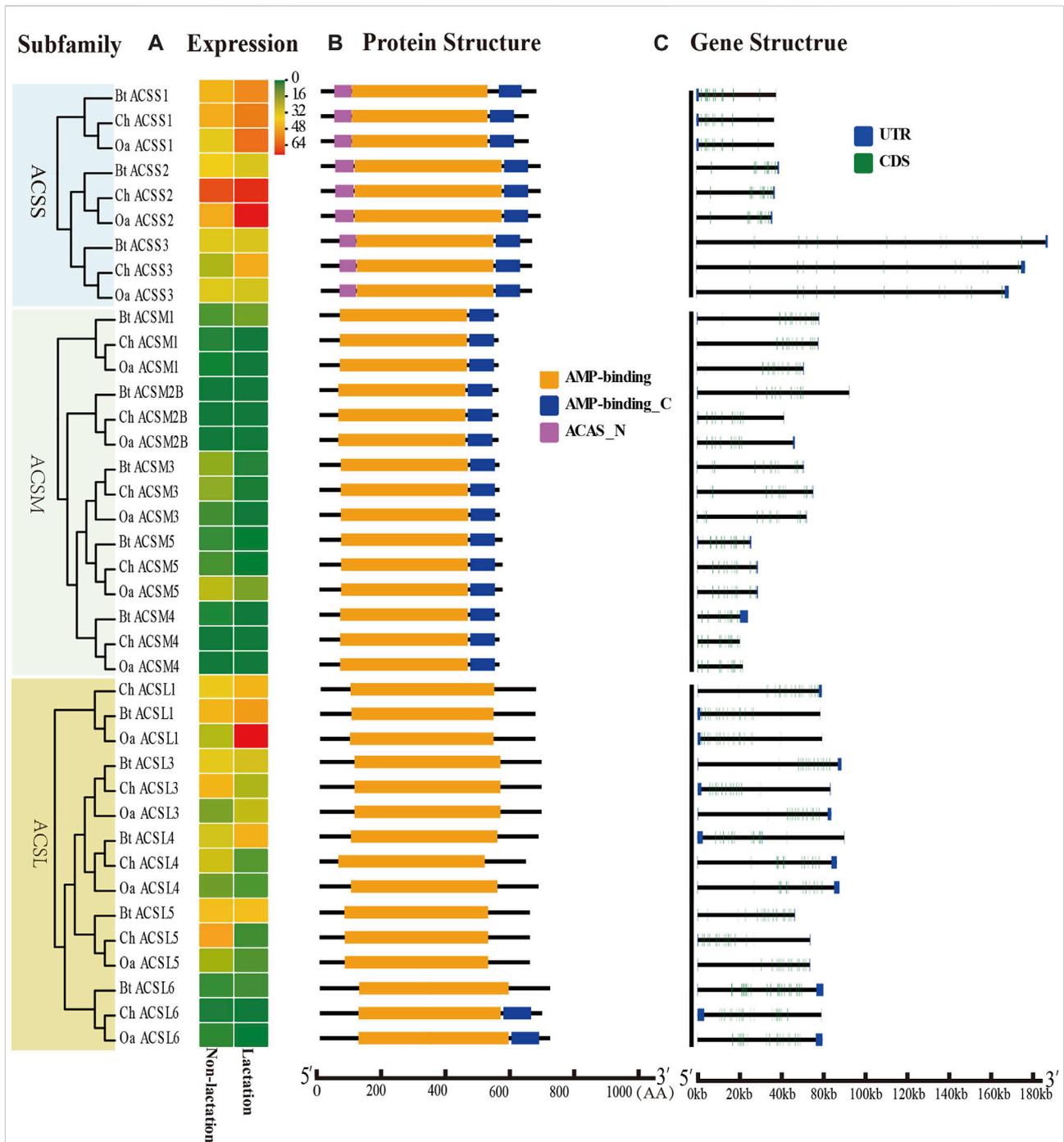
FIGURE 2
Expression profiles, gene structure and protein structure of ACS genes from the ACSS, ACSM, and ACSL subfamilies. Legends: **(A)** Heatmap showing the expression profiles of ACS genes in the non-lactation and lactation period of three species (dairy goat, sheep, and cow). Color gradient from green-to-red indicates expression values change from low to high. **(B)** Structures of ACS proteins with the AMP-binding domain represented by orange boxes, the AMP_C domain in red and ACAS_N domain in purple boxes. **(C)** Structure of ACS genes with exons in green, UTR regions in blue, and solid lines between the colored boxes representing introns.

ACS proteins and the same gene name, except that the *ACSM2A* protein in the ACSM group. The ACS proteins of goat and sheep showed closer evolutionary relationships among three species. The ACSVL group was the largest branch of the ChACS phylogeny and contained six ChACS proteins.
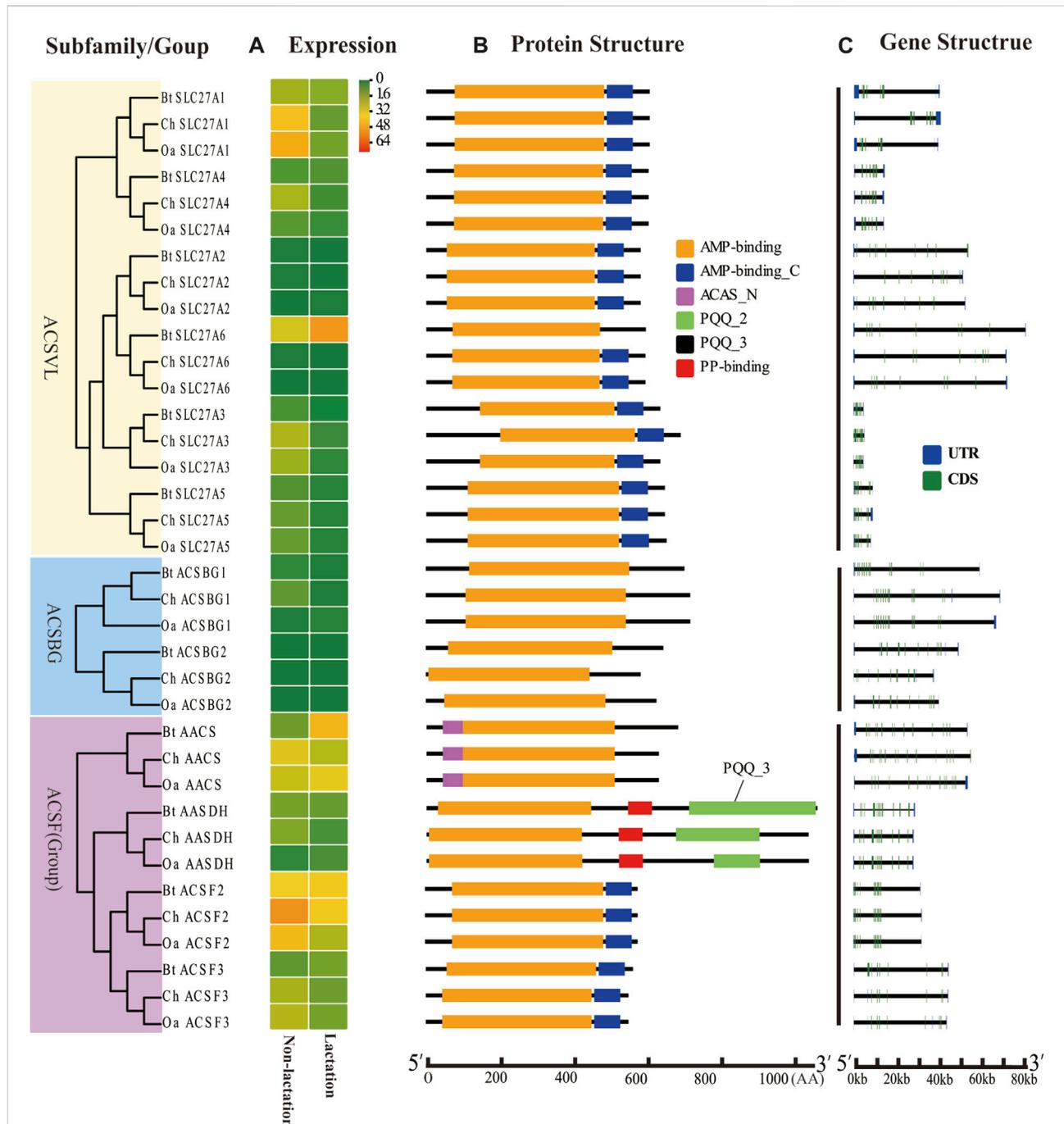
FIGURE 3
Expression profiles, gene structure, and protein structure of ACS genes from the ACSVL and ACSBG subfamilies and ACSF group. Legends: **(A)** Heatmap showing the expression profiles of ACS genes in the non-lactation and lactation period of three species (dairy goat, sheep, and cow). Color gradient from green-to-red indicates expression values change from low to high. **(B)** Structures of ACS proteins with the AMP-binding domain represented by orange boxes, the AMP_C domain in red, ACAS_N domain in purple, PQQ_2 domain in green, PQQ_3 domain in black and PP-binding domain in red. **(C)** Structure of ACS genes with exons in green, UTR regions in blue, and solid lines between the colored boxes representing introns.

To further characterize the ACS genes in goat, the protein sequences were aligned using the MAFFT-7.429. The conserved sequences of goat, sheep and cow ACS proteins were predicted by Pfam, and all proteins were found to contain AMP-binding domains (Figures 2B, 3B). In addition, the distribution of other protein domains was generally group specific. For instance, the

FIGURE 4
The chromosomal distribution of ChACS genes. Legends: The chromosomal position of each ACS gene was mapped to the goat genome. The chromosome number is indicated at the top of each chromosome.

acetyl-coenzyme A synthetase N-terminus (ACAS_N) domain was shared by all proteins in AACS subfamily. The AMP-binding enzyme C-terminal (AMP-binding_C) domain appeared in ACSS, ACSM, and ACSVL subfamily. Other domains, including phosphopantetheine attachment site (PP-binding), PQQ_2 and PQQ_3, were exclusive to AASDH proteins. Notably, *ChSLC27A6* and *ChACSL6* contained the AMP-binding_C domain, while *BtSLC27A6* and *BtACSL6* did not, indicating the function of *SLC27A6* and *ACSL6* genes might have changed in goat and cow (Figure 3B).

acyl-coenzyme A synthetase gene structure, chromosomal location, and conserved motif analysis

The structural diversity of exon-introns is considered to play an important role in genetic evolution. Therefore, we performed exon–intron structure analysis to explore the structural evolution mechanism of the ACS genes using the GSDS tool. The result showed that the number of introns in the ChACS genes contained from 9 to 21 introns (Tables 1, 2), which was
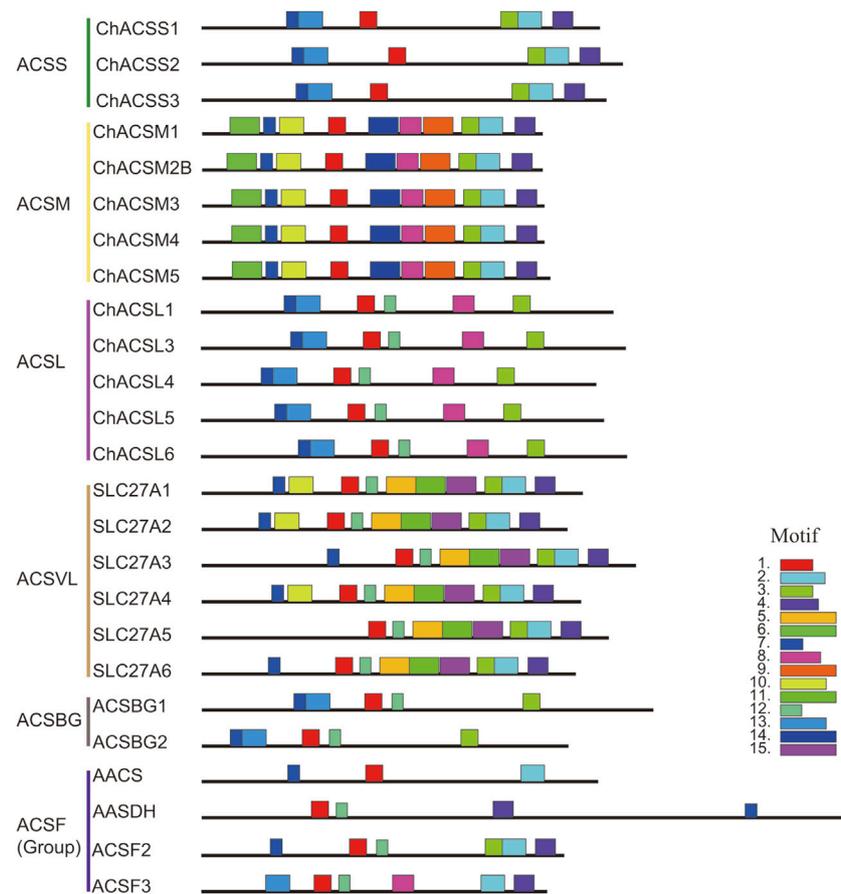
FIGURE 5
Motif distribution in ACS genes from goat. Legends: Motifs were predicted using the MEME web server (https://www.swissmodel.expasy.org/). The motifs are represented by different colors. The length of each box in the figure does not represent the actual motif size.

essentially consistent with what was found for the BtACS (Figures 2C, 3C). The exon-intron structure data also supported the phylogenetic tree topological structure. For example, the ACSVL subfamily contained 9–12 introns, and the ACSM subfamily had 12–14 introns (Table 1). The majority of ChACS genes had both 5′-and 3′-untranslated regions (UTRs), the ChSLC27A3 contained a 5′-UTR only, and the ChACM4 contained no UTR region (Figures 2C, 3C).
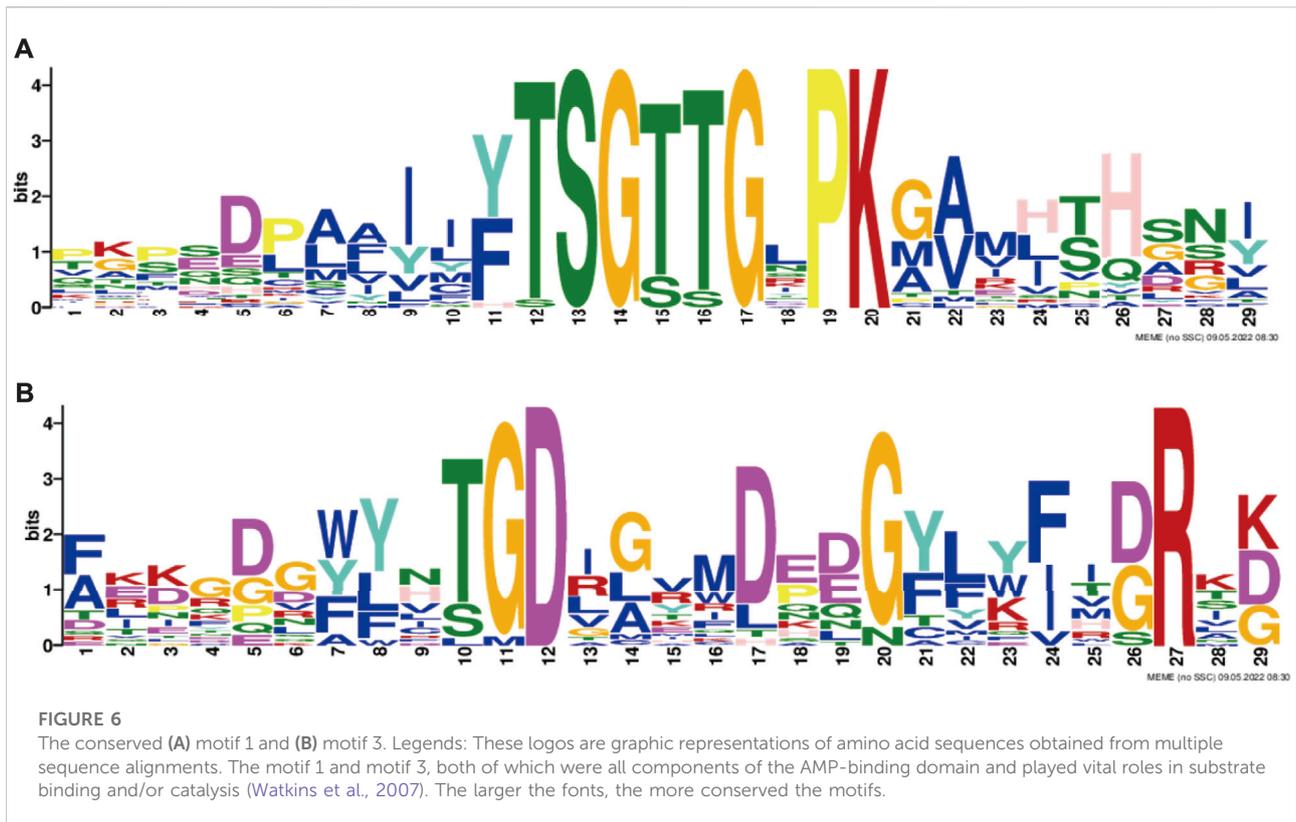
To investigate the distribution of ChACS genes in the goat genome, the ChACS genes were mapped to individual chromosomes. The 25 ChACSs were distributed on 16 chromosomes, with five ChACSs on chromosome 25, four ChACSs on chromosome 7, two genes on chromosomes 25 and 28, and one ChACS on each of chromosomes 2, 3, 5, 6, 10, 11, 17, 19, 21, 26, 27, and X (Figure 4).

In this study, the MEME tool was used to identify conserved motifs. A total of 15 different conserved motifs were predicted in the ChACS proteins (Figure 5). All of the 25 ACS proteins contained Motif 1, ChACSF3 did not contain Motif 7, while three proteins (ChAACS, ChAACDH, and ChACSF3) did not contain Motif 3. The closest ACS proteins in the phylogenetic tree had similar motifs. For example, each ACSM group member contained 10 motifs with similar motif composition pattern, which was very distinct from that of proteins in the other groups. Notably, more than 22 ACS proteins contained motif 1 and motif 3 (Figure 6), which were components of the AMP-binding domain and played vital roles in substrate binding and/or catalysis (Watkins et al., 2007).

Expression analysis of acyl-coenzyme A synthetase genes

To analyze the expression profiles of ACS genes in distinct tissues of dairy goat, we investigated the FPKM values of the ACS genes in heart, kidney, liver, mammary gland, and muscle tissues. The majority of the 25 ACS genes were expressed in the test tissues, and the expression of the ACSBG2 was not detected in

FIGURE 6
The conserved **(A)** motif 1 and **(B)** motif 3. Legends: These logos are graphic representations of amino acid sequences obtained from multiple sequence alignments. The motif 1 and motif 3, both of which were all components of the AMP-binding domain and played vital roles in substrate binding and/or catalysis (Watkins et al., 2007). The larger the fonts, the more conserved the motifs.
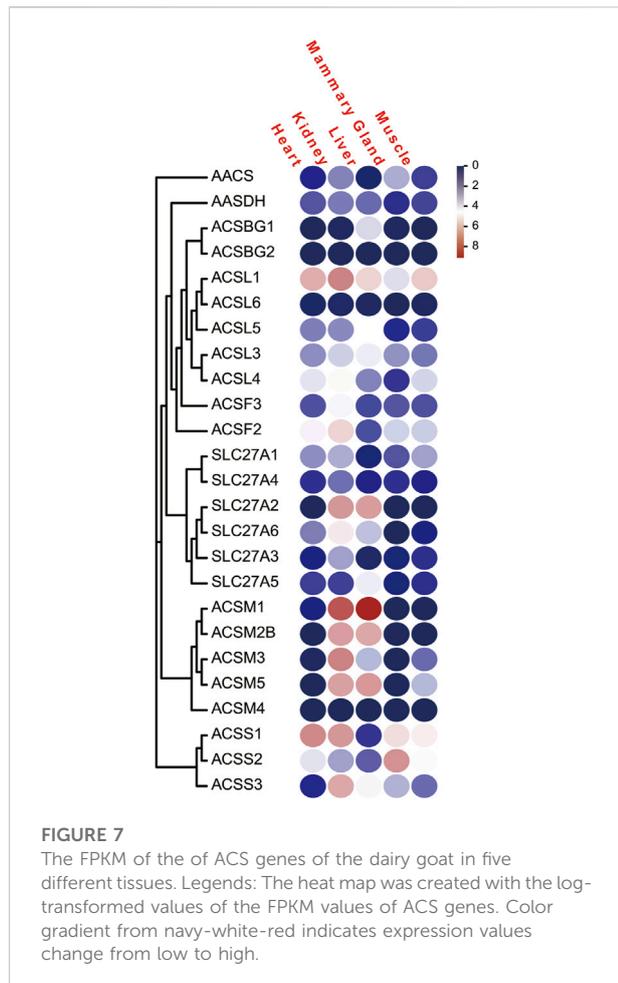
multi tissues. Only *ACSL1* was highly expressed in all tissues. In general, the expression profiles analysis suggested that the expression of ACS genes of the dairy goat was related to distinct tissues and the expression patterns were also distinct among each ACS subfamily (Figure 7; Supplementary Sheet S3). For example, the members of the ACSM subfamily, except for *ACSM4*, exhibited higher expression level in kidney and liver than in other tissues. In heart, *ACSS1* was most highly expressed, and the ACS genes, including *ACSL1*, *ACSL4*, *ACSF2*, *ACSS1*, and *ACSS2*, were highly expressed. Most of ACS genes were highly expressed in kidney with *ACSM1* expression was prominent. Similarly, *ACSM1* was also highly expressed in liver. The majority of ACS genes exhibited lower gene expression in mammary gland and muscle tissues. The highest expression was observed in mammary gland and muscle for *ACSS2* and *ACSL1*, respectively. These finding suggested ACS genes had different functions in distinct tissues.

To explore the ACS genes expression profiles in mammary glands among dairy goat, sheep and cow, we compared their transcript abundance in the non-lactation and lactation period (Figures 2A, 3A). The expression pattern indicated that the expression levels of ACS genes was associated with each ACS protein group. For instance, the proteins of ACSM and ACSBG groups showed lower expression in three species, and members of ACSS and ACSL groups exhibited high expressions, except for *ACSL6*. Some ACS genes showed identical expression trend in

dairy goat and sheep, while exhibited distinct expression patterns in cow (Supplementary Sheet S4). For example, the expression of *ChACSS2* and *OaACSS2* was the most highly expressed relative to other members in lactation stage, while *BtACSS2* was not preferentially expressed. In addition, the *ACSS2* exhibited higher transcript abundance level in lactation stage than in non-lactation stage in dairy goats and sheep, while the opposite was observed in cow. Whereas, the transcript abundance of *ACSL4* and *ACSF3* was significantly lower in lactation stage than in non-lactation stage in dairy goat and sheep, which showed opposite result in cow. Some genes such as *ACSM3*, *ACSM5*, and *ACSS1* exhibited the same expression trend in dairy goat and cow, which were lower expression level in non-lactation period than in lactation period. These differences in expression profiles of goat and cow suggested that the ASC genes might have an effect on composition of fatty acids in goat milk.

Discussion

Previous studies have shown that fatty acids are activated to fatty acyl-CoAs by ACSs before involving in both anabolic and catabolic processes. Processes such as the synthesis of acylated protein and comple lipids, fatty acid extension or unsaturation, and fatty acid oxidation all require activated fatty acid substrates (Steinberg et al., 2000). This study found that a total of 25 full-length ACSs (without

FIGURE 7
The FPKM of the of ACS genes of the dairy goat in five different tissues. Legends: The heat map was created with the log-transformed values of the FPKM values of ACS genes. Color gradient from navy-white-red indicates expression values change from low to high.

ACSM2A), representing five subfamilies (ACSS, ACSM, ACSL, ACSVL, ACSBG) and one group (ACSF) in the *Capra hircus* reference genome. In addition, we performed analyses of their phylogeny relationships, gene structures, conserved domains and motifs, expression profiling in multi tissues, and expressed difference in the non-lactation and lactation mammary glands of three species (goat, sheep, and cow).

There are 26 members of the ACS gene family reported in mammals (Watkins et al., 2007). Surprisingly, *ACSM2A* is missing in goat, sheep and cow genomes in the present study. The biological function of *ACSM2A* has rarely been reported in the current literature (Rencia and Erasmus, 2016; van der sluis et al., 2018). Several reports have shown that the *ACSM2A* and *ACSM2B* are nearly identical, with nucleotide homology of 98.8% and an amino acid identity of 97.1%. However, evidence also suggests that they are distinct genes (Rencia and Erasmus, 2016). Our findings showed the *ChACSM2B* was homologous to the *HsACSM2A* and *HsACSM2B* (Watanabe et al., 2020), with the amino acid identity of 80.6% and 80.4% respectively. Thus, *ChACSM2B* is highly conserved in the evolutionary process and might have similar function to *HsACSM2A*. Phylogenetic

analysis exhibited that each evolutionary branch of the ACS family contained goat, sheep, cow and human proteins, suggesting that possible functions may be conserved among species (Liu et al., 2020). The four members of ACSF group, including *AACS*, *ACSF2*, *ACSF3*, and *AASDH*, clustered into four distinct clades, which was in consistent with the previous studies (Watkins et al., 2007). All the coding sequence of *ACSs* are disrupted by 9–21 introns, and the intron positions of goat and cow are distinct, suggesting that intron insertion may be result from independent events (Karan et al., 2003). All of ACS members contain related AMP-binding domains and FA binding motifs (Watkins et al., 2007). In this study, we found all ACS proteins shared a similar AMP-binding domain, which also illustrated that the AMP-binding functional domain was a conserved sequence and directly participated in the catalytic reaction (Wang et al., 2022). In addition to AMP-binding domain, the distribution of other protein domains was generally group specific. The conserved motif analysis showed that ACSs motifs also shared group specific, and the closest ACS proteins in the phylogenetic tree had similar motifs. These findings suggest that a wide range and diversity of ACSs might be result from suffering selective pressures to adapt to the metabolism of numerous and complex fatty acids in the evolutionary process (Watkins et al., 2007).

Each ACS gene plays a unique role, channeling its CoA derivatives to a specific metabolic pathway (Rencia and Erasmus, 2016). Fatty acyl-CoA molecules are the important regulatory molecules and metabolic intermediates (He et al., 2022), which have a variety of functions in the metabolism. Acyl-CoAs are oxidized to provide cellular energy, and are instrumental in the synthesis of acylated protein, and complex lipids such as triacylglycerols and phospholipids (Wang et al., 2022). Although *ACSL1* is highly expressed in heart, kidney, liver, mammary gland, and skeletal muscle, it has been shown to have different functions in different tissues because of the dual location on both the mitochondria and the endoplasmic reticulum (Grevengoed et al., 2014). The ACSM proteins are considered as liver mitochondrial enzymes (Rencia and Erasmus, 2016), which were also confirmed in this study. We also observed that the ACSM proteins were highly expressed in kidney, indicating that the ACSM family members might also play vital roles in kidney fatty acids metabolism. Several of the ACSVL family members located in the internal cellular membranes may have separate transport and activation functions, while ACSVLs not located on the plasma membrane are thought to enhance cellular uptake of fatty acids (Grevengoed et al., 2014). Multiple ACS proteins were expressed simultaneously in the same tissue, suggesting that these genes may coordinate together to perform a similar function. Several genes in the same subfamily exhibited distinct expression pattern from other members, such as *ACSL1* in ACSL subfamily and *SLC27A6* in ACSVL subfamily, suggesting that they may be involved in different biological functions.

vBCFAs are the cause of the goaty flavour of goat milk (Teng et al., 2018) and are more likely to be synthesized in goat tissues than in rumen microbes (Berthelot et al., 2001). It is believed that methylmalonate formed by carboxylation of propionic acid is an essential substance for the synthesis of vBCFAs (Priolo et al., 2001). Thus, *ACSs* might be involved in the synthesis or regulation of vBCFAs. By investigating the differences of expressed profiles in non-lactation and lactation mammary glands between goat and cow, it was possible to explore which ACS genes participate in formation of the goaty flavour of goat milk. In this study, the analysis of expression profiles suggests that *ACSF3* showed identical expression trend in dairy goat and sheep, while exhibited distinct expression patterns in cow. *ACSF3* is able to activate malonic acid to malonyl CoA (Chen et al., 2011), and also catalyzes methylmalonic acid to methylmalonyl-CoA (Monteuuis et al., 2017). It was reasonable to speculate that *ACSF3* provided the substrate for the synthesis of vBCFAs. ACS catalyzes propionate to propionyl-CoA which is a primer for the synthesis of 4-methyloctanoic acid (4-Et-8:0), and butyryl-CoA is the primer for the synthesis of 4-ethyloctanoic acid (4-Et-8:0). In this study, ACSS subfamily members were highly expressed, and the *ChACSS2* was the gene with the highest expression among all *ChACSs*, but *BtACSS2* was not preferentially expressed in cow. ACSSs typically activate short-chain fatty acids like acetate, propionate, or butyrate, involving in energy metabolism (Watkins et al., 2007). Our subcellular localization results have also proven that *ACSS1* and *ACSS3* were the mitochondrial proteins (Moffett et al., 2020; He et al., 2022), and *ACSS2* was localized in cytoplasm and nucleus (Li et al., 2017). As the fatty acids are activated in cytoplasm, our results suggest that *ACSS2* might participate in the synthesis of vBCFAs.

Goat milk contains high amounts of short-chain and medium-chain fatty acids (Luna et al., 2008). The oxidation of fatty acids is initiated in the cytoplasm by the formation of acyl-CoA by ACSs that are located in the endoplasmic reticulum and mitochondrial outer membrane (Watkins and Ellis, 2012). ACSMs typically activate medium-chain fatty acids, while ACSLs have a preference for long-chain fatty acids (Rencia and Erasmus, 2016). In this study, ACSMs were expressed at relatively low levels in lactation stage. We suggest that low expression of ACSMs is a possible reason for the high concentration of free medium-chain fatty acids in goat milk. We also observed that the genes including *ChACSL4*, *ChACSL5*, *OaACSL4*, and *OaACSL5* exhibited a lower transcript abundance in lactation stage than in non-lactation stage, but *BtACSL4* and *BtACSL5* showed opposite trend. Thus, we propose that *ACSL4* and *ACSL5* may have an effect on the long-chain fatty acid content between goat milk and cow's milk.

ACSVL subfamily members are integral transmembrane proteins, which play a vital role in the absorption of long-chain fatty acids into cells (Gallardo et al., 2013). *ACSF2* is a mitochondrial matrix enzyme involved in the tricarboxylic acid cycle and fatty acid synthesis (Yang et al., 2019). *AASDH* is a protein of unknown function and homologous to bacterial non-ribosomal peptide synthetase (Drozak et al., 2014). Previous studies and the current research suggest these genes may not related to the metabolic process of vBCFAs.

Conclusion

A total of 25 ACS genes were characterized in goats and subdivided into five subfamilies. The ACS proteins all had the conserved the AMP-binding domain and motif1. The phylogenetic relationships of *ACSs* were also supported by gene structures, motifs and protein domain. The majority of the ACS genes were expressed in the multi tissues, with similar or different expression levels. These findings provide reference information to further understand the classification and putative functions of ACS genes in goats. Two genes, *ACSS2* and *ACSF3*, may take part in the synthesis of vBCFAs. This study also provides genomic and expression information for *ACSs* in goat, and the findings may be useful for further research on the formation mechanisms of the goaty flavour in goat milk.

Data availability statement

The sequences information analyzed in this study are available in the GenBank (https://www.ncbi.nlm.nih.gov/), and the Ensemble database (http://ftp.ensembl.org/pub/release-104/). The RNA-seq data was from the NCBI's SRA database (Accession: PRJNA309284, PRJNA309345, PRJNA339650, PRJNA637690, and PRJNA482783).

Ethics statement

The animal study was reviewed and approved by the Animal Care and Use Committee of the Northwest A&F University.

Author contributions

Funding

and Development Plan of Shaanxi Province, China (No. 2021ZDLNY02-02).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene. 2022.980463/full#supplementary-material

References

Ansari, I. H., Longacre, M. J., Stoker, S. W., Kendrick, M. A., O'Neill, L. M., Zitur, L. J., et al. (2017). Characterization of Acyl-CoA synthetase isoforms in pancreatic beta cells: Gene silencing shows participation of ACSL3 and ACSL4 in insulin secretion. *Arch. Biochem. Biophys.* 618, 32–43. doi:10.1016/j.abb.2017.02.001

Balakrishnan, S., Gao, S., Lercher, M. J., Hu, S., and Chen, W. H. J. N. A. R. (2019). Evolview v3: A webserver for visualization, annotation, and management of phylogenetic trees. *Nucleic Acids Res.* 47, W270–W275. doi:10.1093/nar/gkz357

Berthelot, V., Bas, P., Schmidely, P., and Duvaux-Ponter, C. J. S. R. R. (2001). Effect of dietary propionate on intake patterns and fatty acid composition of adipose tissues in lambs. *Small Rumin. Res.* 40 (1), 29–39. doi:10.1016/s0921-4488(00)00217-0

Black, P. N., Zhang, Q., Weimar, J. D., and DiRusso, C. C. (1997). Mutational analysis of a fatty acyl-coenzyme A synthetase signature motif identifies seven amino acid residues that modulate fatty acid substrate specificity. *J. Biol. Chem.* 272 (8), 4896–4903. doi:10.1074/jbc.272.8.4896

Brennand, C. P., Kim, H. J., and Lindsay, R. C. J. (2010). Aroma properties and thresholds of some branched-chain and other minor volatile fatty acids occurring in milkfat and meat lipids. *J. Sens. Stud.* 4 (2), 105–120. doi:10.1111/j.1745-459X.1989.tb00461.x

Chen, H., Kim, H. U., Weng, H., and Browse, J. J. P. C. (2011). Malonyl-CoA synthetase, encoded by ACYL ACTIVATING ENZYME13, is essential for growth and development of arabidopsis. *Plant Cell* 23 (6), 2247–2262. doi:10.1105/tpc.111.086140

Drozak, J., Veiga-Da-Cunha, M., Kadziolka, B., and Van Schaftingen, E. J. F. J. (2014). Vertebrate AcylCoA synthetase family member4 (ACSF4-U26) is a β-alanine-activating enzyme homologous to bacterial non-ribosomal peptide synthetase. *FEBS J.* 281 (6), 1585–1597. doi:10.1111/febs.12725

Fernandez, R. F., Kim, S. Q., Zhao, Y., Foguth, R. M., Weera, M. M., Counihan, J. L., et al. (2018). Acyl-CoA synthetase 6 enriches the neuroprotective omega-3 fatty acid DHA in the brain. *Proc. Natl. Acad. Sci. U. S. A.* 115 (49), 12525–12530. doi:10.1073/pnas.1807958115

Finn, R. D., Clements, J., and Eddy, S. R. J. N. A. R. (2011). HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Res.* 39, 29–37. doi:10.1093/nar/gkr367

Gallardo, D., Amills, M., Quintanilla, R., and Pena, R. N. (2013). Mapping and tissue mRNA expression analysis of the pig solute carrier 27A (SLC27A) multigene family. *Gene* 515 (1), 220–223. doi:10.1016/j.gene.2012.11.029

Grevengoed, T. J., Klett, E. L., and Coleman, R. A. (2014). Acyl-CoA metabolism and partitioning. *Annu. Rev. Nutr.* 34, 1–30. doi:10.1146/annurev-nutr-071813-105541

Ha, J. K., and Lindsay, R. C. (1993). Release of volatile branched-chain and other fatty acids from ruminant milk fats by various lipases. *J. Dairy Sci.* 76 (3), 677–690. doi:10.3168/jds.s0022-0302(93)77391-9

He, W., Zhou, X., Wu, Q., Zhou, L., Zhang, Z., Zhang, R., et al. (2022). Acetyl CoA synthase 2 potentiates ATG5-induced autophagy against neuronal apoptosis after subarachnoid hemorrhage. *J. Mol. Histol.* 10, 511–521. doi:10.1007/s10735-022-10057-x

Kaffarnik, S., Kayademir, Y., Heid, C., and Vetter, W. J. J. (2014). Concentrations of volatile 4-alkyl-branched fatty acids in sheep and goat milk and dairy products. *J. Food Sci.* 79 (11), C2209–C2214. doi:10.1111/1750-3841.12673

Kanter, J. E., Tang, C., Oram, J. F., and Bornfeldt, K. E. (2012). Acyl-CoA synthetase 1 is required for oleate and linoleate mediated inhibition of cholesterol efflux through ATP-binding cassette transporter A1 in macrophages. *Biochim. Biophys. Acta* 1821 (3), 358–364. doi:10.1016/j.bbalip.2011.10.008

Katoh, K., and Standley, D. M (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* 30 (4), 772–780. doi:10.1093/molbev/mst010

Karan, D., Lesbats, M., David, J. R., and Capy, P. J. J. (2003). Evolution of the AMP-forming acetyl-CoA synthetase gene in the Drosophilidae family. *J. Mol. Evol.* 57 (1), S297–S303. doi:10.1007/s00239-003-0040-1

Kochan, G., Pilka, E. S., Delft, F. V., Oppermann, U., and Yue, W. W. (2009). Structural snapshots for the conformation-dependent catalysis by human medium-chain acyl-coenzyme A synthetase ACSM2A. *J. Mol. Biol.* 388 (5), 997–1008. doi:10.1016/j.jmb.2009.03.064

Lam-Tung, N., Schmidt, H. A., Arndt, V. H., and Quang, M. B. J. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 1, 268–274. doi:10.1093/molbev/msu300

Li, C., Zhu, J., Shi, H., Luo, J., and Loor, J. J. J. (2020). Comprehensive transcriptome profiling of dairy goat mammary gland identifies genes and networks crucial for lactation and fatty acid metabolism. *Front. Genet.* 11, 878. doi:10.3389/fgene.2020.00878

Li, X., Yu, W., Qian, X., Xia, Y., Zheng, Y., Lee, J. H., et al. (2017). Nucleus-translocated ACSS2 promotes gene transcription for lysosomal biogenesis and autophagy. *Mol. Cell* 66 (5), 684–697. doi:10.1016/j.molcel.2017.04.026

Liu, Z., Ge, X. X., Wu, X. M., Xu, Q., and Guo, W. W. J. B. G. (2020). Genome-wide analysis of the citrus B3 superfamily and their association with somatic embryogenesis. *BMC Genomics* 21 (1), 305. doi:10.1186/s12864-020-6715-9

Luna, P., Bach, A., Juárez, M., and de la Fuente, M. A. (2008). Effect of a diet enriched in whole linseed and sunflower oil on goat milk fatty acid composition and conjugated linoleic acid isomer profile. *J. Dairy Sci.* 91 (1), 20–28. doi:10.3168/jds.2007-0447

Mashek, D. G., Bornfeldt, K. E., Coleman, R. A., Berger, J., Bernlohr, D. A., Black, P., et al. (2004). Revised nomenclature for the mammalian long-chain acyl-CoA synthetase gene family. *J. Lipid Res.* 45 (10), 1958–1961. doi:10.1194/jlr.E400002-JLR200

Meller, N., Morgan, M. E., Wong, W. P., Altemus, J. B., and Sehayek, E. J. L. (2013). Targeting of Acyl-CoA synthetase 5 decreases jejunal fatty acid activation with no effect on dietary long-chain fatty acid absorption. *Lipids Health Dis.* 12 (1), 88–94. doi:10.1186/1476-511X-12-88

Moffett, J. R., Puthillathu, N., Vengilote, R., Jaworski, D. M., and Namboodiri, A. M. (2020). Acetate revisited: A key biomolecule at the nexus of metabolism, epigenetics, and oncogenesis - Part 2: Acetate and ACSS2 in health and disease. *Front. Physiol.* 11, 580171. doi:10.3389/fphys.2020.580171

Monteuuis, G., Suomi, F., Kertr, J. M., Masud, A. J., and Kastaniotis, A. J. J. B. J. (2017). A conserved mammalian mitochondrial isoform of acetyl-CoA carboxylase ACC1 provides the malonyl-CoA essential for mitochondrial biogenesis in tandem with ACSF3. *Biochem. J.* 474 (22), 3783–3797. doi:10.1042/BCJ20170416

Ndiaye, H., Liu, J. Y., Hall, A., Minogue, S., Morgan, M. Y., and Waugh, M. G. (2020). Immunohistochemical staining reveals differential expression of ACSL3 and ACSL4 in hepatocellular carcinoma and hepatic gastrointestinal metastases. *Biosci. Rep.* 40 (4), BSR20200219. doi:10.1042/BSR20200219

O'Brien, M. J., Beijerink, N. J., Sansom, M., Thornton, S. W., Chew, T., and Wade, C. M. (2020). A large deletion on CFA28 omitting ACSL5 gene is associated with intestinal lipid malabsorption in the Australian Kelpie dog breed. *Sci. Rep.* 10 (1), 18223. doi:10.1038/s41598-020-75243-x

Perez-Nunez, I., Karaky, M., Fedetz, M., Barrionuevo, C., Izquierdo, G., Matesanz, F., et al. (2019). Splice-site variant in ACSL5: A marker promoting opposing effect on cell viability and protein expression. *Eur. J. Hum. Genet.* 27 (12), 1836–1844. doi:10.1038/s41431-019-0414-5

Priolo, A., Micol, D., and Agabriel, J. J. A. R. (2001). Effects of grass feeding systems on ruminant meat colour and flavour. A review. *Anim. Res.* 50 (3), 185–200. doi:10.1051/animres:2001125

Rencia, V., and Erasmus, E. J. E. O. D. M. T. (2016). Xenobiotic/medium chain fatty acid: CoA ligase – a critical review on its role in fatty acid metabolism and the detoxification of benzoic acid and aspirin. *Expert Opin. Drug Metab. Toxicol.* 12 (10), 1169–1179. doi:10.1080/17425255.2016.1206888

Rossi Sebastiano, M., and Konstantinidou, G. (2019). Targeting long chain acyl-CoA synthetases for cancer therapy. *Int. J. Mol. Sci.* 20 (15), E3624. doi:10.3390/ijms20153624

Salles, C., Sommerer, N., Septier, C., Issanchou, S., Science, J. L. Q., Garem, A., et al. (2010). Goat cheese flavor: Sensory evaluation of branched-chain fatty acids and small peptides. *J. Food Sci.* 67 (2), 835–841. doi:10.1111/j.1365-2621.2002.tb10686.x

Steinberg, S. J., Morgenthaler, J., Heinzer, A. K., Smith, K. D., and Watkins, P. A. (2000). Very long-chain acyl-CoA synthetases. Human "bubblegum" represents a new family of proteins capable of activating very long-chain fatty acids. *J. Biol. Chem.* 275 (45), 35162–35169. doi:10.1074/jbc.M006403200

Takahiro, F., and Tokuo, Y. J. o. B. (1992). Cloning and functional expression of a novel long-chain acyl-CoA synthetase expressed in brain. *J. Biochem.* 111 (2), 197–203. doi:10.1093/oxfordjournals.jbchem.a123737

Tang, Y., Zhou, J., Hooi, S. C., Jiang, Y. M., and Lu, G. D. (2018). Fatty acid activation in carcinogenesis and cancer development: Essential roles of long-chain acyl-CoA synthetases. *Oncol. Lett.* 16 (2), 1390–1396. doi:10.3892/ol.2018.8843

Teng, F., Reis, M. G., Ma, Y., and Day, L. (2018). Effects of season and industrial processes on volatile 4-alkyl-branched chain fatty acids in sheep milk. *Food Chem.* 260, 327–335. doi:10.1016/j.foodchem.2018.04.011

van der Sluis, R., Genetics, R. J. M., and Mgg, G. (2018). Analyses of the genetic diversity and protein expression variation of the acyl: CoA medium-chain ligases, ACSM2A and ACSM2B. *Mol. Genet. Genomics* 293 (5), 1279–1292. doi:10.1007/s00438-018-1460-3

Wang, Y., Guo, L., Zhao, Y., Zhao, X., and Yuan, Z. (2022). Systematic analysis and expression profiles of the 4-coumarate: CoA ligase (4CL) gene family in pomegranate (punica granatum L.). *Int. J. Mol. Sci.* 23 (7), 3509. doi:10.3390/ijms23073509

Watanabe, H., Paxton, R. L., Tolerico, M. R., Nagalakshmi, V. K., Gomez, R. A., Okusa, M. D., et al. (2020). Expression of Acsm2, a kidney-specific gene, parallels the function and maturation of proximal tubular cells. *Am. J. Physiol. Ren. Physiol.* 319 (4), F603–F611. doi:10.1152/ajprenal.00348.2020

Watkins, P. A., and Ellis, J. M. (2012). Peroxisomal acyl-CoA synthetases. *Biochim. Biophys. Acta* 1822 (9), 1411–1420. doi:10.1016/j.bbadis.2012.02.010

Watkins, P. A., Maiguel, D., Jia, Z., and Pevsner, J. (2007). Evidence for 26 distinct acyl-coenzyme A synthetase genes in the human genome. *J. Lipid Res.* 48 (12), 2736–2750. doi:10.1194/jlr.M700378-JLR200

Yang, Y., An, C., Yao, Y., Cao, Z., Gu, T., Xu, Q., et al. (2019). Intron polymorphisms of MAGI-1 and ACSF2 and effects on their expression in different goose breeds. *Gene* 701, 82–88. doi:10.1016/j.gene.2019.02.102