The rumen microbiome is crucial for converting feed into absorbable nutrients used for milk synthesis, and the efficiency of this process directly impacts the profitability and sustainability of the dairy industry. Recent studies have found that the rumen microbial composition explains part of the variation in feed efficiency traits, including dry matter intake, milk energy, and residual feed intake. The main goal of this study was to reveal relationships between the host genome, rumen microbiome, and dairy cow feed efficiency using structural equation models. Our specific objectives were to (i) infer the mediation effects of the rumen microbiome on feed efficiency traits, (ii) estimate the direct and total heritability of feed efficiency traits, and (iii) calculate the direct and total breeding values of feed efficiency traits. Data consisted of dry matter intake, milk energy, and residual feed intake records, SNP genotype data, and 16S rRNA rumen microbial abundances from 448 mid-lactation Holstein cows from 2 research farms. We implemented structural equation models such that the host genome directly affects the phenotype ($G \rightarrow P$) and the rumen microbiome ($G \rightarrow M$), while the microbiome affects the phenotype ($M \rightarrow P$), partially mediating the effect of the host genome on the phenotype ($G \rightarrow M \rightarrow P$). We found that 7 to 30% of microbes within the rumen microbial community had structural coefficients different from zero. We classified these microbes into 3 groups that could have different uses in dairy farming. Microbes with heritability <0.10 but significant causal effects on feed efficiency are attractive for external interventions. On the other hand, 2 groups of microbes with heritability ≥0.10, significant causal effects, and genetic covariances and causal effects with the same or opposite sign to feed efficiency are attractive for selective breeding, improving or decreasing the trait heritability and response to selection, respectively. In general, the inclusion of the different microbes in genomic models tends to decrease the trait heritability rather than increase it, ranging from −15% to +5%, depending on the microbial group and phenotypic trait. Our findings provide more understanding to target rumen microbes that can be manipulated, either through selection or management interventions, to improve feed efficiency traits. 

Key words: Dry Matter Intake, Microbiome, Structural Equation Model, Total heritability

INTRODUCTION

In dairy cows, the rumen microbiome is crucial for digesting and transforming feedstuffs into volatile fatty acids and microbial amino acids needed for maintenance, growth, and lactation (Weimer, 1998). The effectiveness of the process to convert feed into milk is what we know as feed efficiency, which greatly impacts the profitability and sustainability of the dairy industry (VandeHaar et al., 2016). Feed efficiency is commonly measured using residual feed intake (RFI) as the difference between the actual dry matter intake (DMI) and predictions from regression on known energy sinks, including metabolic body weight, change in body weight, and milk energy (Berry and Crowley, 2013; VandeHaar et al., 2016).

The rumen microbiome has been associated with milk traits (Lima et al., 2015; Buitenhuis et al., 2019; Martinez
Boggio et al., 2023a), methane emissions (Difford et al., 2018; Saborío-Montero et al., 2020) and, more recently, with feed efficiency traits in dairy cows (Martinez Boggio et al., 2023b). In the latter study, we showed that the rumen microbiome explains up to 20% of the phenotypic variance of 3 feed efficiency traits, namely DMI, milk energy, and RFI, and the inclusion of the rumen microbiome composition into prediction models improved the predictive ability of those models (Martinez Boggio et al., 2023b). We also found that the direct heritability, defined as the proportion of the phenotypic variance explained by the host cow genes acting directly on the phenotype, was consistently 10% smaller than the total heritability, which included mediated effects, suggesting that the rumen microbiome mediates part of the host genetic effect on feed efficiency traits (Martinez Boggio et al., 2023b). Quantifying the mediation effect of the rumen microbiome on feed efficiency could improve the prediction of estimated breeding values (EBV) for direct host effect and total effect (including mediated effects) and contribute to optimizing herd management decisions.

The potential benefits of including omics traits (e.g., microbiome) in genetic evaluations and of splitting the genetic effect into direct and mediated effects were evaluated by Christensen et al. (2021) using a 2-joint model. Recursive models (RM) are structural equation models (SEM) that assume unidirectional causality between variables in a multivariate system (Gianola and Sorensen, 2004). Note that a causal effect of a variable \( y_1 \) on another variable \( y_2 \) exists if and only if a change in \( y_1 \) produces a change in the expected value of \( y_2 \) (Pearl, 2009). This means that 2 conditions must be met: changes in \( y_1 \) cause changes in \( y_2 \), and changes in \( y_2 \) do not cause changes in \( y_1 \). In the case of a causal relationship between microbiome and feed efficiency, an external intervention affecting microbial abundance and composition would produce a change in feed efficiency. Causality is important for understanding the underlying mechanisms that determine the expression of complex traits (Wu et al., 2010; Valente et al., 2015). Interestingly, given the equivalence between multi-trait models and recursive models, it is possible to unravel the mediation effect of the microbiome on the phenotype of interest and to estimate the direct and mediated contributions of the host genome to the variance components and EBV for direct and total genetic effects.

In this study, we aimed to infer (i) the mediation effects of rumen microbiome on 3 feed efficiency traits, namely DMI, milk energy, and RFI, (ii) the direct and total heritability of feed efficiency traits, and (iii) the direct and total breeding values of feed efficiency traits in mid-lactation Holstein cows.

### MATERIALS AND METHODS

Trials were conducted at the University of Florida (Gainesville, FL, USA) and the University of Guelph (Guelph, ON, Canada). Experimental protocols were approved by the Institutional Animal Care Committees of Florida (protocol n°202300000187) and Guelph (protocol n°4064).

Data comprised phenotypes, genotypes, and rumen microbiome records from 451 lactating Holstein cows. Cows were from 6 experiments, 237 cows enrolled in 5 experiments conducted at the University of Florida Dairy Research Unit (Alachua, FL, USA), and 214 cows enrolled in a single experiment at the Ontario Dairy Research Centre (Elora, ON, Canada), between January 2019 and July 2020. More details about the cows, housing, and feeding system are provided in Martinez Boggio et al. (2023b).

#### Feed efficiency traits

Three feed efficiency traits were evaluated: DMI, net energy secreted in milk (NESec), and RFI. All calculations were performed for 31 to 136 d in milk (DIM). Milk energy (Mcal/day) was calculated according to the NRC equations for nutrient requirements of dairy cattle (NRC, 2001) as follows:

\[
NESec = [(9.29 \times \text{kg Fat}) + (5.47 \times \text{kg Protein}) + (3.95 \times \text{kg Lactose})].
\]

We generated weekly means of body weight (BW) using daily records. Weekly changes in BW were calculated as the difference between 2 consecutive weeks and divided by the number of days between measurements to generate a daily change in BW (\( \Delta BW \)) in kg/day.

The DMI records were regressed on the 3 main energy sinks to calculate RFI values using a linear model as follows:

\[
DMI = mDIM + Parity + cohort + b_1\text{NESec} + b_2\text{mBW} + b_3\Delta BW + e,
\]

where \( mDIM \) represents the effect of midpoint days in milk for the period from 31 to 136 DIM (2 levels: \( \leq 67 \text{ DIM and } >67 \text{ DIM}), Parity \) represents the effect of parity with 2 levels (primiparous and multiparous), \( cohort \) represents the fixed effect of treatment nested within experiment and season (13 levels), \( \text{NESec} \) is the net energy for lactation secreted in milk with partial regression coefficient \( b_1 \), \( mBW \) is the metabolic BW with partial regression coefficient \( b_2 \), \( \Delta BW \) is the change in BW with partial regression coefficient \( b_3 \), and \( e \) is the random residual of the model, representing RFI. The coefficient...
of determination of the model was 0.81. All variables included in the model significantly \( (P < 0.05) \) contributed to variation in DMI.

**Rumen fluid sampling**

Rumen fluid sampling was performed on a single day in the middle of each experimental period when cows were between 66 and 105 DIM. The rumen fluid sample for each cow was obtained using an orogastric tube attached to a glass vial and equipped with a pump. Samples were snap-frozen in liquid nitrogen immediately after sample collection. The microbial DNA was extracted with subsequent amplification by PCR of the 16S rRNA gene V4 region. The sequencing process was done using an Illumina MiSeq platform (Illumina, San Diego, CA, USA) set for a 16S Metagenomics Workflow. More details about DNA extraction, amplification, and sequencing are provided in Martinez Boggio et al. (2023b). The raw sequence reads were deposited in the Sequence Read Archive of the National Center for Biotechnology Information under the Bio Project PRJNA962991.

Amplicon sequences of the 451 rumen samples were processed using the DADA2 pipeline version 1.8 (Callahan et al., 2016) in the R software (version 4.1.2; R Core Team, 2022). Briefly, denoising analysis was performed by demultiplexing sequencing reads and inspecting for errors. Sequences were trimmed and filtered, chimeras were removed, and an amplicon sequence variant (ASV) table with 9,707 ASV was created. Taxonomy assignment was performed using 16S rRNA SILVA version 138 database until the genus level (Quast et al., 2012).

From the 451 rumen samples, we discarded one sample during wet lab processing and 2 samples with very few \(<3,361\) sequences and few \(<120\) ASV. We retained 448 samples and 1,636 ASV after the removal of those ASV present in less than 5\% of samples. Considering the compositionality of the microbial abundances, we added 1 to all counts in the abundance table to avoid zeros and then applied a centered log-ratio transformation (CLR) with the function \texttt{clr} implemented in the R package compositions (Van Den Boogaart and Tolosana-Delgado, 2008).

**Genotyping**

All 448 cows had genotype data for 78,933 SNP across the genome. The SNP data were kindly provided by the Council on Dairy Cattle Breeding (Bowie, MD). We retained 78,016 SNP after quality control, including call rates of \( \geq 95\% \) for SNP and \( \geq 95\% \) for individuals and excluding SNP with minor allele frequency \(< 1\%\).

**Mediation analysis with structural equation modeling**

We aimed to estimate the mediation effect of the rumen microbiome on feed efficiency. We used a recursive model to assess the causal relationship of each ASV (M) with feed efficiency traits (P), as depicted in Figure 1. In our model, the host genome (G) directly affects the phenotype \( (G_p \rightarrow P) \) and the rumen microbiome \( (G_M \rightarrow M) \), while the microbiome also affects the phenotype \( (M \rightarrow P) \), acting as a mediator of indirect effects of the host genome on the phenotype \( (G \rightarrow M \rightarrow P) \). The structural coefficient \( (\lambda_{P-M}) \) represents the causal effect of M on P while keeping G constant. Environmental factors \( (E_p \text{ and } E_M) \) affect P and M, respectively, and we assume these effects are uncorrelated (zero covariance).

The structural equation models were implemented as a 2-trait model (Equation 1), including one feed efficiency trait and one ASV at a time, as follows:

\[
\Lambda y = Wb + a + e, \quad (1)
\]

where \( y \) is a vector of phenotypic records for P (DMI, NESec, or RFI) and M \((1 \text{ to } 1,636 \text{ ASV})\); \( \Lambda \) is the matrix of structural coefficients, defined as:

\[
\Lambda = \begin{bmatrix} 1 & \lambda_{P-M} \\ -\lambda_{P-M} & 1 \end{bmatrix}, \quad (2)
\]

where \( \lambda_{P-M} \) is the effect of M on P; \( b \) is a vector of fixed effects that include cohort (13 levels) and lactation (4 levels: 1, 2, 3, and 4+), which was not included in RFI models; \( W \) is a design matrix relating fixed effects to phenotypic records; \( a \) is a vector of additive genetic effects, and \( e \) is a vector of residuals. The distributional assumptions are multivariate Gaussian distributions, \( a \sim N (0, G_0 \otimes G) \) and \( e \sim N (0, R_0 \otimes I_n) \), where \( I_n \) is the identity matrix of order n-traits, and \( G \) is the additive genomic relationship matrix, computed as \( G = ZZ' / k \), where \( Z \) is a matrix of centered and standardized SNP genotypes and \( k \) represents the number of SNP (VanRaden, 2008). \( G_0 \) and \( R_0 \) are symmetric covariance matrices among the 2 traits (P and M) for genomic and residual effects, respectively, with elements equal to:

\[
G_0 = \begin{bmatrix} \sigma^2_{a,p} & \sigma_{a,pm} \\ \sigma_{a,pm} & \sigma^2_{a,m} \end{bmatrix}, \quad (3)
\]

\[
R_0 = \begin{bmatrix} \sigma^2_{e,p} & \sigma_{e,pm} \\ \sigma_{e,pm} & \sigma^2_{e,m} \end{bmatrix}, \quad (4)
\]
where \( \sigma^2_{a_P} \) and \( \sigma^2_{a_M} \) are the genetic variances of the direct effects on the phenotype and microbiome, respectively, and \( \sigma^2_{a_P M} \) is the covariance between additive direct effects, which represent the effects of genes that directly affect both traits (i.e., via linkage disequilibrium or pleiotropy); \( \sigma^2_{e_P} \) and \( \sigma^2_{e_M} \) are the residual variances of the phenotype and microbiome, respectively, and \( \sigma^2_{e_P M} \) is the residual covariance between direct effects. Note that the parameters of SEM are identifiable because \( R_0 \) is diagonal \( \sigma_{e_P M} = 0 \) (Wu et al., 2010). This recursive model can be reparametrized as a standard mixed model by multiplying all terms of Equation 1 by \( \Lambda^{-1} \), as follows:

\[
y = \Lambda^{-1} W b + \Lambda^{-1} a + \Lambda^{-1} e = \Lambda^{-1} W b + a^* + e^*, \tag{5}
\]

where \( a^* \) and \( e^* \) are the vectors of breeding values and residuals under the standard mixed model, and \( G_0^* = \Lambda^{-1} G_0 \) and \( R_0^* = \Lambda^{-1} R_0 \).

In our recursive model, the vector \( a \) has 2 effects: \( a_P \) as the direct additive genetic effects, and \( a_M \) as the microbiome effects. Furthermore, the indirect genetic effects mediated by the rumen microbiome are represented by \( \lambda_P \rightarrow M a_M \), such that the reparametrized \( a^*_P = a_p + \lambda_P \rightarrow M a_M \) is the vector of overall genetic effects exerted by the host genome on a phenotype through all the causal paths. The elements of \( G_0^* \) can be written as:

\[
\sigma^2_{a_P} = \sigma^2_{a_P} + 2\lambda_P \rightarrow M \sigma a_{PM} + \lambda_P \rightarrow M \sigma^2_{a_M}, \tag{8}
\]

\[
\sigma^2_{a_M} = \sigma^2_{a_M}, \tag{9}
\]

\[
\sigma^2_{a_{PM}} = \sigma a_{PM} + \lambda_P \rightarrow M \sigma^2_{a_M}, \tag{10}
\]

where \( \sigma^2_{a_P} \) and \( \sigma^2_{a_M} \) are the total genetic variances of the phenotype (DMI, NESec, or RFI) and microbiome (1 to 1,636 ASV), respectively; Equation 10 represents the covariance between phenotype and microbiome. Note that if \( \lambda_P \rightarrow M \sigma^2_{a_M} \) is of opposite sign, but of equal magnitude to \( \sigma_{a_{PM}} \), the total genetic covariance could be null, and given the Gaussian distribution assumed for the effects, that means that in this case, the effects are independent.
**Enrichment analysis**

We regrouped all ASV in their corresponding genus level and performed an enrichment analysis based on the sign (positive or negative) of their effect on the phenotypes. We computed a 2-by-2 contingency table for each microbial genus. The contingency table contained the number of those significant ASV with positive or negative effects and affiliated or not to a specific genus. We evaluated the enrichment using a Fisher’s exact test declaring significance at $P < 0.05$.

**Estimation of genetic parameters and breeding values**

We estimated the direct heritability ($h_{dp}^2$) and total heritability ($h_p^2$) of each feed efficiency trait, and the heritability of each ASV ($h_{MS}^2$), as follows:

$$h_{dp}^2 = \frac{\sigma_{a,p}^2}{\sigma_{a,p}^2 + \sigma_{e,p}^2 + 2\lambda_{P-M}^2 (\sigma_{a,PM}^2 + \sigma_{e,PM}^2) + \lambda_{P-M}^2 (\sigma_{a,M}^2 + \sigma_{e,M}^2)}$$

(11)

$$h_p^2 = \frac{\sigma_{a,p}^2}{\sigma_{a,p}^2 + \sigma_{e,p}^2 + 2\lambda_{P-M}^2 (\sigma_{a,PM}^2 + \sigma_{e,PM}^2) + \lambda_{P-M}^2 (\sigma_{a,M}^2 + \sigma_{e,M}^2)}$$

(12)

$$h_{MS}^2 = \frac{\sigma_{a,M}^2}{\sigma_{a,M}^2 + \sigma_{e,M}^2}.$$  \hspace{1cm} (13)

We defined a $\lambda_{P-M}$ as significantly different from zero when the absolute value of its estimate divided by its standard error was larger than 2. Note that (estimate/standard error) follows a t-distribution, and thus t-values larger than 2.2, given that the degrees of freedom are large, represent $\lambda_{P-M}$ that are significantly different from zero.

We estimated individual breeding values for DMI, NE-Sec, and RFI, represented as $a_p$, and for each ASV represented as $a_{M}$, using Equation 1. We also calculated the total EBV as $a_t = a_p + \lambda_{P-M} a_M$, where $\lambda_{P-M} a_M$ represents the mediated EBV. We calculated the Spearman’s correlation between $a_p$ and $a_t$ for each feed efficiency trait and ASV, and we summarized the results by doing an average per feed efficiency trait. The models were implemented using the BLUPF90 family of programs (Misztal et al., 2002), and the variance components were estimated using 200 initial rounds of EM-REML and continuing to reach convergence with AI-REML.

**RESULTS**

**Structural (recursive) effects**

A total of 7 to 30% of microbes of the rumen microbial community (1,636 ASV) had structural coefficients different from zero depending on the phenotype. Specifically, 492 ASV (30%) had an effect on DMI, 393 ASV (24%) had an effect on NE-Sec, and 108 ASV (7%) had an effect on RFI. Of these 983 ASV, 15 showed effects simultaneously on all 3 traits, and others showed effect on 2 traits, i.e., 158 affected DMI and NE-Sec, 46 affected DMI and RFI, and 17 affected NE-Sec and RFI (Figure 2A).

Structural coefficient estimates are expressed as the expected change in the number of phenotypic standard deviations per one-unit increase in CLR-transformed ASV units by dividing the recursive effect by the standard deviations of each phenotypic trait. The structural coefficients for DMI ranged from −0.42 ± 0.06 to 0.22 ± 0.04 (kg/d CLR-transformed ASV units), structural coefficients for NE-Sec ranged from −0.43 ± 0.05 to 0.27 ± 0.06 (Mcal/d CLR-transformed ASV units), and structural coefficients for RFI ranged from −0.17 ± 0.05 to 0.14 ± 0.03 (kg/d CLR-transformed ASV units).

**Enrichment analysis**

Structural coefficients with positive and negative effects were identified, including 209 positive and 283 negative for DMI, 193 positive and 200 negative for milk energy, and 59 positive and 49 negative for RFI. Similar microbial genera encompassed ASV with positive effects on feed efficiency traits: *Prevotella* (15% on RFI; and 28% on DMI and NE-Sec), an unknown genus of the family *F082* (3.5% on NE-Sec; 8% on DMI; 12% on RFI), *Prevotellaceae UCG-003* (4% on DMI), *Prevotella_7* (3.5% on NE-Sec), and *[Eubacterium] hallii* group (5% on RFI). On the other hand, ASV with negative effects on feed efficiency traits were encompassed by *Prevotella* (6% on all traits) but most of them were from different microbial genera: DMI showed an unknown genus of the family *Lachnospiraceae* (7%), *Lachnospiraceae NK3A20* group (5%), and *[Eubacterium] nodatum* group (2.5%); NE-Sec showed *Christensenellaceae R-7* group (11%), and *Rikenellaceae RC9 gut* group (4%); and RFI showed *Prevotella_7* (10%). However, as shown in Figure 2B, not all those genera were overrepresented in the rumen microbiome with positive or negative effects. Note that *Prevotellaceae* and *F082* families had mostly positive effects on feed efficiency, while *Lachnospiraceae* and *Christensenellaceae* had mostly negative effects.
The heritability ($h^2_M$) estimates for the 1,636 ASV ranged from 0 to 0.43 ± 0.12, with a mean of 0.07. We classified and grouped rumen microbial ASV based on their $h^2_M$, phenotypic effect ($\lambda_P$), and genetic covariance with the different phenotypic traits (Supplemental Table S1; https://data.mendeley.com/datasets/t4n8v9bdm5/2; Martinez Boggio, 2024). We generated 4 groups of microbes for each feed efficiency trait (Figure 3): Group 1) ASV with non-significant phenotypic effect; Group 2) ASV with low host genetic control ($h^2_M < 10\%$) and significant phenotypic effect; Group 3) ASV with moderate to high host genetic control ($h^2_M \geq 10\%$), significant phenotypic effect, and $\lambda_{P-M}$ and genetic covariance ($\sigma_{aM}$) with the same sign; and Group 4) ASV with moderate to high host genetic control ($h^2_M \geq 10\%$), significant phenotypic effect, and $\lambda_{P-M}$ and genetic covariance ($\sigma_{aM}$) with opposite signs. Note that microbes from group 2 could be useful for external interventions, e.g., nutritional or feeding management, while microbes from group 3 and group 4 could have potential use in selective breeding.

**Microbial heritability**

We estimated the direct ($h^2_{dP}$) and total heritability ($h^2_{tP}$) for each feed efficiency trait across those ASV with structural coefficients different from 0. The mean $h^2_{dP}$ and $h^2_{tP}$ of DMI were 0.41 ± 0.11 and 0.39 ± 0.11, estimates for NESec were 0.49 ± 0.11 and 0.46 ± 0.11, and estimates for RFI were 0.09 ± 0.08 and 0.07 ± 0.08, respectively. In Figure 4, we show that the total heritability of feed efficiency traits increased when $\lambda_{P-M}$ had the same sign as the genetic covariance ($\sigma_{aM}$), and the stronger the genetic covariance, the larger the increase in $h^2_{tP}$. Interestingly, the inclusion of rumen microbiome tends to decrease rather than increase $h^2_{tP}$, and that suggests, as shown in Figure 3, a large number of ASV counteracting the direct genetic effect on feed efficiency.

**Direct and total breeding values**

We estimated the direct ($a_P$) and total ($a_t$) EBV for each feed efficiency trait. The $a_P$ and $a_t$ represent the additive genetic effects contributed directly by host genes to the phenotype of interest and the rumen microbiome, respectively, while the $a_t$ represents the entire additive effects.
of the host genes on feed efficiency, combining both the direct and indirect (i.e., mediated by the microbiome) contributions to the final phenotype. The average Spearman correlations between the direct and total EBV were 0.98 for DMI (range from 0.63 to 1), 0.92 for NESec (range from −1 to 1), and 0.99 for RFI (range from 0.92 to 1). More ASV were found in group 4 for NESec than for DMI and RFI. Hence, the correlations obtained were also lower for NESec than for DMI and RFI. Figure 5 shows the regression of indirect EBV ($\lambda_{P\rightarrow M}$) on total EBV for each feed efficiency trait using some of the most relevant rumen microbes. For example, ASV826 was from group 3 for RFI with a positive relationship between $\lambda_{P\rightarrow M}$ and $a_P$, but from group 4 for DMI and NESec. However, ASV579 and ASV704 were from groups 3 and 4, showing positive and negative relationships between $\lambda_{P\rightarrow M}$ and $a_P$, respectively. Note that rumen microbes from group 3 always have positive relationship between indirect and total EBV, so they are potentially useful in improving the response to selection.

**DISCUSSION**

We investigated the use of structural equation models to reveal relationships between the host genome, rumen microbiome, and feed efficiency in lactating Holstein cows. The rumen microbiome is crucial in dairy cow feed efficiency because it transforms feedstuffs into absorbable nutrients such as volatile fatty acids and microbial amino acids, supplying the cow with nutrients needed for maintenance of the host and for milk synthesis. It has been demonstrated that efficient and inefficient cows have different rumen microbial composition (Delgado et al., 2019; Monteiro et al., 2022). Furthermore, the inclusion of microbiome data into genomic models improves phenotypic prediction of feed efficiency traits (Martinez Boggio et al., 2023b). It should be noted that causal inference is required to learn how feed efficiency traits are expected to respond to microbiome interventions and how microbiome data could be used to guide selection and management decisions.
Structural equation models have been applied to numerous traits, but recently, special attention has been given to their use in unraveling the relationships between the host genome, microbiome, and phenotypes of interest in livestock species (Saborío-Montero et al., 2020; Mora et al., 2022; Martinez Boggio et al., 2023a). In dairy cows, mediation effects of the rumen microbiome on methane emissions were reported by Saborío-Montero et al. (2020), but there are no reports yet for feed efficiency traits. To quantify the mediation effect of the microbiome on phenotypes ($G \rightarrow M \rightarrow P$), we estimated the genetic covariance between phenotypes and microbiome ($\sigma_{a,PM}$), the microbial heritability ($h_m^2$, $G_M \rightarrow P$), which ranged from 0 to 0.43 ± 0.12 and was similar to that reported in other studies using dairy cows (Difford et al., 2018; Li et al., 2019; Saborío-Montero et al., 2020), and the causal effects of microbiome on phenotypes ($\lambda_{P,M}$; $M \rightarrow P$). We identified that approximately 30% of the microbial community had significant causal effects on DMI and milk energy, but only 11% on RFI. Residual feed intake is defined as the residuals from the regression of DMI on its main energy sinks, which could explain the low number of ASV with causal effects. Note that only 17 ASV simultaneously affect RFI and milk energy (i.e., a predictor of DMI), while 158 ASV simultaneously affect DMI and milk energy.

Interestingly, we classified rumen microbes into 3 groups that could have different uses in dairy farming. Microbes from group 2, which affect feed efficiency traits and have low host genetic control ($h_m^2$), are attractive for external interventions such as nutritional management, such as changing forage/concentrate ratio, adding feed additives or direct-fed microbial. Interventions that could change the microbial abundances of group 2 microbes would improve dairy cow feed efficiency. Conversely, ASV from groups 3 and 4, which have significant phenotypic effects ($\lambda_{P,M}$) and moderate heritability ($h_m^2$ up to 0.43), are attractive for their use in animal breeding because they are transmitted and mediate genetic effects on DMI, milk energy, and RFI. Groups 3 and 4 could be included in selection indices based on their selection response or used for indirect selection targeting those microbes with a favorable genetic correlation with the phenotype of interest. Note that we as-

![Figure 4](image-url)
sumed no microbial interactions within or between groups in the rumen.

The total heritability \( h^2_{tP} \) estimates for DMI and milk energy are similar to those obtained with single trait models, including only the host genome (Martinez Boggio et al., 2023b). That was expected because the total heritability includes all direct and indirect paths to the phenotype of interest. On average, \( h^2_{tP} \) was 2 to 3% lower than the direct heritability \( h^2_{dP} \), suggesting no mediation effects of the rumen microbiome on feed efficiency traits, as was showed by Martinez Boggio et al. (2023a) for milk traits in dairy sheep. But as shown in Figure 4, \( h^2_{tP} \) changes as a function of the ASV included in the recursive model. Note that the use of SEM allowed us to dis-

*Figure 5. Regression of total breeding values (EBV) on mediated breeding values for dry matter intake (DMI), net energy secreted in milk (NESec) and residual feed intake (RFI). Green dots represent breeding values when using ASV from group 3 (moderate to high heritability, casual effect on the phenotype, and genetic covariance and causal effect of same signs), and red dots represent breeding values when using ASV from group 4 (moderate to high heritability, casual effect on the phenotype, and genetic covariance and causal effect of opposite signs).*
entangle the relationships between the host genome and microbes in the rumen microbial community. Thus, the $h_{iP}^2$ can be greater than $h_{dP}^2$ if we include rumen microbes from group 3, or $h_{iP}^2$ can be lower than $h_{dP}^2$ if we include rumen microbes from group 4. Figure 3 shows that most of the ASV with a moderate heritability ($h_{m}^2$) are from group 4, which reduces the total heritability of milk energy secretion and DMI, and they could be a potential problem for genetic selection. Note that in general, the inclusion of different microbes tends to reduce more than increase $h_{iP}^2$ (Figure 4), as reported by Varona and González-Recio (2023) using simulated data, and that could explain the overall differences obtained in this study between $h_{iP}^2$ and $h_{dP}^2$. We hypothesize that for feed efficiency traits, as was proposed for methane emissions in beef cattle (Martínez-Álvaro et al., 2022), the cow genome can regulate rumen microbial activities through mechanisms such as substrate limitation, microbial inhibition, and/or playing a role in coordinating actions among microbial communities.

Interestingly, we split breeding values into direct genetic effects ($a_P$) attributed to the genes that directly affect feed efficiency and mediated genetic effects ($\lambda_{PM}a_M$), which are caused by host genes that, by affecting the rumen microbiome, produce a phenotypic change in feed efficiency. If selection is based on $a_P$, the response to selection will occur through the direct genetic effects. However, as Varona and González-Recio (2023) noted, a correlated genetic response could be caused in the rumen microbial abundances due to additive genetic covariance. On the other hand, if selection is based on $a_M$, the response to selection will be affected by the microbes included as mediated traits. If they are rumen microbes from group 3 (such as ASV579), we will obtain an improvement in selection response, but if they are from group 4 (such as ASV714), we could reduce the response on feed efficiency. We should also consider those cases where the same ASV, e.g., ASV826, has a different impact on each feed efficiency trait (Figure 5). Note that ASV presented here only illustrates their potential use in animal breeding; hence, future studies in mid-lactation dairy cows should confirm their contribution to feed efficiency.

Overall, the identification of microbial groups that differentially contribute to the response of selection suggest that parsimonious SEM assuming a Gaussian prior distribution for the set of recursive effects on phenotypes and no correlation between the direct and mediated effects (Christensen et al., 2021) could not fit complex traits such as feed efficiency. However, it is still unclear how to reduce the large number of microbiome causal effects without losing information about their relationship with the host genome. The use of selection indices to improve traits by optimizing the microbiome composition could be a solution, as proposed by Weishaar et al. (2020) working with pigs. Furthermore, we suggest further work is needed to understand the causal network between the rumen microbiome and complex traits.

**CONCLUSIONS**

Structural equation models (SEM) offer an alternative to disentangle relationships between the host genome, rumen microbiome, and feed efficiency traits in lactating Holstein cows. We classified the rumen microbes into 3 groups, each of which could have different uses in dairy farming. For example, we found microbes that could be useful for external interventions because they have a causal effect on feed efficiency traits and low heritability. In addition, we found 2 more groups of microbes that could change the total heritability and response to selection. The total and direct heritability estimates were similar for DMI, NESec, and RFI. Therefore, one group of microbes with moderate host genetic control, significant phenotypic effects, and genetic covariance and phenotypic effect with the same sign improves total heritability and the response to selection, and the other group with genetic covariance and phenotypic effects with opposite signs decreases the heritability and the response to selection. In summary, SEM provide guidance to target microbes that can be manipulated using selective breeding or feeding management to improve dairy cattle feed efficiency.

**ACKNOWLEDGMENTS**

The authors acknowledge funding from USDA National Institute of Food Agriculture (grant n° 2019-67015-32114) and the Foundation for Food and Agriculture Research (FFAR; grant n° RC109491). Collection of rumen fluid samples and phenotypes in Canadian cows were supported by the Ontario Agri-Food Innovation Alliance (project n° 030277) and the Food from Thought Thematic Research Fund (project n° 499121), and their feed efficiency phenotypes and genotypes were provided by contributions from the Resilient Dairy Genome Project (Genome Canada, Genome Alberta, and organizations listed at http://www.resilientdairy.ca/funders-and-partners/) as well as the Dairy Research Cluster 3 (Lactanet and Agriculture and Agri-Food Canada). The authors would like to thank all farm staff at the University of Florida Dairy Unit (Alachua, FL, USA) and the Ontario Dairy Research Centre (Elora, ON, Canada), graduate students and interns for all their help with animal handling, sample collection and data acquisition, and Drs. Delaine de Melo and Erika Bonsaglia for their help.
with sample processing. The authors have not stated any conflicts of interest.

REFERENCES


ORCIDs

Guillermo Martínez Boggio https://orcid.org/0000-0003-3597-9426

Hugo F. Monteiro https://orcid.org/0000-0003-3305-4497

Fabio S. Lima https://orcid.org/0000-0001-8377-6469

Caio C. Figueiredo https://orcid.org/0000-0003-0890-0743

Rafael S. Bisinotto https://orcid.org/0000-0001-6144-2919

José E. P. Santos https://orcid.org/0000-0003-3403-1465

Bruna Mion https://orcid.org/0000-0002-7061-7415

Flavio S. Schenkel https://orcid.org/0000-0001-8700-0633

Eduardo S. Ribeiro https://orcid.org/0000-0002-4201-7383

Kent A. Weigel https://orcid.org/0000-0002-2391-6260

Guilherme J. M. Rosa https://orcid.org/0000-0001-9172-6461

Francisco Peñagaricano https://orcid.org/0000-0001-6661-3991

Journal of Dairy Science Vol. TBC No. TBC, TBC