Metabolomics meets machine learning: Longitudinal metabolite profiling in serum of normal versus overconditioned cows and pathway analysis

Morteza H. Ghaffari,[1] Amirhossein Jahanbekam,[2] Hassan Sadri,[3]* Katharina Schuh,[1,4] Georg Dusel,[4] Cornelia Prehn,[5] Jerzy Adamski,[5,6,7] Christian Koch,[8] and Helga Sauerwein[1]†

[1]Institute of Animal Science, Physiology and Hygiene Unit, University of Bonn, 53115 Bonn, Germany
[2]Department of Epileptology, University of Bonn, Bonn 53127, Germany
[3]Department of Clinical Science, Faculty of Veterinary Medicine, University of Tabriz, 516616471 Tabriz, Iran
[4]Department of Life Sciences and Engineering, Animal Nutrition and Hygiene Unit, University of Applied Sciences Bingen, 55411 Bingen am Rhein, Germany
[5]Research Unit Molecular Endocrinology and Metabolism, Genome Analysis Center, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg 85764, Germany
[6]Lehrstuhl für Experimentelle Genetik, Technische Universität München, Freising-Weihenstephan 85350, Germany
[7]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, 8 Medical Drive, Singapore 117597, Singapore
[8]Educational and Research Centre for Animal Husbandry, Hofgut Neumuehle, 67728 Muenchweileran der Alsenz, Germany

ABSTRACT

This study aimed to investigate the differences in the metabolic profiles in serum of dairy cows that were normal or overconditioned when dried off for elucidating the pathophysiological reasons for the increased health disturbances commonly associated with overconditioning. Fifteen weeks antepartum, 38 multiparous Holstein cows were allocated to either a high body condition (**HBCS**; n = 19) group or a normal body condition (**NBCS**; n = 19) group and were fed different diets until dry-off to amplify the difference. The groups were also stratified for comparable milk yields (NBCS: 10,361 ± 302 kg; HBCS: 10,315 ± 437 kg; mean ± standard deviation). At dry-off, the cows in the NBCS group (parity: 2.42 ± 1.84; body weight: 665 ± 64 kg) had a body condition score (**BCS**) <3.5 and backfat thickness (**BFT**) <1.2 cm, whereas the HBCS cows (parity: 3.37 ± 1.67; body weight: 720 ± 57 kg) had BCS >3.75 and BFT >1.4 cm. During the dry period and the subsequent lactation, both groups were fed identical diets but maintained the BCS and BFT differences. A targeted metabolomics (AbsoluteIDQ p180 kit, Biocrates Life Sciences AG, Innsbruck, Austria) approach was performed in serum samples collected on d −49, +3, +21, and +84 relative to calving for identifying and quantifying up to 188 metabolites from 6 different compound classes (acylcarnitines, AA, biogenic amines, glycerophospholipids, sphingolipids, and hexoses). The concentrations of 170 metabolites were above the limit of detection and could thus be used in this study. We used various machine learning (**ML**) algorithms (e.g., sequential minimal optimization, random forest, alternating decision tree, and naïve Bayes–updatable) to analyze the metabolome data sets. The performance of each algorithm was evaluated by a leave-one-out cross-validation method. The accuracy of classification by the ML algorithms was lowest on d 3 compared with the other time points. Various ML methods (partial least squares discriminant analysis, random forest, information gain ranking) were then performed to identify those metabolites that were contributing most significantly to discriminating the groups. On d 21 after parturition, 12 metabolites (acetylcarnitine, hexadecanoyl-carnitine, hydroxyhexadecenoyl-carnitine, octadecanoyl-carnitine, octadecenoyl-carnitine, hydroxybutyryl-carnitine, glycine, leucine, phosphatidylcholine-diacyl-C40:3, *trans*-4-hydroxyproline, carnosine, and creatinine) were identified in this way. Pathway enrichment analysis showed that branched-chain AA degradation (before calving) and mitochondrial β-oxidation of long-chain fatty acids along with fatty acid metabolism, purine metabolism, and alanine metabolism (after calving) were significantly enriched in HBCS compared with NBCS cows. Our results deepen the insights into the phenotype related to overconditioning from the preceding lactation and the pathophysiological sequelae such as increased lipolysis and ketogenesis and decreased feed intake.

Key words: metabolomics, machine learning, metabolic pathway, transition cow

INTRODUCTION

Dairy cows undergo comprehensive metabolic and physiological changes during the transition from late pregnancy to lactation (Drackley, 1999; Ceciliani et al., 2018). In this critical life stage, when feed intake cannot meet the nutrient requirements to cope with the nutrient deficit postpartum, body reserves, mainly lipids and to some extent proteins, are mobilized (de Vries and Veerkamp, 2000; Roche et al., 2009; Roche et al., 2013). Body condition score is among the most important factors affecting productivity, reproduction, health, and longevity of dairy cattle. Overconditioning around calving may predispose cows to a greater risk for periparturient metabolic disorders associated with insufficient nutrient intake and a massive mobilization of body reserves (Roche et al., 2009, 2013; Bjerre-Harpøth et al., 2015). Thus, they are more likely to have difficulties at and after calving as a result of the suboptimal transition from pregnancy to lactation, leading to substantial economic losses for the dairy industry.

Metabolic profiling, known as metabolomics, is increasingly used in dairy science and was applied for predicting the risk of diseases (Zhang et al., 2017a; Dervishi et al., 2018) and for biomarker and pathway discovery in some metabolic diseases in transition cows (Hailemariam et al., 2014; Zhang et al., 2017b; Zandkarimi et al., 2018). Humer et al. (2016) reported metabolic alterations (using a targeted metabolomics approach) in the serum of dairy cows with low (<0.4 mmol/L), medium (0.4–0.7 mmol/L), and high (>0.7 mmol/L) postpartal nonesterified fatty acid (**NEFA**) concentrations but with similar milk production, BW, and BCS. Thus, to the best of our knowledge, potential differences in the metabolic profiles of dairy cows of divergent body condition were not reported. Recently, plasma metabolomic profiling using untargeted metabolomics revealed that most of the metabolites with increasing concentrations from late pregnancy to early lactation were primarily involved in lipid and energy metabolism, whereas most of the metabolites with decreasing concentrations were related to AA metabolism in transition cows (Luo et al., 2019).

The identification of metabolites and metabolic pathways associated with overconditioning around calving may help to elucidate the mechanisms involved in the pathophysiology of overconditioning in dairy cows. We used a targeted metabolomics approach (AbsoluteIDQ p180 kit; Biocrates Life Sciences AG, Innsbruck, Austria), which allows for the quantification of up to 180 metabolites from 6 compound classes [hexoses, AA, biogenic amines, acylcarnitines (**AcylCN**), glycerophospholipids, and sphingolipids], for metabolic profiling of periparturient dairy cows to get a more in-depth view on the changes in the serum metabolites of overconditioned cows compared with cows with normal BSC.

Acylcarnitines, glycerophospholipids (lyso-, diacyl-, and acyl-alkyl-phosphatidylcholines; **PC**), and sphingomyelins (**SM**) were of core interest in our study due to their pivotal role in lipid metabolism (Cole et al., 2012). Among lipids, PC are essential for very low density lipoprotein synthesis in the liver, thus playing a crucial role in triglyceride export by the liver (Cole et al., 2012). Shifts in the serum concentrations of PC are linked to fatty liver disease in transition cows (Imhasly et al., 2015). The elevated concentrations of serum long-chain AcylCN result from incomplete fatty acid oxidation (**FAO**) in response to an increased load of fatty acids in dairy cows around calving (Rico et al., 2018; Yang et al., 2019). In a human study, AcylCN was shown to have the potential to stimulate proinflammatory signaling pathways through the development of insulin resistance (Adams et al., 2009).

Besides body fat, labile protein reserves, mainly from skeletal muscle protein, are also used to provide AA for milk protein synthesis, direct oxidation, and gluconeogenesis in early lactation (Plaizier et al., 2000; Kuhla et al., 2011; Sadri et al., 2016). In addition, biogenic amines are a class of nitrogenous compounds that are involved in a wide variety of physiological processes (Ghaffari et al., 2019), although their role in transition cows is poorly understood.

For fully exploiting metabolomics data, conceptually different approaches of evaluation provide added value. In system biology, machine learning (**ML**) approaches are becoming of interest to provide actionable knowledge from large data sets and to improve metabolic profiling endeavors. Different supervised ML techniques such as support vector machine (**SVM**), random forest (**RF**), partial least squares discriminant analysis (**PLS-DA**), and decision tree are commonly applied to analyze metabolomics data (Kell, 2006; Mahadevan et al., 2008; Issaq et al., 2009; Cuperlovic-Culf, 2018).

In this study, we thus aimed to create a deep phenotype that by means of ML would allow for differentiating NBCS and HBCS cows according to their metabolic profile and the pathways enriched at different stages of the transition period. Applying ML, we aimed at moving on from descriptive studies toward elucidating the differentially altered pathways in more detail to improve our understanding of the pathophysiological processes involved. Overconditioning is most likely to happen during late lactation, when decreasing milk yields are not adequately considered by offering less nutrient-dense diets; therefore, we used an experimental approach reenacting this situation.

MATERIALS AND METHODS

The farm trial was conducted at the Educational and Research Centre for Animal Husbandry, Hofgut Neumuehle, Muenchweiler a.d. Alsenz, Germany. All animal experiments were carried out according to the German Animal Welfare Act and approved by the local authority for animal welfare affairs [Landesuntersuchungsamt Rheinland-Pfalz (G 14-20-071)], Koblenz, Germany.

Animal Grouping

To characterize the metabolic signature of the periparturient cows, an animal model of normal versus increased lipomobilization around calving based on a targeted divergence in body condition was used to identify significantly enriched metabolic pathways in serum (Figure 1). As reported earlier (Schuh et al., 2019), 38 multiparous Holstein cows were allocated to 2 groups 15 wk antepartum (**ap**), each consisting of 19 cows that were fed differently to reach either normal body condition (**NBCS**; n = 19; average parity: 2.42 $\pm$ 1.84; average BW: 665.4 $\pm$ 63.7 kg; mean $\pm$ SD) or an overconditioned status (**HBCS**; n = 19; average parity: 3.37 $\pm$ 1.67; average BW: 720.1 $\pm$ 56.9 kg; mean $\pm$ SD) at dry-off. The targeted BCS and backfat thickness (**BFT**) at dry-off (HBCS: >3.75 and >1.4 cm, respectively; NBCS: <3.5 and <1.2 cm, respectively) were reached. The preselected cows were also stratified for comparable milk yields (NBCS: 10,361 $\pm$ 302 kg; HBCS: 10,315 $\pm$ 437 kg). The HBCS cows received a high-energy ration (NE_L = 7.2 MJ/kg of DM), and the NBCS cows received a reduced-energy ration (NE_L = 6.8 MJ/kg of DM; Supplemental Table S1, https:// doi.org/10.3168/jds.2019-17114). After the differential feeding, both groups received the same diets during the dry period and the subsequent lactation. The differences in body condition at 15 wk ap (Δ = 0.7 BCS points and 1.1 cm of BFT) were largely maintained until calving (Figure 2a and b). More precisely, the BCS of cows in both groups declined during lactation, but the losses were greater in HBCS cows than in NBCS cows (Figure 2c and d). The calculated energy balance (i.e., the energy intake minus the energy needed or demand for milk synthesis and maintenance) was more negative ($P < 0.001$) in HBCS cows than in NBCS cows, and positive values were reached later for HBCS cows than for NBCS cows (Figure 2e).

Sampling and Laboratory Analyses

Blood samples were collected from the vena caudalis mediana before the morning feeding on d −49,

+3, +21, and +84 relative to calving. After clotting and subsequent centrifugation (10 min, 2,000 $\times$ g at 4°C), the sera were obtained and stored at −80°C until analysis. The metabolite profiles in serum were determined by liquid chromatography (**LC**)–electrospray ionization–tandem MS and flow injection electrospray ionization–tandem MS profiling through targeted metabolomics using the AbsoluteIDQ p180 kit (Biocrates Life Sciences AG). Mass detection and compound identification were performed by multiple reaction monitoring. The kit plates were used for the quantification of up to 188 metabolites belonging to 6 compound classes (Supplemental Table S2, https://doi.org/10.3168/jds .2019-17114), including AcylCN (40), proteinogenic and modified AA (21), glycerophospholipids (76 PC, 14 lysoPC), SM (15), biogenic amines (21), and hexoses (1).

This kit was validated according to the European Medicines Agency guidelines (EMA, 2011), which implies proof of reproducibility within a given error range. Equal volumes (10 μL) of the serum samples were applied directly to the assay. Analytical specifications for the limit of detection (**LOD**) and evaluated quantification ranges, further LOD for semiquantitative measurements, identities of quantitative and semiquantitative metabolites, specificity, potential interferences, linearity, precision and accuracy, reproducibility, and stability were according to the manufacturer's manual AS-P180 and as described in Zukunft et al. (2013). The LOD was set to 3 times the values of zero samples. Sample preparation and LC-MS/MS measurements of metabolites were performed according to the manufacturer's manual UM-P180 (Biocrates Life Sciences AG) and as previously described in Zukunft et al. (2013). The assay procedures of the AbsoluteIDQ p180 kit and the metabolite nomenclature have been described in detail previously (Romisch-Margl et al., 2012; Zukunft et al., 2013). Sample handling was performed by a Hamilton Microlab STAR robot (Hamilton Bonaduz AG, Bonaduz, Switzerland) and an Ultravap nitrogen evaporator (Porvair Sciences, Leatherhead, UK) along with standard laboratory equipment.

Mass spectrometric analyses were performed on an API 4000 triple quadrupole system (Sciex Deutschland GmbH, Darmstadt, Germany) equipped with a 1200 Series HPLC (Agilent Technologies Deutschland GmbH, Böblingen, Germany) and an HTC PAL autosampler (CTC Analytics, Zwingen, Switzerland) controlled by Analyst 1.6.1 software (AB SCIEX Deutschland GmbH, Darmstadt, Germany). Mass detection and compound identification were performed by multiple reaction monitoring. Measurements of metabolites were performed according to the manufacturer's manual.
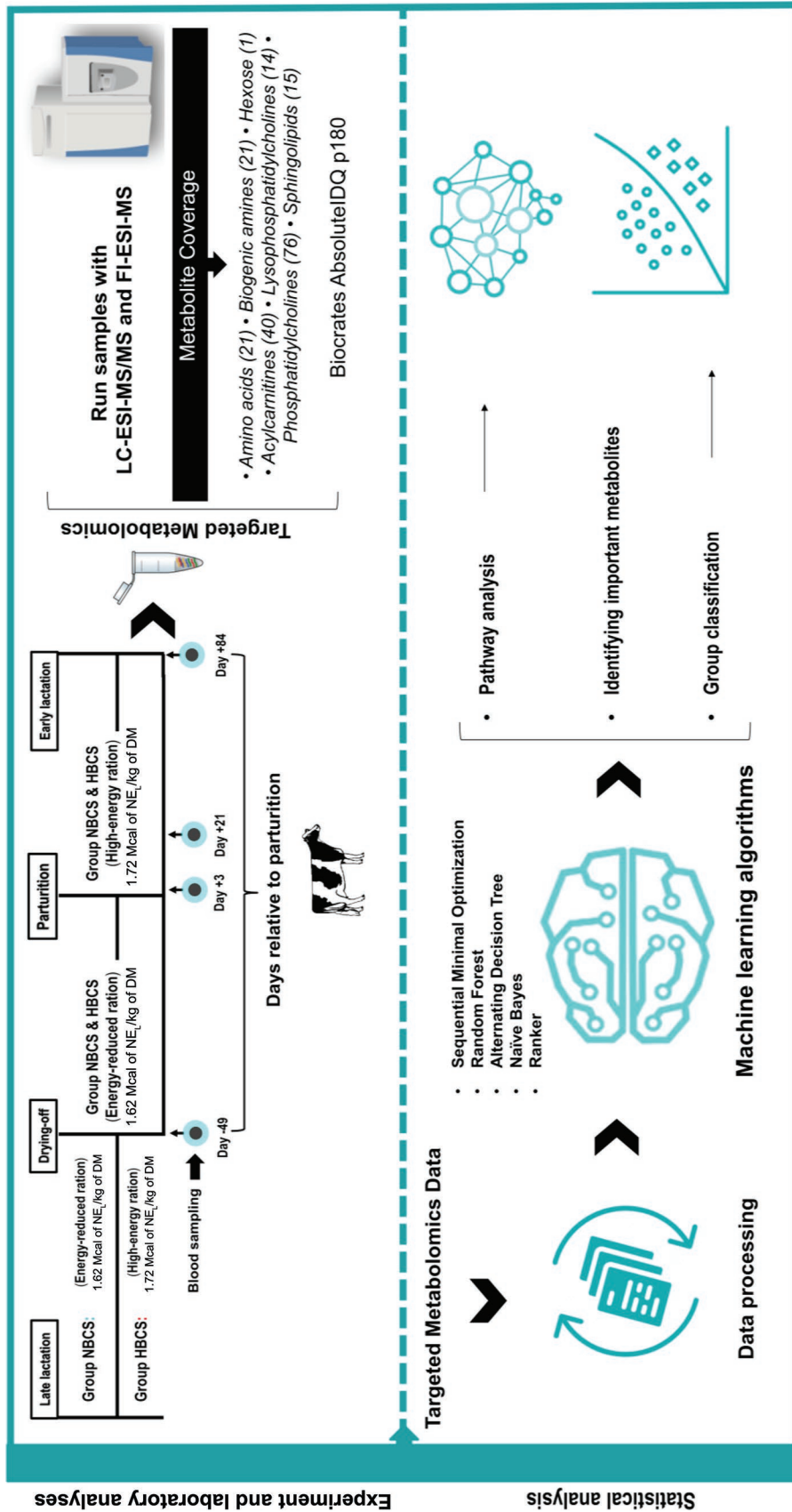
Figure 1. Overview of the experimental setup to characterize the metabolic signature of normal cows (NBCS; <3.5 BCS and <1.2-cm backfat thickness at dry-off) and over-conditioned cows (HBCS; >3.75 BCS and >1.4-cm backfat thickness at dry-off) during the experimental period (n = 19/group). The experimental setup was recently described by Schuh et al. (2019). LC-ESI-MS/MS = liquid chromatography electrospray ionization tandem MS; FI-ESI-MS = flow-injection electrospray MS. AbsoluteIDQ p180 kit manufactured by Biocrates Life Sciences AG (Innsbruck, Austria).

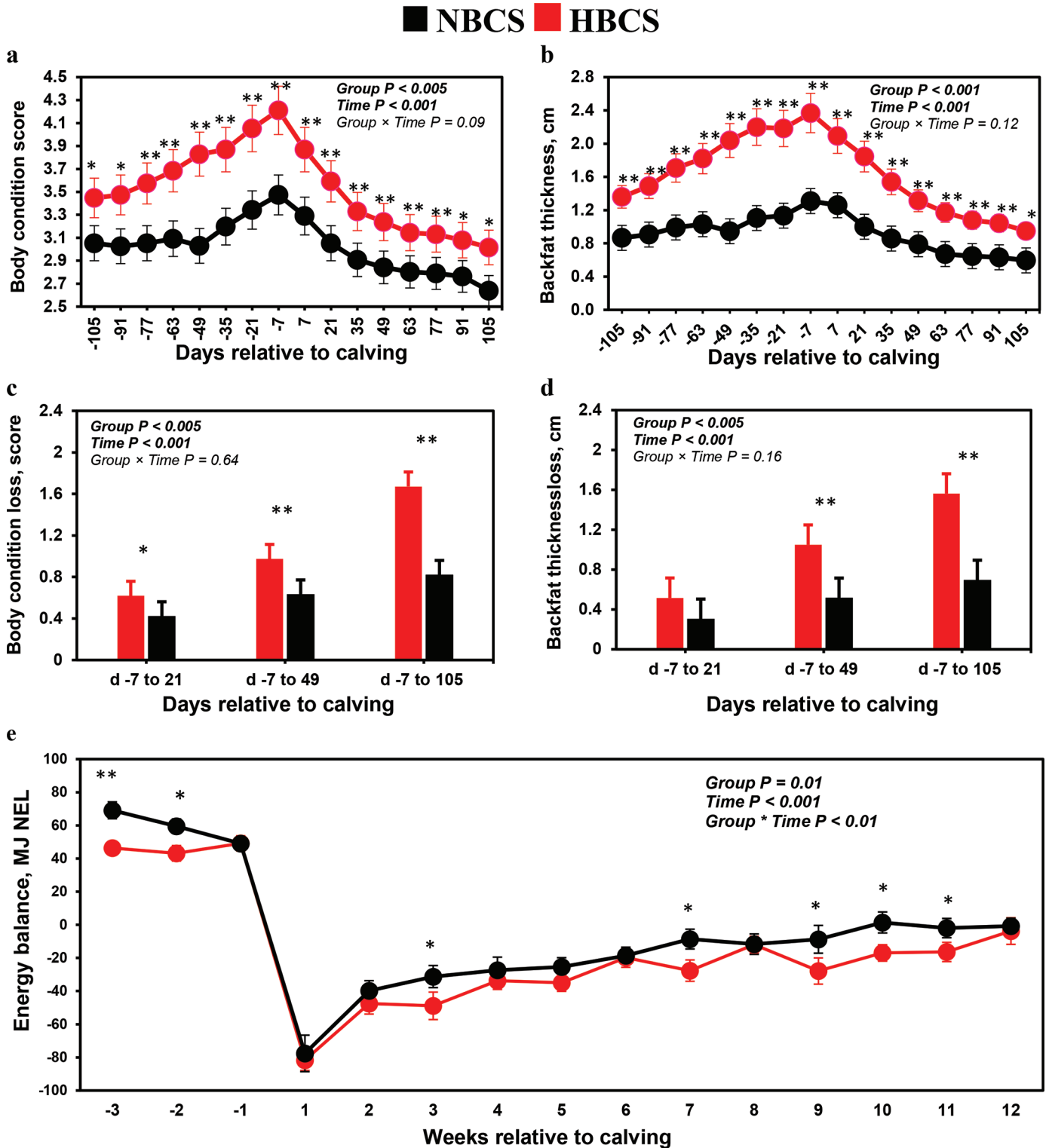**Figure 2.** Changes in (a) BCS, (b) backfat thickness (BFT), (c) BCS loss, (d) BFT loss, and (e) energy balance of normal cows (NBCS; <3.5 BCS and <1.2-cm BFT at dry-off) and overconditioned cows (HBCS; >3.75 BCS and >1.4-cm BFT at dry-off) during the experimental period (n = 19/group). Asterisks indicate a significant difference between the groups at a given time point (*P < 0.05; **P < 0.01). Data are presented as means ± SEM. Data for BCS, BFT, BCS loss, BFT loss, and energy balance are from Schuh et al. (2019).

Briefly, internal standards for the LC-MS/MS procedure were pipetted onto the filter inserts of the 96-well plate. This plate is prespiked with internal standards for the FIA-MS/MS procedure by the manufacturer. Ten microliters of serum was placed into the cavities of the 96-well filter plate of the p180 kit. Serum samples were dried in a nitrogen stream for 30 min. Amino acids and biogenic amines in the samples were derivatized with an excess of 5% phenylisothiocyanate (Sigma-Aldrich, Steinheim, Germany) for 20 min and dried under a nitrogen stream. Metabolites and internal standards were then extracted with 300 µL of methanol (AppliChem, Darmstadt, Germany) containing 5 mM ammonium acetate (Sigma-Aldrich) by incubation for 30 min at room temperature with continuous shaking. Metabolites were eluted by a centrifugation step (5 min at 500 × g at room temperature). One part of the eluate was diluted with Biocrates kit running solvent (1/5 vol/vol) for FIA-MS/MS analysis, and the other part was diluted with water (50/50 vol/vol) for LC-MS/MS analysis. The LC run was performed using an XDB-C18 column (3 × 100 mm, 3.5 µm; Agilent Technologies, Waldbronn, Germany). The serum concentrations of the metabolites were calculated using internal standards and are given as micromoles per liter. (Biogenic amines and AA are absolutely quantified with corresponding individual internal standards. For other chemical classes such as lipids, only 1 internal standard is used per class, and therefore the estimation of their concentrations is semiquantitative.) After considering the LOD, 170 metabolites were used in this study.

Data evaluation for serum metabolite concentrations and quality assessment was performed with the software MultiQuant 3.0.1 (Sciex Deutschland GmbH) and the MetIDQ software package (Biocrates Life Sciences AG), which is an integral part of the AbsoluteIDQ p180 kit. The potential associations of metabolomics data with classical variables [NEFA, BHB, leptin, insulin, glucose, IGF-1, revised quantitative insulin sensitivity check index (**RQUICKI**)] were assessed.

Statistics Analysis and Bioinformatics

Different ML approaches were performed to calculate the posterior probability $p(y|X)$ for each cow, where y is the predicting cow's label (HBCS or NBCS) and X is the cow's given feature vectors (serum metabolite data). The ML codes were implemented in Matlab equipped with Weka toolbox (Hall et al., 2009). To estimate the probability, ML algorithms look for a function h that maps the feature vector (serum metabolite data) X to output Y: $h: X \rightarrow Y$. Vector X comprises 170 features x1 . . . x170, and the output variable $y \in Y$ can be a binary

case of HBCS or NBCS. The function h initializes itself with some (random) parameters for mapping from X to Y and measures the amount of misclassification (loss error), and then searches repeatedly for the best combination of mapping parameters that minimizes the loss error the most. We let 4 distinct ML algorithms learn from the serum metabolite data and predict the posterior probability. Namely, we made use of sequential minimal optimization (**SMO**), RF, alternating decision tree (**ADTree**), and naïve Bayes–Updatable (**NB**) to measure the probability in 4 distinctive ways.

Sequential minimal optimization is basically a flavor of support vector machine algorithm that is optimized for speed. The SMO classifier is set to use a polynomial kernel, normalized training data set, and the C value of 1. Random forest is an ensemble learning method comprising numerous random decision trees. Each random decision tree samples the features with replacement and trains a tree classifier. Random forests search for the optimum decision tree among all randomly generated decision trees. Random forests are known to generalize well on training data. The RF in our work was set to 100 random trees and 10 maximum depth and 10 features. Alternating decision tree is a rule-based classification algorithm to produce a decision tree of training data by applying boosting techniques. We set the number of boosting iteration to 10 and the search path to expandable. Naïve Bayes was also used because, in theory, it performs well on a limited number of samples. Naïve Bayes classifiers work based on Bayes' theorem to measure the posterior probability. An updateable version of NB classifiers was used because it showed a superior classification performance, especially in medical detection systems. To be able to properly evaluate the performance of the ML models, a leave-one-out cross-validation schema was performed. In such a setting, for each ML algorithm, 1 cow is chosen for the performance evaluation (test set) and the remaining 37 cows are used in the learning process (train set). The obtained model will then be evaluated by the test data, and the performance will be measured. Repeating this procedure 38 times will give 38 performance measures. The average of all 38 performance values will be the final performance for the corresponding ML algorithm. The classification results across 4 classifiers indicate how robust the features could represent the HBCS and NBCS cows. The classification performance is a report based on different metrics. For true positive (TP) as the number true positives, false positive (FP) as the number of false positives, true negative (TN) as the number of true negatives, and false negative (FN) as the number of false negatives, we measured the following:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN)$$

$$Sensitivity = TP/(TP + FN)$$

$$Specificity = TN/(TN + FP)$$

$$Precision = TP/(TP + FP)$$

$$F\text{-measure score} = (2 \times sensitivity \times precision)/$$
$$(sensitivity + precision)$$

The entire classification procedure was performed once for d −49 and once for d +21. The serum data set of 38 cows comprised 170 distinct measurements (features). Next, to investigate how each individual feature (serum metabolite) contributes to its corresponding classified group (HBCS or NBCS), 2 algorithms were used: RF and information gain ranking (**IGR**). Serum metabolite feature importance was extracted from both algorithms. Common important features across 3 algorithms (RF, PLS-DA, and IGR) were selected for further analysis (12 features). In a follow-up analysis only, the selected important features were used to classify the HBCS and NBCS cows. Similar to the first classification setting, the same 4 classifiers were used. The classification was accomplished again for the data of d −49 and +21. The results of classification with 170 features were compared with the classification with 12 features. The selected important features were fed to a random tree classifier to obtain a traceable tree for serum metabolite classification.

The multivariate statistical analysis of the serum metabolite data was performed using the web-based metabolomic data processing tool MetaboAnalyst 4.0 (Chong et al., 2018; see http://www.metaboanalyst.ca for detailed methodology). Briefly, as quality control, variables containing more than 50% missing values (i.e., values lower than LOD) were not considered for the statistical analysis. The metabolite data were transformed using the generalized log-transformation and then Pareto scaled to correct for heteroscedasticity, reduce the skewness of the data, and reduce mask effects (van den Berg et al., 2006). The PLS-DA and variable importance in projection (**VIP**) were performed using the plsr function provided by the R pls package (Mevik and Wehrens, 2007) to identify the differential metabolites between the groups and to rank the metabolites according to their importance in discriminating groups. Permutation tests were used to assess the significance of the class discrimination determined by PLS-DA. The classification and cross-validation were performed using the corresponding wrapper function offered by the caret package. In each permutation, a PLS-DA model

was built between the data (x) and the permuted class labels (y) using the optimal number of components determined by cross-validation for the model based on the original class assignment. The heatmap clustered by Euclidean distance and Ward's minimum variance method (ward.D), which were calculated with Metabo-Analyst 4.0.

For the multilevel data sets that contain multiple types of variation, ANOVA-simultaneous component analysis was used to find metabolic changes with time, groups, and their interaction using MetaboAnalyst 4.0. For metabolite set enrichment analysis, metabolite data were mapped according to the Human Metabolome Database (HMDB, www.hmdb.ca), and the quantitative enrichment analysis was performed using the global test package by Metaboanalyst 4.0 (Goeman et al., 2004) to estimate a Q-statistic for each metabolite set. Furthermore, network analysis was constructed using NetworkAnalyst online software (www.networkanalyst.ca) based on the Kyoto Encyclopedia of Genes and Genomes (https://www.genome.jp/kegg/). This application allowed the visualization and interpretation of metabolomic data in the context of identifying and visualizing the enriched pathways.

The data of 12 serum metabolites identified by different ML were analyzed with estimation methods and presented as mean difference estimation plots (Claridge-Chang and Assam, 2016). Effect size was measured using Hedges' g (Greenland et al., 2016); per standard practice, effect sizes were referred to as trivial ($g < 0.2$), small ($0.2 < g < 0.5$), moderate ($0.5 < g < 0.8$), or large ($g > 0.8$). The effect size was presented as a bootstrap 95% confidence interval on separate but aligned axes. To indicate estimate precision, 95% confidence intervals for the mean differences were calculated using bootstrap methods (resampled 5,000 times, bias corrected, and accelerated) and are displayed with the bootstrap distribution of the mean. The Mann-Whitney U statistic was used to calculate P-values for pro forma reporting exclusively (Crichton, 2000).

A repeated-measures analysis was conducted using the MIXED procedure in SAS (version 9.4; SAS Institute Inc., Cary, NC) for classical metabolic variables (NEFA, BHB, leptin, glucose, insulin, IGF-1, RQUICKI) for group, time, and their interactions. Before analysis, all data were tested for normality of distribution by evaluating the Shapiro-Wilk statistic using the UNIVARIATE procedure of SAS, and, where appropriate, they were transformed using a $\log_{10}$ transformation. The model consisted of group, time (sampling date), and interaction of group and time as fixed effects and cow as the random effect. All means were compared using the PDIFF statement of SAS. Statistical significance was declared at $P \leq 0.05$, and a

trend to significance was declared at $0.05 < P \leq 0.10$. Pearson correlations were used to evaluate associations between NEFA, BHB, and serum metabolomics data ($P < 0.05$; false discovery ratio ≤ 0.1).

RESULTS

Classification Models and Feature Selection

Various supervised ML-based classifiers were performed to differentiate NBCS from HBCS cows based on the metabolomic data and to identify the important serum metabolites that contributed to the group separation. An overview of the ML procedures is shown in Figure 3. In step A, 4 distinct ML classification algorithms (i.e., SMO, RF, ADTree, and NB) were built and optimized. For assessing the performance of the classifiers, the sensitivity, specificity, positive predictive value, F measure, and accuracy were calculated, and the validation results were summarized in Table 1. The accuracies of SMO, RF, ADTree, and NB classifiers on d -49 ap were greater than 70%. Model accuracies obtained from SMO, RF, ADTree, and NB on d $+3$ postpartum (**pp**) were near the random level (50%).

In step B, a feature selection step was included in the ML analysis using RF and IGR algorithms with the aim of identifying the important metabolic features underlying the group separation (HBCS vs. NBCS). Figure 4 displays the top 34 serum metabolites that were identified by RF selection on d -49 ap (Figure 4a) and d $+21$ pp (Figure 4b) and their contributions (VIP and frequency of selection) to the group separation in ascending order. The highest-ranking serum metabolites on d -49 ap were His, hydroxybutyryl-carnitine (**C3-DC-C4-OH**), Ser, octadecenoyl-carnitine (C18:1), methylglutaryl-carnitine (C5-M-DC), and butenyl-carnitine (C4:1), and the highest-ranking metabolites on d $+21$ pp were octadecanoyl-carnitine (C18:0), butyryl-carnitine (C4:0), decadienyl-carnitine (C10:2), Asp, and acetylcarnitine (C2). As an outcome of the feature selection process, using the IGR algorithm, 34 metabolic features were identified and ranked according to their information gain rank on d -49 ap (Figure 4c) and d $+21$ pp (Figure 4d). Among the top 20 metabolites identified by IGR selection on d -49 ap, there were 8 AA (His, Val, Lys, Leu, Ile, Met, Orn, and Pro), 9 phosphatidylcholine diacyls (PC aa C32:0, C36:0, C38:0, C38:1, C38:6, C40:1, C40:2, C40:3, C42:5), and 1 sphingomyelin (SM C24:0). However, the highest-ranking metabolites on d $+21$ pp were 7 diacyl phosphatidylcholines (PC aa C32:0, C36:0, C42:1, C40:2, C40:3, C42:5, C42:6), 4 acyl-alkyl phosphatidylcholines (PC ae C42:5, C38:3, C36:1, C42:3), 2 sphingomyelins (SM C24:0, C26:0), Gly, creatinine, carnitine,

and some AcylCN [C2, hexadecanoyl-carnitine (C16), octadecanoyl-carnitine (C18), C18:1].

A PLS-DA was also performed to assess the significance of class discrimination and to extract the most important metabolites in the group separation. The PLS-DA score plots based on the analysis of complete digital maps for HBCS and NBCS cows showed a complete (Figure 5a and d) and significant separation of these 2 groups on d -49 ap ($P = 0.011$ by permutation test; Supplemental Figure S1, https://doi.org/10.3168/jds.2019-17114) and d $+21$ pp ($P = 0.004$ by permutation test; Supplemental Figure S2, https://doi.org/10.3168/jds.2019-17114). According to the VIP scores using the above PLS-DA model, 15 variables (metabolites) with a VIP >1.5 were selected. On d -49 ap (Figure 5b), 6 metabolites belonging to AA (Orn, Lys, Leu, Val, His, Met), 2 biogenic amines [α-aminoadipic acid, asymmetric dimethyl-arginine (**ADMA**)], carnitine (C0), C2, propionyl-carnitine (C3), and butyryl-carnitine (C4) contributed most significantly to the separation between the HBCS and NBCS groups. On d $+21$ pp (Figure 5e), 7 metabolites belonging to AcylCN [C2, C3-DC-C4-OH, tetradecanoyl-carnitine (C14:0), C16, hydroxyhexadecenoyl-carnitine (C16:1-OH), C18, and C18:1], 2 diacyl-PC (PC aa C38:0, C40:3), 2 acyl-alkyl-PC (PC ae C38:1, C42:4), 2 AA (Gly and Ile), and creatinine contributed most significantly to the separation between the HBCS and NBCS groups.

Further, the serum metabolites were analyzed by hierarchical clustering with a heat map to visualize the effect of overconditioning around calving. Figure 5c and f shows the levels of the top 70 metabolites based on PLS-DA VIP scores at 2 time points (d -49 ap and $+21$ pp) in HBCS and NBCS cows. The top 70 metabolites that differed significantly between groups on d $+3$ and $+84$ pp are shown in Supplemental Figure S3a and b (https://doi.org/10.3168/jds.2019-17114). In a follow-up analysis, common important features (serum metabolites) across the 3 algorithms (RF, IGR, and PLS-DA) were selected for further analysis.

Taken together, 12 metabolites, including His, Leu, Ile, Val, Lys, Orn, Pro, Ser, C0, ADMA, PC aa C36:0, and PC aa C40:2, were the common important features that were identified on d -49 ap by the ML algorithms (Figure 6a). The serum concentrations of His (effect size, $g = 1.64$, $P < 0.01$), Leu ($g = 1.51$, $P < 0.01$), Ile ($g = 1.03$, $P < 0.01$), Val ($g = 1.3$, $P < 0.01$), Lys ($g = 1.48$, $P < 0.01$), Orn ($g = 1.26$, $P < 0.01$), C0 ($g = 1.28$, $P < 0.01$), and ADMA ($g = 1.09$, $P < 0.01$) were greater but those of PC aa C36:0 ($g = -1.03$, $P < 0.01$) and PC aa C40:2 ($g = -1.11$, $P < 0.01$) were lower in HBCS cows compared with NBCS cows with large effect ($g > 0.8$, Figure 6a). The serum concentrations of
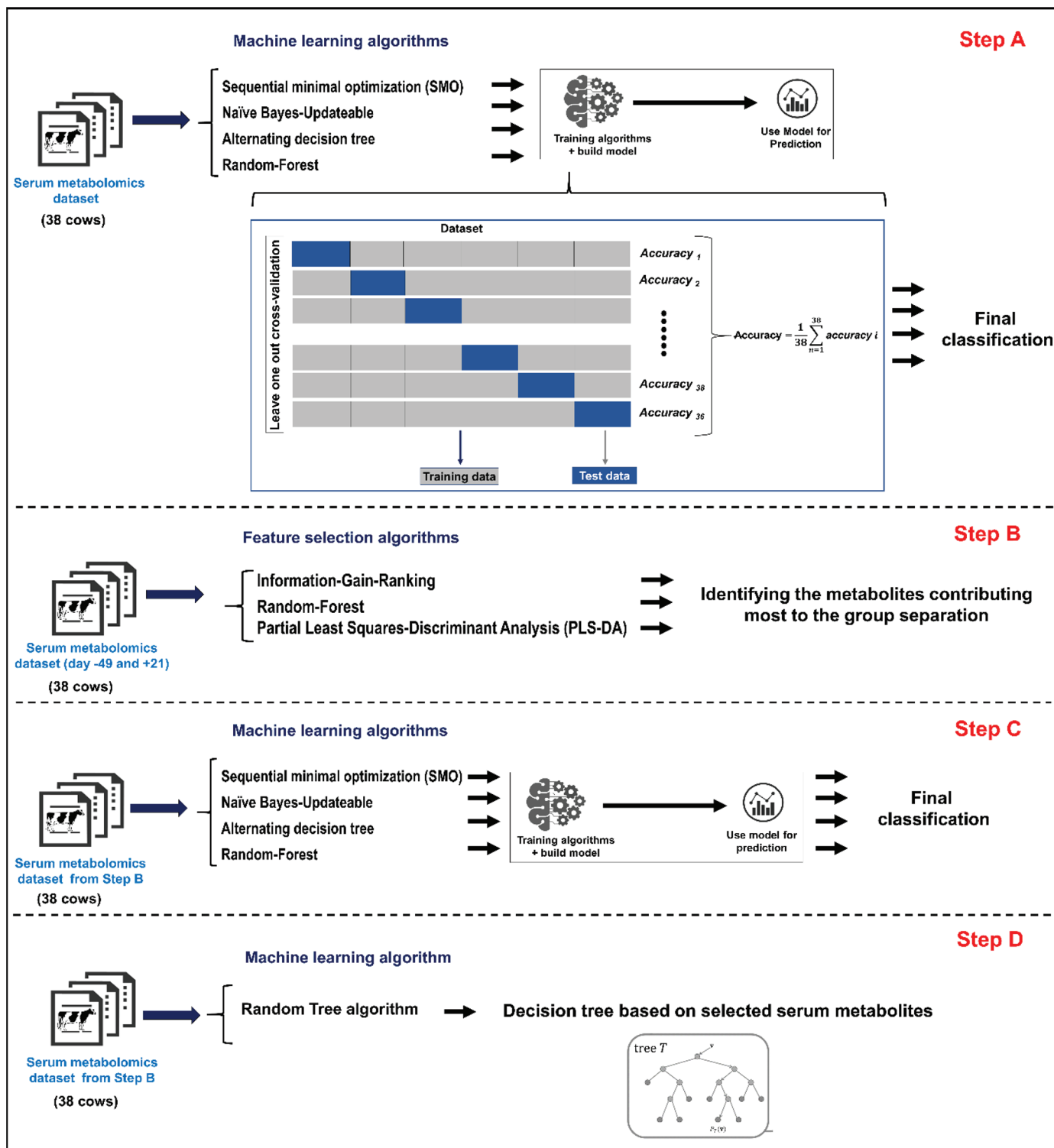
Figure 3. The workflow of machine learning analysis procedures.

Pro ($g = 0.60$, $P = 0.05$) and Ser ($g = 0.66$, $P = 0.05$) were greater in HBCS cows compared with NBCS cows with moderate effect ($0.5 < g < 0.8$, Figure 6a).

On d +21 pp, 12 metabolites, including C2, C16, C16:1-OH, C18:0, C18:1, C3-DC-C4-OH, Gly, Leu, PC aa C40:3, *trans*-4-hydroxyproline (**t4-OH-Pro**),

Table 1. Diagnostic performance of each machine learning classification model using 170 serum metabolites

Model	Day relative to calving	Sensitivity, %	Specificity, %	Positive predictive value	F-measure	Accuracy, %
Sequential minimal optimization	d −49	68	73	72	70	71
	d +3	68	63	65	66	65
	d +21	52	68	62	57	60
	d +84	73	78	77	75	76
Random forest	d −49	73	73	73	73	73
	d +3	52	68	62	57	60
	d +21	68	73	72	70	71
	d +84	68	57	61	65	63
Alternating decision tree	d −49	68	84	81	74	76
	d +3	57	47	52	55	52
	d +21	52	57	55	54	55
	d +84	68	68	68	68	68
Naïve Bayes	d −49	68	73	72	70	71
	d +3	57	63	61	59	60
	d +21	57	73	68	62	65
	d +84	63	63	63	63	63

carnosine, and creatinine, were the common important features that were identified by the algorithms (Figure 6b). The serum concentrations of C2 ($g = 1.25$, $P < 0.01$), C16 ($g = 1.31$, $P < 0.01$), C16:1-OH ($g = 1.17$, $P < 0.01$), C18:0 ($g = 1.34$, $P < 0.01$), C18:1 ($g = 0.90$, $P = 0.01$), Gly ($g = 1.04$, $P < 0.01$), and carnosine ($g = 0.82$, $P = 0.02$) were greater in HBCS cows compared with NBCS cows with large effect ($g > 0.8$, Figure 6b). The serum concentrations of creatinine ($g = 0.72$, tendency $P = 0.08$), C4-OH (C3-DC; $g = 0.79$, $P = 0.02$), and t4-OH-Pro ($g = 0.65$, $P = 0.04$) were greater in HBCS cows compared with NBCS cows with moderate effect ($0.5 < g < 0.8$). Serum concentrations of Leu were greater ($g = 0.43$, $P = 0.02$) in HBCS than in NBCS with small effect ($0.2 < g < 0.5$, Figure 6b). The serum concentrations of PC aa C40:3 ($g = 1.12$, $P < 0.01$) were lower in HBCS cows compared with NBCS cows with large effect ($g > 0.8$, Figure 6b).

In step C, the 12 aforementioned metabolites were tested to classify between NBCS and HBCS cows for validating the effectiveness of these subset metabolites for the group classification. Similar to the first classification setting (step A), the same 4 ML classification algorithms (SMO, RF, ADTree, and NB) were used for the data of d −49 ap and d +21 pp; the validation results are summarized in Table 2. Model accuracies obtained from the ML classifiers on d −49 ap for SMO, RF, ADTree, and NB were 78.9, 78.9, 73.7, and 76.3%, respectively. The classification accuracies on d +21 pp for SMO, RF, ADTree, and NB were 78.9, 76.3, 60.5, and 78.9%, respectively. Thus, the above-mentioned subset features have strongly contributed to the separation of HBCS and NBCS cows.

In step D, the subset of serum metabolites was used to construct decision trees. Figure 6 c and d shows the predictor metabolites and the cut-off points for each

metabolite to discriminate between HBCS and NBCS cows. On d −49 ap, the classifier selected Leu as the starting point, and 53% of cows (10/19) with serum Leu levels <0.26 µmol/L were identified as NBCS. The decision tree algorithm identified 84% of cows (16/19) as HBCS, with serum concentrations ≥0.26 µmol/L for Leu, ≥0.52 µmol/L for His, and <0.63 µmol/L for PC aa C40:2. On d +21 pp, the classifier selected PC aa C40:3 as the starting point, and 53% of cows (10/19) with serum PC aa C40:3 levels ≥0.08 µmol/L were identified as NBCS. The decision tree algorithm identified 68% of cows (13/19) as HBCS, with serum levels <0.08 µmol/L for PC aa C40:3, ≥0.32 µmol/L for Leu, and ≥0.21 µmol/L for C2.

Classical Metabolic Variables

The data of classical variables (metabolites and hormones) assessed weekly from 7 wk ap to 12 wk pp in blood serum were recently described by Schuh et al. (2019). In this study, the data of classical variables were reported on d −49, +3, +21, and +84 relative to calving. The serum NEFA concentrations were higher ($P < 0.05$) in HBCS compared with NBCS cows in early lactation (Supplemental Figure S4a, https://doi.org/10.3168/jds.2019-17114). The serum BHB concentrations increased after parturition ($P < 0.001$, d +21) and were higher ($P < 0.01$) in HBCS cows compared with NBCS cows on d +21 pp (Supplemental Figure S4b). The serum leptin concentrations decreased during lactation and were higher ($P < 0.05$) in HBCS compared with NBCS cows before and after calving (d −49 ap, +3 pp, and +21 pp; Supplemental Figure S4c). The serum concentrations of glucose were not different between the groups but fluctuated around calving (Supplemental Figure S4d). The serum concentrations of insulin and
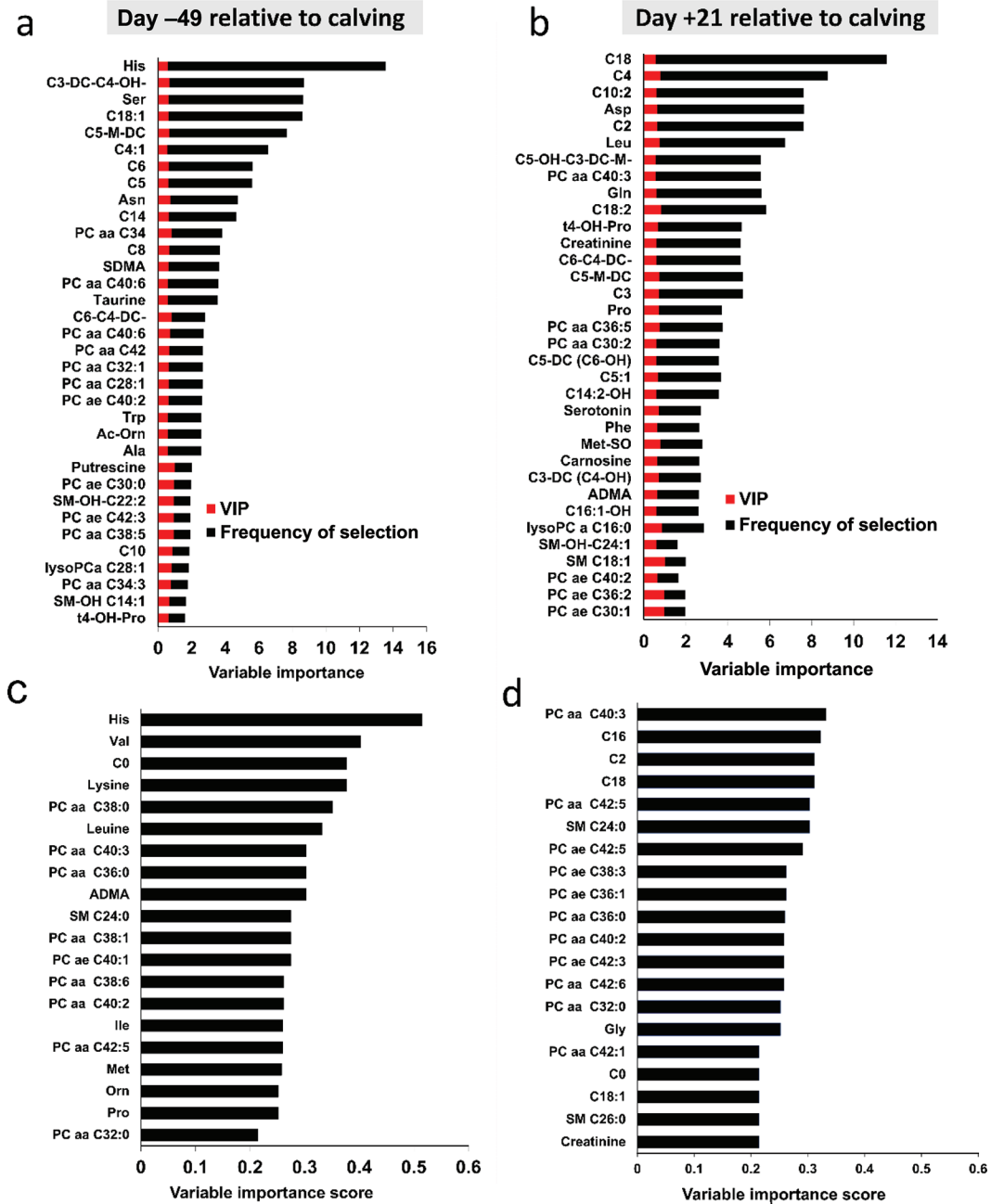
Figure 4. Important serum metabolites identified by random forest and information gain ranking on (a, c) d −49 antepartum and (b, d) d +21 postpartum. The x-axis indicates the variable importance in projection (VIP) scores corresponding to each variable on the y-axis. C4-OH (C3-DC) = hydroxybutyryl-carnitine; C18:1 = octadecenoyl-carnitine; C5-M-DC = methylglutaryl-carnitine; C4:1 = butenyl-carnitine; C6 = hexenoyl-carnitine; C5 = valeryl-carnitine; C14 = tetradecanoyl-carnitine; PC aa = di-acyl phosphatidylcholine; C8 = octanoyl-carnitine; SDMA = symmetric dimethyl-arginine; C6 (C4:1-DC) = hexanoylcarnitineumaryl-carnitine; PC ae = acyl alkyl phosphatidylcholine; Ac-Orn = acetylornithine; SM (OH) = hydroxysphingomyeline; C10 = decanoyl-carnitine; LysoPC a = lysophosphatidylcholine acyl; t4-OH-Pro = trans-4-hydroxyproline; C18 = octadecanoyl-carnitine; C4 = butyryl-carnitine/isobutyryl-carnitine; C10:2 = decadienyl-carnitine; C2 = acetyl-carnitine; C5-OH (C3-DC-M) = hydroxyvaleryl-/-isovaleryl-/-methylbutyryl-carnitine (methylmalonyl-carnitine); C18:2 = octadecadienyl-carnitine; C3 = propionyl-carnitine; C5-DC (C6-OH) = glutaryl-carnitine (hydroxyhexanoyl-carnitine); C5:1 = tiglyl-carnitine/methylcrotonyl-carnitine; C14:2-OH = hydroxytetradecadienyl-carnitine; Met-SO = methionine sulfoxide; ADMA = asymmetric dimethyl-arginine; C16:1-OH = hydroxy-hexadecenoyl-carnitine; SM = sphingomyeline; C0 = carnitine; Orn = ornithine; C16 = hexadecanoyl-carnitine.
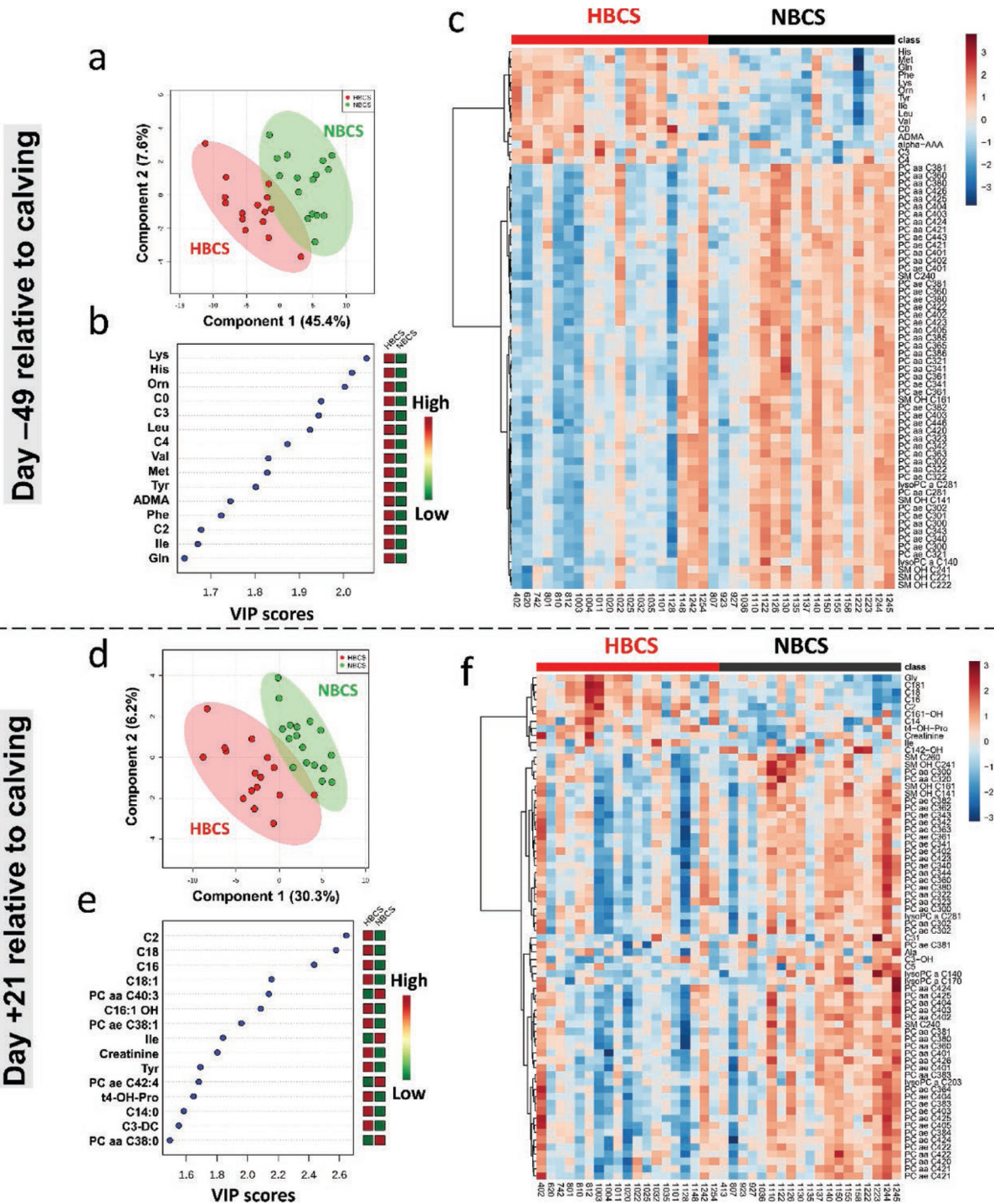
Figure 5. Partial least squares discriminant analysis showing 2 clusters for normal cows [NBCS; <3.5 BCS and <1.2-cm backfat thickness (BFT) at dry-off] and overconditioned cows (HBCS; >3.75 BCS and >1.4-cm BFT at dry-off) on (a) d −49 antepartum (ap) and (d) d +21 postpartum (pp), ranked by variable importance in projection (VIP) on (b) d −49 ap and (e) d +21 pp. The numbers in the score plots represent the code numbers of the individual animals. The heatmap of serum metabolites illustrates statistically significant serum metabolite differences in NBCS and HBCS cows on (c) d −49 ap and (f) d +21 pp. The colors in the heatmap reflect the serum metabolite abundance (mean cantered and divided by the range of each variable). Orn = ornithine; C0 = carnitine; C3 = propionyl-carnitine; C4 = butyryl-carnitine/isobutyryl-carnitine; ADMA = asymmetric dimethyl-arginine; C2 = acetylcarnitine; alpha-AAA = alpha-aminoadipic acid; PC aa = di-acyl phosphatidylcholine; PC ae = acyl alkyl phosphatidylcholine; SM = sphingomyelin; LysoPC a = lysophosphatidylcholine acyl; SM (OH) = hydroxysphingomyeline; C18:1 = octadecenoyl-carnitine; C18 = octadecanoyl-carnitine; C16 = hexadecanoyl-carnitine; C16:1-OH = hydroxyhexadecenoyl-carnitine; C14 = tetradecanoyl-carnitine; t4-OH-Pro = trans-4-hydroxyproline; C14:2-OH = hydroxytetradecadienyl-carnitine; C3:1 = propenoyl-carnitine; C3-OH = hydroxypropionyl-carnitine; C5 = valeryl-carnitine.

Figure 6. (a, b) The serum concentrations of 12 metabolites identified by different machine learning algorithms on (a) d −49 antepartum and (b) d +21 postpartum. The effect size (Hedges' g) between normal cows (NBCS; <3.5 BCS and <1.2-cm backfat thickness at dry-off) and over-conditioned cows (HBCS; >3.75 BCS and >1.4-cm backfat thickness at dry-off) is shown in the Gardner-Altman estimation plot. Both groups were plotted on the left axes; the mean difference was plotted on a floating axis on the right as a bootstrap sampling distribution. A Hedges' g of 1 shows that the 2 groups differ by 1 SD. The mean difference was depicted as a dot; the 95% CI is indicated by the ends of the vertical error bar and reported above as effect size [CI width lower bound; upper bound]. The P-value reported is the likelihood of observing the effect size. (c, d) Random forest decision model based on 12 serum metabolites identified by different machine learning techniques on cows' data on (c) d −49 antepartum and (d) d +21 postpartum. PC aa = di-acyl phosphatidylcholine; ADMA = asymmetric dimethyl-arginine; C0 = carnitine; C2 = acetylcarnitine; C16 = hexadecanoyl-carnitine; C16:1-OH = hydroxyhexadecenoyl-carnitine; t4-OH-Pro = trans-4-hydroxyproline.
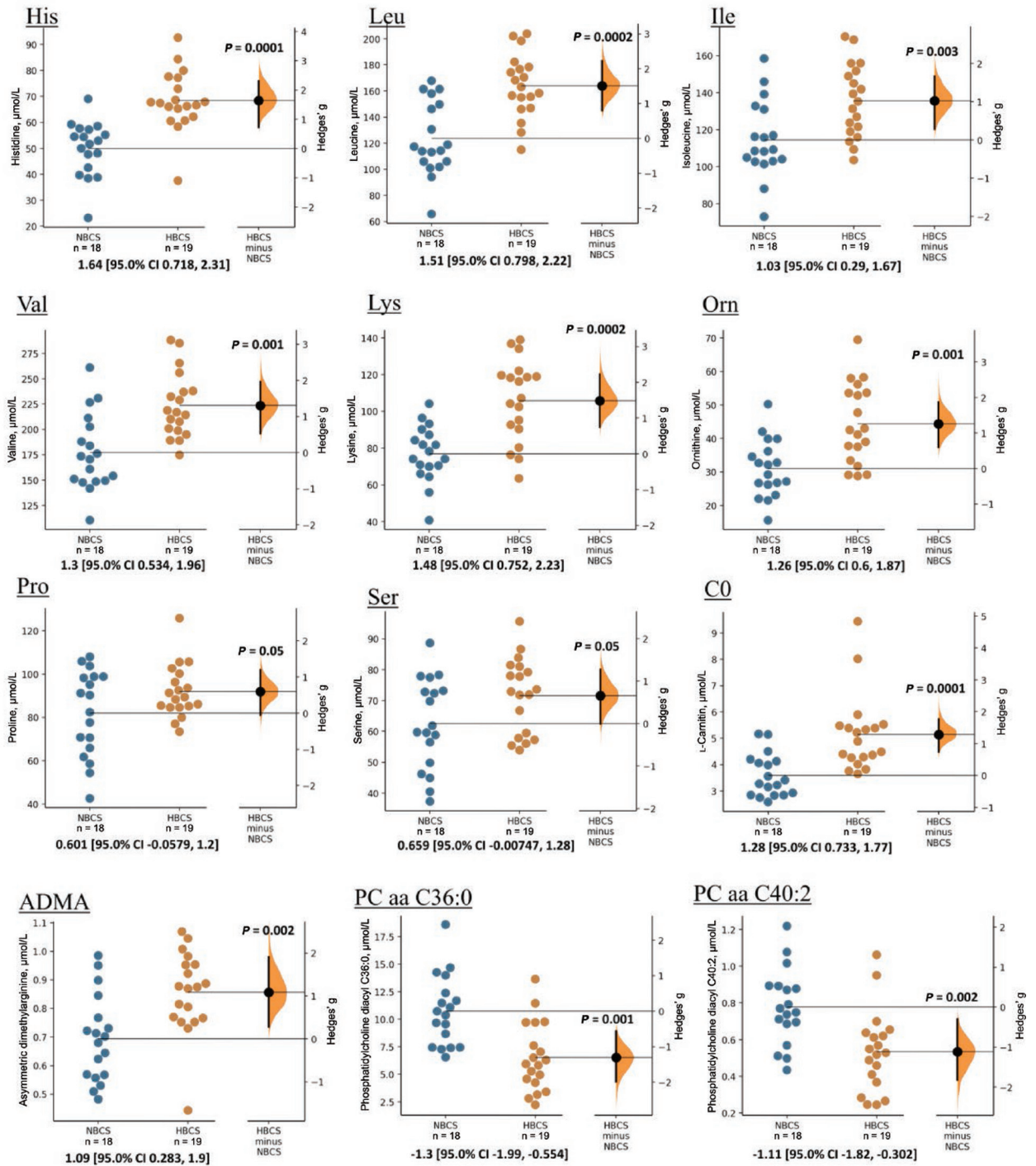
b

d +21 relative to calving

Figure 6 (Continued). (a, b) The serum concentrations of 12 metabolites identified by different machine learning algorithms on (a) d −49 antepartum and (b) d +21 postpartum. The effect size (Hedges' g) between normal cows (NBCS; <3.5 BCS and <1.2-cm backfat thickness at dry-off) and overconditioned cows (HBCS; >3.75 BCS and >1.4-cm backfat thickness at dry-off) is shown in the Gardner-Altman estimation plot. Both groups were plotted on the left axes; the mean difference was plotted on a floating axis on the right as a bootstrap sampling distribution. A Hedges' g of 1 shows that the 2 groups differ by 1 SD. The mean difference was depicted as a dot; the 95% CI is indicated by the ends of the vertical error bar and reported above as effect size [CI width lower bound; upper bound]. The P-value reported is the likelihood of observing the effect size. (c, d) Random forest decision model based on 12 serum metabolites identified by different machine learning techniques on cows' data on (c) d −49 antepartum and (d) d +21 postpartum. PC aa = di-acyl phosphatidylcholine; ADMA = asymmetric dimethyl-arginine; C0 = carnitine; C2 = acetylcarnitine; C16 = hexadecanoyl-carnitine; C16:1-OH = hydroxyhexadecenoyl-carnitine; t4-OH-Pro = trans-4-hydroxyproline.

c

d −49 relative to calving
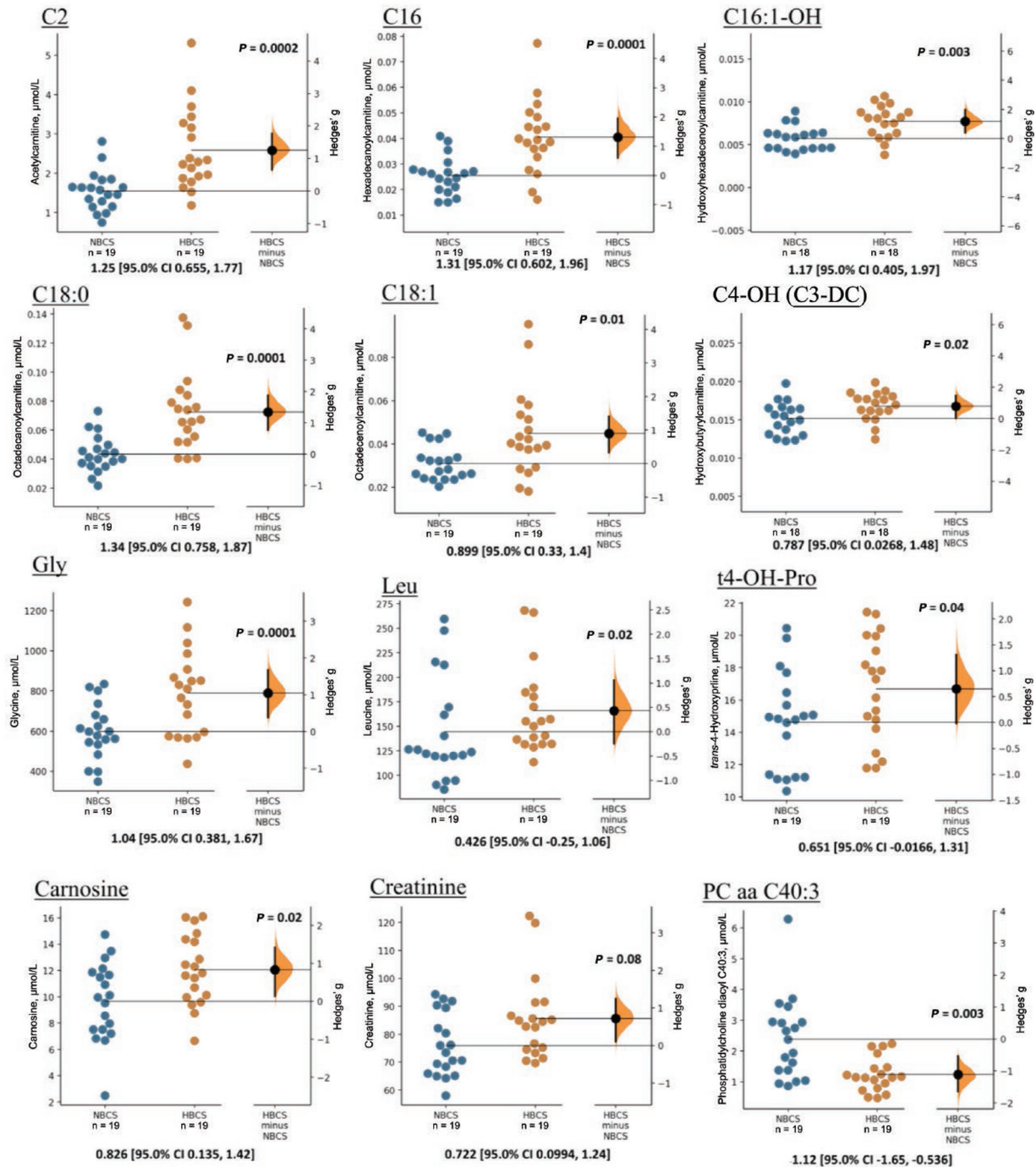

d

d +21 relative to calving

Figure 6 (Continued). (a, b) The serum concentrations of 12 metabolites identified by different machine learning algorithms on (a) d −49 antepartum and (b) d +21 postpartum. The effect size (Hedges' *g*) between normal cows (NBCS; <3.5 BCS and <1.2-cm backfat thickness at dry-off) and overconditioned cows (HBCS; >3.75 BCS and >1.4-cm backfat thickness at dry-off) is shown in the Gardner-Altman estimation plot. Both groups were plotted on the left axes; the mean difference was plotted on a floating axis on the right as a bootstrap sampling distribution. A Hedges' *g* of 1 shows that the 2 groups differ by 1 SD. The mean difference was depicted as a dot; the 95% CI is indicated by the ends of the vertical error bar and reported above as effect size [CI width lower bound; upper bound]. The *P*-value reported is the likelihood of observing the effect size. (c, d) Random forest decision model based on 12 serum metabolites identified by different machine learning techniques on cows' data on (c) d −49 antepartum and (d) d +21 postpartum. PC aa = di-acyl phosphatidylcholine; ADMA = asymmetric dimethyl-arginine; C0 = carnitine; C2 = acetylcarnitine; C16 = hexadecanoyl-carnitine; C16:1-OH = hydroxyhexadecenoyl-carnitine; t4-OH-Pro = trans-4-hydroxyproline.
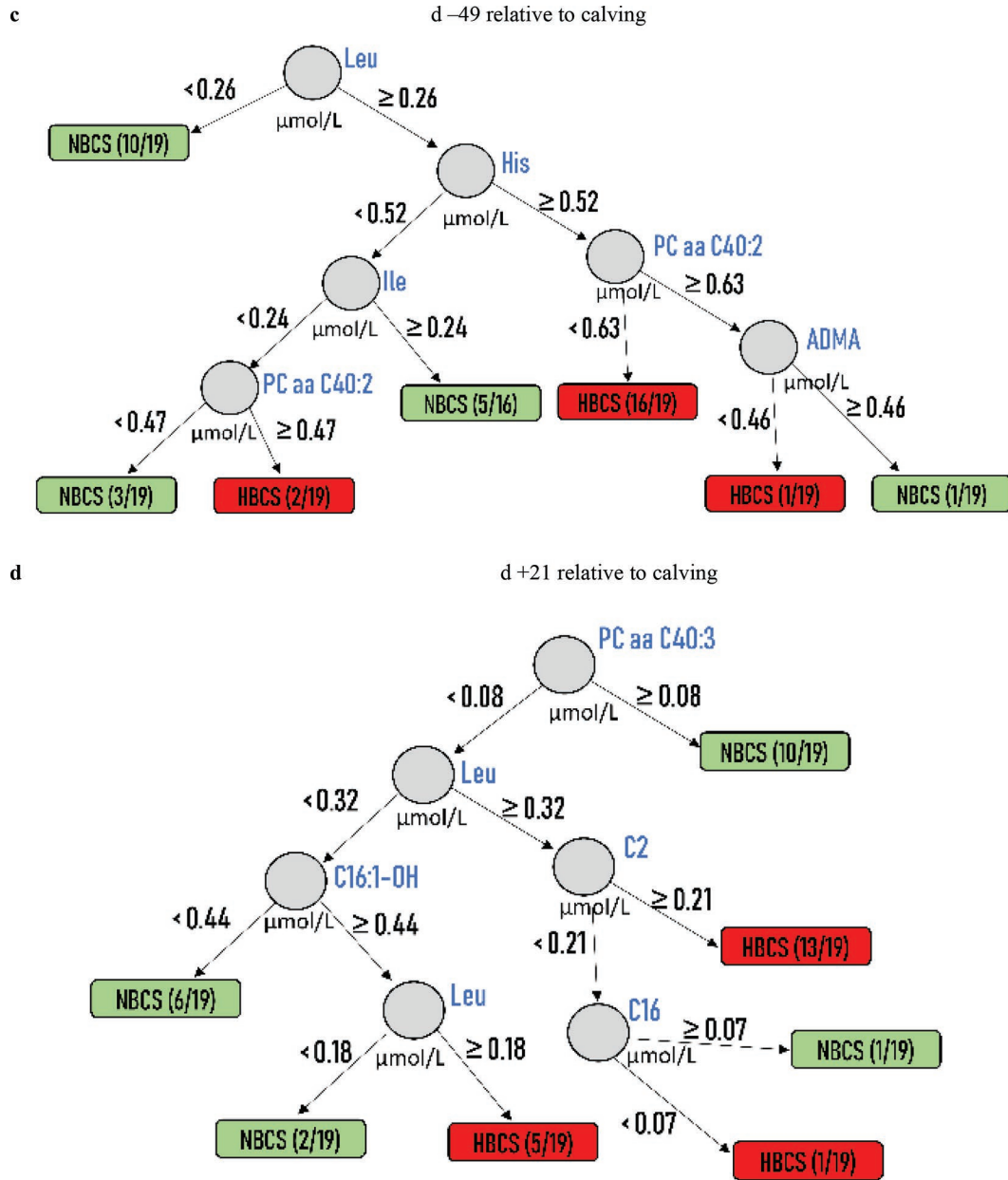
IGF-1 were lower in lactation than in the dry period in both groups (Supplemental Figure S5a and b, https://doi.org/10.3168/jds.2019-17114). When considering all time points, HBCS cows had greater insulin concentrations than NBCS cows, but differences could not be assigned to individual time points when doing Bonferroni-corrected multiple comparisons (Supplemental Figure S5a and b). Insulin sensitivity was estimated by calculating the RQUICKI from the data of blood glucose, insulin, and NEFA (Holtenius and Holtenius,

2007). Greater insulin but also glucose concentrations in HBCS cows indicate reduced insulin sensitivity. We performed a Spearman correlation analysis to identify significant correlations between individual metabolites and NEFA or BHB values. The serum NEFA value was positively correlated with C16:0- and C18:0-AcylCN and negatively correlated with multiple long-chain PC species (i.e., PC ae 44:3, PC ae 44:2, PC aa 34:4, PC ae 42:5, PC ae 42:3, PC ae 36:0; Supplemental Figure S6a, https://doi.org/10.3168/jds.2019-17114). Serum BHB was positively correlated with C2 and negatively correlated with multiple long-chain PC species (i.e., PC ae 38:2, PC aa 42:1, PC ae 38:4, PC ae 42:0, PC ae 38:2, lysoPC a 20:3; Supplemental Figure S6b).

Time-Resolved Analysis of Metabolomics Data

For the time-series metabolomics, ANOVA-simultaneous component analysis was performed to identify the major patterns associated with each factor and to describe the variation between animals (HBCS and NBCS), the variation in time (d −49, +3, +21, and +84 relative to calving) and their interaction. This is based on Manly's unrestricted permutation of observation, which calculates the permuted variation associated with each factor. A permutation approach was used to validate the model as demonstrated by significance levels of $P < 0.01$ for the groups and the time (Supplemental Figure S7, https://doi.org/10.3168/jds.2019-17114). The significant variables were identified based on the leverage and the squared prediction errors (**SPE**) associated with each variable (Figure 7a). The leverage and the SPE were used to evaluate the importance of the metabolite to the model and the fitness of the model for the particular metabolite, respectively. Variables with low SPE and higher leverage were modeled well after the major patterns. Based on the leverage and SPE plots, the significant variables associated with a specific factor were identified (Figure 7a). Figure 7b shows those features that were well modeled by group, time, and their interaction. Four

metabolites (C2, PC ae C38:1, PC aa C40:3, PC aa C36:0) were most affected by group (HBCS vs. NCBS), and 3 metabolites (PC aa C40:3, PC ae C38:1, PC ae C38:3) were most altered during the transition period. The time-course profiles of representative metabolites with high rankings by the multivariate empirical Bayes approach (**MEBA**) are shown in Figure 7c. Hotelling's T^2 was used to rank the serum metabolites by their difference in temporal profiles across different biological conditions (HBCS vs. NBCS) over the transition period. Among the top metabolites, 11 phosphatidylcholine diacyls (C36:0, C38:0, C38:1, C40:2, C40:3, C40:4, C42:0, C42:1, C42:4, C42:5, C42:6), 13 phosphatidylcholine acyl-alkyls (C30:0, C34:0, C36:0, C36:2, C38:0, C38:2, C38:3, C40:1, C40:2, C42:2, C42:3, C44:3, C44:6), 3 sphingomyelins (C24:0, C26:0, C24:1), 2 AA (Lys, His), and 4 AcylCN (C2, C16, C18, C18:2) were identified by MEBA. Acetylcarnitine had the highest statistical value (Figure 7c).

Pathway Mapping

Following ML classification analysis, quantitative enrichment analysis was performed to identify the metabolic pathways that were most affected by overconditioning around calving. Enrichment tests were performed using metabolite set enrichment analysis software (www.metaboanalyst.ca). Mapping the serum metabolites on d −49 ap identified branched-chain AA (**BCAA**) degradation, propanoate metabolism, methyl-histidine metabolism, betaine metabolism, His metabolism, Gly, and Ser metabolism as the most affected metabolic pathways in HBCS compared with NBCS cows (Figure 8a). Mitochondrial β-oxidation of long-chain fatty acids along with fatty acid metabolism, purine metabolism, and Ala metabolism were the main metabolic pathways affected by overconditioning on d + 21 pp (Figure 8b). Furthermore, network analysis was constructed using NetworkAnalyst online software to visualize the enriched metabolic pathways as shown in Figure 8c and d.

Table 2. Diagnostic performance of each machine learning classification model using a subset (n = 12 metabolites) of serum metabolites

Model	Day relative to calving	Sensitivity, %	Specificity, %	Positive predictive value	F-measure	Accuracy, %
Sequential minimal optimization	d −49	0.84	0.74	0.76	0.80	78.9
	d +21	0.82	0.74	0.76	0.80	78.9
Random forest	d −49	0.84	0.74	0.76	0.80	78.9
	d +21	0.74	0.79	0.78	0.76	76.3
Alternating decision tree	d −49	0.68	0.79	0.76	0.72	73.7
	d +21	0.53	0.68	0.63	0.57	60.5
Naïve Bayes	d −49	0.79	0.74	0.75	0.77	76.3
	d +21	0.84	0.74	0.76	0.80	78.9

a

c

Serum metabolite	Hotelling's T^2
Acetylcarnitine (C2)	26.921
Phosphatidylcholine diacyl C42:5	23.598
Phosphatidylcholine diacyl C38:0	21.389
Phosphatidylcholine diacyl C36:0	21.029
Sphingomyelin C24:0	19.367
Phosphatidylcholine diacyl C40:3	18.207
Phosphatidylcholine diacyl C38:1	17.932
Phosphatidylcholine diacyl C42:6	16.234
Phosphatidylcholine acyl-alkyl C40:1	15.868
Phosphatidylcholine acyl-alkyl C42:3	14.993
Hexadecanoylcarnitine (C16:0)	14.989
Octadecanoylcarnitine (C18:0)	14.545
Phosphatidylcholine diacyl C40:2	14.054
Phosphatidylcholine diacyl C42:1	13.884
Phosphatidylcholine acyl-alkyl C36:0	13.542
Lysine	13.187
Phosphatidylcholine diacyl C42:0	13.029
Phosphatidylcholine diacyl C40:4	12.974
Phosphatidylcholine acyl-alkyl C34:0	12.73
Phosphatidylcholine acyl-alkyl C38:2	12.689
Phosphatidylcholine acyl-alkyl C30:0	12.479
Phosphatidylcholine diacyl C42:4	12.12
Phosphatidylcholine acyl-alkyl C44:3	11.629
Phosphatidylcholine acyl-alkyl C40:2	11.619
Phosphatidylcholine acyl-alkyl C42:2	11.606
Octadecadienylcarnitine (C18:2)	11.594
Phosphatidylcholine acyl-alkyl C44:6	11.593
Hydroxysphingomyeline C24:1	11.397
Sphingomyelin C26:0	11.284
Phosphatidylcholine acyl-alkyl C38:0	11.2
Phosphatidylcholine acyl-alkyl C36:2	11.199
Phosphatidylcholine acyl-alkyl C38:3	11.123
lysoPhosphatidylcholine acyl C14:0	10.765
Histidine	10.758

b

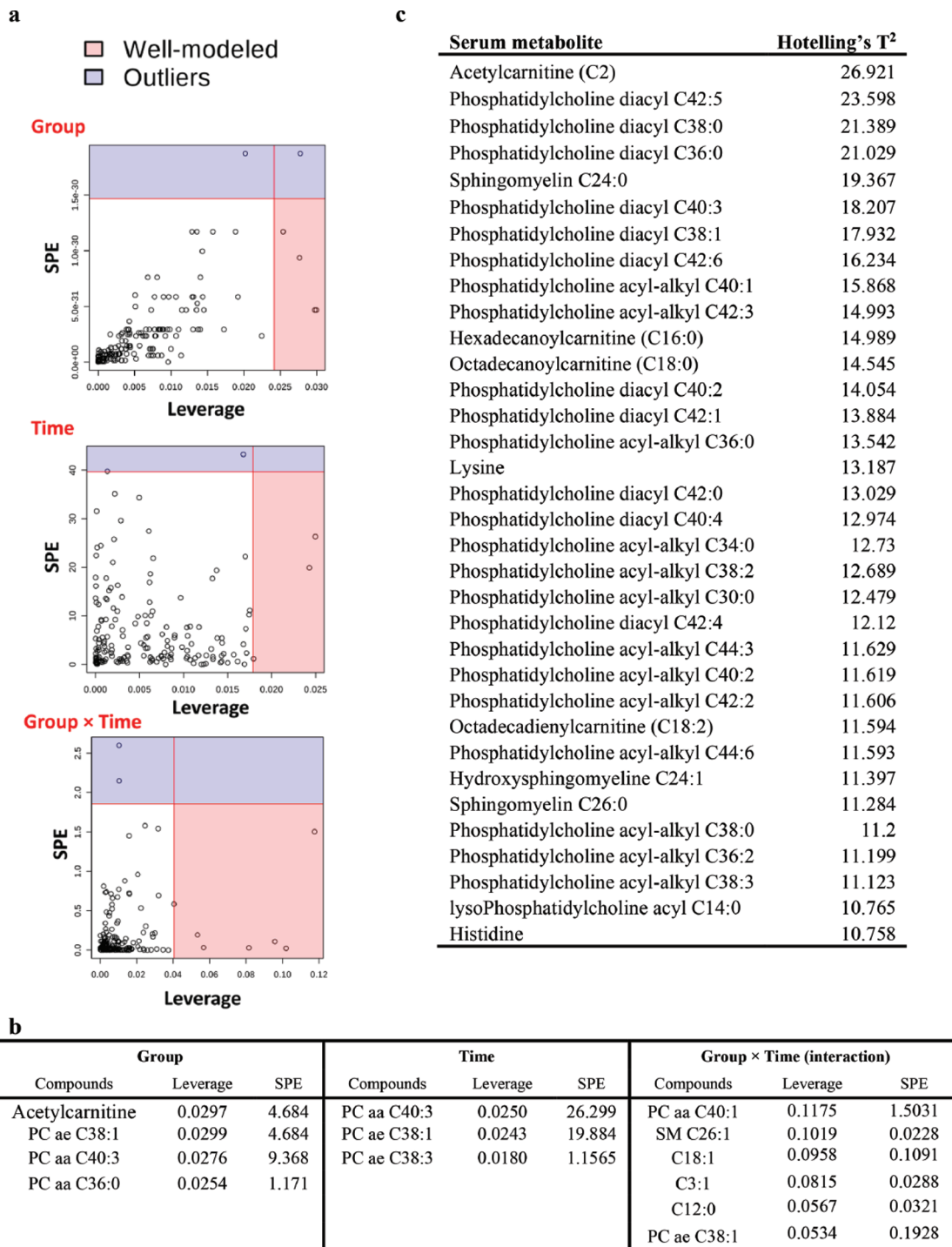| Group			Time			Group × Time (interaction)		
Compounds	Leverage	SPE	Compounds	Leverage	SPE	Compounds	Leverage	SPE
Acetylcarnitine	0.0297	4.684	PC aa C40:3	0.0250	26.299	PC aa C40:1	0.1175	1.5031
PC ae C38:1	0.0299	4.684	PC ae C38:1	0.0243	19.884	SM C26:1	0.1019	0.0228
PC aa C40:3	0.0276	9.368	PC ae C38:3	0.0180	1.1565	C18:1	0.0958	0.1091
PC aa C36:0	0.0254	1.171				C3:1	0.0815	0.0288
						C12:0	0.0567	0.0321
						PC ae C38:1	0.0534	0.1928

Figure 7. (a) Leverage and squared prediction error (SPE) scatter plots of the ANOVA-simultaneous component analysis (ASCA) variables submodels for normal (NBCS; <3.5 BCS and <1.2-cm backfat thickness at dry-off) and overconditioned (HBCS; >3.75 BCS and >1.4-cm backfat thickness at dry-off) cows, time, and their interactions. Vertical and horizontal lines indicate cut-off leverage and SPE values, respectively. The leverage evaluates the importance of the metabolite to the model, and SPE tests the fitness of the model for particular metabolites. Metabolites with a high leverage value and a low SPE value were considered to be differential metabolites (the well-modeled group). Metabolites in blue have patterns that are different from the major patterns. (b) Details of compounds in the leverage–SPE scatter plots of the ASCA variables submodels for group (HBCS vs. NBCS cows), time (d −49, +3, +21, and +84 relative to calving), and their interactions. (c) Hotelling's T^2 distribution is a generalization of Student's t-statistic that is used in multivariate hypothesis testing. Hotelling's T^2 was used to rank the serum metabolites by their difference in temporal profiles across different biological conditions (HBCS vs. NBCS cows) over the transition period. Metabolites with higher Hotelling's T^2 values include those whose profiles are more different across the time series. PC aa = di-acyl phosphatidylcholine; PC ae = acyl alkyl phosphatidylcholine; SM = sphingomyeline.
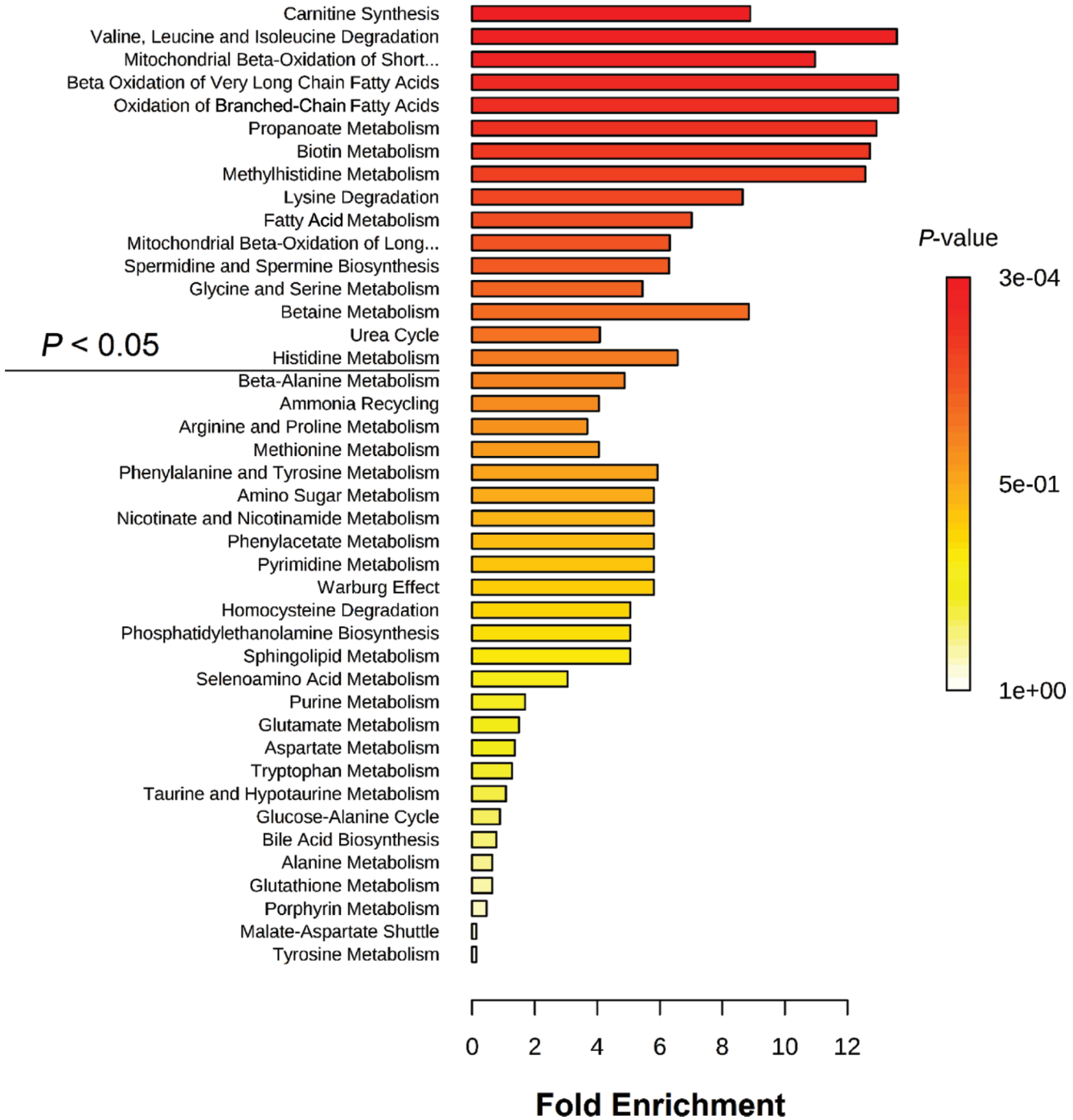
a

Figure 8. (a, b) Summary plot of meaningful metabolic pathways in serum from the metabolite sets enrichment analysis (MSEA) that are ranked by Holm *P*-value on (a) d −49 antepartum (ap) and (b) d +21 postpartum (pp). (c, d) A total of 170 serum metabolites were imported to MSEA to identify the most enriched pathways. Network analysis was constructed to visualize the enriched metabolic pathways. The nodes in blue are the metabolic pathways that are enriched across normal cows (NBCS; <3.5 BCS and <1.2-cm backfat thickness at dry-off) and overconditioned cows (HBCS; >3.75 BCS and >1.4-cm backfat thickness at dry-off) on (c) d −49 ap and (d) d +21 pp.

DISCUSSION

A deeper understanding of transition cow biology can help reduce the incidence of metabolic diseases in early lactation (Drackley, 1999; Grummer et al., 2004). Overconditioning is related to an increased risk for production diseases and often results from energy intakes exceeding the actual needs in late lactation. We herein used an experimental model for overcondition developing in late lactation. In contrast to other previous studies comparing cows of different BCS by applying different feeding regimens during the dry period, we did this during the last 8 wk before dry-off, similar to the approach taken by Roche et al. (2013, 2015), who

b

d +21 relative to calving

Metabolite Sets Enrichment Overview HBCS vs. NBCS

Mitochondrial Beta-Oxidation of Long-Chain Saturated Fatty Acids
Beta Oxidation of Very Long Chain Fatty Acids
Oxidation of Branched-Chain Fatty Acids
Fatty Acid Metabolism
Alanine Metabolism
Ammonia Recycling
Bile Acid Biosynthesis
Glutamate Metabolism
Glutathione Metabolism
Porphyrin Metabolism
Purine Metabolism
Methionine Metabolism
Carnitine Synthesis
Glycine and Serine Metabolism
Arginine and Proline Metabolism

P < 0.05

Valine, Leucine and Isoleucine Degradation
Propanoate Metabolism
Aspartate Metabolism
Biotin Metabolism
Lysine Degradation
Beta-Alanine Metabolism
Histidine Metabolism
Mitochondrial Beta-Oxidation of Short...
Urea Cycle
Spermidine and Spermine Biosynthesis
Phospholipid Biosynthesis
Catecholamine Biosynthesis
Thyroid hormone synthesis
Tyrosine Metabolism
Phenylalanine and Tyrosine Metabolism

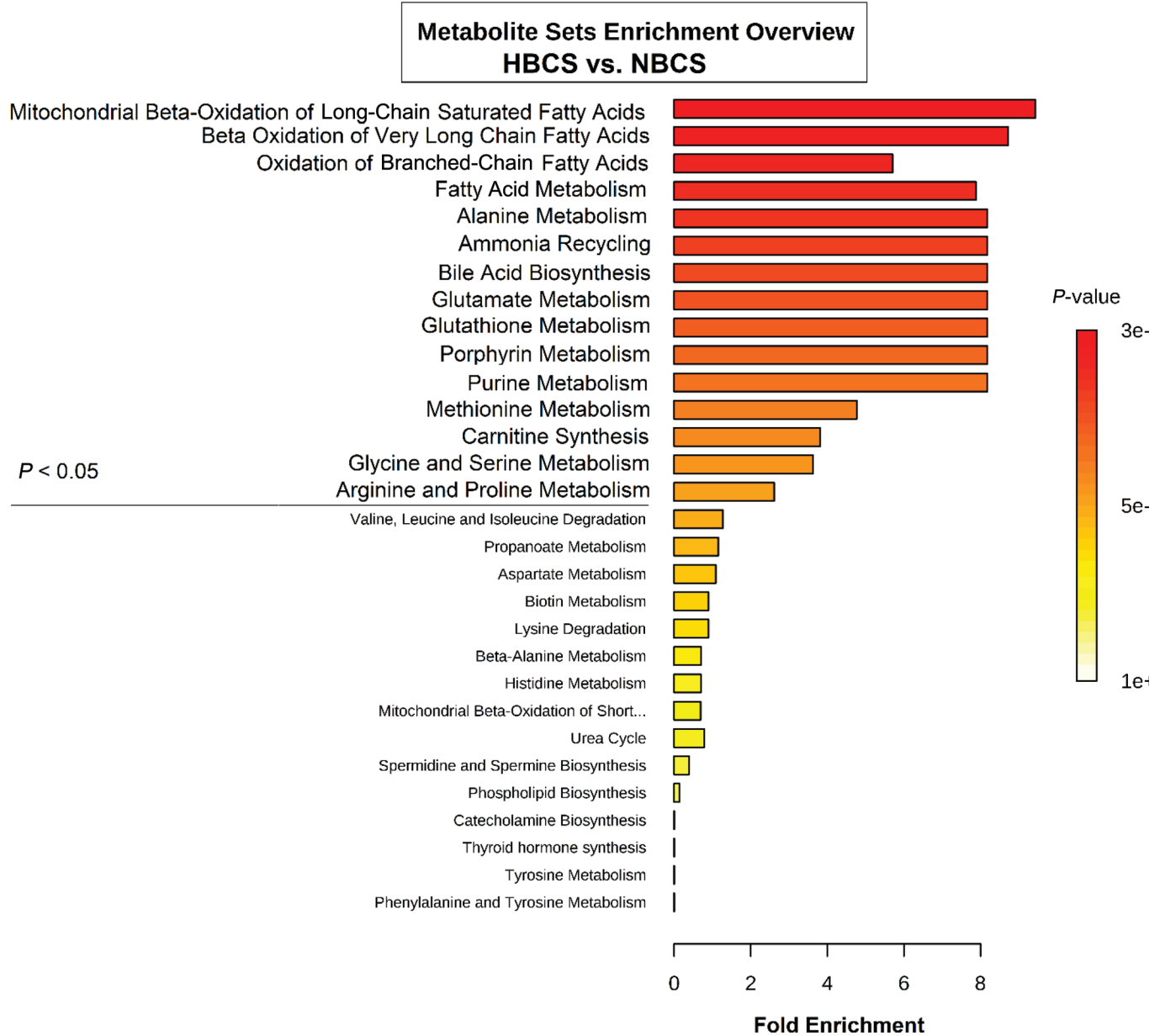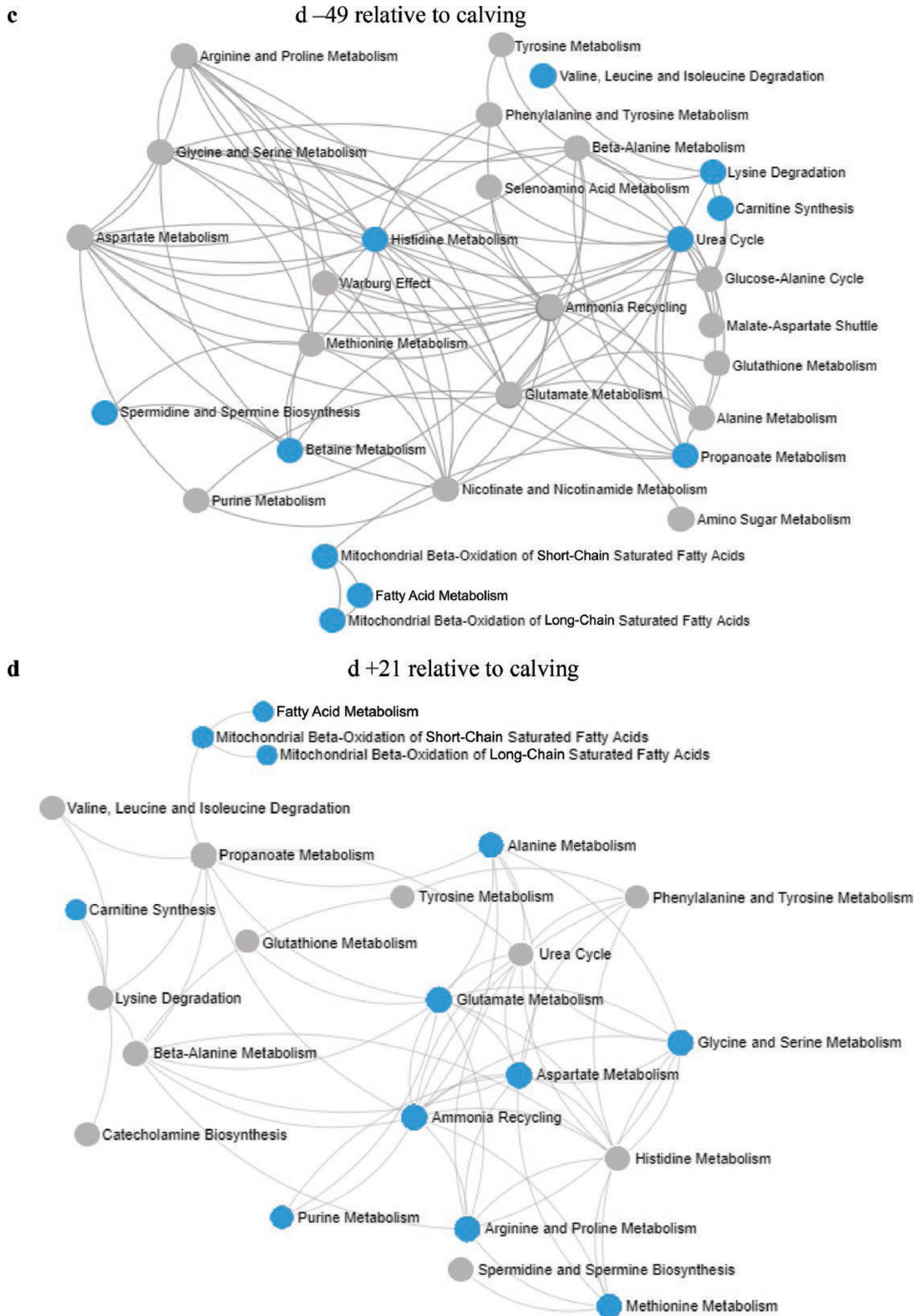*P*-value
3e-
5e-
1e-

Fold Enrichment
0 2 4 6 8

Figure 8 (Continued). (a, b) Summary plot of meaningful metabolic pathways in serum from the metabolite sets enrichment analysis (MSEA) that are ranked by Holm *P*-value on (a) d −49 antepartum (ap) and (b) d +21 postpartum (pp). (c, d) A total of 170 serum metabolites were imported to MSEA to identify the most enriched pathways. Network analysis was constructed to visualize the enriched metabolic pathways. The nodes in blue are the metabolic pathways that are enriched across normal cows (NBCS; <3.5 BCS and <1.2-cm backfat thickness at dry-off) and overconditioned cows (HBCS; >3.75 BCS and >1.4-cm backfat thickness at dry-off) on (c) d −49 ap and (d) d +21 pp.

Figure 8 (Continued). (a, b) Summary plot of meaningful metabolic pathways in serum from the metabolite sets enrichment analysis (MSEA) that are ranked by Holm P-value on (a) d −49 antepartum (ap) and (b) d +21 postpartum (pp). (c, d) A total of 170 serum metabolites were imported to MSEA to identify the most enriched pathways. Network analysis was constructed to visualize the enriched metabolic pathways. The nodes in blue are the metabolic pathways that are enriched across normal cows (NBCS; <3.5 BCS and <1.2-cm backfat thickness at dry-off) and overconditioned cows (HBCS; >3.75 BCS and >1.4-cm backfat thickness at dry-off) on (c) d −49 ap and (d) d +21 pp.

assessed classical metabolites via individual analyses. Moreover, we combined the differential feeding with a preselection of the cows aiming to amplify the variation in BCS that was already in existence in the herd. As shown in detail earlier (Schuh et al., 2019) and summarized in Figure 2, the phenotypes in the HBCS and NBCS groups obtained maintained these differences throughout the entire study period (i.e., until wk 15 of lactation; Schuh et al., 2019). We used an innovative approach to enhance the current understanding of metabolic profiling in the serum of overconditioned dairy cows. We used a targeted metabolomics approach and different ML algorithms to identify important metabolite differentiators between HBCS and NBCS cows and to identify the key metabolic pathways that are affected by overconditioning around calving.

Using the ML approach, we found that serum metabolite profiles could successfully separate the 2 groups with high accuracy on d −49 and +21 relative to calving. Using ML, we herein report a panel of serum metabolites (12/170) that distinguished HBCS from NBCS cows before calving, including His, Ile, Leu, Val, Lys, Orn, Pro, Ser, C0, PC aa C36:0, PC aa C40:2, and ADMA (Figure 6a). However, due to the different diets used in these cows from wk 15 to 7 before calving, the changes observed in the serum metabolites on d −49 were subject to influences from both body condition and the diets. The changes observed in the serum AA concentrations of HBCS and NBCS cows on d −49 might be due, at least in part, to the greater microbial protein yield and MP supply for the higher-energy diet because the observed differences largely disappeared on d +3 (i.e., after receiving the identical diets). We found greater serum BCAA concentrations in HBCS cows than in NBCS cows before calving.

In this study, the above-mentioned subset of serum metabolites (n = 12) strongly contributed to the separation of HBCS and NBCS cows with higher accuracy compared with all 170 metabolites. Appending more predictors improves classification performance if and only if the predictors carry relevant information. Otherwise, they undermine the classification performance by adding more variance and noise to the data, making it harder for the classifier to converge to a generalized model. The other point to consider is the number of animals (n = 38) compared with the number of features and predictors (serum metabolites = 170). For most of the classical ML algorithms to obtain a model, the number of features should be lower than the sample size.

Interestingly, the single metabolite with the most significant association with overconditioning before calving was Leu, which was greater in HBCS cows than in NBCS cows (Figure 6a). In addition, by mapping the serum metabolites to their respective metabolic pathways, some metabolic pathways such as BCAA degradation, propanoate metabolism, biotin metabolism, and methyl-histidine metabolism were differentially enriched before calving. However, the changes in the serum concentrations of BCAA and the enrichment of the BCAA degradation pathway associated with overconditioning tended to be reduced after calving (Figure 8b; Supplemental Figure S8a and b, https://doi.org/10.3168/jds.2019-17114). Thus, as stated above, we cannot determine whether these observations were independent of differential feeding that cows received from wk 15 to 7 before calving. However, elevation in the blood concentrations of BCAA, including Leu, has been reported in obese individuals (She et al., 2007; Newgard, 2017) and may contribute to the development of insulin resistance and type 2 diabetes. Based on human and laboratory animal studies, it is more likely that increased circulating BCAA is due to decreased expression and activity of BCAA catabolic enzyme in obesity (She et al., 2007; Pietiläinen et al., 2008; Leskinen et al., 2010), and thus, the elevated circulating levels of BCAA are probably a consequence rather than a cause of insulin resistance (Muniandy et al., 2019). Whether greater serum BCAA concentrations act in the same manner in dairy cows is currently unknown and warrants further investigation.

Propanoate metabolism was also found to be another metabolic pathway that was affected in response to overconditioning before calving in the current study. The propanoate metabolism begins with the conversion of propionic acid to propionyl coenzyme A (propionyl-CoA), which is the usual first step in the metabolism of carboxylic acids (Frayn, 2010). In a study by Chen et al. (2015), the propanoate metabolism pathway was enriched in abnormal obesity in humans and was reported to be linked with mitochondrial dysfunction.

Acetylcarnitine, C16:0- and C18:0-AcylCN, SM C24:0, and long-chain PC were the metabolites ranked highest based on Hotelling's T^2 in MEBA in both negative and positive modes and were found to change significantly over the entire time course based on ANOVA. We also identified a subset of serum metabolites (12/170 metabolites) that contributed most significantly to discriminating the groups after calving on d 21 pp. These metabolites that differentiated between HBCS and NBCS cows were AcylCN (C2, C3-DC-C4-OH, C16, C16:1-OH, C18:0, and C18:1), AA (Gly, Leu), and biogenic amines (t4-OH-Pro, carnosine, and creatinine). Acylcarnitines are intermediate fatty acid esters that are formed from their respective acyl-CoA by both mitochondrial and peroxisomal enzymes for FAO (Rinaldo and Matern, 2002). We found higher levels of C2 and long-chain AcylCN in HBCS cows

than in NBCS cows after calving. Accordingly, similar changes in serum concentrations of C2 and long-chain AcylCN were observed in high-mobilizing cows compared with low-mobilizing cows (Humer et al., 2016). It is likely that fatty acid oxidation should be in excess relative to oxidation in tricarboxylic acid (**TCA**) and the respiratory chain to guarantee a continuous supply of energy. In support of this, our data showed an increase of long-chain fatty acyl-CoA entry into mitochondria, but this did not seem to be paralleled by an upregulation of the downstream metabolic pathways, such as the TCA cycle and respiratory chain. Thus, perturbations of the TCA cycle such as the depletion of several TCA cycle intermediates could be one explanation for the incomplete oxidation of the long-chain fatty acyl-CoA through β-oxidation in tissues of HBCS cows. By mapping the serum metabolites, some pathways were found to be closely related to overconditioning on d 21 pp. These included mitochondrial β-oxidation of fatty acids and fatty acid metabolism. Elevated lipolysis, indicated by increasing serum NEFA concentrations in early lactation, was observed in this study (Supplemental Figure S4a, https://doi.org/10.3168/jds .2019-17114). We speculate that this increase in long-chain AcylCN might mirror the increase in lipolysis and the concomitant increase in FAO rates. The observed greater accumulation of long-chain AcylCN in the serum of HBCS cows is consistent with previous findings that mitochondrial stress induced by lipid overload is associated with incomplete FAO and spillover of AcylCN into the circulation (Sigauke et al., 2003; Mihalik et al., 2010).

The 3 biogenic amines (t4-OH-Pro, carnosine, and creatinine) identified by ML algorithms are worth further discussion for their associations with overconditioning after calving. Carnosine, creatinine, and t4-OH-Pro were increased in the HBCS cows. Carnosine is an endogenous dipeptide (β-alanine and L-His) with many functions, such as pH buffering, metal ion chelation, and antioxidant properties; the latter is mediated both directly and indirectly by different mechanisms, including metal ion chelation and scavenging reactive oxygen species and peroxyl radicals (Boldyrev et al., 2013) with systemic rather than local effects (Stegen et al., 2015). As reported previously from this experiment, HBCS cows were metabolically challenged during early lactation due to more severe negative energy balance and intense mobilization of body fat, associated with reduced DMI and compromised antioxidative capacity as reflected by reduced total ferric reducing antioxidant power (Schuh et al., 2019). Therefore, in view of results obtained in humans and laboratory animals (Boldyrev et al., 2013; Stegen et al., 2015; Regazzoni et al., 2016) and our present observations, the greater serum car-

nosine concentrations might have a potential to support antioxidant defenses in HBCS cows. Similarly, elevated serum carnosine concentrations in cows in early lactation were also reported in previous studies (Huber et al., 2016; Zhang et al., 2017c). We also found higher levels of creatinine and t4-OH-Pro in HBCS cows than in NBCS cows. In the current study, the group differences in serum creatinine indicated that creatinine could discern the greater mobilization of body reserves (body protein in particular) in HBCS cows after calving. Increased serum creatinine was shown to result from deficits in dietary protein in sheep (Sahoo et al., 2009) and may reflect increased muscle protein degradation.

The metabolic profiling used herein also identified a strong phospholipid-sphingomyelin metabolism-related signature associated with overconditioning in periparturient cows. Sphingomyelins are formed of a ceramide core linked to 1 fatty acid and phosphatidylethanolamine or PC (Floegel et al., 2013). Notably, overconditioning and insulin resistance have been associated with altered sphingolipid metabolism and elevated circulating concentrations of ceramide in overweight cows (Rico et al., 2015, 2018). In periparturient cows, circulating ceramide concentrations (C18:0, C18:1, C20:0, C22:0, C22:1, and C24:1) were higher in overweight (BCS >4.0) compared with lean (BCS <3.0) cows, suggesting impaired sphingolipid metabolism (Rico et al., 2015). In the current study, the serum concentrations of glycerophospholipids were inversely associated with overconditioning around calving. Glycerophospholipids (lyso-, diacyl-, and acyl-alkyl-PC) are primarily synthesized in the liver either by the cytidine diphosphate–choline pathway or via sequential methylation of PE and secreted into the circulation as part of blood lipoproteins (Cole et al., 2011, 2012). As reported previously from this experiment (Schuh et al., 2019), the circulating NEFA concentrations in the HBCS group increased earlier and to greater levels than those in the NBCS group, pointing to an increased risk for high lipid mobilization in the HBCS cows. This is further supported by several changes observed in the serum PC. Diacyl-PC have long been of interest to scientists in view of their critical involvement in the secretion of hepatic triacylglycerides as very low density lipoprotein particles (Cole et al., 2012). The liver triglyceride contents of animals have not been assessed in the current study, but the accumulation of hepatic triglycerides is directly related to blood NEFA concentrations (Yuan et al., 2012). It is thus likely that a large influx of NEFA into the liver accompanied by a reduced level of PC may lead to a greater accumulation of triacylglycerides in the liver of HBCS cows. Similar to our findings, others have found decreasing circulating PC levels in obese humans (Pi-

etiläinen et al., 2007) and dairy cows affected by fatty liver disease during the transition period (Imhasly et al., 2014). A further study showed that cows with excessive lipid mobilization (serum NEFA concentrations >0.7 mmol/L) had lower circulating concentrations of long-chain PC (C40:3, C42:5, C42:6) compared with a low-lipolysis group (serum NEFA concentrations <0.4 mmol/L; Humer et al., 2016). Thus, the abnormal decline of certain specific PC and SM could serve as a potential biomarker indicative of fatty liver disease in dairy cows (Imhasly et al., 2014, 2015).

Nutritional management of dairy cows to reach a desired level of BCS at calving should be considered well before the start of the dry period. Thus, late lactation is an important yet often overlooked part of the lactation cycle for controlling body condition efficiently. In this study, we thus aimed to create a deep phenotype that by means of ML would allow for differentiating NBCS and HBCS cows according to their metabolic profile and identify the important metabolites associated with overconditioning. Applying ML, we aimed to move on from descriptive studies toward elucidating the metabolic changes in more detail to improve our understanding of the pathophysiological processes involved. Our results support a working model for cows already differing in BCS at dry-off, which most resembles the farm situation in which cows are not assigned to diets according to their stage of lactation. Our data showed mismatched acetyl-CoA generation versus entry into the TCA cycle, or mitochondrial fatty acid overload, as well as perturbations in blood PC in overconditioned cows that were reflected in higher circulating levels of AcylCN of short- and long-chain length and lower circulating levels of PC. Thus, it is not the rate of β-oxidation per se that is important but rather the way the resultant metabolite acetyl-CoA is further metabolized in the tissues. The implications of these findings are important given the therapeutic potential of targeting mitochondrial metabolism. Targeting mitochondrial metabolism might be associated with concomitant enhancement of hepatic ketone bodies, CO_2, and ATP production as previously documented in human studies. Besides aiming to avoid overcondition in late lactation, targeting these metabolic assets might provide tools for protecting against the metabolic abnormalities associated with overconditioning and consequently leading to a successful transition of overconditioned cows from late pregnancy to lactation.

CONCLUSIONS

Our data provide valuable information about the metabolic footprint of overconditioning in a well-controlled study in periparturient dairy cows. We identified a subset of serum metabolites (mainly BCAA and long-chain AcylCN) that were significantly associated with overconditioning in cows. This finding opens up a new avenue of research toward identifying the individual contributions of these key metabolites in overconditioning. Among the affected pathways, BCAA degradation (before calving) and mitochondrial β-oxidation (after calving) were the most enriched pathways. Overconditioned cows had a greater concentration of long-chain AcylCN in serum, pointing to mitochondrial overload and impairment of β-oxidation rate in early lactation. Some of these metabolic pathways were previously identified in relation to obesity in humans. This study can be extended further to include other data types (i.e., miRNA and mRNA or miRNA and protein), potentially providing a further step in understanding overconditioning in periparturient cows.

REFERENCES

Adams, S. H., C. L. Hoppel, K. H. Lok, L. Zhao, S. W. Wong, P. E. Minkler, D. H. Hwang, J. W. Newman, and W. T. Garvey. 2009. Plasma acylcarnitine profiles suggest incomplete long-chain fatty acid beta-oxidation and altered tricarboxylic acid cycle activity in type 2 diabetic African-American women. J. Nutr. 139:1073–1081.

Bjerre-Harpøth, V., A. C. Storm, M. Eslamizad, B. Kuhla, and M. Larsen. 2015. Effect of propylene glycol on adipose tissue mobilization in postpartum over-conditioned Holstein cows. J. Dairy Sci. 98:8581–8596.

Boldyrev, A. A., G. Aldini, and W. Derave. 2013. Physiology and pathophysiology of carnosine. Physiol. Rev. 93:1803–1845.

Ceciliani, F., C. Lecchi, C. Urh, and H. Sauerwein. 2018. Proteomics and metabolomics characterizing the pathophysiology of adaptive reactions to the metabolic challenges during the transition from late pregnancy to early lactation in dairy cows. J. Proteomics 178:92–106.

Chen, H. H., Y. J. Tseng, S. Y. Wang, Y. S. Tsai, C. S. Chang, T. C. Kuo, W. J. Yao, C. C. Shieh, C. H. Wu, and P. H. Kuo. 2015. The metabolome profiling and pathway analysis in metabolic healthy and abnormal obesity. Int. J. Obes. (Lond.) 39:1241–1248.

Chong, J., O. Soufan, C. Li, I. Caraus, S. Z. Li, G. Bourque, D. S. Wishart, and J. G. Xia. 2018. MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis. Nucleic Acids Res. 46(W1):W486–W494.

Claridge-Chang, A., and P. N. Assam. 2016. Estimation statistics should replace significance testing. Nat. Methods 13:108–109.

Cole, L. K., V. W. Dolinsky, J. R. B. Dyck, and D. E. Vance. 2011. Impaired phosphatidylcholine biosynthesis reduces atherosclerosis and prevents lipotoxic cardiac dysfunction in ApoE(−/−) mice. Circ. Res. 108:686–694.

Cole, L. K., J. E. Vance, and D. E. Vance. 2012. Phosphatidylcholine biosynthesis and lipoprotein metabolism. Biochim. Biophys. Acta 1821:754–761.

Crichton, N. 2000. Information point: Mann-Whitney test. J. Clin. Nurs. 9:583.

Cuperlovic-Culf, M. 2018. Machine learning methods for analysis of metabolic data and metabolic pathway modeling. Metabolites 8:E4.

de Vries, M. J., and R. F. Veerkamp. 2000. Energy balance of dairy cattle in relation to milk production variables and fertility. J. Dairy Sci. 83:62–69.

Dervishi, E., G. Zhang, D. Hailemariam, R. Mandal, D. S. Wishart, and B. N. Ametaj. 2018. Urine metabolic fingerprinting can be used to predict the risk of metritis and highlight the pathobiology of the disease in dairy cows. Metabolomics 14:83. https://doi.org/10.1007/s11306-018-1379-z.

Drackley, J. K. 1999. Biology of dairy cows during the transition period: The final frontier? J. Dairy Sci. 82:2259–2273.

EMA (European Medicines Agency). 2011. Guideline on bioanalytical method validation. Accessed Aug. 23, 2019. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-bioanalytical-method-validation_en.pdf.

Floegel, A., N. Stefan, Z. H. Yu, K. Muhlenbruch, D. Drogan, H. G. Joost, A. Fritsche, H. U. Haring, M. H. de Angelis, A. Peters, M. Roden, C. Prehn, R. Wang-Sattler, T. Illig, M. B. Schulze, J. Adamski, H. Boeing, and T. Pischon. 2013. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. Diabetes 62:639–648.

Frayn, K. N. 2010. Metabolic Regulation: A Human Perspective. 3rd ed. Blackwell Science, Oxford, UK.

Ghaffari, M. H., H. Sadri, K. Schuh, G. Dusel, D. Frieten, C. Koch, C. Prehn, J. Adamski, and H. Sauerwein. 2019. Biogenic amines: Concentrations in serum and skeletal muscle from late pregnancy until early lactation in dairy cows with high versus normal body condition score. J. Dairy Sci. 102:6571–6586.

Goeman, J. J., S. A. van de Geer, F. de Kort, and H. C. van Houwelingen. 2004. A global test for groups of genes: Testing association with a clinical outcome. Bioinformatics 20:93–99.

Greenland, S., S. J. Senn, K. J. Rothman, J. B. Carlin, C. Poole, S. N. Goodman, and D. G. Altman. 2016. Statistical tests, P values, confidence intervals, and power: A guide to misinterpretations. Eur. J. Epidemiol. 31:337–350.

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.

Hailemariam, D., R. Mandal, F. Saleem, S. M. Dunn, D. S. Wishart, and B. N. Ametaj. 2014. Identification of predictive biomarkers of disease state in transition dairy cows. J. Dairy Sci. 97:2680–2693.

Hall, M., E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. 2009. The WEKA data mining software. SIGKDD Explor. 11:10–18.

Holtenius, P., and K. Holtenius. 2007. A model to estimate insulin sensitivity in dairy cows. Acta Vet. Scand. 49:29.

Huber, K., S. Danicke, J. Rehage, H. Sauerwein, W. Otto, U. Rolle-Kampczyk, and M. von Bergen. 2016. Metabotypes with properly functioning mitochondria and anti-inflammation predict extended productive life span in dairy cows. Sci. Rep. 6:24642. https://doi.org/10.1038/srep24642.

Humer, E., A. Khol-Parisini, B. U. Metzler-Zebeli, L. Gruber, and Q. Zebeli. 2016. Alterations of the lipid metabolome in dairy cows experiencing excessive lipolysis early postpartum. PLoS One 11:e0158633. https://doi.org/10.1371/journal.pone.0158633.

Imhasly, S., C. Bieli, H. Naegeli, L. Nystrom, M. Ruetten, and C. Gerspach. 2015. Blood plasma lipidome of dairy cows during the transition period. BMC Vet. Res. 11:252. https://doi.org/10.1186/s12917-015-0565-8.

Imhasly, S., H. Naegeli, S. Baumann, M. von Bergen, A. Luch, H. Jungnickel, S. Potratz, and C. Gerspach. 2014. Metabolomic biomarkers correlating with hepatic lipidosis in dairy cows. BMC Vet. Res. 10:122. https://doi.org/10.1186/1746-6148-10-122.

Issaq, H. J., Q. N. Van, T. J. Waybright, G. M. Muschik, and T. D. Veenstra. 2009. Analytical and statistical approaches to metabolomics research. J. Sep. Sci. 32:2183–2199.

Kell, D. B. 2006. Metabolomics, modelling and machine learning in systems biology—Towards an understanding of the languages of cells. FEBS J. 273:873–894.

Kuhla, B., G. Nurnberg, D. Albrecht, S. Gors, H. M. Hammon, and C. C. Metges. 2011. Involvement of skeletal muscle protein, glycogen, and fat metabolism in the adaptation on early lactation of dairy cows. J. Proteome Res. 10:4252–4262.

Leskinen, T., R. Rinnankoski-Tuikka, M. Rintala, T. Seppanen-Laakso, E. Pollanen, M. Alen, S. Sipila, J. Kaprio, V. Kovanen, P. Rahkila, M. Oresic, H. Kainulainen, and U. M. Kujala. 2010. Differences in muscle and adipose tissue gene expression and cardio-metabolic risk factors in the members of physical activity discordant twin pairs. PLoS One 5:e12609. https://doi.org/10.1371/journal.pone.0012609.

Luo, Z. Z., L. H. Shen, J. Jiang, Y. X. Huang, L. P. Bai, S. M. Yu, X. P. Yao, Z. H. Ren, Y. X. Yang, and S. Z. Cao. 2019. Plasma metabolite changes in dairy cows during parturition identified using untargeted metabolomics. J. Dairy Sci. 102:4639–4650.

Mahadevan, S., S. L. Shah, T. J. Marrie, and C. M. Slupsky. 2008. Analysis of metabolomic data using support vector machines. Anal. Chem. 80:7562–7570.

Mevik, B. H., and R. Wehrens. 2007. The pls package: Principal component and partial least squares regression in R. J. Stat. Softw. 18:1–23.

Mihalik, S. J., B. H. Goodpaster, D. E. Kelley, D. H. Chace, J. Vockley, F. G. S. Toledo, and J. P. DeLany. 2010. Increased levels of plasma acylcarnitines in obesity and type 2 diabetes and identification of a marker of glucolipotoxicity. Obesity (Silver Spring) 18:1695–1700.

Muniandy, M., V. Velagapudi, A. Hakkarainen, J. Lundbom, N. Lundbom, A. Rissanen, J. Kaprio, K. H. Pietiläinen, and M. Ollikainen. 2019. Plasma metabolites reveal distinct profiles associating with different metabolic risk factors in monozygotic twin pairs. Int. J. Obes. (Lond.) 43:487–502.

Newgard, C. B. 2017. Metabolomics and metabolic diseases: Where do we stand? Cell Metab. 25:43–56.

Pietiläinen, K. H., J. Naukkarinen, A. Rissanen, J. Saharinen, P. Ellonen, H. Keranen, A. Suomalainen, A. Gotz, T. Suortti, H. Yki-Jarvinen, M. Oresic, J. Kaprio, and L. Peltonen. 2008. Global transcript profiles of fat in monozygotic twins discordant for BMI: Pathways behind acquired obesity. PLoS Med. 5:e51.

Pietiläinen, K. H., M. Sysi-Aho, A. Rissanen, T. Seppanen-Laakso, H. Yki-Jarvinen, J. Kaprio, and M. Oresic. 2007. Acquired obesity is associated with changes in the serum lipidomic profile independent of genetic effects—A monozygotic twin study. PLoS One 2:e218.

Plaizier, J. C., J. P. Walton, A. Martin, T. Duffield, R. Bagg, P. Dick, and B. W. McBride. 2000. Short communication: Effects of monensin on 3-methylhistidine excretion in transition dairy cows. J. Dairy Sci. 83:2810–2812.

Regazzoni, L., B. de Courten, D. Garzon, A. Altomare, C. Marinello, M. Jakubova, S. Vallova, P. Krumpolec, M. Carini, J. Ukropec, B. Ukropcova, and G. Aldini. 2016. A carnosine intervention study in overweight human volunteers: Bioavailability and reactive carbonyl species sequestering effect. Sci. Rep. 6:27224.

Rico, J. E., V. V. R. Bandaru, J. M. Dorskind, N. J. Haughey, and J. W. McFadden. 2015. Plasma ceramides are elevated in overweight Holstein dairy cows experiencing greater lipolysis and insulin resistance during the transition from late pregnancy to early lactation. J. Dairy Sci. 98:7757–7770.

Rico, J. E., Y. Zang, N. J. Haughey, A. G. Rius, and J. W. McFadden. 2018. Short communication: Circulating fatty acylcarnitines are elevated in overweight periparturient dairy cows in association with sphingolipid biomarkers of insulin resistance. J. Dairy Sci. 101:812–819.

Rinaldo, P., and D. Matern. 2002. Fatty acid oxidation disorders. Annu. Rev. Physiol. 64:477–502.

Roche, J. R., N. C. Friggens, J. K. Kay, M. W. Fisher, K. J. Stafford, and D. P. Berry. 2009. Invited review: Body condition score and its association with dairy cow productivity, health, and welfare. J. Dairy Sci. 92:5769–5801.

Roche, J. R., K. A. Macdonald, K. E. Schutz, L. R. Matthews, G. A. Verkerk, S. Meier, J. J. Loor, A. R. Rogers, J. McGowan, S. R. Morgan, S. Taukiri, and J. R. Webstert. 2013. Calving body condition score affects indicators of health in grazing dairy cows. J. Dairy Sci. 96:5811–5825.

Roche, J. R., S. Meier, A. Heiser, M. D. Mitchell, C. G. Walker, M. A. Crookenden, M. V. Riboni, J. J. Loor, and J. K. Kay. 2015. Effects of precalving body condition score and prepartum feeding level on production, reproduction, and health parameters in pasture-based transition dairy cows. J. Dairy Sci. 98:7164–7182.

Romisch-Margl, W., C. Prehn, R. Bogumil, C. Rohring, K. Suhre, and J. Adamski. 2012. Procedure for tissue sample preparation and metabolite extraction for high-throughput targeted metabolomics. Metabolomics 8:133–142.

Sadri, H., F. Giallongo, A. N. Hristov, J. Werner, C. H. Lang, C. Parys, B. Saremi, and H. Sauerwein. 2016. Effects of slow-release urea and rumen-protected methionine and histidine on mammalian target of rapamycin (mTOR) signaling and ubiquitin proteasome-related gene expression in skeletal muscle of dairy cows. J. Dairy Sci. 99:6702–6713.

Sahoo, A., A. K. Pattanaik, and T. K. Goswami. 2009. Immunobiochemical status of sheep exposed to periods of experimental protein deficit and realimentation. J. Anim. Sci. 87:2664–2673.

Schuh, K., H. Sadri, S. Haussler, L. A. Webb, C. Urh, M. Wagner, C. Koch, J. Frahm, S. Danicke, G. Dusel, and H. Sauerwein. 2019. Comparison of performance and metabolism from late pregnancy to early lactation in dairy cows with elevated v. normal body condition at dry-off. Animal 13:1478–1488.

She, P., C. Van Horn, T. Reid, S. M. Hutson, R. N. Cooney, and C. J. Lynch. 2007. Obesity-related elevations in plasma leucine are associated with alterations in enzymes involved in branched-chain amino acid metabolism. Am. J. Physiol. Endocrinol. Metab. 293:E1552–E1563.

Sigauke, E., D. Rakheja, K. Kitson, and M. J. Bennett. 2003. Carnitine palmitoyltransferase II deficiency: A clinical, biochemical, and molecular review. Lab. Invest. 83:1543–1554.

Stegen, S., B. Stegen, G. Aldini, A. Altomare, L. Cannizzaro, M. Orioli, S. Gerlo, L. Deldicque, M. Ramaekers, P. Hespel, and W. Derave. 2015. Plasma carnosine, but not muscle carnosine, attenuates high-fat diet-induced metabolic stress. Appl. Physiol. Nutr. Metab. 40:868–876.

van den Berg, R. A., H. C. J. Hoefsloot, J. A. Westerhuis, A. K. Smilde, and M. J. van der Werf. 2006. Centering, scaling, and transformations: Improving the biological information content of metabolomics data. BMC Genomics 7:142. https://doi.org/10.1186/1471-2164-7-142.

Yang, Y., H. Sadri, C. Prehn, J. Adamski, J. Rehage, S. Danicke, B. Saremi, and H. Sauerwein. 2019. Acylcarnitine profiles in serum and muscle of dairy cows receiving conjugated linoleic acids or a control fat supplement during early lactation. J. Dairy Sci. 102:754–767.

Yuan, K., R. D. Shaver, S. J. Bertics, M. Espineira, and R. R. Grummer. 2012. Effect of rumen-protected niacin on lipid metabolism, oxidative stress, and performance of transition dairy cows. J. Dairy Sci. 95:2673–2679.

Zandkarimi, F., J. Vanegas, X. Fern, C. S. Maier, and G. Bobe. 2018. Metabotypes with elevated protein and lipid catabolism and inflammation precede clinical mastitis in prepartal transition dairy cows. J. Dairy Sci. 101:5531–5548.

Zhang, G., Q. Deng, R. Mandal, D. S. Wishart, and B. N. Ametaj. 2017a. Metabolomics-based profiling identifies serum signatures that predict the risk of metritis in transition dairy cows. J. Anim. Sci. 95(Suppl. 4):168. (Abstr.)

Zhang, G., Q. L. Deng, R. Mandal, D. S. Wishart, and B. N. Ametaj. 2017b. DI/LC-MS/MS-based metabolic profiling for identification of early predictive serum biomarkers of metritis in transition dairy cows. J. Agric. Food Chem. 65:8510–8521.

Zhang, G. S., E. Dervishi, S. M. Dunn, R. Mandal, P. Liu, B. Han, D. S. Wishart, and B. N. Ametaj. 2017c. Metabotyping reveals distinct metabolic alterations in ketotic cows and identifies early predictive serum biomarkers for the risk of disease. Metabolomics 13:43.

Zukunft, S., M. Sorgenfrei, C. Prehn, G. Moller, and J. Adamski. 2013. Targeted metabolomics of dried blood spot extracts. Chromatographia 76:1295–1305.