



Genome-wide association for metabolic clusters in early-lactation Holstein dairy cows

H. Atashi,^{1,2} M. Salavati,³ J. De Koster,¹ M. A. Crowe,⁴ G. Opsomer,¹ the GplusE consortium, and M. Hostens^{1*}

¹Department of Reproduction, Obstetrics and Herd Health, Ghent University, Merelbeke 9820, Belgium

²Department of Animal Science, Shiraz University, Shiraz 71441-65186, Iran

³The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian EH25 9RG, UK

⁴University College Dublin, 4 Dublin, Ireland

ABSTRACT

The aim of this study was to detect the genomic region or regions associated with metabolic clusters in early-lactation Holstein cows. This study was carried out in 2 experiments. In experiment I, which was carried out on 105 multiparous Holstein cows, animals were classified through k-means clustering on log-transformed and standardized concentrations of blood glucose, insulin-like growth factor I, free fatty acids, and β -hydroxybutyrate at 14 and 35 d in milk (DIM), into metabolic clusters, either balanced (BAL) or other (OTR). Forty percent of the animals were categorized in the BAL group, and the remainder were categorized as OTR. The cows were genotyped for a total of 777,962 SNP. A genome-wide association study was performed, using a case-control approach through the GEMMA software, accounting for population structure. We found 8 SNP (BTA11, BTA23, and BTAX) associated with the predicted metabolic clusters. In experiment II, carried out on 4,267 second-parity Holstein cows, milk samples collected starting from the first week until 50 DIM were used to determine Fourier-transform mid-infrared (FT-MIR) spectra and subsequently to classify the animals into the same metabolic clusters (BAL vs. OTR). Twenty-eight percent of the animals were categorized in the BAL group, and the remainder were classified in the OTR category. Although daily milk yield was lower in BAL cows, we found no difference in daily fat- and protein-corrected milk yield in cows from the BAL metabolic cluster compared with those in the OTR metabolic cluster. In the next step, a single-step genomic BLUP was used to identify the genomic region(s) associated with the predicted meta-

bolic clusters. The results revealed that prediction of metabolic clusters is a highly polygenic trait regulated by many small-sized effects. The region of 36,258 to 36,295 kb on BTA27 was the highly associated region for the predicted metabolic clusters, with the closest genes to this region (*ANK1* and *miR-486*) being related to hematopoiesis, erythropoiesis, and mammary gland development. The heritability for metabolic clustering was 0.17 (SD 0.03), indicating that the use of FT-MIR spectra in milk to predict metabolic clusters in early-lactation across a large number of cows has satisfactory potential to be included in genetic selection programs for modern dairy cows.

Key words: metabolic adaptation, transition period, genome-wide association study, dairy cow

INTRODUCTION

The transition period, defined as the last 3 wk before and the first 3 wk after calving, is critically important to the health, production, and profitability of dairy cows (Drackley, 1999; LeBlanc, 2010). The early-lactation cow needs comprehensive adaptive reactions to cope with the challenges caused by the transition from the pregnant, nonlactating state to the nonpregnant, lactating state (Bell, 1995; Goff and Horst, 1997; Esposito et al., 2014). Dry matter intake starts to decrease a few weeks before parturition, and its lowest level occurs at calving (Bell, 1995; Ingvarstsen and Andersen, 2000). Although most cows rapidly increase DMI in the first weeks after calving (Ingvarstsen and Andersen, 2000), high-yielding dairy cows enter a state of negative energy balance (NEB) around calving, when the energy demand for maintenance and lactation exceeds the dietary intake potential (Bauman and Bruce Currie, 1980; Patton et al., 2006; Alawneh et al., 2012). The NEB results in a mobilization of body reserves to provide energy requirements for milk synthesis (van der Drift et al., 2012). Although lipid mobilization is a

Received July 31, 2019.

Accepted January 22, 2020.

*Corresponding author: miel.hostens@ugent.be

normal and necessary process in metabolic adaptation to support lactation, many health disorders are associated with uncontrolled lipid mobilization (Contreras et al., 2018). Increased levels of metabolic and infectious disorders (mastitis, ketosis, fatty liver, hypocalcemia, and retained placenta) and decreased reproductive performance (anestrus, decreased conception rate, and delayed reproduction) have been linked with longer and more severe periods of NEB (Drackley, 1999; Collard et al., 2000; LeBlanc, 2010).

The ability of the animal to cope with the NEB in the transition period varies considerably between individuals and may be partly explained by the genetic background of animals (Kessel et al., 2008; Liinamo et al., 2012; Spurlock et al., 2012; Pryce et al., 2016). Although some countries have already initiated genetic evaluation of metabolic disorders, opportunities exist to use predictor traits and genomic information to strengthen genetic evaluations for metabolic health in dairy cows. Challenges such as lack of a standard recording system and accurate data collection, and difficulty in diagnosing subclinical cases and determining economic values for metabolic disorders, are among the reasons that interest is growing in using easily measurable predictors of metabolic diseases, either recorded on-farm using sensors and milk tests or off-farm using data collected from routine milk recording (Zwald et al., 2004; Pryce et al., 2016). Clustering early-lactation cows based on different blood or milk metabolites can be an effective way to group individual cows based on their ability to cope with the altered metabolic challenges of lactation (De Koster et al., 2019). This clustering methodology can be used as a diagnostic or herd-management tool but, on a larger scale, can also be used to identify superior animals with high genetic values for targeted breeding purposes. Metabolic imbalance is defined as “a condition where the regulatory mechanisms are insufficient for the animals to function optimally leading to a high risk of digestive, metabolic and infectious problems” (Ingvarsen, 2006, p. 176). Elevated levels of free fatty acids (**FFA**) and BHB, and decreased levels of glucose and IGF-I, are reported as blood biomarkers of metabolically imbalanced cows (Ingvarsen et al., 2003; Puppel and Kuczynska, 2016).

Although physiological adaptations of dairy cows in the transition period are well described, knowledge about the genetic background remains insufficient (Bell and Bauman, 1997; Drackley et al., 2005). Genome-wide association studies provide a fine strategy for mapping the underlying genetic backgrounds of complex traits in humans, animals, and plants (Zhang et al., 2012; Wang et al., 2015a). The aim of the present study was to use GWAS to identify genetic markers linked with

metabolic clusters in early-lactation Holstein cows. To predict metabolic clusters of animals, we used blood metabolites (experiment I) and Fourier-transform mid-infrared (**FT-MIR**) spectra of milk (experiment II).

MATERIALS AND METHODS

Experiment I

Animals and Phenotype. The data in experiment I were collected as part of the Genotype plus Environment (**GplusE**) FP7-Project (<http://www.gpluse.eu>). Detailed descriptions of the experiments, laboratory analysis, phenotypic data, and k-means clustering methodology are provided in De Koster et al. (2019). In short, a total of 105 multiparous Holstein dairy cows from 4 research herds (Aarhus University, Denmark; UCD Lyons Research farm, University College Dublin, Ireland; Agri-Food and Biosciences Institute, Northern Ireland, UK; and Leibniz Institute for Farm Animal Biology, Germany) were combined. Glucose, IGF-1, FFA, and BHB concentrations were determined in blood samples taken around 14 and 35 DIM, as described in Krogh et al. (2019). K-means clustering based on the log-transformed and standardized blood metabolite concentrations (mean = 0 and SD = 1) was used to describe the metabolic profiles of the included cows. The cluster of cows with high concentrations of glucose and IGF-I and low concentrations of FFA and BHB at 14 and 35 DIM was defined as a balanced (**BAL**) metabolic cluster and was compared with the “other” (**OTR**) category. The mean glucose (mM), IGF-I (ng/mL), FFA (mM), and BHB (mM) values for BAL cows at 14 DIM were 3.62, 121.66, 0.52, and 0.44, respectively. The corresponding values for cows belonging to the OTR cluster were 3.28, 49.63, 0.81, and 0.71, respectively.

Genome-Wide Association Study. All cows were genotyped using Illumina BovineHD BeadChip (San Diego, CA) for a total of 777,962 SNP. Quality control for markers was performed using PLINK (Purcell et al., 2007). Those SNP with no position information, call rate less than 95%, minor allele frequency less than 5%, or Hardy-Weinberg equilibrium P -value less than 6.4×10^{-8} were excluded. Single-SNP association analysis was performed using a mixed-model approach in GEMMA software (Zhou and Stephens, 2012), which implements the genome-wide efficient mixed-model association algorithm and accounts for population stratification. The centered relatedness matrix was calculated from all genotypes. The association test was then performed with phenotype, genotype, and centered relatedness matrix files, using a univariate linear mixed model.

Each SNP was fitted as a covariate, and the likelihood ratio test was used for each SNP against the null hypothesis of $g = 0$, using the following statistical model:

$$\mathbf{y} = \mathbf{W}\boldsymbol{\alpha} + \mathbf{x}\beta + \mathbf{u} + \boldsymbol{\varepsilon},$$

where \mathbf{y} is an $n \times 1$ vector of phenotype values for all individuals (the predicted metabolic clusters and concentrations of BHB, glucose, IGF-1, and FFA); \mathbf{W} is an $n \times c$ matrix of covariates (fixed effects containing herd, parity, and mean); $\boldsymbol{\alpha}$ is a $c \times 1$ vector of the corresponding coefficients, including the intercept; \mathbf{x} is an $n \times 1$ vector of genotypes of a marker at the locus tested; β is the effect size of the marker; \mathbf{u} is a $n \times 1$ vector of random polygenic effects with a covariance structure as $\mathbf{u} \sim N(0, \mathbf{K}\mathbf{V}\mathbf{g})$, where \mathbf{K} is an $n \times n$ marker-based additive genetic relatedness matrix and $\mathbf{V}\mathbf{g}$ is the polygenic additive variance; and $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of residual errors with $\boldsymbol{\varepsilon} \sim N(0, \mathbf{I}\mathbf{V}\mathbf{e})$, where \mathbf{I} is an $n \times n$ identity matrix and $\mathbf{V}\mathbf{e}$ is the residual variance. The thresholds of the Bonferroni-corrected P -values for suggestive and 5% genome-wide significance associations were set as 1.74×10^{-6} (1 divided by the number of SNP) and 8.68×10^{-6} (0.05 divided by the number of SNP), respectively (Lander and Kruglyak, 1995).

Experiment II

Animals and Phenotype. Data in experiment II were collected as part of Work Package 4 from the Gpluse FP7-Project (<http://www.gpluse.eu>). This experiment was performed on 4,267 second-parity Holstein cows with calving year 2015 to 2018, distributed over 50 herds in Belgium and the Netherlands. De Koster et al. (2019) compared different models to predict the metabolic clusters of early-lactation dairy cows using milk metabolites or FT-MIR spectra of milk and reported that the use of FT-MIR spectra had a high accuracy in predicting metabolic clusters of individual cows. A detailed description of the procedure is provided in De Koster et al. (2019). In short, milk samples were collected starting from the first week in milk until 50 DIM, and FT-MIR spectra of samples were determined. Then, FT-MIR spectra were used to divide animals into either the BAL (cows with favorable metabolic profiles; $n = 1,201$, 28%) or the OTR group ($n = 3,066$, 72%).

Genotypic Data. Individuals ($n = 31,895$) were genotyped using the BovineLD ($n = 20,462$), Bovine SNP50K ($n = 10,638$), or BovineHD SNP panels (795 animals; all panels from Illumina). Genotypes of ani-

mals were imputed to HD with a reference population of 795 HD individuals (46 males and 749 females) using FImpute software, version 2.2 (Sargolzaei et al., 2014). In total 12,367 out of 31,895 genotyped individuals either had phenotypic data or were in the pedigree file and were used in the subsequent association analysis. (The number of animals with records was 4,267; the number of animals with records and with genotypic data was 3,725; the number of animals with records and no genotypes was 542; and the number of animals with genotypes and no records was 8,642.) Quality control was performed on the imputed data, and SNP with minor allele frequency less than 5% were excluded. After genomic data quality control, 566,345 out of 730,539 SNP were available for association analysis.

Association Between Predicted Metabolic Clusters and Production Performance. The association of 305-d milk yield and the predicted metabolic clusters was determined using a linear mixed model through inclusion of herd-year-season of calving (HYS), metabolic clusters (BAL and OTR), covariate effects of age at first calving (FCA) in both linear and quadratic forms, and random effect of dam's sire. The association of the predicted metabolic clusters with production traits [including daily milk yield, daily fat- and protein-corrected milk (FPCM) yield, milk fat percentage, milk protein percentage, daily fat yield, daily protein yield, daily fat and protein yield, fat-to-protein ratio, and milk SCS during the entire lactation] was investigated using linear mixed models constructed with week postcalving as a repeated observation within the random factor of cow. Interaction effects were removed from the model if nonsignificant ($P > 0.05$). Multiple testing correction was performed using Tukey's post hoc test, and significance was declared at $P < 0.05$ (R Core Team, 2017).

Variance Component Estimation. Pedigree information was collected for all phenotyped animals and contained a total of 43,181 individuals (12,367 and 4,267 out of 43,181 animals had genotype and phenotype data, respectively). The variance components were estimated by Bayesian inference, considering a linear single trait animal model. The linear model included HYS as systematic effect, covariate effects of FCA in both linear and quadratic forms, DIM and milk yield at sampling as well as animal direct genetic effect and residual effect as random effects. The complete model can be represented as follows:

$$y_{ij} = \mu + \text{HYS}_i + b_1(\text{FCA}_j) + b_2(\text{FCA}_j)^2 + b_3(\text{DIM}_j) + b_4(\text{milk}_j) + a_j + e_{ij},$$

where y_{ij} represents the response variable of animal j (BAL vs. OTR metabolic cluster); μ is the overall mean; HYS_i is the fixed effect of i th herd-year-season of calving; b_1 and b_2 are the linear and quadratic regression coefficients of the dependent variable on the age at first calving; FCA_j is the age at first calving of j th cow; b_3 and b_4 are the linear regression coefficients of the dependent variable on DIM and milk yield at sampling; DIM_j and $milk_j$ are, respectively, DIM and milk yield at sampling on j th cow; a_j is the additive genetic effect; and e_{ij} is the random residual error. The additive genetic and residual variances were obtained as follows:

$$\text{var} \begin{bmatrix} \mathbf{a} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{H}\sigma_a^2 & 0 \\ 0 & \mathbf{I}\sigma_e^2 \end{bmatrix},$$

where σ_a^2 and σ_e^2 are, respectively, total additive genetic and residual variances, \mathbf{a} is the vector of direct additive genetic effects, \mathbf{e} is a vector of residual effects, \mathbf{I} is the identity matrix, and \mathbf{H} is a matrix that combines pedigree and genomic relationships, and its inverse consists of the integration of the additive and genomic relationship matrices \mathbf{A} and \mathbf{G} , respectively (Aguilar et al., 2010):

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

where \mathbf{A} is the numerator relationship matrix based on pedigree for all animals; \mathbf{A}_{22} is the numerator relationship matrix for genotyped animals; and \mathbf{G} is the genomic relationship matrix which was obtained using following function described by VanRaden (2008):

$$\mathbf{G} = \frac{\mathbf{ZDZ}'}{\sum_{i=1}^M 2p_i(1-p_i)},$$

where \mathbf{Z} is a matrix of gene content adjusted for allele frequencies (0, 1, or 2 for aa , Aa , and AA , respectively); \mathbf{D} is a diagonal matrix of weights for SNP variances (initially $\mathbf{D} = \mathbf{I}$); M is the number of SNP, and p_i is the minor allele frequency of the i th SNP. The \mathbf{H} matrix was built scaling \mathbf{G} based on \mathbf{A}_{22} , considering that the average of diagonal of \mathbf{G} is equal to average of the diagonal of \mathbf{A}_{22} , and the average of off-diagonal \mathbf{G} is equal to average off-diagonal \mathbf{A}_{22} . The analysis, consisting of a single chain of 350,000 cycles with a burn-in of 100,000 cycles, taking a sample every 50 iterations, was performed using THRGIBBS1F90 (Misztal et al., 2002). Chain convergence was assessed by visual inspection. The posterior estimates were obtained retrospec-

tively, using the POSTGIBBSF90 program (Misztal et al., 2002).

Single-Step GWAS. Analyses were performed using the weighted single-step GWAS methodology (Wang et al., 2012), considering the same linear animal model used to estimate the variance or covariance components previously mentioned. The animal effects were decomposed into those for genotyped (\mathbf{a}_g) and ungenotyped animals (\mathbf{a}_n). The animal effects of genotyped animals are a function of SNP effects, $\mathbf{a}_g = \mathbf{Z}\mathbf{u}$, where \mathbf{Z} is a matrix relating genotypes of each locus and \mathbf{u} is a vector of SNP marker effect. The variance of animal effects was assumed thus:

$$\text{Var}(\mathbf{a}_g) = \text{Var}(\mathbf{Z}\mathbf{u}) = \mathbf{ZDZ}'\sigma_u^2 = \mathbf{G}^*\sigma_a^2,$$

where \mathbf{D} is a diagonal matrix of weights for variances of markers ($\mathbf{D} = \mathbf{I}$ for GBLUP) and σ_u^2 is the genetic additive variance captured by each SNP marker when the weighted relationship matrix (\mathbf{G}^*) was built with no weight.

The SNP effects were obtained using the following equation:

$$\hat{\mathbf{u}} = \lambda \mathbf{DZ}' \mathbf{G}^{*-1} \hat{\mathbf{a}}_g = \mathbf{DZ}' (\mathbf{ZDZ}')^{-1} \hat{\mathbf{a}}_g,$$

where λ was defined by VanRaden (2008) as a normalizing constant, described as

$$\lambda = \frac{\sigma_u^2}{\sigma_a^2} = \frac{1}{\sum_{i=1}^M 2p_i(1-p_i)},$$

for which M is the number of SNP and p_i is the frequency of the second allele in the i th SNP. The following iterative process described by Wang et al. (2012) was used to estimate the SNP effects: (1) $\mathbf{D} = \mathbf{I}$ in the first step; (2) calculate \mathbf{G} matrix; (3) calculate genetic (G)EBV for the entire data set using single-step GBLUP; (4) convert GEBV to SNP effects ($\hat{\mathbf{u}}$), $\hat{\mathbf{u}} = \lambda \mathbf{DZ}' \mathbf{G}^{*-1} \hat{\mathbf{a}}_g$; (5) calculate the variance of each SNP, $\mathbf{d}_i = \hat{\mathbf{u}}_i^2 2p_i(1-p_i)$, where i is the i th SNP; (6) normalize SNP weights to remain the total genetic variance constant; exit or loop to step 2. The effects of markers were obtained by 2 iterations from step 2 to 6, as proposed by Wang et al. (2012). The percentage of genetic variance explained by the i th region has been calculated as follows:

$$\frac{\text{Var}(\hat{a}_i)}{\sigma_a^2} \times 100\% = \frac{\text{Var}\left(\sum_{j=1}^n \mathbf{Z}_j \hat{\mathbf{u}}_j\right)}{\sigma_a^2} \times 100,$$

where a_i is the genetic value of the i th region that consists of n consecutive SNP ($n = 1, 5, 10, 15, 20,$ and 25), σ_a^2 is the total genetic variance, \mathbf{Z}_j is the vector of SNP content of the j th SNP for all individuals, and \hat{u}_j is the marker effect of the j th SNP within the i th region. The results were presented by the proportion of variance explained by each window of n consecutive SNP.

RESULTS AND DISCUSSION

Experiment I

The cows ($n = 105$) were divided into either BAL ($n = 42, 40\%$) or OTR ($n = 63, 60\%$) metabolic clusters using k -means clustering based on concentrations of blood glucose, IGF-1, BHB, and FFA at 14 and 35 DIM. The cows were genotyped for a total of 777,962 SNP, of which 576,092 passed all quality control measures and were used for the GWAS analysis.

Manhattan and Q-Q plots of SNP associated with the predicted metabolic clusters, along with those for the blood metabolites (concentrations of BHB, glucose, IGF-1, and FFA) used to create the clusters, are presented in Figure 1. Although no SNP were associated with the predicted metabolic clusters at the 5% genome-wide significance threshold, 8 [BTA11 ($n = 2$), BTA23 ($n = 2$), BTAX ($n = 4$)] were associated at the suggestive genome-wide significance threshold. The identified SNP in BTA11 and BTAX were also associated with IGF-1 at the suggestive genome-wide significance threshold (Figure 1). The identified SNP, along with their position and 100-kb flanking genes, are presented in Table 1 (assembly UMD 3.1, annotation release 103; https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Bos_taurus/103/). The SNP identified on BTA11 were mapped inside QTL for IgG level (Maltecca et al., 2009), body weight (McClure et al., 2010), and milk α -lactalbumin percentage (Zhou et al., 2019). The immune response profile in transition dairy cows has been addressed in several studies (Herr et al., 2011; Sordillo, 2016). Herr et al. (2011) reported a significant decrease of the serum IgG level beginning at wk 8 before parturition and recovering by wk 4 after parturition. The SNP BovineHD1100022178 on BTA11 was mapped inside the *MIR2285AF-1* gene, which is a microRNA (miRNA). It has been reported that miRNA have important effects on mammary gland development (Piao and Ma, 2012). The SNP detected on BTA23, mapped inside genes including *LOC101906324*, *NEDD9*, *SMIM13*, and *LOC104969831*, overlap with QTL for milk fatty acid content (Bouwman et al., 2012), body weight (Michenet et al., 2016), and DMI (Sherman et al., 2009). Linkage disequilibrium analysis

was performed between 26 SNP (on BTAX position 17.2 to 17.3 Mb), including 4 associated with the predicted metabolic clusters. The estimated pair-wise r^2 values showed that the SNP identified as associated with the predicted metabolic clusters are highly correlated (Figure 2).

In conclusion, the aim of this experiment was to use GWAS to identify genetic markers linked with metabolic clusters in early-lactation Holstein cows. However, it seems that a larger sample size is needed for detecting significant SNP in a GWAS. Although k -means clustering based on blood metabolites can be considered an interesting method to predict metabolic clusters of transition dairy cows, the application of this method on a large sample size to perform an accurate GWAS analysis would be very challenging and expensive. De Koster et al. (2019) suggested a method of clustering early-lactation cows based on FT-MIR spectra of milk to predict metabolic clusters of animals, which was used in experiment II.

Experiment II

In this experiment, milk samples were used to determine FT-MIR spectra and subsequently predict the metabolic clusters for each animal. The cows ($n = 4,267$) were grouped into either BAL ($n = 1,201, 28\%$) or OTR ($n = 3,066, 72\%$) metabolic clusters. De Koster et al. (2019) reported that the use of FT-MIR spectra had a high accuracy (80%) to predict metabolic clusters of individual cows. Therefore, BAL cows were described as having higher concentrations of blood glucose and IGF-I and lower concentrations of blood FFA and BHB compared with OTR cows. We found a significant association between 305-d milk yield and the metabolic clusters ($P < 0.05$). The average 305-d milk yields (\pm SE) for cows in the BAL metabolic cluster versus those in OTR were 9,622 (± 51.8) and 9,840 (± 36.2) kg, respectively. Daily milk and protein yields were lower in BAL cows compared with OTR. Daily fat yield, daily fat and protein yield, fat-to-protein ratio, milk fat percentage, and milk protein percentage were higher in BAL cows than OTR; however, no difference occurred for milk SCS in BAL versus OTR cows (Figure 3; Table 2). Although we discovered no difference for daily FPCM during the entire lactation, FPCM was higher during early lactation in BAL cows compared with OTR (Table 2; Figure 3). This contrasts with the original clustering study (De Koster et al., 2019), which reported no difference in FPCM between BAL cows ($n = 43$; mean = 38.42 kg) and OTR cows ($n = 64$; mean = 37.76 kg). van Hoeij et al. (2019) reported that better metabolic status in dairy cows in early lactation,

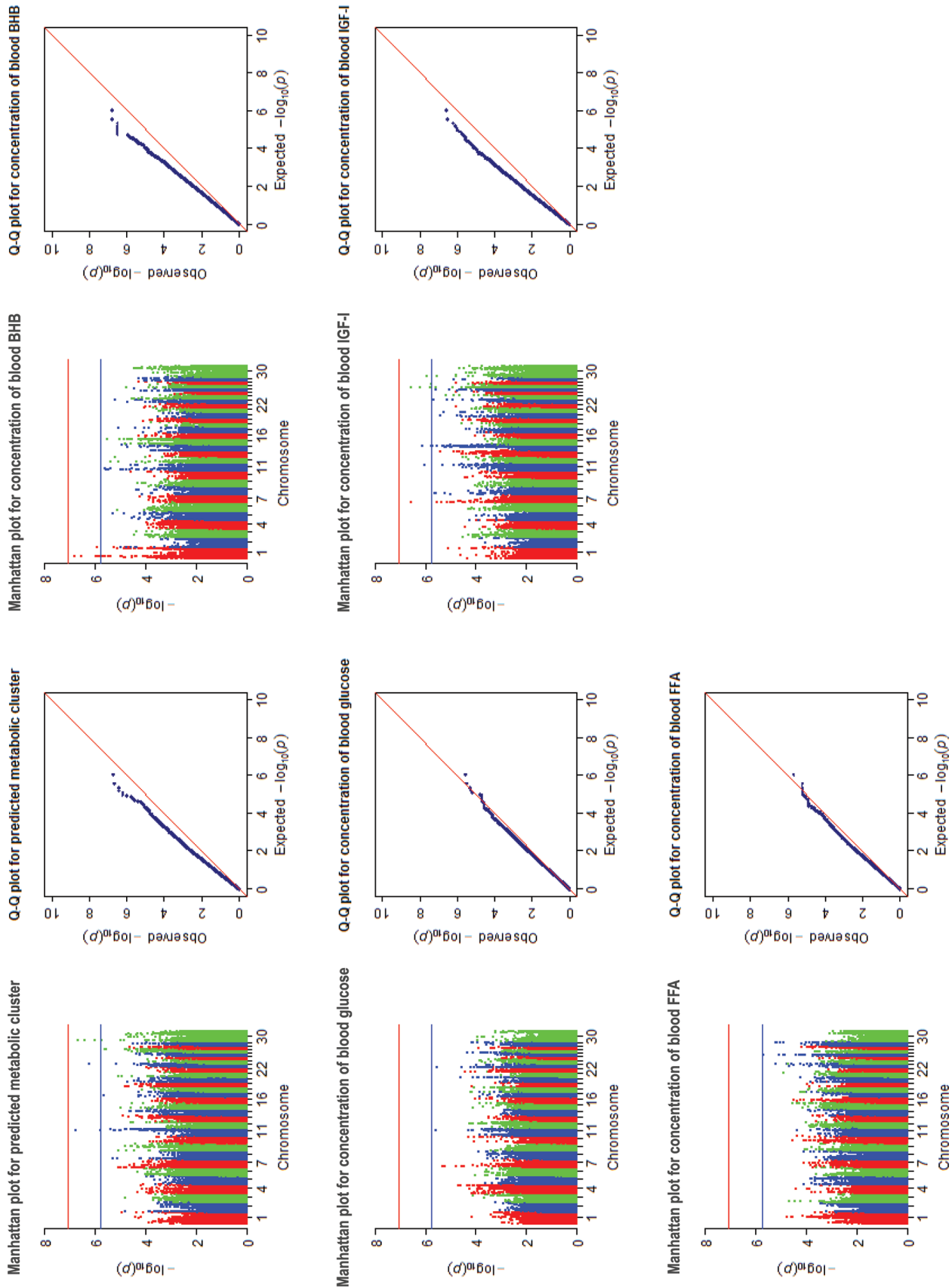


Figure 1. Manhattan and Q-Q plots of SNP associated with the predicted metabolic clusters, along with those for the metabolites [blood concentrations of BHB, glucose, IGF-1, and free fatty acids (FFA)] used to create the clusters. Horizontal red and blue lines in the Manhattan plots indicate the genome-wide significance threshold $[-\log_{10}(8.7 \times 10^{-8})]$ and suggestive significance threshold $[-\log_{10}(1.74 \times 10^{-6})]$. Metabolic clusters were created using k-means clustering based on log-transformed and standardized concentrations of glucose, IGF-1, FFA, and BHB in blood samples taken from 105 multiparous Holstein dairy cows around 14 and 35 DIM (experiment I).

Table 1. Experiment I: SNP detected by genome-wide association study (GWAS) for balanced or imbalanced (BAL vs. OTR) metabolic clusters¹

SNP name	BTA no.	Chromosome position	P-value ²	MAF ³	100-kb flanking genes ⁴
BovineHD1100021914	11	76586236	1.8×10^{-7}	0.367 (TC)	—
BovineHD1100022178	11	77436574	1.6×10^{-7}	0.371 (TC)	MIR2285AF-1 (27,582), LOC104979434 (inside), LOC785416 (56,419)
BovineHD2300013006	23	44852748	5.9×10^{-7}	0.371 (GA)	LOC101906324 (88,276), NEDD9 (inside)
BovineHD2300013034	23	44937247	5.9×10^{-7}	0.371 (TC)	SMIM13 (83,227), LOC104969831 (51,090), NEDD9 (inside)
BovineHD3000005706	30	17232546	2.1×10^{-7}	0.252 (GA)	MIR6526-3 (19,914), GPC4 (37,007), GPC3 (72,982)
BovineHD3000005716	30	17257114	3.8×10^{-7}	0.252 (AG)	MIR6526-3 (44,482), GPC4 (61,575), GPC3 (48,414)
BovineHD3000005723	30	17277730	3.8×10^{-7}	0.371 (AG)	MIR6526-3 (65,098), GPC4 (82,191), GPC3 (27,798)
BovineHD3000005726	30	17216072	9.6×10^{-7}	0.257 (GT)	MIR6526-3 (3,440), GPC4 (20,533), GPC3 (89,456)

¹BAL and OTR metabolic clusters were created using k-means clustering based on log-transformed and standardized concentrations of glucose, IGF-1, free fatty acids and BHB in blood samples taken from 105 multiparous Holstein dairy cows around 14 and 35 DIM.

²P-value of the mixed model for single-SNP association analysis.

³Minor allele frequency and 2 alleles for the SNP.

⁴Genes in 100-kb flanking regions of SNP position. Numbers in parentheses represent the physical distance (in base pairs) between the SNP and the corresponding genes. Official gene symbol (Assembly UMD 3.1, annotation release 103; https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Bos_taurus/103/).

as indicated by plasma metabolites, is associated with higher DMI and energy balance. Beerda et al. (2004) reported no significant difference for the adaptive capacity in high-producing compared with low-producing cows.

The posterior mean of heritability estimates for predicted metabolic clusters was 0.17 (SD 0.03), indicating that the genetic variation for the predicted metabolic clusters is large enough for genetic selection. The heritabilities of DMI and energy balance, respectively, over the entire period of lactation in Nordic Red dairy cattle have been reported to be 0.23 and 0.10 (Liinamo et al., 2012). Spurlock et al. (2012) reported heritability estimates ranging from 0.27 to 0.63 and 0.12 to 0.49 for DMI and energy balance, respectively, in Holstein cows.

In the present study, a window-based GWAS was used to identify genetic markers linked with metabolic clusters predicted using FT-MIR of milk in early-lactating Holstein dairy cows. The absence of a universal approach for hypothesis testing is an important challenge of window-based GWAS (Aguilar et al., 2019), even though it is quite a common procedure in genetic studies. Window-based GWAS may use different window types (distinct or sliding windows) and variable window sizes (defined as the number of SNP or the number of base pairs). The common form for declaring significance is to use a threshold on the additive genetic variance explained by individual windows (Aguilar et al., 2019). However, it is unclear what window size is optimal, and no standard presently exists to define the threshold on explained genetic variance (Aguilar et al., 2019). Therefore, determining the proper window size is usually subjective, and researchers often have not justified their choices or sometimes have acknowledged that their choices are arbitrary (Myles et al., 2008; Beissinger et al., 2015). Han and Peñagaricano (2016) considered 1.5-Mb SNP windows that explained more than 0.50% of the genetic variance as the threshold to declare significance. Suwannasing et al. (2018), using Porcine SNP60k BeadChip (Illumina), considered 5-adjacent SNP windows that explained more than 1% of the total genetic variance as the threshold to declare significance. Medeiros de Oliveira Silva et al. (2017), using the BovineHD SNP panel (Illumina), considered 50-adjacent SNP windows (with an average of 280 kb) that explained more than 0.5% of additive genetic variance as the threshold to declare significance. Fragomeni et al. (2014) reported that small and large window sizes are accompanied, respectively, by large noise and absence of peaks. Beissinger et al. (2015) reported that small sliding windows had the most favorable ratio of detection rate to false-positive rate compared with large window sizes. In the present study, sliding windows of 1, 5, 10, 15, 20, and 25 consecutive SNP were used to

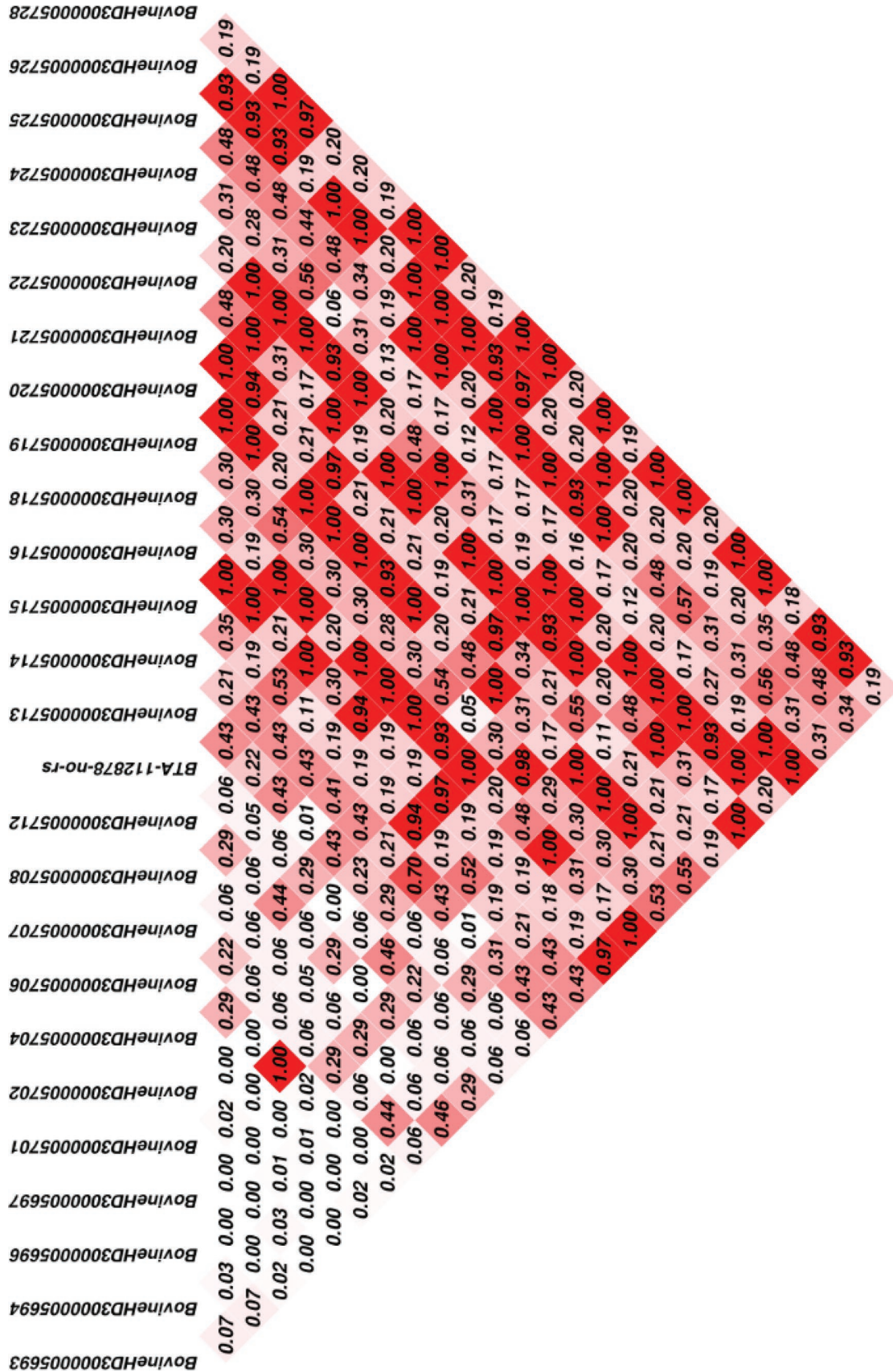


Figure 2. Linkage disequilibrium between 26 SNP (on BTAX position 17.2 to 17.3 Mb), including 4 SNP associated with predicted metabolic clusters. Color scale ranges from red to white; color intensity decreases with decreasing r^2 value (experiment I).

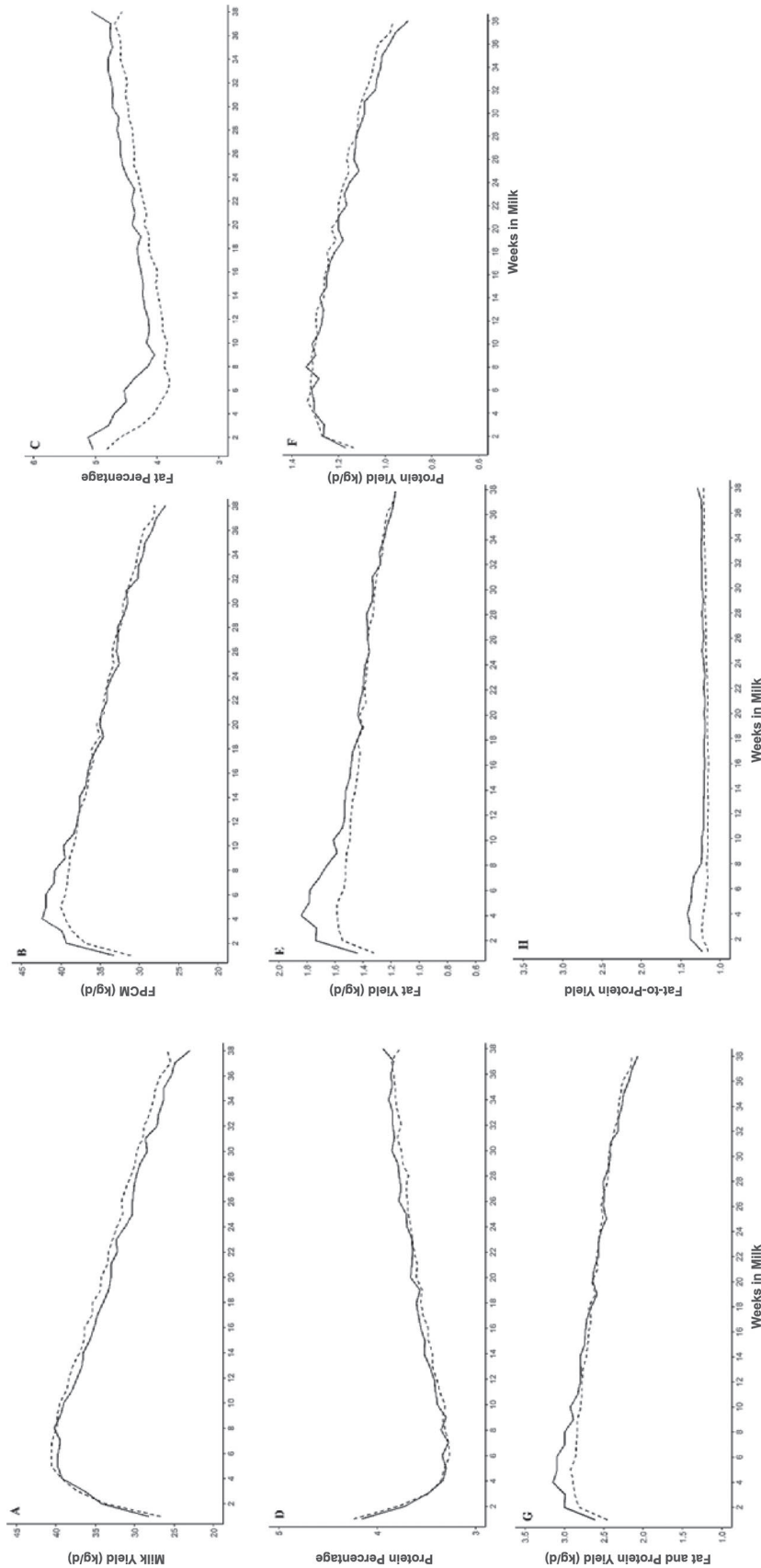


Figure 3. Daily milk yield (A), daily fat- and protein-corrected milk (FPCM) yield (B), milk fat percentage (C), milk protein percentage (D), daily fat yield (E), daily protein yield (F), daily fat and protein yield (G), and fat-to-protein ratio (H) in cows with a balanced metabolic cluster (BAL, solid line) and those with an imbalanced metabolic cluster (OTR, dashed line). Milk samples collected from the first week in milk until 50 DIM were used to determine Fourier-transform mid-infrared spectra to classify the animals in identical metabolic clusters, BAL or OTR (experiment II).

Table 2. Association of predicted metabolic clusters¹ (balanced or imbalanced, BAL vs. OTR) with daily milk yield, fat- and protein-corrected milk (FPCM) yield, fat percentage, protein percentage, fat yield, protein yield, fat and protein yield, fat-to-protein ratio, and milk SCS during entire period of lactation

Item	BAL	OTR
Milk yield (kg/d) LSM	32.8 (1.89) ^b	33.5 (1.89) ^a
FPCM (kg/d) LSM	35.1 (1.70) ^a	34.7 (1.69) ^a
Fat percentage (%)	4.51 (0.17) ^a	4.24 (0.17) ^b
Protein percentage (%)	3.64 (0.09) ^a	3.61 (0.09) ^b
Fat yield (kg/d) LSM	1.46 (0.07) ^a	1.40 (0.07) ^b
Protein yield (kg/d) LSM	1.18 (0.06) ^b	1.19 (0.06) ^a
Fat + protein yield (kg/d) LSM	2.64 (0.12) ^a	2.59 (0.12) ^b
Fat-to-protein ratio	1.24 (0.03) ^a	1.18 (0.03) ^b
Milk SCS	3.72(0.18) ^a	3.73 (0.18) ^a

^{a,b}LSM with different superscripts differ ($P < 0.05$).

¹Milk samples collected starting from the first week in milk until 50 DIM were used to determine Fourier-transform mid-infrared spectra and subsequently used to predict metabolic clusters of animals (BAL vs. OTR). Values in parentheses are SE.

identify genetic markers associated with the metabolic clusters in early-lactation Holstein dairy cows and to determine whether the regions identified might change depending on the window size (Figure 4). The results of single-SNP windows showed small individual SNP variances for the predicted metabolic clusters (with mean and SD of 1.3×10^{-4} and 6.6×10^{-8} , respectively), with just 16 individual SNP explaining more than 0.005% of the additive genetic variance (35 times more than the mean). The top 16 individual SNP combined explained more than 0.09% of the total additive variance and were mapped on BTA27 in position 36,258 to 36,295 kb (a region with length <38,000 bp). The mean genetic variance explained by an individual 5-adjacent SNP window was 8.4×10^{-4} . The first and second top 5-adjacent SNP windows explained more than 0.04 and 0.03% of the genetic variance and were mapped on BTA27 in position 36,265 to 36,271, and 36,259 to 36,264 kb, respectively. The first and second top 10-adjacent SNP windows explained more than 0.07 and 0.06% of the genetic variance and were mapped on BTA27 in position 36,258 to 36,268, and BTA20 in position 54,493 to 54,508 kb, respectively. The first and second top 15-adjacent SNP windows explained more than 0.11 and 0.09% of the genetic variance and were mapped on BTA27 in position 36,258 to 36,280, and BTA20 in position 54,485 to 54,207 kb, respectively. The first top 20-adjacent SNP window explained more than 0.15% of the total additive genetic variance and was mapped on BTA27 in position 36,258 to 36,290 kb. The second top 20-adjacent SNP window was mapped on BTA20 in position 54,470 to 54,520 kb and explained approximately 0.09% of the total additive genetic variance. The first top 25-adjacent SNP window explained more than

0.15% of the additive genetic variance and was mapped on BTA27 in position 36,258 to 36,296 kb (a region with length <38,000 bp). All the top 16 individual SNP identified in the single-SNP analysis were presented in this window. The second top 25-adjacent SNP window was mapped on BTA7 in position 64,537 to 64,568 kb and explained 0.11% of additive genetic variance. The results of different sliding window sizes showed that the region identified on BTA27 in position 36,258 to 36,295 kb overlaps all windows sizes considered and can be suggested to be associated with the predicted metabolic clusters in early-lactation Holstein cows. The identified region overlaps with QTL regions reported for meat quality (Horodyska et al., 2015), milk fatty acid profiles (Bouman et al., 2011), and DMI in cattle (Tetens et al., 2014). The identified region was mapped inside the *ANK1* and *miR-486* genes. The association between the *ANK1* and *miR-486* genes and the predicted metabolic clusters may be partly explained by the regulating roles of these genes in hematopoiesis (Sureshchandra et al., 2016), erythropoiesis (Wang et al., 2015b), and mammary gland development (Li et al., 2015). The *ANK1* gene is a large (~240 kb) gene that encodes the adapter protein ankyrin-1. The ankyrins act as adaptors among a variety of integral membrane proteins and the spectrin skeleton (Bennett and Baines, 2001), and play important roles in cell motility, proliferation, and activation (Nelson and Lazarides, 1984). Ankyrin1, the prototype of the ankyrins, has been reported to be a hematopoiesis-specific regulator (Sureshchandra et al., 2016). In humans, genetic variants located within *ANK1* have been reported to be associated with glycemic traits, impaired insulin release, and type 2 diabetes (Imamura et al., 2012; Diabetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, 2012). In cattle, the *ANK1* gene reportedly been associated with meat quality, milk protein, and rump conformation (Kolbehdari et al., 2008; Horodyska et al., 2015; Cai et al., 2018). Tetens et al. (2014) reported a QTL for DMI, as a correlated trait with metabolic adaptation in dairy cows, mapped on *ANK1*, and SNP in the *ANK1* promoter have also been reported to be associated with intramuscular fat in bovine and porcine tissues (Aslan et al., 2010, 2012). The miRNA, such as *miR-486*, are small (~22 nucleotides) noncoding RNA that regulate many fundamental biological processes primarily by affecting both the stability and the translation of mRNA (Bartel, 2004). In particular, *miR-486* is located within the last intron of *ANK1* that is common to all transcripts and could be cotranscribed with any or all variants (Tessema et al., 2017). Evidence indicates potential cotranscription and coregulation of *ANK1* and *miR-486* (Gallagher et al., 2000; Shaham et al., 2015;

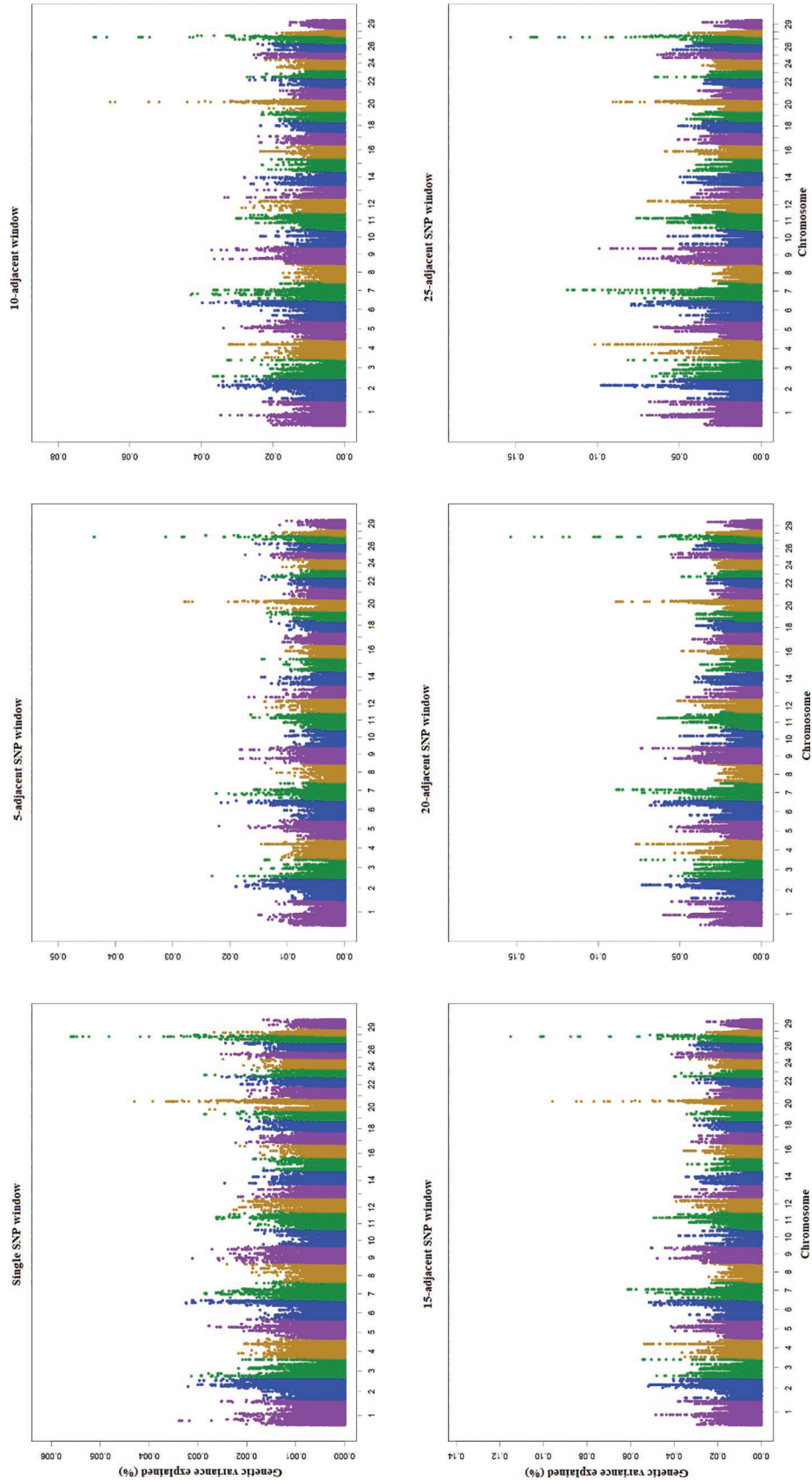


Figure 4. Additive genetic variance explained by windows of 1-, 5-, 10-, 15-, 20-, and 25-adjacent SNP across chromosomes for metabolic adaptations in early-lactation Holstein cows (experiment II).

Tessema et al., 2017). In humans, biological links have been reported between the *miR-486* locus and metabolic disorders, including type 2 diabetes and obesity (Valsesia et al., 2019). A significant association between residual feed intake and the expression level of *miR-486* has been reported in cattle (De Oliveira et al., 2018) and pigs (Jing et al., 2015). Mammary gland development is controlled by multiple genes, including miRNA (Piao and Ma, 2012). Li et al. (2015) reported that *miR-486* is expressed in both bovine mammary gland tissues and in mammary epithelial cells and regulates lactation. The same study further identified *miR-486* as a downstream regulator of PTEN (phosphatase and tensin homolog) that is required for the development of the bovine mammary gland (Li et al., 2015).

Relevant information supporting the results not presented here is provided in supplemental files (Excel files S1 to S6; <https://github.com/Bovi-analytics/Atashi-et-al.2020>).

CONCLUSIONS

The aim of the present study was to detect the genomic region or regions associated with metabolic clusters in early-lactation Holstein cows. Blood metabolites (experiment I) and FT-MIR spectra of milk (experiment II) were used to predict metabolic clusters (BAL vs. OTR) for, respectively, 105 and 4,267 Holstein dairy cows. The results of experiment I revealed 8 SNP (on BTA11, BTA23, and BTAX) associated with the predicted metabolic clusters at the suggestive genome-wide significance threshold. Considering the results of experiment II, it can be concluded that metabolic clusters predicted based on FT-MIR spectra in milk samples collected during early lactation are highly polygenic traits regulated by many small-sized effects. The region of 36,258 to 36,295 kb on BTA27 was the most highly associated region for the predicted metabolic clusters, with the closest genes to this region (*ANK1* and *miR-486*) being related to hematopoiesis, erythropoiesis, and mammary gland development, and reportedly also related to DMI and residual feed intake in cattle and swine. The heritability (0.17) of the predicted metabolic clusters indicates that its genetic variation is large enough for genetic selection. The SNP of importance detected in this study could be used to provide weight information for SNP in future genomic prediction of metabolic clusters.

ACKNOWLEDGMENTS

This project received funding from the European Union's Seventh Framework Program (Brussels, Bel-

gium) for research, technological development, and demonstration, under grant agreement no. 613689. The views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission. Authors within the GplusE consortium: Mark Crowe, Niamh McLoughlin, Alan Fahey, Fiona Carter, Elizabeth Matthews, Andreia Santoro, Colin Byrne, Pauline Rudd, Roisin O'Flaherty, Sinead Hallinan, Claire Wathes, Mazdak Salavati, Zhangrui Cheng, Ali Fouladi, Geoff Pollott, Dirk Werling, Beatriz Sanz Bernardo, Conrad Ferris, Alistair Wylie, Matt Bell, Mieke Vaneetvelde, Kristof Hermans, Miel Hostens, Geert Opsomer, Sander Moerman, Jenne De Koster, Hannes Bogaert, Jan Vandepitte, Leila Vandeveld, Bonny Vanranst, Klaus Ingvartsen, Martin Tang Sorensen, Johanna Hoglund, Susanne Dahl, Soren Ostergaard, Janne Rothmann, Mogens Krogh, Else Meyer, Leslie Foldager, Charlotte Gaillard, Jehan Ettema, Tine Rousing, Torben Larsen, Victor Silva de Oliveira, Cinzia Marchitelli, Federica Signorelli, Francesco Napolitano, Bianca Moioli, Alessandra Crisà, Luca Buttazzoni, Jennifer McClure, Daragh Matthews, Francis Kearney, Andrew Cromie, Matt McClure, Shujun Zhang, Xing Chen, Huanchun Chen, Junlong Zhao, Liguang Yang, Guohua Hua, Chen Tan, Guiqiang Wang, Michel Bonneau, Marlène Sciarretta, Armin Pearn, Arnold Evertson, Linda Kosten, Anders Fogh, Thomas Andersen, Matthew Lucy, Chris Elvik, Gavin Conant, Jerry Taylor, Deborah Triant, Nicolas Gengler, Michel Georges, Frédéric Colinet, Marilou Ramos Pamplona, Hedi Hammami, Catherine Bastin, Haruko Takeda, Aurelie Laine, Anne-Sophie Van Laere, Rodrigo Mota, Saied Naderi Darbagshahi, Frederic Dehareng, Clement Grelet, Amelie Vanlierde, Eric Froidmont, Frank Becker, Martin Schulze, and Sergio Palma Vera. The authors declare no potential conflict of interest associated with this research.

REFERENCES


- Aguilar, I., A. Legarra, F. Cardoso, Y. Masuda, D. Lourenco, and I. Misztal. 2019. Frequentist p-values for large-scale-single step genome-wide association, with an application to birth weight in American Angus cattle. *Genet. Sel. Evol.* 51:28. <https://doi.org/10.1186/s12711-019-0469-3>.
- Aguilar, I., I. Misztal, D. Johnson, A. Legarra, S. Tsuruta, and T. 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>.
- Alawneh, J., M. Stevenson, N. Williamson, N. Lopez-Villalobos, and T. Otle. 2012. The effect of liveweight change on reproductive performance in a seasonally calving, pasture fed dairy herd. *Livest. Sci.* 145:131–139. <https://doi.org/10.1016/j.livsci.2011.12.025>.
- Aslan, O., R. M. Hamill, A. M. Mullen, G. C. Davey, M. Gil, C. D. Gladney, and T. Sweeney. 2012. Association between promoter polymorphisms in a key cytoskeletal gene (*Ankyrin 1*) and intra-

- muscular fat and water-holding capacity in porcine muscle. *Mol. Biol. Rep.* 39:3903–3914. <https://doi.org/10.1007/s11033-011-1169-4>.
- Aslan, O., T. Sweeney, A. M. Mullen, and R. M. Hamill. 2010. Regulatory polymorphisms in the bovine *Ankyrin 1* gene promoter are associated with tenderness and intramuscular fat content. *BMC Genet.* 11:111. <https://doi.org/10.1186/1471-2156-11-111>.
- Bartel, D. P. 2004. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 116:281–297. [https://doi.org/10.1016/S0092-8674\(04\)00045-5](https://doi.org/10.1016/S0092-8674(04)00045-5).
- Bauman, D. E., and W. Bruce Currie. 1980. Partitioning of nutrients during pregnancy and lactation—A review of mechanisms involving homeostasis and homeorhesis. *J. Dairy Sci.* 63:1514–1529. [https://doi.org/10.3168/jds.S0022-0302\(80\)83111-0](https://doi.org/10.3168/jds.S0022-0302(80)83111-0).
- Beerda, B., J. Kornalijnslijper, J. Van der Werf, E. Noordhuizen-Stassen, and H. Hopster. 2004. Effects of milk production capacity and metabolic status on HPA function in early postpartum dairy cows. *J. Dairy Sci.* 87:2094–2102. [https://doi.org/10.3168/jds.S0022-0302\(04\)70027-2](https://doi.org/10.3168/jds.S0022-0302(04)70027-2).
- Beissinger, T. M., G. J. Rosa, S. M. Kaepler, D. Gianola, and N. De Leon. 2015. Defining window-boundaries for genomic analyses using smoothing spline techniques. *Genet. Sel. Evol.* 47:30. <https://doi.org/10.1186/s12711-015-0105-9>.
- Bell, A. W. 1995. Regulation of organic nutrient metabolism during transition from late pregnancy to early lactation. *J. Anim. Sci.* 73:2804–2819. <https://doi.org/10.2527/1995.7392804x>.
- Bell, A. W., and D. E. Bauman. 1997. Adaptations of glucose metabolism during pregnancy and lactation. *J. Mammary Gland Biol. Neoplasia* 2:265–278. <https://doi.org/10.1023/A:1026336505343>.
- Bennett, V., and A. J. Baines. 2001. Spectrin and ankyrin-based pathways: Metazoan inventions for integrating cells into tissues. *Physiol. Rev.* 81:1353–1392. <https://doi.org/10.1152/physrev.2001.81.3.1353>.
- Bouwman, A. C., H. Bovenhuis, M. H. Visker, and J. A. 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>.
- Cai, Z., B. Guldbrandtsen, M. S. Lund, and G. Sahana. 2019. Dissecting closely linked association signals in combination with the mammalian phenotype database can identify candidate genes in dairy cattle. *BMC Genet.* 20:15. <https://doi.org/10.1186/s12863-019-0717-0>.
- Collard, B. L., P. J. Boettcher, J. C. Dekkers, D. Pettilerc, and L. R. Schaeffer. 2000. Relationships between energy balance and health traits of dairy cattle in early lactation. *J. Dairy Sci.* 83:2683–2690. [https://doi.org/10.3168/jds.S0022-0302\(00\)75162-9](https://doi.org/10.3168/jds.S0022-0302(00)75162-9).
- Contreras, G. A., C. Strieder-Barboza, and J. De Koster. 2018. Symposium review: Modulating adipose tissue lipolysis and remodeling to improve immune function during the transition period and early lactation of dairy cows. *J. Dairy Sci.* 101:2737–2752. <https://doi.org/10.3168/jds.2017-13340>.
- De Koster, J., M. Salavati, C. Grelet, M. A. Crowe, E. Matthews, R. O'Flaherty, G. Opsomer, L. Foldager, and M. Hostens. 2019. Prediction of metabolic clusters in early-lactation dairy cows using models based on milk biomarkers. *J. Dairy Sci.* 102:2631–2644. <https://doi.org/10.3168/jds.2018-15533>.
- De Oliveira, P. S., L. L. Coutinho, P. C. Tizioto, A. S. Cesar, G. B. de Oliveira, W. J. S. Diniz, A. O. De Lima, J. M. Reecy, G. B. Mourão, A. Zerlotini, and L. C. A. Regitano. 2018. An integrative transcriptome analysis indicates regulatory mRNA-miRNA networks for residual feed intake in Nelore cattle. *Sci. Rep.* 8:17072. <https://doi.org/10.1038/s41598-018-35315-5>.
- DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium. 2012. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* 44:981–990. <https://doi.org/10.1038/ng.2383>.
- Drackley, J. K. 1999. Biology of dairy cows during the transition period: The final frontier? *J. Dairy Sci.* 82:2259–2273. [https://doi.org/10.3168/jds.S0022-0302\(99\)75474-3](https://doi.org/10.3168/jds.S0022-0302(99)75474-3).
- Drackley, J. K., H. M. Dann, G. N. Douglas, N. A. J. Guretzky, N. B. Litherland, J. P. Underwood, and J. J. Looor. 2005. Physiological and pathological adaptations in dairy cows that may increase susceptibility to periparturient diseases and disorders. *Ital. J. Anim. Sci.* 4:323–344. <https://doi.org/10.4081/ijas.2005.323>.
- Esposito, G., P. C. Irons, E. C. Webb, and A. Chapwanya. 2014. Interactions between negative energy balance, metabolic diseases, uterine health and immune response in transition dairy cows. *Anim. Reprod. Sci.* 144:60–71. <https://doi.org/10.1016/j.anireprosci.2013.11.007>.
- Fragomeni, B. O., I. Misztal, D. L. Lourenco, I. Aguilar, R. Okimoto, and W. M. Muir. 2014. Changes in variance explained by top SNP windows over generations for three traits in broiler chicken. *Front. Genet.* 5:332. <https://doi.org/10.3389/fgene.2014.00332>.
- Gallagher, P. G., M. Romana, W. T. Tse, S. E. Lux, and B. G. Forget. 2000. The human *ankyrin-1* gene is selectively transcribed in erythroid cell lines despite the presence of a housekeeping-like promoter. *Blood* 96:1136–1143. <https://doi.org/10.1182/blood.V96.3.1136>.
- Goff, J. P., and R. Horst. 1997. Physiological changes at parturition and their relationship to metabolic disorders. *J. Dairy Sci.* 80:1260–1268. [https://doi.org/10.3168/jds.S0022-0302\(97\)76055-7](https://doi.org/10.3168/jds.S0022-0302(97)76055-7).
- Han, Y., and F. Peñagaricano. 2016. Unravelling the genomic architecture of bull fertility in Holstein cattle. *BMC Genet.* 17:143. <https://doi.org/10.1186/s12863-016-0454-6>.
- Herr, M., H. Bostedt, and K. Failing. 2011. IgG and IgM levels in dairy cows during the periparturient period. *Theriogenology* 75:377–385. <https://doi.org/10.1016/j.theriogenology.2010.09.009>.
- Horodyska, J., T. Sweeney, M. Ryan, and R. Hamill. 2015. Novel SNPs in the *Ankyrin 1* gene and their association with beef quality traits. *Meat Sci.* 108:88–96. <https://doi.org/10.1016/j.meatsci.2015.04.019>.
- Imamura, M., S. Maeda, T. Yamauchi, K. Hara, K. Yasuda, T. Morizono, A. Takahashi, M. Horikoshi, M. Nakamura, H. Fujita, T. Tsunoda, M. Kubo, H. Watada, H. Maegawa, M. Okada-Iwabu, M. Iwabu, N. Shojima, T. Ohshige, S. Omori, M. Iwata, H. Hirose, K. Kaku, C. Ito, Y. Tanaka, K. Tobe, A. Kashiwagi, R. Kawamori, M. Kasuga, N. Kamatani, Y. Nakamura, and T. Kadowaki. 2012. A single-nucleotide polymorphism in ANK1 is associated with susceptibility to type 2 diabetes in Japanese populations. *Hum. Mol. Genet.* 21:3042–3049. <https://doi.org/10.1093/hmg/dds113>.
- Ingvartsen, K. L. 2006. Feeding- and management-related diseases in the transition cow: Physiological adaptations around calving and strategies to reduce feeding-related diseases. *Anim. Feed Sci. Technol.* 126:175–213. <https://doi.org/10.1016/j.anifeeds.2005.08.003>.
- Ingvartsen, K. L., and J. B. Andersen. 2000. Integration of metabolism and intake regulation: A review focusing on periparturient animals. *J. Dairy Sci.* 83:1573–1597. [https://doi.org/10.3168/jds.S0022-0302\(00\)75029-6](https://doi.org/10.3168/jds.S0022-0302(00)75029-6).
- Ingvartsen, K. L., R. J. Dewhurst, and N. C. Friggens. 2003. On the relationship between lactational performance and health: Is it yield or metabolic imbalance that cause production diseases in dairy cattle? A position paper. *Livest. Prod. Sci.* 83:277–308. [https://doi.org/10.1016/S0301-6226\(03\)00110-6](https://doi.org/10.1016/S0301-6226(03)00110-6).
- Jing, L., Y. Hou, H. Wu, Y. Miao, X. Li, J. Cao, J. Michael Brameld, T. Parr, and S. Zhao. 2015. Transcriptome analysis of mRNA and miRNA in skeletal muscle indicates an important network for differential Residual Feed Intake in pigs. *Sci. Rep.* 5:11953. <https://doi.org/10.1038/srep11953>.
- Kessel, S., M. Stroehl, H. H. D. Meyer, S. Hiss, H. Sauerwein, F. J. Schwarz, and R. M. Bruckmaier. 2008. Individual variability in physiological adaptation to metabolic stress during early lactation in dairy cows kept under equal conditions. *J. Anim. Sci.* 86:2903–2912. <https://doi.org/10.2527/jas.2008-1016>.

- Kolbehdari, D., Z. Wang, J. Grant, B. Murdoch, A. Prasad, Z. Xiu, E. Marques, P. Stothard, and S. Moore. 2008. A whole-genome scan to map quantitative trait loci for conformation and functional traits in Canadian Holstein bulls. *J. Dairy Sci.* 91:2844–2856. <https://doi.org/10.3168/jds.2007-0585>.
- Krogh, M. A., M. Hostens, M. Salavati, C. Grelet, M. T. Sorensen, D. C. Wathes, C. P. Ferris, C. Marchitelli, F. Signorelli, F. Napolitano, F. Becker, T. Larsen, E. Matthews, F. Carter, A. Vanlierde, G. Opsomer, N. Gengler, F. Dehareng, M. A. Crowe, K. L. Ingvarsen, and L. Foldager. 2019. Between- and within-herd variation in blood and milk biomarkers in Holstein cows in early lactation. *Animal* 7:1–9. <https://doi.org/10.1017/S1751731119002659>.
- Lander, E., and L. Kruglyak. 1995. Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. *Nat. Genet.* 11:241–247. <https://doi.org/10.1038/ng1195-241>.
- LeBlanc, S. 2010. Monitoring metabolic health of dairy cattle in the transition period. *J. Reprod. Develop.* 56(Suppl.):S29–S35.
- Li, D., X. Xie, J. Wang, Y. Bian, Q. Li, X. Gao, and C. Wang. 2015. MiR-486 regulates lactation and targets the PTEN gene in cow mammary glands. *PLoS One* 10:e0118284. <https://doi.org/10.1371/journal.pone.0118284>.
- Liinamo, A.-E., P. Mäntysaari, and E. Mäntysaari. 2012. Genetic parameters for feed intake, production, and extent of negative energy balance in Nordic Red dairy cattle. *J. Dairy Sci.* 95:6788–6794. <https://doi.org/10.3168/jds.2012-5342>.
- Maltecca, C., K. Weigel, H. Khatib, M. Cowan, and A. Bagnato. 2009. Whole-genome scan for quantitative trait loci associated with birth weight, gestation length and passive immune transfer in a Holstein × Jersey crossbred population. *Anim. Genet.* 40:27–34. <https://doi.org/10.1111/j.1365-2052.2008.01793.x>.
- McClure, M. C., N. S. Morsci, R. D. Schnabel, J. W. Kim, P. Yao, M. M. Rolf, S. D. McKay, S. J. Gregg, R. H. Chapple, S. L. Northcutt, and J. F. Taylor. 2010. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. *Anim. Genet.* 41:597–607. <https://doi.org/10.1111/j.1365-2052.2010.02063.x>.
- Medeiros de Oliveira Silva, R., N. Bonvino Stafuzza, B. de Oliveira Fragomeni, G. Miguel Ferreira de Camargo, T. Matos Ceacero, J. Noely dos Santos Gonçalves Cyrillo, F. Baldi, A. Augusti Boligon, M. E. Zerlotti Mercadante, D. Lino Lourenco, I. Misztal, and L. Galvão de Albuquerque. 2017. Genome-wide association study for carcass traits in an experimental Nelore cattle population. *PLoS One* 12:e0169860. <https://doi.org/10.1371/journal.pone.0169860>.
- Michenet, A., M. Barbat, R. Saintilan, E. Venot, and F. Phocas. 2016. Detection of quantitative trait loci for maternal traits using high-density genotypes of Blonde d'Aquitaine beef cattle. *BMC Genet.* 17:88. <https://doi.org/10.1186/s12863-016-0397-y>.
- Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. Lee. 2002. BLUPF90 and related programs (BGF90). Pages 743–744 in *Proc. 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France.
- Myles, S., K. Tang, M. Somel, R. E. Green, J. Kelso, and M. Stoneking. 2008. Identification and analysis of genomic regions with large between-population differentiation in humans. *Ann. Hum. Genet.* 72:99–110.
- Nelson, W. J., and E. Lazarides. 1984. Goblin (ankyrin) in striated muscle: Identification of the potential membrane receptor for erythroid spectrin in muscle cells. *Proc. Natl. Acad. Sci. USA* 81:3292–3296. <https://doi.org/10.1073/pnas.81.11.3292>.
- Patton, J., D. A. Kenny, J. F. Mee, F. P. O'Mara, D. C. Wathes, M. Cook, and J. J. Murphy. 2006. Effect of milking frequency and diet on milk production, energy balance, and reproduction in dairy cows. *J. Dairy Sci.* 89:1478–1487. [https://doi.org/10.3168/jds.S0022-0302\(06\)72215-9](https://doi.org/10.3168/jds.S0022-0302(06)72215-9).
- Piao, H. L., and L. Ma. 2012. Non-coding RNAs as regulators of mammary development and breast cancer. *J. Mammary Gland Biol. Neoplasia* 17:33–42. <https://doi.org/10.1007/s10911-012-9245-5>.
- Pryce, J. E., K. L. Parker Gaddis, A. Koeck, C. Bastin, M. Abdel-sayed, N. Gengler, F. Miglier, B. Heringstad, C. Egger-Danner, K. F. Stock, A. J. Bradley, and J. B. Cole. 2016. Invited review: Opportunities for genetic improvement of metabolic diseases. *J. Dairy Sci.* 99:6855–6873. <https://doi.org/10.3168/jds.2016-10854>.
- Puppel, K., and B. Kuczynska. 2016. Metabolic profiles of cow's blood: A review. *J. Sci. Food Agric.* 96:4321–4328. <https://doi.org/10.1002/jsfa.7779>.
- Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. De Bakker, M. J. Daly, and P. C. Sham. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81:559–575. <https://doi.org/10.1086/519795>.
- R Core Team. 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. *BMC Genomics* 15:478. <https://doi.org/10.1186/1471-2164-15-478>.
- Shaham, L., E. Vendramini, Y. Ge, Y. Goren, Y. Birger, M. R. Tijssen, M. McNulty, I. Geron, O. Schwartzman, L. Goldberg, S. T. Chou, H. Pitman, M. J. Weiss, S. Michaeli, B. Sredni, B. Göttgens, J. D. Crispino, J. W. Taub, and S. Izraeli. 2015. MicroRNA-486-5p is an erythroid oncomiR of the myeloid leukemias of Down syndrome. *Blood* 125:1292–1301. <https://doi.org/10.1182/blood-2014-06-581892>.
- Sherman, E. L., J. D. Nkrumah, C. Li, R. Bartusiak, B. Murdoch, and S. S. Moore. 2009. Fine mapping quantitative trait loci for feed intake and feed efficiency in beef cattle. *J. Anim. Sci.* 87:37–45. <https://doi.org/10.2527/jas.2008-0876>.
- Sordillo, L. M. 2016. Nutritional strategies to optimize dairy cattle immunity. *J. Dairy Sci.* 99:4967–4982. <https://doi.org/10.3168/jds.2015-10354>.
- Spurlock, D. M., J. C. M. Dekkers, R. Fernando, D. A. Koltes, and A. Wolc. 2012. Genetic parameters for energy balance, feed efficiency, and related traits in Holstein cattle. *J. Dairy Sci.* 95:5393–5402. <https://doi.org/10.3168/jds.2012-5407>.
- Sureshchandra, S., M. Rais, C. Stull, K. Grant, and I. Messaoudi. 2016. Transcriptome profiling reveals disruption of innate immunity in chronic heavy ethanol consuming female rhesus macaques. *PLoS One* 11:e0159295. <https://doi.org/10.1371/journal.pone.0159295>.
- Suwannasing, R., M. Duangjinda, W. Boonkum, R. Taharnklaew, and K. Tuangsihtanon. 2018. The identification of novel regions for reproduction trait in Landrace and Large White pigs using a single step genome-wide association study. *Asian-Australas. J. Anim. Sci.* 31:1852–1862. <https://doi.org/10.5713/ajas.18.0072>.
- Tessema, M., C. M. Yingling, M. A. Picchi, G. Wu, T. Ryba, Y. Lin, A. O. Bungum, E. S. Edell, A. Spira, and S. A. Belinsky. 2017. ANK1 methylation regulates expression of MicroRNA-486-5p and discriminates lung tumors by histology and smoking status. *Cancer Lett.* 410:191–200. <https://doi.org/10.1016/j.canlet.2017.09.038>.
- Tetens, J., G. Thaller, and N. Krattenmacher. 2014. Genetic and genomic dissection of dry matter intake at different lactation stages in primiparous Holstein cows. *J. Dairy Sci.* 97:520–531. <https://doi.org/10.3168/jds.2013-7301>.
- Valesia, A., Q.-P. Wang, N. Gheldof, J. Carayol, H. Ruffieux, T. Clark, V. Shenton, L. J. Oyston, G. Lefebvre, S. Metairon, C. Chabert, O. Walter, P. Mironova, P. Lau, P. Descombes, N. Viguerie, D. Langin, M.-E. Harper, A. Astrup, W. H. Saris, R. Dent, G. G. Neely, and J. Hager. 2019. Genome-wide gene-based analyses of weight loss interventions identify a potential role for *NKX6.3* in metabolism. *Nat. Commun.* 10:540. <https://doi.org/10.1038/s41467-019-08492-8>.
- van der Drift, S. G. A., M. Houweling, J. T. Schonewille, A. G. M. Tielens, and R. Jorritsma. 2012. Protein and fat mobilization and associations with serum β -hydroxybutyrate concentrations in dairy cows. *J. Dairy Sci.* 95:4911–4920. <https://doi.org/10.3168/jds.2011-4771>.
- van Hoeij, R., A. Kok, R. Bruckmaier, M. Haskell, B. Kemp, and A. van Knegsel. 2019. Relationship between metabolic status and behavior in dairy cows in week 4 of lactation. *Animal* 13:640–648.

- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423. <https://doi.org/10.3168/jds.2007-0980>.
- Wang, H., I. Misztal, I. Aguilar, A. Legarra, and W. 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, K., D. Liu, J. Hernandez-Sanchez, J. Chen, C. Liu, Z. Wu, M. Fang, and N. Li. 2015a. Genome wide association analysis reveals new production trait genes in a male Duroc population. *PLoS One* 10:e0139207. <https://doi.org/10.1371/journal.pone.0139207>.
- Wang, L.-S., L. Li, L. Li, S. Chu, K.-D. Shiang, M. Li, H.-Y. Sun, J. Xu, F.-J. Xiao, G. Sun, J. J. Rossi, Y. W. Ho, and R. Bhatia. 2015b. MicroRNA-486 regulates normal erythropoiesis and enhances growth and modulates drug response in CML progenitors. *Blood* 125:1302–1313. <https://doi.org/10.1182/blood-2014-06-581926>.
- Zhang, H., Z. Wang, S. Wang, and H. Li. 2012. Progress of genome wide association study in domestic animals. *J. Anim. Sci. Biotechnol.* 3:26. <https://doi.org/10.1186/2049-1891-3-26>.
- Zhou, C., C. Li, W. Cai, S. Liu, H. W. Yin, S. Shi, Q. Zhang, and S. Zhang. 2019. Genome-wide association study for milk protein composition traits in a Chinese Holstein population using a single-step approach. *Front. Genet.* 10:72. <https://doi.org/10.3389/fgene.2019.00072>.
- Zhou, X., and M. Stephens. 2012. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* 44:821–824. <https://doi.org/10.1038/ng.2310>.
- Zwald, N. R., K. A. Weigel, Y. M. Chang, R. D. Welper, and J. S. Clay. 2004. Genetic selection for health traits using producer-recorded data. II. Genetic correlations, disease probabilities, and relationships with existing traits. *J. Dairy Sci.* 87:4295–4302. [https://doi.org/10.3168/jds.S0022-0302\(04\)73574-2](https://doi.org/10.3168/jds.S0022-0302(04)73574-2).

ORCIDS

- H. Atashi  <https://orcid.org/0000-0002-6853-6608>
M. Salavati  <https://orcid.org/0000-0002-7349-2451>
J. De Koster  <https://orcid.org/0000-0002-2891-0298>
M. A. Crowe  <https://orcid.org/0000-0003-1999-815X>
G. Opsomer  <https://orcid.org/0000-0002-6131-1000>
M. Hostens  <https://orcid.org/0000-0001-5376-976X>