Meta-analysis of mammary RNA seq datasets reveals the molecular understanding of bovine lactation biology

Article *in* Genome · May 2019

DOI: 10.1139/gen-2018-0144

CITATIONS
0

READS
82

5 authors, including:

Periyasamy Vijayakumar
Tamil Nadu Veterinary and Animal Sciences University
15 PUBLICATIONS **71** CITATIONS
SEE PROFILE

Bakyaraj Sanniyasi
Tamil Nadu Veterinary and Animal Sciences University
2 PUBLICATIONS **39** CITATIONS
SEE PROFILE

Singaravadivelan Arunasalam
Tamil Nadu Veterinary and Animal Sciences University
8 PUBLICATIONS **0** CITATIONS
SEE PROFILE

NRC Research Press

ARTICLE

Meta-analysis of mammary RNA seq datasets reveals the molecular understanding of bovine lactation biology

Periyasamy Vijayakumar, Sanniyasi Bakyaraj, Arunasalam Singaravadivelan, Thangavelu Vasanthakumar, and Ramalingam Suresh

Abstract: A better understanding of the biology of lactation, both in terms of gene expression and the identification of candidate genes for the production of milk and its components, is made possible by recent advances in RNA seq technology. The purpose of this study was to understand the synthesis of milk components and the molecular pathways involved, as well as to identify candidate genes for milk production traits within whole mammary transcriptomic datasets. We performed a meta-analysis of publically available RNA seq transcriptome datasets of mammary tissue/milk somatic cells. In total, 11 562 genes were commonly identified from all RNA seq based mammary gland transcriptomes. Functional annotation of commonly expressed genes revealed the molecular processes that contribute to the synthesis of fats, proteins, and lactose in mammary secretory cells and the molecular pathways responsible for milk synthesis. In addition, we identified several candidate genes responsible for milk production traits and constructed a gene regulatory network for RNA seq data. In conclusion, this study provides a basic understanding of the lactation biology of cows at the gene expression level.

Key words: cow, mammary transcriptome, meta-analysis, pathways, candidate genes, milk synthesis.

Résumé : Une meilleure connaissance de la biologie de la lactation, tant en ce qui concerne l'expression des gènes que l'identification des gènes candidats pour la production du lait et de ses composantes, est rendue possible par les récentes avancées en matière de technologie RNA-seq. Le but de cette étude était de comprendre la synthèse des composantes du lait et les sentiers moléculaires impliqués ainsi que d'identifier des gènes candidats pour la production laitière au sein de jeux de données transcriptomiques pour les tissus mammaires. Les auteurs ont réalisé une méta-analyse des données RNA-seq disponibles pour les tissus mammaires et les cellules somatiques du lait. Au total, 11 562 gènes étaient partagés au sein de tous les transcriptomes de la glande mammaire. Une annotation fonctionnelle des gènes exprimés communément a contribué à une compréhension des processus moléculaires contribuant à la synthèse des lipides, des protéines et du lactose au sein des cellules sécrétrices mammaires et des sentiers moléculaires responsables de la synthèse du lait. De plus, les auteurs ont identifié plusieurs gènes candidats contrôlant la production du lait et ont construit un réseau de régulation génique à partir des données RNA-seq. En conclusion, cette étude contribue à une meilleure compréhension de la lactation chez les vaches au niveau de l'expression génique. [Traduit par la Rédaction]

Mots-clés : vache, transcriptome des tissus mammaires, méta-analyse, sentiers, gènes candidats, synthèse du lait.

Introduction

Cow milk is an important source of dietary food and provides essential nutrients like energy, fats, proteins, calcium, magnesium, selenium, riboflavin, vitamin B12, and pantothenic acid to infant mammals (http://www.fao.org/dairy-production-products/products/milk-composition/en/). In general, the gross composition of cow's milk is constituted of 87.7% water, 4.9% lactose, 3.4% fat, 3.3% protein, and 0.7% minerals. The milk composition varies according to breed, lactation stage, dietary intake, parity, farming system, the physical environment, and seasonal effect.

Milk components like lactose, proteins, and fats are synthesized from glucose, amino acids, triglycerides, or fatty acids of dietary nutrient origin or body resources in the secretory cells of the mammary gland (Hurley 2010). The molecular mechanisms and pathways for the synthesis of milk and its components are not completely under-

Received 14 August 2018. Accepted 5 May 2019.

Corresponding Editor: L.L. Guan.

P. Vijayakumar, A. Singaravadivelan, T. Vasanthakumar, and R. Suresh. Veterinary College and Research Institute, TANUVAS, Orathanadu-614 625, Thanjavur, Tamil Nadu, India.
S. Bakyaraj. College of Poultry Production and Management, TANUVAS, Hosur-635 110, Krishnagiri, Tamil Nadu, India.
Corresponding author: Periyasamy Vijayakumar (email: nayaganviji@gmail.com).

stood. Detailed knowledge of lactation biology at the molecular level will aid in the identification of direct causative genes responsible for milk traits, and the application of these causative genes in genomic selection programmes will increase the accuracy of dairy cattle selection methods (Wiggans et al. 2017). Further, based on this molecular mechanism, breeders can attempt to tackle the detrimental effects of high milk production on other functional traits by altering the actual dynamics of lactation biology. For example, the idea of decreasing peak milk production through a slower increase in early lactation and that will increase the persistence of production in late lactation (Strucken et al. 2015).

We used a meta-analysis approach to understand the molecular mechanisms and pathways responsible for milk synthesis. A meta-analysis is a statistical approach that combines results from independent but related studies (Ramasamy et al. 2008). The identification of differentially expressed genes showed large inconsistencies in each study because of small sample sizes, low sample quality, and differences in laboratory protocol and platform (Lee et al. 2016). The meta-analysis approach overcomes these limitations. This approach enhances the statistical power in studies with small sample sizes, is a relatively inexpensive option, and increases the reliability and generalizability of results by combining information from multiple related existing studies (Ramasamy et al. 2008; Rung and Brazma 2013). This approach has been widely used to understand complex pathophysiological conditions, including infectious diseases (Pennings et al. 2008; Chang et al. 2011) and cancer (Chen et al. 2014; Xu et al. 2015).

Here, we performed a meta-analysis of publicly available cow RNA seq datasets, which covers three independent studies containing 16 mammary tissue/milk somatic cells samples. Bioinformatics analysis of RNA seq datasets revealed biological pathways responsible for milk synthesis and candidate genes responsible for milk production. This meta-analysis approach consolidates previous findings of lactation biology and provides a proof of principle for future studies combining information from diverse heterogeneous sources to understand the complete lactation biology.

Materials and methods

Selection of eligible RNA seq expression datasets

A search for RNA seq datasets was completed on NCBI Sequence Read Archive (SRA) and European Nucleotide Archive database using the following keywords: cow, mammary gland tissue, somatic cells, milk fat globules, milk yield, milk fat, and milk protein. Finally, we selected three independent studies that examined differential gene expression in cows at various stages of lactation under different conditions (Li et al. 2016; Yang et al. 2016; https://www.ncbi.nlm.nih.gov/sra/SRX3341240[accn]). Yang et al. (2016) studied differential gene expression in milk fat globules at days 10 and 70 post-calving between high and low 305-day milk yield, milk fat yield, and milk protein yield of Chinese Holstein cows. Li et al. (2016) studied the bovine transcriptome from the mammary tissue of six extremely high and six low phenotypic values for milk protein percentage at peak lactation. The third group studied the mammary transcriptome of six Chinese Holstein cows in peak lactation (https://www.ncbi.nlm.nih.gov/sra/?term=SRP122763). In total, 121.65 GB of data belonging to 16 RNA seq samples were included in the final meta-analysis (Table S1[1]).

RNA sequence genome mapping

We used HISAT (hierarchical indexing for spliced alignment of transcripts) spliced aligner for genome mapping of each RNA seq sample against the cow genome (Kim et al. 2015). The SAM/BAM file generated from the previous step and the GTF file downloaded from Ensembl database were input to cufflinks software for transcript assembly and Fragments per Kilobase of transcript per Million mapped reads (FPKM) estimates for each mapped RNA-Seq data (Trapnell et al. 2012). Our main objective was to understand the molecular biology of lactation in cows, so we did not perform any differential expression analysis. Instead, we compiled commonly expressed genes in all RNA seq samples, and the compiled commonly expressed genes profile was further used for functional and network analysis.

Functional annotation

Commonly expressed genes identified in this meta-analysis were functionally annotated with a publicly available annotation database. Initially, the commonly expressed genes were analyzed for gene ontology terms in Database for Annotation, Visualization and Integrated Discovery (DAVID, https://david.ncifcrf.gov/) and pathway analysis in Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis (http://www.genome.jp/kegg/).

Candidate genes

QTL mapping (Georges et al. 1995) and genome-wide association study (GWAS) (Mai et al. 2010; Bouwman et al. 2011) have detected a large number of candidate genes associated with milk yield and milk composition in dairy cows. A candidate genes list was downloaded from the cattle QTL database (QTLdb, http://www.animalgenome.org/cgi-bin/QTLdb/BT/index). This candidate genes list was then compared with a list of commonly expressed genes to identify reported candidate genes responsible for milk production in RNA Seq datasets.

[1]Supplementary data are available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/gen-2018-0144.

Gene regulatory networks construction

Genome-scale gene network

Gene regulatory networks have an important role to gain a more comprehensive understanding of the regulation of cellular processes and events (Li et al. 2013). In this study, we reconstructed the whole genome-scale gene network by using the high-performance web server DeGN Server (Li et al. 2013). We prepared the dataset for commonly expressed genes in a tab-delimited text file and uploaded it onto the DeGN Server. The network was constructed by context likelihood of relatedness method with the following parameter settings, including gene–gene association estimation method by mutual information and a cut-off threshold of 3.8. After the whole genome-scale network was reconstructed, we submitted identified candidate genes in the DeGN Server as a seed gene list to mine different subnetworks present in the whole genome-scale network. These subnetworks were visualized using SN Builder with a threshold cut-off of 3.8.

Protein–protein interaction (PPI) network

Protein interactions play fundamental roles in structuring and mediating essentially all biological processes. We used the online web server NetworkAnalyst for construction of protein–protein interaction (PPI) networks (Xia et al. 2015). The putative candidate genes were identified based on two widely used topological measures: degree and betweenness centrality. The degree of a node is the number of connections it has to other nodes, while betweenness measures the number of shortest paths going through the node. Nodes with a higher betweenness value are potentially important hubs in cellular signal trafficking (Xia et al. 2015).

Results and discussions

Our meta-analysis identified 11 562 genes that were commonly expressed in all 16 RNA-seq datasets. Milk components like lactose, proteins, and fats have to be synthesized in the mammary alveolar secretory cells (Strucken et al. 2015). However, the molecular mechanisms of synthesizing these milk components are not well known. Hence, we described the molecular understanding of milk component synthesis based on meta-analysis dataset.

Milk lactose synthesis

Cow milk contains 4.9% carbohydrate that is predominantly consisting of lactose. Lactose is a milk sugar unique to the mammary gland. Lactose is a disaccharide composed of the monosaccharides D-glucose and D-galactose. Lactose influences the osmotic pressure between blood and alveoli and thereby plays the main role in the quantity of milk secretion (Zhao and Keating 2007).

Lactose synthesis in the mammary gland involves the expression of a large number of genes (Ollier et al. 2007). For the synthesis of one molecule of lactose, glucose is converted to UDP-glucose, which in turn is converted to one UDP-galactose by an enzyme called UDP-glucose 4-epimerase (GALE). The initial step in the synthesis of UDP-galactose is phosphorylation of glucose by hexokinase 2 (HK2) (Zhao et al. 2012). Glucose and UDP-galactose are transported into the Golgi lumen by solute carrier family 35 member A2 (SLC35A2) and solute carrier family 35 member B1 (SLC35B1) (Mohammad et al. 2012), where lactose is synthesized in the Golgi lumen in the presence of a lactose synthase enzyme. Lactose synthase is composed of galactosyltransferase (GT) and α-lactalbumin (α-LA). ß1,4 galactosyltransferase (B4GALT3) is unique among all glycosyltransferases because of its substrate specificity, i.e., α-lactalbumin is expressed in the mammary gland resulting in lactose synthesis exclusively within the mammary gland (Hurley 2010). The protein kinase B alpha (PKB1, also known as AKT1) is an important regulator of lactose synthesis in the mammary gland of the cow (Lin et al. 2016). AKT1 acts as a potent survival factor for secretory epithelial cells (Kennedy et al. 1997) and the transport and metabolism of glucose (Plas and Thompson 2005). The *GALE*, *B4GALT3*, *SLC35A2*, *SLC35B1*, *HK2*, and *AKT1* encoding genes involved in lactose synthesis were expressed and presented expression values of 20.35, 7.81, 7.34, 45.14, 22.33, and 60.15, respectively, (except α-LA) and PI3K/AKT signaling pathway (Fig. 1), GO terms of cellular carbohydrate metabolic process (P-value 3.95E-07), galactosyltransferase activity (P-value 2.63E−08), and cellular component for Golgi membrane (P-value 3.00E-10) were enriched in our meta-analysis dataset (Tables 1 and 2). The protein encoded genes 6-phosphogluconolactonase (*PGLS*), ADP dependent glucokinase (*ADPGK*), beta-1,4-galactosyltransferase 3 (*B4GALT3*), beta-1,4-galactosyltransferase 7(*B4GALT7*), galactosidase alpha(*GLA*), fucosyltransferase 4(*FUT4*), and alpha-N-acetylgalactosaminidase (*NAGA*) involved in cellular carbohydrate metabolic process were expressed and presented expression values of 41.77, 5.65, 7.81, 4.47, 17.18, 1.34, and 46.97, respectively. Collectively, our results demonstrate that the expression of *GALE, B4GALT3*, *SLC35A2*, *SLC35B1*, *HK2*, and *AKT1* genes and activation of the PI3K/AKT signaling pathway contributes to lactose synthesis in the Golgi membrane of alveolar secretory cells.

Milk fat synthesis

Fats is a highly variable component of milk. Triglycerides are a major component (98%) of milk fat. Milk also contains other milk lipids including diacylglycerides (0.25%–0.48%), monoacylglycerides (0.02%–0.04%), phospholipids (0.6%–1.0%), cholesterol (0.2%–0.4%), glycolipids (0.006%), and free fatty acids (0.1%–0.4%) (Hurley 2010).

Milk fat triglycerides are synthesized in mammary epithelial cells. The precursor required for triglycerides are derived from blood or de novo synthesized within the

Fig. 1. Molecular pathway analysis of the mammary meta-analysis dataset. [Colour online.]

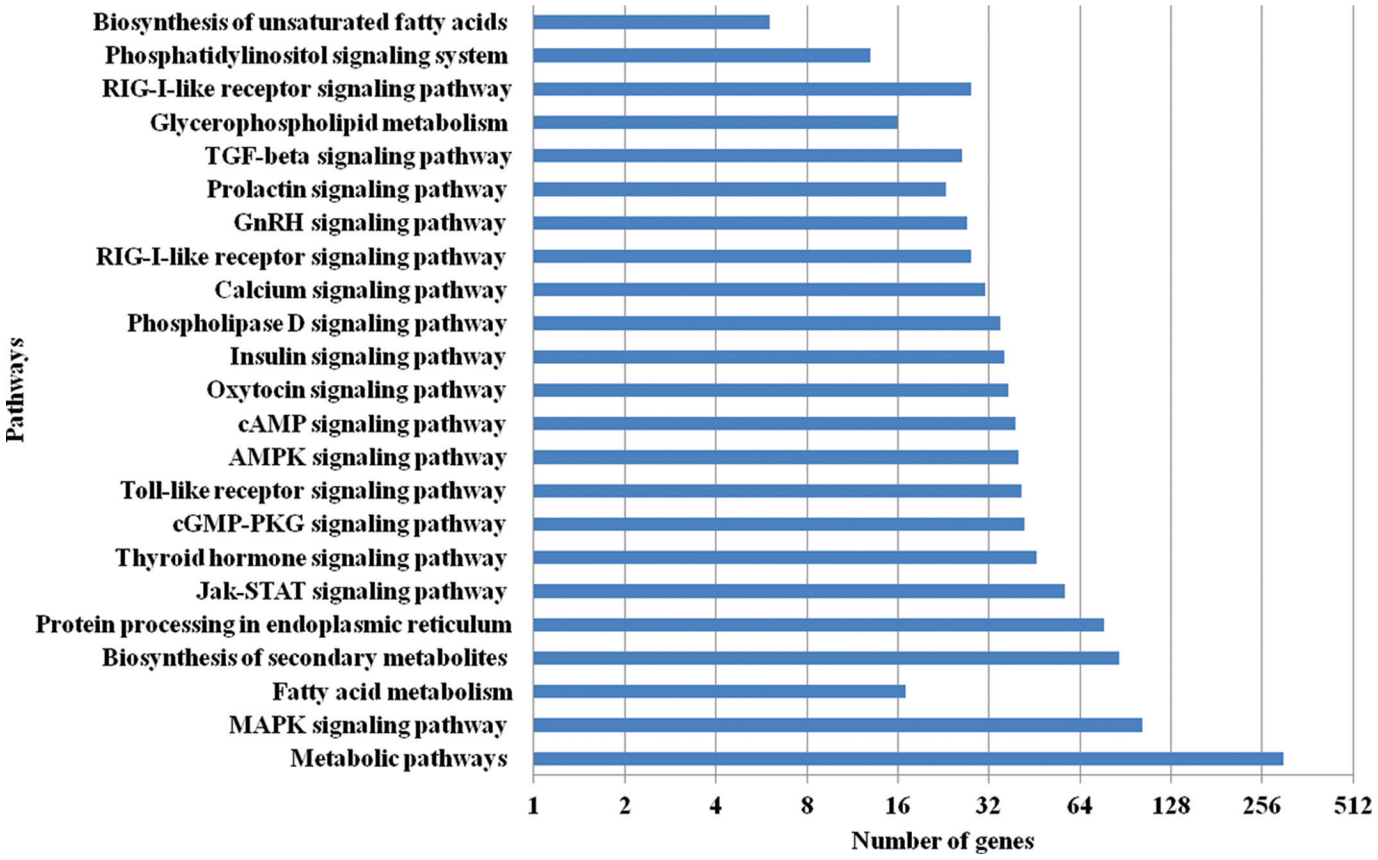


Table 1. Gene ontology terms related to mammary biological process enriched from the meta-analysis dataset.

Biological process	No. of genes	*P* value
Regulation of transcription from RNA polymerase II promoter	305	2.46E-36
Regulation of cytokine biosynthetic process	377	6.35E-27
Energy reserve metabolic process	144	6.86E-26
Lipid biosynthetic process	223	1.91E-24
Nitrogen compound metabolic process	146	1.04E-20
Steroid metabolic process	197	5.69E-18
Phosphatidylinositol biosynthetic process	161	6.03E-18
Fatty acid oxidation	168	1.46E-16
JAK-STAT cascade	164	4.51E-15
Tyrosine phosphorylation of STAT protein	137	5.61E-15
Inorganic anion transport	80	9.17E-15
Regulation of secretion	89	1.65E-13
Carbohydrate transport	90	9.84E-12
Rho protein signal transduction	61	1.16E-09
Positive regulation of transferase activity	92	1.49E-08
Regulation of intracellular transport	86	1.80E-08
Regulation of protein metabolic process	82	6.09E-08
Protein tetramerization	23	6.86E-08
Cellular carbohydrate metabolic process	49	3.95E-07
Glycolipid metabolic process	41	1.86E-06
One-carbon metabolic process	18	6.31E-06
Endoplasmic reticulum unfolded protein response	24	1.64E-05
Protein ubiquitination	10	2.21E-05

Table 2. Gene ontology terms related to mammary molecular function and cellular components enriched from the meta-analysis dataset.

Gene ontology terms	No. of genes	*P* value
Molecular function		
Structural molecule activity	598	1.45E-53
Lipase activity	193	6.99E-41
Ras GTPase binding	254	2.93E-31
Steroid dehydrogenase activity	149	1.27E-21
Amine transmembrane transporter activity	129	3.07E-20
Transcription cofactor activity	221	2.51E-15
Hydrolase activity	57	1.11E-13
Transferase activity	63	9.02E-13
Protein tyrosine phosphatase activity	130	1.08E-12
Protein serine/threonine kinase activity	26	4.17E-12
Pyrophosphatase activity	80	1.16E-11
Acetyltransferase activity	40	2.40E-11
Carbonate dehydratase activity	56	3.93E-11
Protein serine/threonine phosphatase activity	84	4.07E-11
Oxidoreductase activity	193	1.86E-10
Calcium channel activity	64	6.76E-09
Sodium channel activity	24	9.72E-09
Galactosyltransferase activity	83	2.63E-08
G-protein coupled receptor binding	30	1.26E-07
Phosphatase binding	46	5.47E-07
Carbohydrate kinase activity	161	4.43E-06
Growth factor binding	11	0.000569
Phospholipase activity	95	0.00112
Cellular component		
Chromosome	1810	4.00E-103
Nucleus	1720	4.16E-100
Nucleoplasm	765	8.44E-71
Proteasome complex	46	1.21E-16
Endocytic vesicle	135	1.61E-15
Organelle inner membrane	66	1.96E-14
Acetylcholine-gated channel complex	75	2.12E-11
Synaptic vesicle	23	5.63E-11
Golgi membrane	215	3.00E-10
Voltage-gated potassium channel complex	82	5.27E-10
Oligosaccharyltransferase complex	161	2.69E-09
Lysosomal membrane	70	6.66E-09
Microtubule associated complex	30	2.37E-08
Cytosol	156	2.44E-06
Transcription factor complex	15	3.16E-05
Endoplasmic reticulum	267	8.90E-07

mammary epithelial cells. The sources of blood-derived fatty acids are very low-density lipoproteins (VLDL) or chylomicrons (Moore and Christie 1979). The enzyme lipoprotein lipase (LPL) present in the mammary capillaries hydrolyzes the triglycerides in the VLDL, resulting in the release of free fatty acids, diacylglycerides, monoacylglycerides, or glycerol in the blood (Noble 1978; Moore and Christie 1979). These precursors are transported into mammary epithelial cells from blood for free fatty acids or triglycerides synthesis in mammary alveolar secretory cells (Hurley 2010).

Acetyl-CoA carboxylase (ACACA) and fatty acid synthetase (FASN) are two key enzymes required for fatty acid synthesis in the mammary gland. ACACA is the rate-limiting enzyme for the fatty acid synthesis pathway in mammary epithelial cells, and FASN is a large complex enzyme responsible for the chain elongation of the fatty acid chains. Milk fat triglycerides are synthesized on the smooth endoplasmic reticulum of epithelial cells and secreted as small fat droplets (Hurley 2010). Sterol regulatory element binding protein 1 (SREBP-1) is the master regulator of lipid metabolism in the ruminant mammary gland (Bionaz and Loor 2008; Ma and Corl 2012; Xu et al. 2016). Acyl-CoA synthetase short-chain family member 2 (ACSS2) and ATP citrate lyase (ACLY) both play a key role in lipogenesis by synthesizing acetyl-CoA from acetate

and converts glucose-derived citrate into acetyl-CoA, respectively, in goat mammary cells (Xu et al. 2018). The expression of *FASN*, *ACACA*, fatty acid desaturase 1(*FADS1*), elongase of very long chain fatty acid 6 (*ELOVL6*), fatty acid binding protein 3 (*FABP3*), cluster of differentiation 36 (*CD36*), glycerol-3-phosphate acyltransferase 1 (*GPAM*), Acyl Coenzyme A oxidase (*ACOX*), carnitine palmitoyltransferase 1A (*CPT1A*), Perilipin 3 (*PLIN3*), Lipin1 (*LPIN1*), and Peroxisome proliferator-activated receptor alpha (*PPARA*) genes were observed in mammary gland samples and presented expression values of 100.15, 16.47, 9.51, 0.39, 1112.43, 1730.25, 178.63, 9.03, 20.52, 26.56, and 10.35. Xu et al. (2018) reported the expression of lipogenic genes related to de novo fatty acid synthesis (*FASN*; *ACACA*), fatty acid desaturation (*FADS1*) and elongation (*ELOVL6*), long-chain fatty acid transportation (*FABP3*; *CD36*), TAG synthesis (*GPAM*), lipid oxidation (*ACOX*; *CPT1A*) and lipid droplet formation and secretion (*PLIN3*; *LPIN1*), and *PPARA* during simultaneous knockdown of *ACSS2* and *ACLY* genes in goat mammary cells. Different GO terms (lipid biosynthetic process, lipase activity, and endoplasmic reticulum) involved in the fat metabolism were enriched in this meta-analysis dataset (Tables 1 and 2). These results indicate that the expression of critical enzymes (LPL, ACACA, and FASN) and transcription factors (SREBP1 and PPARA) contribute to the lipid biosynthetic process in the endoplasmic reticulum of mammary alveolar secretory cells.

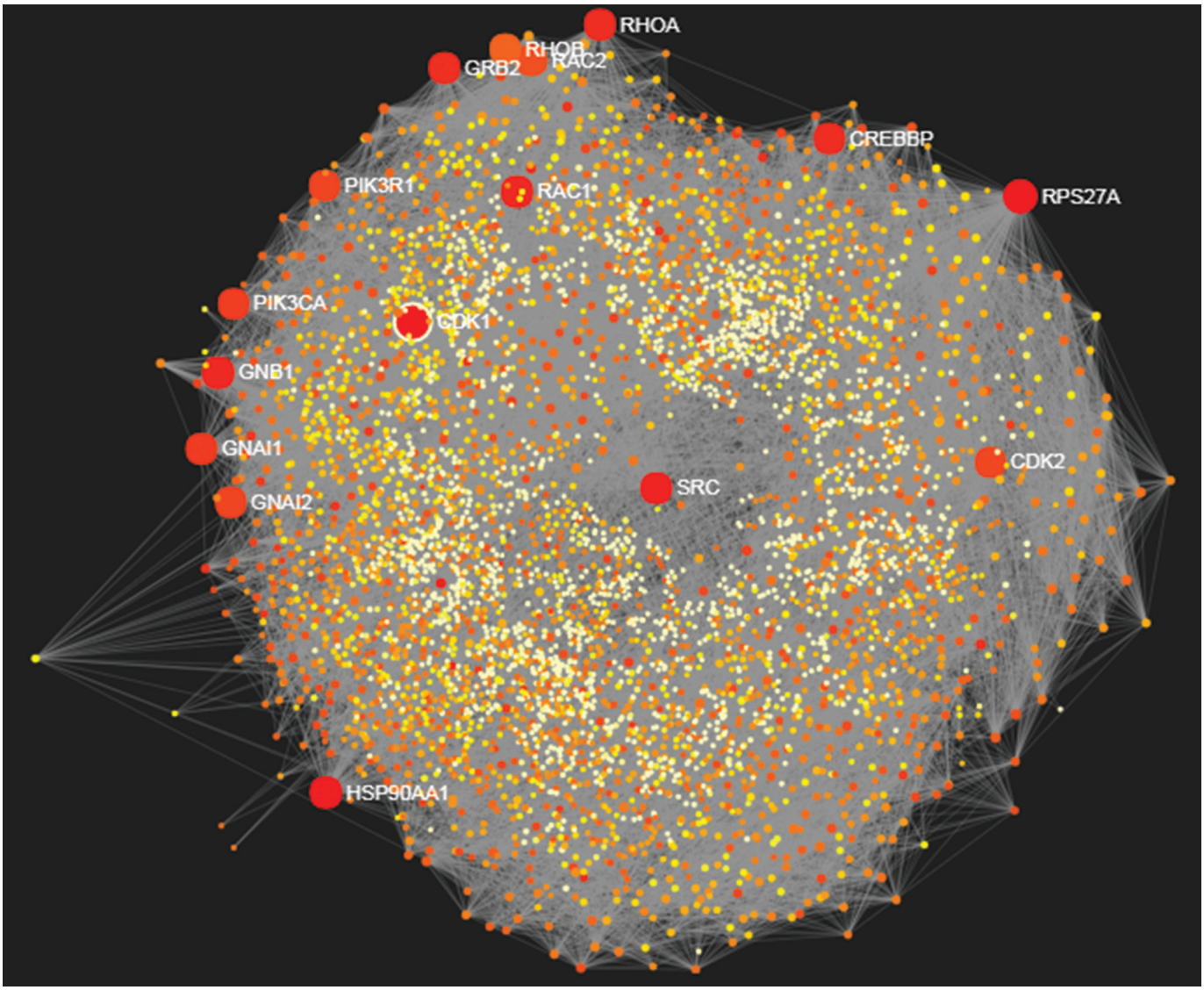
Milk protein synthesis

Cow milk contains 3.3% of protein, and these milk proteins are divided into caseins and whey proteins. Casein is one of the most abundant organic components and high-quality proteins of cow milk and is essential nutrients for neonatal growth (Whitney 1988). Alpha-S1-casein (CSN1S1), Beta-casein (CSN2), and Kappa-casein (CSN3) are synthesized in the mammary epithelial cells, and the amino acids for protein synthesis are from cows dietary or rumen microbial sources. The L-type amino acid transporter 1 (LAT1, encoded by SLC7A5) acts as a transporter of essential AA into mammary epithelial cells to maintain cell growth and casein protein synthesis (Lin et al. 2018). In mammary epithelial cells, protein synthesis occurs in the rough endoplasmic reticulum and is then transported into the Golgi apparatus for further post-translational processing (Hurley 2010). Glucose and amino acids may activate the mTOR signaling pathway leading to higher expression of LAT1 transporter, thus affecting milk casein protein synthesis (Appuhamy et al. 2014; Duan et al. 2017). Deficiency of glucose and amino acids reduced the casein gene transcription via inhibition of the Jak2/Stat5 pathway, and reduced translation via suppression of the mTOR pathway by activation of an AMPK signaling pathway (Zhang et al. 2018). The major milk proteins (CSN1S1, CSN2, and CSN3), minor milk proteins like beta-lactoglobulin, lactoperoxidase, and other proteases, *LAT1*, protease activators, nucleases, glycosidases, etc protein encoded genes expressed in the alveolar secretory cells. Further, JAK2/STAT5 and AMPK/mTOR signaling pathways and GO terms related to protein metabolism such as tyrosine phosphorylation of STAT protein, regulation of protein metabolic process, endoplasmic reticulum unfolded protein response, amine transmembrane transporter activity, and protein tyrosine phosphatase activity were enriched in our meta-analysis dataset (Tables 1 and 2).

Molecular pathway analysis

Commonly expressed genes were mapped into the KEGG pathway database. In total, 8550 genes were mapped into 365 KEGG reference pathways. From the 365 pathways, we manually selected those pathways relevant to lactation biology based on literature review and the number of mapped genes >10. From our meta-analysis dataset, KEGG pathway analysis enriched milk synthesis related pathways included the prolactin signaling pathway (number of genes involved n = 23), MAPK signaling pathway (n = 103), JAK/STAT signalling pathway (n = 57), mTOR signaling pathway (n = 51), AMPK signaling pathway (n = 40), oxytocin signaling pathway (n = 37), PI3K-Akt signaling pathway (n = 111), thyroid hormone signaling pathway (n = 46), fatty acid metabolism (n = 17), protein processing in endoplasmic reticulum (n = 77), and calcium signaling pathway (n = 31) (Fig. 1). In the thyroid hormone signaling pathway, a total of 46 genes were mapped into this pathway. This pathway is comprised of the MAPK signaling pathway, PI3K-Akt signaling pathway, calcium signaling pathway, HIF-1 signaling pathway, and mTOR signaling pathway. Genes including *SRC*, *THRB*, *THRA*, *PDPK1*, *CASP9*, *FOXO1*, *BAD*, *TBC1D4*, *MTOR*, *TSC2*, *HRAS*, *PRKCA*, *MAP2K2*, and *STAT1* were involved in this thyroid hormone signaling pathway. Previous studies have shown that thyroxine (or iodinated proteins) increased milk production by 10%–25% depending on a general increase in body metabolism (Blaxter et al. 1949; Capuco et al. 1999; Davis et al. 1988; Quevedo-Corona et al. 2000). This increased body metabolism may be due to activation of the above-mentioned pathways that will lead to more milk production in animals. However, further detailed studies are required to verify the thyroid hormone functional role in milk production. The stimulation of d-Glucose and amino acids phosphorylates and activates Jak2 at Tyr1007/1008 residues, and activated Jak2 phosphorylates latent STAT5 monomers on a conserved tyrosine region (Nan et al. 2014; Yang et al. 2015; Zhang et al. 2018). The phosphorylated STAT5 undergoes dimerization and then binds with specific elements of the casein promoter and induces the transcription of casein genes (Yamashita et al. 2001). Cellular energy stress activates AMP-activated protein kinase (AMPK) and activated AMPK inhibits ATP-consuming processes like protein synthesis. In bovine mammary epithelial cells, activated AMPK suppresses the global protein synthesis by inhibiting mTOR signal-

Fig. 2. Protein–protein interaction network constructed from the meta-analysis dataset. Genes that have a degree centrality score of ≥66 and a betweenness centrality score of ≥34 166 are consider as candidate genes involved in cellular signal trafficking and are highlighted in red.



ing (Burgos et al. 2013; Zhang et al. 2018). The PI3K/AKT pathway is involved in the regulation of diverse cellular functions, like apoptotic, metabolism, and cell cycle progression (Huang et al. 2018). PI3K activates AKT, activated AKT2 induces glycogen synthesis through inhibition of glycogen synthase kinase 3 (GSK-3) and FoxO transcription factors (FoxO1), protein synthesis via mTOR and downstream elements, and regulates fatty acid synthesis through mTORC1 and SREBP (Cross et al. 1995; Krycer et al. 2010; Hay 2011; Kousteni 2012). The prolactin signaling pathway plays an important role in lactation biology. Prolactin mediates its action through a transmembrane protein of the prolactin receptor, resulting in the activation of various cellular signaling cascades including Jak2/Stat, the major cascade, Src kinase, phosphatidylinositol-3-kinase (PI3K)/AKT, and mitogen-activated protein kinase (MAPK) pathways (https://www.genome.jp/dbget-bin/www_bget?ko04917). An oxytocin signaling pathway is responsible for milk letdown. The actions of oxytocin are mediated by oxytocin receptor binding resulted in the activation of the main signaling pathways like Gq/PLC/Ins3 pathway, MAPK, and the RhoA/Rho kinase pathways, contributing to direct contractile effect on myoepithelial cells (https://www.genome.jp/dbget-bin/www_bget?ko04921). Most of the RNA seq samples were included in the meta-analysis, the study samples were collected from either milk somatic cells or whole mammary tissue, hence a high number of immunity-related pathways were enriched, including the Toll-like receptor signaling pathway, NOD-like receptor signaling pathway, RIG-I-like receptor signaling pathway, TGF-beta signaling pathway, Jak-STAT signaling pathway, T cell receptor signaling pathway, TNF signaling pathway, cytokine-cytokine receptor interaction, chemokine signal-

Table 3. Candidate genes identified from protein–protein interaction (PPI) networks based on the degree of centrality and betweenness of centrality measures.

Candidate gene	Degree centrality	Betweenness centrality	Molecular function
RPS27A	373	2503600	Ribonucleoprotein
GNB1	185	616990	Transducer
RAC1	172	622170	GTPase activity
HSP90AA1	162	1746300	Chaperone
RHOA	144	442780	GTPase activity
CDK1	143	547050	Serine/threonine-protein kinase
GRB2	137	384210	Epidermal growth factor receptor binding
RAC2	136	80151	GTPase activity
SRC	136	814830	Serine/threonine-protein kinase
GNAI1	131	209980	Transducer
CREBBP	127	431500	Acyltransferase
GNAI2	123	146510	Transducer
RHOB	112	34166	GTPase activity
PIK3CA	111	189800	Serine/threonine-protein kinase
CDK2	109	128130	Serine/threonine-protein kinase
PIK3R1	104	133650	Phosphatidylinositol 3-kinase regulator activity
RPS3	98	134460	Ribonucleoprotein
CDKN1A	70	96555	Protein kinase inhibitor
PTK2	66	125970	Tyrosine-protein kinase

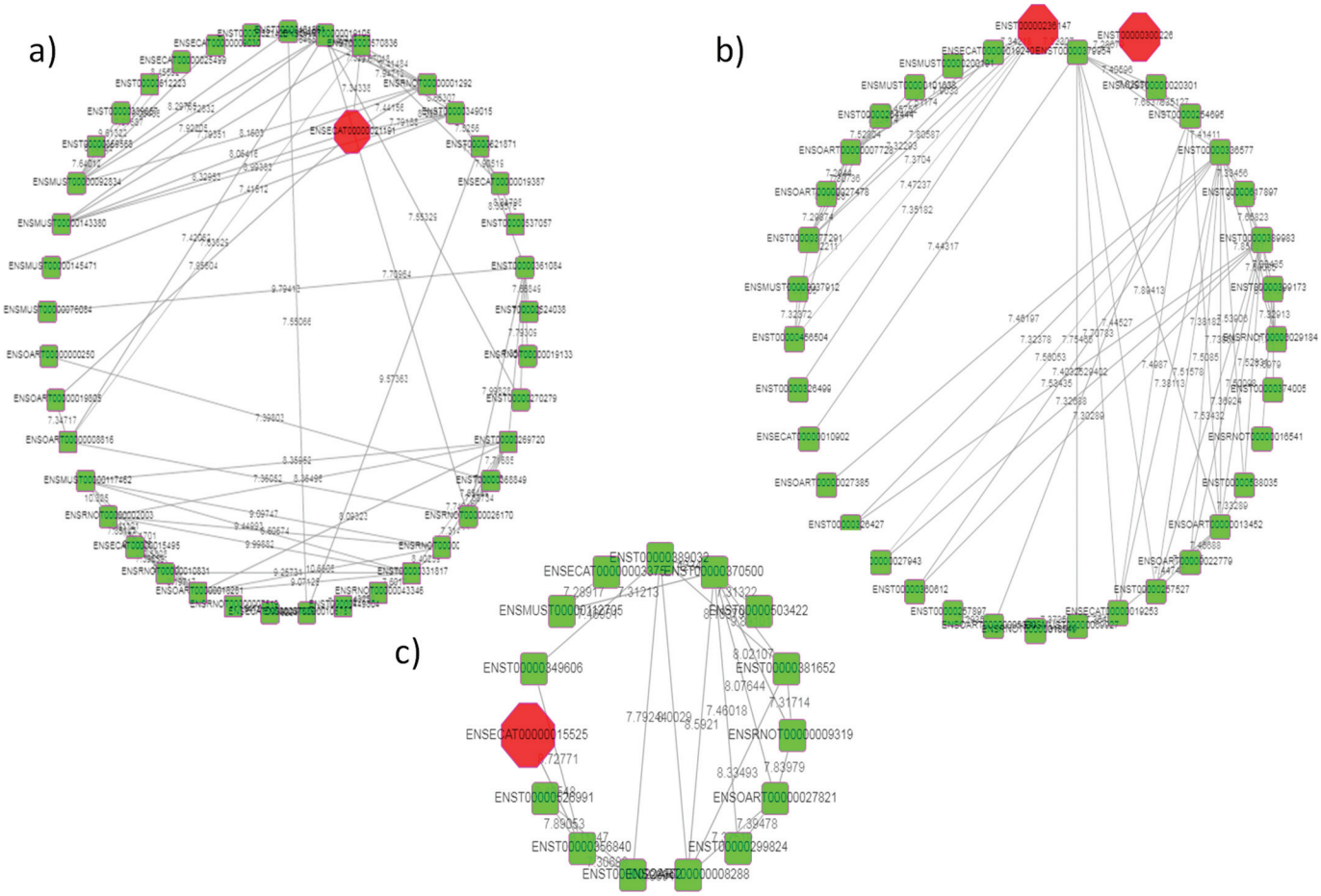
ing pathway, and NF-kappa B signaling pathwayin the meta-analysis dataset.

Candidate genes for milk production traits

Candidate genes responsible for milk yield, fat, and protein percentage (yield) were identified from our meta-analysis dataset. By comparing the meta-analysis dataset with QTLdb (http://www.animalgenome.org/cgi-bin/QTLdb/BT/index), we identified 130 candidate genes responsible for various milk traits in the meta-analysis dataset (Table S2[1]). Genes identified in the meta-analysis dataset that are responsible for milk yield include *IGF1*, *MAP4K4*, *SRC*, *ARL4A*, *SELL*, *SLC27A1*, *TXNDC5*, *ZBTB7A*, *ARFGEF1*, *CEP63*, *PEX2*, *PKIA*, *ABCG2*, and *ATP1A1* genes with expression values of 1.08, 1905.13, 1.20, 6.26, 30.56, 10.01, 47.00, 4.64, 5.68, 2.17, 2.92, 0.67, 197.83, and 93.83, respectively. Insulin-like growth factor 1 (IGF1) proteins are structurally and functionally related to the insulin hormone. IGF1 plays critical roles in the control of lactation, mammary gland development, growth processes, and fertility in cattle (Akers 2006; Lucy 2008). IGF1 is required for the critical control of nutrient utilization and partitioning (Chagas et al. 2007), thus enabling high milk production (Rose et al. 2005; Lucy 2008; Mullen et al. 2011). *IGF1* plays an important role in enhancing glucose uptake in cells and mediating cell growth and development (Jensen et al. 2018). IGF1 binds with the alpha subunit of IGF1R receptors initiating the IGFR1 signaling leading to activation of the PI3K-AKT/PKB and the Ras-MAPK pathways. Zhang et al. (2011) reported epidermal growth factor receptor (EGFR)-regulated MAPK/ERK with insulin-like signaling to control systemic glucose homeostasis. Mitogen-activated protein kinase 4 (MAP4K4) is a member of the serine/threonine protein kinase family. MAP4K4 plays a key role in the NF-kB signaling pathway for regulating the expression of pro-inflammatory genes in macrophages and could be associated with mastitis resistance (Gao et al. 2016; Bhattarai et al. 2017). Proto-oncogene tyrosine-protein kinase Src (SRC) is a non-receptor protein tyrosine kinase, which plays a pivotal role in cell signaling. SRC regulates the transcription/translation of milk casein protein genes expression in mammary epithelial cells (Watkin et al. 2008; Liu et al. 2010). Both *MAP4K4* and *SRC* genes have been used as biomarkers for cancer progression (Gao et al. 2016; Ke et al. 2016). Solute carrier family 27 member 1 (SLC27A1) plays an important role in mediating long-chain fatty acids trafficking across the plasma membrane, and a mutation in *SLC27A1* has potential effects on milk yield traits (Lv et al. 2011). ATP binding cassette subfamily G member 2 (*ABCG2*) has been detected in alveolar epithelial cells in the mammary gland and transports various xenobiotics, riboflavin, vitamin K3, and cholesterol into milk (Cohen-Zinder et al. 2005; van Herwaarden et al. 2007; Wu et al. 2008). In summary, the identified candidate genes were involved in nutrient partitioning, secretion of nutrients, nutrients uptake, and activation of various signaling pathways of the mammary gland. In addition, these genes were may act as casual gene/mutation responsible for high milk production; however, specific molecular mechanisms of these genes for milk yield is not known well.

The candidate genes responsible for milk protein percentage (yield) that were identified in our meta-analysis dataset include *GOSR2*, *SLC9A9*, *STAT5B*, *UBR5*, *RAC2*, *SLC38A3*, *LPIN1*, *SNX13*, *TBC1D22A*, *CA8*, *CTBP2*, and *HIF1A* genes with expression values of 6.00, 2.23, 4.59, 4.17,

Fig. 3. Subnetworks identified from the whole genome-scale network. Genes highlighted in red (ENSECAT00000021191-FASN; ENST00000236147-SELL; ENST00000300228-MS4A8; ENSECAT00000015525-GABARAPL1) are identified as candidate genes.



86.96, 9.87, 18.78, 1.5, 5.74, 1.91, 16.32, and 63.11, respectively. Ras-related C3 botulinum toxin substrate 2 (RAC2) is a member of the Ras superfamily of small GTPase and regulates diverse cellular processes like secretory processes, phagocytosis, and epithelial cell polarization. Protein transport from the cis/medial-Golgi to the trans-Golgi network is mediated by Golgi SNAP receptor complex member 2 (GOSR2). Signal transducer and activator of transcription 5B (STAT5B) is a member of the STAT family of transcription factors. These act as transcription activators for adult mammary gland development and TCR signaling (Mohankumar et al. 2008). Prolactin binding to a specific prolactin membrane receptor activates JAK2 kinases and subsequent phosphorylation STAT5 transcription factors. This leads to the activation of the JAK/STAT signaling pathway, which may be responsible for lactation and reproduction in mammals (Watson 2001; Khatib et al. 2008). E3 ubiquitin-protein ligase UBR5 (UBR5) is a targeting-specific protein for ubiquitin-mediated proteolysis and plays an important role in the control of cell progression. Phosphatidate phosphatase (LPIN1) is a magnesium-dependent phosphatidate phosphatase enzyme that catalyzes the dephosphorylation of phosphatidic acid to yield diacylglycerol during triglyceride biosynthesis in the reticulum endoplasmic membrane (Temprano et al. 2016; Viale et al. 2017). The *SNX13* gene belongs to the SNX family involved in intracellular trafficking and is proposed as a candidate genes for milk production traits (Worby and Dixon 2002; Rincón et al. 2009).

The following genes were identified as candidate genes responsible for fat percentage (yield) in our meta-analysis dataset: *MFGE8*, *PTK2*, *SCARB1*, *PDE1B*, *SLC8A1*, *TRAPPC9, IFIH1*, and *IGF1R* genes with expression values of 782.05, 8.56, 12.23, 1.88, 2.37, 8.03, 13.85, and 4.98, respectively. Protein-tyrosine kinase 2 (*PTK2*) gene encodes a cytoplasmic non-receptor protein tyrosine kinase and plays a prominent role in maintaining mammary gland development and function (Nagy et al. 2007). Further, the genetic variants in *PTK2* gene were implicated in milk production ability of dairy cattle (Wang et al. 2013). The activation of PTK2 is required for cell growth and numerous intracellular signaling pathways including MAPK1/ERK2, MAPK3/ERK1, and MAP. *PDE1B* gene belongs to the cyclic nucleotide phosphodiesterase family and regulates the activity of second messengers (cAMP and cGMP) (Bender and Beavo 2006). Scavenger receptor class B member 1 (SCARB1) is a plasma membrane receptor

for phospholipids, cholesterol ester, lipoproteins, high-density lipoprotein cholesterol, phosphatidylserine, and apoptotic cells (Morel et al. 2018). Milk fat globule-EGF factor 8 (MFGE8) protein is a preproprotein that is proteolytically processed into different protein products, and these proteins are involved in various cellular activity like wound healing, epithelial homeostasis, and phagocytic removal of apoptotic cells (https://www.ncbi.nlm.nih.gov/gene?cmd=Retrieve&dopt=full_report&list_uids=4240). Insulin-like growth factor 1 receptor (IGF1R) is a receptor for insulin-like growth factor 1, and activated IGF1R is involved in cell growth and survival control (https://www.uniprot.org/uniprot/P08069). IGF1R is primarily involved in mammary gland development and milk production performance of cows (Lawrence et al. 2007; Szewczuk 2017). The IGF-I/IGF1R signaling pathway is important for normal development of mammary gland tissue, pregnancy, and lactation of dairy cows (Plath-Gabler et al. 2001; Hvid et al. 2011; Yonekura et al. 2015).

Gene interaction networks analysis

Protein–protein interaction (PPI) networks have emerged as an important resource to understand high-throughput genomics data. Protein interactions play fundamental roles in structuring and mediating fundamentally all cell biological processes. PPI networks are often presented as undirected graphs with proteins as nodes and edges indicating interactions between two connecting proteins. PPI networks were constructed for the meta-analysis dataset (Fig. 2). Candidate genes were also identified from PPI networks using two widely used topological measures—degree and betweenness centrality (Table 3). To consider genes as candidate genes involved in cellular signal trafficking, we applied a cutoff score of ≥66 for degree centrality and ≥34 166 for betweenness . With these criteria, we identified a few previously reported candidate genes (*PTK2*, *RAC2*, *SRC*, and *CDKN1A*) and several putative candidate genes (*RPS27A*, *GNB1*, *RAC1*, *HSP90AA1*, *RHOA*, *CDK1*, *GRB2*, etc.) related to milk yield, milk protein, and milk fat percentage (yield).

Analyses of whole genome-scale networks of mammary gland provides a holistic view of all transcription regulations among and within different subnetworks. The genome-scale networks was comprised of 11 564 genes and 370 551 links were reconstructed with a z-score threshold of 3.8 and a mutual information-based association method. The subnetworks were identified from whole genome-scale networks, and further, we identified important candidate genes (*FASN*, *MS4A8*, *SELL*, and *GABARAPL1*) in the subnetworks (Fig. 3). The identified candidate genes in the subnetworks have been reported to be responsible for milk production traits.

Conclusion

In conclusion, this study described the molecular understanding of milk fat, milk protein, and lactose synthesis process in mammary alveolar secretory cells. Further, this study reported several molecular pathways responsible for milk synthesis and identified many candidate genes responsible for milk production traits. Overall this study explained the application of whole transcriptome techniques of lactation biology for the molecular understanding and identification of causal genes responsible for the milk production traits and provides basic information for lactation biology. Candidate genes/mutations assisted selection will provide biologically and functionally meaningful selection methods, because this selection method can act on the causative genes/mutations directly instead of having to rely on linkage equilibrium between molecular markers and causative genes/mutations. Through this approach, genetic improvement of animal production, health, and welfare will be faster and a more sustainable manner can be achieved. However, the exact functional roles of candidate genes in various milk production traits are not known. Casual mutation and gene signatures responsible for high and low milk yield, milk fat yield, and milk protein yield are not yet completely determined. Hence, further detailed functional studies are required for the identification and understanding of the specific functional roles of candidate genes/mutations responsible for various milk production traits.

Competing interests

The authors have declared that no competing interests exist.

References

Akers, R.M. 2006. Major advances associated with hormone and growth factor regulation of mammary growth and lactation in dairy cows. J. Dairy Sci. **89**: 1222–1234. doi:10.3168/jds.S0022-0302(06)72191-9. PMID:16537955.

Appuhamy, J.A., Nayananjalie, W.A., England, E.M., Gerrard, D.E., Akers, R.M., and Hanigan, M.D. 2014. Effects of AMP-activated protein kinase (AMPK) signaling and essential amino acids on mammalian target of rapamycin (mTOR) signaling and protein synthesis rates in mammary cells. J. Dairy Sci. **97**(1): 419–429. doi:10.3168/jds.2013-7189. PMID:24183687.

Bender, A.T., and Beavo, J.A. 2006. PDE1B2 regulates cGMP and a subset of the phenotypic characteristics acquired upon macrophage differentiation from a monocyte. Proc. Natl. Acad. Sci. U.S.A. **103**(2): 460–465. doi:10.1073/pnas.0509972102. PMID:16407168.

Bhattarai, D., Chen, X., Ur Rehman, Z., Hao, X., Ullah, F., Dad, R., et al. 2017. Association of MAP4K4 gene single nucleotide polymorphism with mastitis and milk traits in Chinese Holstein cattle. J. Dairy Res. **84**(1): 76–79. doi:10.1017/S0022029916000832. PMID:28252361.

Bionaz, M., and Loor, J.J. 2008. Gene networks driving bovine milk fat synthesis during the lactation cycle. BMC Genomics, **9**: 366. doi:10.1186/1471-2164-9-366. PMID:18671863.

Blaxter, K.L., Reineke, E.P., Crampton, E.S., and Petersen, W.E. 1949. The role of thyroidal materials and of synthetic goitrogens in animal production and an appraisal of their practical use. J. Anim. Sci. **8**(3): 307–352. doi:10.2527/jas1949.83307x.

Bouwman, A.C., Bovenhuis, H., Visker, M.H.P.W., and van Arendonk, J.A.M. 2011. Genome wide association of milk fatty acids in Dutch dairy cattle. BMC Genet. **12**: 43. doi:10.1186/1471-2156-12-43. PMID:21569316.

Burgos, S.A., Kim, J.J., Dai, M., and Cant, J.P. 2013. Energy depletion of bovine mammary epithelial cells activates AMPK and suppresses protein synthesis through inhibition of mTORC1 signaling. Horm. Metab. Res. **45**(3): 183–189. doi:10.1055/s-0032-1323742. PMID:22972179.

Capuco, A.V., Kahl, S., Jack, L.J., Bishop, J.O., and Wallace, H. 1999. Prolactin and growth hormone stimulation of lactation in mice requires thyroid hormones. Proc. Soc. Exp. Biol. Med. **221**(4): 345–351. doi:10.3181/00379727-221-44417. PMID: 10460696.

Cattle QTL Database - Animal Genome Databases. 2018. Available from http://www.animalgenome.org/cgi-bin/QTLdb/BT/index [accessed 1 January 2018].

Chagas, L.M., Bass, J.J., Blache, D., Burke, C.R., Kay, J.K., Lindsay, D.R., et al. 2007. Invited review: new perspectives on the roles of nutrition and metabolic priorities in the subfertility of high-producing dairy cows. J. Dairy Sci. **90**: 4022–4032. doi:10.3168/jds.2006-852. PMID:17699018.

Chang, S.T., Tchitchek, N., Ghosh, D., Benecke, A., and Katze, M.G. 2011. A chemokine gene expression signature derived from meta-analysis predicts the pathogenicity of viral respiratory infections. BMC Syst. Biol. **5**: 202. doi:10.1186/1752-0509-5-202. PMID:22189154.

Chen, R., Khatri, P., Mazur, P.K., Polin, M., Zheng, Y., Vaka, D., et al. 2014. A meta-analysis of lung cancer gene expression identifies PTK7 as a survival gene in lung adenocarcinoma. Cancer Res. **74**(10): 2892–2902. doi:10.1158/0008-5472.CAN-13-2775. PMID:24654231.

Cohen-Zinder, M., Seroussi, E., Larkin, D.M., Loor, J.J., Everts-van der Wind, A., Lee, J.H., et al. 2005. Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. Genome Res. **15**(7): 936–944. doi:10.1101/gr.3806705. PMID:15998908.

Cross, D.A., Alessi, D.R., Cohen, P., Andjelkovich, M., and Hemmings, B.A. 1995. Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase B. Nature, **378**(6559): 785–789. doi:10.1038/378785a0. PMID:8524413.

Dairy production and products: Milk composition - FAO. 2018. Available from http://www.fao.org/dairy-production-products/products/milk-composition/en/ [accessed 3 May 2018].

DAVID Functional Annotation Bioinformatics Microarray Analysis. 2018. Available from https://david.ncifcrf.gov [accessed 16 February 2018].

Davis, S.R., Collier, R.J., McNamara, J.P., Head, H.H., and Sussman, W. 1988. Effects of thyroxine and growth hormone treatment of dairy cows on milk yield, cardiac output and mammary blood flow. J. Anim. Sci. **66**(1): 70–79. doi:10.2527/jas1988.66170x. PMID:3366719.

Duan, X., Lin, Y., Lv, H., Yang, Y., Jiao, H., and Hou, X. 2017. Methionine induces LAT1 expression in dairy cow mammary gland by activating the mTORC1 signaling pathway. DNA Cell Biol. **36**(12): 1126–1133. doi:10.1089/dna.2017.3792. PMID: 29040000.

Gao, X., Gao, C., Liu, G., and Hu, J. 2016. MAPK4: an emerging therapeutic target in cancer. Cell Biosci. **6**: 56. doi:10.1186/s13578-016-0121-7. PMID:27800153.

Georges, M., Nielsen, D., Mackinnon, M., Mishra, A., Okimoto, R., Pasquino, A.T., et al. 1995. Mapping quantitative traitloci controlling milk production in dairy cattle by exploiting progeny testing. Genetics, **139**(2): 907–920. PMID: 7713441.

Hay, N. 2011. Interplay between FOXO, TOR, and Akt. Biochim. Biophys. Acta, **1813**(11): 1965–1970. doi:10.1016/j.bbamcr.2011.03.013. PMID:21440577.

Huang, X., Liu, G., Guo, J., and Su, Z. 2018. The PI3K/AKT pathway in obesity and type 2 diabetes. Int. J. Biol. Sci. **14**(11): 1483–1496. doi:10.7150/ijbs.27173. PMID:30263000.

Hurley, W.L. 2010. Lactation Biology Website. Available from http://ansci.illinois.edu/static/ansc438/Milkcompsynth/milkcompsynthresources.html [accessed 2 June 2018].

Hvid, H., Klopfleisch, R., Vienberg, S., Hansen, B.F., Thorup, I., Jensen, H.E., and Oleksiewicz, M.B. 2011. Unique expression pattern of the three insulin receptor family members in the rat mammary gland: dominance of IGF-1R and IRR over the IR, and cyclical IGF-1R expression. J. Appl. Toxicol. **31**(4): 312–328. doi:10.1002/jat.1627. PMID:21259294.

Jensen, R.B., Thankamony, A., Holst, K.K., Janssen, J.A.M.J.L., Juul, A., Dunger, D., et al. 2018. Genetic influence on the associations between IGF-I and glucose metabolism in a cohort of elderly twins. Eur. J. Endocrinol. **178**(2): 155–163. doi:10.1530/EJE-17-0754.

Ke, L., Xiang, Y., Guo, X., Lu, J., Xia, W., Yu, Y., et al. 2016. c-Src activation promotes nasopharyngeal carcinoma metastasis by inducing the epithelial-mesenchymal transition via PI3K/Akt signaling pathway: a new and promising target for NPC. Oncotarget, **7**(19): 28340–28355. doi:10.18632/oncotarget.8634. PMID:27078847.

KEGG: Kyoto Encyclopedia of Genes and Genomes - GenomeNet. 2018*a*. Available from http://www.genome.jp/kegg/ [accessed 6 February 2018].

KEGG: Kyoto Encyclopedia of Genes and Genomes - GenomeNet. 2018*b*. Available from https://www.genome.jp/dbget-bin/www_bget?ko04917 [accessed 15 April 2018].

KEGG: Kyoto Encyclopedia of Genes and Genomes - GenomeNet. 2018. Available from https://www.genome.jp/dbget-bin/www_bget?ko04921 [accessed 15 April 2018].

Kennedy, S.G., Wagner, A.J., Conzen, S.D., Jordán, J., Bellacosa, A., Tsichlis, P.N., and Hay, N. 1997. The PI 3-kinase/Akt signaling pathway delivers an anti-apoptotic signal. Genes Dev. **11**(6): 701–713. doi:10.1101/gad.11.6.701. PMID: 9087425.

Khatib, H., Monson, R.L., Schutzkus, V., Kohl, D.M., Rosa, G.J., and Rutledge, J.J. 2008. Mutations in the STAT5A gene are associated with embryonic survival and milk composition in cattle. J. Dairy Sci. **91**(2): 784–793. doi:10.3168/jds.2007-0669. PMID:18218766.

Kim, D., Langmead, B., and Salzberg, S.L. 2015. HISAT: a fast spliced aligner with low memory requirements. Nat. Methods, **12**(4): 357–360. doi:10.1038/nmeth.3317. PMID:25751142.

Kousteni, S. 2012. FoxO1, the transcriptional chief of staff of energy metabolism. Bone, **50**(2): 437–443. doi:10.1016/j.bone.2011.06.034. PMID:21816244.

Krycer, J.R., Sharpe, L.J., Luu, W., and Brown, A.J. 2010. The Akt-SREBP nexus: cell signaling meets lipid metabolism. Trends Endocrinol. Metab. **21**(5): 268–276. doi:10.1016/j.tem.2010.01.001. PMID:20117946.

Lawrence, M.C., McKern, N.M., and Ward, C.W. 2007. Insulin receptor structure and its implications for the IGF-1 receptor. Curr. Opin. Struct. Biol. **17**(6): 699–705. doi:10.1016/j.sbi.2007.07.007. PMID:17851071.

Lee, Y.S., Hwang, S.G., Kim, J.K., Park, T.H., Kim, Y.R., Myeong, H.S., et al. 2016. Identification of novel therapeutic target genes in acquired lapatinib-resistant breast cancer by integrative meta-analysis. Tumour Biol. **37**(2): 2285–2297. doi:10.1007/s13277-015-4033-7. PMID:26361955.

Li, C., Cai, W., Zhou, C., Yin, H., Zhang, Z., Loor, J.J., et al. 2016. RNA-Seq reveals 10 novel promising candidate genes affecting milk protein concentration in the Chinese Holstein population. Sci. Rep. **6**: 26813. doi:10.1038/srep26813. PMID: 27254118.

Li, J., Wei, H., and Zhao, P.X. 2013. DeGN Server: deciphering genome-scale gene networks through high performance re-

verse engineering analysis. Biomed. Res. Int. **2013**: 856325. doi:10.1155/2013/856325. PMID:24328032.

Lin, Y., Sun, X., Hou, X., Qu, B., Gao, X., and Li, Q. 2016. Effects of glucose on lactose synthesis in mammary epithelial cells from dairy cow. BMC Vet. Res. **12**: 81. doi:10.1186/s12917-016-0704-x. PMID:27229304.

Lin, Y., Duan, X., Lv, H., Yang, Y., Liu, Y., Gao, X., and Hou, X. 2018. The effects of L-type amino acid transporter 1 on milk protein synthesis in mammary glands of dairy cows. J. Dairy Sci. **101**(2): 1687–1696. doi:10.3168/jds.2017-13201. PMID:29224866.

Liu, W., Wang, J., Li, Q., Ju, Z., Huang, J., Wang, H., et al. 2010. Correlation analysis between three novel SNPs of the Src gene in bovine and milk production traits. Mol. Biol. Rep. **37**(8): 3771–3777. doi:10.1007/s11033-010-0031-4. PMID:20213510.

Lucy, M.C. 2008. Functional differences in the growth hormone and insulin-like growth factor axis in cattle and pigs: implications for post-partum nutrition and reproduction. Reprod. Domest. Anim. **43**(Suppl. 2): 31–39. doi:10.1111/j.1439-0531.2008.01140.x. PMID:18638098.

Lv, Y., Wei, C., Zhang, L., Lu, G., Liu, K., and Du, L. 2011. Association between polymorphisms in the SLC27A1 gene and milk production traits in Chinese Holstein cattle. Anim. Biotechnol. **22**(1): 1–6. doi:10.1080/10495398.2011.527567. PMID:21328100.

Ma, L., and Corl, B.A. 2012. Transcriptional regulation of lipid synthesis in bovine mammary epithelial cells by sterol regulatory element binding protein-1. J. Dairy Sci. **95**(7): 3743–3755. doi:10.3168/jds.2011-5083. PMID:22720931.

Mai, M.D., Sahana, G., Christiansen, F.B., and Guldbrandtsen, B. 2010. A genome-wide association study for milk production traits in Danish Jersey cattle using a 50 K single nucleotide polymorphism chip. J. Anim. Sci. **88**(11): 3522–3528. doi:10.2527/jas.2009-2713. PMID:20656975.

Mohammad, M.A., Hadsell, D.L., and Haymond, M.W. 2012. Gene regulation of UDP-galactose synthesis and transport: potential rate-limiting processes in initiation of milk production in humans. Am. J. Physiol. Endocrinol. Metab. **303**(3): E365–E376. doi:10.1152/ajpendo.00175.2012. PMID:22649065.

Mohankumar, K.M., Perry, J.K., Kannan, N., Kohno, K., Gluckman, P.D., Emerald, B.S., and Lobie, P.E. 2008. Transcriptional activation of signal transducer and activator of transcription (STAT) 3 and STAT5B partially mediate homeobox A1-stimulated oncogenic transformation of the immortalized human mammary epithelial cell. Endocrinology, **149**(5): 2219–2229. doi:10.1210/en.2007-1320. PMID:18276758.

Moore, J.H., and Christie, W.W. 1979. Lipid metabolism in the mammary gland of ruminant animals. Prog. Lipid Res. **17**(4): 347–395. doi:10.1016/0079-6832(79)90012-0. PMID:38463.

Morel, E., Ghezzal, S., Lucchi, G., Truntzer, C., Pais de Barros, J.P., Simon-Plas, F., et al. 2018. Cholesterol trafficking and raft-like membrane domain composition mediate scavenger receptor class B type 1-dependent lipid sensing in intestinal epithelial cells. Biochim. Biophys. Acta, **1863**(2): 199–211. doi:10.1016/j.bbalip.2017.11.009.

Mullen, M.P., Lynch, C.O., Waters, S.M., Howard, D.J., O'Boyle, P., Kenny, D.A., et al. 2011. Single nucleotide polymorphisms in the growth hormone and insulin-like growth factor-1 genes are associated with milk production, body condition score and fertility traits in dairy cows. Genet. Mol. Res. **10**(3): 1819–1830. doi:10.4238/vol10-3gmr1173. PMID:21948746.

Nagy, T., Wei, H., Shen, T.L., Peng, X., Liang, C.C., Gan, B., and Guan, J.L. 2007. Mammary epithelial-specific deletion of the focal adhesion kinase gene leads to severe lobulo-alveolar hypoplasia and secretory immaturity of the murine mammary gland. J. Biol. Chem. **282**(43): 31766–31776. doi:10.1074/jbc.M705403200. PMID:17716968.

Nan, X., Bu, D., Li, X., Wang, J., Wei, H., Hu, H., et al. 2014. Ratio of lysine to methionine alters expression of genes involved in milk protein transcription and translation and mTOR phosphorylation in bovine mammary cells. Physiol. Genomics, **46**(7): 268–275. doi:10.1152/physiolgenomics.00119.2013. PMID:24474444.

NCBI: Gene Database. 2018. Available from https://www.ncbi.nlm.nih.gov/gene?cmd=Retrieve&dopt=full_report&list_uids=4240 [accessed 25 June 2018].

NCBI: Sequence Read Archive (SRA). 2017. Available from https://www.ncbi.nlm.nih.gov/sra/SRX3341240[accn] [accessed 20 December 2017].

Noble, R.C. 1978. Digestion, absorption and transport of lipids in ruminant animals. Prog. Lipid Res. **17**(1): 55–91. doi:10.1016/0079-6832(78)90005-8. PMID:34168.

Ollier, S., Robert-Granié, C., Bernard, L., Chilliard, Y., and Leroux, C. 2007. Mammary transcriptome analysis of food-deprived lactating goats highlights genes involved in milk secretion and programmed cell death. J. Nutr. **137**(3): 560–567. doi:10.1093/jn/137.3.560. PMID:17311940.

Pennings, J., Kimman, T.G., and Janssen, R. 2008. Identification of a common gene expression response in different lung inflammatory diseases in rodents and macaques. PLoS ONE, **3**(7): e2596. doi:10.1371/journal.pone.0002596. PMID:18612392.

Plas, D.R., and Thompson, C.B. 2005. Akt-dependent transformation: there is more to growth than just surviving. Oncogene, **24**(50): 7435–7442. doi:10.1038/sj.onc.1209097. PMID:16288290.

Plath-Gabler, A., Gabler, C., Sinowatz, F., Berisha, B., and Schams, D. 2001. The expression of the IGF family and GH receptor in the bovine mammary gland. J. Endocrinol. **168**(1): 39–48. doi:10.1677/joe.0.1680039. PMID:11139768.

Quevedo-Corona, L., Franco-Colin, M., Caudillo-Romero, M., Pacheco-Rosado, J., Zamudio-Hernandez, S., and Racotta, R. 2000. 3,5,3′-Triiodothyronine administered to rat dams during lactation increases milk yield and triglyceride concentration and hastens pups growth. Life Sci. **66**(21): 2013–2121. PMID:10823341.

Ramasamy, A., Mondry, A., Holmes, C.C., and Altman, D.G. 2008. Key issues in conducting a meta-analysis of gene expression microarray datasets. PLoS Med. **5**(9): e184. doi:10.1371/journal.pmed.0050184. PMID:18767902.

Rincón, G., Islas-Trejo, A., Casellas, J., Ronin, Y., Soller, M., Lipkin, E., and Medrano, J.F. 2009. Fine mapping and association analysis of a quantitative trait locus for milk production traits on *Bos taurus* autosome 4. J. Dairy Sci. **92**(2): 758–764. doi:10.3168/jds.2008-1395. PMID:19164688.

Rose, M.T., Weekes, T.E., and Rowlinson, P. 2005. Correlation of blood and milk components with the milk yield response to bovine somatotropin in dairy cows. Domest. Anim. Endocrinol. **28**(3): 296–307. doi:10.1016/j.domaniend.2004.12.001. PMID:15760670.

Rung, J., and Brazma, A. 2013. Reuse of public genome-wide gene expression data. Nat. Rev. Genet. **14**(2): 89–99. doi:10.1038/nrg3394. PMID:23269463.

Strucken, E.M., Laurenson, Y.C.S.M., and Brockmann, G.A. 2015. Go with the flow—biology and genetics of the lactation cycle. Front. Genet. **6**: 118. doi:10.3389/fgene.2015.00118. PMID:25859260.

Szewczuk, M. 2017. Polymorphism in exon 2 encoding the putative ligand binding pocket of the bovine insulin-like growth factor 1 receptor affects milk traits in four different cattle breeds. J. Anim. Breed Genet. **134**(1): 34–42. doi:10.1111/jbg.12216. PMID:27112238.

Temprano, A., Sembongi, H., Han, G.S., Sebastián, D., Capellades, J., Moreno, C., et al. 2016. Redundant roles of the phosphatidate phosphatase family in triacylglycerol synthesis in human adipocytes. Diabetologia, **59**(9): 1985–1994. doi:10.1007/s00125-016-4018-0. PMID:27344312.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., et al. 2012. Differential gene and transcript expression anal-

ysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. **7**(3): 562–578. doi:10.1038/nprot.2012.016. PMID:22383036.

UniProt database. 2018. Available from https://www.uniprot.org/uniprot/P08069 [accessed 23 June 2018].

van Herwaarden, A.E., Wagenaar, E., Merino, G., Jonker, J.W., Rosing, H., Beijnen, J.H., and Schinkel, A.H. 2007. Multidrug transporter ABCG2/breast cancer resistance protein secretes riboflavin (vitamin B2) into milk. Mol. Cell Biol. **27**(4): 1247–1253. doi:10.1128/MCB.01621-06. PMID:17145775.

Viale, E., Tiezzi, F., Maretto, F., De Marchi, M., Penasa, M., and Cassandro, M. 2017. Association of candidate gene polymorphisms with milk technological traits, yield, composition, and somatic cell score in Italian Holstein-Friesian sires. J. Dairy Sci. **100**(9): 7271–7281. doi:10.3168/jds.2017-12666. PMID:28711251.

Wang, H., Jiang, L., Liu, X., Yang, J., Wei, J., Xu, J., et al. 2013. A post-GWAS replication study confirming the PTK2 gene associated with milk production traits in Chinese Holstein. PLoS ONE, **8**(12): e83625. doi:10.1371/journal.pone.0083625. PMID:24386238.

Watkin, H., Richert, M.M., Lewis, A., Terrell, K., McManaman, J.P., and Anderson, S.M. 2008. Lactation failure in Src knockout mice is due to impaired secretory activation. BMC Dev. Biol. **8**: 6. doi:10.1186/1471-213X-8-6. PMID:18215306.

Watson, C.J. 2001. Stat transcription factors in mammary gland development and tumorigenesis. J. Mammary Gland Biol. Neoplasia, **6**(1): 115–127. doi:10.1023/A:1009524817155. PMID:11467447.

Whitney, R.McL. 1988. Proteins of milk. *In* Fundamentals of dairy chemistry. *Edited by* N.P. Wong, R. Jenness, M. Keeney, and E.H. Marth. Van Nostrand Reinhold, New York.

Wiggans, G.R., Cole, J.B., Hubbard, S.M., and Sonstegard, T.S. 2017. Genomic selection in dairy cattle: the USDA experience. Annu. Rev. Anim. Biosci. **5**: 309–327. doi:10.1146/annurev-animal-021815-111422. PMID:27860491.

Worby, C.A., and Dixon, J.E. 2002. Sorting out the cellular functions of sorting nexins. Nat. Rev. Mol. Cell Biol. **3**(12): 919–931. PMID:12461558.

Wu, H.J., Luo, J., Wu, N., Matand, K., Zhang, L.J., Han, X.F., and Yang, B.J. 2008. Cloning, sequence and functional analysis of goat ATP-binding cassette transporter G2 (ABCG2). Mol. Biotechnol. **39**(1): 21–27. doi:10.1007/s12033-007-9024-5. PMID:18256940.

Xia, J., Gill, E., and Hancock, R.E.W. 2015. NetworkAnalyst for statistical, visual and network-based approaches for meta-analysis of expression data. Nat. Protoc. **10**(6): 823–844. doi:10.1038/nprot.2015.052. PMID:25950236.

Xu, H.F., Luo, J., Zhao, W.S., Yang, Y.C., Tian, H.B., Shi, H.B., and Bionaz, M. 2016. Overexpression of SREBP1 (sterol regulatory element binding protein 1) promotes de novo fatty acid synthesis and triacylglycerol accumulation in goat mammary epithelial cells. J. Dairy Sci. **99**(1): 783–795. doi:10.3168/jds.2015-9736. PMID:26601584.

Xu, H., Luo, J., Ma, G., Zhang, X., Yao, D., Li, M., and Loor, J.J. 2018. Acyl-CoA synthetase short-chain family member 2 (ACSS2) is regulated by SREBP-1 and plays a role in fatty acid synthesis in caprine mammary epithelial cells. J. Cell Physiol. **233**(2): 1005–1016. doi:10.1002/jcp.25954. PMID:28407230.

Xu, W., Huang, H., Yu, L., and Cao, L. 2015. Meta-analysis of gene expression profiles indicates genes in spliceosome pathway are up-regulated in hepatocellular carcinoma (HCC). Med. Oncol. **32**(4): 96. doi:10.1007/s12032-014-0425-6. PMID:25731616.

Yamashita, H., Nevalainen, M.T., Xu, J., LeBaron, M.J., Wagner, K.U., Erwin, R.A., et al. 2001. Role of serine phosphorylation of Stat5a in prolactin-stimulated β-casein gene expression. Mol. Cell. Endocrinol. **183**: 151–163. doi:10.1016/S0303-7207(01)00546-9. PMID:11604235.

Yang, J.X., Wang, C.H., Xu, Q.B., Zhao, F.Q., Liu, J.X., and Liu, H.Y. 2015. Methionyl-methionine promotes α-s1 casein synthesis in bovine mammary gland explants by enhancing intracellular substrate availability and activating JAK2-STAT5 and mTOR-mediated signaling pathways. J. Nutr. **145**(8): 1748–1753. doi:10.3945/jn.114.208330. PMID:26108540.

Yang, J., Jiang, J., Liu, X., Wang, H., Guo, G., Zhang, Q., and Jiang, L. 2016. Differential expression of genes in milk of dairy cattle during lactation. Anim. Genet. **47**(2): 174–180. doi:10.1111/age.12394. PMID:26692495.

Yonekura, S., Miyazaki, H., and Tokutake, Y. 2015. Comparative expression profiling of lactogenic hormone receptor and it's signaling molecules of bovine mammary glands during lactation. Open J. Anim. Sci. **5**: 106–113. doi:10.4236/ojas.2015.52013.

Zhang, M.C., Zhao, S.G., Wang, S.S., Luo, C.C., Gao, H.N., Zheng, N., and Wang, J.Q. 2018. d-Glucose and amino acid deficiency inhibits casein synthesis through JAK2/STAT5 and AMPK/mTOR signaling pathways in mammary epithelial cells of dairy cows. J. Dairy Sci. **101**(2): 1737–1746. doi:10.3168/jds.2017-12926. PMID:29248227.

Zhang, W., Thompson, B.J., Hietakangas, V., and Cohen, S.M. 2011. MAPK/ERK signaling regulates insulin sensitivity to control glucose metabolism in Drosophila. PLoS Genet. **7**(12): e1002429. doi:10.1371/journal.pgen.1002429. PMID:22242005.

Zhao, F.Q., and Keating, A.F. 2007. Expression and regulation of glucose transporters in the bovine mammary gland. J. Dairy Sci. **90**(1): E76–E86. doi:10.3168/jds.2006-470. PMID:17517754.

Zhao, K., Liu, H.Y., Wang, H.F., Zhou, M.M., and Liu, J.X. 2012. Effect of glucose availability on glucose transport in bovine mammary epithelial cells. Animal, **6**(3): 488–493. doi:10.1017/S1751731111001893. PMID:22436228.