Ambike, V., VandeHaar, M., Parker Gaddis, K., Peñagaricano, F., White, H., Weigel, K., Baldwin, R., Santos, J., Koltes, J., Tempelman, R. (2026) Genome-wide association analysis accounting for genotype by diet interaction for feed efficiency in Holstein cows, *Proceedings 13th World Congress on Genetics Applied to Livestock Production* 24228: https://doi.org/10.31274/wcgalp.24228 ©2026 *Ambike, et al.* CC BY-NC-ND

Genome-wide association analysis accounting for genotype by diet interaction for feed efficiency in Holstein cows

V. Ambike[1*], M.J. VandeHaar[1], K.L. Parker Gaddis[2], F. Peñagaricano[3], H.M. White[3], K.A. Weigel[3], R. Baldwin[4], J.E.P. Santos[5], J.E. Koltes[6], and R.J. Tempelman[1]

[1]Michigan State University, East Lansing, MI, USA; [2]Council on Dairy Cattle Breeding, Bowie, MD, USA; [3]University of Wisconsin-Madison, Madison, WI, USA; [4]Animal Genomics and Improvement Laboratory, ARS USDA, Beltsville, MD, USA; [5]University of Florida, Gainesville, FL, USA; [6]Iowa State University, Ames, IA, USA; [*]Corresponding author: ambikevr@msu.edu

Introduction

Feed efficiency (**FE**) characterizes a dairy cow's ability to convert feed into milk and thus impacts her economic and environmental sustainability. Key FE traits include dry matter intake (**DMI**) and residual feed intake (**RFI**), the latter being the difference between actual and predicted DMI (Connor, 2015). Dietary components like fat, starch, neutral detergent fiber (**NDF**), and crude protein (**CP**) may influence the genetic expression of milk production and DMI and may influence the expression of RFI, thereby making the assessment of genotype-by-diet (**G×D**) interactions potentially important for precision genetic management of livestock.

Random regression models (**RRM**) capture genetic merit as linear or higher order polynomial functions of continuous environmental gradients, such as temperature or diet. Genome-wide association (**GWA**) studies are used to identify potentially important genomic regions associated with complex traits by inferring statistical associations between phenotypes and individual single nucleotide polymorphisms (**SNP**) across the genome. In BLUP-GWA, estimated SNP effects are back solved from genomic prediction models, which consider all markers simultaneously, thereby improving detection of quantitative trait loci (QTL) (Aguilar *et al*., 2019). Some G×D interactions can be due to scaling, whereby animals rank similarly across different environments albeit with heterogeneous variability, whereas some interaction is due to reranking of animals across environments, thereby impacting selection programs. Extending the GWA framework to account for G×D potentially allows the detection of genomic regions that are robust versus sensitive to different dietary components for feed efficiency. These inferences could facilitate selection of robust animals, and discovery of putative candidate genes and gene networks for robustness. Thus, the objective of this study was to conduct a GWA study for DMI and RFI within a RRM framework for modeling GxD.

Materials & Methods

Data. Data was collected from 87 trials at five US research stations namely Michigan State University (MSU: East Lansing, MI), University of Florida (UF: Gainesville, FL), University of Wisconsin-Madison (UW: Madison, WI), Iowa State University (IA: Ames, IA), and the Animal Genomics and Improvement Laboratory (AGIL: Beltsville, MD). A total of 4,670 performance records on 2,951 Holstein cows were used for modeling GxD as a function of CP and NDF whereas 3,049 records on 2,191 cows were available for modeling GxD as a function of fat and starch. These records were based on 28-day intervals between 50 to 150 days in milk (**DIM**) collected from 2007 to 2024. All cows included in the analysis were genotyped for 64,038 autosomal SNP

markers that remained after quality control on the CDCB 69,200 SNP set. These markers were subsequently used to construct the genomic relationship matrix (**GRM**) between animals.

Analysis. For each of the four dietary components (fat, CP, starch, and NDF), RRM was fitted up to cubic orders for both genetic (**G**) and permanent environment (**PE**) effects for both DMI and RFI. Fixed effects included up to fourth order polynomials for diet components, DIM and temperature-humidity index, as well as classification factors of parity (primiparous vs multiparous), two-year intervals for each station, and various interactions. Contemporary groups (i.e., diets within trials) were also fitted as random effects. For RFI, a single step modeling approach (Tempelman *et al*., 2015) was used such that body weight (**BW**), change in BW, and fat and protein yields in milk were additionally included as energy sink covariates.
In all cases, the best fitting G and PE orders for RRM were determined based on likelihood-ratio tests and Akaike's information criteria (**AIC**). Genomic estimated breeding values (**GEBV**) for animal-specific intercepts and slopes were backsolved to get corresponding SNP intercept and slope estimates (Waters *et al*., 2022) with respective standard errors computed based on mixed model theory (Tolhurst, 2024). Multiple testing was corrected using the Bonferroni adjustment based on the effective number of SNP markers (Gao *et al*., 2008). Classical BLUP-GWA (i.e., no GxD) was also conducted for both DMI and RFI as a reference analysis. All analyses were conducted using ASReml-R (Butler *et al*., 2023). Putative candidate genes were identified within ± 500 kb using Ensembl (ARS-UCD1.2). Genomic inflation factors were computed for each analysis to assess the potential inflation or deflation of Type I error.

Results

Random regression models. Dietary components varied widely across experiments and diets based on following ranges for fat (2.7 to 7.9 % DM), CP (14.3 to 20.8 % DM), starch (20 to 33.6 % DM), and NDF (23.7 to 39.8 % DM). Estimated overall mean heritabilities for DMI and RFI based on standard genomic models (i.e., without RR) were 0.32 ± 0.03 and 0.26 ± 0.03, respectively. The best fitting G and PE polynomial orders for fitted RRMs and the corresponding ranges in heritability estimates obtained from these models are presented in Table 1. The RRM for DMI generally required higher order polynomials for optimal fits compared to RFI. GxD with respect to fat, starch and NDF was deemed to be important for DMI whereas GxD with respect to starch and NDF was deemed to be important for RFI.

Table 1. Optimal RRM specifications and estimated heritability ranges for DMI and RFI across dietary components.

Diet Component	DMI			RFI		
	Polynomial Order (G)	Polynomial Order (PE)	Heritability range	Polynomial Order (G)	Polynomial Order (PE)	Heritability range
Fat	Linear	Quadratic	0.12 – 0.37	Intercept	Quadratic	0.16 – 0.26
CP	Intercept	Quadratic	0.17 – 0.32	Intercept	Quadratic	0.17 – 0.36
Starch	Cubic	Intercept	0.19 – 0.62	Linear	Linear	0.16 – 0.26
NDF	Cubic	Intercept	0.23 – 0.44	Quadratic	Intercept	0.16 – 0.30

Genome-wide association analysis. Using a standard non-RR BLUP-GWA analysis, there was no evidence of an association of DMI or RFI with any SNP marker. However, after modeling GxD on starch with a better fitting RRM, a SNP intercept effect on chromosome 26 (marker

rs109238377) was determined (Bonferroni adjusted $P < 0.05$) for RFI. Within a ±1 Mb region surrounding rs109238377, 7 out of 44 neighboring markers exhibited mild evidence of association ($P < 10^{-3}$) and moderate to high linkage disequilibrium with this SNP ($r^2 \geq 0.3$). This pattern thereby indicated a coherent regional association, rather than an isolated single-marker association. This association shows that RRM has the potential to detect QTL that might not be identified otherwise. The nearest annotated gene was RAB11FIP2 (~108 kb from the SNP), along with several others in the window representing potential candidates for mediating the observed GxD effect of starch on RFI. The genomic inflation factors were 1.00 for the traditional GWA whereas they ranged between 0.97 to 1.00 for SNP intercepts and between 0.93 to 1.02 for SNP slopes and higher-order RRM-GWA terms in RRM models. Hence, test statistics were well behaved and tended to be more conservative in their control of Type I error, especially for RRM. As a potential indication of a difference in statistical power for detecting SNP effects across intercepts versus slopes, there were on average 22 more SNP intercept effects compared to SNP linear slope effects having P-values < 10^{-3} across both traits and all 4 dietary components.

Discussion

Our heritability estimates for DMI and RFI were consistent with previously reported values in dairy cattle (Berry and Crowley, 2013). Genotype by diet interactions were generally deemed to be important for DMI and RFI. The underlying genetic architecture, for both average performance (i.e. random intercepts) and robustness/ sensitivity (i.e. random polynomials), was determined to be highly polygenic for both traits. That is, across all 4 dietary components and both intercept and slope effects, only one SNP was determined to have a significant association. Modeling GxD may sharpen GWA inference for FE traits allowing SNP effects to vary across a nutritional continuum. Nevertheless, the size of our study was relatively small by GWA standards such that these associations should be reassessed in the future with a larger dataset.

Our results also suggested that there was a difference in the nature of GxD for different diet components. For example, GxD was generally important to model for starch and NDF, whereas GxD was not detected for CP for either trait. In other words, the genetic sensitivity of DMI to dietary composition appears to be more variable across a dietary starch gradient whereas genetic robustness is more characteristic of CP. This distinction may also explain why a statistically significant SNP was detected for RFI for the intercept on starch after accounting for the complex GxD modeled using the better fitting RRM. SNP detection power for RR-GWAS likely differs between estimating the average (i.e. intercept) and GxD (i.e. linear or higher random regression orders) on the covariate as the intercept captures the baseline genetic differences across all environments. Slope and higher order polynomials capture environment-specific genetic sensitivity such that the power of detection does depend upon the range of dietary components. Genomic inflation factors were generally < 1 indicating slightly deflated and conservative test-statistics.

The gene closest to the significant SNP, RAB11FIP2 is involved in vesicle trafficking and membrane recycling (Cullis *et al*., 2002) that influences how a cell handles nutrient and signaling receptors. Such mechanisms could plausibly impact metabolic responses to starch or energy substrates and therefore modulate RFI under varying starch levels.

Conclusions

This study demonstrates the importance of explicitly modeling GxD when evaluating FE traits and shows that these interactions can be captured using RRM. By developing and implementing a

GWA framework for RRM that allows for a more precise and efficient determination of standard errors for GWA tests based on Tolhurst (2024), we extend the traditional GWA framework, enabling the identification of SNP markers not only associated with intercept effects but also with sensitivity to dietary components. The corresponding test statistics are also well behaved with genomic inflation factors close to unity. This work can be used to identify genotypes and genomic regions that are either sensitive or robust to environmental perturbations.

Acknowledgements

All the collaborating institutions and FFAR and CDCB for research funding (Award # CA18-SS-0000000236 Improving Dairy Feed Efficiency, Sustainability and Profitability by Impacting Farmer's Breeding and Culling Decisions).

References

Aguilar, I., Legarra, A., Cardoso, F., Masuda, Y., Lourenco, D., et al. (2019) Genet. Sel. Evol. 51(1): 28. doi.org/10.1186/s12711-019-0469-3

Berry, D.P., Crowley, J.J. (2013) J. Anim. Sci. 91: 1594-1613. doi.org/10.2527/jas.2012-5862

Butler, D.G., Cullis, B.R., Gilmour, A.R., Gogel, B.J., Thompson, R. ASReml-R Reference Manual Version 4. VSN International Ltd, Hemel Hempstead, UK. Available at: https://asreml.kb.vsni.co.uk/wp-content/uploads/sites/3/ASReml-R-Reference-Manual-4.2.pdf, visited on 09-04-2026.

Connor, E.E. (2015) Animal 9: 395-408. doi.org/10.1017/S1751731114002997

Cullis, D.N., Philip, B., Baleja, J.D., Feig, L.A. (2002) J. Biol. Chem. 277: 49158-49166. doi.org/10.1074/jbc.M206316200

de las Heras-Saldana, S., Clark, S.A., Duijvesteijn, N., Gondro, C., van der Werf, J.H.J., et al. (2019) BMC Genomics 20: 939. doi.org/10.1186/s12864-019-6270-4

Gao, X., Starmer, J., Martin, E.R. (2008) Genet. Epidemiol. 32: 361-369. doi.org/10.1002/gepi.20310

Tempelman, R.J., Spurlock, D.M., Coffey, M., Veerkamp, R.F., Armentano, L.E., et al. (2015) J. Dairy Sci. 98: 2013-2026. doi.org/10.3168/jds.2014.8510

Tolhurst, D.J. (2024) PhD dissertation. University of Edinburgh. Available at: https://era.ed.ac.uk/server/api/core/bitstreams/810e76a9-b0c7-4f8a-8238-ec689d39bac3/content, visited on 09-04-2026.

Waters, D.L., Clark, S.A., Moghaddar, N., van der Werf, J.H. (2022) Genet. Sel. Evol. 54: 40. doi.org/10.1186/s12711-022-00734-6