



Genome-wide association studies for citric and lactic acids in dairy sheep milk in a New Zealand flock

An Zongqi, Ana C. Marshall, J. M. D. R. Jayawardana, Mike Weeks, Simon M. Loveday, Warren McNabb & Nicolas Lopez-Villalobos

To cite this article: An Zongqi, Ana C. Marshall, J. M. D. R. Jayawardana, Mike Weeks, Simon M. Loveday, Warren McNabb & Nicolas Lopez-Villalobos (2024) Genome-wide association studies for citric and lactic acids in dairy sheep milk in a New Zealand flock, *Animal Biotechnology*, 35:1, 2379897, DOI: [10.1080/10495398.2024.2379897](https://doi.org/10.1080/10495398.2024.2379897)

To link to this article: <https://doi.org/10.1080/10495398.2024.2379897>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 05 Aug 2024.



[Submit your article to this journal](#)



Article views: 360



[View related articles](#)



[View Crossmark data](#)

Genome-wide association studies for citric and lactic acids in dairy sheep milk in a New Zealand flock

An Zongqi^a, Ana C. Marshall^{b,c}, J. M. D. R. Jayawardana^{b,d}, Mike Weeks^e, Simon M. Loveday^{c*}, Warren McNabb^c and Nicolas Lopez-Villalobos^b

^aSichuan Agricultural University, College of Science and Technology, Farm Animal Genetic Resources Exploration and Innovation Key Laboratory of Sichuan Province, Chengdu, P. R. China; ^bSchool of Agriculture and Environment, Massey University, Palmerston North, New Zealand; ^cThe Riddet Institute, Massey University, Palmerston North, New Zealand; ^dDepartment of Animal Science, Faculty of Animal Science and Export Agriculture, Uva Wellassa University, Badulla, Sri Lanka; ^eSmart Foods & Bioproducts Group, AgResearch Ltd, Massey University, Palmerston North, New Zealand

ABSTRACT

The objectives of this study were to estimate genetic parameters for citric acid content (CA) and lactic acid content (LA) in sheep milk and to identify the associated candidate genes in a New Zealand dairy sheep flock. Records from 165 ewes were used. Heritability estimates based on pedigree records for CA and LA were 0.65 and 0.33, respectively. The genetic and phenotypic correlations between CA and LA were strong-moderate and negative. Estimates of genomic heritability for CA and LA were also high (0.85, 0.51) and the genomic correlation between CA and LA was strongly negative (-0.96 ± 0.11). No significant associations were found at the Bonferroni level. However, one intragenic SNP in C1QTNF1 (chromosome 11) was associated with CA, at the chromosomal significance threshold. Another SNP associated with CA was intergenic (chromosome 15). For LA, the most notable SNP was intragenic in CYTH1 (chromosome 11), the other two SNPs were intragenic in MGAT5B and TIMP2 (chromosome 11), and four SNPs were intergenic (chromosomes 1 and 24). The functions of candidate genes indicate that CA and LA could potentially be used as biomarkers for energy balance and clinical mastitis. Further research is recommended to validate the present results.

KEYWORDS

Sheep milk; citric acid; lactic acid; genome-wide-association study


Introduction

Sheep dairy farming is growing in New Zealand, and some dairy cattle farms in the North Island have recently shifted to dairy sheep farms.¹ New Zealand accounted for 0.1% of the total sheep milk production worldwide until 2020.² Sheep milk has a high content of protein (as well as casein), fat, lactose, vitamins, and minerals,³ which makes it ideal for processing, particularly into cheese.⁴ Sheep milk is also a richer source of unsaturated fatty acids (UFA) when compared to cow and goat milk.⁵ Moreover, the bioactive substances present in sheep milk, such as organic acids, prebiotics, and probiotics, can be used for the production of cosmetics and for promoting human health.⁶

Citric acid or citrate (CA) and lactic acid (LA) are considered anti-ageing factors⁷ and are some of the least studied components in sheep milk. Sheep milk

has been reported to have a higher content of CA and LA than goat milk,^{8,9} but lower CA content than cow milk.¹⁰ Citric acid is one of the main metabolites synthesized in the mammary gland, whose carbon is derived from glucose and acetate.^{11,12} Formed within the secretory cell, CA enters milk by exocytosis from the Golgi vesicles, in a similar way to lactose and protein.¹² Approximately 7% of total CA in cow milk is in the non-aqueous phase, combined with calcium-casein complexes,¹³ but the majority of CA in milk is soluble.¹⁴ Furthermore, CA is an intermediate in the tricarboxylic acid cycle, also known as the Krebs cycle, which is one step of cellular metabolism to generate energy.¹⁵ Citric acid is also indirectly involved in fat synthesis by providing the reducing equivalents in the form of nicotinamide adenine dinucleotide phosphate (NADPH).¹⁵

CONTACT Nicolas Lopez-Villalobos  N.Lopez-Villalobos@massey.ac.nz  School of Agriculture and Environment, Massey University, Palmerston North, New Zealand.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/10495398.2024.2379897>.

© Commonwealth Scientific and Industrial Research Organisation (CSIRO), Agriculture and Food, 671 Sneydes Road, Werribee, VIC 3030, Australia
© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

The CA content has been proposed to be used as an indicator of the energy status of dairy cows.¹⁶ A higher amount of CA in the milk was found in cows with negative energy balance (NEB). When the energy requirements for maintenance and milk production are not satisfied by feed intake, the cow enters a status of NEB, especially in early lactation when the energy requirement for milk production is at its highest peak.¹⁷ It has been suggested that increases in the lipid metabolism of mammary gland cells affect the milk content of CA in cows with negative energy balance. The reduced levels of CA from mid-lactation have been suggested to be due to improved energy balance.¹⁶

This is further supported by the association of CA content with the synthesis of fatty acids. Citrate concentration in milk was increased when *de novo* synthesis of fatty acids in the mammary cells increased. The changes in CA content in milk were related to changes in intra-mammary energy metabolism and the level of fatty acid synthesis.¹⁵

Citric acid is also known to influence the technological properties of milk. Citric acid content in milk has been linked with milk heat stability, which is relevant in the processing of milk into dairy products that require high-temperature treatment (such as UHT, and milk powder). A deficiency of citrate leads to an excess of soluble calcium, which destabilizes the casein micelles.¹⁸ The effect of CA on cheesemaking has also been investigated.^{19,20} Increased citrate levels in skim milk from individual cows have previously been associated with impaired milk coagulation properties for cheesemaking,¹⁹ but the opposite effect has also been documented with an improvement of milk coagulation with a higher CA.²⁰ The concentration of CA varies not only between breeds of dairy cattle,^{20–22} but also between animals of different genotypes. For example, a high level of CA was shown for animals carrying the A1 allele type of β -casein,²³ and a negative correlation between CA and fat percentage in cow milk has been reported.²¹ Estimates of heritability for CA were reported to be high; 0.54 to 0.82.^{20,21}

Lactic acid, on the other hand, is the main organic acid representing titratable acidity in milk²⁴ and is mainly produced by lactic acid bacteria (LAB) breaking down lactose,²⁵ which means that fresh milk contains very low LA (not over 0.002%).²⁶ Furthermore, the estimate of heritability for LA has been previously reported to be close to zero.²¹ Lactic acid produced by bacterial carbohydrate fermentation contributes to acidification and flavor development in cheesemaking. An increase of 0.03% of LA content in milk has been related to a protein content increase of over 1%.²⁷

Contents of CA and LA have been proposed as metabolite biomarkers for the health status of dairy animals and for mammary gland disease due to the correlation with somatic cell count (SCC) and therefore, with clinical mastitis.^{28–31} Additionally, the correlations between milk composition (fat, protein, and lactose) and these metabolites could have a wide range of applications in food processing.²⁸

The genomic background of these traits has yet to be explored in dairy sheep, while investigations have only recently begun in dairy cows, making this study's contribution noteworthy. In addition, performing GWAS in dairy sheep presents unique challenges compared to dairy cows. Sheep typically produce smaller volumes of milk, making milk sampling more difficult. Additionally, most dairy sheep farms lack milk meters due to their high costs relative to the value of a sheep.

Considering the above investigations and possible applications of CA and LA, the objective of this study was to estimate variance components, genetic and genomic parameters of CA and LA content in sheep milk and perform genome-wide analysis to identify SNPs associated with genes potentially affecting the concentration of CA and LA in dairy sheep milk.

Materials and methods

Animals and milk samples

This study was conducted at a commercial sheep dairy farm in Masterton, New Zealand (latitude: -40.9597 , longitude: 175.6575). Animal ethics approval was obtained for this study (Massey University Animal Ethics Committee-Protocol 21/45). Detailed information on the flock is provided in.³² The age structure of the flock was 15% first-, 28% second-, 23% third- and 34% \geq fourth-parity ewes. The median lambing date of the flock was the 20th of August 2021. Collections of test-day records started after the lambs were fully weaned. A total of 521 test-day records from once-a-day milking were gathered from 1 November 2021 until 31 January 2022 from 169 ewes.³² At least 2 milk tests were performed for each ewe, and 92% of the flock had at least 3 milk test records. On each test-day, milk volume was measured, and a representative milk sample was taken from each animal. The composition of individual milk samples was analyzed using a Fourier-transform Infrared (FTIR) milk-analyser (MilkoScan FT6000) calibrated for sheep milk components. The following milk components were extracted: percentages of fat, protein, lactose, total milk solids, and contents of CA and LA.

Information on lambing date, litter size, and age of animals were supplied to the study.

Genotyping and quality control

Animals were genotyped using the OvineSNP50 Beadchip array (Illumina, San Diego, CA) with a medium-density SNP panel (50k SNPs). Ear tissue samples from 323 ewes (including the 169 ewes and 154 ewe lambs) and 6 rams were collected for DNA extraction and scanning using an iScan[®] at the Equine Parentage Testing Lab (Massey University, Palmerston North). A total of 64,734 SNPs were obtained for quality control using the SNP & Variation Suite (SVS 8.8)³³ software. In the filtering process, genomic records were removed for 23 animals with a call rate < 95% across all the SNPs, from which 4 animals were ewes that had phenotype records. Also, SNPs with > 1% missing genotypes across all individuals (call rate < 95%), and that had a significant deviation from the Hardy-Weinberg equilibrium threshold of $p < 10^{-6}$ or that had minor allele frequency < 1% were also removed. After these quality control edits, a total of 306 animals with SNPs (including 165 ewes with phenotypes), and 45,801 SNPs remained for association analyses.

Statistical analysis

Descriptive statistics were obtained using the MEANS procedure of SAS 9.4.³⁴ Analysis of variance for concentration of CA and LA was performed using the MIXED procedure with a linear model that included the fixed effects of ewe coat color (black or white) as an indicator of variety within the breed, litter size (1 or ≥ 2 lambs), parity number (1st, 2nd, 3rd or ≥ 4 th parity), milking month (Nov, Dec or Jan), and days in milk as a covariate, and the random effect of ewe to account for repeated records on the same ewe. Least-squares means for each class of fixed effects and standard errors were obtained and used for mean comparisons using Fisher's least significant difference test. Lactation curves of CA and LA contents for the flock were modeled with a quadratic and a linear polynomial, respectively. The best order of the polynomial was based on the R^2 values.

Estimation of genetic parameters

ASReml release 4.1³⁵ was used to estimate the variance components, heritability, and genetic and phenotypic correlations between CA and LA. Heritability

estimates for CA and LA were obtained through single-trait animal models, and estimates of genetic correlation were obtained through a bivariate animal model. The single-trait animal model was represented as follows:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wc} + \mathbf{e}$$

where \mathbf{y} is the vector of phenotypic traits; \mathbf{b} is the vector of fixed effects of ewe coat, litter size, parity, milking month, and days in milk as covariate; \mathbf{a} is the vector of animal genetic effects; \mathbf{c} is the vector of permanent (ewe) environment effect, \mathbf{e} was the vector of random residual effects; \mathbf{X} , \mathbf{Z} and \mathbf{W} are the design matrices relating to the records of fixed, additive genetic, and permanent environment effects, respectively.

The distributional properties of this model are the following:

$$E \begin{bmatrix} \mathbf{y} \\ \mathbf{a} \\ \mathbf{c} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{Xb} \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{var} \begin{bmatrix} \mathbf{a} \\ \mathbf{c} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{A}\sigma_a^2 & 0 & 0 \\ 0 & \mathbf{I}_1\sigma_c^2 & 0 \\ 0 & 0 & \mathbf{I}_2\sigma_e^2 \end{bmatrix}$$

where \mathbf{A} is the numerator relationship matrix between all the animals in the pedigree, \mathbf{I}_1 is an identity matrix of order equal to the number of ewes with records, \mathbf{I}_2 is an identity matrix of order equal to the number of records, σ_a^2 is the animal genetic additive variance, σ_c^2 is the permanent environment effect variance and σ_e^2 is residual error variance. The pedigree file included 165 animals with records, which were progeny of 21 sires and 105 dams, with 11 maternal grandsires and 25 maternal granddams.

The heritability (h^2) estimate was obtained as follows:³⁶

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_c^2 + \sigma_e^2}$$

The repeatability (t) estimate was obtained as follows:³⁶

$$t = \frac{\sigma_a^2 + \sigma_c^2}{\sigma_a^2 + \sigma_c^2 + \sigma_e^2}$$

The bivariate model was represented in matrix notation, as follows:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

where y_1 and y_2 are the vectors of phenotypic measures of CA and LA; b_1 and b_2 are the vectors of fixed effect of coat color, litter size, age, milking month, linear and quadratic effects of days in milk; X_1 , X_2 , Z_1 , Z_2 , W_1 and W_2 are design matrices relating the fixed, animal additive genetic and permanent environment effects related to the y_1 and y_2 phenotypes, respectively, a_1 and a_2 are the vectors of random effects of animal for each trait; c_1 and c_2 are the vectors of random permanent environment effects, and e_1 and e_2 are vectors of residual errors for each trait. The expected values of the variables were assumed $E(y_1) = X_1 b_1$; $E(y_2) = X_2 b_2$ with variances and covariances of random effects as follows:

$$\text{var} \begin{bmatrix} a_1 \\ a_2 \\ c_1 \\ c_2 \\ e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} A\sigma_{a1}^2 & A\sigma_{a12} & 0 & 0 & 0 & 0 \\ A\sigma_{a12} & A\sigma_{a2}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_1\sigma_{c1}^2 & I_1\sigma_{c12} & 0 & 0 \\ 0 & 0 & I_1\sigma_{c12} & I_1\sigma_{c2}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_2\sigma_{e1}^2 & I_2\sigma_{e12} \\ 0 & 0 & 0 & 0 & I_2\sigma_{e12} & I_2\sigma_{e2}^2 \end{bmatrix}$$

where σ_{a1}^2 is the additive genetic variance of the CA content, σ_{a2}^2 is the additive genetic variance of LA content and σ_{a12} is the additive genetic covariance between CA and LA contents; σ_{c1}^2 is the permanent environment variance of the CA content, σ_{c2}^2 is the permanent environment variance of LA content and σ_{c12} is the permanent environment covariance between CA and LA contents; σ_{e1}^2 is the residual variance for CA content; σ_{e2}^2 is the residual variance of LA content and σ_{e12} is the residual covariance between CA and LA contents.

The genetic correlation (r_g) between CA and LA was estimated as follows:³⁶

$$r_g = \frac{\sigma_{a12}}{\sigma_{a1} \times \sigma_{a2}}$$

where σ_{a1} and σ_{a2} are genetic additive standard deviations for CA and LA contents, respectively.

The phenotypic correlations (r_p) were estimated as follows:³⁶

$$r_p = \frac{\sigma_{p12}}{\sigma_{p1} \times \sigma_{p2}}$$

where σ_{p1} and σ_{p2} are the phenotypic standard deviation for CA and LA contents, respectively, and σ_{p12} was the phenotypic covariance between CA and LA contents, with $\sigma_{p1}^2 = \sigma_{a1}^2 + \sigma_{c1}^2 + \sigma_{e1}^2$ and $\sigma_{p2}^2 = \sigma_{a2}^2 + \sigma_{c2}^2 + \sigma_{e2}^2$

Genomic-wide association analysis

The phenotypic traits of CA and LA were pre-corrected for fixed effects using the single-trait animal repeatability model in the ASReml 4.1 software package.³⁵ The models included the fixed effects of coat color, litter size, age, milking month, linear and quadratic effects of days in milk as covariates, the random effect of ewe permanent environment, and random residual. Principal component analyses for population stratification correction were performed. PC1 was highly correlated ($r=0.95$) with the coat-color effect and therefore PCs were not included, instead, the genomic relationship matrix was used for adjustment of population structure. The genome-wide association study (GWAS) was performed using a mixed linear model in the software package.³⁷ The following model was fitted for each trait:

$$y = \mu + X\beta + g + e$$

where y was the vector of adjusted phenotype for each ewe, μ was the vector of overall mean, β was the vector of fixed effects and X was the incidence matrix of β to y with the SNPs genotypes of BB, AB, or AA, respectively, g was the random additive polygenic effect and e was the random residual error. The assumptions for the model were: $g \sim N(0, G\sigma_g^2)$, where G was the genomic relationship matrix between the ewe and σ_g^2 was the additive genetic variance explained by SNPs, and $e \sim N(0, I\sigma_e^2)$, where I was the identity matrix of order $n = 165$ and σ_e^2 was the residual variance.

Bonferroni multiple-test correction and the suggestive significance threshold were used for multiple tests, which is considered conservative, to avoid the type I error.³⁸ Bonferroni multiple-test correction was estimated at 5% genome-wide significance as $0.05/m$ ($0.05/45,801 = 1.09 \times 10^{-6}$), which corresponds to 5.96 on a $-\log_{10}$ (p-value) scale. The suggestive significance threshold was estimated at $1/m$ ($1/45,801 = 2.18 \times 10^{-5}$), which was 4.66 on a $-\log_{10}$

(p-value) scale. The chromosomal significance threshold was calculated as $0.05/\text{number of SNPs}$ in each chromosome on a $-\log_{10}$ (p-value) scale. Manhattan plots in which $-\log_{10}$ (p-values) were plotted against their genomic locations of the markers for each trait using the qqman package in R software 4.2.1.³⁹

Candidate genes and functional analysis

Ensembl Release 109, based on the *Ovis aries* (sheep) reference genome in Oar_rambouillet_v1.0 genome assembly (<http://www.ensembl.org/index.html>, accessed on Feb 2023), was used to identify potential candidate genes in association analysis.⁴⁰ Gene annotation boundaries were set at 150 kbp on either side of significantly associated SNP. This range was chosen due to the density of the chip array. The biological functions of the associated candidate genes were reviewed using the Gene Ontology (GO) tool in Ensembl.

For LA candidate genes, enrichment of gene ontology (GO) was investigated using the STRING Genomics 10.0 database.⁴¹ This analysis of protein-protein interaction was based on text mining, experiments, database, co-expression, neighborhood, gene fusion, and co-occurrence, and the minimum required interaction score was set to 0.40 and with a false discovery rate (FDR) <0.05 .

Results

Descriptive statistics

Descriptive statistics of milk components and CA and LA contents for dairy sheep in the 2021–2022 season ($n=165$ individuals) are presented in Table 1 and data distribution plots for CA and LA contents are presented in Figures S1 and S2.

Ewes produced milk with high percentage of fat (6.15%) followed by protein (5.40%) and lactose (4.76%) percentages during the production season. The content of LA was higher than the content of CA, but LA had a smaller coefficient of variation. The lactation curves for LA and CA are shown in Figure 1. Throughout the lactation, LA content slowly

increased, whereas CA content decreased during the first 70 days, and then stabilized, rising slowly until the end of the study period.

Factors affecting the contents of citric and lactic acids

Least-squares means and standard errors of LA and CA for fixed effects of coat color, age, litter size, and milking month are presented in Table 2.

The milking month of ewes was significant ($p<0.01$) for LA concentration and there was no significant effect on these fixed effects for CA composition. However, the coat color and ewe age had nearly significant effects on the CA and LA content, respectively.

Genetic parameters for contents of citric and lactic acids

Table 3 shows the variances and heritabilities for CA and LA, and genetic and phenotypic correlations between CA and LA, using pedigree-based variance and genomic variance components.

Higher heritability values were found for CA than for LA using both genomic- and pedigree-based approaches. A strong negative genomic correlation (-0.96 ± 0.11) between CA and LA was found in this study, whereas the pedigree-based genetic correlation between CA and LA was strong-moderate (-0.67 ± 0.33).

Genome-wide association studies for contents of citric and lactic acids

The Manhattan plots for CA and LA in dairy sheep milk are presented in Figure 2. The QQ-plots of CA and LA before and after data correction for stratification are presented in a supplementary file in Figures S3–S6.

No genome-wide significant associations were found for LA or CA contents in this population using the significance threshold of Bonferroni. However, at the suggestive threshold, a total of five SNPs were significantly associated with LA content.

Table 1. Descriptive statistics¹ of milk components including citric and lactic acids in dairy sheep sampled during the production season 2021–2022 in a commercial flock in New Zealand ($N=622$ test-day milk samples from 165 ewes).

Trait	Mean	SD	Min.	Max.	CV (%)
Fat, %	6.15	0.69	4.23	8.55	11
Protein, %	5.40	0.44	4.31	6.98	8
Lactose, %	4.76	0.16	4.00	5.13	3
Total milk solids, %	16.9	0.96	14.2	20.0	6
Citric acid, mg/100 g	83.7	29.0	20.0	170.0	35
Lactic acid, mg/100 g	186.9	20.2	140.0	250.0	11

¹SD=standard deviation, Min. = minimum value, Max. = maximum value and CV=coefficient of variation.

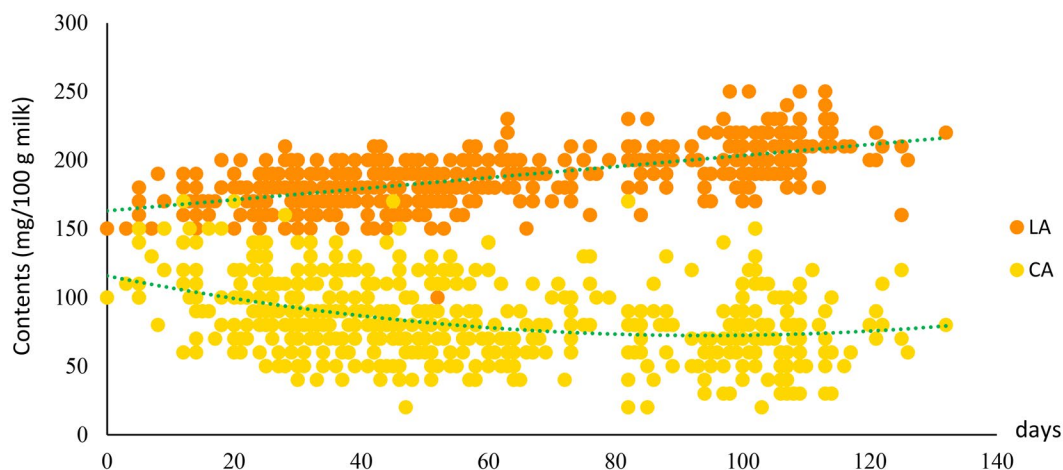


Figure 1. Variation of citric acid (CA; mg/100 g milk) and lactic acid (LA; mg/100 g milk) contents throughout the study period. Linear and quadratic models were used for LA and CA, respectively; days=days in milk, was presented by days in lactation minus 50.

Table 2. Least-squares means, standard errors (SE), and p-value for fixed effects on contents of citric acid (CA; mg/100 g milk) and lactic acid (LA; mg/100 g milk) in milk from ewes ($n=165$) of different coat colors, parities, litter sizes and milking months.

Effect	Citric acid			Lactic acid		
	Mean	SE	p-value	Mean	SE	p-value
Coat color			0.05			0.79
Black	76.1	4.4		184.1	2.4	
White	85.8	2.4		184.8	1.2	
Litter size			0.41			0.10
1	82.6	3.3		182.6	1.8	
2	79.3	3.2		186.3	1.7	
Parity			0.16			0.05
1	85.4	5.8		180.7	3.2	
2	76.9	4.1		188.9	2.2	
3	76.2	4.1		182.5	2.3	
≥ 4	85.2	3.4		185.8	1.9	
Lactation month			0.16			<0.001
November	84.3	3.8		181.0 ^b	2.2	
December	78.8	3.1		178.6 ^b	1.8	
January	79.6	3.5		193.8 ^a	2.0	

^{a,b}Means with different superscripts within effect are significantly different ($p < 0.05$).

Candidate genes and functional analysis

The significant SNPs with respective candidate genes and gene annotation for CA and LA contents are presented in Tables 4 and 5, respectively. The gene ontology functional enrichment results for LA content are presented in Table 6.

Two SNPs significant at the chromosomal significance threshold, and seven associated candidate genes, located in the suggested region, were identified for CA content, in chromosomes 11 and 15. The most relevant SNP for CA was within gene C1QTNF1 in chromosome 11, in a protein-coding region. The second most relevant SNP was intergenic and close to the novel gene ENSOARG00020014135.

Seven SNPs and 53 associated candidate genes located in the suggested region were identified for LA content, in chromosomes 1, 11, and 24. Three SNPs out of seven were mapped to introns. Five out of seven SNPs were significantly associated at the suggestive significance threshold ($-\log_{10}(p\text{-value}) \geq 4.66$). Six SNPs out of seven were significant at the chromosomal significance threshold. Four out of seven SNPs were significant at both suggestive and chromosomal significance thresholds. The most significant SNP for LA content was in chromosome 11, an intron variant of gene cytohesin 1 (CYTH1). The other two SNPs in chromosome 11 were intron variants of the genes MGAT5B and TIMP2. Another strong SNP for LA content, significant at the chromosomal threshold, was intergenic in chromosome 24, close to gene MAZ. The gene C1QTNF1 was a common gene for CA and LA contents in the suggested region.

Discussion

Content and genetic parameters of citric and lactic acids

The average citric acid content in this flock was 83.7 mg/100 g milk. A study from Austria reported a higher CA content (154 ± 25 g/100 g milk).⁹ Contents of CA in dairy goat milk were reported at 102 ± 11 mg/100 g milk.⁹ The content of CA in cow milk (11.4 ± 0.4 mmoles/kg) was greater than in sheep milk (10.7 ± 1.4 mmoles/kg) and goat milk (4.2 ± 0.7 mmoles/kg).¹⁰ In Danish Jersey cows' milk, the CA content was higher than in Danish Holstein's milk,²⁰ whereas in China, the LA content was higher in Holstein's milk than in Jersey's milk.²² In sheep milk,

Table 3. Estimates of variance components and heritabilities with their respective standard errors (in brackets) were obtained from pedigree-based analysis considering the additive genetic relationship matrix and genomic-based analysis considering the genomic relationship matrix for citric acid (CA) and lactic acid (LA).

Pedigree-based variance component estimates ¹							Correlations
Trait	σ_a^2	σ_c^2	σ_e^2	σ_p^2	h^2	t	
CA	511 (71)	0.0 (0.0)	276 (22)	786 (72)	0.65 (0.04)	0.65 (0.04)	$r_p = -0.40$ (0.05)
LA	83 (57)	53 (48)	115 (9)	252 (23)	0.33 (0.21)	0.54 (0.05)	
Genomic-based variance component estimates ¹							Correlations
Trait	σ_a^2	-	σ_e^2	σ_p^2	h^2		
CA	360 (90)	-	0.06 (0.05)	0.42 (0.06)	0.85 (0.13)		$r_g = -0.96$ (0.11)
LA	50 (20)	-	0.05 (0.02)	0.11 (0.01)	0.51 (0.17)		

¹All variance components correspond to contents of CA and LA (mg/100 g of milk): σ_a^2 =additive genetic variance; σ_e^2 = residual variance; σ_c^2 = permanent environment variance; σ_p^2 = total phenotypic variance; h^2 = heritability; t= repeatability; r_g = genetic correlation between CA and LA; r_p = phenotypic correlation between CA and LA.

a significant difference in LA content was reported for Lacaune sheep raised in different regions of south Brazil.²⁴ LA concentration in Lacaune sheep milk (240 mg/100 g milk) was higher than the LA content estimated in our study. A New Zealand study⁸ reported that sheep milk had a higher content of LA (220 to 250 mg/100 g milk) than cow milk (150 to 180 mg/100 g milk) and goat milk (140 to 230 mg/100 g milk).

Reports on the estimates of genetic parameters for CA and LA in sheep milk are scarce in the literature. However, heritability estimates for CA content in Danish Holstein (0.59 ± 0.20) and in Danish Jersey (0.82 ± 0.31)²⁰ were greater than our estimates. Heritability for LA reported in the literature was reported to be close to zero,²¹ which contradicts the present reported heritability values for LA. The estimates of genetic and genomic correlations between CA and LA contents estimated in our study were strongly negative. This indicates that ewes with genetic merit for producing milk with higher CA content are of low genetic merit for LA content in milk. It is probably related to the co-metabolism of CA and glucose, which starts with the degradation of CA to pyruvate, generating LA as the end-product of glucose metabolism.⁴² Also, lactic acid bacteria produce LA from lactose and diacetyl from CA.^{42,43}

In the present study, an opposite trend was observed between CA content and LA content throughout the lactation of dairy sheep. Citric acid content decreased in earlier stages and increased slowly until the end of lactation, agreeing with the previous study for cow and sow.¹⁰ Also, energy balance has been negatively correlated with milk fat yield ($r = -0.78$) in dairy cows, and CA yield has been positively correlated with fat yield throughout the lactation ($r = 0.53$), indicating

that citrate is used for milk fat synthesis in cows in negative energy balance and therefore, could be used as an indicator of energy status.^{15,16,44} The underlying context of this is that the energy deficit leads to increased lipolysis, and the uptake of fatty acids mobilized from body fat tissue is increased resulting in a slowly increased fat synthesis in the mammary cells.^{16,45,46} Samples in our study were collected from around 50 days of lactation (the first 50 days were the lamb suckling period), thus the elevated content that is generally reported in the first 2-6 weeks could not be evaluated in our study.

Citric acid and LA have also been proposed as metabolic biomarkers for mastitis, with increased LA content and decreased CA content in clinical mastitis.²⁸⁻³¹ Records on clinical mastitis were not available in this study and it was not possible to explore this relationship.

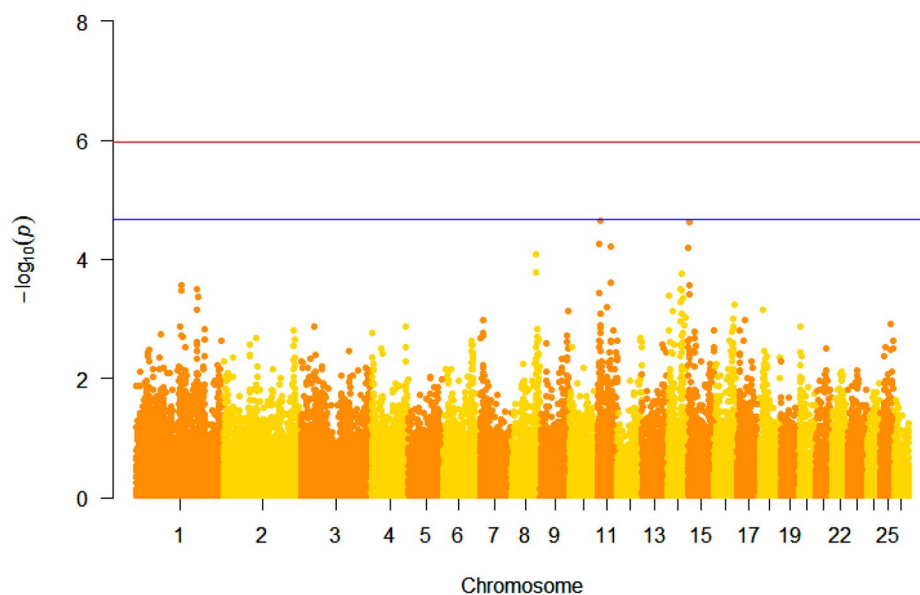
Genome-wide association for the content of citric acid

The QQ-plots showed a proper fit for stratification and a deviation to the right which suggests a true association between SNPs and concentration of CA and LA.

In agreement with previous findings of for cow milk,⁴⁷ a genomic locus in chromosome 11 was associated with CA content. However, studies in cow milk have also revealed strong candidate genes associated with CA content in bovine chromosome 20.^{47,48}

The specific biological functions of the gene C1q and TNF Related 1 (C1QTNF1) are mainly related to the regulation of glucose metabolic processes (GO:0010906) and positive regulation of the MAPK

Citric Acid



Lactic Acid

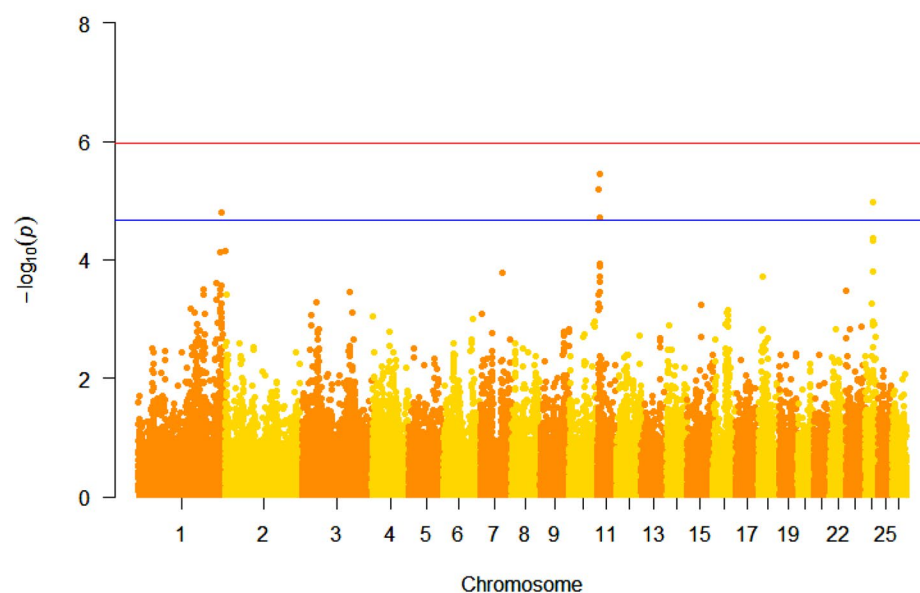


Figure 2. Manhattan plots for $-\log_{10}(p)$ -values of marker effects for citric acid and lactic acid in dairy sheep milk. The genome-wide significance threshold of Bonferroni correction is represented by the red line at $-\log_{10}(p)$ -value = 5.96, and the suggestive significance threshold is represented by the blue line at $-\log_{10}(p)$ -value = 4.66.

Table 4. The single nucleotide polymorphisms (SNPs) were identified as significant at the chromosomal significance threshold (0.05/number of SNPs in chr) for citric acid (CA) content in dairy sheep milk.

SNP	Chr	Chr $-\log_{10}(p)$	Position	Ref./MA	Ref. Freq.	Effect (SE)	$-\log_{10}(p)$	Annotation	Candidate genes ¹
OAR11_9907480.1	11	4.34	9907480	A/G	838.93	-14.42 (3.40)	4.65	Protein coding	C1QTNF1 , TIMP2, LGALS3BP, CANT1, ENGASE, RBFOX3
OAR15_3892898.1	15	4.47	3892898	A/G	373.33	-10.38 (2.46)	4.62	Intergenic	ENSOARG00020014135

Chr=chromosome; Chr $-\log_{10}(p)$ = chromosomal significance threshold; Ref. = reference allele; MA=minor allele; Ref. freq. = reference allele frequency; SE=standard error. Reference allele frequency, effect and standard error presented were multiplied by 1000.¹Candidate genes include those within 150 kbp distance either side of the significant SNP. Gene names marked in bold are those that include the intragenic SNP. *Closest gene within 5kbp distance either side of the significant SNP.

Table 5. The single nucleotide polymorphisms (SNPs) were identified as significant at the suggestive significance threshold ($-\log_{10}(p\text{-value}) \geq 4.66$) and/or at the chromosomal significance threshold (0.05/number of SNPs in chr) for lactic acid (LA) content in dairy sheep milk.

SNP	Chr	Chr – log ₁₀ (p)	Position	Ref./ MA	Ref. Freq.	Effect (SE)	–log ₁₀ (p)	Annotation	Candidate genes ¹
s64611.1	1	5.01	284089015	A/G	260.42	6.26 (1.45)	4.81	Intergenic	ENSOARG00020002543
s25858.1	11	4.34	9641214	A/G	690.00	6.01 (1.30)	5.44	Intron	CYTH1 , DNAH17, TIMP2, ENSOARG00020011905
s56469.1	11	4.34	8356745	A/C	310.00	–5.84 (1.29)	5.19	Intron	MGAT5B , MFSD11, METTL23, SRSF2, JMJD6 ENSOARG00020016717
OAR11_9760146.1	11	4.34	9760146	A/G	630.87	5.50 (1.29)	4.71	Intron	TIMP2 , ENSOARG00020011905, CYTH1, CANT1, LGALS3BP, C1QTNF1
s28518.1	24	4.11	26968037	A/G	676.67	–6.97 (1.58)	4.98	Intergenic	MAZ*, CDIPT, C16ORF54, KIF22, ZG16, QPRT, ENSOARG00020020033, ENSOARG00020019988, TAOX2, KCTD13, SEZ6L2, MVP, PAGR1, PRRT2, TMEM219, ENSOARG00020016749, ENSOARG00020019032, ENSOARG00020019335
OAR24_26621114.1	24	4.11	26621114	G/A	653.33	–6.12 (1.50)	4.36	Intergenic	NUPR1*, TUFM, ATP2A1, SH2B1, ATXN2L, CLN3, IL27, EIF3C, ENSOARG00020026028, ENSOARG00020026072, SGF29, MAPK3, GDPD3, YPEL3, TBX6, ALDOA, CORO1A, PPP4C, ENSOARG00020014726, ENSOARG00020014644, ENSOARG00020014591
OAR24_25044819.1	24	4.11	25044819	G/A	613.33	–5.50 (1.35)	4.33	Intergenic	–

Chr = chromosome; Chr – log₁₀ (p) = chromosomal significance threshold; Ref. = reference allele; MA = minor allele; Ref. freq. = reference allele frequency; SE = standard error. Reference allele frequency, effect and standard error presented were multiplied by 1000. –log₁₀ (p) = SNP significance, marked in bold if significant at the chromosomal significance threshold. ¹Candidate genes include those within 150 kbp distance either side of the significant SNP. Gene names marked in bold are those that include the intragenic SNP. *Closest gene within 5kbp distance either side of the significant SNP.

Table 6. Gene Ontology (GO) functional enrichment for candidate genes for lactic acid content, with false discovery rate (FDR) <0.05.

Trait	Gene Ontology	Biological process description	Strength	FDR
LA	GO:0018279	Protein N-linked glycosylation via asparagine	2.17	0.0012
	GO:006487	Protein N-linked glycosylation	2.03	1.73e-11
	GO:006486	Protein glycosylation	1.72	3.50e-16
	GO:009101	Glycoprotein biosynthetic process	1.67	9.48e-17

cascade (GO:0043410). The gene C1QTNF1 participates in lipid metabolism as a novel adipokine.⁴⁹ The increased uptake of glucose leads to accelerated production of citrate, and citrate is then converted by ACLY to acetyl-CoA, which is an essential biosynthetic precursor for lipid synthesis.⁵⁰ In sheep, C1QTNF1 was significantly downregulated in the adipocytokine signaling pathway in the tail fat tissue of Altay sheep,⁵¹ and differentially expressed in the longissimus muscle tissue of Polled Dorset sheep.⁵² Owing to C1QTNF1 being positioned in the

protein-coding region, further studies are needed to validate the association between citrate metabolism and the molecular function of C1QTNF. There are no specific biological functions annotated for the novel gene ENSOARG00020014135.

Genome-wide association for the content of lactic acid

The enrichment analyses performed for the LA candidate genes uncovered four biological processes, with protein N-linked glycosylation via asparagine (GO:0018279) standing out as the most notable function. The specific functions of the most probable candidate gene for LA content, Cytohesin 1 (CYTH1), are related to the regulation of cell adhesion (GO:0030155), and regulation of ARF protein signal transduction (GO:0032012). CYTH1 could be involved in the regulation of casein synthesis (particularly through Glu) in bovine mammary epithelial cells.^{53,54}

The second most probable candidate gene MGAT5B codes for protein N-linked glycosylation (GO:0006487) and protein O-linked glycosylation via serine (GO:0018242) processes. Interestingly, a study⁵⁵

reported MGAT5B to be positively correlated with acetate, which is the first product of citrate metabolism and the intermediate of co-metabolism of glucose and citrate in milk.^{41,55} The co-metabolism of both glucose and citrate starts with the degradation of citrate into acetate and oxaloacetate, following that oxaloacetate is catalyzed to pyruvate which is the intermediate for co-metabolism and an essential product of glucose metabolism.⁵⁶ It has been identified that co-metabolism results in a decreased uptake of glucose and increased acetate and D-lactate production, in other metabolic functions; citrate had a slightly inhibitory effect on the enzymes of the 'ethanol' leg of glucose metabolism.⁵⁶ The gene MGAT5B was found to be differentially expressed in Glycan biosynthesis and lipid metabolism processes in cervical cancer cells, regulating the decrease of lipid peroxidation, and the increase in glucose consumption.⁵⁷ In dairy cattle milk, MGAT5B was associated with clinical mastitis and SCC.⁵⁸

In dairy cows, TIMP2 is involved in the regulation of intra-mammary infections and fibrosis process.⁵⁹ The gene MAZ participates in the regulation of calcium-dependent activation of synaptic vesicle fusion (GO:0150037), and synaptic vesicle fusion to presynaptic active zone membrane (GO:0031629). Also, it has been reported that the MAZ gene is a transcription factor that is highly upregulated in chronic inflammatory disease.⁶⁰

The gene NUPR1 is associated with acute inflammatory response (GO:0002526), and negative regulation of glycolytic process (GO:0045820). Also, NUPR1 is related to the degradation of insulin storages and subsequent secretion during inflammatory and obesity-related tissue stress.⁶¹

The association between genes MAZ and NUPR1 and LA content in milk has not been identified in previous GWAS studies, but these genes are associated with mastitis in cows⁶² and buffalo.⁶³ It has been reported that LA content can also be used as a signal molecule of the immune system and has novel signal transduction functions in preventing excessive inflammatory responses, both intracellular and extracellular.^{64,65} It has been suggested that LA content in milk is a strong indicator of mastitis in early lactating cows and goats, with a positive correlation between LA and SCC.^{28–30} The identified genes in this study can be useful for future GWAS related to sheep mastitis traits.

Conclusions

High heritability was estimated for both LA and CA traits. This study was the first to report genomic

variance for CA and LA, and a strong negative genetic correlation was found between CA and LA. The genomic background of these traits has not yet been investigated in dairy sheep and is only recently being investigated for dairy cows, which highlights the unique contribution of this study. Despite the small sample size, the genome-wide association analysis in our study detected a total of 9 SNPs and 56 candidate genes associated with CA and LA content. Of notable importance for both CA and LA contents in sheep milk were the genetic variations in chromosome 11. The functions of genes are associated with energy balance, metabolic processes, protein glycosylation, immunity response, and clinical mastitis, which may provide more information for research on LA and CA metabolism in sheep and on using LA and CA as metabolic biomarkers for energy balance and clinical mastitis. The findings of the present study lack robustness for large-scale applications due to the small sample size. Therefore, further genomic studies must be performed with larger populations of dairy sheep. Studies on gene expression could also validate the present results.

Acknowledgements

The authors would like to thank the commercial farm and the scientific communities from the NZ3M group, Massey University, and AgResearch.

Complying with ethics of experimentation

The experiment and all animal handling procedures were approved by the Massey University Animal Ethics Committee (Massey University Animal Ethics Committee-Protocol 21/45).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the New Zealand Ministry of Business, Innovation and Employment (MBIE), through the New Zealand Milks Mean More (NZ3M) Endeavor Programme (Contract MAUX1803). This work was also supported by the Riddet Institute National Centre of Research Excellence through a PhD scholarship to A.C.M.

Data availability statement

The data presented in this experiment are available within the article.

References

- Peterson S, Prichard C. The sheep dairy industry in New Zealand: A review. *N. Z. Soc. Anim. Prod.* 2015;75:119–126.
- Ministry of Primary Industries – Massey University. New Zealand Sheep Dairy Survey; MPI-Massey University: Wellington and Palmerston North; 2020. Accessed March 10, 2023. <https://www.mpi.govt.nz/dm/document/43183-Sheep-Dairy-Survey-Report-July-2020>.
- Park Y, Juárez M, Ramos M, Haenlein G. Physico-chemical characteristics of goat and sheep milk. *Small Rumin. Res.* 2007;68(1–2):88–113.
- Balthazar C, Pimentel T, Ferrão L, et al. Sheep milk: Physicochemical characteristics and relevance for functional food development. *Compr Rev Food Sci Food Saf.* 2017;16(2):247–262.
- Raynal-Ljutovac K, Lagriffoul G, Paccard P, Guillet I, Chilliard Y. Composition of goat and sheep milk products: An update. *Small Rumin. Res.* 2008;79(1):57–72.
- Mohapatra A, Shinde AK, Singh R. Sheep milk: A pertinent functional food. *Small Rumin. Res.* 2019;181:6–11.
- Smith W. Comparative effectiveness of α -hydroxy acids on skin properties. *Int J Cosmet Sci.* 1996;18(2):75–83.
- Watson RR, Collier RJ, Preedy VR. *Nutrients in dairy and their implications on health and disease.* Cambridge, MA: Academic Press; 2017:345–362.
- Mayer H, Fiechter G. Physical and chemical characteristics of sheep and goat milk in Austria. *Int. Dairy J.* 2012;24(2):57–63.
- Konar A, Thomas PC, Rook JAF. The concentrations of some water-soluble constituents in the milk of cows, sows, ewes and goats. *J. Dairy Res.* 1971;38(3):333–341.
- Hardwick DC, Linzell JL, Mepham TB. The metabolism of acetate and glucose by the isolated perfused udder. 2. The contribution of acetate and glucose to carbon dioxide and milk constituents. *Biochem J.* 1963;88(2): 213–220.
- Linzell JL, Mepham TB, Peaker M. The secretion of citrate into milk. *J Physiol.* 1976;260(3):739–750.
- Davies DT, White JCD. The use of ultrafiltration and dialysis in isolating the aqueous phase of milk and in determining the partition of milk constituents between the aqueous and disperse phases. *J. Dairy Res.* 1960;27(2):171–190.
- Gaucheron F. The minerals of milk. *Reprod Nutr Dev.* 2005;45(4):473–483.
- Garnsworthy P, Masson L, Lock A, Mottram T. Variation of milk citrate with stage of lactation and de novo fatty acid synthesis in dairy cows. *J Dairy Sci.* 2006;89(5):1604–1612.
- Xu W, van Kneysel A, Saccenti E, van Hoeij R, Kemp B, Vervoort J. Metabolomics of milk reflects a negative energy balance in cows. *J Proteome Res.* 2020;19(8): 2942–2949.
- De Vries M, Veerkamp R. Energy balance of dairy cattle in relation to milk production variables and fertility. *J Dairy Sci.* 2000;83(1):62–69.
- Timlin M, Tobin JT, Brodkorb A, et al. The impact of seasonality in pasture-based production systems on milk composition and functionality. *Foods.* 2021; 10(3):607.
- Sundekilde UK, Frederiksen PD, Clausen MR, Larsen LB, Bertram HC. Relationship between the metabolite profile and technological properties of bovine milk from two dairy breeds elucidated by NMR-based metabolomics. *J Agric Food Chem.* 2011;59(13):7360–7367.
- Poulsen NA, Buitenhuis AJ, Larsen LB. Phenotypic and genetic associations of milk traits with milk coagulation properties. *J Dairy Sci.* 2015;98(4):2079–2087.
- Buitenhuis AJ, Sundekilde UK, Poulsen NA, Bertram HC, Larsen LB, Sørensen P. Estimation of genetic parameters and detection of quantitative trait loci for metabolites in Danish Holstein milk. *J Dairy Sci.* 2013;96(5):3285–3295.
- Yang Y, Zheng N, Zhao X, et al. Metabolomic biomarkers identify differences in milk produced by Holstein cows and other minor dairy animals. *J Proteomics.* 2016;136:174–182.
- Lv Z, Liu H, Yang Y, et al. Changes in metabolites from bovine milk with beta-casein variants revealed by metabolomics. *Animals.* 2020;10(6):954.
- Merlin Junior IA, Sifuentes dos Santos J, Grecco Costa L, et al. Sheep milk: physical-chemical characteristics and microbiological quality. *Arch Latinoam Nutr.* 2015;65(3):193–198.
- Harding F. Hygienic quality. In: *Milk Quality.* Boston, MA: Springer US; 1995:40–59.
- Troy H, Sharp PF. Detection of lactic acid in milk and cream. *J Dairy Sci.* 1934;17(12):759–762.
- Schmidt KA, Stupar J, Shirley JE, Adapa S. Factors affecting titratable acidity in raw milk. *Kans Agric Exp Stn Res Rep.* 1996;(2):60–62.
- Caboni P, Murgia A, Porcu A, et al. A metabolomics comparison between sheep's and goat's milk. *Food Res Int.* 2019;119:869–875.
- Davis SR, Farr VC, Prosser CG, et al. Milk L-lactate concentration is increased during mastitis. *J Dairy Res.* 2004;71(2):175–181.
- Farr VC, Prosser CG, Nicholas GD, Lee J, Hart AL, Davis SR. Increased milk lactic acid concentration is an early indicator of mastitis. *N Z Soc Anim Prod.* 2002;62:22–23.
- Oshima M, Fuse H. Citric acid concentration in sub-clinical mastitic milk. *J Dairy Res.* 1981;48(3):387–392.
- Marshall AC, Lopez-Villalobos N, Loveday SM, Ellis A, McNabb W. Modelling lactation curves for dairy sheep in a New Zealand flock. *Animals (Basel).* 2023;13(3):349.
- SVS. *SNP & Variation Suite™ [Software].* 8th ed. Bozeman MT, Ed. Boston, MA: SVS Inc.; Bozeman, MT: Golden Helix Inc; 2023. Accessed March 10, 2023. https://www.goldenhelix.com/products/SNP_Variation/index.html.
- SAS Statistical Analysis System, Version 9.4.* Cary, NC: SAS Institute Inc; 2004.
- Gilmour A, Gogel B, Cullis B, Welham S, Thompson R. *ASReml User Guide Release 4.1 Structural Specification.* Hemel Hempstead, UK: VSN International Ltd; 2015.
- Falconer DS, Mackay M. *Introduction to Quantitative Genetics.* 4th ed. Harlow, UK: Addison Wesley Longman, 1996.
- Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for genome-wide complex trait analysis. *Am J Hum Genet.* 2011;88(1):76–82.

38. Armstrong RA. When to use the Bonferroni correction. *Ophthalmic Physiol Opt.* 2014;34(5):502–508.
39. Stephen DT. qqman: An R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv.* 2014:005165.
40. Yates AD, Achuthan P, Akanni W, et al. Ensembl 2020. *Nucleic Acids Res.* 2020;48(D1):D682–D688.
41. Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 2015;43(Database issue):D447–452.
42. García-Quintans N, Blancato VS, Repizo GD, Magni C, López P. Citrate metabolism and aroma compound production in lactic acid bacteria. In: Mayo B, López P, Pérez-Martínez G, eds. *Molecular Aspects of Lactic Acid Bacteria for Traditional and New Applications*; Kerala, India: Research Signpost; 2008.
43. Fox PF, McSweeney PL, Paul L. Dairy chemistry and biochemistry. London, UK: Blackie Academic and Professional, Thomson Science; 1998.
44. Baticz O, Tömösközi S, Vida L, et al. Relationship between concentration of citrate and ketone bodies in cow's milk. *Acta Vet Hung.* 2002;50(3):253–261.
45. Churakov M, Karlsson J, Rasmussen AE, Holtenius K. Milk fatty acids as indicators of negative energy balance of dairy cows in early lactation. *Animal.* 2021;15(7):100253.
46. Buttchereit N, Stamer E, Junge W, Thaller G. Evaluation of five lactation curve models fitted for fat: protein ratio of milk and daily energy balance. *J Dairy Sci.* 2010;93(4):1702–1712.
47. Sanchez MP, Rocha D, Charles M, et al. Sequence-based GWAS and post-GWAS analyses reveal a key role of SLC37A1, ANKH, and regulatory regions on bovine milk mineral content. *Sci Rep.* 2021;11(1):7537.
48. Chen Y, Hu H, Atashi H, et al. Genetic analysis of milk citrate predicted by milk mid-infrared spectra of Holstein cows in early lactation. *J Dairy Sci.* 2023;107(5):3047–3061.
49. Geiger K, Muendlein A, Leiherer A, et al. The association between the adipokine C1QTNF1 and type 2 diabetes is significantly modulated by obesity. *Diabet.* 2018;67(Supplement_1): A14846.
50. Granchi C. ATP citrate lyase (ACLY) inhibitors: An anti-cancer strategy at the crossroads of glucose and lipid metabolism. *Eur J Med Chem.* 2018;157:1276–1291.
51. Zhang W, Xu M, Wang J, et al. Comparative transcriptome analysis of key genes and pathways activated in response to fat deposition in two sheep breeds with distinct tail phenotype. *Front Genet.* 2021;12:639030.
52. Liu N, He J, Yu W, et al. Transcriptome analysis of skeletal muscle at prenatal stages in Polled Dorset versus Small-tailed Han sheep. *Genet Mol Res.* 2015;14(1):1085–1095.
53. Chen P-W, Gasilina A, Yadav MP, Randazzo PA. Control of cell signaling by Arf GTPases and their regulators: Focus on links to cancer and other GTPase families. *Biochim Biophys Acta Mol Cell Res.* 2022;1869(2):119171.
54. Luo C, Peng W, Kang J, et al. Glutamine regulates cell growth and casein synthesis through the CYTHs/ARFGAP1-Arf1-mTORC1 pathway in bovine mammary epithelial cells. *J Agric Food Chem.* 2021;69(24):6810–6819.
55. Yang X, Tai Y, Ma Y, et al. Cecum microbiome and metabolism characteristics of Silky Fowl and White Leghorn chicken in late laying stages. *Front Microbiol.* 2022;13:984654.
56. Cogan T. Co-metabolism of citrate and glucose by *Leuconostoc* spp.: effects on growth, substrates and products. *J Appl. Bacteriol.* 1987;63(6):551–558.
57. Riera Leal A, Ortiz-Lazareno PC, Jave-Suárez LF, et al. 17beta-estradiol-induced mitochondrial dysfunction and Warburg effect in cervical cancer cells allow cell survival under metabolic stress. *Int J Oncol.* 2020;56(1):33–46.
58. Sahana G, Guldbbrandtsen B, Thomsen B, et al. Genome-wide association study using high-density single nucleotide polymorphism arrays and whole-genome sequences for clinical mastitis traits in dairy cattle. *J Dairy Sci.* 2014;97(11):7258–7275.
59. Miao Z, Ding Y, Bi Y, Chen M, Cao X, Wang F. Staphylococcus aureus on the effect of expression of MMPs/TIMPs and uPA system in bovine mammary fibroblasts. *J Microbiol Immunol Infect.* 2021;54(3):411–419.
60. Triner D, Castillo C, Hakim JB, et al. Myc-associated zinc finger protein regulates the proinflammatory response in colitis and colon cancer via STAT3 signaling. *Mol Cell Biol.* 2018;38(22):E00386.
61. Páth G, Mehana AE, Pilz IH, et al. NUPR1 preserves insulin secretion of pancreatic β -cells during inflammatory stress by multiple low-dose streptozotocin and high-fat diet. *Am J Physiol Endocrinol Metab.* 2020;319(2):E338–E344.
62. Han H. Identification of several key genes by microarray data analysis of bovine mammary gland epithelial cells challenged with *Escherichia coli* and *Staphylococcus aureus*. *Gene.* 2019;683:123–132.
63. Ahlawat S, Choudhary V, Singh T, et al. First report on delineation of differentially expressed genes and pathways in milk somatic cells of mastitic and healthy Murrah buffaloes. *Gene.* 2022;831:146575.
64. Biswas SK, Chittezhath M, Shalova IN, Lim J-Y. Macrophage polarization and plasticity in health and disease. *Immunol Res.* 2012;53(1–3):11–24.
65. Zhou HC, Xin-Yan Y, Yu WW, Liang XQ, Du XY, Liu ZC, Long JP, Zhao GH, Liu HB. Lactic acid in macrophage polarization: The significant role in inflammation and cancer. *Int Rev Immunol.* 2022;41(1):4–18.