

Machine learning approach reveals a metabolic signature of over-conditioned cows

**M. H. Ghaffari¹, A. Jahanbekam², H. Sadri³, K. Schuh^{1,4},
G. Dusel⁴, C. Prehn⁵, J. Adamski⁵, C. Koch⁶, H. Sauerwein¹**

¹ Institute of Animal Science, University of Bonn, Germany.

² Department of Epileptology, University of Bonn, Germany.

³ Department of Clinical Science, University of Tabriz, Iran.

⁴ Department of Life Sciences and Engineering, University of Applied Sciences Bingen, Germany.

⁵ Institute of Experimental Genetics, Neuherberg, Germany.

⁶ Educational and Research Centre for Animal Husbandry, Hofgut Neumuehle, Germany.



Background

- **Transition period:** Metabolic and physiological changes
- **Overconditioning:** Greater risk of metabolic disorders

(Roche et al., 2009, 2013, Bjerre-Harpøth et al., 2015).

Analytical techniques:

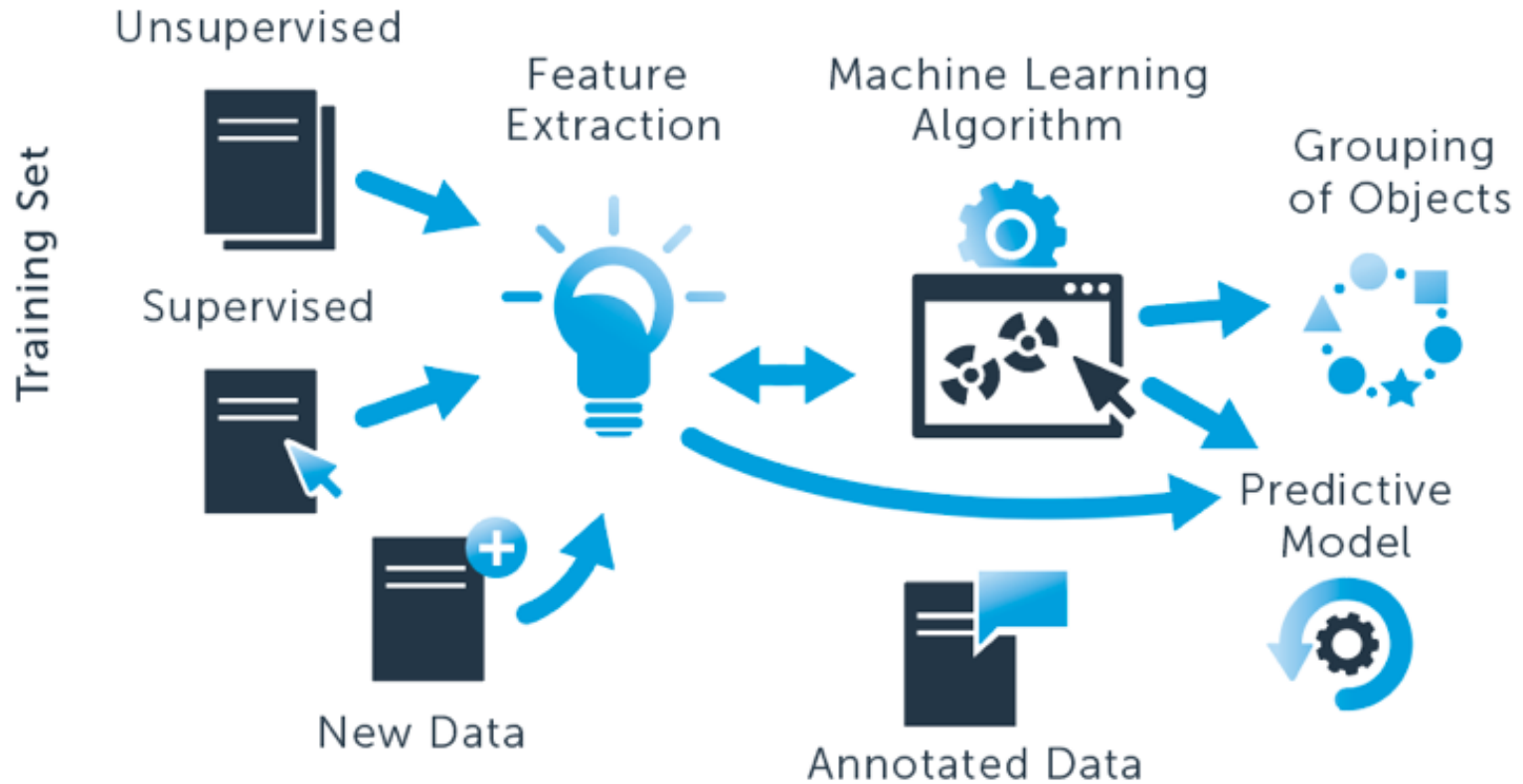
- largely focused on serum classical variables such as BHB, NEFA, insulin, leptin and etc.
- Approaches that may not telling the whole story about the metabolic dysfunction in transition dairy cows.

Metabolomics approaches:

- A wider range of metabolite classes
- Identification of key metabolites within metabolic pathways

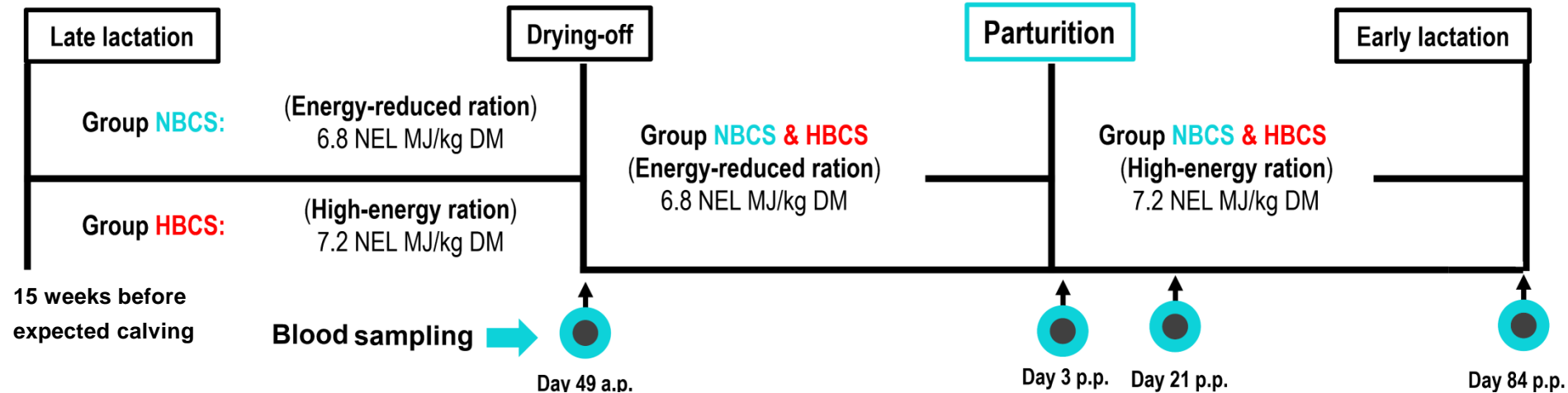
Background

Machine Learning



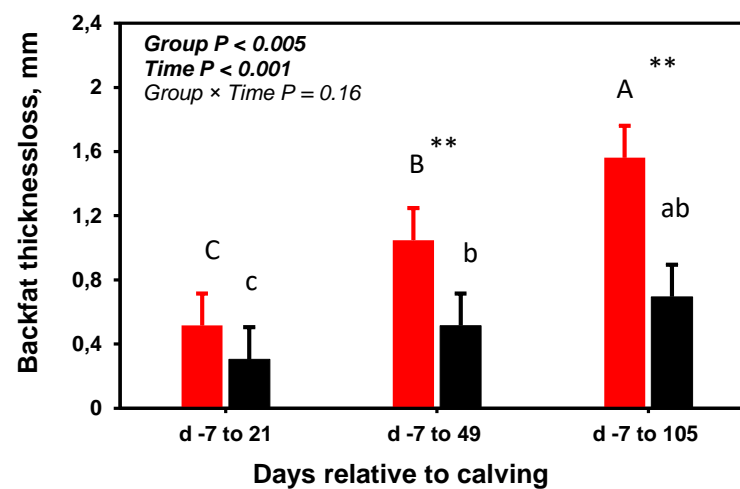
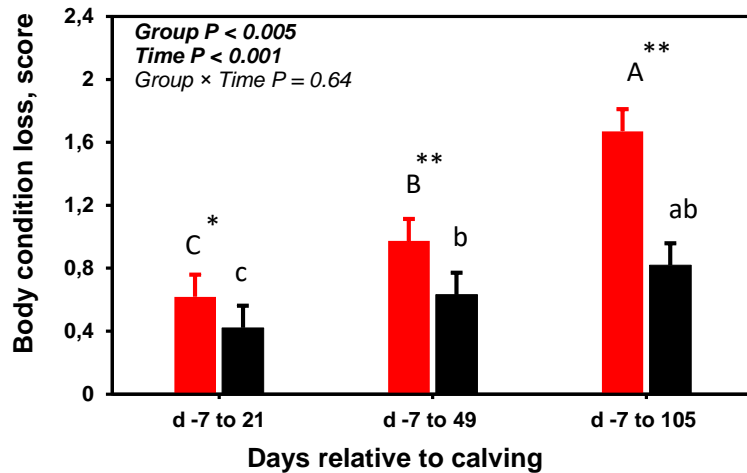
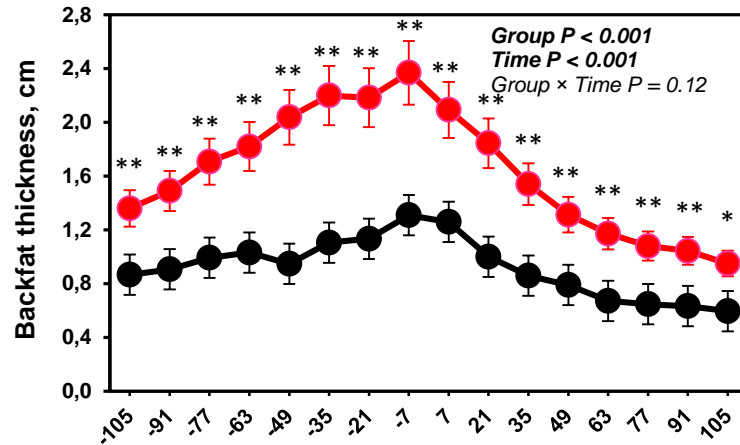
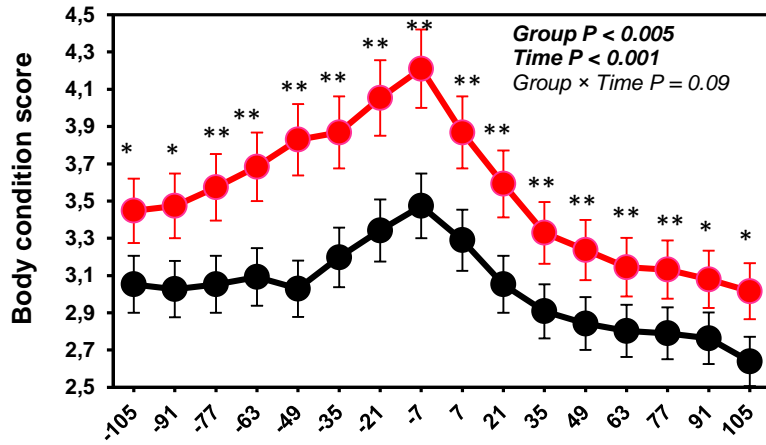
<http://mysoftheaven.com/service/machine-learning>

Study design



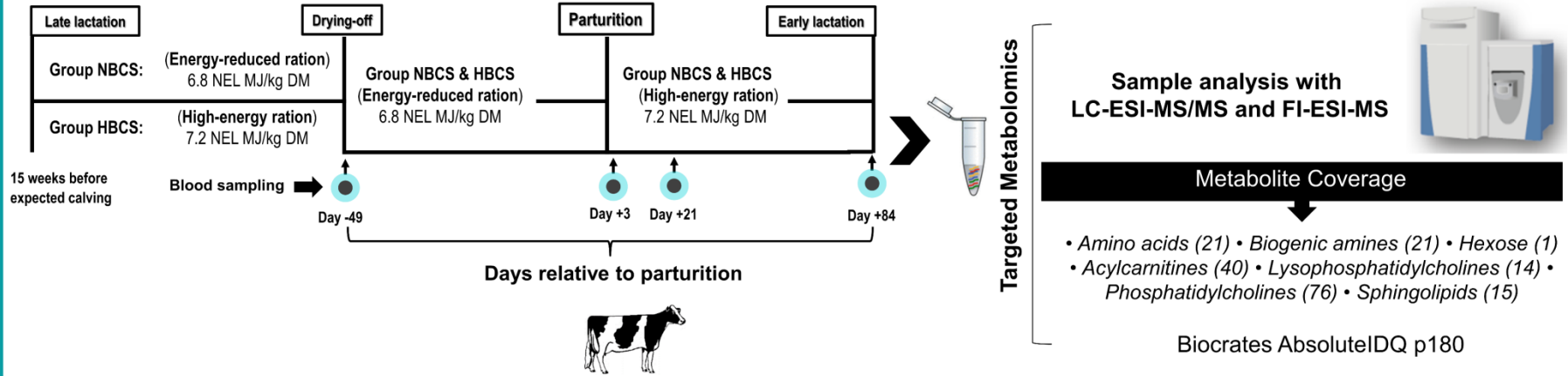
The targeted body condition score (BCS) and backfat thickness (BFT) at dry-off:
HBCS: BCS >3.75 and BFT >1.4 cm
NBCS: BCS <3.5 and < BFT 1.2 cm

Study design



Study design

Experiment and laboratory analyses

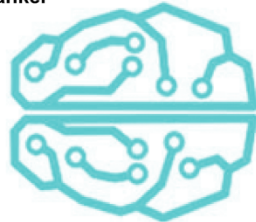


Statistical analysis

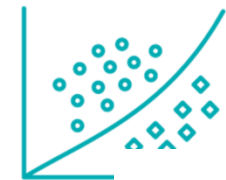
Targeted Metabolomics Data



- Sequential Minimal Optimization
- Random Forest
- Alternating Decision Tree
- Naive Bayes
- Ranker



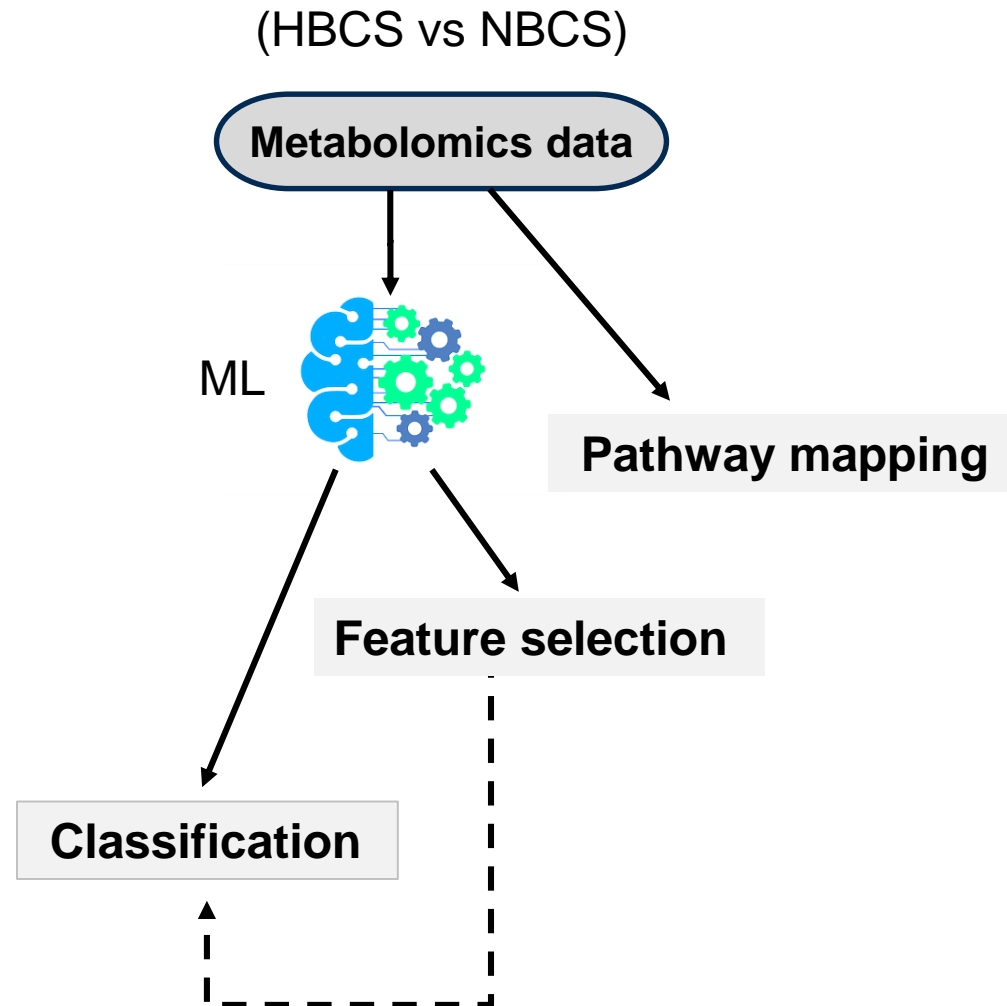
- Pathway analysis
- Identifying important metabolites
- Group classification



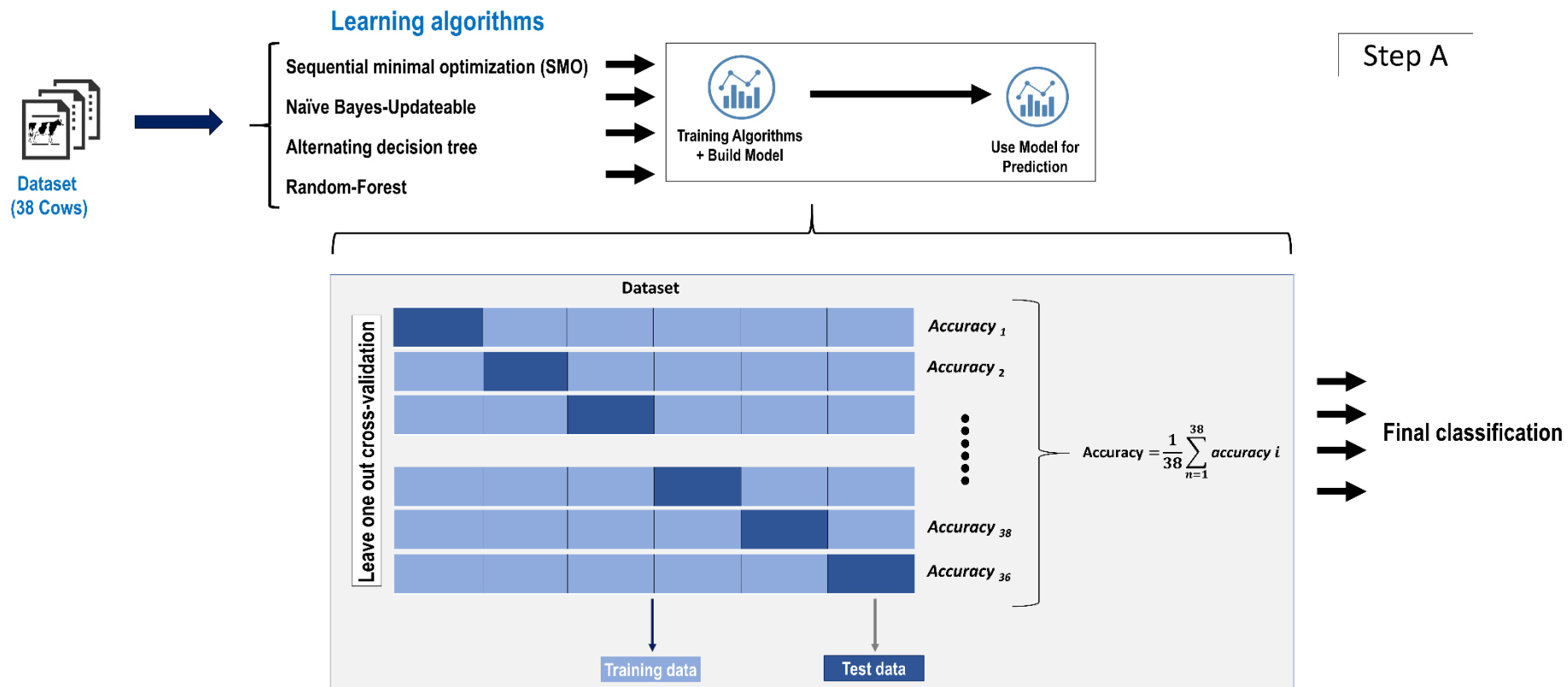
Objectives

- To investigate the differences in the metabolic profiles in serum of dairy cows that were normal or over-conditioned around calving.
- To identify important serum metabolites that are associated with over-conditioning using different machine learning algorithms.
- To identify metabolic pathways that are significantly enriched in the over-conditioned cows.

Overview of statistics analysis and bioinformatics



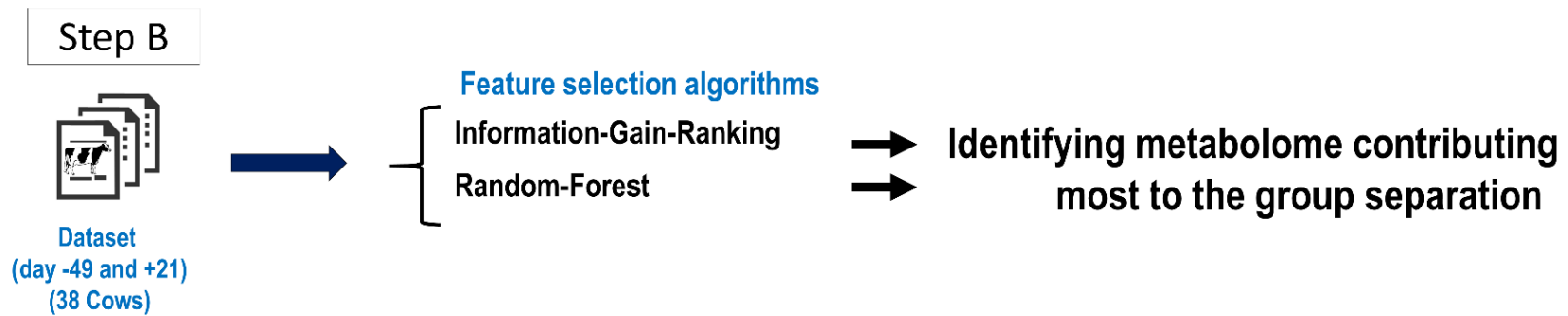
Methods: Supervised machine learning

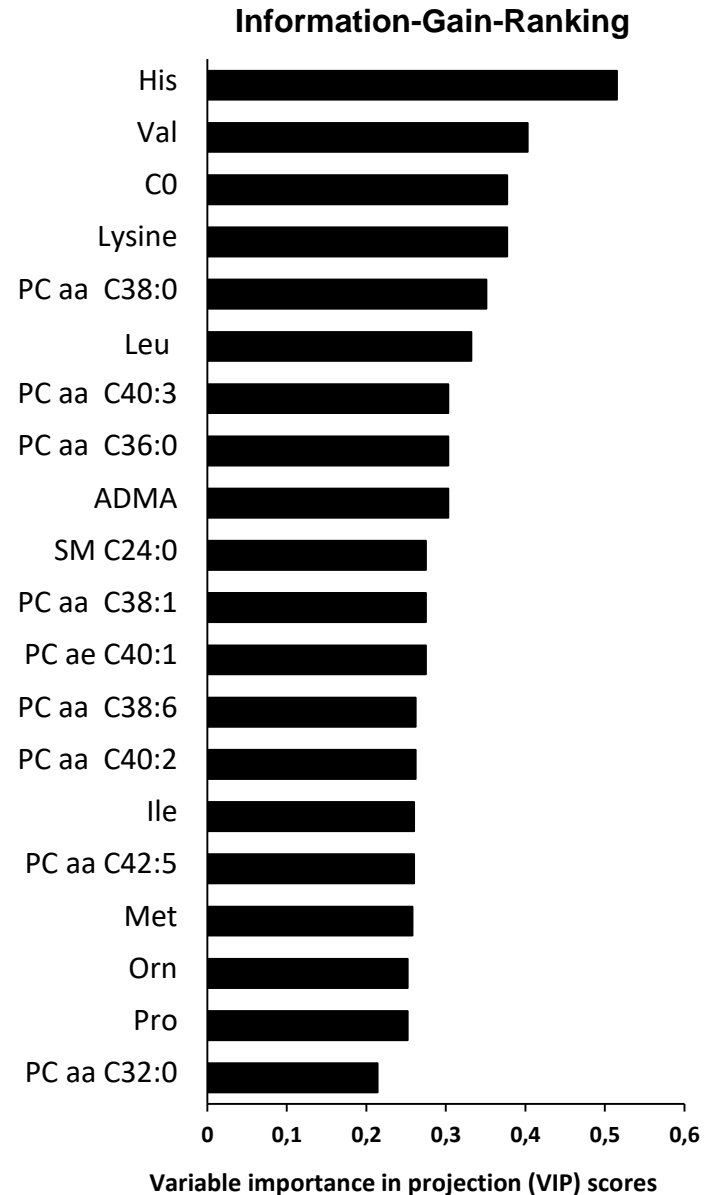
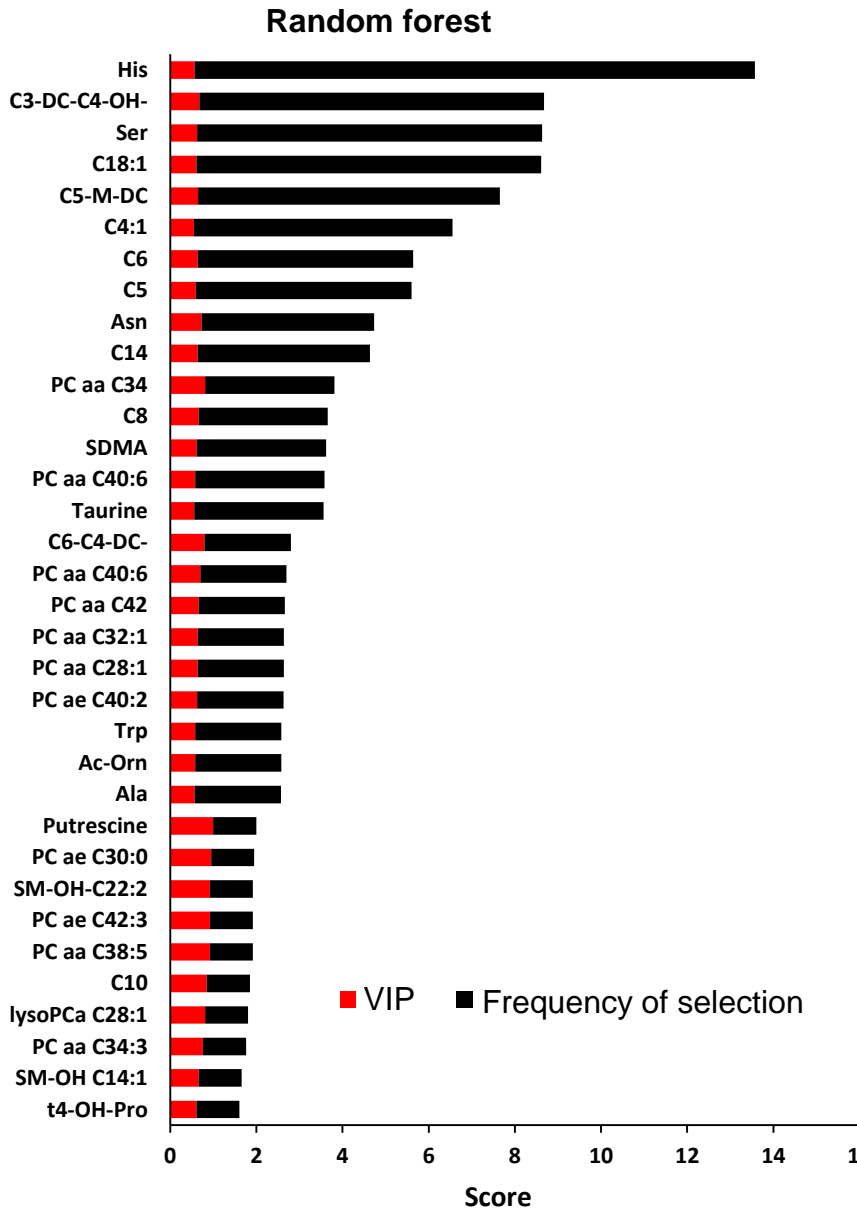


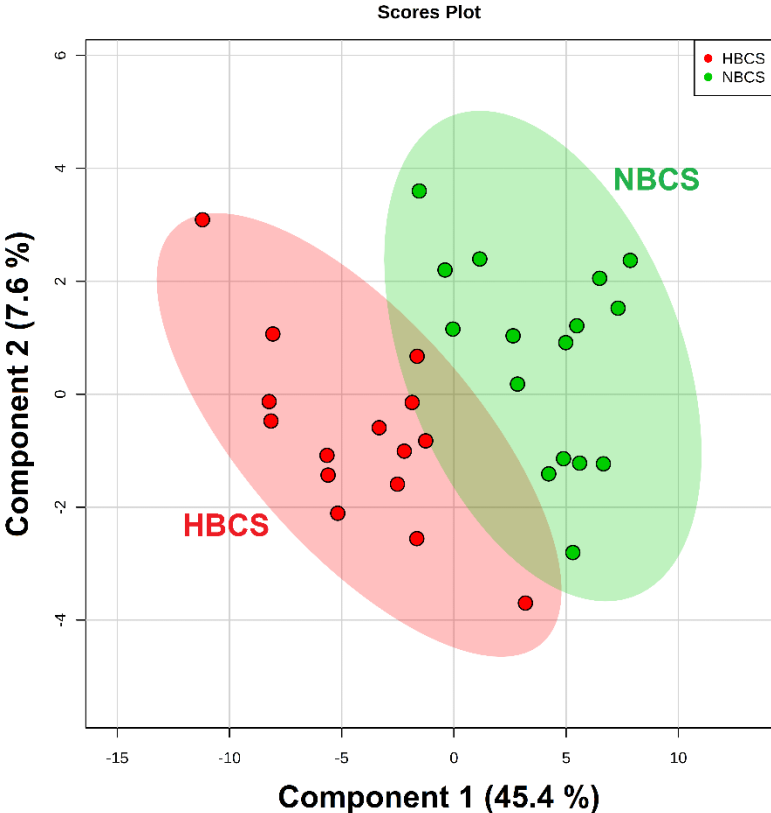
Results: Machine learning classification

Model	Day relative to calving	Sensitivity (%)	Specificity (%)	Accuracy (%)
Sequential Minimal Optimization (SMO)	D -49	68	73	71
	D +3	68	63	65
	D +21	52	68	60
	D +84	73	78	76
Random Forest	D -49	73	73	73
	D +3	52	68	60
	D +21	68	73	71
	D +84	68	57	63
Alternating Decision Tree	D -49	68	84	76
	D +3	57	47	52
	D +21	52	57	55
	D +84	68	68	68
Naïve Bayes-Updatable	D -49	68	73	71
	D +3	57	63	60
	D +21	57	73	65
	D +84	63	63	63

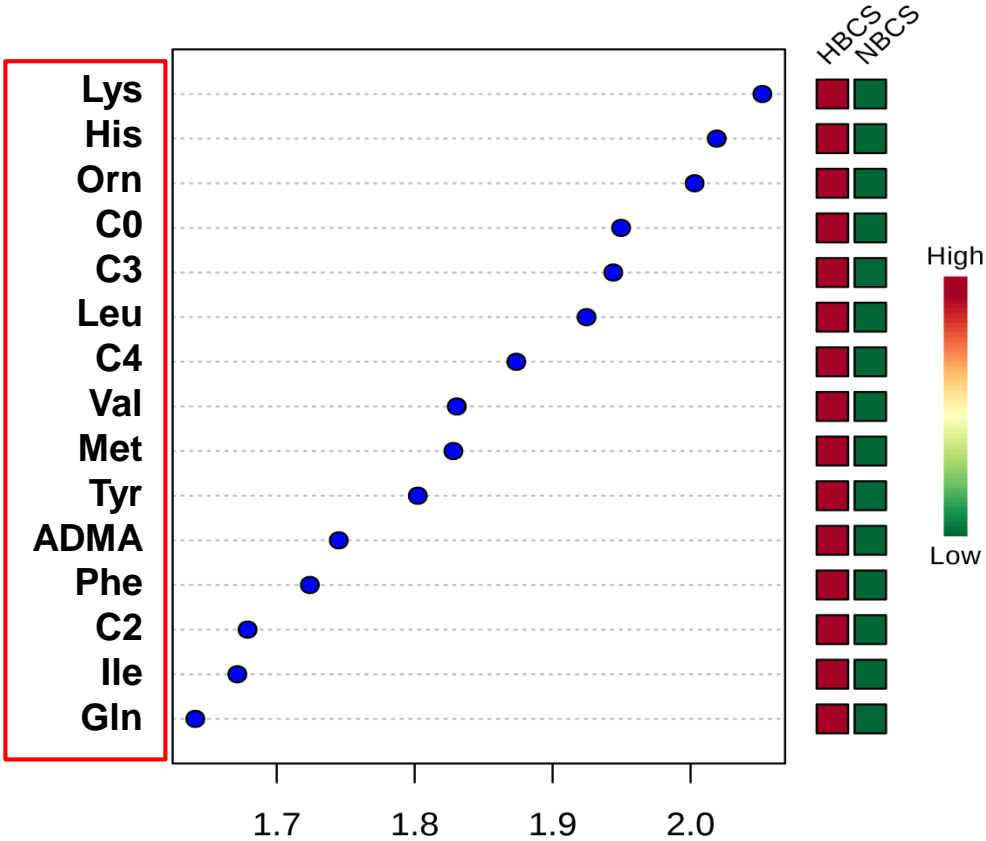
Method: Feature selection





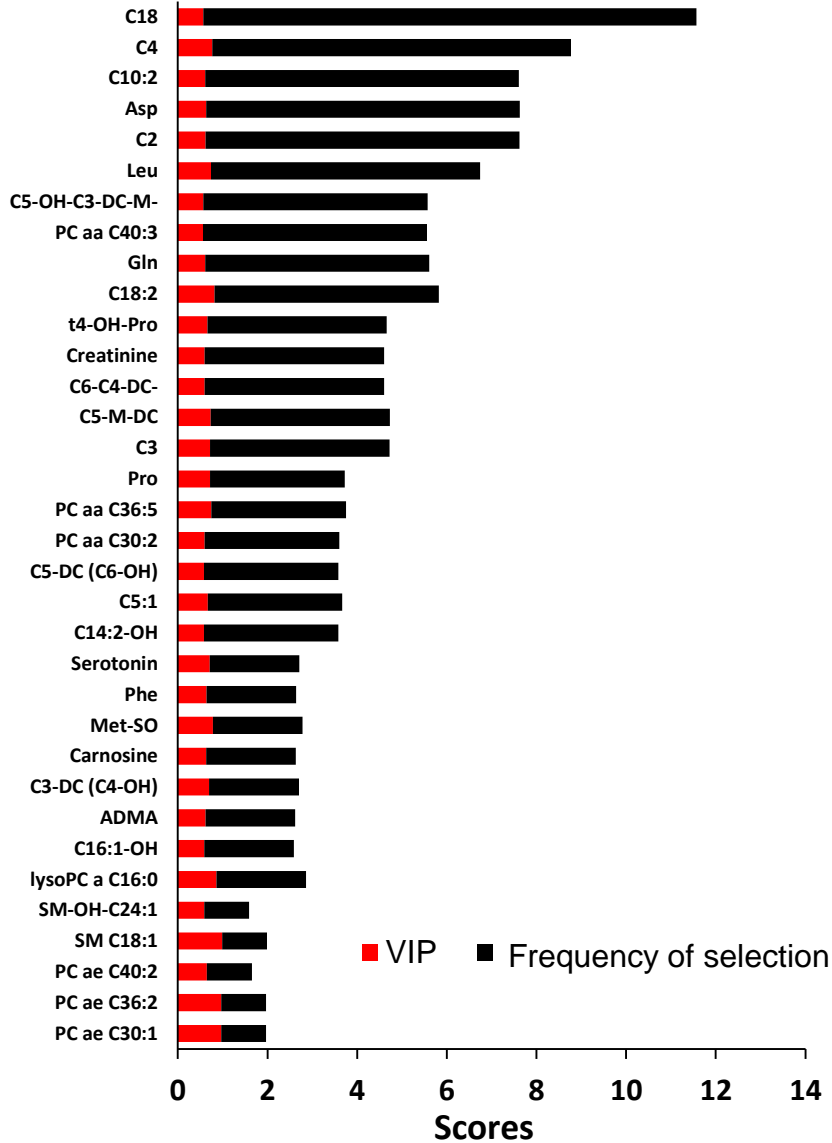


Partial Least Squares - Discriminant Analysis (PLS-DA)

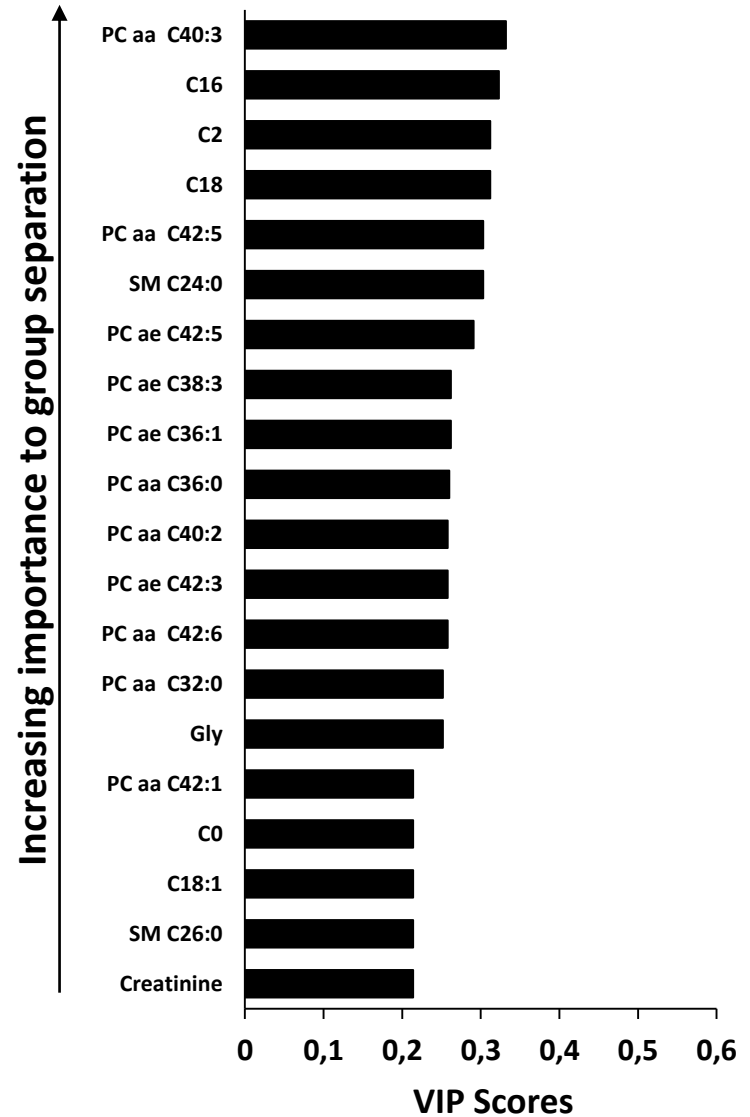


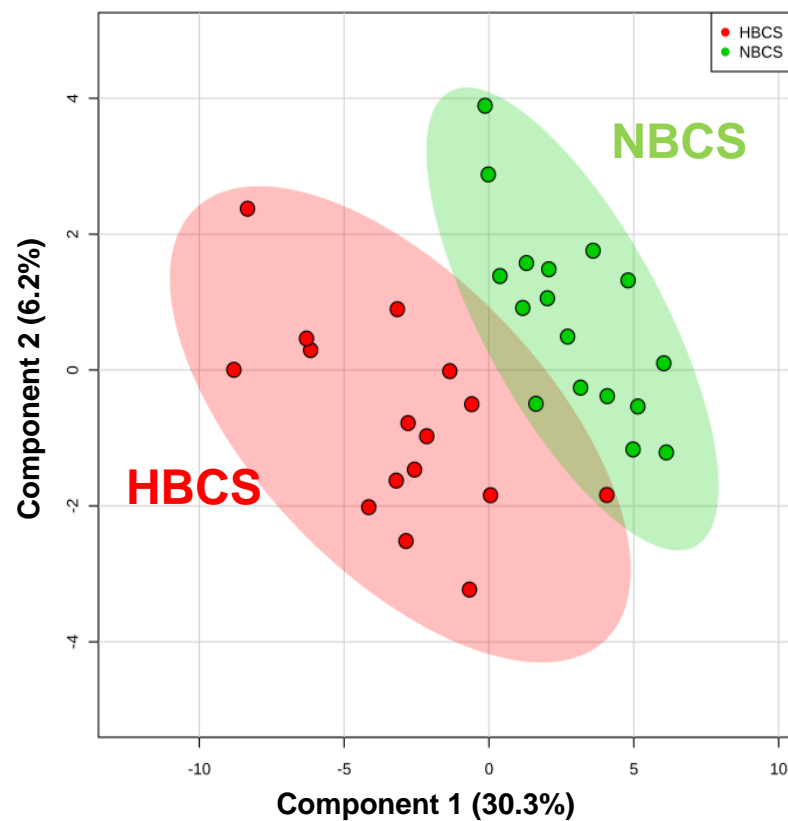
Variable importance in projection (VIP)

Random forest

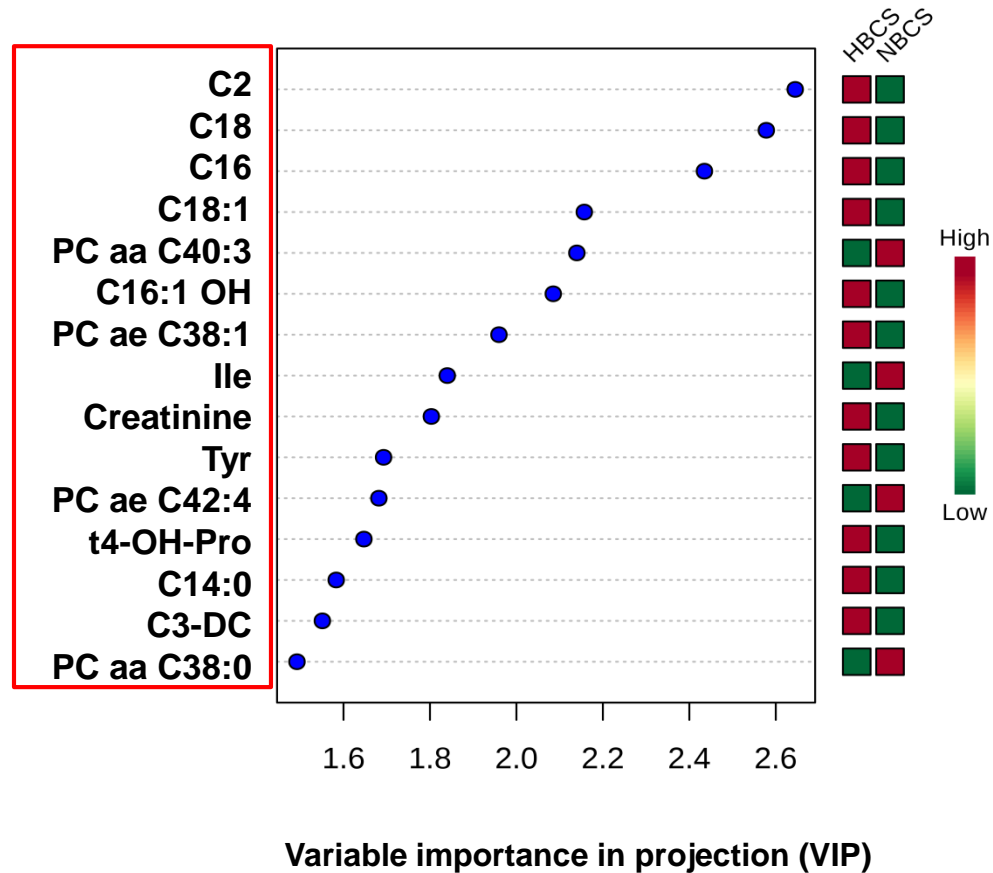


Information-Gain-Ranking





Partial Least Squares - Discriminant Analysis (PLS-DA)



Results: Feature selection

In a follow-up analysis, common important features (serum metabolites) across the three ML algorithms (RF, IGR, and PLS-DA) were selected for further analysis.

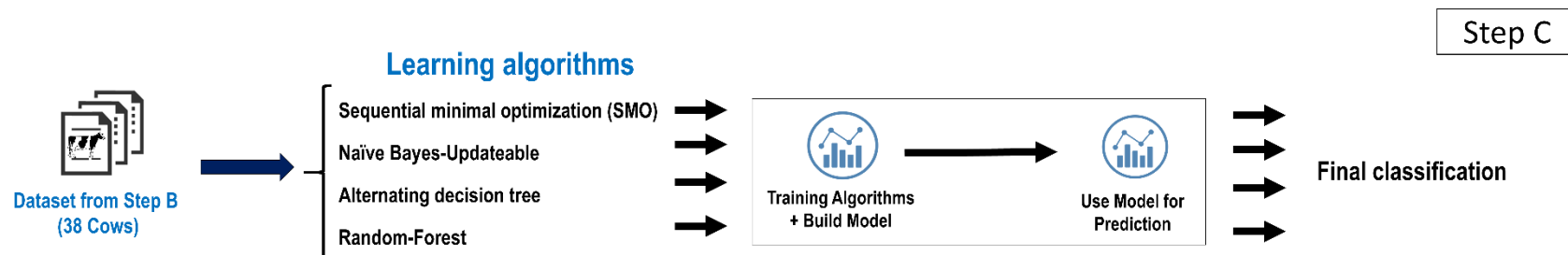
Selected features: d -49

ADMA (Asymmetric dimethylarginine)
Carnitine (C0)
Histidine
Isoleucine
leucine
Lysine
Ornithine
Proline
Serine
Valine
PC aa C36:0
PC aa C40:2

Selected features: d +21

Hexadecanoyl-L-carnitine (C16)
Octadecanoyl-L-carnitine (C18)
Hydroxyhexadecanoyl-L-carnitine (C16:1 OH)
Octadecenoyl-L-carnitine (C18:1)
Acetylcarnitine (C2)
Malonylcarnitine
Carnosine
Creatinine
Glycine
Isoleucine
PC aa C40:3
Hydroxyproline

Results: Machine learning classification



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

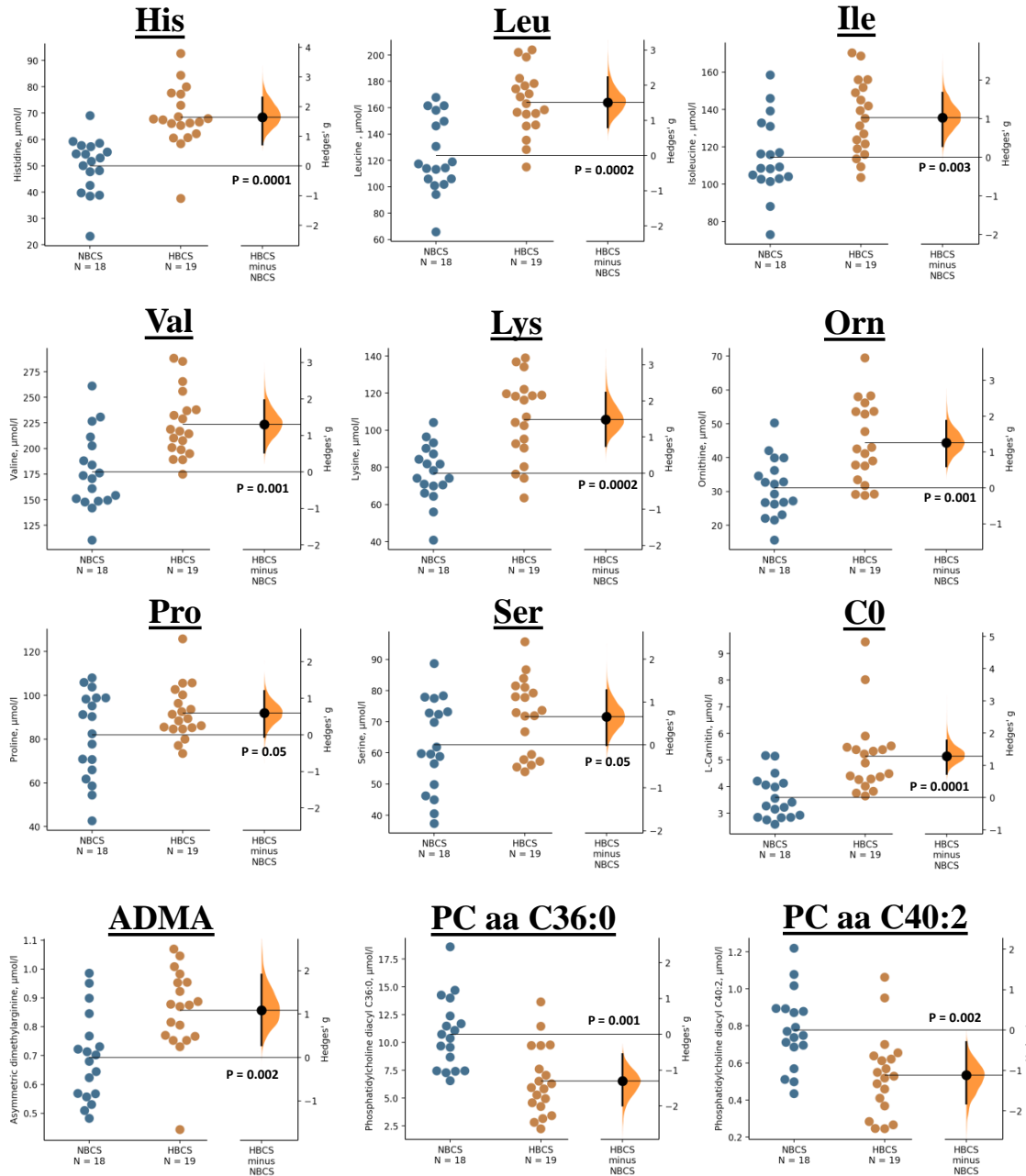
Model	Days relative to calving	Sensitivity (%)	Specificity (%)	Accuracy (%)
Sequential Minimal Optimization (SMO)	D -49	0.84	0.74	78.9
	D +21	0.82	0.74	78.9
Random Forest	D -49	0.84	0.74	78.9
	D +21	0.74	0.79	76.3
Alternating Decision Tree	D -49	0.68	0.79	73.7
	D +21	0.53	0.68	60.5
Naïve Bayes-Updateable	D -49	0.79	0.74	76.3
	D +21	0.84	0.74	78.9

Results: d -49

Selected metabolites

Serum concentration, $\mu\text{mol/L}$

Effect size (Hedges' g)

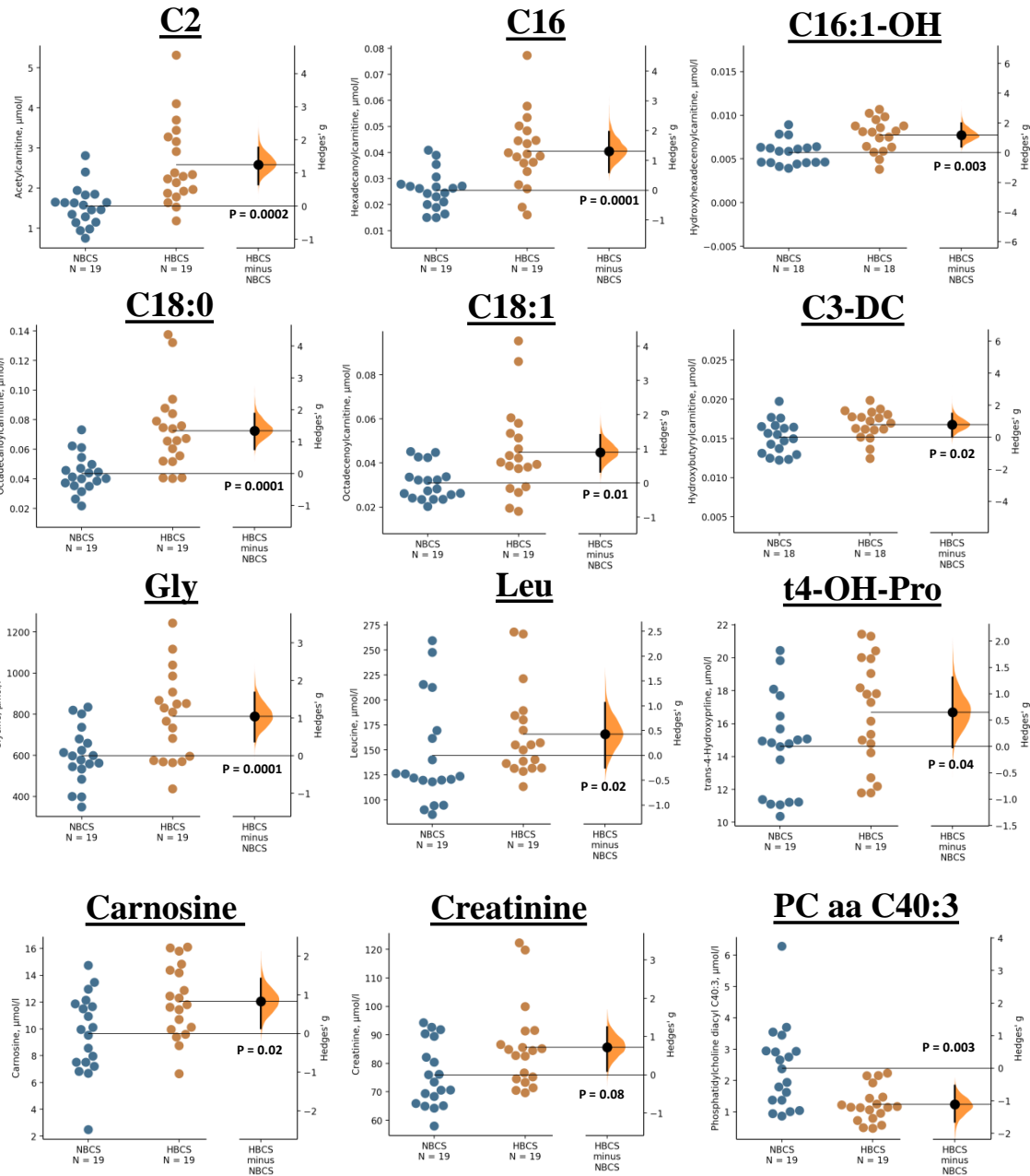


The **effect size (Hedges' g)** between NBCS and HBCS was shown in the above Gardner-Altman estimation plot. **P-value** = the Mann-Whitney U test. Effect size as per standard practice referred to as either 'trivial' ($g < 0.2$), 'small' ($0.2 < g < 0.5$), 'moderate' ($0.5 < g < 0.8$) or 'large' ($g > 0.8$).

Results: d +21

Selected metabolites

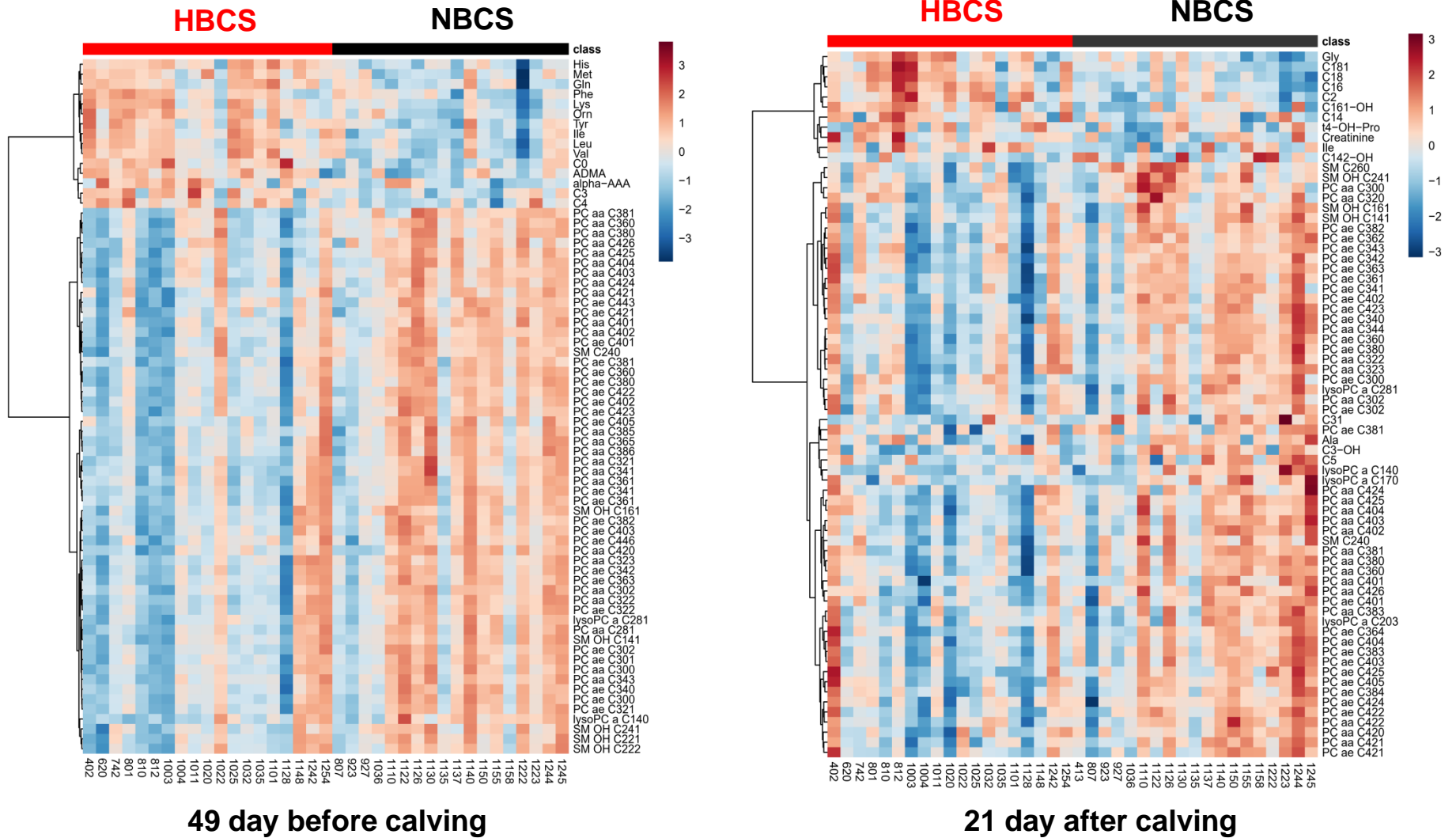
Serum concentration, $\mu\text{mol/L}$



Effect size (Hedges' g)

The effect size (Hedges' g) between NBCS and HBCS was shown in the above Gardner-Altman estimation plot. P-value = the Mann-Whitney U test. Effect size as per standard practice referred to as either 'trivial' ($g < 0.2$), 'small' ($0.2 < g < 0.5$), 'moderate' ($0.5 < g < 0.8$) or 'large' ($g > 0.8$).

Results: Heatmap of serum metabolites (top 70)



Quantitative enrichment analysis (QEA)
was performed using the metabolic pathways library



to identify biologically meaningful patterns that are significantly enriched in quantitative metabolomic data of **HBCS** compared to **NBCS**.

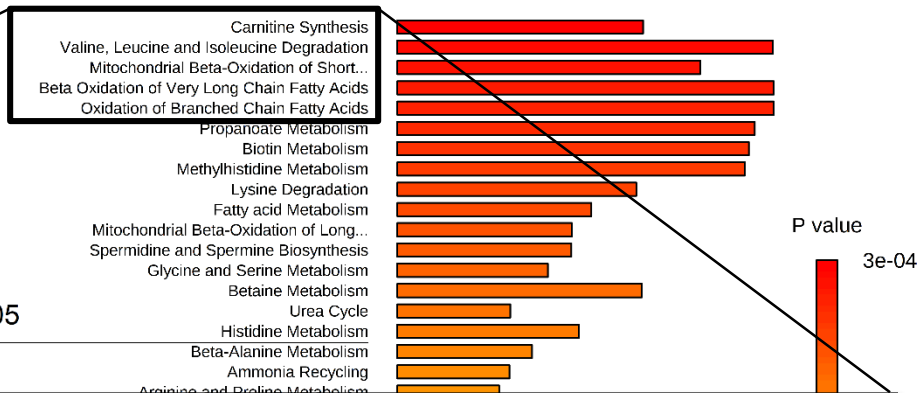
This web-based facility uses

- KEGG metabolic pathways
- Novel algorithms and concepts (i.e., Global Test) with pathway analysis.

Results

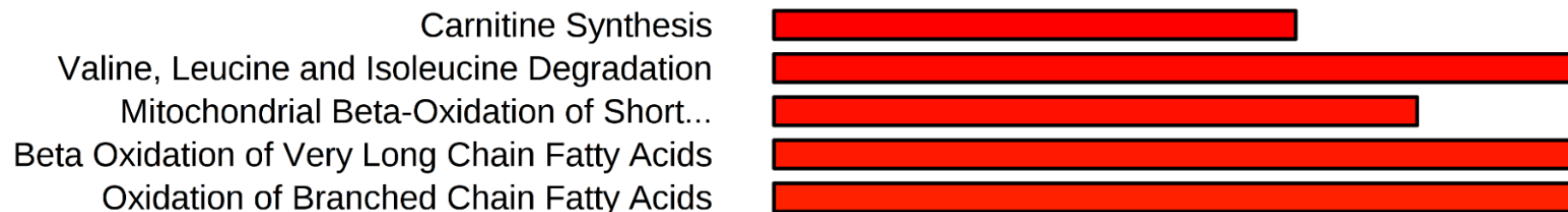
Day -49 relative to calving

Metabolite Sets Enrichment Overview HBCS vs. NBCS



P < 0.05

Metabolite Sets Enrichment Overview HBCS vs. NBCS

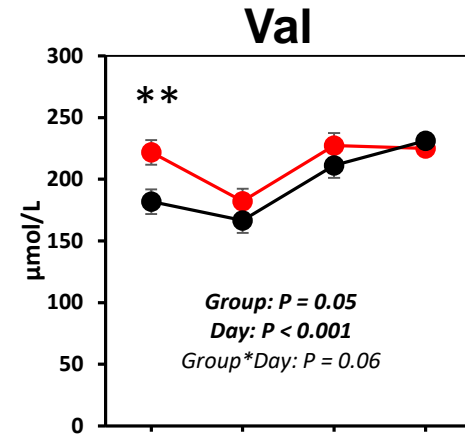
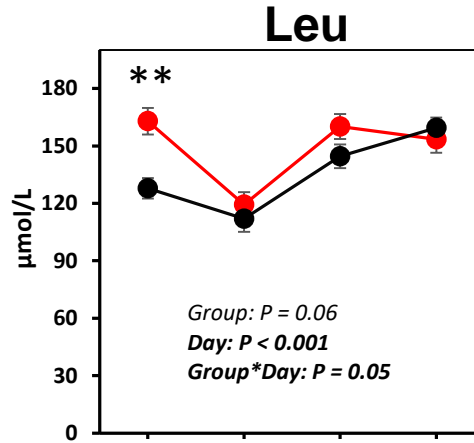
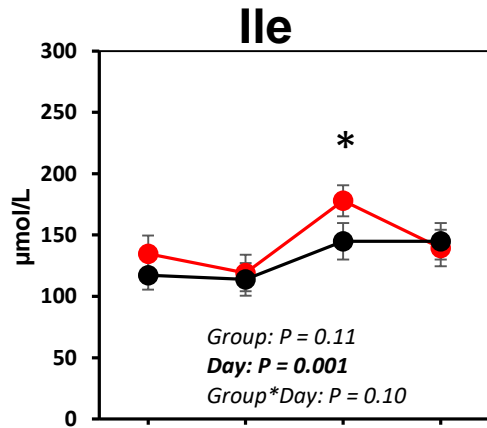


0 2 4 6 8 10 12

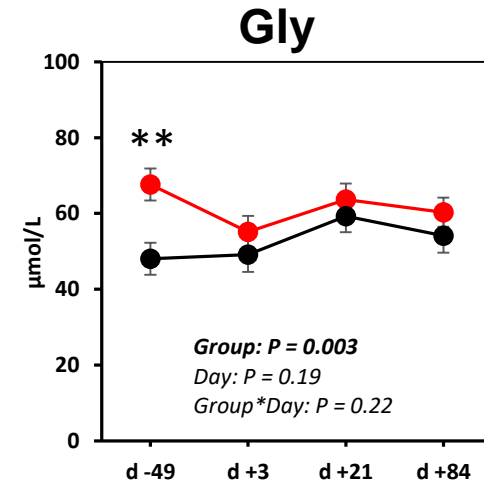
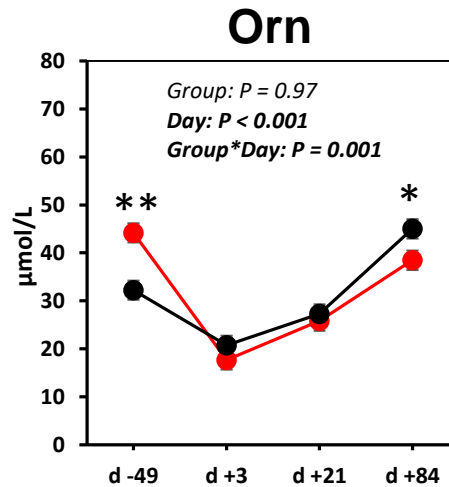
Fold Enrichment

The bar chart size indicates the pathway enrichment, and the color density represents the difference. Pathways with a larger size and darker red color are the more enriched pathways

Results: Alterations of the some serum amino acids



NBCS
 HBCS



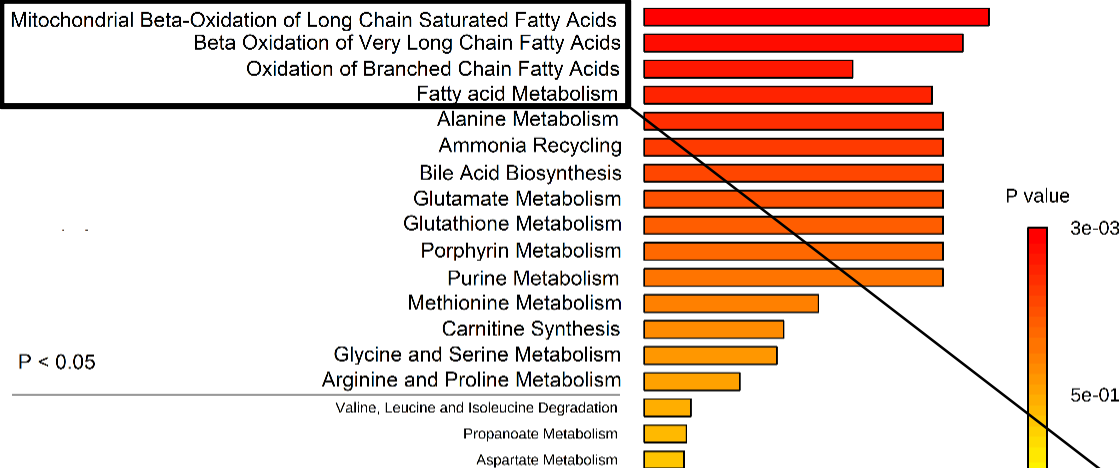
Days relative to calving

Days relative to calving

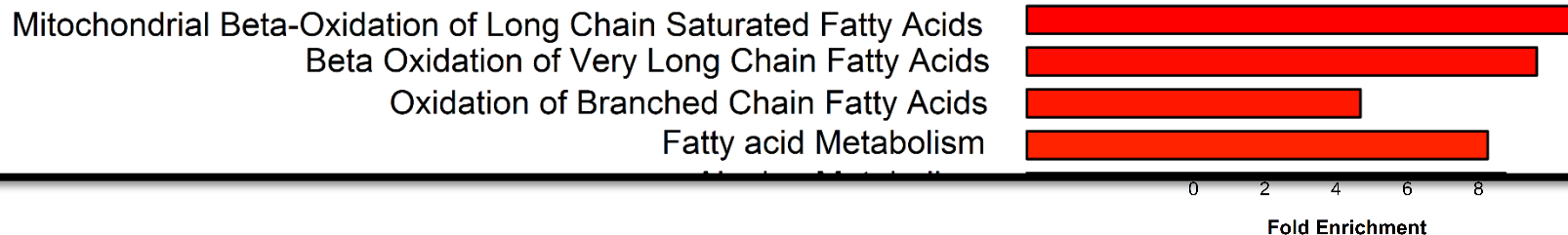
Results

Day +21 relative to calving

Metabolite Sets Enrichment Overview HBCS vs. NBCS



Metabolite Sets Enrichment Overview HBCS vs. NBCS

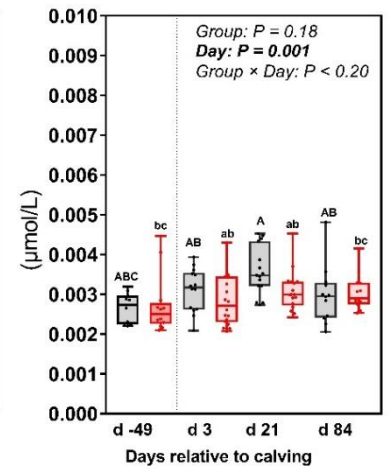
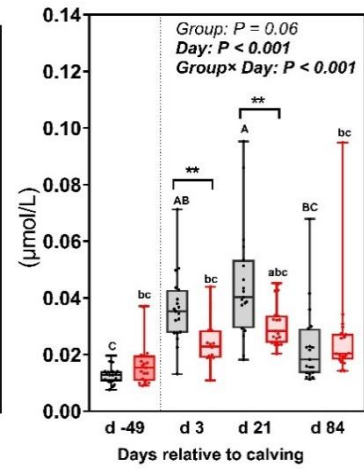
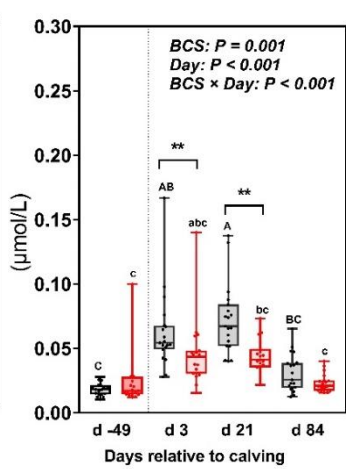
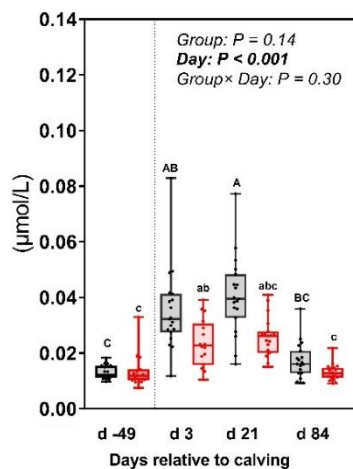
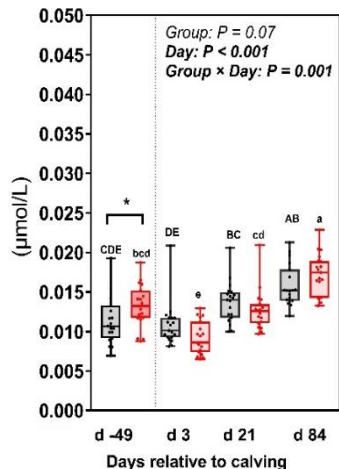
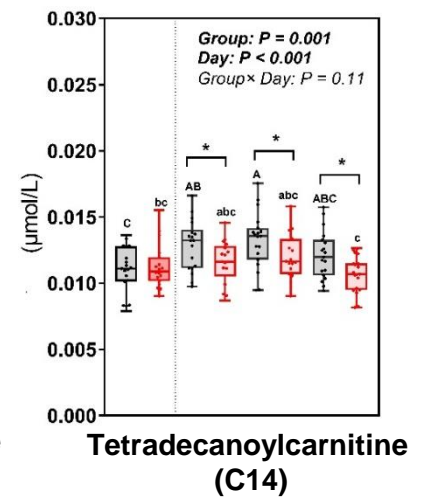
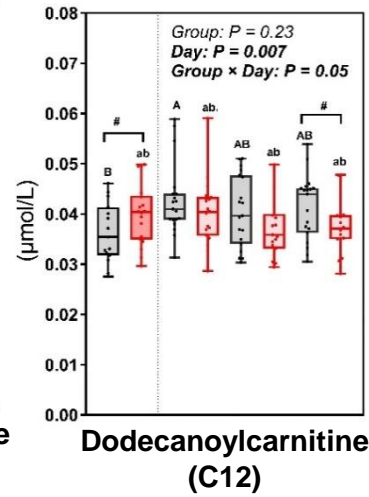
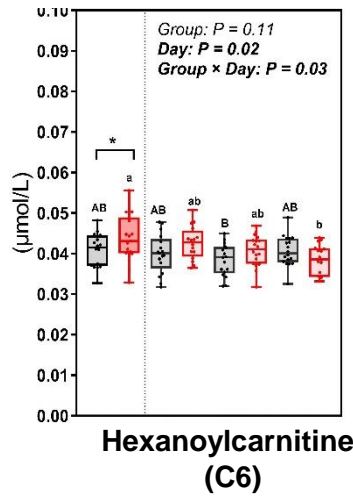
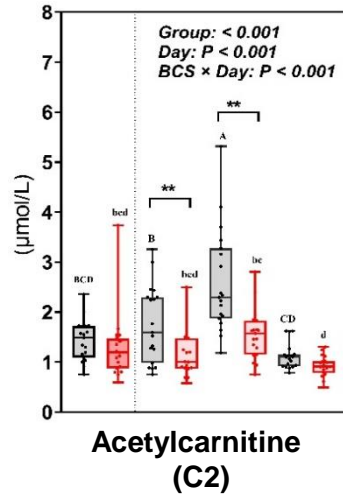
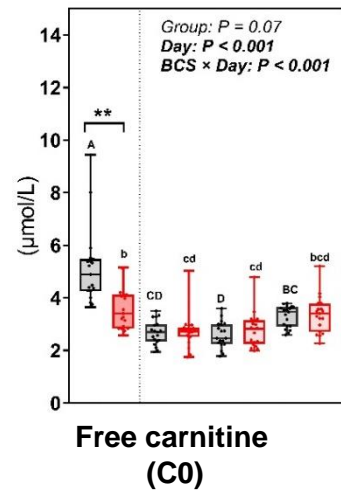


The bar chart size indicates the pathway enrichment, and the color density represents the difference. Pathways with a larger size and darker red color are the more enriched pathways

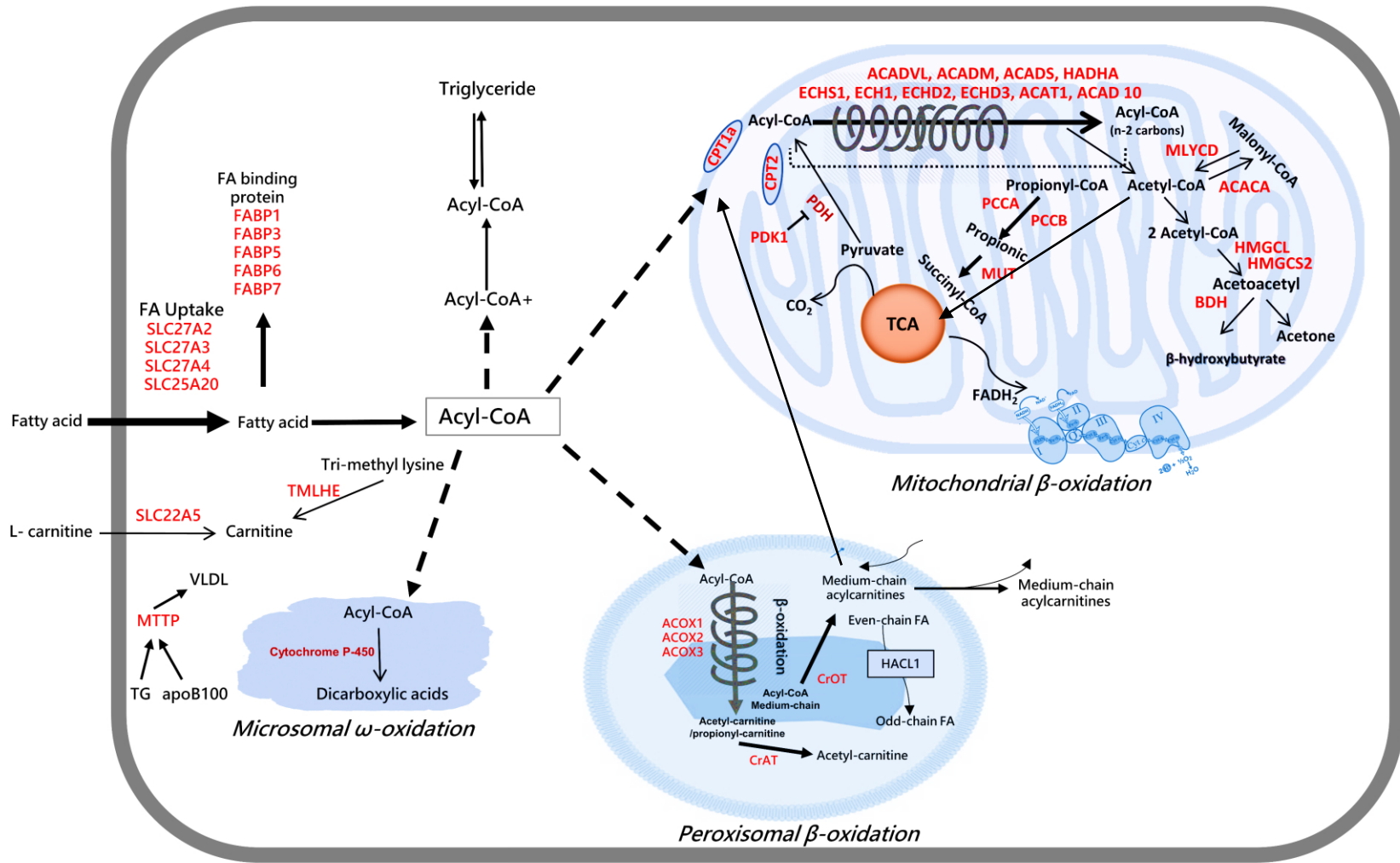
Results:

Alterations of the acylcarnitine profiles in serum

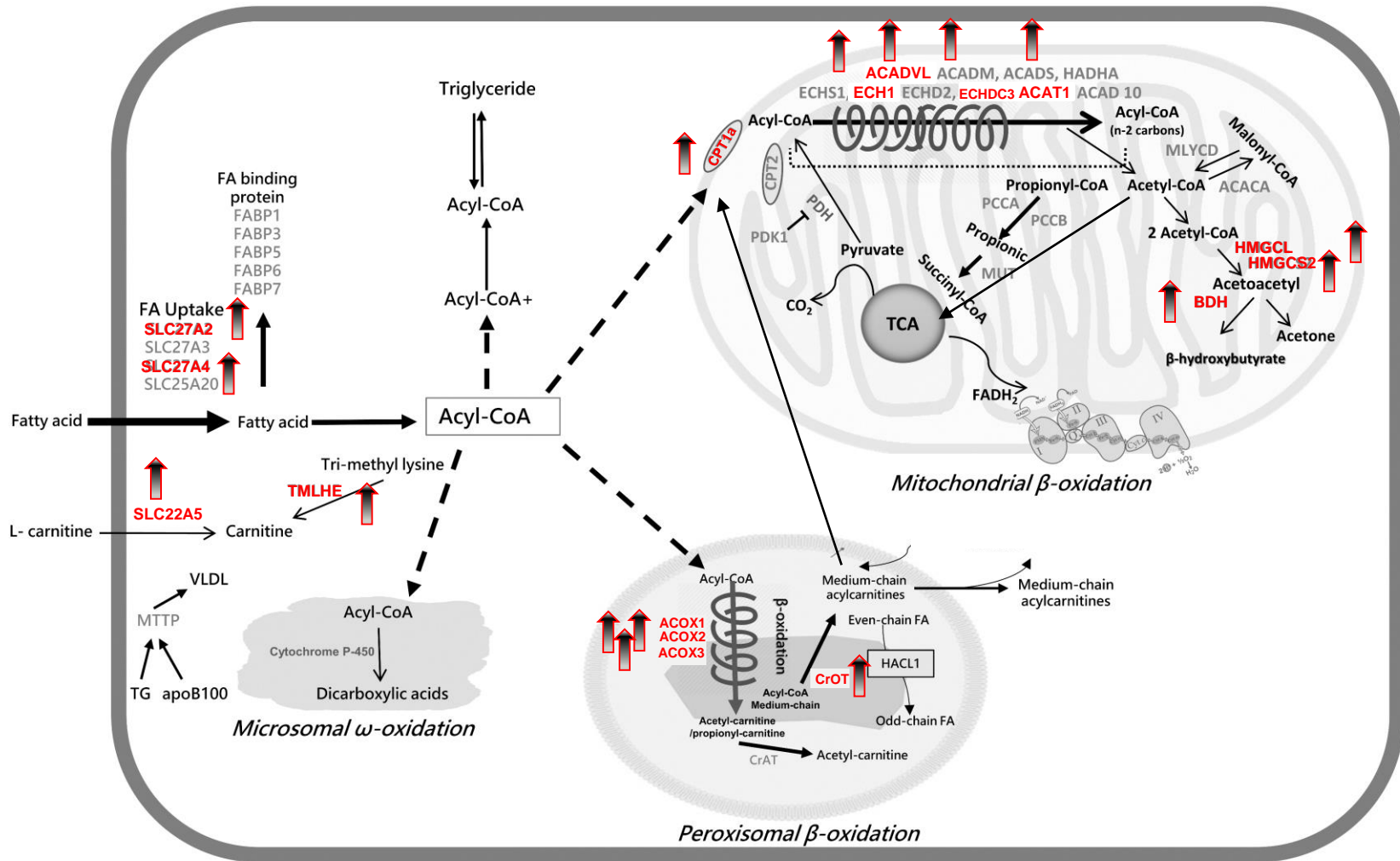
HBCS
 NBCS



Results Hepatic gene expression of fatty acid oxidation and carnitine metabolism



Results Hepatic gene expression of fatty acid oxidation and carnitine metabolism



Conclusion

- ❑ Using machine learning, we identified a subset ($n = 12$) of serum metabolites (mainly BCAA and long-chain acylcarnitines) that were significantly associated with over-conditioning in cows.
- ❑ HBCS cows displayed higher levels of circulating acylcarnitine and during early lactation that may reflect mitochondrial overload and incomplete fatty acid oxidation

Conclusion

- ❑ Mapping the serum metabolites identified alterations in multiple pathways including Carnitine synthesis, BCAA degradation before calving.
- ❑ **Mitochondrial FA oxidation** pathways were the most functionally enriched pathways in HBCS compared with NBCS cows after calving.
- ❑ Hepatic gene expression of FA oxidation during early lactation may reflect **mitochondrial and peroxisomal** overload and incomplete fatty acid oxidation being associated with over-conditioning around calving.

Conclusion

Proper management

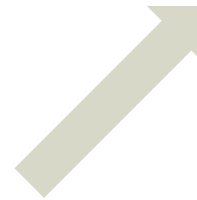
- aiming at preventing overconditioning during the late lactation.



leading to a successful transition of over-conditioned cows from late pregnancy to lactation.

Treatment of metabolic disorders

- targeting these metabolic processes might provide tools for protecting against the metabolic abnormalities associated with over-conditioning.



QUESTIONS?

*Thank
you*



Hedges' g is a measure of effect size. Effect size tells you how much one group differs from another — usually a difference between an experimental group and control group.

$$\text{Hedges' } g = \frac{\bar{x}_{Test} - \bar{x}_{Control}}{\text{StDev}_{pooled}} \times F$$

where F can be approximated by

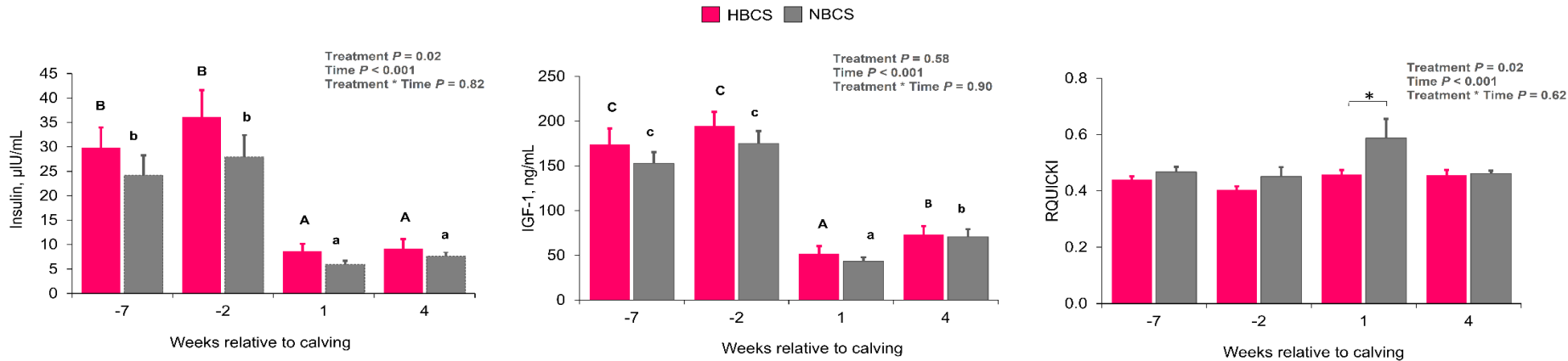
$$F \approx 1 - \frac{3}{4(n_{Test} + n_{Control}) - 9}$$

Here at EstimationStats.com, we use the exact form for F , which employs the [gamma function](#) (Γ):

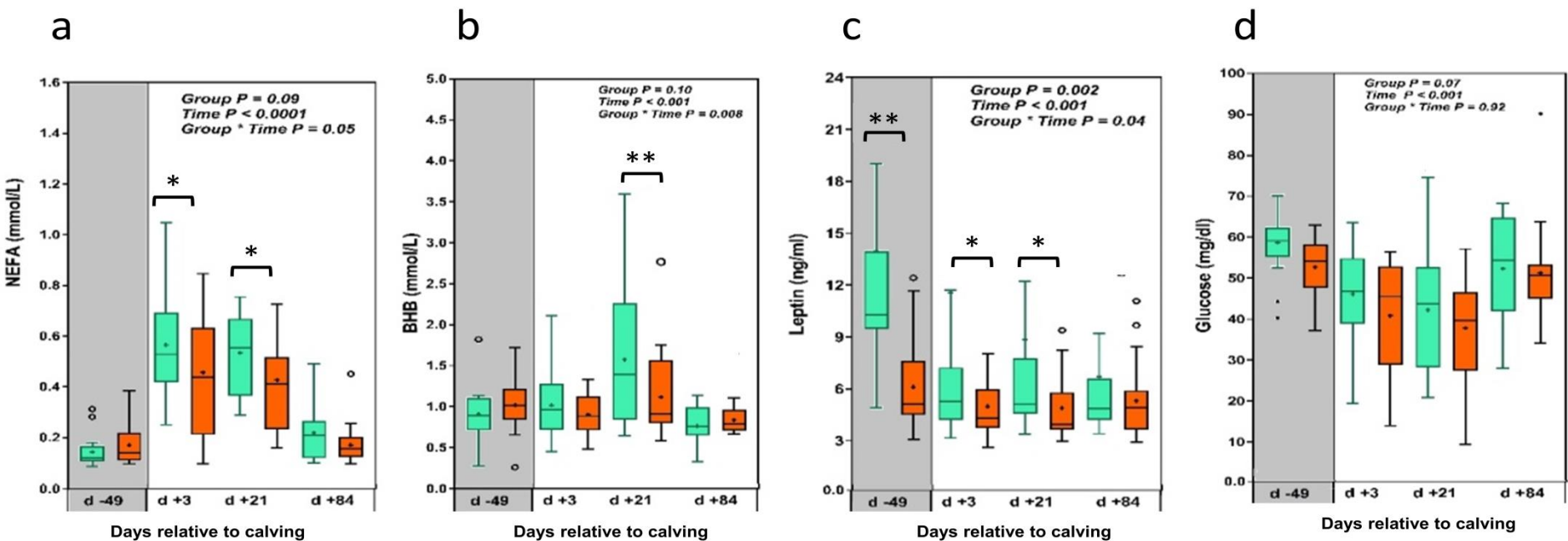
$$F = \frac{\Gamma(e/2)}{\sqrt{e/2} \times \Gamma((e-1)/2)}$$

and e refers to the degrees of freedom:

$$e = n_{Test} + n_{Control} - 2$$



Supplemental Figure 1. Longitudinal changes of serum insulin, IGF-1, and RQUICKI in normal-conditioned (NBCS; < 3.5 BCS and < 1.2 cm BFT at dry-off) and over-conditioned (HBCS; > 3.75 BCS and > 1.4 cm BFT at dry-off) cows during the experimental period (n = 18 per treatment). Different letters indicate differences ($P < 0.05$) between the time points in the HBCS (A, B), and in the NBCS (a, b) group, respectively. Symbols indicate differences ($* P < 0.05$) between HBCS and NBCS at a given time-point. Data are presented as means \pm SEM and are from Schuh et al., (2019).



Supplemental Figure 2. Longitudinal changes of non-esterified fatty acids (NEFA), β -hydroxybutyrate (BHB), leptin, and glucose in serum of normal-conditioned (NBCS; < 3.5 BCS and < 1.2 cm BFT at dry-off) and over-conditioned (HBCS; > 3.75 BCS and > 1.4 cm BFT at dry-off) cows during the experimental period ($n = 18$ per treatment). Different letters indicate differences ($P < 0.05$) between the time points. Symbols indicate a significant difference ($*P < 0.05$; $**P < 0.01$) between the groups at a given time. Data are presented as means \pm SEM and were published previously (Schuh et al., 2018).