



Genome-wide association study for milk production traits in a Brazilian Holstein population

L. H. S. lung,¹ J. Petrini,¹ J. Ramírez-Díaz,¹ M. Salvian,¹ G. A. Rovadoscki,¹ F. Pilonetto,¹ B. D. Dauria,¹ P. F. Machado,¹ L. L. Coutinho,¹ G. R. Wiggans,² and G. B. Mourão^{1*}

¹Department of Animal Science, University of São Paulo (USP)/Luiz de Queiroz College of Agriculture (ESALQ), Piracicaba, São Paulo 13418900, Brazil

²Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350

ABSTRACT

Advances in the molecular area of selection have expanded knowledge of the genetic architecture of complex traits through genome-wide association studies (GWAS). Several GWAS have been performed so far, but confirming these results is not always possible due to several factors, including environmental conditions. Thus, our objective was to identify genomic regions associated with traditional milk production traits, including milk yield, somatic cell score, fat, protein and lactose percentages, and fatty acid composition in a Holstein cattle population producing under tropical conditions. For this, 75,228 phenotypic records from 5,981 cows and genotypic data of 56,256 SNP from 1,067 cows were used in a weighted single-step GWAS. A total of 46 windows of 10 SNP explaining more than 1% of the genetic variance across 10 *Bos taurus* autosomes (BTA) harbored well-known and novel genes. The *MGST1* (BTA5), *ABCG2* (BTA6), *DGAT1* (BTA14), and *PAEP* (BTA11) genes were confirmed within some of the regions identified in our study. Potential novel genes involved in tissue damage and repair of the mammary gland (*COL18A1*), immune response (*LTTTC19*), glucose homeostasis (*SLC37A1*), synthesis of unsaturated fatty acids (*LTBP1*), and sugar transport (*SLC37A1* and *MFS4A*) were found for milk yield, somatic cell score, fat percentage, and fatty acid composition. Our findings may assist genomic selection by using these regions to design a customized SNP array to improve milk production traits on farms with similar environmental conditions.

Key words: dairy cattle, genome-wide association study, milk composition

INTRODUCTION

Selection based on phenotype and pedigree information has enabled great improvements in genetic merit in several livestock species. With the advances in the molecular area of selection, it is possible to obtain a wider knowledge about the genetic architecture of complex quantitative traits and use it as an additional information source in breeding programs. The progress in genome sequencing and high-throughput genotyping technologies has made feasible the identification of SNP associated with phenotypes of interest through genome-wide association studies (GWAS).

In dairy cattle, GWAS have identified genetic variants associated with complex phenotypes, such as milk yield and quality traits (e.g., Bouwman et al., 2011, 2012; Schopen et al., 2011; Meredith et al., 2013). However, despite the large number of GWAS conducted so far, the validation of SNP effects is not always possible. Confirming GWAS results in an independent population is the most reliable way to validate associations previously found, as well as providing a clearer view about the genetic architecture underlying complex traits, as large genomic regions can be refined to a few candidate genes (Visscher, 2008; Raven et al., 2014). Factors such as breed, breeding goals, statistical method, linkage disequilibrium (LD) pattern, and definition of the traits can contribute to the variability of GWAS results, mainly in species and breeds widely studied (e.g., Holstein cattle).

Environmental conditions, such as season, nutrition, and management, are also variables that can modify gene expression of most of the traits of interest. Recently, studies revealed that polymorphisms in diacylglycerol O-acyltransferase 1 (*DGAT1*) and stearoyl-CoA desaturase 1 (*SCD1*) genes showed different effects on milk production traits and fatty acid (FA) composition according to the season and temperature (Duchemin et al., 2013; Komisarek and Kolenda, 2016). In addition, a joint GWAS of the Chinese and Danish Holstein populations showed potential genotype by environment

Received March 23, 2018.

Accepted October 19, 2018.

*Corresponding author: gbmourao@usp.br

Table 1. Summary of the data set used in this study

Trait name	Trait abbreviation	N ¹	Mean (SD)	Minimum	Maximum	Heritability ²
Milk yield (kg/d)	MY	53,592	35.41 (9.56)	6.75	63.90	0.13
Somatic cell score	SCS	57,279	2.79 (2.17)	-3.64	9.29	0.12
Fat percentage	FP	56,377	3.43 (0.69)	1.37	5.49	0.19
Protein percentage	PP	56,377	3.07 (0.29)	2.20	3.94	0.35
Lactose percentage	LP	56,377	4.65 (0.20)	4.04	5.26	0.31
Palmitic acid (%)	C16:0	32,379	0.84 (0.20)	0.24	1.45	0.26
Stearic acid (%)	C18:0	32,378	0.61 (0.14)	0.20	1.01	0.14
Oleic acid (%)	C18:1	32,372	0.65 (0.19)	0.10	1.21	0.07
Saturated fatty acids (%)	SFA	32,379	2.23 (0.46)	0.86	3.59	0.25
Unsaturated fatty acids (%)	UFA	32,379	1.01 (0.26)	0.24	1.77	0.08
Monounsaturated fatty acids (%)	MUFA	32,379	0.85 (0.22)	0.19	1.52	0.07
Polyunsaturated fatty acids (%)	PUFA	32,369	0.15 (0.04)	0.03	0.28	0.11

¹N = number of observations.

²Obtained using combined pedigree-genomic relationship matrix (Petrini et al., 2016).

interaction for milk FA traits (Li et al., 2015). The influence of heat stress on milk yield and composition traits is well described in the review by Tao et al. (2018). Knowing that different GWAS results can be found according to the environmental conditions will help us to make better selection decisions. Therefore, the aim of our study was to perform a GWAS to identify genomic regions affecting milk yield and composition, including FA profile, in a Holstein cattle population producing under tropical conditions.

MATERIALS AND METHODS

Phenotypic Data

We used 75,228 phenotypic records collected between May 2012 and December 2016 of 5,981 Holstein cows located on 4 Brazilian farms from first to sixth lactation in our study. A detailed description of the phenotypic quality control, environment, and production system in each farm can be seen in Petrini et al. (2016), and a summary of the phenotypic data is shown in Table 1. These cows were assigned to 355 contemporary groups (**CG**) and were daughters of 317 sires. The CG were formed by concatenating information from season and year of calving as well as farm and month of analysis. The pedigree included 10,491 animals. The traits studied were test-day milk yield (**MY**; kg/d), SCS (transformed according to Ali and Shook, 1980), fat percentage (**FP**), protein percentage (**PP**), lactose percentage (**LP**), palmitic acid (**C16:0**), stearic acid (**C18:0**), oleic acid (**C18:1**), SFA, UFA, MUFA, and PUFA. The milk components were expressed as grams per 100 g of milk and measured by mid-infrared spectroscopy (Delta Instruments CombiScope Filter, Advanced Instruments Inc., Norwood, MA; Rodriguez et al., 2014). Comparing milk samples measured by mid-infrared spectroscopy and GC, with fatty acids

expressed as grams per 100 g of milk fat, we observed a high association according to the Pearson correlations (0.6–0.9) and a slight discrepancy (94% of the samples were within the concordance range) measured by Bland-Altman test, which indicated that both methods show a similar pattern. Thus, the measurement method would not affect conclusions.

The SFA group included C4:0, C6:0, C8:0, C10:0, C11:0, C12:0, C13:0, C14:0, C15:0, C16:0, C17:0, C18:0, C19:0, C23:0, and C24:0; the MUFA group included C14:1, C15:1, C16:1, C17:1, C18:1n-9, C18:1n-7, C19:1, C22:1n-9, and C24:1; the PUFA group included C18:2n-6, C18:3n-3, *cis*-9,*trans*-11-CLA, C20:3n-3, C22:2, eicosapentaenoic acid, C20:5; and the MUFA plus PUFA composed the UFA group.

Genotypes

Hair root samples from 1,136 cows were collected for DNA extraction using the NucleoSpin Tissue Kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany). Genotypes from 747 cows were obtained with the Illumina Bovine LD BeadChip (Illumina, San Diego, CA), containing 6,909 SNP, and the remaining 389 cows with the GeneSeek Genomic Profiler Bovine 50k (Neogen Corporation, Lincoln, NE), containing 47,843 SNP. All genotypes were imputed to 60,671 SNP using findhap.f90 (VanRaden et al., 2013). The reference population was composed of 1,584,539 females and 237,570 males, all Holstein, genotyped with a variety of chips manufactured largely by Illumina and provided by Illumina, GeneSeek (Neogen Agrigenomics, Lexington, KY) and Zoetis (Zoetis, Parsippany, NJ). The individual chips contributed from 2,710 to 56,643 of the 60,671 SNP used to form the haplotype library used to impute the genotypes. Markers located in sex chromosomes ($n = 1,681$), monomorphic ($n = 70$), with proportion of missing genotypes >10% ($n = 1,160$) and with minor allele

frequency <0.02 ($n = 1,504$), and samples with call rate <90% ($n = 69$) were excluded. After this, 1,067 cows and 56,256 SNP remained for association analysis.

GWAS

The association analysis was performed using the weighted single-step genomic BLUP (**WssGBLUP**) method (Wang et al., 2012). The WssGBLUP allows the use of all available phenotypic, genotypic, and pedigree information and does not require adjusted measures as response variables. Additionally, using different weights for SNP also mimics the polygenic nature of the studied traits by taking the presence of major genes into account. The fitted model was

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{S}\mathbf{c} + \mathbf{e},$$

where \mathbf{y} is the vector of phenotypic records; $\boldsymbol{\beta}$ is the vector of fixed effects including CG, lactation order, and DIM (cubic regression); \mathbf{a} is the vector of random additive genetic effects; \mathbf{c} is the vector of random permanent environmental effects; \mathbf{e} is the vector of random residual effects; and \mathbf{X} , \mathbf{Z} , and \mathbf{S} are the incidence matrices relating phenotypic records to fixed, random animal, and permanent environmental effects, respectively. We assumed that $\mathbf{a} \sim N(0, \mathbf{H}\sigma_a^2)$, $\mathbf{c} \sim N(0, \mathbf{I}\sigma_c^2)$, and $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$, where σ_a^2 , σ_c^2 , and σ_e^2 are the additive genetic, permanent environmental, and residual variances, respectively; \mathbf{I} is an identity matrix; and \mathbf{H} is the matrix that combines pedigree and genomic information (Aguilar et al., 2010). Its inverse is given by

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}_w^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix},$$

where \mathbf{A}^{-1} is the inverse of the pedigree relationship matrix, \mathbf{G}_w^{-1} is the inverse of the genomic relationship matrix, and \mathbf{A}_{22}^{-1} is the inverse of the pedigree relationship matrix for genotyped animals only. To ensure full rank, the \mathbf{G}_w^{-1} matrix was created as $\mathbf{G}_w = \alpha\mathbf{G} + \beta\mathbf{A}_{22}$, where \mathbf{G} was constructed as described in VanRaden (2008) and using the default parameterization in the preGSf90 ($\alpha = 0.95$ and $\beta = 0.05$). The WssGBLUP was performed using BLUPF90 family programs (Misztal et al., 2002). The SNP effects (\hat{u}) were calculated iteratively following scenario S2 as described by Wang et al. (2012):

$$\hat{u} = \mathbf{DZ}'[\mathbf{ZDZ}']^{-1}\hat{a}_g,$$

where \mathbf{D} is the diagonal matrix of weights for variances of SNP, \mathbf{Z} is the matrix that relates the genotypes to each locus, and \hat{a}_g is the genomic EBV of genotyped animals. The iterative process was repeated 2 times. In the first iteration, the same weight (1) for all SNP was assumed (i.e., $\mathbf{D} = \mathbf{I}$), and in the second the weight for each SNP was obtained as

$$d_i = \hat{u}_i^2 2p_i(1 - p_i),$$

where d is the diagonal elements of the \mathbf{D} matrix, i is the i th SNP, and p_i is the frequency of the second allele of the i th SNP.

The genetic variance explained by each 10-SNP window ($n = 5,640$) was calculated based on the SNP effect (\hat{u}_i) and allele frequencies (p_i and q_i) of 10 consecutive markers and then divided by the total genetic variance explained by all SNP to convert to a proportion. The concept of grouping SNP into windows was adopted as a way to better capture the effect of a QTL instead of a single SNP (Habier et al., 2011). The 10-SNP windows that explained more than 1% of the genetic variance were considered promising regions and used to identify potential candidate genes by using the National Center for Biotechnology Information (NCBI) database and Ensembl Genome Browser (<http://www.ensembl.org/index.html>) through the biomaRt R package (<http://www.bioconductor.org>; Ensembl release 90). The threshold of 1% was chosen to identify the most important windows and to use the level used in other studies (e.g., Gonzalez-Pena et al., 2016; Zhu et al., 2017; Marques et al., 2018; Pereira et al., 2018). The Manhattan plots were created using qqman package (Turner, 2014).

RESULTS AND DISCUSSION

The Manhattan plots of the proportion of genetic variance explained by 10-SNP windows for each trait are shown in Figure 1. The peak on BTA14 explained the largest proportion of variance for FP and most of the FA traits. Additional regions explaining large proportions of the variance were found on BTA6 (37.86 to 40.62 Mbp) for LP and BTA11 (103.04 to 106.43 Mbp) for PUFA. The 10-SNP windows that explained more than 1% of the genetic variance are listed in Table 2. A total of 46 windows distributed over 10 chromosomes (BTA1, 5, 6, 8, 9, 11, 12, 14, 16, and 19) explained more than 1% of the genetic variance. Taken together, these windows explained from 1.20 to 18.60% of the genetic variance across all traits. The results will be further discussed by chromosome.

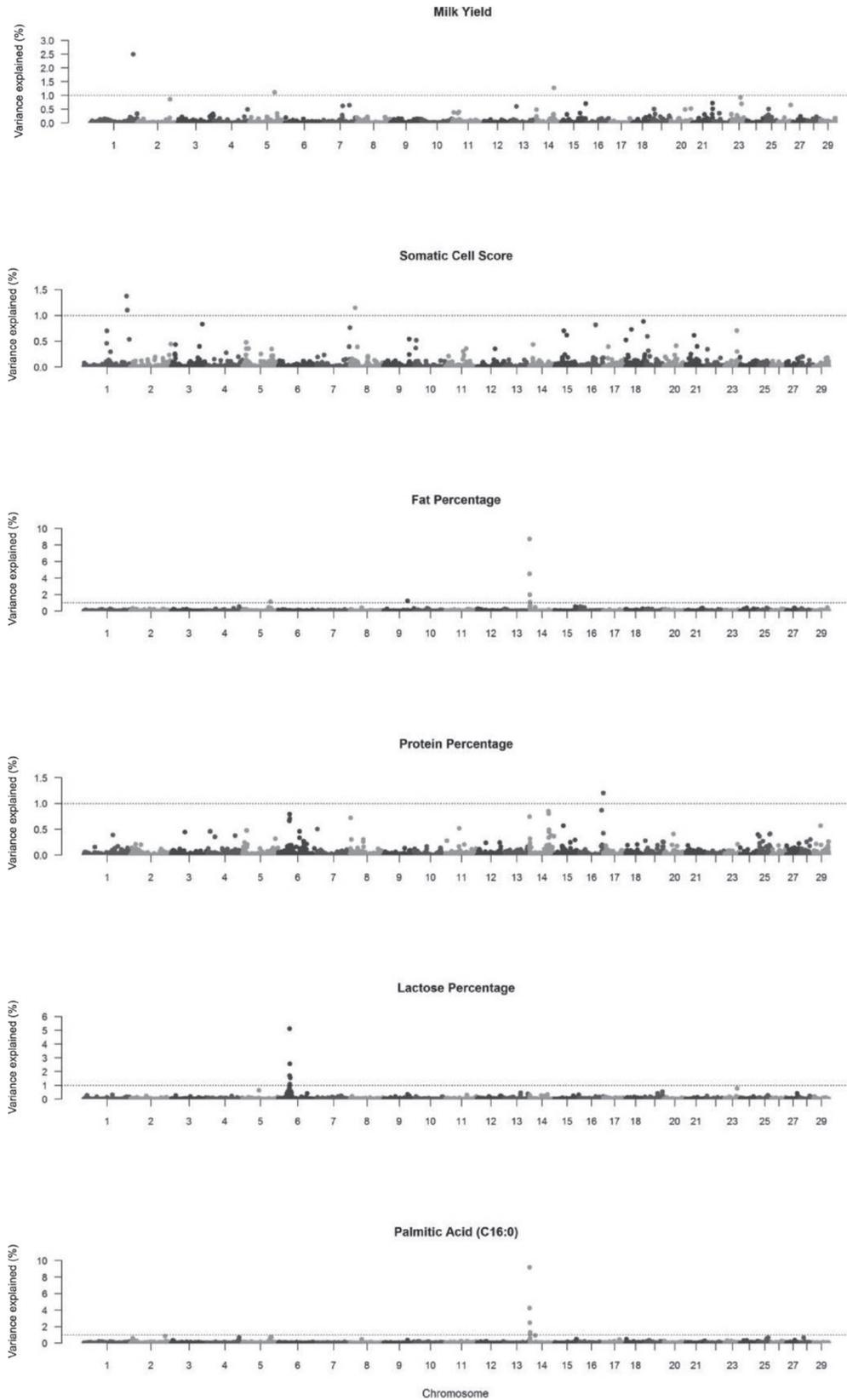


Figure 1. Manhattan plot of the proportion of genetic variance explained by 10-SNP windows for each trait. The dotted black line indicates the adopted threshold (1%).

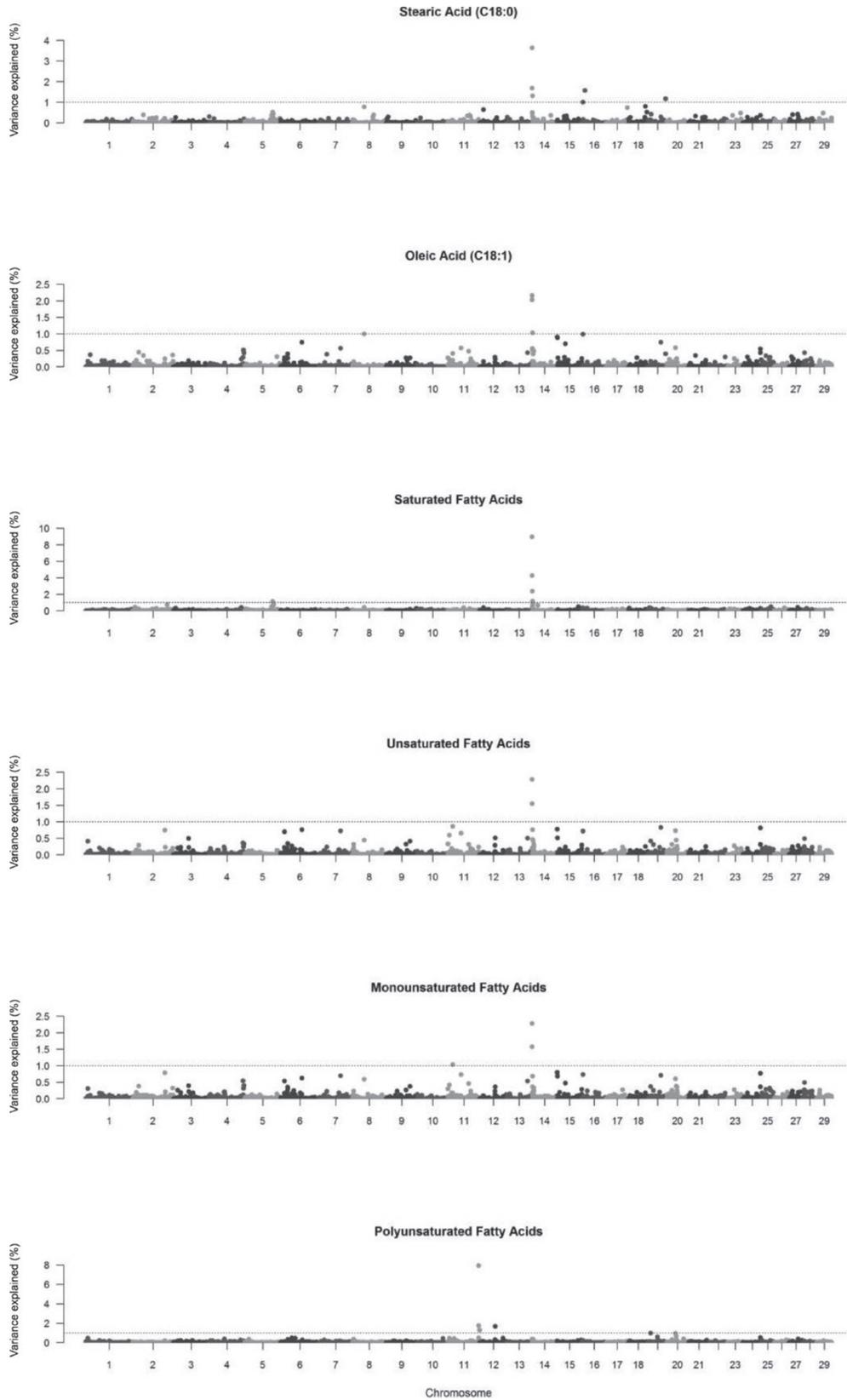


Figure 1 (Continued). Manhattan plot of the proportion of genetic variance explained by 10-SNP windows for each trait. The dotted black line indicates the adopted threshold (1%).

On BTA1, a window from 144.38 to 145.13 Mbp was associated simultaneously with MY and SCS, where the solute carrier family 37 member 1 (*SLC37A1*) and phosphodiesterase 9A (*PDE9A*) genes are located. The *SLC37A1* gene encodes a glucose-6-phosphate transporter involved in the glucose blood homeostasis (Pan et al., 2011) and is highly expressed in the mammary gland (Kemper et al., 2015; Raven et al., 2016). Previously, *SLC37A1* was associated with MY, FP, and PP (Kemper et al., 2016; Raven et al., 2016; Pausch et al., 2017), supporting our results for MY, and with other milk components such as α_{S1} -CN and α -LA (Sanchez et al., 2017), phosphorus concentration (Buitenhuis et al., 2015), and FA composition (Olsen et al., 2017). The *PDE9A* gene is a cGMP-specific phosphodiesterase also highly expressed in mammary gland and previously associated with milk production traits (Yang et al., 2015).

Furthermore, on BTA1 an additional window was associated with SCS at about 147 Mbp. The collagen type XVIII α 1 chain (*COL18A1*) gene is located less than 0.10 Mbp downstream from this window and was

considered the most likely candidate. Collagen genes are among those differentially expressed in mammary gland in response to mastitis, highlighting their role in tissue damage and repair (Huang et al., 2014; Wang et al., 2016).

Two windows on BTA5 were found, the first at about 86 Mbp for MY and the second from 93.66 to 94.10 Mbp for FP and SFA. No potential candidate gene was identified in the first window; however, SNP previously associated with milk, protein, and fat production were found in this region (Nayeri et al., 2016). The second region harbors the microsomal glutathione S-transferase 1 (*MGST1*) gene, indicated as the most likely causal mutation despite its unclear role in regulating milk composition (Littlejohn et al., 2016). The *MGST1* gene encodes a protein that protects against oxidative stress and participates in glutathione metabolism. The effect of *MGST1* on FP and other milk composition traits has been reported in several studies (e.g., Jiang et al., 2010; Littlejohn et al., 2016; Raven et al., 2016).

A peak formed by 5 contiguous windows from 37.86 to 40.62 Mbp was associated with LP on BTA6. The

Table 2. The 10-SNP windows that explained more than 1% of the genetic variance for milk yield and composition traits

Window ¹	Trait ²												PCG ³	
	MY	SCS	FP	PP	LP	C16:0	C18:0	C18:1	SFA	UFA	MUFA	PUFA		
1_144.38:145.12	2.49	1.37												<i>SLC37A1</i> , <i>PDE9A</i>
1_147.12:147.38		1.10												<i>COL18A1</i>
5_86.24:86.50	1.11													—
5_93.66:94.10			1.13						1.12					<i>MGST1</i>
6_37.86:38.13					5.11									<i>ABCG2</i>
6_38.13:38.30					1.69									<i>ABCG2</i>
6_38.62:38.93					2.55									<i>ABCG2</i>
6_39.03:39.34					1.07									<i>ABCG2</i>
6_40.32:40.62					1.55									<i>ABCG2</i>
8_16.91:17.46		1.15												<i>LRRRC19</i>
9_80.88:81.32			1.21											—
11_15.75:16.02											1.03			<i>LTBP1</i>
11_103.04:103.28												1.75		<i>PAEP</i>
11_103.28:103.47												7.93		<i>PAEP</i>
11_106.09:106.43												1.29		<i>PAEP</i>
12_52.49:52.82												1.69		—
14_1.18:1.65			4.49			4.24	1.68	2.03	4.25	1.55	1.57			<i>DGAT1</i>
14_1.67:2.00			8.72			9.18	3.63	2.16	8.96	2.28	2.27			<i>DGAT1</i>
14_2.05:2.27			1.97			2.46			2.39					<i>DGAT1</i>
14_3.07:3.43			1.08			1.07	1.31	1.03	1.17					<i>DGAT1</i>
14_3.46:3.70						1.28								<i>DGAT1</i>
14_61.44:62.08	1.27													—
16_3.07:3.42							1.00							<i>MFSD4A</i> , <i>SLC35A3</i>
16_10.30:10.74							1.56							—
16_80.95:81.18				1.20										—
19_60.24:60.44							1.17							—
Sum	4.87	3.62	18.60	1.20	11.97	18.23	10.35	5.22	17.89	3.83	4.87	12.66		

¹Represented by BTA and physical position (in Mbp) (BTA_Mbp:Mbp).

²MY = milk yield; FP = fat percentage; PP = protein percentage; LP = lactose percentage; C16:0 = palmitic acid; C18:0 = stearic acid; C18:1 = oleic acid.

³PCG = potential candidate genes.

ATP-binding cassette subfamily G member 2 (*ABCG2*) gene is located in the beginning of this large region. The effects of *ABCG2* on milk production traits have been well reported in the literature (e.g., Nayeri et al., 2016; Lopdell et al., 2017; Pausch et al., 2017; Sanchez et al., 2017).

On BTA8, a window from 16.91 to 17.54 Mbp was associated with SCS. Among the genes found in this region, leucine-rich repeat containing 19 (*LRRC19*) was thought as the most likely candidate for belonging to the LRR family as well as toll-like receptors, which are involved in immune response. Evidence supporting the association between *LRRC19* and immune response was previously observed in mice and humans (Chai et al., 2009; Su et al., 2014; Cao et al., 2016).

A window in the beginning (15.75–16.02 Mbp) and 3 others at the end (103.04–106.43 Mbp) of BTA11 were associated with MUFA and PUFA, respectively. The first region harbors the latent transforming growth factor β -binding protein 1 (*LTBP1*) gene, which regulates transforming growth factor- β activity (Saharinen et al., 1999). In a human study, the expression of the stearoyl-CoA desaturase, an important enzyme involved in the synthesis of UFA, shown to be controlled by transforming growth factor- β , supports our finding (Samuel et al., 2002). The second region on BTA11 was formed by 2 subsequent windows from 103.04 to 103.47 Mbp and another at about 106 Mbp. The progesterone-associated endometrial protein (*PAEP*) gene is located at 103.3 Mbp and is well known to affect protein composition (e.g., Raven et al., 2016; Sanchez et al., 2017); however, recent evidence suggesting its relation with FA was reported (Knutsen et al., 2018). Those authors reported that its relationship can be due to the potential transporter effect of *PAEP* by binding with SFA and UFA found in vitro (Le Maux et al., 2014) or due to the high LD with glycosyltransferase 6 domain containing 1 (*GLT6D1*). Other genes involved in lipid metabolism were found in this region, including carboxyl ester lipase (*CEL*; 103 Mbp), globoside α -1,3-N-acetylgalactosaminyltransferase 1 (*GBGT1*; 103 Mbp), ATP-binding cassette subfamily A member 2 (*ABCA2*; 106 Mbp), and prostaglandin D2 synthase (*PTGDS*; 106 Mbp).

The region that explained the highest proportions of genetic variance for FP and all FA, except PUFA, was observed on BTA14 (1.19 to 3.70 Mbp, comprising 5 windows). The *DGAT1* gene is located at 1.8 Mbp, which is well known to affect milk production and composition (e.g., Bovenhuis et al., 2015; 2016; Buitenhuis et al., 2016; Sanchez et al., 2017). Partially, these results can be also assigned to genes neighboring *DGAT1* that were previously related to milk production traits, such as *CYHR1*, *VPS28*, *MROH1*, *OPLAH*, and *GPR20* (e.g.,

Jiang et al., 2014; Frąszczak and Szyda, 2016; Nayeri et al., 2016; Suchocki et al., 2016). Bennewitz et al. (2004) showed that *DGAT1* is not exclusively responsible for genetic variation related to milk production. A study that applied 3 different methods for sample stratification correction to evaluate the effects on dairy GWAS identified *DGAT1* as the most significant for fat yield, and in one method the A5D786-CYHR1-VPS28-DGAT1 region was identified as the most significant for MY (Ma et al., 2012); this suggests that there is an LD region affecting milk production traits. In our study, the window from 1.67 to 2.0 Mbp was the most important in this region for FP, C16:0, C18:0, C18:1, SFA, UFA, and MUFA, which highlights *DGAT1* as the main candidate by capturing most of the signal from the QTL in this large region.

On BTA16, a window from 3.07 to 3.42 Mbp was associated with C18:0, where *MFSD4A*, *SLC41A1*, and *SLC45A3*, all solute carrier family members, are located. The *MFSD4A* gene was identified as a transporter expressed in food regulatory brain areas in mice (Perland et al., 2017). Those authors also suggested *MFSD4A* can be a sugar transporter for sharing 20% of the AA with *MFSD4B*, a sodium-dependent sugar transporter (Horiba et al., 2003). The *SLC41A1* gene is a magnesium transporter (Goytain and Quamme, 2005), and *SLC45A3* belong to the SLC45 family known to encode sugar transporters and play an important role in myelin maintenance by modulating glucose and lipid metabolism in oligodendrocytes (Shin et al., 2012; Vitavska and Wiczorek, 2013). These findings suggest that *MFSD4A* and *SLC45A3*, both sugar transporters, can be good candidates for C18:0 regulating loci, as sugar may be used for the synthesis of FA.

Novel regions without potential candidate genes were found on BTA9 (80.88–81.32 Mbp) for FP, BTA12 (52.49–52.82 Mbp) for PUFA, BTA14 (61.44–62.08 Mbp) for MY, and BTA16 (10.30–10.74 Mbp) and BTA19 (60.18–60.32 Mbp) for C18:0. In some of these regions, QTL were previously associated with milk production and composition traits, including FA (Bouman et al., 2011, 2012; Meredith et al., 2012). For PP, only 1 window explained more than 1% of the genetic variance (BTA16, 80.95–81.18 Mbp), but no promising gene was identified.

The regions shared between FP and FA (BTA5 and 14) represent the genetic correlations among these traits. If 2 traits are positively and highly correlated with each other, it is expected that the same genes affect both. Fat percentage had more windows in common with the groups of SFA (C16:0, C18:0, and SFA) than the group of UFA (C18:1, UFA, MUFA, and PUFA). These findings are supported by the higher genetic correlation estimates found between FP and SFA compared with

UFA (Bastin et al., 2011; Penasa et al., 2015; Petrini et al., 2016).

In general, we identified regions previously reported and novel, even using data from a Holstein population raised in a tropical condition. Partially, these known regions may have been identified by the intense use of AI, facilitating the dissemination of genetic material from influential bulls, mainly from United States and Canada, and making our population genetically closer to others previously studied. This suggests that major genomic regions for milk traits can be better captured than regions with minor effects, regardless of the environmental condition of the analyzed population.

Besides helping to reveal the genetics behind milk production traits, our findings can also be used to improve traits through genomic selection. The cost of high-density genotyping is still a limitation in Brazil; therefore, a low-density customized SNP array incorporating GWAS results may improve accuracy of genomic prediction, as previously reported (e.g., Zhang et al., 2014; Song et al., 2018). The magnitude of the gain in accuracy will depend on the genetic architecture of the trait; for example, FP is known to be strongly determined by a major region, *DGAT1*, whereas SCS seems to be influenced by regions with small effects according to studies published to date. Therefore, the validation of the effects found in our study in independent populations is needed to determine their value in the selection process.

CONCLUSIONS

A genome-wide association study using data from a Brazilian Holstein cattle population revealed 46 genomic regions on 10 BTA associated with milk production traits, including FA. Four of these regions harbor genes previously associated with milk production traits (*MGST1*, *ABCG2*, *PAEP*, and *DGAT1*). Genes involved in tissue damage and repair, immune response, glucose homeostasis, synthesis of UFA, and sugar transport are located in the novel regions. Besides helping to unravel the genetic mechanisms underlying milk production traits, our findings may also be useful to compose a custom SNP array as a cost-effective approach to improve these traits through genomic selection.

ACKNOWLEDGMENTS

We thank the Coordination for the Improvement of Higher Education Personnel (CAPES, Brasília, Brazil) and National Council for Scientific Technological Development (CNPq, Brasília, Brazil) for providing financial support (grants number: 553966/2009-0, 484560/2012-3, 433056/2016-9), fellowships to Paulo Machado, Luiz

Coutinho and Gerson Mourão, and São Paulo Research Foundation (FAPESP, São Paulo, Brazil) for providing financial support (grant number: 2010/12929-6) and scholarship to Laiza Iung and Juliana Petrini (grant numbers: 2012/24788-3, 2012/15948-7). The authors thank Clínica do Leite and Department of Animal Science, 'Luiz de Queiroz' College of Agriculture, University of São Paulo (ESALQ/USP) for assistance and for providing the database, and the Cooperative Dairy DNA Repository (CDDR, Beltsville, MD) for providing the genotypes used as reference population in imputation analysis.

REFERENCES

- Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* 93:743–752. <https://doi.org/10.3168/jds.2009-2730>.
- Ali, A. K. A., and G. E. Shook. 1980. An optimum transformation for somatic cell concentration in milk. *J. Dairy Sci.* 63:487–490. [https://doi.org/10.3168/jds.S0022-0302\(80\)82959-6](https://doi.org/10.3168/jds.S0022-0302(80)82959-6).
- Bastin, C., N. Gengler, and H. Soyeurt. 2011. Phenotypic and genetic variability of production traits and milk fatty acid contents across days in milk for Walloon Holstein first-parity cows. *J. Dairy Sci.* 94:4152–4163. <https://doi.org/10.3168/jds.2010-4108>.
- Bennowitz, J., N. Reinsch, S. Paul, C. Looft, B. Kaupe, C. Weimann, G. Erhardt, G. Thaller, C. Kühn, M. Schwerin, H. Thomsen, F. Reinhardt, R. Reents, and E. Kalm. 2004. The *DGAT1* K232A mutation is not solely responsible for the milk production quantitative trait locus on the bovine chromosome 14. *J. Dairy Sci.* 87:431–442. [https://doi.org/10.3168/jds.S0022-0302\(04\)73182-3](https://doi.org/10.3168/jds.S0022-0302(04)73182-3).
- Bouwman, A. C., H. Bovenhuis, M. H. P. W. Visker, and J. A. M. van Arendonk. 2011. Genome-wide association of milk fatty acids in Dutch dairy cattle. *BMC Genet.* 12:43. <https://doi.org/10.1186/1471-2156-12-43>.
- Bouwman, A. C., M. H. Visker, J. A. van Arendonk, and H. Bovenhuis. 2012. Genomic regions associated with bovine milk fatty acids in both summer and winter milk samples. *BMC Genet.* 13:93. <https://doi.org/10.1186/1471-2156-13-93>.
- Bovenhuis, H., M. H. P. W. Visker, N. A. Poulsen, J. Sehested, H. J. F. van Valenberg, J. A. M. van Arendonk, L. B. Larsen, and A. J. Buitenhuis. 2016. Effects of the diacylglycerol o-acyltransferase 1 (*DGAT1*) K232A polymorphism on fatty acid, protein, and mineral composition of dairy cattle milk. *J. Dairy Sci.* 99:3113–3123. <https://doi.org/10.3168/jds.2015-10462>.
- Bovenhuis, H., M. H. P. W. Visker, H. J. F. van Valenberg, A. J. Buitenhuis, and J. A. M. van Arendonk. 2015. Effects of the *DGAT1* polymorphism on test-day milk production traits throughout lactation. *J. Dairy Sci.* 98:6572–6582. <https://doi.org/10.3168/jds.2015-9564>.
- Buitenhuis, B., N. A. Poulsen, G. Gebreyesus, and L. B. Larsen. 2016. Estimation of genetic parameters and detection of chromosomal regions affecting the major milk proteins and their post-translational modifications in Danish Holstein and Danish Jersey cattle. *BMC Genet.* 17:114. <https://doi.org/10.1186/s12863-016-0421-2>.
- Buitenhuis, B., N. A. Poulsen, L. B. Larsen, and J. Sehested. 2015. Estimation of genetic parameters and detection of quantitative trait loci for minerals in Danish Holstein and Danish Jersey milk. *BMC Genet.* 16:52. <https://doi.org/10.1186/s12863-015-0209-9>.
- Cao, S., X. Su, B. Zeng, H. Yan, Y. Huang, E. Wang, H. Yun, Y. Zhang, F. Liu, W. Li, H. Wei, Y. Che, and R. Yang. 2016. The gut epithelial receptor LRRC19 promotes the recruitment of immune cells and gut inflammation. *Cell Reports* 14:695–707. <https://doi.org/10.1016/j.celrep.2015.12.070>.

- Chai, L., L. Dai, Y. Che, J. Xu, G. Liu, Z. Zhang, and R. Yang. 2009. LRR19, a novel member of the leucine-rich repeat protein family, activates NF- κ B and induces expression of proinflammatory cytokines. *Biochem. Biophys. Res. Commun.* 388:543–548. <https://doi.org/10.1016/j.bbrc.2009.08.043>.
- Duchemin, S., H. Bovenhuis, W. M. Stoop, A. C. Bouwman, J. A. M. van Arendonk, and M. H. P. W. Visker. 2013. Genetic correlation between composition of bovine milk fat in winter and summer, and DGAT1 and SCD1 by season interactions. *J. Dairy Sci.* 96:592–604. <https://doi.org/10.3168/jds.2012-5454>.
- Frąszczak, M., and J. Szyda. 2016. Comparison of significant single nucleotide polymorphisms selections in GWAS for complex traits. *J. Appl. Genet.* 57:207–213. <https://doi.org/10.1007/s13353-015-0305-6>.
- Gonzalez-Pena, D., G. Gao, M. Baranski, T. Moen, B. M. Cleveland, P. Brett Kenney, R. L. Vallejo, Y. Palti, and T. D. Leeds. 2016. Genome-wide association study for identifying loci that affect fillet yield, carcass, and body weight traits in rainbow trout (*Oncorhynchus mykiss*). *Front. Genet.* 7:203. <https://doi.org/10.3389/fgene.2016.00203>.
- Goytain, A., and G. A. Quamme. 2005. Functional characterization of human SLC41A1, a Mg²⁺ transporter with similarity to prokaryotic MgtE Mg²⁺ transporters. *Physiol. Genomics* 21:337–342. <https://doi.org/10.1152/physiolgenomics.00261.2004>.
- Habier, D., R. L. Fernando, K. Kizilkaya, and D. J. Garrick. 2011. Extension of the Bayesian alphabet for genomic selection. *BMC Bioinformatics* 12:186. <https://doi.org/10.1186/1471-2105-12-186>.
- Horiba, N., S. Masuda, C. Ohnishi, D. Takeuchi, M. Okuda, and K. I. Inui. 2003. Na⁺-dependent fructose transport via rNaGLT1 in rat kidney. *FEBS Lett.* 546:276–280. [https://doi.org/10.1016/S0014-5793\(03\)00600-8](https://doi.org/10.1016/S0014-5793(03)00600-8).
- Huang, J., G. Luo, Z. Zhang, X. Wang, Z. Ju, C. Qi, Y. Zhang, C. Wang, R. Li, J. Li, W. Yin, Y. Xu, S. J. Moisés, J. J. Looor, and J. Zhong. 2014. ITRAQ-proteomics and bioinformatics analyses of mammary tissue from cows with clinical mastitis due to natural infection with *Staphylococcus aureus*. *BMC Genomics* 15:839. <https://doi.org/10.1186/1471-2164-15-839>.
- Jiang, L., J. Liu, D. Sun, P. Ma, X. Ding, Y. Yu, and Q. Zhang. 2010. Genome wide association studies for milk production traits in Chinese Holstein population. *PLoS One* 5:e13661. <https://doi.org/10.1371/journal.pone.0013661>.
- Jiang, L., X. Liu, J. Yang, H. Wang, J. Jiang, L. Liu, S. He, X. Ding, J. Liu, and Q. Zhang. 2014. Targeted resequencing of GWAS loci reveals novel genetic variants for milk production traits. *BMC Genomics* 15:1105. <https://doi.org/10.1186/1471-2164-15-1105>.
- Kemper, K. E., M. D. Littlejohn, T. Lopdell, B. J. Hayes, L. E. Bennett, R. P. Williams, X. Q. Xu, P. M. Visscher, M. J. Carrick, and M. E. Goddard. 2016. Leveraging genetically simple traits to identify small-effect variants for complex phenotypes. *BMC Genomics* 17:858. <https://doi.org/10.1186/s12864-016-3175-3>.
- Kemper, K. E., C. M. Reich, P. J. Bowman, C. J. Vander Jagt, A. J. Chamberlain, B. A. Mason, B. J. Hayes, and M. E. Goddard. 2015. Improved precision of QTL mapping using a nonlinear Bayesian method in a multi-breed population leads to greater accuracy of across-breed genomic predictions. *Genet. Sel. Evol.* 47:29. <https://doi.org/10.1186/s12711-014-0074-4>.
- Knutsen, T. M., H. G. Olsen, V. Tafintseva, M. Svendsen, A. Kohler, M. P. Kent, and S. Lien. 2018. Unravelling genetic variation underlying de novo-synthesis of bovine milk fatty acids. *Sci. Rep.* 8:2179. <https://doi.org/10.1038/s41598-018-20476-0>.
- Komisarek, J., and M. Kolenda. 2016. The effect of DGAT1 polymorphism on milk production traits in dairy cows depending on environmental temperature. *Turk. J. Vet. Anim. Sci.* 40:251–254. <https://doi.org/10.3906/vet-1508-7>.
- Le Maux, S., S. Bouhallab, L. Giblin, A. Brodtkorb, and T. Croguenec. 2014. Bovine β -lactoglobulin/fatty acid complexes: Binding, structural, and biological properties. *Dairy Sci. Technol.* 94:409–426. <https://doi.org/10.1007/s13594-014-0160-y>.
- Li, X., A. J. Buitenhuis, M. S. Lund, C. Li, D. Sun, Q. Zhang, N. A. Poulsen, and G. Su. 2015. Joint genome-wide association study for milk fatty acid traits in Chinese and Danish Holstein populations. *J. Dairy Sci.* 98:8152–8163. <https://doi.org/10.3168/jds.2015-9383>.
- Littlejohn, M. D., K. Tiplady, T. A. Fink, K. Lehnert, T. Lopdell, T. Johnson, C. Couldrey, M. Keehan, R. G. Sherlock, C. Harland, A. Scott, R. G. Snell, S. R. Davis, and R. J. Spelman. 2016. Sequence-based association analysis reveals an MGST1 eQTL with pleiotropic effects on bovine milk composition. *Sci. Rep.* 6:25376. <https://doi.org/10.1038/srep25376>.
- Lopdell, T. J., K. Tiplady, M. Struchalin, T. J. J. Johnson, M. Keehan, R. Sherlock, C. Couldrey, S. R. Davis, R. G. Snell, R. J. Spelman, and M. D. Littlejohn. 2017. DNA and RNA-sequence based GWAS highlights membrane-transport genes as key modulators of milk lactose content. *BMC Genomics* 18:968. <https://doi.org/10.1186/s12864-017-4320-3>.
- Ma, L., G. R. Wiggans, S. Wang, T. S. Sonstegard, J. Yang, B. A. Crooker, J. B. Cole, C. P. Van Tassel, T. J. Lawlor, and Y. Da. 2012. Effect of sample stratification on dairy GWAS results. *BMC Genomics* 13:536. <https://doi.org/10.1186/1471-2164-13-536>.
- Marques, D. B. D., J. W. M. Bastiaansen, M. L. W. J. Broekhuijse, M. S. Lopes, E. F. Knol, B. Harlizius, S. E. F. Guimarães, F. F. Silva, and P. S. Lopes. 2018. Weighted single-step GWAS and gene network analysis reveal new candidate genes for semen traits in pigs. *Genet. Sel. Evol.* 50:40. <https://doi.org/10.1186/s12711-018-0412-z>.
- Meredith, B. K., D. P. Berry, F. Kearney, E. K. Finlay, A. G. Fahey, D. G. Bradley, and D. J. Lynn. 2013. A genome-wide association study for somatic cell score using the Illumina high-density bovine beadchip identifies several novel QTL potentially related to mastitis susceptibility. *Front. Genet.* 4:229. <https://doi.org/10.3389/fgene.2013.00229>.
- Meredith, B. K., F. J. Kearney, E. K. Finlay, D. G. Bradley, A. G. Fahey, D. P. Berry, and D. J. Lynn. 2012. Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland. *BMC Genet.* 13:21. <https://doi.org/10.1186/1471-2156-13-21>.
- Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. H. Lee. 2002. BLUPF90 and related programs (BGF90). Communication No. 28–07 in Proc. 7th World Congr. Genet. Appl. Livest. Prod., Montpellier, France. INRA (Institut National de la Recherche Agronomique), Cedex, France.
- Nayeri, S., M. Sargolzaei, M. K. Abo-Ismael, N. May, S. P. Miller, F. Schenkel, S. S. Moore, and P. Stothard. 2016. Genome-wide association for milk production and female fertility traits in Canadian dairy Holstein cattle. *BMC Genet.* 17:75. <https://doi.org/10.1186/s12863-016-0386-1>.
- Olsen, H. G., T. M. Knutsen, A. Kohler, M. Svendsen, L. Gidskehaug, H. Grove, T. Nome, M. Sodeland, K. K. Sundaassen, M. P. Kent, H. Martens, and S. Lien. 2017. Genome-wide association mapping for milk fat composition and fine mapping of a QTL for de novo synthesis of milk fatty acids on bovine chromosome 13. *Genet. Sel. Evol.* 49:20. <https://doi.org/10.1186/s12711-017-0294-5>.
- Pan, C. J., S. Y. Chen, H. S. Jun, S. R. Lin, B. C. Mansfield, and J. Y. Chou. 2011. SLC37A1 and SLC37A2 are phosphate-linked, glucose-6-phosphate antiporters. *PLoS One* 6:e23157. <https://doi.org/10.1371/journal.pone.0023157>.
- Pausch, H., R. Emmerling, B. Gredler-Grandl, R. Fries, H. D. Daetwyler, and M. E. Goddard. 2017. Meta-analysis of sequence-based association studies across three cattle breeds reveals 25 QTL for fat and protein percentages in milk at nucleotide resolution. *BMC Genomics* 18:853. <https://doi.org/10.1186/s12864-017-4263-8>.
- Penasa, M., F. Tiezzi, P. Gottardo, M. Cassandro, and M. De Marchi. 2015. Genetics of milk fatty acid groups predicted during routine data recording in Holstein dairy cattle. *Livest. Sci.* 173:9–13. <https://doi.org/10.1016/j.livsci.2014.12.014>.
- Pereira, G. L., L. A. Chardulo, J. A. I. Silva, R. Faria, and R. A. Curi. 2018. Genomic regions associated with performance in racing line of Quarter Horses. *Livest. Sci.* 211:42–51. <https://doi.org/10.1016/j.livsci.2018.02.015>.
- Perland, E., S. V. Hellsten, N. Schweizer, V. Arapi, F. Rezayee, M. Bushra, and R. Fredriksson. 2017. Structural prediction of two

- novel human atypical SLC transporters, MFSD4A and MFSD9, and their neuroanatomical distribution in mice. *PLoS One* 12:e0186325. <https://doi.org/10.1371/journal.pone.0186325>.
- Petrini, J., L. H. S. Iung, M. A. P. Rodriguez, M. Salvian, F. Pértille, G. A. Rovadoski, L. D. Cassoli, L. L. Coutinho, P. F. Machado, G. R. Wiggans, and G. B. Mourão. 2016. Genetic parameters for milk fatty acids, milk yield and quality traits of a Holstein cattle population reared under tropical conditions. *J. Anim. Breed. Genet.* 133:384–395. <https://doi.org/10.1111/jbg.12205>.
- Raven, L. A., B. G. Cocks, and B. J. Hayes. 2014. Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. *BMC Genomics* 15:62. <https://doi.org/10.1186/1471-2164-15-62>.
- Raven, L. A., B. G. Cocks, K. E. Kemper, A. J. Chamberlain, C. J. Vander Jagt, M. E. Goddard, and B. J. Hayes. 2016. Targeted imputation of sequence variants and gene expression profiling identifies twelve candidate genes associated with lactation volume, composition and calving interval in dairy cattle. *Mamm. Genome* 27:81–97. <https://doi.org/10.1007/s00335-015-9613-8>.
- Rodriguez, M. A. P., J. Petrini, E. M. Ferreira, L. R. M. B. Mourão, M. Salvian, L. D. Cassoli, A. V. Pires, P. F. Machado, and G. B. Mourão. 2014. Concordance analysis between estimation methods of milk fatty acid content. *Food Chem.* 156:170–175. <https://doi.org/10.1016/j.foodchem.2014.01.092>.
- Saharinen, J., M. Hyytiäinen, J. Taipale, and J. Keski-Oja. 1999. Latent transforming growth factor- β binding proteins (LTBPs)—structural extracellular matrix proteins for targeting TGF- β action. *Cytokine Growth Factor Rev.* 10:99–117. [https://doi.org/10.1016/S1359-6101\(99\)00010-6](https://doi.org/10.1016/S1359-6101(99)00010-6).
- Samuel, W., C. N. Nagineni, R. Krishnan Kutty, W. Tony Parks, J. S. Gordon, S. M. Prouty, J. J. Hooks, and B. Wiggert. 2002. Transforming growth factor- α regulates stearyl coenzyme A desaturase expression through a Smad signaling pathway. *J. Biol. Chem.* 277:59–66. <https://doi.org/10.1074/jbc.M108730200>.
- Sanchez, M. P., A. Govignon-Gion, P. Croiseau, S. Fritz, C. Hozé, G. Miranda, P. Martin, A. Barbat-Leterrier, R. Letaïef, D. Rocha, M. Brochard, M. Boussaha, and D. Boichard. 2017. Within-breed and multi-breed GWAS on imputed whole-genome sequence variants reveal candidate mutations affecting milk protein composition in dairy cattle. *Genet. Sel. Evol.* 49:68. <https://doi.org/10.1186/s12711-017-0344-z>.
- Schopen, G. C. B., M. H. P. W. Visker, P. D. Koks, E. Mullaart, J. A. M. van Arendonk, and H. Bovenhuis. 2011. Whole-genome association study for milk protein composition in dairy cattle. *J. Dairy Sci.* 94:3148–3158. <https://doi.org/10.3168/jds.2010-4030>.
- Shin, D., S. Y. B. Howng, L. J. Ptáček, and Y. H. Fu. 2012. MiR-32 and its target SLC45A3 regulate the lipid metabolism of oligodendrocytes and myelin. *Neuroscience* 213:29–37. <https://doi.org/10.1016/j.neuroscience.2012.03.054>.
- Song, H., L. Li, P. Ma, S. Zhang, G. Su, M. S. Lund, Q. Zhang, and X. Ding. 2018. Short communication: Improving the accuracy of genomic prediction of body conformation traits in Chinese Holsteins using markers derived from high-density marker panels. *J. Dairy Sci.* 101:5250–5254. <https://doi.org/10.3168/jds.2017-13456>.
- Su, X., S. Min, S. Cao, H. Yan, Y. Zhao, H. Li, L. Chai, S. Mei, J. Yang, Y. Zhang, Z. Zhang, F. Liu, W. Sun, Y. Che, and R. Yang. 2014. LRRC19 expressed in the kidney induces TRAF2/6-mediated signals to prevent infection by uropathogenic bacteria. *Nat. Commun.* 5:4434. <https://doi.org/10.1038/ncomms5434>.
- Suchocki, T., K. Wojdak-Maksymiec, and J. Szyda. 2016. Using gene networks to identify genes and pathways involved in milk production traits in Polish Holstein dairy cattle. *Czech J. Anim. Sci.* 61:526–538. <https://doi.org/10.17221/43>.
- Tao, S., R. M. Orellana, X. Weng, T. N. Marins, G. E. Dahl, and J. K. Bernard. 2018. Symposium review: The influences of heat stress on bovine mammary gland function. *J. Dairy Sci.* 101:5642–5654. <https://doi.org/10.3168/jds.2017-13727>.
- Turner, S.-D. 2014. qqman: an R package for visualizing GWAS results using Q-Q and Manhattan plots. *bioRxiv*. Accessed Jul. 25, 2018. <http://biorxiv.org/content/early/2014/05/14/005165.short>.
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423. <https://doi.org/10.3168/jds.2007-0980>.
- VanRaden, P. M., D. J. Null, M. Sargolzaei, G. R. Wiggans, M. E. Tooker, J. B. Cole, T. S. Sonstegard, E. E. Connor, M. Winters, J. B. C. H. M. van Kaam, A. Valentini, B. J. Van Doormaal, M. A. Faust, and G. A. Doak. 2013. Genomic imputation and evaluation using high-density Holstein genotypes. *J. Dairy Sci.* 96:668–678. <https://doi.org/10.3168/jds.2012-5702>.
- Visscher, P. M. 2008. Sizing up human height variation. *Nat. Genet.* 40:489–490. <https://doi.org/10.1038/ng0508-489>.
- Vitavska, O., and H. Wiczorek. 2013. The SLC45 gene family of putative sugar transporters. *Mol. Aspects Med.* 34:655–660. <https://doi.org/10.1016/j.mam.2012.05.014>.
- Wang, H., I. Misztal, I. Aguilar, A. Legarra, and W. M. Muir. 2012. Genome-wide association mapping including phenotypes from relatives without genotypes. *Genet. Res. (Camb.)* 94:73–83. <https://doi.org/10.1017/S0016672312000274>.
- Wang, X. G., Z. H. Ju, M. H. Hou, Q. Jiang, C. H. Yang, Y. Zhang, Y. Sun, R. L. Li, C. F. Wang, J. F. Zhong, and J. M. Huang. 2016. Deciphering transcriptome and complex alternative splicing transcripts in mammary gland tissues from cows naturally infected with *Staphylococcus aureus* mastitis. *PLoS One* 11:e0159719. <https://doi.org/10.1371/journal.pone.0159719>.
- Yang, S. H., X. J. Bi, Y. Xie, C. Li, S. L. Zhang, Q. Zhang, and D. X. Sun. 2015. Validation of PDE9A gene identified in GWAS showing strong association with milk production traits in Chinese Holstein. *Int. J. Mol. Sci.* 16:26530–26542. <https://doi.org/10.3390/ijms161125976>.
- Zhang, Z., U. Ober, M. Erbe, H. Zhang, N. Gao, J. He, J. Li, and H. Simianer. 2014. Improving the accuracy of whole genome prediction for complex traits using the results of genome wide association studies. *PLoS One* 9:e93017. <https://doi.org/10.1371/journal.pone.0093017>.
- Zhu, B., H. Niu, W. Zhang, Z. Wang, Y. Liang, L. Guan, P. Guo, Y. Chen, L. Zhang, Y. Guo, H. Ni, X. Gao, H. Gao, L. Xu, and J. Li. 2017. Genome wide association study and genomic prediction for fatty acid composition in Chinese Simmental beef cattle using high density SNP array. *BMC Genomics* 18:464. <https://doi.org/10.1186/s12864-017-3847-7>.