



Prediction of metabolic clusters in early-lactation dairy cows using models based on milk biomarkers

J. De Koster,¹ M. Salavati,² C. Grelet,³ M. A. Crowe,⁴ E. Matthews,⁴ R. O’Flaherty,⁵ G. Opsomer,¹ L. Foldager,^{6,7} GplusE,* and M. Hostens^{1†}

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

²Royal Veterinary College, NW1 0TU London, United Kingdom

³Walloon Agricultural Research Center, Valorisation of Agricultural Products Department, B-5030 Gembloux, Belgium

⁴University College Dublin, 4 Dublin, Ireland

⁵GlycoScience Group, NIBRT, Fosters Avenue, Mount Merion, 4 Dublin, Ireland

⁶Department of Animal Science, Aarhus University, DK-8830 Tjele, Denmark

⁷Bioinformatics Research Centre, Aarhus University, DK-8000 Aarhus, Denmark

ABSTRACT

The aim of this study was to describe metabolism of early-lactation dairy cows by clustering cows based on glucose, insulin-like growth factor I (IGF-I), free fatty acid, and β -hydroxybutyrate (BHB) using the k-means method. Predictive models for metabolic clusters were created and validated using 3 sets of milk biomarkers (milk metabolites and enzymes, glycans on the immunoglobulin fraction of milk, and Fourier-transform mid-infrared spectra of milk). Metabolic clusters are used to identify dairy cows with a balanced or imbalanced metabolic profile. Around 14 and 35 d in milk, serum or plasma concentrations of BHB, free fatty acids, glucose, and IGF-I were determined. Cows with a favorable metabolic profile were grouped together in what was referred to as the “balanced” group ($n = 43$) and were compared with cows in what was referred to

as the “other balanced” group ($n = 64$). Cows with an unfavorable metabolic profile were grouped in what was referred to as the “imbalanced” group ($n = 19$) and compared with cows in what was referred to as the “other imbalanced” group ($n = 88$). Glucose and IGF-I were higher in balanced compared with other balanced cows. Free fatty acids and BHB were lower in balanced compared with other balanced cows. Glucose and IGF-I were lower in imbalanced compared with other imbalanced cows. Free fatty acids and BHB were higher in imbalanced cows. Metabolic clusters were related to production parameters. There was a trend for a higher daily increase in fat- and protein-corrected milk yield in balanced cows, whereas that of imbalanced cows was higher. Dry matter intake and the daily increase in dry matter intake were higher in balanced cows and lower in imbalanced cows. Energy balance was continuously higher in balanced cows and lower in imbalanced cows. Weekly or twice-weekly milk samples were taken and milk metabolites and enzymes (milk glucose, glucose-6-phosphate, BHB, lactate dehydrogenase, *N*-acetyl- β -D-glucosaminidase, isocitrate), immunoglobulin glycans (19 peaks), and Fourier-transform mid-infrared spectra (1,060 wavelengths reduced to 15 principal components) were determined. Milk biomarkers with or without additional cow information (days in milk, parity, milk yield features) were used to create predictive models for the metabolic clusters. Accuracy for prediction of balanced (80%) and imbalanced (88%) cows was highest using milk metabolites and enzymes combined with days in milk and parity. The results and models of the present study are part of the GplusE project and identify novel milk-based phenotypes that may be used as predictors for metabolic and performance traits in early-lactation dairy cows.

Key words: metabolic clustering, dairy cow, prediction, milk biomarker

Received August 10, 2018.

Accepted November 25, 2018.

*List of authors in the GplusE consortium: Niamh McLoughlin, Alan Fahey, Elizabeth Matthews, Andreia Santoro, Colin Byrne, Pauline Rudd, Roisin O’Flaherty, Sinead Hallinan, Claire Wathes, Zhangrui Cheng, Ali Fouladi, Geoff Pollott, Dirk Werling, Beatriz Sanz Bernardo, Alistair Wylie, Matt Bell, Mieke Vaneetvelde, Kristof Hermans, Geert Opsomer, Sander Moerman, Jenne De Koster, Hannes Bogaert, Jan Vandepitte, Leila Vandeveld, Bonny Vanranst, Johanna Hoglund, Susanne Dahl, Soren Ostergaard, Janne Rothmann, Mogens Krogh, Else Meyer, Charlotte Gaillard, Jehan Ettema, Tine Rousing, Federica Signorelli, Francesco Napolitano, Bianca Moiola, 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, Andrea Pompozzi, Armin Pearn, Arnold Evertson, Linda Kosten, Anders Fogh, Thomas Andersen, Matthew Lucy, Chris Elsik, Gavin Conant, Jerry Taylor, Nicolas Gengler, Michel Georges, Frédéric Colinet, Marilou Ramos Pamplona, Hedi Hammami, Catherine Bastin, Haruko Takeda, Aurelie Laine, Anne-Sophie Van Laere, Martin Schulze, and Sergio Palma Vera.

†Corresponding author: miel.hostens@ugent.be

INTRODUCTION

Physiological adaptations of dairy cows in the transition period are well described (Bell and Bauman, 1997; Drackley et al., 2005). As parturition approaches, growth hormone (**GH**) concentrations start to increase and insulin and IGF-I concentrations decrease (Lucy et al., 2009). Growth hormone enhances milk production through partitioning of nutrients toward the mammary gland (Lucy, 2008), lipolysis increases, lipogenesis is almost completely downregulated, and hepatic gluconeogenesis is stimulated (Bauman, 1999; Lucy, 2008). Growth hormone induces the production of IGF-I by the liver, which inhibits GH secretion by hypophysis (Lucy, 2008). Homeorhetic adaptation mechanisms in the periparturient period decrease the expression of GH receptors by the liver, which downregulate hepatic IGF-I production, known as uncoupling of the GH-IGF-I axis (Lucy, 2008). Decreased IGF-I concentrations in the periparturient period relieve negative feedback on pituitary GH production (Lucy, 2008). As lactation progresses, recoupling of the GH-IGF-I axis is characterized by repleted IGF-I levels and effectuated by slightly elevated insulin levels (Lucy, 2008).

Homeorhetic mechanisms commonly observed during the transition period support the metabolic prioritization of the lactating mammary gland, leading to a sharp increase in milk production (Bauman and Currie, 1980). The increase in DMI in postpartum cows lags behind the increased milk production, which results in a period of negative energy balance (**EB**; Grummer et al., 2004). Severe negative EB is a risk factor for metabolic, infectious, and reproductive disorders (Moyes et al., 2013). Metabolic indicators for the severity of the negative EB are used to detect individual cows that are unable to cope with the altered metabolic challenge of lactation. Elevated free fatty acids (**FFA**) and BHB and decreased glucose and IGF-I are reported as biomarkers of metabolically imbalanced cows, which are more at risk for an unsuccessful transition from the dry period to lactation (Ingvarlsen et al., 2003; Puppel and Kuczynska, 2016). Cut-off values for FFA and BHB have been determined to predict negative health consequences in the periparturient period (McArt et al., 2013; Ospina et al., 2013). However, the metabolic profile of individual cows within a herd shows a high level of variability (Ingvarlsen et al., 2003). The correlation between FFA and BHB is especially difficult to explain. Although some cows with normal BHB levels have high FFA levels, other cows with normal FFA levels have high BHB levels, demonstrating unexplained interindividual variability within a group of cows (Ospina et al., 2013; McCarthy et al., 2015). This interindividual variability represents the interindi-

vidual ability of cows to adapt to the altered metabolic challenge of lactation and demonstrates the limitations of monitoring metabolism based on a single metabolic indicator (Ingvarlsen et al., 2003; Bjerre-Harpøth et al., 2012; Moyes et al., 2013). Moreover, monitoring early-lactation metabolism requires early identifiable biomarkers. Due to better accessibility and the ease of automated repeat sampling owing to recently developed in-line sampling and analytical technologies, milk is the preferred medium in which biomarkers can be measured (Ingvarlsen and Friggens, 2005; Nielsen et al., 2005a; Egger-Danner et al., 2015). It is highly probable that certain glycan structures (Zhao and Keating, 2007), milk metabolites and enzymes (Weekes et al., 1983; Wallace and Matthews, 2002; Bjerre-Harpøth et al., 2012), or mid-infrared spectra (Voelker and Allen, 2003; Maury et al., 2007; Soyeurt et al., 2011) of bovine milk can serve as biomarkers to monitor early-lactation metabolism and performance, but comparable prediction methodologies are lacking.

The aim of this study was to combine concentrations of blood metabolites to describe metabolism of dairy cows in early and peak lactation by clustering cows based on concentrations of glucose, IGF-I, FFA, and BHB. Metabolic clusters are used to identify dairy cows with a balanced or imbalanced metabolic profile. Furthermore, the relationship between metabolic clusters and production parameters [DMI, fat- and protein-corrected milk (**FPCM**) production, BW, EB, and BCS] are determined. Finally, 3 sets of milk biomarkers—milk metabolites and enzymes [milk biomarker (**MBM**) 1], glycans on the IgG fraction of milk (**MBM**2), and Fourier-transform mid-infrared spectra (**FT-MIR**) of milk (**MBM**3)—are used to create and validate predictive models for the different metabolic clusters.

MATERIALS AND METHODS

The experiments were carried out in accordance with the standards recommended by EU Directive 2010/63/EU for animal experiments. Detailed description of the experiments, laboratory analysis of blood and milk, production data, and so on are given in Foldager et al. (2019) and described in brief below.

Animals and Sampling

Samples and data between calving and 50 d post-calving (1–50 DIM) were obtained from 130 Holstein Friesian cows (parity 2: $n = 42$; parity 3: $n = 51$; parity ≥ 4 : $n = 37$) in 4 research herds: Aarhus University, Denmark; UCD Lyons Research Farm, University College Dublin, Ireland; Agri-Food and Biosciences Insti-

tute, Northern Ireland, UK; and Leibniz Institute for Farm Animal Biology, Dummerstorf, Germany. Cows were milked twice daily.

Analyses of Glucose, IGF-I, FFA, and BHB

Blood samples were taken around 14 DIM (15 ± 0.1 DIM; **D14**) and 35 DIM (37 ± 0.1 DIM; **D35**) in serum and heparin tubes by jugular or coccygeal venipuncture. Blood plasma glucose concentrations were determined using an enzymatic method (Advia 1800 Clinical Chemistry System, Siemens Diagnostics, Erlangen, Germany). Plasma FFA concentrations were determined using an enzymatic method (NEFA C ACS-ACOD assay method, Wako Diagnostics, Mountain View, CA). Plasma BHB concentrations were determined by measuring absorbance at 340 nm due to the production of NADH at alkaline pH in the presence of BHB dehydrogenase. Serum concentrations of IGF-I were determined using an RIA following acid-ethanol extraction using the method previously described by Beltman et al. (2010).

Analyses of Milk Metabolites and Enzymes, Milk Glycans, and FT-MIR

Metabolites and Enzymes. Weekly, 0200 h milk samples were collected starting from the first week in milk until 50 DIM. At each day of sampling, 2 separate samples of approximately 8 mL were obtained and stored at -18°C . Fluorometric assays were used to determine milk glucose and glucose-6-phosphate (Larsen, 2015), BHB (Larsen and Nielsen, 2005), lactate dehydrogenase (Larsen, 2005), *N*-acetyl- β -D-glucosaminidase (Larsen et al., 2010), and isocitrate (Larsen, 2014). Urea was determined by spectrophotometry (Nielsen et al., 2005b). Finally, 6 milk metabolites and enzymes were available as the first set of milk biomarker (MBM1).

IgG Glycans. Weekly, 0200 h milk samples were collected starting from the first week in milk until 50 DIM. Milk samples were centrifuged at $4,000 \times g$ for 30 min at 4°C . Five milliliters of whole milk supernatant was recovered and frozen at -20°C . In duplicate, 300 μL of thawed sample was filtered through a 1- μm glass fiber filter plate (Acroprep, VWR International Ltd., Radnor, PA) at $3,000 \times g$ for 10 min at room temperature and collected in a 96-well Greiner plate (Cruinn Diagnostics, Dublin, Ireland). All processes relating to IgG purification and IgG-glycan release were carried out on a Hamilton Robotics StarLet liquid-handling platform (Hamilton Robotics, Reno, NV) using a protocol adapted from Stöckmann et al. (2013). The *N*-glycans were separated on a Waters Acquity UPLC

instrument (Waters, Milford, MA) and analyzed using Empower V3 (Waters).

The IgG was captured by passing 290 μL of filtered sample through a Protein G matrix (Phytip G; 200- μL column, 200- μL resin bed; Phynexus, San Jose, CA). The column tips were washed 5 times with a 0.1 *M* sodium phosphate binding buffer (pH 7.4), eluted with a 0.2 *M* glycine-hydrochloride buffer (pH 2.5), and neutralized with 1 *M* Tris-hydrochloride buffer (pH 9.0). The purified protein was enriched by pooling from 2 identical plates.

One hundred microliters of the enriched IgG was transferred to a 96-well ultrafiltration plate (Acroprep, Omega membrane, 10 kDa; VWR International Ltd.) and centrifuged at $3,000 \times g$ for 30 min at room temperature. Fifty microliters of a dithiothreitol denaturing buffer was added to each well, mixed, and left at room temperature for 10 min. The sample was then transferred to a 96-well thermobalanced denaturation plate (Armadillo High Performance 96-well PCR Plate; Thermo Fisher Scientific, Waltham, MA), incubated at 95°C for 10 min, and cooled at room temperature for 20 min. The denatured sample was transferred back onto the ultrafiltration plate and incubated in 20 μL of 1 *M* iodoacetamide buffer for 10 min. The sample was washed with 20 μL of 25 *mM* sodium bicarbonate and filtered. Then, 0.4 μL of PNGase F (2.5 U/mL; New England Biolabs, Ipswich, UK) was added to each well and incubated for 30 min at 38°C with agitation. The released *N*-glycans were recovered by centrifugation at $3,000 \times g$ through a 10-kDa filter for 10 min at room temperature.

Eight microliters of the *N*-glycan sample was incubated with 12 μL of the fluorescent tag 6-aminoquinolyl-*N*-hydroxysuccinimidyl carbamate (3 mg/mL MeCN) at room temperature. Sixteen microliters of each sample was separated using hydrophilic separation chromatography (Glycan BEH Amide 130Å Column; Waters), and 19 peaks were manually identified and integrated. Each peak's percentage of the total area under the 19 peaks was used as the IgG glycan measure for the statistical analyses. This set of 19 peaks was used as a second set of milk biomarker (MBM2).

FT-MIR Spectra. Twice weekly, a.m. and p.m. milk samples were collected starting from the first week in milk until 50 DIM, preserved with bronopol 0.02%, and stored at 4°C . Analyses were done locally on FT2 and FT6000 spectrometers (Foss, Hillerød, Denmark) or at Centre Wallon de Recherches Agronomiques (Belgium) by a Standard Lactoscope FT-MIR Automatic (Delta Instruments, Drachten, the Netherlands). The a.m. and p.m. FT-MIR spectra were combined into a daily spectrum by a weighted average taking into

account the a.m. and p.m. milk yields. The FT-MIR spectra of the different instruments were standardized and merged into a common data set following the procedure described in Grelet et al. (2015). Finally, absorbance values at 1,060 wavenumbers were available as a third set of milk biomarkers (MBM3). Samples were analyzed locally for fat, protein, and lactose content by Fourier-transform infrared spectroscopy with FT2 and FT6000 spectrometers (Foss) or a Standard Lactoscope FT-MIR Automatic (Delta Instruments) and using the predictive models provided by the manufacturers.

FPCM, DMI, EB, and BW Measurements

All cows were milked twice daily and daily yields were recorded. Milk samples were collected twice weekly until 50 DIM and analyzed for composition of protein, fat, and lactose by mid-infrared analysis and for SCC by flow cytometry. Bronopol (0.02%; Sigma-Aldrich, St. Louis, MO) was added as a preservative to all samples. Milk weighted average for milk components was used in the subsequent analysis. Fat- and protein-corrected milk production was calculated as $(0.337 + 0.116 \times \text{milk fat \%} + 0.06 \times \text{milk protein \%}) \times \text{kg of milk}$ (CVB, 2007).

Body condition score was recorded at the moment of blood sampling using a 1-to-5 scale with 0.25-point increments (Edmonson et al., 1989). Live weights were recorded twice weekly using weight scales until 50 DIM. Daily individual DMI was recorded using an electronic feeding system (Insentec, Markneesse, the Netherlands). Over the course of the sampling period, weekly ration samples were collected, dried, and shipped for NE_L analysis in a single run (Cumberland Valley Agricultural Services, Franklin County, PA). The EB was calculated by calculating the daily difference between energy input and output. The energy input was calculated by multiplying the weekly NE_L density of the ration with the daily DMI of the animal. The energy output was calculated according to the NRC (2001) calculation correcting for NE_M using the last recorded BW of the animal and most recent fat, protein, and lactose content together with the daily milk yield for the NE_L .

Statistical Analyses

Metabolic Clustering. Due to missing observations and technical and logistical limitations, only 107 cows were included in the final statistical analyses. The k-means clustering method was used to group cows based on the log-transformed and standardized concentrations (mean = 0; SD = 1) of glucose, IGF-I, FFA, and BHB on D14 and D35. Sum of squares plots were used

to determine the optimal number of clusters. Three clusters were created in both periods: A (n = 51), B (n = 51), and C (n = 5) on D14 and D (n = 58), E (n = 25), and F (n = 24) on D35. Pairwise comparisons of glucose, IGF-I, FFA, and BHB between A, B, and C on D14 and D, E, and F on D35 were done using an ANOVA. The clustering of the individual cows on D14 and D35 was compared. Cows in cluster A on D14 and cluster D on D35 were considered to be metabolically balanced (n = 43; referred to as “balanced” cows). Cows not in cluster A on D14 and cluster D on D35 were grouped together (n = 64; referred to as “other balanced” cows). Likewise, cows in cluster B on D14 and cluster F on D35 (n = 14) together with cows in cluster C on D14 and cluster F on D35 (n = 5) were considered to be metabolically imbalanced (n = 19; referred to as “imbalanced” cows). Cows not in these clusters on D14 or D35 were grouped together (n = 88; referred to as “other imbalanced” cows).

Metabolic and Production Performance of Metabolic Clusters. Pairwise comparisons between balanced and other balanced and between imbalanced and other imbalanced cows were done using a linear mixed effect model after log transformation of the metabolite and hormone concentrations with period (D14 and D35) as repeated observation within the random factor cow. Linear mixed effect models were constructed for the FPCM, DMI, EB, and BW of the cows in the metabolic cluster (balanced vs. other balanced and imbalanced vs. other imbalanced), with day postcalving as a repeated observation within the random factor cow. All pairwise comparisons were done using Tukey's post hoc test. Residuals of the models were checked and found to be normally distributed. Interaction effects were removed from the model if nonsignificant ($P < 0.05$). Significance and tendency were declared at $P < 0.05$ and $0.05 \leq P < 0.10$, respectively. Results are presented as least squares means \pm standard error of the mean unless otherwise stated.

Prediction of Metabolic Clusters

Feature Preparation of FT-MIR. Due to the nature of the FT-MIR data, a dimension reduction step was needed to reduce the high number of variables (n = 1,060) in contrast with glycan (n = 19) or blood metabolites (n = 6). First, wavenumbers known to be noninformative due to the water component in milk (Grelet et al., 2016) were removed. Next, FT-MIR spectra were reduced using a principal component analysis. Exploration of the variance plot revealed 15 principal components that contributed most to the entire variation in FT-MIR data. After the filter and reduction step, 1,060 wavenumbers were reduced to 15 principal

components, which were used in the remainder of the analysis.

Biomarker Preparation. Each set of biomarkers was subsequently split into 3 sets: 1 without additional cow information, 1 including the DIM at sampling and the parity of the animal, and 1 including the DIM at sampling, the parity of the animal, and milk yield features (number of milkings and the minimum, maximum, mean, standard deviation, and sum of milk yield in the period up to the sampling).

Random Sampler. Three sets of milk biomarkers (MBM1, MBM2, and MBM3) were available as predictors to classify the animals according to their metabolic cluster. However, for each of the biomarkers and according to the research protocol, multiple samples at varying DIM were available. A random sampler was created to select 1 sample per animal from 1 to 50 DIM. As such, the sampler works as if an official milk recording organization entered each of the participating research herds and sampled all cows at random stages in lactation on a given day.

Random Forest. In the next step, 10 models were created using the aforementioned random milk recording sampler as a data set and a random forest classifier to predict balanced and imbalanced cows in each run. Each random forest used 1,000 trees and a maximum depth of 5. One-third of the features were selected to be used as candidates for splitting at each tree node. A separate training and test data set was created for balanced and imbalanced cows by randomly splitting the animals by a ratio of 75/25. All features were standardized before entering the model to have a mean of 0 and a standard deviation of 1. Accuracy (% of cows with a correctly predicted metabolic cluster) was selected as the evaluation metric to rank the performance of each of the models. The final minimum, maximum, and average accuracy of the models were reported. A schematic overview of the different steps in the data preparation, model creation and validation, and statistical analysis is given in Figure 1.

RESULTS

Metabolic Clusters

Metabolic clustering based on concentrations of glucose, IGF-I, FFA, and BHB using the k-means method resulted in 3 distinct metabolic clusters on D14 and D35. On D14, glucose concentrations were highest in A, intermediate in B, and lowest in C; IGF-I concentrations were highest in A, intermediate in C, and lowest in B; and FFA and BHB concentrations were lowest in A, intermediate in B, and highest in C (Table 1). On D35, glucose and IGF-I concentrations were higher in

D compared with E and F; FFA concentrations were higher in F compared with D and E; and BHB concentrations were lowest in D, intermediate in E, and highest in F (Table 1).

Production and Metabolic Performance in Balanced Versus Other Balanced Cows

By comparing the metabolic clusters of individual cows on D14 and D35, 43 cows were identified to have a balanced metabolic profile (cluster A on D14 and cluster D on D35; Table 2) during the postpartum period and were grouped together in the balanced group. All other cows ($n = 64$) were grouped in the other balanced group. Glucose and IGF-I concentrations increased toward D35 in both groups and were higher in balanced cows compared with other balanced cows (Table 3). The FFA concentrations decreased toward D35 in both groups and were lower in balanced cows compared with other balanced cows (Table 3). The BHB concentrations were lower in balanced cows compared with other balanced cows but did not change between periods (Table 3). There was no difference in FPCM yield in balanced cows compared with other balanced cows (Figure 2; Table 4). The DMI and the daily increase in DMI were higher in balanced cows (Figure 2; Table 4). The decrease in BW was more pronounced in other balanced cows. Energy balance was continuously higher in balanced cows compared with other balanced cows (Figure 2; Table 4). The BCS decreased from D14 to D35 in other balanced cows but not in balanced cows (Table 3).

Production and Metabolic Performance in Imbalanced Versus Other Imbalanced Cows

By comparing the metabolic clusters of individual cows on D14 and D35, 19 cows were identified as having an imbalanced metabolic profile (5 cows in cluster C on D14 and cluster F on D35 and 14 cows in cluster B on D14 and cluster F on D35; Table 2) during the postpartum period and were grouped together in the imbalanced group. All other cows ($n = 88$) were grouped in the other imbalanced group. Glucose concentrations were lower in imbalanced cows and were not different on D14 compared with D35 (Table 5). The IGF-I concentrations were lower in imbalanced cows compared with other imbalanced cows and increased in both groups on D35 (Table 5). The FFA and BHB concentrations were higher in imbalanced cows in both periods. The FFA concentrations decreased on D35 compared with D14, whereas BHB concentrations were not different between periods (Table 5). The FPCM yield of imbalanced cows was higher compared with

Table 1. Glucose, IGF-I, free fatty acid (FFA), and BHB concentrations in metabolic clusters¹ around 14 DIM (D14; clusters A, B, and C) and 35 DIM (D35; clusters D, E, and F)²

Item	D14			D35		
	A	B	C	D	E	F
Glucose (mM)	3.59 ± 0.04 ^a	3.36 ± 0.04 ^b	2.60 ± 0.10 ^c	3.83 ± 0.04 ^a	3.46 ± 0.06 ^b	3.35 ± 0.06 ^b
IGF-I* (ng/mL)	115.35 ± 6.66 ^a	43.26 ± 2.52 ^b	63.22 ± 10.95 ^b	136.48 ± 6.77 ^a	49.44 ± 3.77 ^b	59.12 ± 4.60 ^b
FFA** (mM)	0.46 ± 0.03 ^a	0.87 ± 0.06 ^b	1.47 ± 0.32 ^b	0.33 ± 0.03 ^a	0.26 ± 0.03 ^a	0.77 ± 0.11 ^b
BHB (mM)	0.46 ± 0.02 ^a	0.65 ± 0.03 ^b	2.19 ± 0.32 ^c	0.43 ± 0.02 ^a	0.52 ± 0.03 ^b	1.14 ± 0.07 ^c

^{a-c}Least squares means with different superscripts within the same time period differ ($P < 0.05$).

¹Clusters were created using the k-means clustering method based on the log-transformed and standardized concentrations of glucose, IGF-I, FFA, and BHB.

²Data are presented as LSM ± SEM.

***Trend for a difference between B and C at * $P = 0.10$ or ** $P = 0.06$.

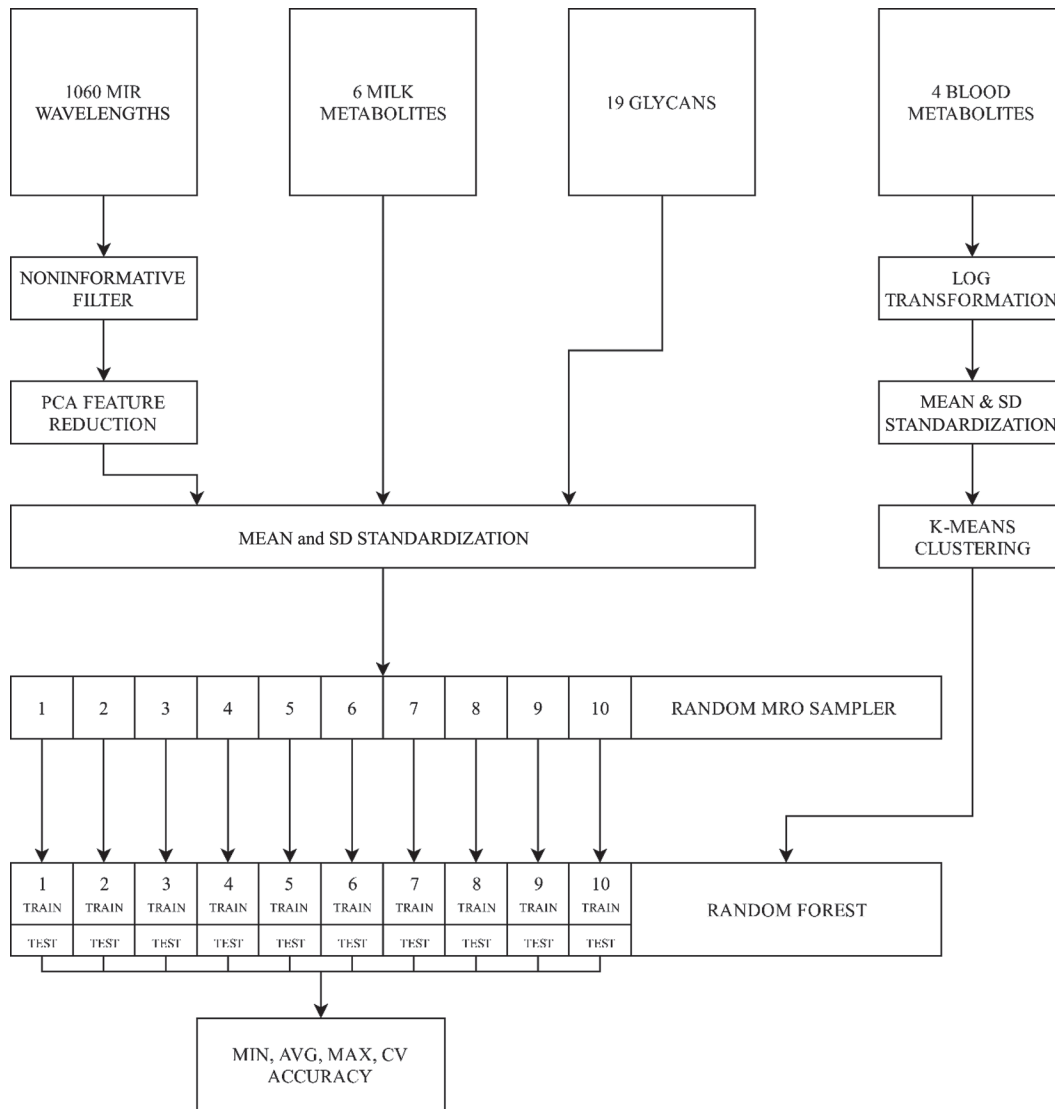


Figure 1. Statistical prediction methodology for metabolic clustering in dairy cows. MRO = milk recording organization; MIN = minimum; AVG = average; MAX = maximum.

Table 2. Number of cows in metabolic clusters A, B, and C around 14 DIM (D14) and metabolic clusters D, E, and F around 35 DIM (D35)¹

D14 cluster	D35 cluster			Total
	D	E	F	
A	43	3	5	51
B	15	22	14	51
C	0	0	5	5
Total	58	25	24	107

¹Clusters were created using the k-means clustering method based on the log-transformed and standardized concentrations of glucose, IGF-I, FFA, and BHB.

that of other imbalanced cows (Figure 3; Table 6). The DMI and the daily increase in DMI were lower in imbalanced cows (Figure 3; Table 6). The decrease in BW was more pronounced in imbalanced cows. Energy balance was continuously lower in imbalanced cows (Figure 3; Table 6). Imbalanced cows had a higher BCS and the decrease in BCS toward D35 was more pronounced compared with other imbalanced cows (Table 5).

Prediction of Metabolic Clusters

A total of 5,400 cross-validation models were summarized in Figures 4 and 5. The average prediction accuracy of balanced cows was highest (76%) for FT-MIR spectra. The average prediction of balanced cows outperformed the prediction of the separate metabolic clusters on D14 and D35. Overall, the highest accuracy was found at 80% accuracy using milk metabolites and enzymes combined with DIM and parity, followed by FT-MIR spectra combined with DIM, parity, and milk yield features (79% accuracy). The lowest coefficient of variation in accuracy between the different runs was

found in FT-MIR spectra predictions with DIM and parity.

The average prediction accuracy of imbalanced cows was highest (81%) for metabolites and enzymes with DIM and parity. In contrast to the balanced cows, the prediction of imbalanced cows outperformed the prediction of the separate metabolic clusters on D14 and D35 except for milk metabolites and enzymes combined with DIM, parity, and milk yield features; IgG glycans combined with DIM, parity, and milk yield features; and predictions using FT-MIR spectra. The highest accuracy was found at 88% accuracy using metabolites and enzymes in combination with DIM and parity or FT-MIR spectra (87% accuracy).

DISCUSSION

This research is part of an European Union-funded project “Genotype plus Environment” (GplusE) aiming to identify novel milk-based phenotypes that may be used as predictors for health traits in dairy cows (Crowe et al., 2018). The present study describes how milk-based phenotypes are linked with clustering of cows in metabolic balance and imbalance and the relationship with production performance parameters.

Different metabolites and hormones can be used to characterize the metabolism of cows in early lactation. The altered metabolic environment in the periparturient period increases the rate of adipose tissue lipolysis. A certain level of FFA is considered normal and necessary. However, FFA concentrations increase when lipolysis is excessive; this is associated with a detrimental effect on immunity, metabolism, and milk production (Roche et al., 2013; Contreras et al., 2018). Part of the circulating FFA are converted into ketone

Table 3. Glucose, IGF-I, free fatty acid (FFA), and BHB concentrations in the balanced and other balanced metabolic clusters¹ around 14 DIM (D14) and 35 DIM (D35)²

Item	D14		D35		Cluster	P-value	Cluster × period
	Balanced	Other balanced	Balanced	Other balanced			
Glucose ³ (mM)	3.62 ± 0.05 ^a	3.28 ± 0.04 ^b	3.84 ± 0.05 ^a	3.49 ± 0.04 ^b	<0.0001	<0.0001	NS
IGF-I ³ (ng/mL)	121.66 ± 7.60 ^a	49.63 ± 2.55 ^b	152.93 ± 9.55 ^a	62.39 ± 3.24 ^b	<0.0001	<0.0001	NS
FFA ⁴ (mM)	0.52 ± 0.05 ^a	0.81 ± 0.06 ^b	0.29 ± 0.03 ^a	0.45 ± 0.03 ^b	<0.001	<0.0001	NS
BHB (mM)	0.44 ± 0.03 ^a	0.71 ± 0.04 ^b	0.42 ± 0.03 ^a	0.69 ± 0.04 ^b	<0.0001	0.42	NS
BCS ⁵	2.75 ± 0.06	2.87 ± 0.05	2.71 ± 0.05	2.71 ± 0.05	0.43	<0.0001	<0.05

^{a,b}Least squares means with different superscripts within the same period differ (P < 0.05).

¹Three clusters were created in both periods: A (n = 51), B (n = 51), and C (n = 5) on D14 and D (n = 58), E (n = 25), and F (n = 24) on D35. Cows in cluster A on D14 and cluster D on D35 were considered to be metabolically balanced (n = 43; “balanced” cluster). Cows not in cluster A on D14 and cluster D on D35 were grouped together (n = 64; “other balanced” cluster).

²Data are presented as LSM ± SEM.

³Period effect indicates an increase in glucose and IGF-I concentrations on D35 compared with D14 in both clusters.

⁴Period effect indicates a decrease in FFA concentrations on D35 compared with D14 in both clusters.

⁵Cluster × period interaction effect indicates a decrease in BCS on D35 compared with D14 in other balanced cows but not in balanced cows.

bodies by the liver. Excessive production of ketone bodies may lead to clinical or subclinical ketosis (Roche et al., 2013). Glucose concentrations are tightly controlled by homeostatic mechanisms (De Koster and Opsomer, 2013). Cows suffering from severe negative EB have lower glucose concentrations (Wathes et al., 2011). The usefulness of glucose as a single indicator of metabolic imbalance has been questioned (Mulligan et al., 2006). However, studies by Bjerre-Harpøth et al. (2012) and Moyes et al. (2013) identified glucose as an important metabolite to be included in an index of metabolic imbalance together with FFA and BHB. Late pregnancy and early lactation are marked by a decrease in the concentration of IGF-I. The nadir IGF-I concentration is reached in the first week after calving (Butler et al.,

2003; Radcliff et al., 2003); after the second week, IGF-I concentrations are markedly influenced by the energy status of the cows (Fenwick et al., 2008; Wathes et al., 2011). Animals with severe negative EB have lower IGF-I concentrations (Fenwick et al., 2008), and IGF-I has been suggested to be an indicator of nutritional status (Cohick, 1998; Zulu et al., 2002).

Metabolic imbalance is defined as “a condition where the regulating mechanisms are insufficient for the animals to function optimally leading to a high risk of a complex of digestive, metabolic, and infectious problems” (Ingvarsen, 2006). Metabolic clustering of dairy cows based on concomitant changes in the concentration of different metabolites improves the identification of metabolically balanced cows compared with the use

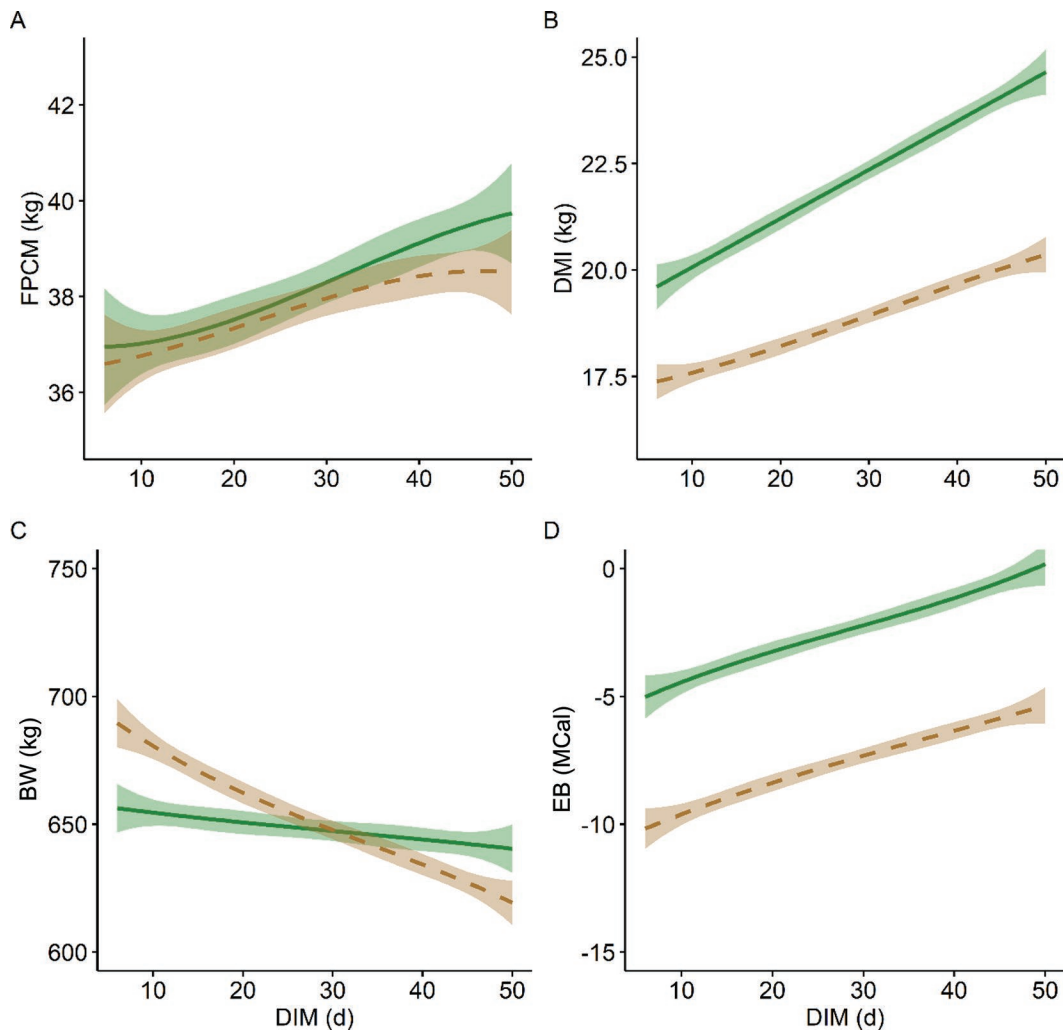


Figure 2. Fat- and protein-corrected milk (FPCM) yield (A), DMI (B), BW (C), and energy balance (EB; D) in balanced (green, continuous line) and other balanced (orange, dashed line) metabolic clusters. Three clusters were created in both periods: A ($n = 51$), B ($n = 51$), and C ($n = 5$) on D14 (around 14 DIM) and D ($n = 58$), E ($n = 25$), and F ($n = 24$) on D35 (around 35 DIM). Cows in cluster A on D14 and cluster D on D35 were considered to be metabolically balanced ($n = 43$; “balanced” cluster). Cows not in cluster A on D14 and cluster D on D35 were grouped together ($n = 64$; “other balanced” cluster). Lines represent the LSM of the models, and the colored areas represent the 95% confidence limits.

Table 4. Fat- and protein-corrected milk (FPCM) yield, DMI, BW, and energy balance (EB) in balanced and other balanced metabolic clusters¹ and the effect of day after calving (slope DIM) in the postpartum period

Item	Cluster		P-values		
	Balanced	Other balanced	DIM	Cluster	DIM × cluster
FPCM (kg/d)					
LSM ± SEM	38.42 ± 0.99	37.76 ± 0.80	<0.05	0.60	NS
Slope DIM ± SEM	0.043 ± 0.018	0.043 ± 0.018			
DMI (kg/d)					
LSM ± SEM	22.12 ± 0.48 ^a	18.72 ± 0.39 ^b	<0.0001	<0.01	<0.01
Slope DIM ± SEM	0.11 ± 0.01 ^a	0.06 ± 0.01 ^b			
BW (kg)					
LSM ± SEM	647.97 ± 9.56	650.51 ± 7.66	<0.0001	<0.05	<0.0001
Slope DIM ± SEM	-0.34 ± 0.14 ^a	-1.41 ± 0.11 ^b			
EB (Mcal/d)					
LSM ± SEM	-2.40 ± 0.71 ^a	-7.59 ± 0.61 ^b	<0.0001	<0.0001	NS
Slope DIM ± SEM	0.11 ± 0.01	0.11 ± 0.01			

^{a,b}Least squares means and slopes with different superscripts differ ($P < 0.05$).

¹Three clusters were created in both periods: A (n = 51), B (n = 51), and C (n = 5) on D14 and D (n = 58), E (n = 25), and F (n = 24) on D35. Cows in cluster A on D14 and cluster D on D35 were considered to be metabolically balanced (n = 43; “balanced” cluster). Cows not in cluster A on D14 and cluster D on D35 were grouped together (n = 64; “other balanced” cluster).

of a single indicator (Bjerre-Harpøth et al., 2012; Moyes et al., 2013; Grelet et al., 2018). Metabolic clustering using the k-means method in the present study resulted in 3 distinct metabolic groups in early and peak lactation. Based on the comparison of the glucose, IGF-I, FFA, and BHB concentrations between the metabolic groups, groups A and D showed a balanced metabolic profile on D14 and D35, respectively (high glucose, high IGF-I, low FFA, and low BHB), whereas groups B and C on D14 and group F on D35 showed an imbalanced metabolic profile (low glucose, low IGF-I, high FFA, or high BHB). Dairy cows in cluster A on D14 and cluster D on D35 were identified as metabolically balanced

throughout the study and had higher glucose and IGF-I and lower FFA and BHB compared with other cows in both early and peak lactation. Dairy cows in cluster C on D14 and cluster F on D35 together with cows in cluster B on D14 and cluster F on D35 were identified as metabolically imbalanced throughout the study and had lower glucose and IGF-I and higher FFA and BHB compared with other cows on D14 and D35.

The different metabolic clusters were characterized by differences in FPCM, DMI, BW, BCS, and EB in the postpartum period. The relationship between metabolic clusters and FPCM was weak. The FPCM yield was not different in balanced cows. Metabolically

Table 5. Glucose, IGF-I, free fatty acid (FFA), and BHB concentrations in imbalanced and other imbalanced metabolic clusters¹ around 14 DIM (D14) and 35 DIM (D35)²

Item	D14		D35		P-value		
	Imbalanced	Other imbalanced	Imbalanced	Other imbalanced	Cluster	Period	Cluster × period
Glucose ³ (mM)	3.19 ± 0.06 ^a	3.49 ± 0.04 ^b	3.38 ± 0.07 ^a	3.69 ± 0.04 ^b	<0.0001	<0.0001	NS
IGF-I ⁴ (ng/mL)	43.61 ± 5.53 ^a	78.56 ± 4.73 ^b	55.07 ± 6.98 ^a	99.21 ± 6.01 ^b	<0.0001	<0.0001	NS
FFA ⁵ (mM)	1.32 ± 0.17 ^a	0.56 ± 0.04 ^b	0.76 ± 0.10 ^a	0.32 ± 0.02 ^b	<0.0001	<0.0001	NS
BHB (mM)	1.15 ± 0.09 ^a	0.50 ± 0.02 ^b	1.11 ± 0.09 ^a	0.49 ± 0.02 ^b	<0.0001	0.39	NS
BCS ⁶	3.16 ± 0.09 ^a	2.75 ± 0.04 ^b	2.89 ± 0.09 ^a	2.66 ± 0.04 ^b	<0.0001	<0.0001	<0.001

^{a,b}Least squares means with different superscripts within the same period differ ($P < 0.05$).

¹Three clusters were created in both periods: A (n = 51), B (n = 51), and C (n = 5) on D14 and D (n = 58), E (n = 25), and F (n = 24) on D35. Cows in cluster B on D14 and cluster F on D35 (n = 14) together with cows in cluster C on D14 and cluster F on D35 (n = 5) were considered to be metabolically imbalanced (n = 19; “imbalanced” cluster). Cows not in these clusters on D14 or D35 were grouped together (n = 88; “other imbalanced” cluster).

²Data are presented as LSM ± SEM.

³Glucose concentrations were not different between periods for imbalanced cows and increased on D35 for other imbalanced cows.

⁴Period effect indicates an increase in IGF concentrations on D35 compared with D14 in both clusters.

⁵Period effect indicates a decrease in FFA concentrations on D35 compared with D14 in both clusters.

⁶Period effect indicates a decrease in BCS on D35 compared with D14 in both clusters. There was a more pronounced decrease in BCS from D14 to D35 in imbalanced cows compared with other imbalanced cows.

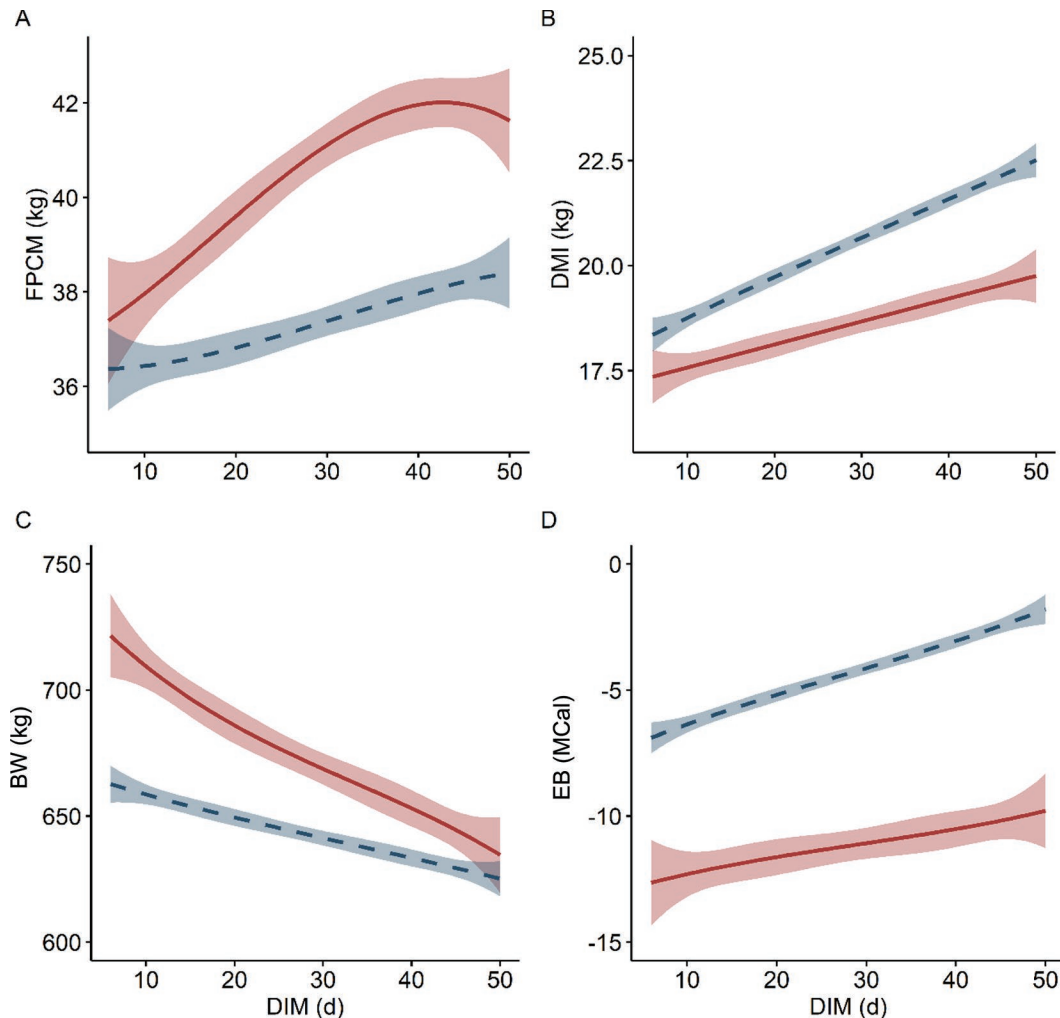


Figure 3. Fat- and protein-corrected milk (FPCM) yield (A), DMI (B), BW (C), and energy balance (EB; D) in imbalanced (red, continuous line) and other imbalanced (blue, dashed line) metabolic clusters. Three clusters were created in both periods: A ($n = 51$), B ($n = 51$), and C ($n = 5$) on D14 (around 14 DIM) and D ($n = 58$), E ($n = 25$), and F ($n = 24$) on D35 (around 35 DIM). Cows in cluster B on D14 and cluster F on D35 ($n = 14$) together with cows in cluster C on D14 and cluster F on D35 ($n = 5$) were considered to be metabolically imbalanced ($n = 19$; “imbalanced” cluster). Cows not in these clusters on D14 or D35 were grouped together ($n = 88$; “other imbalanced” cluster). Lines represent the LSM of the models, and the colored areas represent the 95% confidence limits.

imbalanced cows had higher FPCM yield. The weak relationship between milk yield, metabolic balance, and disorders has been described by Ingvarstsen et al. (2003), Ingvarstsen (2006), and Bjerre-Harpøth et al. (2012). Ingvarstsen (2006) stated that it is not milk yield as such that is the cause of metabolic imbalance or disorders but rather the individual cow's inability to cope with the metabolic challenges of early lactation. Dry matter intake was consistently higher in metabolically balanced cows and lower in imbalanced cows. Daily increase in DMI was higher in cows with a favorable metabolic profile (balanced and other imbalanced). The relationship between DMI and metabolism may be explained by the fact that certain metabolites (FFA) may regu-

late feed intake in ruminants by the hepatic oxidation of these metabolites, thereby causing a satiety signal and depressing feed intake (Ingvarstsen and Andersen, 2000; Allen et al., 2009). Energy balance is influenced mainly by energy intake and less by milk production (Santos et al., 2010). The improved nutritional status in balanced and other imbalanced cows had a positive effect on the EB of the animals, whereas imbalanced cows were in severe negative EB throughout the study. Body weight is affected by different factors in postpartum cows: frame size, DMI, stage of lactation, and EB. As such, BW is not a good predictor for the severity of the negative EB or the mobilization of energy (Schröder and Staufenbiel, 2006). However, automated

daily measurement of BW has been used to assess the energy status of dairy cows (van Straten et al., 2008; Thorup et al., 2012). The more pronounced decrease in BW in other balanced and imbalanced cows compared with balanced and other imbalanced cows, respectively, can be explained by a higher degree of mobilization of energy reserves due to the severe negative EB, a lower gut fill due to the decreased DMI, or a combination of both factors. Body condition score is a subjective indicator for the mobilization of subcutaneous adipose tissue. In particular, the postpartum decrease in BCS is associated with metabolic and infectious disorders (Roche et al., 2013). Cows with an unfavorable metabolic profile (other balanced and imbalanced) had a more pronounced decrease in BCS from D14 to D35, indicative of a higher degree of body fat mobilization.

The current study furthermore focused on comparing multiple biomarkers to predict the metabolic clusters as a novel phenotypic trait in early-lactating dairy cows. To our knowledge, this is the first study to position the different sets of milk biomarkers relatively across one another according to their predictive accuracy. The current scope of the study focused on comparing different milk biomarkers rather than the individual fine-tuning of the individual biomarkers to increase the predictive accuracy as described by Grelet et al. (2018) for FT-MIR and metabolic clustering. From our results, it can be noted that within each set of biomarkers, adding DIM, parity, and milk yield features did not improve the predictive accuracy much. Hence, it is debatable whether collecting such information is worth the effort

compared with the relatively small gain in predictive accuracy. Overall, we were able to predict balanced and imbalanced animals in early lactation with varying prediction accuracy across the 3 sets of biomarkers. Prediction of imbalanced cows was more accurate across all milk biomarkers compared with balanced cows, where FT-MIR outperformed the other milk biomarkers. Further efforts are made within the GplusE project for industry-wide application of the metabolic clustering technique. External validation of the balanced cows can lead to establishment of a novel phenotypic trait for genetic selection as suggested by Egger-Danner et al. (2015) and Crowe et al. (2018). External validation of the imbalanced cows can help identify cows for specific management strategies, such as elective propylene glycol treatment versus group treatment, as proposed by others (Lomander et al., 2012; Jenkins et al., 2015).

CONCLUSIONS

The k-means clustering of blood metabolites was found to effectively identify balanced and imbalanced cows across the participating countries within the GplusE project. Furthermore, production parameters revealed marked differences in DMI and energy balance, underlining the phenotypic validity of metabolic clusters. Finally, prediction using both FT-MIR or milk metabolites and enzymes allows implementation of metabolic clusters across larger cow numbers as a novel trait for genetic selection or identification of imbalanced early-lactating dairy cows.

Table 6. Fat- and protein-corrected milk (FPCM) yield, DMI, BW, and energy balance (EB) in imbalanced and other imbalanced metabolic clusters¹ and the effect of day after calving (slope DIM) in the postpartum period

Item	Cluster		P-value		
	Imbalanced	Other imbalanced	DIM	Cluster	DIM × cluster
FPCM (kg/d)					
LSM ± SEM	41.14 ± 1.47 ^a	37.29 ± 0.66 ^b	<0.01	<0.05	NS
Slope DIM ± SEM	0.049 ± 0.018	0.049 ± 0.018			
DMI (kg/d)					
LSM ± SEM	18.56 ± 0.77 ^a	20.51 ± 0.35 ^b	<0.0001	0.37	0.05
Slope DIM ± SEM	0.05 ± 0.02 ^a	0.10 ± 0.01 ^b			
BW (kg)					
LSM ± SEM*	672.29 ± 14.12	642.73 ± 6.42	<0.0001	<0.01	<0.001
Slope DIM ± SEM	-1.69 ± 0.22 ^a	-0.80 ± 0.10 ^b			
EB (Mcal/d)					
LSM ± SEM	-11.14 ± 1.25 ^a	-4.32 ± 0.50 ^b	<0.0001	<0.0001	NS
Slope DIM ± SEM	0.11 ± 0.01	0.11 ± 0.01			

^{a,b}Least squares means and slope with different superscripts differ (*P* < 0.05).

¹Three clusters were created in both periods: A (n = 51), B (n = 51), and C (n = 5) on D14 and D (n = 58), E (n = 25), and F (n = 24) on D35. Cows in cluster B on D14 and cluster F on D35 (n = 14) together with cows in cluster C on D14 and cluster F on D35 (n = 5) were considered to be metabolically imbalanced (n = 19; “imbalanced” cluster). Cows not in these clusters on D14 or D35 were grouped together (n = 88; “other imbalanced” cluster).

*Trend for a difference between imbalanced and other imbalanced cows (*P* = 0.06).

Milk biomarker predictor set	Class	Run ₁	Run ₂	Run ₃	Run ₄	Run ₅	Run ₆	Run ₇	Run ₈	Run ₉	Run ₁₀	Mean	Min	Max	CV
Metabolites and enzymes	EARLY and PEAK	70	74	63	70	64	67	58	60	70	64	66	58	74	8
	EARLY	69	59	68	54	68	63	62	65	57	61	63	54	69	8
	PEAK	60	74	60	61	58	61	71	56	65	47	61	47	74	12
+ DIM & Parity	EARLY and PEAK	68	63	70	71	71	72	58	72	80	71	69	58	80	9
	EARLY	67	59	66	72	65	65	59	71	65	64	65	59	72	7
	PEAK	69	64	71	63	59	60	67	59	60	60	63	59	71	7
+ DIM & Parity + Milk yield features	EARLY and PEAK	63	70	61	72	67	73	61	65	51	70	65	51	73	10
	EARLY	57	60	75	68	70	59	66	67	60	65	65	57	75	9
	PEAK	73	68	72	64	64	58	55	59	57	68	64	55	73	10
IgG Glycans	EARLY and PEAK	67	52	66	59	57	56	68	73	53	58	61	52	73	12
	EARLY	43	43	47	50	52	43	45	39	42	50	45	39	52	9
	PEAK	68	58	66	58	58	47	52	67	58	66	60	47	68	12
+ DIM & Parity	EARLY and PEAK	63	53	73	58	63	61	61	64	55	51	60	51	73	11
	EARLY	50	62	49	52	51	40	41	62	58	42	51	40	62	16
	PEAK	55	59	57	47	57	61	63	62	55	62	58	47	63	8
+ DIM & Parity + Milk yield features	EARLY and PEAK	69	64	57	62	60	62	67	50	65	63	62	50	69	9
	EARLY	45	55	37	56	42	52	45	58	56	55	50	37	58	15
	PEAK	37	65	54	56	60	66	63	64	65	60	59	37	66	15
FT-MIR spectra	EARLY and PEAK	67	76	72	70	77	73	78	78	68	75	73	67	78	5
	EARLY	70	68	68	64	71	70	73	68	63	67	68	63	73	5
	PEAK	65	71	63	62	64	65	66	80	69	66	67	62	80	8
+ DIM & Parity	EARLY and PEAK	76	77	70	78	77	74	78	78	74	74	76	70	78	4
	EARLY	68	70	59	67	61	79	71	70	74	70	69	59	79	8
	PEAK	68	67	71	69	47	63	63	65	70	64	65	47	71	10
+ DIM & Parity + Milk yield features	EARLY and PEAK	76	76	69	65	79	76	69	73	72	74	73	65	79	6
	EARLY	71	67	72	74	63	66	66	65	71	61	68	61	74	6
	PEAK	61	57	63	56	64	67	65	66	65	76	64	56	76	9

Figure 4. Prediction accuracy within each run and summary for different classification models to predict balanced cows in the early (14 DIM) and peak (35 DIM), early, or peak periods. Three clusters were created in both periods: A (n = 51), B (n = 51), and C (n = 5) on D14 and D (n = 58), E (n = 25), and F (n = 24) on D35 (around 35 DIM). Cows in cluster A on D14 (around 14 DIM) and cluster D on D35 were considered to be metabolically balanced (n = 43; “balanced” cluster). FT-MIR = Fourier-transform mid-infrared; Min = minimum; Max = maximum.

Milk biomarker predictor set	Class	Run ₁	Run ₂	Run ₃	Run ₄	Run ₅	Run ₆	Run ₇	Run ₈	Run ₉	Run ₁₀	Mean	Min	Max	CV
Metabolites and enzymes	EARLY and PEAK	81	82	84	77	71	79	80	81	83	79	80	71	84	5
	EARLY	79	82	80	79	81	77	81	80	80	81	80	77	82	2
	PEAK	85	85	75	85	76	71	81	82	79	83	80	71	85	6
+ DIM & Parity	EARLY and PEAK	81	87	76	80	87	88	77	76	76	78	81	76	88	6
	EARLY	81	80	84	75	80	84	76	78	79	83	80	75	84	4
	PEAK	83	83	76	76	77	76	81	80	81	79	79	76	83	4
+ DIM & Parity + Milk yields features	EARLY and PEAK	87	78	82	79	79	82	79	79	80	80	80	78	87	3
	EARLY	84	83	76	78	82	86	79	82	83	72	81	72	86	5
	PEAK	81	88	79	87	80	78	87	77	80	81	82	77	88	5
IgG Glycans	EARLY and PEAK	81	78	83	80	80	81	82	84	70	80	80	70	84	5
	EARLY	73	71	84	75	77	71	81	74	80	78	76	71	84	6
	PEAK	76	73	79	74	73	71	80	73	74	80	75	71	80	4
+ DIM & Parity	EARLY and PEAK	81	82	78	79	82	79	74	77	81	80	79	74	82	3
	EARLY	80	77	75	81	72	77	79	79	72	81	77	72	81	4
	PEAK	75	75	73	72	75	76	77	80	77	77	76	72	80	3
+ DIM & Parity + Milk yields features	EARLY and PEAK	81	79	76	80	81	76	68	76	78	79	77	68	81	5
	EARLY	76	82	79	78	77	81	82	76	82	79	79	76	82	3
	PEAK	63	72	69	74	71	76	74	68	70	78	72	63	78	6
FT-MIR Spectra	EARLY and PEAK	84	81	78	79	76	87	76	81	83	79	80	76	87	4
	EARLY	77	84	83	81	86	84	76	86	87	79	82	76	87	5
	PEAK	69	72	79	76	73	75	76	85	72	78	76	69	85	6
+ DIM & Parity	EARLY and PEAK	78	78	73	74	77	81	73	82	80	78	77	73	82	4
	EARLY	85	87	86	80	82	84	86	80	87	85	84	80	87	3
	PEAK	73	74	77	79	76	76	73	74	79	80	76	73	80	3
+ DIM & Parity + Milk yields features	EARLY and PEAK	80	77	71	75	73	87	86	82	84	79	79	71	87	7
	EARLY	86	87	84	80	87	83	81	87	85	85	84	80	87	3
	PEAK	69	70	73	76	78	77	73	80	73	76	74	69	80	5

Figure 5. Prediction accuracy within each run and summary for different classification models to predict imbalanced cows in the early (14 DIM) and peak (35 DIM), early, or peak periods. Three clusters were created in both periods: A (n = 51), B (n = 51), and C (n = 5) on D14 (around 14 DIM) and D (n = 58), E (n = 25), and F (n = 24) on D35 (around 35 DIM). Cows in cluster B on D14 and cluster F on D35 (n = 14) together with cows in cluster C on D14 and cluster F on D35 (n = 5) were considered to be metabolically imbalanced (n = 19; “imbalanced” cluster). FT-MIR = Fourier-transform mid-infrared; Min = minimum; Max = maximum.

ACKNOWLEDGMENTS

This project received funding from the European Union's Seventh Framework Programme (Brussels, Belgium) 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.

REFERENCES

- Allen, M. S., B. J. Bradford, and M. Oba. 2009. Board-invited review: The hepatic oxidation theory of the control of feed intake and its application to ruminants. *J. Anim. Sci.* 87:3317–3334.
- Bauman, D. E. 1999. Bovine somatotropin and lactation: From basic science to commercial application. *Domest. Anim. Endocrinol.* 17:101–116.
- Bauman, D. E., and W. B. Currie. 1980. Partitioning of nutrients during pregnancy and lactation—A review of mechanisms involving homeostasis and homeorhesis. *J. Dairy Sci.* 63:1514–1529.
- Bell, A. W., and D. E. Bauman. 1997. Adaptations of glucose metabolism during pregnancy and lactation. *J. Mammary Gland Biol. Neoplasia* 2:265–278.
- Beltman, M. E., N. Forde, P. Furney, F. Carter, J. F. Roche, P. Lonergan, and M. A. Crowe. 2010. Characterisation of endometrial gene expression and metabolic parameters in beef heifers yielding viable or non-viable embryos on Day 7 after insemination. *Reprod. Fertil. Dev.* 22:987–999.
- Bjerre-Harpøth, V., N. C. Friggens, V. M. Thorup, T. Larsen, B. M. Damgaard, K. L. Ingvarstsen, and K. M. Moyes. 2012. Metabolic and production profiles of dairy cows in response to decreased nutrient density to increase physiological imbalance at different stages of lactation. *J. Dairy Sci.* 95:2362–2380.
- Butler, S. T., A. L. Marr, S. H. Pelton, R. P. Radcliff, M. C. Lucy, and W. R. Butler. 2003. Insulin restores GH responsiveness during lactation-induced negative energy balance in dairy cattle: Effects on expression of IGF-I and GH receptor 1A. *J. Endocrinol.* 176:205–217.
- Cohick, W. S. 1998. Role of the insulin-like growth factors and their binding proteins in lactation. *J. Dairy Sci.* 81:1769–1777.
- 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.
- Crowe, M. A., M. Hostens, and G. Opsomer. 2018. Reproductive management in dairy cows—The future. *Ir. Vet. J.* 71:1.
- CVB. 2007. Tabellenboek Veevoeding 2007. Centraal Veevoederbureau Lelystad, Lelystad, the Netherlands.
- De Koster, J. D., and G. Opsomer. 2013. Insulin resistance in dairy cows. *Vet. Clin. North Am. Food Anim. Pract.* 29:299–322.
- 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.
- Edmonson, A. J., I. J. Lean, L. D. Weaver, T. Farver, and G. Webster. 1989. A body condition scoring chart for Holstein dairy cows. *J. Dairy Sci.* 72:68–78.
- Egger-Danner, C., J. B. Cole, J. E. Pryce, N. Gengler, B. Heringstad, A. Bradley, and K. F. Stock. 2015. Invited review: Overview of new traits and phenotyping strategies in dairy cattle with a focus on functional traits. *Animal* 9:191–207.
- Fenwick, M. A., R. Fitzpatrick, D. A. Kenny, M. G. Diskin, J. Patton, J. J. Murphy, and D. C. Wathes. 2008. Interrelationships between negative energy balance (NEB) and IGF regulation in liver of lactating dairy cows. *Domest. Anim. Endocrinol.* 34:31–44.
- Foldager, L., M. Hostens, M. Salavati, C. Grelet, M. T. Sorensen, D. C. Wathes, C. P. Ferris, C. Marchitelli, F. Becker, T. Larsen, E. Matthews, R. O'Flaherty, G. Opsomer, N. Gengler, M. A. Crowe, and E. Gplus. Consortium, and K. L. Ingvarstsen. 2019. Predicting risk of physiological imbalance in Holstein dairy cows from milk biomarkers. *Animal* (in review).
- Grelet, C., C. Bastin, M. Gele, J. B. Daviere, M. Johan, A. Werner, R. Reding, J. A. F. Pierna, F. G. Colinet, P. Dardenne, N. Gengler, H. Soyeurt, and F. Dehareng. 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, beta-hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. *J. Dairy Sci.* 99:4816–4825.
- Grelet, C., J. A. F. Pierna, P. Dardenne, V. Baeten, and F. Dehareng. 2015. Standardization of milk mid-infrared spectra from a European dairy network. *J. Dairy Sci.* 98:2150–2160.
- Grelet, C., A. Vanlierde, M. Hostens, L. Foldager, M. Salavati, K. Ingvarstsen, M. Crowe, M. Sorensen, E. Froidmont, and C. Ferris. 2018. Potential of milk mid-IR spectra to predict metabolic status of cows through blood components and an innovative clustering approach. *Animal* <https://doi.org/10.1017/S1751731118001751>.
- Grummer, R. R., D. G. Mashek, and A. Hayirli. 2004. Dry matter intake and energy balance in the transition period. *Vet. Clin. North Am. Food Anim. Pract.* 20:447–470.
- Ingvarstsen, 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.
- Ingvarstsen, 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.
- Ingvarstsen, 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.
- Ingvarstsen, K. L., and N. C. Friggens. 2005. To what extent do variabilities in hormones, metabolites and energy intake explain variability in milk yield? *Domest. Anim. Endocrinol.* 29:294–304.
- Jenkins, N. T., G. Pena, C. Risco, C. C. Barbosa, A. Vieira-Neto, and K. N. Galvao. 2015. Utility of inline milk fat and protein ratio to diagnose subclinical ketosis and to assign propylene glycol treatment in lactating dairy cows. *Can. Vet. J.* 56:850–854.
- Larsen, T. 2005. Determination of lactate dehydrogenase (LDH) activity in milk by a fluorometric assay. *J. Dairy Res.* 72:209–216.
- Larsen, T. 2014. Fluorometric determination of free and total isocitrate in bovine milk. *J. Dairy Sci.* 97:7498–7504.
- Larsen, T. 2015. Fluorometric determination of free glucose and glucose 6-phosphate in cows' milk and other opaque matrices. *Food Chem.* 166:283–286.
- Larsen, T., and N. I. Nielsen. 2005. Fluorometric determination of beta-hydroxybutyrate in milk and blood plasma. *J. Dairy Sci.* 88:2004–2009.
- Larsen, T., C. M. Rontved, K. L. Ingvarstsen, L. Vels, and M. Bjerring. 2010. Enzyme activity and acute phase proteins in milk utilized as indicators of acute clinical *E. coli* LPS-induced mastitis. *Animal* 4:1672–1679.
- Lomander, H., J. Frossling, K. L. Ingvarstsen, H. Gustafsson, and C. Svensson. 2012. Supplemental feeding with glycerol or propylene glycol of dairy cows in early lactation—Effects on metabolic status, body condition, and milk yield. *J. Dairy Sci.* 95:2397–2408.
- Lucy, M. C. 2008. Functional differences in the growth hormone and insulin-like growth factor axis in cattle and pigs: Implications for post-partum nutrition and reproduction. *Reprod. Domest. Anim.* 43:31–39.
- Lucy, M. C., G. A. Verkerk, B. E. Whyte, K. A. Macdonald, L. Burton, R. T. Cursons, J. R. Roche, and C. W. Holmes. 2009. Somatotrophic axis components and nutrient partitioning in genetically diverse dairy cows managed under different feed allowances in a pasture system. *J. Dairy Sci.* 92:526–539.
- Maury, E., K. Ehala-Aleksejev, Y. Guiot, R. Detry, A. Vandenhooff, and S. M. Brichard. 2007. Adipokines oversecreted by omental adi-

- pose tissue in human obesity. *Am. J. Physiol. Endocrinol. Metab.* 293:E656–E665.
- McArt, J. A. A., D. V. Nydam, G. R. Oetzel, T. R. Overton, and P. A. Ospina. 2013. Elevated non-esterified fatty acids and β -hydroxybutyrate and their association with transition dairy cow performance. *Vet. J.* 198:560–570.
- McCarthy, M. M., S. Mann, D. V. Nydam, T. R. Overton, and J. A. A. McArt. 2015. Short communication: Concentrations of nonesterified fatty acids and β -hydroxybutyrate in dairy cows are not well correlated during the transition period. *J. Dairy Sci.* 98:6284–6290.
- Moyes, K. M., T. Larsen, and K. L. Ingvarsen. 2013. Generation of an index for physiological imbalance and its use as a predictor of primary disease in dairy cows during early lactation. *J. Dairy Sci.* 96:2161–2170.
- Mulligan, F. J., L. O'Grady, D. A. Rice, and M. L. Doherty. 2006. A herd health approach to dairy cow nutrition and production diseases of the transition cow. *Anim. Reprod. Sci.* 96:331–353.
- Nielsen, N. I., N. C. Friggens, M. G. Chagunda, and K. L. Ingvarsen. 2005a. Predicting risk of ketosis in dairy cows using in-line measurements of beta-hydroxybutyrate: A biological model. *J. Dairy Sci.* 88:2441–2453.
- Nielsen, N. I., T. Larsen, M. Bjerring, and K. L. Ingvarsen. 2005b. Quarter health, milking interval, and sampling time during milking affect the concentration of milk constituents. *J. Dairy Sci.* 88:3186–3200.
- NRC. 2001. *Nutrient Requirements of Dairy Cattle*. Vol. 1. 7th ed. Natl. Acad. Press, Washington, DC.
- Ospina, P. A., J. McArt, T. Overton, T. Stokol, and D. Nydam. 2013. Using nonesterified fatty acids and β -hydroxybutyrate concentrations during the transition period for herd-level monitoring of increased risk of disease and decreased reproductive and milking performance. *Vet. Clin. North Am. Food Anim. Pract.* 29:387–412.
- Puppel, K., and B. Kuczynska. 2016. Metabolic profiles of cow's blood; A review. *J. Sci. Food Agric.* 96:4321–4328.
- Radcliff, R. P., B. L. McCormack, B. A. Crooker, and M. C. Lucy. 2003. Plasma hormones and expression of growth hormone receptor and insulin-like growth factor-I mRNA in hepatic tissue of periparturient dairy cows. *J. Dairy Sci.* 86:3920–3926.
- Roche, J. R., J. K. Kay, N. C. Friggens, J. J. Looor, and D. P. Berry. 2013. Assessing and managing body condition score for the prevention of metabolic disease in dairy cows. *Vet. Clin. North Am. Food Anim. Pract.* 29:323–336.
- Santos, J. E., R. S. Bisinotto, E. S. Ribeiro, F. S. Lima, L. F. Greco, C. R. Staples, and W. W. Thatcher. 2010. Applying nutrition and physiology to improve reproduction in dairy cattle. *Soc. Reprod. Fertil. Suppl.* 67:387–403.
- Schröder, U. J., and R. Staufenbiel. 2006. Invited review: Methods to determine body fat reserves in the dairy cow with special regard to ultrasonographic measurement of backfat thickness. *J. Dairy Sci.* 89:1–14.
- Soyeurt, H., F. Dehareng, N. Gengler, S. McParland, E. Wall, D. P. Berry, M. Coffey, and P. Dardenne. 2011. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. *J. Dairy Sci.* 94:1657–1667.
- Stöckmann, H., B. Adamczyk, J. Hayes, and P. M. Rudd. 2013. Automated, high-throughput IgG-antibody glycoprofiling platform. *Anal. Chem.* 85:8841–8849.
- Thorup, V. M., D. Edwards, and N. C. Friggens. 2012. On-farm estimation of energy balance in dairy cows using only frequent body weight measurements and body condition score. *J. Dairy Sci.* 95:1784–1793.
- van Straten, M., N. Y. Shpigel, and M. Friger. 2008. Analysis of daily body weight of high-producing dairy cows in the first one hundred twenty days of lactation and associations with ovarian inactivity. *J. Dairy Sci.* 91:3353–3362.
- Voelker, J. A., and M. S. Allen. 2003. Pelleted beet pulp substituted for high-moisture corn: 3. Effects on ruminal fermentation, pH, and microbial protein efficiency in lactating dairy cows. *J. Dairy Sci.* 86:3562–3570.
- Wallace, T. M., and D. R. Matthews. 2002. The assessment of insulin resistance in man. *Diabet. Med.* 19:527–534.
- Wathes, D. C., Z. Cheng, M. A. Fenwick, R. Fitzpatrick, and J. Patton. 2011. Influence of energy balance on the somatotrophic axis and matrix metalloproteinase expression in the endometrium of the postpartum dairy cow. *Reproduction* 141:269–281.
- Weekes, T. E., Y. Sasaki, and T. Tsuda. 1983. Enhanced responsiveness to insulin in sheep exposed to cold. *Am. J. Physiol.* 244:E335–E345.
- Zhao, F. Q., and A. F. Keating. 2007. Expression and regulation of glucose transporters in the bovine mammary gland. *J. Dairy Sci.* 90(Suppl. 1):E76–E86.
- Zulu, V. C., T. Nakao, and Y. Sawamukai. 2002. Insulin-like growth factor-I as a possible hormonal mediator of nutritional regulation of reproduction in cattle. *J. Vet. Med. Sci.* 64:657–665.