



OPEN

## Multi-breed host rumen epithelium transcriptome and microbiome associations and their relationship with beef cattle feed efficiency

P. A. S. Fonseca<sup>1</sup>, S. Lam<sup>1</sup>, Y. Chen<sup>2</sup>, S. M. Waters<sup>3</sup>, L. L. Guan<sup>2</sup> & A. Cánovas<sup>1</sup>✉

Understanding host-microbial interactions in the rumen and its influence on desirable production traits may lead to potential microbiota manipulation or genetic selection for improved cattle feed efficiency. This study investigated the host transcriptome and its correlation with the rumen archaea and bacteria differential abundance of two pure beef cattle breeds (Angus and Charolais) and one composite beef hybrid (Kinsella) divergent for residual feed intake (RFI; low-RFI vs. high-RFI). Using RNA-Sequencing of rumen tissue and 16S rRNA gene amplicon sequencing, differentially expressed genes ( $FDR \leq 0.05$ ,  $|\log_2(\text{Fold-change})| > 2$ ) and differentially abundant ( $p\text{-value} < 0.05$ ) archaea and bacteria amplicon sequence variants (ASV) were determined. Significant correlations between gene expression and ASVs ( $p\text{-value} < 0.05$ ) were determined using Spearman correlation. Interesting associations with muscle contraction and the modulation of the immune system were observed for the genes correlated with bacterial ASVs. Potential functional candidate genes for feed efficiency status were identified for Angus (*CCL17*, *CCR3*, and *CXCL10*), Charolais (*KCNK9*, *GGT1* and *IL6*), and Kinsella breed (*ESR2*). The results obtained here provide more insights regarding the applicability of target host and rumen microbial traits for the selection and breeding of more feed efficient beef cattle.

The projected future global demands for meat and dairy production have pressured agricultural sectors to provide more efficient and sustainable food production systems, leading to opportunities and challenges for improving cattle production systems<sup>1,2</sup> such as improving cattle feed efficiency. Feed efficiency is a complex trait influenced by multiple biological mechanisms mainly attributed to metabolizable energy use and energy partitioning<sup>3</sup>. This suggests a systematic approach to studying feed efficiency may help better understand its underlying biological mechanisms.

Approaches using high-throughput -OMICs technologies, including RNA-Sequencing (RNA-Seq), have allowed for improved understanding of host genome and transcriptome features<sup>4,5</sup>. Consequently, bringing insight to associations between the host transcriptome and desirable production traits to help better understand the underlying genetic architecture and biological complexity of cattle feed efficiency<sup>6–11</sup>.

Due to the known relationship between the metabolic capacity for ruminants to expend energy and their symbiotic rumen microbial community<sup>12,13</sup>, systematic approaches have studied the rumen microbiome in association with feed efficiency. This highlights the need to better understand host genetics with other major energy sinks, like the rumen, with feed efficiency. Prior studies in beef cattle have emerged, providing a better understanding of how rumen microbial features and host genetics interact<sup>14–18</sup>.

The host genetic influence on rumen microbial features and its association with feed efficiency in beef cattle has been recently reported, where 19 single nucleotide polymorphism (SNPs) in the host genome were associated with 14 rumen microbial taxa. Among the identified SNPs, five were co-localized with known quantitative trait loci (QTL) for cattle feed efficiency<sup>14</sup>. In addition, Li et al.<sup>14</sup> revealed distinguishable differences in the rumen microbial features across beef cattle breeds and residual feed intake (RFI) groups. The RFI is defined the difference between the actual and the predicted feed intake required to obtain rate of gain<sup>19</sup>. Consequently, animals with negative RFI values are consuming less than the predicted consumption and consequently, are more feed

<sup>1</sup>Centre for Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, Guelph, ON N1G 2W1, Canada. <sup>2</sup>Livestock Gentec, Department of Agriculture, Food & Nutritional Science, University of Alberta, Edmonton, AB T6H 2P5, Canada. <sup>3</sup>Teagasc, Animal and Bioscience Research Department, Animal and Grassland Research and Innovation Centre, Grange, Dunsany C15 PW93, Co. Meath, Ireland. ✉email: acanovas@uoguelph.ca

efficient. On the other hand, animals with positive RFI consume more than the predicted consumption and are less feed efficient. Li et al.<sup>14</sup> was the first study to characterize bovine genotype with heritable microbial features with feed efficiency in cattle. Some studies have also estimated the heritability of rumen microbiome features, revealing low to moderate heritabilities<sup>14,20–22</sup> suggesting its ability to be genetically inherited. With evidence of the rumen microbiome being partially influenced by host genetics<sup>14,22</sup> and the potential for rumen microbial traits to be heritable, it is possible that the rumen microbiome can be genetically inherited and a specific rumen microbial profile may be optimal for improved feed efficiency<sup>22</sup>. These findings highlight the importance to study host-microbiome genetic interactions and their influence on feed efficiency in beef and dairy cattle and highlight the potential to either modify the rumen microbiome or select for it to improve efficiency. Understanding host-microbial interactions in the rumen and its influence on desirable production traits may lead to potential microbiota manipulation or genetic selection for improved cattle feed efficiency.

The abundance of studies aiming to discover the arcane interactions underlying feed efficiency suggests that knowledge is still limited on the impacts of variations in the rumen microbiota on host biology and function. Further approaches to understanding host-microbial interactions in the rumen and its influence on desirable production traits may lead to potential microbiota manipulation or genetic selection for improved cattle feed efficiency<sup>23</sup>.

The objectives of this study were to: (1) Determine significantly differentially expressed (DE) genes across divergent feed efficiency groups in two pure beef cattle breeds and one composite hybrid; (2) Determine differentially abundant rumen archaeal and bacterial amplicon sequence variant (ASV) abundance between divergent feed efficiency groups in two pure beef cattle breeds and one composite hybrid; (3) Identify significant correlations between the archaea or bacteria differentially abundant ASV (rumen microbial metagenome) and gene expression (host functional transcriptome) between divergent feed efficient cattle to determine 'host-microbiome genetic interactions' that may be influencing feed efficiency; and 4) Perform functional annotation of the genes that were significantly associated with the differentially abundant ASVs.

## Results and discussion

### RNA-sequencing and Amplicon-sequencing alignment statistics

On average, the total number of reads mapped for all samples in low and high-RFI animals were 303,428,164 and 318,958,134, for Angus (AN); 317,776,686 and 299,963,316 for Charolais (CH); and 328,016,010 and 332,007,994 for Kinsella (KN). The majority of reads were successfully uniquely mapped to the bovine reference genome (ARS-UCD2.1 release 97) which was on average, 86.68% for low-RFI and 91.02% for high-RFI for AN, 91.22% for low-RFI and 90.22% for high-RFI for CH, and 90.64% for low-RFI and 90.65% for high-RFI for KN (Supplementary Table 1). The average number of paired-end reads per sample was 39,586,465 ± 5,152,459 and the percent of uniquely mapped reads to the bovine reference genome (ARS-UCD2.1 release 97) was 90.07 ± 5.12%. As expected, an increase in the percentage of uniquely mapped reads was observed in this study using the new bovine reference genome (ARS-UCD2.1 release 97) compared to previous studies<sup>24,25</sup>, due to the improved gene annotation and sequence quality of the new reference.

The alignment statistics of the Amplicon-Seq analysis for each rumen content sample is shown in Supplementary Table 2. The average number of sequenced reads, total reads mapped, and percent uniquely mapped per sample for AN was 59,784.38 ± 8763.46, 48,535.63 ± 7168.28, and 81.20%, respectively. For CH, the average number of sequenced reads, total reads mapped, and percent uniquely mapped was 63,585.00 ± 5929.53, 51,635.44 ± 4767.78, and 81.22%, respectively. Finally, for KN the average number of sequenced reads, total reads mapped, and percent uniquely mapped was 59,831.63 ± 8976.15, 48,413.31 ± 7207.97, and 80.93%, respectively. Similarly, Li et al.<sup>15</sup> studied the rumen microbial metagenome and identified 54.63 ± 1.42 million sequence reads after quality control, and reported 78.47 ± 0.26% of reads mapped back to assembly contigs.

### Differentially expressed genes between high-RFI and low-RFI groups

The statistical and biological significance thresholds (FDR < 0.05, and  $|\log_2(\text{FC})| > 2$ ) resulted in the identification of DEG exclusively for the Kinsella breed. For Angus and Charolais, only the biological significance threshold was reached. The DEG identified for Kinsella breed and the top 20 genes with a  $|\log_2(\text{FC})| > 2$  for Charolais and Angus breeds are shown on Table 1. In total, 578 (445 downregulated and 133 upregulated), 426 (274 downregulated and 152 upregulated) and 641 (358 downregulated and 283 upregulated) genes showed an  $|\log_2(\text{FC})| > 2$  for Angus, Charolais and Kinsella, respectively (Supplementary Table 3).

#### Downregulated genes in low-RFI group

The 22 DEG (FDR < 0.05,  $\log_2(\text{FC}) > 2$ ) identified between high and low-RFI animals for the Kinsella breed were all downregulated in the low-RFI group (Table 1). The top 3 enriched GO terms identified for the DEG in Kinsella breed were *muscle system process* (adjusted p-value =  $4.28 \times 10^{-37}$ ), *muscle contraction* (adjusted p-value =  $1.87 \times 10^{-32}$ ), and *myofibril assembly* (adjusted p-value =  $5.31 \times 10^{-26}$ ). This close relationship between the DEG downregulated in the low-RFI group and the regulation of muscle tissue development and contraction is highlighted by the presence of genes coding for myosin (*MYH1*, *MYLPF*, *MYL2*, *MYH2*, *MYL1*, *MYH7*, and *MYBPC2*), actin (*ACTN3* and *ACTA1*) and troponin (*TNNI2*, *TNNC2*, *TNNT3*, and *TNNC1*) related proteins. In addition, genes coding for proteins directly related with the regulation of biological processes associated with muscular activity such as *CKM*, *ENO3*, *CASQ1* and *MB* were also identified as differentially expressed between high and low-RFI groups for Kinsella breed (downregulated in the low-RFI group). The enrichment analysis of GO terms for downregulated genes in the low-RFI group indicated that similar biological processes were involved with these genes in the three breeds. The functional grouping performed through the Jaccard correlation

Ensembl ID	Gene symbol	P-value	FDR	log2(FC)	Breed
ENSBTAG00000046725	<i>TNNC2</i>	$5.61 \times 10^{-7}$	$5.62 \times 10^{-3}$	-5.93	Kinsella
ENSBTAG00000018204	<i>MYH1</i>	$6.20 \times 10^{-7}$	$5.62 \times 10^{-3}$	-5.85	Kinsella
ENSBTAG00000013921	<i>CKM</i>	$7.30 \times 10^{-7}$	$5.62 \times 10^{-3}$	-4.05	Kinsella
ENSBTAG00000021218	<i>MYLPF</i>	$9.53 \times 10^{-7}$	$5.62 \times 10^{-3}$	-5.65	Kinsella
ENSBTAG00000018369	<i>MYL2</i>	$1.72 \times 10^{-6}$	$5.62 \times 10^{-3}$	-5.77	Kinsella
ENSBTAG00000052709	-	$1.91 \times 10^{-6}$	$5.62 \times 10^{-3}$	-5.70	Kinsella
ENSBTAG00000007090	<i>MYH2</i>	$2.33 \times 10^{-6}$	$5.62 \times 10^{-3}$	-5.41	Kinsella
ENSBTAG00000010880	<i>TNNI2</i>	$2.81 \times 10^{-6}$	$5.62 \times 10^{-3}$	-5.41	Kinsella
ENSBTAG00000009707	<i>MYL1</i>	$3.00 \times 10^{-6}$	$5.62 \times 10^{-3}$	-5.80	Kinsella
ENSBTAG00000022158	<i>TNNT3</i>	$3.06 \times 10^{-6}$	$5.62 \times 10^{-3}$	-4.92	Kinsella
ENSBTAG00000045757	<i>TNNC1</i>	$3.08 \times 10^{-6}$	$5.62 \times 10^{-3}$	-5.24	Kinsella
ENSBTAG00000014547	<i>PGAM2</i>	$3.15 \times 10^{-6}$	$5.62 \times 10^{-3}$	-5.15	Kinsella
ENSBTAG00000009703	<i>MYH7</i>	$3.25 \times 10^{-6}$	$5.62 \times 10^{-3}$	-5.48	Kinsella
ENSBTAG00000022244	<i>ACTN3</i>	$3.26 \times 10^{-6}$	$5.62 \times 10^{-3}$	-5.07	Kinsella
ENSBTAG00000048585	-	$3.28 \times 10^{-6}$	$5.62 \times 10^{-3}$	-5.31	Kinsella
ENSBTAG00000006541	<i>ATP2A1</i>	$3.30 \times 10^{-6}$	$5.62 \times 10^{-3}$	-3.96	Kinsella
ENSBTAG00000046332	<i>ACTA1</i>	$3.60 \times 10^{-6}$	$5.78 \times 10^{-3}$	-5.05	Kinsella
ENSBTAG00000031573	<i>NMRK2</i>	$8.23 \times 10^{-6}$	$1.25 \times 10^{-2}$	-5.21	Kinsella
ENSBTAG00000020080	<i>MYBPC2</i>	$9.62 \times 10^{-6}$	$1.38 \times 10^{-2}$	-5.04	Kinsella
ENSBTAG00000005534	<i>ENO3</i>	$1.15 \times 10^{-5}$	$1.57 \times 10^{-2}$	-3.70	Kinsella
ENSBTAG00000005333	<i>MB</i>	$3.11 \times 10^{-5}$	$4.04 \times 10^{-2}$	-3.84	Kinsella
ENSBTAG00000020223	<i>CASQ1</i>	$3.30 \times 10^{-5}$	$4.09 \times 10^{-2}$	-3.99	Kinsella
ENSBTAG00000018369	<i>MYL2</i>	$1.58 \times 10^{-3}$	1.00	-4.22	Angus
ENSBTAG00000018204	<i>MYH1</i>	$2.84 \times 10^{-3}$	1.00	-4.05	Angus
ENSBTAG00000046332	<i>ACTA1</i>	$2.74 \times 10^{-3}$	1.00	-4.03	Angus
ENSBTAG00000021218	<i>MYLPF</i>	$2.73 \times 10^{-3}$	1.00	-4.00	Angus
ENSBTAG00000045757	<i>TNNC1</i>	$1.37 \times 10^{-3}$	1.00	-3.98	Angus
ENSBTAG00000009707	<i>MYL1</i>	$1.65 \times 10^{-3}$	1.00	-3.97	Angus
ENSBTAG00000031573	<i>NMRK2</i>	$2.62 \times 10^{-3}$	1.00	-3.97	Angus
ENSBTAG00000007090	<i>MYH2</i>	$1.42 \times 10^{-3}$	1.00	-3.96	Angus
ENSBTAG00000052488	<i>SPINK9</i>	$5.28 \times 10^{-4}$	1.00	-3.96	Angus
ENSBTAG00000052709	-	$1.28 \times 10^{-3}$	1.00	-3.95	Angus
ENSBTAG00000011392	<i>MYBPC1</i>	$4.06 \times 10^{-3}$	1.00	-3.95	Angus
ENSBTAG00000010880	<i>TNNI2</i>	$1.78 \times 10^{-3}$	1.00	-3.94	Angus
ENSBTAG00000009703	<i>MYH7</i>	$1.55 \times 10^{-3}$	1.00	-3.94	Angus
ENSBTAG00000011734	<i>ANKRD1</i>	$3.45 \times 10^{-3}$	1.00	-3.92	Angus
ENSBTAG00000046725	<i>TNNC2</i>	$2.99 \times 10^{-3}$	1.00	-3.92	Angus
ENSBTAG00000022158	<i>TNNT3</i>	$1.20 \times 10^{-3}$	1.00	-3.80	Angus
ENSBTAG00000022244	<i>ACTN3</i>	$1.39 \times 10^{-3}$	1.00	-3.80	Angus
ENSBTAG00000005333	<i>MB</i>	$1.57 \times 10^{-3}$	1.00	-3.74	Angus
ENSBTAG00000008394	<i>MYL3</i>	$1.32 \times 10^{-2}$	1.00	-3.72	Angus
ENSBTAG00000007782	<i>MYOT</i>	$6.05 \times 10^{-3}$	1.00	-3.70	Angus
ENSBTAG00000037768	<i>MMP3</i>	$1.93 \times 10^{-4}$	$4.82 \times 10^{-1}$	-4.46	Charolais
ENSBTAG00000032642	-	$1.09 \times 10^{-3}$	1.00	-4.18	Charolais
ENSBTAG00000046633	-	$1.13 \times 10^{-3}$	1.00	-3.87	Charolais
ENSBTAG00000021565	-	$1.34 \times 10^{-2}$	1.00	-3.23	Charolais
ENSBTAG00000005353	<i>DES</i>	$4.90 \times 10^{-4}$	$6.68 \times 10^{-1}$	-2.81	Charolais
ENSBTAG00000039406	-	$4.56 \times 10^{-4}$	$6.61 \times 10^{-1}$	2.69	Charolais
ENSBTAG00000015441	<i>ACTG2</i>	$4.49 \times 10^{-4}$	$6.61 \times 10^{-1}$	-2.67	Charolais
ENSBTAG00000051421	<i>PCP4</i>	$7.05 \times 10^{-3}$	1.00	-2.65	Charolais
ENSBTAG00000033803	<i>FABP7</i>	$4.54 \times 10^{-3}$	1.00	-2.64	Charolais
ENSBTAG00000027279	-	$1.23 \times 10^{-2}$	1.00	-2.61	Charolais
ENSBTAG00000006546	<i>GSTA2</i>	$3.57 \times 10^{-3}$	1.00	-2.57	Charolais
ENSBTAG00000011207	<i>CNN1</i>	$3.37 \times 10^{-4}$	$6.40 \times 10^{-1}$	-2.49	Charolais
ENSBTAG00000046124	<i>KRT36</i>	$4.19 \times 10^{-5}$	$3.74 \times 10^{-1}$	-2.18	Charolais
ENSBTAG00000047121	-	$7.73 \times 10^{-3}$	1.00	-2.13	Charolais

Continued

Ensembl ID	Gene symbol	P-value	FDR	log <sub>2</sub> (FC)	Breed
ENSBTAG00000015988	MYH11	5.49 × 10 <sup>-5</sup>	3.74 × 10 <sup>-1</sup>	- 2.12	Charolais
ENSBTAG00000011424	TPM2	7.77 × 10 <sup>-5</sup>	3.96 × 10 <sup>-1</sup>	- 2.11	Charolais
ENSBTAG000000044123	HAND2	3.79 × 10 <sup>-2</sup>	1.00	- 2.09	Charolais
ENSBTAG00000017305	-	1.01 × 10 <sup>-1</sup>	1.00	- 2.09	Charolais
ENSBTAG000000049878	-	1.18 × 10 <sup>-1</sup>	1.00	- 2.03	Charolais
ENSBTAG00000011473	MYL9	6.20 × 10 <sup>-6</sup>	1.69 × 10 <sup>-1</sup>	- 2.03	Charolais

**Table 1.** Differentially expressed genes or top 20 genes based on absolute log<sub>2</sub>(fold-change) values between the rumen tissue samples of high- and low-RFI groups for Angus, Charolais and Kinsella breeds.

coefficient identified that among the top 30 GO terms associated with the downregulated genes in the three breeds a strong association with muscular development was present (Supplementary Fig. 1).

Similarly, the top 3 enriched GO terms identified for the downregulated genes in the Angus low-RFI group were also *muscle system process* (adjusted p-value = 1.02 × 10<sup>-36</sup>), *muscle contraction* (adjusted p-value = 5.47 × 10<sup>-35</sup>), and *myofibril assembly* (adjusted p-value = 2.09 × 10<sup>-28</sup>). It is important to highlight that for Angus, there were no genes which passed the statistical (FDR < 0.05) and biological (|log<sub>2</sub>(FC)| > 2) significance thresholds simultaneously. However, the analysis of the genes which passed the biological significance threshold suggest a strong association of these genes with regulation of the muscular tissue development. For example, for the Angus breed all the top 20 genes with highest |log<sub>2</sub>(FC)| were downregulated (Table 1). Among these genes, eight were from the gene families associated with the myosin superfamily of motor proteins (*MYL2*, *MYH1*, *MYLPF*, *MYL1*, *MYH2*, *MYBPC1*, *MYH7*, and *MYL3*), four were associated with troponin gene family (*TNNC1*, *TNNC2*, *TNNT3*, and *TNNI2*), and two were associated with actinin filaments (*ACTA1* and *ACTN3*). Additionally, interesting genes such as *MYOT* and *MB* were also identified among these genes.

There were no genes differentially expressed between high and low-RFI groups for the Charolais breed based on both statistical and biological significance criteria, as well as for Angus breed. Similar to the results obtained for Kinsella and Angus, genes responsible to encode myosin (*MYH11* and *MYL9*) and actin (*ACTG2*) related proteins were identified among the top 20 genes with highest |log<sub>2</sub>(FC)|.

The myosin and troponin related proteins are key regulators of the biochemical efficiency of the striate muscle<sup>26</sup>. The expression pattern or polymorphisms in genes related with myosin proteins were associated with feed efficiency status in diverse species, such as chickens, rainbow trout and cattle<sup>14,27-29</sup>. Actin, myosin and troponin are crucial actors in the processes of muscle contraction<sup>30,31</sup> and might influence the rumen contractions and consequently the digesta passage rate through the rumen. The passage rate of digesta through the rumen was previously associated with the microbiota content<sup>32</sup> and levels of produced methane (CH<sub>4</sub>)<sup>32-34</sup>. Interestingly, the increase of feeding, which raises the digesta passage rate, is associated with the reduction of digestibility and increasing of efficiency of the microbiota community to synthesize gasses and volatile fatty acids from fermented carbon<sup>35,36</sup>. Additionally, associations between CH<sub>4</sub> levels and cellulose digestibility, neutral detergent fiber digestibility, energy digestibility and cell wall digestion were identified<sup>37-39</sup>. A detailed review about the relationship between the digesta passage rate, CH<sub>4</sub> emissions and feed efficiency was published by Lovendahl et al.<sup>40</sup>.

Additionally, an interesting association between some of downregulated genes in the low-RFI group and the GO terms *feeding behavior* (adjusted p-value = 0.09) and *regulation of appetite* (adjusted p-value = 0.12) was observed (Supplementary Fig. 1). The downregulated genes *TCF15*, *TACR1*, *HAND2*, *CARTPT*, and *NPW* were associated with *feeding behavior*. On the other hand, the *HTR4*, *SLC22A3*, and *CARTPT* genes were associated with *regulation of appetite*. Among the downregulated genes associated with feeding behavior and regulation of appetite, the most distinguish functional roles were identified for *NPW* and *CARTPT* genes. The *NPW* codifies a neuropeptide which, in rats, increases the serum levels of prolactin, corticosterone, and stimulate food and water intake<sup>41,42</sup>. Additionally, experimental evidence through the intracerebroventricular administration of two residuals of *NPW* (*NPW*<sub>23</sub> and *NPW*<sub>30</sub>) in rats resulted in an increased body temperature, suppression of feeding and body weight gain<sup>43</sup>. Regarding *CARTPT*, Alpha melanocyte-stimulating hormone (α-MSH) derives from *CARTPT* product (cocainx10- and amphetamine-regulated transcript) and proopiomelanocortin (POMC) leads to an enhanced state of food intake and appetite inhibition<sup>44</sup>. Additionally, in humans, polymorphisms on *CARTPT* are associated with obesity and an alteration of the leptin effect on thermogenesis and energy expenditure<sup>45-47</sup>.

These results suggest a lower expression of genes with potential functional role in the regulation of rumen contraction and feeding behaviour. The frequency of rumen contractions is associated with the passage of the rumen digesta through the rumen which might influence the microbiota diversity and the concentration of vitality fatty acids (REF). Regarding feed behavior, significant correlation was observed between feed duration and frequency, meal size and eating rate with RFI in beef cattle<sup>48-51</sup>.

#### Upregulated genes in low-RFI group

In total, 133, 152, and 283 genes were upregulated (based on |log<sub>2</sub>(FC)| > 2) for Angus, Charolais and Kinsella, respectively. Enriched GO terms for upregulated genes were identified exclusively for Kinsella breed. A total of 177 GO terms were enriched for the upregulated genes in the low-RFI group in the Kinsella breed (GO-BP:105, GO-MF: 22, and GO-CC: 50). The top three enriched GO terms were *modulation of chemical synaptic transmission*

(adjusted p-value =  $4.49 \times 10^{-06}$ ), *regulation of trans-synaptic signaling* (adjusted p-value =  $4.49 \times 10^{-06}$ ), and *neuron cellular homeostasis* (adjusted p-value =  $2.43 \times 10^{-04}$ ). The functional grouping of the top 30 GO terms annotated for the upregulated genes in the low-RFI group for Kinsella breed reinforce the association of these genes with regulation of neuronal activity with the largest functional group classified as *amine neurotransmitter GABAergic glutamatergic* (Supplementary Fig. 1). The effects of GABA in the human body results in an increase in the plasma levels of growth hormone, acting as an anabolic agent<sup>52</sup>. Interestingly, the inclusion of GABA producing bacteria as a feed additive demonstrated a potential impact to improve growth performance in ruminants through the reduction of amines and increasing anti-oxidative activity<sup>53</sup>. In addition, the utilization of rumen-protected GABA has beneficial effects over food intake and milk production in dairy cows<sup>54,55</sup>. Consequently, suggesting a potential role of these genes over food intake and energy homeostasis which might directly impact the feed efficiency status.

The upregulated genes in the low-RFI group in Angus and Charolais breeds were not associated with any enriched GO terms. The largest functional groups obtained through the Jaccard correlation coefficient were *regulation gonad and morphogenesis assembly* and *meiotic acid catabolic process* for Angus and Charolais, respectively (Supplementary Fig. 1). Despite the absence of enriched terms, the GO terms annotated for these groups of genes suggest an association with regulation cellular proliferation, cell cycle control and cellular homeostasis.

### Differential archaea and bacteria ASVs between high-RFI and low-RFI groups

The abundance of archaea and bacteria ASVs for Angus, Charolais and Kinsella obtained in the current study are shown on Supplementary Table 4. Regarding the archaea species, all breed groups had a significant difference between high and low-RFI animals (Table 2). The differential abundance analysis for bacterial ASVs at the genus level between high-RFI and low-RFI animals resulted in two, five, and five genera with significant differential abundance for Angus, Charolais and Kinsella, respectively (Table 3).

Specie	Mean ( $\pm$ SD) HRFI	Mean ( $\pm$ SD) LRFI	t	P-value	Breed
<i>Methanobrevibacter ruminantium clade</i>	1.153 ( $\pm$ 0.453)	0.975 ( $\pm$ 0.452)	0.789	0.443	Angus
<i>Methanobrevibacter boviskoreani clade</i>	0.172 ( $\pm$ 0.437)	0.338 ( $\pm$ 0.533)	-0.684	0.506	Angus
<i>Methanobrevibacter gottschalkii clade</i>	0.231 ( $\pm$ 0.304)	0.279 ( $\pm$ 0.218)	-0.365	0.721	Angus
<i>Methanosphaera sp. ISO3-F5</i>	0.108 ( $\pm$ 0.042)	0.104 ( $\pm$ 0.057)	0.149	0.884	Angus
<i>Methanosphaera sp. ISO3-F5</i>	0.089 ( $\pm$ 0.012)	0.075 ( $\pm$ 0.02)	1.699	0.117	Charolais
<i>Methanobrevibacter ruminantium clade</i>	1.164 ( $\pm$ 0.341)	1.296 ( $\pm$ 0.089)	-1.062	0.319	Charolais
<i>Methanobrevibacter boviskoreani clade</i>	0.149 ( $\pm$ 0.399)	0 ( $\pm$ 0)	1.056	0.326	Charolais
<i>Methanobrevibacter gottschalkii clade</i>	0.26 ( $\pm$ 0.165)	0.262 ( $\pm$ 0.091)	-0.042	0.967	Charolais
Group12 sp. ISO4-H5	0.067 ( $\pm$ 0.074)	0.042 ( $\pm$ 0.042)	0.828	0.425	Kinsella
<i>Methanosphaera sp. ISO3-F5</i>	0.162 ( $\pm$ 0.043)	0.147 ( $\pm$ 0.046)	0.708	0.491	Kinsella
<i>Methanobrevibacter gottschalkii clade</i>	0.267 ( $\pm$ 0.089)	0.315 ( $\pm$ 0.202)	-0.606	0.558	Kinsella
<i>Methanobrevibacter ruminantium clade</i>	1.242 ( $\pm$ 0.112)	1.213 ( $\pm$ 0.208)	0.358	0.727	Kinsella

**Table 2.** Differentially abundant archaea genus in the rumen digesta of high- and low-RFI groups for Angus, Charolais and Kinsella breeds. *HRFI* high residual feed intake, *LRFI* low residual feed intake, *SD* standard deviation, *t* t-test statistic.

Genus	Mean ( $\pm$ SD) HRFI	Mean ( $\pm$ SD) LRFI	t	P-value	Breed
<i>Bacteroidales RF16 uncultured rumen bacterium</i>	0.026 ( $\pm$ 0.017)	0.079 ( $\pm$ 0.055)	-2.618	0.029	Angus
<i>Ruminococcus 1</i>	0.045 ( $\pm$ 0.013)	0.025 ( $\pm$ 0.016)	2.742	0.017	Angus
<i>Clostridiales vadin BB60 uncultured rumen bacterium</i>	0.006 ( $\pm$ 0.011)	0.032 ( $\pm$ 0.02)	-3.270	0.008	Charolais
<i>Clostridiales Family XIII uncultured</i>	0.008 ( $\pm$ 0.015)	0.031 ( $\pm$ 0.02)	-2.637	0.021	Charolais
<i>Prevotellaceae UCG-001</i>	0.104 ( $\pm$ 0.025)	0.071 ( $\pm$ 0.029)	2.490	0.026	Charolais
<i>Oribacterium</i>	0.092 ( $\pm$ 0.035)	0.06 ( $\pm$ 0.018)	2.300	0.043	Charolais
<i>Erysipelotrichaceae UCG-004</i>	0.07 ( $\pm$ 0.019)	0.106 ( $\pm$ 0.04)	-2.305	0.044	Charolais
<i>Papillibacter</i>	0.072 ( $\pm$ 0.035)	0.123 ( $\pm$ 0.032)	-3.013	0.009	Kinsella
<i>Ruminococcus 1</i>	0.372 ( $\pm$ 0.093)	0.257 ( $\pm$ 0.038)	3.225	0.010	Kinsella
<i>Prevotellaceae UCG-001</i>	0.132 ( $\pm$ 0.043)	0.084 ( $\pm$ 0.022)	2.848	0.017	Kinsella
<i>Ruminococcus gnavreuii group</i>	0.121 ( $\pm$ 0.042)	0.078 ( $\pm$ 0.019)	2.583	0.028	Kinsella
<i>Christensenellaceae R-7 group</i>	0.114 ( $\pm$ 0.016)	0.205 ( $\pm$ 0.106)	-2.406	0.045	Kinsella

**Table 3.** Differentially abundant bacteria genus in the rumen digesta of high- low-RFI groups for Angus, Charolais and Kinsella breeds. *HRFI* high residual feed intake, *LRFI* low residual feed intake, *SD* standard deviation, *t* t-test statistic.



The *Bacteroidales RF16 uncultured rumen bacterium* was differentially abundant between high and low-RFI animals in the Angus breed, with higher mean values in the low-RFI group (Table 3). Previously, the *Bacteroidales RF16* group was identified as differentially abundant in the rumen of Yak animals subject to different feed types, showing higher abundance in animals fed with forage when compared to animals fed with concentrate<sup>56</sup>. In the same study, the *Bacteroidales RF16* group was negatively correlated with important rumen metabolites (isobutyrate, isovalerate, L-Phenylalanine, L-Leucine, Hypoxanthine, L-Glutamate, Adrenic acid and Stearic acid). Oppositely, in sheep, the *Bacteroidales RF16* group showed higher abundance in the rumen of animals fed with concentrate feed supplement when compared with animals fed without supplement<sup>57</sup>. In pig models, the *Bacteroidales RF16* group showed an interesting decreasing pattern of abundance in the gut of animals subjected to heat stress<sup>58</sup>. The phyla Bacteroidete, which includes the genus *Bacteroidales RF16* group, is associated with reduction of endotoxins in the intestine, consequently, acting in the alleviation of inflammation and insulin resistance<sup>59–61</sup>. Therefore, these evidences reinforce a potential role of *Bacteroidales RF16* group in the regulation of feed efficiency in Angus breed.

The genus *Prevotellaceae UCG-001* was identified as differentially abundant between high and low-RFI animals in Charolais and Kinsella breeds (Table 3). In both breeds, high-RFI animals showed higher abundance of *Prevotellaceae UCG-001* when compared with low-RFI animals. The analysis of different portions of Yak rumen epithelium identified *Prevotellaceae UCG-001* as enriched in the rumen ventral epithelium<sup>62</sup>. In Holstein, the feeding of essential oils to neonatal animals resulted in an increased abundance of *Prevotellaceae*<sup>63</sup>. Interestingly, in pigs, high feed efficient animals showed higher abundance of *Prevotellaceae* when compared with low feed efficient animals<sup>64</sup>, while the opposite pattern was observed in the current study. Despite these differences, the *Prevotellaceae UCG-001* seems to play an important role in the mediation of restoration of leptin related pathways in lipid metabolism disorders as a response to insulin-supplemented diets in mice<sup>65</sup>. Consequently, suggesting a potential role of this genus in the control of feed efficiency in cattle through the control of feeding behavior by leptin action and lipid metabolism.

Two genera from the family *Ruminococcus* were identified as differentially abundant between high and low-RFI animals for Angus and Kinsella breeds. The genus *Ruminococcus 1* was observed to be more abundant in high-RFI animals in both breeds, while the *Ruminococcus gauvreauii* group was more abundant in the high-RFI animals for the Kinsella breed (Table 3). Bacteria from the *Ruminococcus* genus are characterized as cellulolytic and acting on the production of acetate, formate and hydrogen<sup>66</sup>. In pigs, the relative abundance of *Ruminococcus* genus in the gut microbiome was positively associated with feed efficiency<sup>67</sup>. In chickens, the relative abundance of *Ruminococcus* were higher in high-RFI animals in duodenum and ileum<sup>68</sup>. Additionally, in crossbreed feed steers the levels of *Ruminococcus* were significantly different between animals with contrasting average daily body weight gain and average daily dry matter intake<sup>69</sup>.

In Charolais, two genera from the *Clostridiales* family were identified as differentially abundant between high and low-RFI animals, showing higher abundance in low-RFI animals (*Clostridiales vadin BB60 uncultured rumen bacterium* and *Clostridiales Family XIII uncultured*). In pigs, members of *Clostridiales* family were identified as more abundant in the fecal microbiota of low-RFI animals<sup>70,71</sup>. Additionally, an integrative analysis of expressed genes and bacterial communities in the jejunal and cecal digesta of chickens suggested that *Clostridiales* modulates the expression levels of cytokines, the tight-junction protein *OCN* and nutrient transporters for glucose and short-chain fatty acid uptake<sup>72</sup>. The other two differentially abundant genus identified for Charolais breed were *Oribacterium* and *Erysipelotrichaceae UCG-004*. The *Oribacterium* genus was identified as more abundant in the high-RFI group while *Erysipelotrichaceae UCG-004* was more abundant in the low-RFI group (Table 3). In Holstein cows, a high-starch total mixed ration treatment to induce milk fat depression resulted in an increased abundance of *Oribacterium* in the ruminal fluid<sup>73</sup>. Additionally, the same study identified a correlation between the abundance of *Oribacterium* and the concentration of amines and amino acids, where a negative correlation between *Oribacterium* abundance and alanine levels stands out. The identification of metabolites differentially abundant between divergent feed efficient beef steers related with amino acids metabolism suggest a relevant role of the rumen amino acid concentrations in the regulation of feed efficiency<sup>74</sup>. The genus abundance levels of *Erysipelotrichaceae UCG-004* in the rumen fluid of Holstein cows was identified as a key predictor of mastitis in a random forest model<sup>75</sup>. Additionally, in pigs fed with soluble fiber during pregnancy showed a remarkable decrease in the abundance of *Erysipelotrichaceae UCG-004* in the progeny of these animals, suggesting an inter-generational effect of the diet over the microbial community<sup>76</sup>.

For Kinsella, other two differentially abundant genus identified between high and low-RFI animals for Kinsella breed were *Prevotellaceae UCG-001* (higher abundance in high-RFI) and *Christensenellaceae R-7* (higher abundance in low-RFI). Previously, in Angus cattle, the relative abundance of *Prevotellaceae* in the ruminal fluid was positively correlated with rumen pH and feed-to-gain ratio<sup>77</sup>. Interestingly, in Holstein, the abundance of *Prevotellaceae UCG-001* in the rumen was positively correlated with total milk solid<sup>78</sup>. In addition, also in Holstein, the inclusion of commence (a blend of active *S. cerevisiae*, *E. lactis*, *B. subtilis*, *E. faecium*, and *L. casei*, and their fermentation products) and RX3 in the diet reduced the relative abundance of *Prevotellaceae UCG-001* in the rumen<sup>79</sup>. In humans, the *Christensenellaceae* genus showed to be a widespread genus with a heritable pattern and associated with health<sup>80</sup>. In the current study, the *Christensenellaceae R-7* group showed a higher mean in the low-RFI group. Interestingly, in pigs, the relative abundance of *Christensenellaceae* had the most significant association with feed efficiency with higher mean in the low-RFI animals as observed here<sup>81</sup>. Similarly, in lambs, the *Christensenellaceae* family showed higher abundance in efficient animals when compared with inefficient animals<sup>82</sup>. In Nellore breed, the relative abundance of *Christensenellaceae R-7* group was significantly different between positive and negative RFI groups<sup>83</sup>. In addition, the relative abundance of *Christensenellaceae* were positively correlated with acetate and negatively correlated with propionate concentrations in the rumen in a multi-breed study composed by the same breeds evaluated in the current study<sup>14</sup>. The acetate:propionate ratio is highly correlated with ruminal pH and methane production in cattle<sup>84</sup>. Considering the importance of the

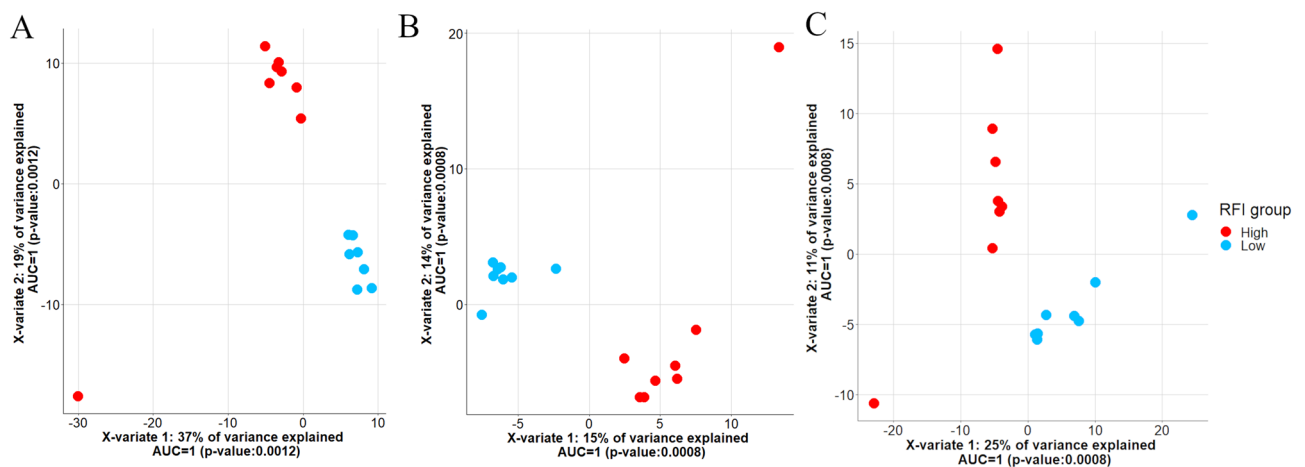
ruminal pH over the feed efficiency in dairy cattle<sup>85</sup>, the *Christensenellaceae R-7 group* is an important ASVs to be monitored in beef cattle.

### Correlation between candidate genes and ASVs for feed efficiency in beef cattle

The sPLS-DA analysis using the RNA-Seq counts for the genes with  $|\log_2(\text{FC})| > 2$  was able to perfectly discriminate high- and low-RFI animals for all the three breeds (Fig. 1). The variant selection based on the loading vectors for the first component of the discriminant analysis resulted in the selection of 462, 321, and 502 genes for Angus, Charolais and Kinsella, respectively (Supplementary Table 5). The top 10 enriched GO terms for the selected genes for each breed are shown on Table 4. The Supplementary Table 6 shows all GO terms annotated for the genes selected in the sparse partial least squares discriminant analysis (sPLS-DA). The enriched GO terms for these subsets of genes were closely related with those identified as enriched for all genes with  $|\log_2(\text{FC})| > 2$  and discussed previously. In summary, several GO terms associated with muscle development and regulation of muscle contraction were identified for all the breeds. As previously discussed, the regulation of muscle contraction in the rumen was associated with feed efficiency status in diverse species, such as chickens, rainbow trout and cattle<sup>14,27–29</sup>.

Several significant correlations were observed between the arcsine square root-transformed relative abundance of the differential abundant bacterial genus and the read counts for the genes with  $|\log_2(\text{FC})| > 2$  for the three breeds (Supplementary Table 7). The top 10 enriched terms for each pair of correlated bacterial genus and genes are shown in Supplementary Fig. 2. In addition of GO enriched terms associated with muscle contraction as identified before, an interesting enrichment pattern for immune system related terms was identified for the genes correlated with several bacterial genus. In Angus, the genes correlated with *Ruminococcus 1* were associated with enriched GO terms such as neutrophil, granulocyte and leukocyte migration. For Charolais, enriched GO terms such as hepatic immune response (*Erysipelotrichaceae UCG-004* and *Clostridiales Family XIII uncultured*), regulation of inflammatory response (*Oribacterium*), immature B cell differentiation (*Prevotellaceae UCG-001*) and myeloid leukocyte mediated immunity (*Clostridiales vadin BB60 uncultured rumen bacterium*) were identified. For Kinsella, immunity-related enriched GO terms were identified for the genes correlated with *Christensenellaceae R7 group* (neutrophil-mediated killing of gram-positive bacterium and neutrophil-mediated killing of bacterium). The important role of the immune system in the control of animal growth and efficiency have been highlighted in several species with a crucial role of the microbiome in the shaping of the immune system<sup>86–89</sup>. Therefore, the latter results might pinpoint potential candidate genes that contribute to the regulation of immune system components associated with the control of ruminal bacterial communities differentially abundant between divergent feed efficient animals.

Interestingly, there were no genes correlated with bacterial genus shared among the three breeds (all of the genes correlated with all of the bacterial genus were different among breeds). However, within breeds, common genes were identified as significantly correlated with two or more bacterial genus (Fig. 2). In addition, the comparative clustering of metabolic pathways among the genes correlated with ASVs within each breed identified relevant biological functions shared across these genes (Fig. 3). It is interesting to highlight that to build the clustering, only metabolic pathways shared between the two differentially abundant genus identified in Angus breed were included. On the other hand, for Charolais and Kinsella, only pathways shared with at least four of the five differentially abundant genus identified for each breed were included in the clustering. For Angus breed, the gene *ENSBTAG0000054941* (SCAN domain-containing protein 1-like) was simultaneously correlated with *Ruminococcus 1* and *Bacteroidales RF16 uncultured rumen bacterium*. However, the most interesting relationship between these two genera was identified when the common metabolic pathways among the correlated genes were analyzed (Fig. 3). Despite the low number of overlapping correlated genes, the two genera are chemokine and cytokine related pathways due to the correlations observed between *Ruminococcus 1* and *CCL17* ( $\rho = 0.62$ ,



**Figure 1.** High- and low-RFI sample clustering based on gene expression levels in the rumen tissue for Angus (A), Charolais (B) and Kinsella (C) breeds based on the first two principal components obtained in the sPLS-DA.

Breed	Description	GO class	P-value	FDR	Gene
Angus	Contractile fiber	CC	$3.18 \times 10^{-53}$	$9.82 \times 10^{-51}$	MYH15, FBXO32, SCN1A, CASQ2, MYH6, KY, LMOD3, MYOZ3, TNNT3, ANKRD1, LMOD2, SYNPO2L, TRIM54, MYOZ1, MYPN, MYBPC1, RYR1, MYL2, TNNC1, MYH2, STYXL2, CMYA5, MYL3, CSR3, MYOT, TNNI2, LRRC39, XIRP2, NRAP, TNNT1, IGFN1, ACTA1, MYLPF, ABRA, MYOM2, CAV3, TNNC2, LDB3, MYL1, ACTN3, MYOM3, MYH1, NEB, MYLK2, CASQ1, ACTN2, CACNA1S, MYOM1, KLHL41, TCAP, SMTNL1, TRIM63, CAPN3, TNNI1, MYH8, FLNC, TMOD4, MYOD1, MYOZ2
Angus	Myofibril	CC	$5.05 \times 10^{-53}$	$9.82 \times 10^{-51}$	MYH15, FBXO32, SCN1A, CASQ2, MYH6, KY, LMOD3, MYOZ3, TNNT3, ANKRD1, LMOD2, SYNPO2L, TRIM54, MYOZ1, MYPN, MYBPC1, RYR1, MYL2, TNNC1, MYH2, STYXL2, CMYA5, MYL3, CSR3, MYOT, TNNI2, LRRC39, XIRP2, NRAP, TNNT1, IGFN1, ACTA1, ABRA, MYOM2, CAV3, TNNC2, LDB3, MYL1, ACTN3, MYOM3, MYH1, NEB, MYLK2, CASQ1, ACTN2, CACNA1S, MYOM1, KLHL41, TCAP, SMTNL1, TRIM63, CAPN3, TNNI1, MYH8, FLNC, TMOD4, MYOD1, MYOZ2
Angus	Sarcomere	CC	$3.81 \times 10^{-51}$	$4.94 \times 10^{-49}$	FBXO32, SCN1A, CASQ2, MYH6, KY, LMOD3, MYOZ3, TNNT3, ANKRD1, LMOD2, SYNPO2L, TRIM54, MYOZ1, MYPN, RYR1, MYL2, TNNC1, MYH2, STYXL2, CMYA5, MYL3, CSR3, MYOT, TNNI2, LRRC39, XIRP2, NRAP, TNNT1, IGFN1, ACTA1, ABRA, MYOM2, CAV3, TNNC2, LDB3, MYL1, ACTN3, MYOM3, MYH1, NEB, MYLK2, CASQ1, ACTN2, CACNA1S, MYOM1, KLHL41, TCAP, SMTNL1, TRIM63, CAPN3, TNNI1, MYH8, FLNC, TMOD4, MYOZ2
Angus	Muscle system process	BP	$7.53 \times 10^{-37}$	$2.22 \times 10^{-33}$	STAC2, FBXO32, SLC8A3, SCN1A, SGCA, CLCN1, CASQ2, MYH6, TRIM72, LMOD3, CHRM3, CKMT2, MYBPH, TNNT3, LMOD2, SCN4B, MYOZ1, MYBPC1, RYR1, MYL2, TNNC1, MYH2, MYL3, CSR3, MYOT, TNNI2, TNNT1, MYBPC2, ACTA1, MYLPF, MYOM2, CAV3, TNNC2, ATP2A1, MYL1, ACTN3, MYOM3, MYH1, MYLK2, CASQ1, SLN, PGAM2, ACTN2, CACNA1S, MYOM1, KLHL41, JSRP1, TCAP, GHSR, TRIM63, SCN4A, TNNI1, MYH8, TMOD4, CHRNA1, MYOD1, MLIP, MYOZ2, TACR2
Angus	Muscle contraction	BP	$3.20 \times 10^{-35}$	$4.72 \times 10^{-32}$	STAC2, SLC8A3, SCN1A, SGCA, CLCN1, CASQ2, MYH6, LMOD3, CHRM3, CKMT2, MYBPH, TNNT3, LMOD2, SCN4B, MYBPC1, RYR1, MYL2, TNNC1, MYH2, MYL3, CSR3, MYOT, TNNI2, TNNT1, MYBPC2, ACTA1, MYLPF, MYOM2, CAV3, TNNC2, ATP2A1, MYL1, ACTN3, MYOM3, MYH1, MYLK2, CASQ1, PGAM2, ACTN2, CACNA1S, MYOM1, KLHL41, JSRP1, TCAP, GHSR, TRIM63, SCN4A, TNNI1, MYH8, TMOD4, CHRNA1, TACR2
Angus	Myofibril assembly	BP	$6.47 \times 10^{-30}$	$6.36 \times 10^{-27}$	SIX4, MYH6, LMOD3, TNNT3, ANKRD1, LMOD2, SYNPO2L, MYOZ1, MYPN, MYL2, CSR3, NRAP, TNNT1, ACTA1, MYOM2, CAV3, LDB3, NEB, CASQ1, ACTN2, KLHL41, TCAP, CAPN3, FLNC, TMOD4, MYOZ2
Angus	Striated muscle cell development	BP	$1.04 \times 10^{-29}$	$7.69 \times 10^{-27}$	SIX4, MYH6, LMOD3, TNNT3, ANKRD1, LMOD2, SYNPO2L, MYOZ1, MYPN, MYL2, CSR3, NRAP, TNNT1, ACTA1, MYOM2, CAV3, LDB3, NEB, CASQ1, ACTN2, KLHL41, TCAP, CAPN3, FLNC, TMOD4, MYOZ2
Angus	I band	CC	$1.27 \times 10^{-26}$	$1.23 \times 10^{-24}$	FBXO32, SCN1A, CASQ2, MYH6, KY, MYOZ3, ANKRD1, SYNPO2L, TRIM54, MYOZ1, MYPN, RYR1, MYL3, CSR3, MYOT, XIRP2, NRAP, IGFN1, CAV3, LDB3, ACTN3, NEB, CASQ1, ACTN2, CACNA1S, TCAP, SMTNL1, TRIM63, CAPN3, FLNC, MYOZ2
Angus	Muscle cell development	BP	$2.77 \times 10^{-25}$	$1.63 \times 10^{-22}$	NFATC2, SIX4, MYH6, SYPL2, LMOD3, TNNT3, ANKRD1, LMOD2, SYNPO2L, MYF6, ALPK2, MYOZ1, MYPN, RYR1, MYL2, CSR3, NRAP, TNNT1, ACTA1, MYOM2, CAV3, LDB3, ACTN3, NEB, CASQ1, ACTN2, KLHL41, TCAP, CAPN3, FLNC, TMOD4, MYOD1, MYOZ2
Angus	Sarcomere organization	BP	$4.48 \times 10^{-25}$	$2.20 \times 10^{-22}$	SIX4, MYH6, TNNT3, ANKRD1, LMOD2, SYNPO2L, MYOZ1, MYPN, CSR3, TNNT1, MYOM2, CAV3, LDB3, CASQ1, ACTN2, KLHL41, TCAP, CAPN3, FLNC, MYOZ2
Charolais	Voltage-gated potassium channel activity	MF	$5.29 \times 10^{-05}$	0.022	KCNK9, KCNH3, HCN3, LRRC55, KCNA2, KCNB1, KCNJ13
Charolais	Voltage-gated cation channel activity	MF	$1.56 \times 10^{-04}$	0.022	KCNK9, KCNH3, HCN3, LRRC55, KCNA2, CACNG1, KCNB1, KCNJ13
Charolais	Voltage-gated ion channel activity	MF	$3.10 \times 10^{-04}$	0.022	KCNK9, KCNH3, HCN3, LRRC55, SCN4A, KCNA2, CACNG1, KCNB1, KCNJ13
Charolais	Voltage-gated channel activity	MF	$3.10 \times 10^{-04}$	0.022	KCNK9, KCNH3, HCN3, LRRC55, SCN4A, KCNA2, CACNG1, KCNB1, KCNJ13
Charolais	Potassium channel activity	MF	$3.16 \times 10^{-04}$	0.022	KCNK9, KCNH3, HCN3, LRRC55, KCNA2, KCNB1, KCNJ13
Charolais	Tachykinin receptor activity	MF	$3.35 \times 10^{-04}$	0.022	TACR1, TACR2
Charolais	Cation channel activity	MF	$3.39 \times 10^{-04}$	0.022	KCNK9, KCNH3, HCN3, LRRC55, SCN4A, MCOLN3, KCNA2, CACNG1, KCNB1, KCNJ13, PKD1L2, NALF1
Charolais	Receptor ligand activity	MF	$7.78 \times 10^{-04}$	0.044	IL6, EPGN, IL17B, SPP1, F2, CCL27, FGF5, IL37, XCL1, SST, TAFA4, CARTPT, LTB, IL27
Charolais	Signaling receptor activator activity	MF	$8.92 \times 10^{-04}$	0.045	IL6, EPGN, IL17B, SPP1, F2, CCL27, FGF5, IL37, XCL1, SST, TAFA4, CARTPT, LTB, IL27
Charolais	Gated channel activity	MF	0.001	0.049	KCNK9, KCNH3, HCN3, LRRC55, SCN4A, MCOLN3, KCNA2, CACNG1, KCNB1, KCNJ13, NALF1
Kinsella	Contractile fiber	CC	$1.82 \times 10^{-38}$	$7.88 \times 10^{-36}$	MYH4, MYPN, MYL6B, KY, MYOZ2, CAPN3, LMOD3, IGFN1, TNNT1, TNNI1, CACNA1S, TNNC1, CAV3, TNNC2, CSR3, MYLPF, MYL2, TMOD4, LRRC39, MYL1, NEB, MYH1, MYH7, MYOM2, ACTN3, MYOT, TNNT3, TNNI2, MYL3, MYOZ1, SMTNL1, MYBPC1, ACTA1, MYH2, NRAP, RYR1, CASQ1, LMOD2, TRIM54, XIRP2, MYH8, LDB3, ANKRD1, CAVIN4, MYLK2, SYNPO2L, CMYA5, CASQ2, MYH6

Continued

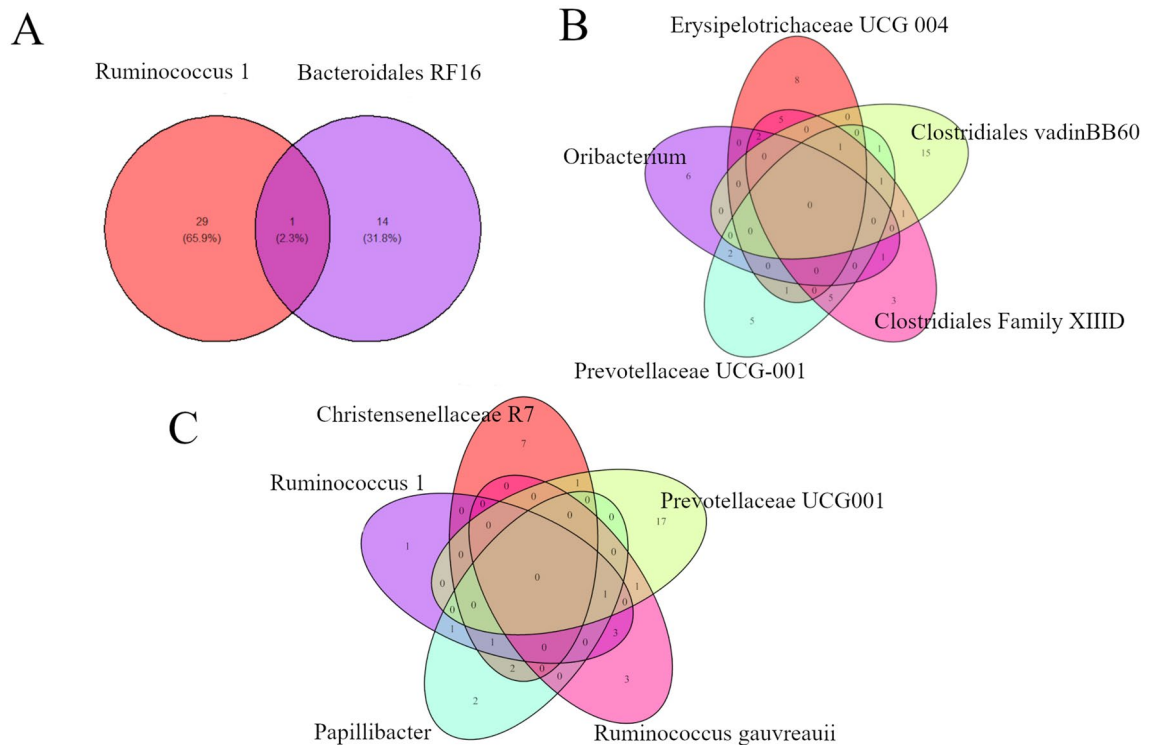


Breed	Description	GO class	P-value	FDR	Gene
Kinsella	Sarcomere	CC	$1.56 \times 10^{-37}$	$3.39 \times 10^{-35}$	MYH4, MYPN, KY, MYOZ2, CAPN3, LMOD3, IGFN1, TNNT1, TNNI1, CACNA1S, TNNC1, CAV3, TNNC2, CSRP3, MYL2, TMOD4, LRRC39, MYL1, NEB, MYH1, MYH7, MYOM2, ACTN3, MYOT, TNNT3, TNNI2, MYL3, MYOZ1, SMTNL1, ACTA1, MYH2, NRAP, RYR1, CASQ1, LMOD2, TRIM54, XIRP2, MYH8, LDB3, ANKRD1, CAVIN4, MYLK2, SYNPO2L, CMYA5, CASQ2, MYH6
Kinsella	Myofibril	CC	$6.20 \times 10^{-37}$	$8.95 \times 10^{-35}$	MYH4, MYPN, KY, MYOZ2, CAPN3, LMOD3, IGFN1, TNNT1, TNNI1, CACNA1S, TNNC1, CAV3, TNNC2, CSRP3, MYL2, TMOD4, LRRC39, MYL1, NEB, MYH1, MYH7, MYOM2, ACTN3, MYOT, TNNT3, TNNI2, MYL3, MYOZ1, SMTNL1, MYBPC1, ACTA1, MYH2, NRAP, RYR1, CASQ1, LMOD2, TRIM54, XIRP2, MYH8, LDB3, ANKRD1, CAVIN4, MYLK2, SYNPO2L, CMYA5, CASQ2, MYH6
Kinsella	Muscle system process	BP	$2.94 \times 10^{-29}$	$1.04 \times 10^{-25}$	STAC2, MYH4, CHRM2, SSTR2, PRKAG3, MYL6B, MLIP, STAC3, P2RX2, MYOZ2, LMOD3, TNNT1, TNNI1, CACNA1S, SCN4A, TRIM72, SLN, TNNC1, CAV3, ATP2A1, TNNC2, CSRP3, MYLPE, MYL2, MYBPC2, TMOD4, MYL1, MYBPH, MYH1, MYH7, MYOM2, ACTN3, MYOT, TNNT3, TNNI2, MYL3, MYOZ1, MYBPC1, ACTA1, MYH2, CHRNA1, RYR1, CASQ1, LMOD2, MYH8, CLCN1, JSRP1, MYLK2, CKMT2, CASQ2, MYH6, SCN4B, MSTN
Kinsella	Muscle contraction	BP	$1.60 \times 10^{-27}$	$2.86 \times 10^{-24}$	STAC2, MYH4, CHRM2, SSTR2, MYL6B, STAC3, P2RX2, LMOD3, TNNT1, TNNI1, CACNA1S, SCN4A, TNNC1, CAV3, ATP2A1, TNNC2, CSRP3, MYLPE, MYL2, MYBPC2, TMOD4, MYL1, MYBPH, MYH1, MYH7, MYOM2, ACTN3, MYOT, TNNT3, TNNI2, MYL3, MYOZ1, MYBPC1, ACTA1, MYH2, CHRNA1, RYR1, CASQ1, LMOD2, MYH8, CLCN1, JSRP1, MYLK2, CKMT2, CASQ2, MYH6, SCN4B
Kinsella	Myofibril assembly	BP	$1.75 \times 10^{-21}$	$2.07 \times 10^{-18}$	MYPN, MYOZ2, CAPN3, LMOD3, TNNT1, CAV3, CSRP3, MYL2, TMOD4, NEB, MYOM2, TNNT3, MYOZ1, ACTA1, NRAP, CASQ1, LMOD2, LDB3, ANKRD1, SYNPO2L, MYH6
Kinsella	Striated muscle cell development	BP	$2.51 \times 10^{-21}$	$2.23 \times 10^{-18}$	MYPN, MYOZ2, CAPN3, LMOD3, TNNT1, CAV3, CSRP3, MYL2, TMOD4, NEB, MYOM2, TNNT3, MYOZ1, ACTA1, NRAP, CASQ1, LMOD2, LDB3, ANKRD1, SYNPO2L, MYH6
Kinsella	I band	CC	$1.08 \times 10^{-19}$	$1.17 \times 10^{-17}$	MYPN, KY, MYOZ2, CAPN3, IGFN1, CACNA1S, CAV3, CSRP3, NEB, MYH7, ACTN3, MYOT, MYL3, MYOZ1, SMTNL1, NRAP, RYR1, CASQ1, TRIM54, XIRP2, LDB3, ANKRD1, CAVIN4, SYNPO2L, CASQ2, MYH6
Kinsella	Skeletal muscle contraction	BP	$1.64 \times 10^{-19}$	$1.17 \times 10^{-16}$	STAC2, STAC3, TNNT1, TNNI1, SCN4A, TNNC1, CAV3, ATP2A1, TNNC2, MYH7, ACTN3, TNNT3, TNNI2, CHRNA1, CASQ1, MYH8, JSRP1
Kinsella	Muscle cell development	BP	$7.61 \times 10^{-19}$	$4.52 \times 10^{-16}$	SHOX2, TBX18, MYPN, STAC3, ALPK2, P2RX2, MYOZ2, CAPN3, LMOD3, TNNT1, CAV3, CSRP3, MYL2, TMOD4, NEB, MYOM2, ACTN3, TNNT3, MYOZ1, ACTA1, NRAP, RYR1, CASQ1, LMOD2, LDB3, ANKRD1, SYNPO2L, MYH6

**Table 4.** Top 10 enriched gene ontology terms for the genes selected which better classify high- and low-RFI animals based on the expression in the rumen tissue in the sPLS-DA for Angus, Charolais and Kinsella breeds.

p-value = 0.013) and *Bacteroidales RF16 uncultured rumen bacterium* with *CCR3* ( $\rho = 0.57$ , p-value = 0.30) and *CXCL10* ( $\rho = 0.60$ , p-value = 0.18). Interestingly, different cytokines and chemokines were already identified as less expressed in the rumen of beef cattle with greater average daily gain (ADG) when compared with animals with lesser ADG<sup>90</sup>.

For Charolais, the *KCNK9* gene was correlated with four out five differential abundant bacterial genera (*Clostridiales vadin BB60 uncultured rumen bacterium*, *Clostridiales Family XIII uncultured*, *Prevotellaceae UCG-001*, and *Erysipelotrichaceae UCG-004*). For *Clostridiales vadin BB60 uncultured rumen bacterium* ( $\rho = -0.546$ , p-value = 0.02), *Clostridiales Family XIII uncultured* ( $\rho = -0.552$ , p-value = 0.02), and *Erysipelotrichaceae UCG-004* ( $\rho = -0.517$ , p-value = 0.04) positive correlations were observed. On the other hand, for *Prevotellaceae UCG-001* a negative correlation was observed ( $\rho = 0.571$ , p-value = 0.02). The *KCNK9* encodes a pH-dependent potassium channel called TASK3, which has a ubiquitous expression, but a remarkable function on T lymphocytes is observed due to its effects on downstream functions such as secretion of proinflammatory cytokines<sup>91</sup>. Consequently, playing an important role of the modulation of the immune system. Additionally, the genes *GGT1* and *IL6* were significantly correlated with *Clostridiales Family XIII uncultured*, *Erysipelotrichaceae UCG-004*, and *Oribacterium*. An interesting opposite correlation pattern between these genes was observed for the three genera. In *Erysipelotrichaceae UCG-004*, *GGT1* was positively correlated with the relative abundance of this genus ( $\rho = 0.731$ , p-value = 0.001), while *IL6* was negatively correlated ( $\rho = -0.548$ , p-value = 0.027). Similarly, for *Clostridiales Family XIII uncultured* *GGT1* ( $\rho = 0.776$ , p-value = 0.0004) and *IL6* ( $\rho = -0.601$ , p-value = 0.013), were also positively and negatively correlated with the relative abundance of this genus, respectively. On the other hand, for *Oribacterium*, a reversed pattern is observed, with *GGT1* ( $\rho = -0.567$ , p-value = 0.021) being negatively correlated and *IL6* ( $\rho = 0.562$ , p-value = 0.023) being positively correlated. The *IL6* (interleukin 6) is an important inflammatory factor and was previously observed to show a significant increase in the rumen of goats and cows presenting subacute rumen acidosis<sup>92,93</sup>. Additionally, the metabolic pathways associated with *IL6* seems to connect the differential abundant genus identified for Charolais breed and their respective correlated genes in a single network (Fig. 3). The *GGT1* ( $\gamma$ -glutamyl transferase) codifies a cell surface enzyme responsible to cleave extracellular glutathione, GSH S-conjugates, among others  $\gamma$ -glutamyl compounds<sup>94</sup>. The glutathione is an antioxidant which plays crucial roles in the maintenance and regulation of thiol-redox status in the cell and contributes to the reduction of oxidative stress in dairy cows<sup>95,96</sup>. Despite this interesting correlation pattern, any direct relationship between *GGT1* and *IL6* was found in the literature.



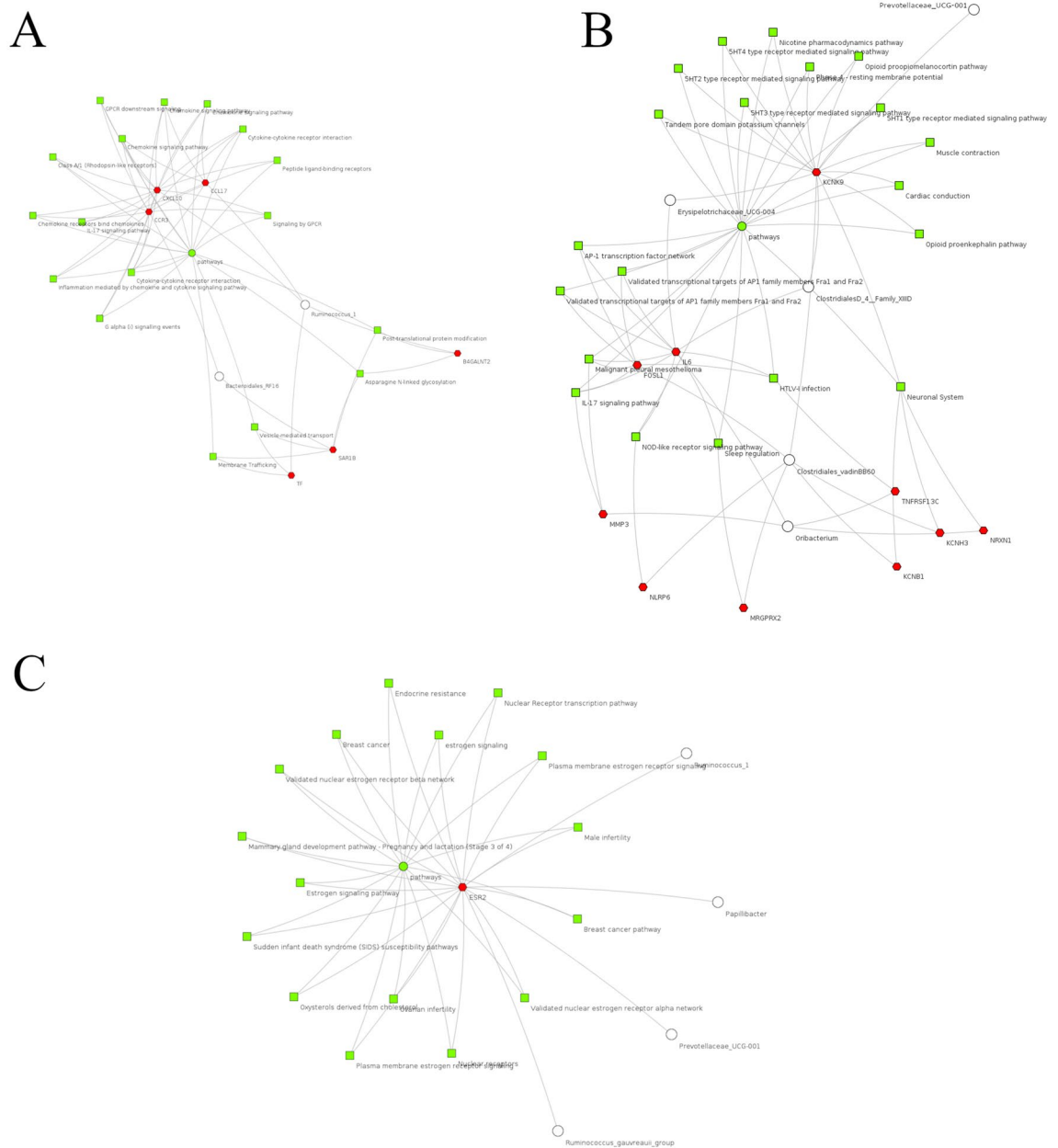
**Figure 2.** Venn diagrams showing the number of correlated genes shared between different bacterial genus for Angus (A), Charolais (B) and Kinsella (C) breeds.

In the Kinsella composite hybrid breed, the gene *ESR2* was positively correlated with three genera, which were *Ruminococcus 1* ( $\rho = 0.750$ ,  $p$ -value = 0.0008), *Prevotellaceae UCG-001* ( $\rho = 0.627$ ,  $p$ -value = 0.009), and *Ruminococcus gauvreauii* group ( $\rho = 0.661$ ,  $p$ -value = 0.005). In addition, a negative correlation was observed between *ESR2* and *Papillibacter* ( $\rho = -0.620$ ,  $p$ -value = 0.010). Interestingly, a pivotal role of *ESR2* was observed in the connection of these differentially abundant genus through different metabolic pathways (Fig. 3). The *ESR2* gene is responsible to codify the estrogen receptor beta, a transcription factor from the steroid hormone receptor superfamily. Among other function, such as reproductive-related functions, *ESR2* plays crucial roles in the glucose transporter *GLUT4* regulation and polymorphisms in this gene are associated with lipid level, insulin resistance and body fat in humans<sup>97,98</sup>. Additionally, the expression of *ESR2* was identified as inhibited in the rumen papilla of weaned cows, when compared with suckling calves<sup>99</sup>. In mice, the knockout of *ESR1*, a paralogous of *ESR2*, resulted in different levels of bacteria in the seminal fluid and gut<sup>100</sup>. Taken together, these results may suggest a potential role of *ESR2* in the regulation of bacterial communities in the rumen and a possible association with the control of feed efficiency status in beef cattle.

The results discussed here were obtained through a multi-breed comparison. Despite all the animals being raised under the same environmental and management conditions it is not possible to exclude the existence of hidden variables which might affect the rumen gene expression and microbiome diversity. These differences might be responsible, at least in part, by the relatively low overlapping of the results between breeds. However, these results also suggest a potential strong breed dependent effect that must be in deep investigated in further studies. Additionally, it is important to highlight that we are investigating the expression profile at rumen epithelium and generalizations regarding the direction and magnitude of the functional effect of the candidate genes in the whole organism scale must be carefully performed. The rumen epithelium is an important part of the components affecting the nutritional status<sup>48</sup> and both ASVs and DEG identified here has the potential to directly act over biological mechanisms responsible to regulate feed efficiency in beef cattle.

## Conclusion

Studies on the genetic architecture of the rumen epithelium microbiome and host interactions and how these interactions influence production traits, such as feed efficiency, are newly emerging. Integration of metagenomics and host-transcriptome can provide the taxonomic and functional profile of important components of the complex system responsible for the differences observed between high and low efficient animals. At the transcriptome level, despite Angus and Charolais don't reach the statistical threshold ( $FDR < 0.05$ ) to identify DEG the analysis of genes with a significant biological threshold ( $|\log_2(FC)| > 2$ ) suggested similar processes differentially expressed between HRFI and LRFI groups for all the three breeds. The regulation of rumen contraction and feed behaviour were associated with the abovementioned genes, mainly through downregulation in the low-RFI group, in all the three cattle. The metagenomic analysis indicated that only bacterial ASVs (at genus level) were differentially abundant between the feed efficiency groups in all the three breeds in the rumen epithelium. Therefore, the current study, using a multi-breed approach, identified significantly correlated host genes associated with RFI



**Figure 3.** Comparative clustering of metabolic pathways (green squares) annotated for the genes (red hexagons) significantly correlated with differentially abundant ASVs (white circles) between high- and low-RFI animals for Angus (A), Charolais (B) and Kinsella (C).

with differentially abundant rumen bacteria ASVs at genus level. The functional profile suggest that these genes may be interacting with the host microbiome, modulating the feed efficiency status across beef cattle breeds. Overall, the results of this study help to better understand the biological processes associated with the microbial dynamics in the rumen epithelium and the production response in the host. Additionally, it may suggest the application of rumen microbiome features that lead to desirable host gene function and phenotypes, providing more insight on targeting both host and rumen epithelium microbial traits that lead to the selection and breeding of more feed efficient beef cattle. Further studies on host-microbiome interactions and complementary systems biology analysis can lead to a specific host and rumen microbiome profile which is characterized for high feed efficient cattle. Additionally, the characterization of breed potential dependent effects over the microbiome and host interaction in the rumen epithelium should be further investigate to provide better management strategies to improve the feed efficiency in each breed.

**Materials and methods**

**Animal, sampling, and RFI measurement information**

The animal and sampling information, as well as the Ethics Committee approval information, has been described in detail by Sun et al.<sup>25</sup>. All animals used in this study were managed based on the guidelines established by the

Canadian Council of Animal Care and the experimental procedures were approved by the University of Alberta Livestock Animal Care and Use Committee (Protocol No.: AUP00000927). The experimental design and the analyses performed in the current study are in accordance with ARRIVE guidelines. Briefly, a total of 48 beef steers from a herd of 738 cattle cohort from the Roy Berg Kinsella Research Ranch, University of Alberta (Alberta, Canada) were used in this study. The beef steer breeds included two pure beef cattle breeds; Angus (AN; n = 16), and Charolais (CH; n = 16), and one composite hybrid Kinsella (KN; n = 16). Within each breed, from the 16 total animals, 8 animals were classified in the high RFI group and 8 in the low RFI group. The Kinsella herd is a composite hybrid, consisting of Angus, Charolais, Galloway, Hereford, Holstein, Brown Swiss, and Simmental breeds<sup>3</sup>.

The RFI estimates were calculated using data recorded for dry matter intake (DMI), which were collected using an automated feeding system (GrowSafe system Ltd. Airdrie, Alberta, Canada), through a 70-d period using the RFI model described by Lancaster et al.<sup>101</sup>. Rumen epithelium samples were collected from all animals within 30 min of slaughter after the experiment, which were immediately submerged in liquid nitrogen and then stored at -80°C until RNA isolation.

### RNA extraction, library construction, and sequencing of rumen epithelium RNA-Seq data

Detailed information on RNA extraction, library construction and sequencing are described by Sun et al.<sup>25</sup>. Briefly, 100 mg of rumen epithelium tissue was ground to fine powder using sterilized mortar, and total RNA was isolated using mirVana total RNA Isolation Kit (Ambion, Carlsbad, CA, USA) following the manufacturer's instructions. The RNA was quantified using the Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA) and checked for purity and integrity using the Agilent 2200 TapeStation (Agilent Technologies, Santa Clara, CA, USA). The RNA with the ratio of 28S/18S rRNA ratio ranging from 1.7 to 2.4 and the RNA integrity number > 7.0 was used for RNA-Seq library construction. One microgram of isolated total RNA per sample was used to construct a cDNA library for RNA-Seq according to the protocol of TruSeq Stranded Total RNA Sample Prep Kit (Illumina, San Diego, CA, USA). The cDNA libraries with qualified concentration ( $\geq 2$  nM) were performed in 16 lanes (12 samples/lane) on Illumina HiSeq 4000 sequencing platform (Illumina, San Diego, CA, USA) to obtain paired-end reads (2 × 100 bp, average Phred quality score  $\geq 33$ ) at the McGill University and Genome Quebec Innovation Centre (Montreal, Quebec, Canada).

### Reads mapping, assembling and annotation of rumen epithelium RNA-Seq data

The mRNA sequence reads were aligned to the new bovine reference genome (*Bos Taurus* ARS-UCD 1.2. bovine reference genome) with the ENSEMBL annotation tool (<http://www.ensembl.org/info/data/ftp/index.html>) using CLC Genomics Workbench (CLC Version 12.0.2., Aarhus, Denmark). Quality control analysis was performed in all samples as described by Cánovas et al.<sup>102</sup>. Briefly, quality control was performed using the 'NGS quality control' tool of CLC Genomics Workbench (CLC Version 12.0.2., Aarhus, Denmark), which assessed GC content (50% GC base content and less than 0.1% over-represented sequences), ambiguous base content, same length reads (100 bp), Phred score (Phred  $\geq 33$ ), base coverage (100% coverage in all bases), nucleotide contributions and over-represented sequence parameters (25% of A, T, G and C nucleotide contributions)<sup>102</sup>. One sample from the low-RFI Angus group (ID = 407) was removed from the study due to poor alignment quality (uniquely mapped reads < 80%). The remaining samples (n = 47) passed the quality control analysis according to the conditions stated above<sup>5</sup>.

### Differential gene expression analysis of rumen epithelium RNA-Seq data for low and high-RFI gene modules

Using the 'Empirical Analysis of differential Gene Expression' tool in CLC Genomics Workbench (CLC Version 12.0.2., Aarhus, Denmark), differentially expressed genes (DEG) were identified between low-RFI (more feed efficient) and high-RFI (less feed efficient) animals for each breed (AN low-RFI n = 7; high-RFI n = 8; CH low-RFI n = 8; high-RFI n = 8; KN low-RFI n = 8; high-RFI n = 8) by t-test. Gene expression levels were quantified in reads per kilo base per million mapped reads (RPKM). Genes were considered significantly differentially expressed between low and high-RFI groups when they met the following statistical and biological significance thresholds: False Discovery Rate (FDR)  $q < 0.05$ , and  $\log_2$  of fold-change ( $|\log_2(\text{FC})| > 2$ ).

### Rumen contents DNA extraction and archaea and bacterial Amplicon-Sequencing analysis

Total genomic DNA was isolated from rumen digesta using the repeated bead beating plus column (RBB + C) method, which is previous described by Yu et al.<sup>103</sup>. Briefly, the quality and quantity of DNA was measured using a NanoDrop Spectrophotometer ND-1000 (Thermo Fisher Scientific Inc., Wilmington, Delaware, USA).

The Amplicon-Seq analysis was performed using Quantitative Insights into Microbial Ecology 2, (QIIME2) (Ref.<sup>104</sup>; <http://qiime.org/>). Amplicon sequence libraries were processed, and low-quality ends of reads and primers were trimmed. The forward primer Arc915aF-CS1 ACACGTGACGACATGGTTCTACAAGGAATTG GCGGGGAGCAC and reverse primer Arc1386R-CS2 TACGGTAGCAGAGACTTGGTCTGCGGTGTGTGC AAGGAGC were used to amplify the archaea 16S rRNA gene amplicon region. The forward primer Bac9F-CS1 ACACGTGACGACATGGTTCTACAGAGTTTGATCMTGGCTCAG and reverse primer Bac515R-CS2 TACGGT AGCAGAGACTTGGTCTCCGCGGCKGCTGGCAC were used to amplify the bacteria 16S rRNA gene region. The database used to align the archaea reads was Rumen and Intestinal Methanogen-DB (RIM-DB)<sup>105</sup> and the database used to align the bacteria reads was Silva ribosomal RNA gene database Release 132<sup>106-109</sup>. A feature table and representative sequences table for bacteria and archaea profiles for each breed was generated for further analysis. Using these tables, a taxonomy file including ASV abundance for archaea and bacteria were generated. Detailed information regarding the animal management and condition during the rumen digesta collection is available in Li et al.<sup>15</sup>.



## Comparison of rumen archaea and bacteria taxonomic profiles between low and high-RFI groups across each pure breed and composite hybrid

Using QIIME2 command 'qiime feature-table relative-frequency', an output of the taxonomic ASV abundance based on total counts of each ASV in each taxa level (Kingdom, Phylum, Class, Order, Family, Genus, Species), for each animal was generated. Tables were then edited to have one table for each breed incorporating the ASV abundance for each taxonomic level, the taxonomic identification names, and the sample information. The ASV were analyzed regarding the differential abundance at genus level for bacteria and species level for archaea. Only taxonomic groups with an overall relative abundance > 0.5% were retained for the differential abundance analysis. Subsequently, for each animal within each breed, the relative abundance was calculated for the retained ASVs. To determine the significantly differentially abundant ASVs between high and low-RFI animals within each breed analysis (p-value < 0.05), the relative abundances were arcsine square root-transformed and a t-test was applied as described by Li et al.<sup>15</sup>.

## Identification of correlations between differentially abundant ASVs and gene expression

Initially, a sPLA-DA was performed for each breed individually using the read counts for those genes with  $|\log_2(\text{FC})| > 2$  in the comparison between high- and low-RFI groups. The sPLA-DA was performed using the mixOmics R package with the read counts standardized to zero means and unit variances<sup>110</sup>. The selectVar() function from mixOmics package was used to select a subset of genes which better differentiated the high and low-RFI groups based on the feature loadings from the first principal component of the sPLA-DA. The Spearman correlation between the read counts of the selected genes for each breed and the arcsine square root-transformed relative abundance of respective differentially abundant ASVs were estimated using the cor.test() function in R<sup>111</sup>. The significant correlations were defined as those with a p-value < 0.05. Subsequently, the ToppCluster software was used to perform a comparative clustering of metabolic pathways annotated for the genes significantly correlated with different ASVs within each breed<sup>112</sup>.

## Gene Ontology annotation and enrichment analysis

The R packages ClusterProfiler and enrichplot were used for gene ontology (GO) term enrichment analysis, graphic representation and functional grouping<sup>113</sup>. The GO enrichment analysis was performed for each breed for all the genes with  $|\log_2(\text{FC})| > 2$ , for the subset of genes selected in the sPLS-DA analysis, and for each subset of genes correlated with each differentially abundant ASV. For GO terms annotated for all the genes with  $|\log_2(\text{FC})| > 2$ , the function pairwise\_termsim() from enrichplot was used to calculate the Jaccard correlation coefficient. The terms were functionally grouped using the similarity matrix built by the Jaccard correlation coefficient to identify terms functionally closely related and reduce redundancy across terms.

## Data availability

The distribution of gene expression data can be accessed at <https://www.cattleomics.com/transcriptome>. The metagenome reads are available at National Center for Biotechnology Information (NCBI) under the accession ID PRJNA448333.

Received: 15 December 2022; Accepted: 19 September 2023

Published online: 27 September 2023

## References

1. Seymour, D. J. et al. Development of a robust, net energy-based measure of feed efficiency in dairy cattle. *J. Dairy Sci.* **102**, 131–131 (2019).
2. Ominski, K. et al. The role of livestock in sustainable food production systems in Canada. *Can. J. Anim. Sci.* **101**, 591–601. <https://doi.org/10.1139/cjas-2021-0005> (2021).
3. Nkrumah, J. D. et al. Genetic and phenotypic relationships of feeding behavior and temperament with performance, feed efficiency, ultrasound, and carcass merit of beef cattle. *J. Anim. Sci.* **85**, 2382–2390 (2007).
4. Wickramasinghe, S., Rincon, G., Islas-Trejo, A. & Medrano, J. F. Transcriptional profiling of bovine milk using RNA sequencing. *BMC Genom.* **13**, 1–14. <https://doi.org/10.1186/1471-2164-13-45> (2012).
5. Cánovas, A. et al. Comparison of five different RNA sources to examine the lactating bovine mammary gland transcriptome using RNA-Sequencing. *Sci. Rep.* **4**, 5297–5297. <https://doi.org/10.1038/srep05297> (2014).
6. De Souza Fonseca, P. A. et al. Combining multi-OMICs information to identify key-regulator genes for pleiotropic effect on fertility and production traits in beef cattle. *PLoS One* **13**(10), e0205295. <https://doi.org/10.1371/journal.pone.0205295> (2018).
7. Cánovas, A. et al. Segregation of regulatory polymorphisms with effects on the gluteus medius transcriptome in a purebred pig population. *PLoS One* **7**(4), e35583. <https://doi.org/10.1371/journal.pone.0035583> (2012).
8. Cardoso, T. F. et al. RNA-seq based detection of differentially expressed genes in the skeletal muscle of Duroc pigs with distinct lipid profiles. *Sci. Rep.* **7**(1), 40005. <https://doi.org/10.1038/srep40005> (2017).
9. Dias, M. M. et al. SNP detection using RNA-sequences of candidate genes associated with puberty in cattle. *Genet. Mol. Res.* **16**(1), gmr16019522. <https://doi.org/10.4238/gmr16019522> (2017).
10. Lam, S. et al. Development and comparison of RNA-sequencing pipelines for more accurate SNP identification: Practical example of functional SNP detection associated with feed efficiency in Nellore beef cattle. *BMC Genom.* **21**, 1–17. <https://doi.org/10.1186/s12864-020-07107-7> (2020).
11. Lam, S. et al. Identification of functional candidate variants and genes for feed efficiency in Holstein and Jersey cattle breeds using RNA-sequencing. *J. Dairy Sci.* **104**, 1928–1950 (2021).
12. Williams, A. G. & Coleman, G. S. The rumen protozoa. In *The Rumen Microbial Ecosystem* (eds Hobson, P. N. & Stewart, C. S.) 73–139 (Springer, 1997). [https://doi.org/10.1007/978-94-009-1453-7\\_3](https://doi.org/10.1007/978-94-009-1453-7_3).
13. Gruninger, R. J., Ribeiro, G. O., Cameron, A. & McAllister, T. A. Invited review: Application of meta-omics to understand the dynamic nature of the rumen microbiome and how it responds to diet in ruminants. *Animal* **13**(9), 1843–1854. <https://doi.org/10.1017/S1751731119000752> (2019).



14. Li, F. *et al.* Host genetics influence the rumen microbiota and heritable rumen microbial features associate with feed efficiency in cattle. *Microbiome* **7**, 1–17 (2019).
15. Li, F., Hitch, T. C. A., Chen, Y., Creevey, C. J. & Guan, L. L. Comparative metagenomic and metatranscriptomic analyses reveal the breed effect on the rumen microbiome and its associations with feed efficiency in beef cattle 06 Biological Sciences 0604 Genetics 06 Biological Sciences 0605 Microbiology. *Microbiome* **7**, 1–21 (2019).
16. Myer, P. R. Bovine genome-microbiome interactions: metagenomic frontier for the selection of efficient productivity in cattle systems. *mSystems* **4**(3), 10–1128 (2019).
17. Zhou, M. *et al.* Assessment of microbiome changes after rumen transfaunation: Implications on improving feed efficiency in beef cattle. *Microbiome* **6**, 11–14 (2018).
18. Clemmons, B. A., Voy, B. H. & Myer, P. R. Altering the gut microbiome of cattle: considerations of host-microbiome interactions for persistent microbiome manipulation. *Microb. Ecol.* **77**, 523–526. <https://doi.org/10.1007/s00248-018-1234-9> (2019).
19. Koch, R. M., Swiger, L. A., Chambers, D. & Gregory, K. E. Efficiency of feed use in beef cattle. *J. Anim. Sci.* **22**(2), 486–494. <https://doi.org/10.2527/jas1963.222486x> (1963).
20. Sasson, G. *et al.* Heritable bovine rumen bacteria are phylogenetically related and correlated with the cow's capacity to harvest energy from its feed. *mBio* **8**, 10–1128 (2017).
21. Difford, G. F. *et al.* Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows. *PLoS Genet.* **14**, e1007580 (2018).
22. John Wallace, R. *et al.* A heritable subset of the core rumen microbiome dictates dairy cow productivity and emissions. *Sci. Adv.* **5**(7), 8391. <https://doi.org/10.1126/sciadv.aav8391> (2019).
23. Malmuthuge, N. & Guan, L. L. Understanding host-microbial interactions in rumen: Searching the best opportunity for microbiota manipulation. *J. Anim. Sci. Biotechnol.* **8**(1), 1–7. <https://doi.org/10.1186/s40104-016-0135-3> (2017).
24. Kong, R. S. G., Liang, G., Chen, Y., Stothard, P. & Guan, L. L. Transcriptome profiling of the rumen epithelium of beef cattle differing in residual feed intake. *BMC Genom.* **17**, 1–16 (2016).
25. Sun, H. Z., Zhao, K., Zhou, M., Chen, Y. & Guan, L. L. Landscape of multi-tissue global gene expression reveals the regulatory signatures of feed efficiency in beef cattle. *Bioinformatics* **35**, 1712–1719 (2019).
26. Perry, S. V. The regulation of contractile activity in muscle. *Biochem. Soc. Trans.* **7**, 593–617 (1979).
27. Alami-Durante, H., Wrutniak-Cabello, C., Kaushik, S. J. & Médale, F. Skeletal muscle cellularity and expression of myogenic regulatory factors and myosin heavy chains in rainbow trout (*Oncorhynchus mykiss*): Effects of changes in dietary plant protein sources and amino acid profiles. *Compar. Biochem. Physiol. Mol. Integr. Physiol.* **156**, 561–568 (2010).
28. Bottje, W. G. *et al.* Gene expression in breast muscle associated with feed efficiency in a single male broiler line using a chicken 44K microarray. II. Differentially expressed focus genes. *Poult. Sci.* **91**, 2535–2547 (2012).
29. Poompramun, C. *et al.* Revealing pathways associated with feed efficiency and meat quality traits in slow-growing chickens. *Animals* **11**, 2977 (2021).
30. Rayment, I. *et al.* Structure of the actin-myosin complex and its implications for muscle contraction. *Science* **1979**(261), 58–65 (1993).
31. Geeves, M. A. The dynamics of actin and myosin association and the crossbridge model of muscle contraction. *Biochem. J.* **274**, 1. <https://doi.org/10.1042/bj2740001> (1991).
32. Roehe, R. *et al.* Bovine host genetic variation influences rumen microbial methane production with best selection criterion for low methane emitting and efficiently feed converting hosts based on metagenomic gene abundance. *PLoS Genet.* **12**, 1–20 (2016).
33. Pinares-Patiño, C. S., Ulyatt, M. J., Lassey, K. R., Barry, T. N. & Holmes, C. W. Rumen function and digestion parameters associated with differences between sheep in methane emissions when fed chaffed lucerne hay. *J. Agric. Sci.* **140**, 205–214 (2003).
34. Goopy, J. P. *et al.* Low-methane yield sheep have smaller rumens and shorter rumen retention time. *Br. J. Nutr.* **111**, 578–585 (2014).
35. Kowalczyk, J. An introduction to rumen studies. *Anim. Feed Sci. Technol.* **22**, 269 (1989).
36. Van Soest, P. J. Nutritional ecology of the ruminant. *Nutr. Ecol. Ruminant* <https://doi.org/10.7591/9781501732355> (2019).
37. Cabezas-García, E. H., Krizsan, S. J., Shingfield, K. J. & Huhtanen, P. Between-cow variation in digestion and rumen fermentation variables associated with methane production. *J. Dairy Sci.* **100**, 4409–4424 (2017).
38. Pinares-Patino, C. S. *et al.* Is rumen retention time implicated in sheep differences in methane emission?. *Proc. New Zealand Soc. Anim. Prod.* **71**, 219–222 (2011).
39. Pinares-Patiño, C. S. & Clark, H. Rumen function and digestive parameters associated with methane emissions in dairy cows. *Proceedings of 4th Australasian Dairy Science Symposium 2010* (2010).
40. Lovendahl, P. *et al.* Review: Selecting for improved feed efficiency and reduced methane emissions in dairy cattle. *Animal* **12**, s336–s349. <https://doi.org/10.1017/S1751731118002276> (2018).
41. Shimomura, Y. *et al.* Identification of neuropeptide W as the endogenous ligand for orphan G-protein-coupled receptors GPR7 and GPR8. *J. Biol. Chem.* **277**, 35826–35832 (2002).
42. Baker, J. R., Cardinal, K., Bober, C., Taylor, M. M. & Samson, W. K. Neuropeptide W acts in brain to control prolactin, corticosterone, and growth hormone release. *Endocrinology* **144**, 2816–2821 (2003).
43. Mondal, M. S. *et al.* A role for neuropeptide W in the regulation of feeding behavior. *Endocrinology* **144**, 4729–4733 (2003).
44. Lau, J. & Herzog, H. CART in the regulation of appetite and energy homeostasis. *Front. Neurosci.* **8**, 313 (2014).
45. Murphy, K. G. Dissecting the role of cocaine- and amphetamine-regulated transcript (CART) in the control of appetite. *Brief. Funct. Genom. Proteom.* **4**, 95–111. <https://doi.org/10.1093/bfpg/4.2.95> (2005).
46. Rigoli, L. *et al.* Molecular analysis of the CART gene in overweight and obese Italian children using family-based association methods. *Acta Paediatr. Int. J. Paediatr.* **99**, 722–726 (2010).
47. del Giudice, E. M. *et al.* Mutational screening of the CART gene in obese children. *Diabetes* **50**, 2157–2160 (2001).
48. Na, S. W. Understanding the role of rumen epithelial host-microbial interactions in cattle feed efficiency. *Anim. Nutr.* **10**, 41–53 (2022).
49. Roehe, R. *et al.* Bovine host genetic variation influences rumen microbial methane production with best selection criterion for low methane emitting and efficiently feed converting hosts based on metagenomic gene abundance. *PLoS Genet.* **12**, e1005846. <https://doi.org/10.1371/journal.pgen.1005846> (2016).
50. Montanholi, Y. R. *et al.* Assessing feed efficiency in beef steers through feeding behavior, infrared thermography and glucocorticoids. *Animal* **4**, 692–701. <https://doi.org/10.1017/S1751731109991522> (2010).
51. Nkrumah, J. D. *et al.* Genetic and phenotypic relationships of feeding behavior and temperament with performance, feed efficiency, ultrasound, and carcass merit of beef cattle. *J. Anim. Sci.* **85**, 2382–2390. <https://doi.org/10.2527/jas.2006-657> (2007).
52. Powers, M. GABA supplementation and growth hormone response. *Med. Sport Sci.* **59**, 36–46. <https://doi.org/10.1159/000341944> (2013).
53. Ku, B. S. *et al.* Effect of  $\gamma$ -aminobutyric acid (GABA) producing bacteria on in vitro rumen fermentation, biogenic amine production and anti-oxidation using corn meal as substrate. *Asian-Australas. J. Anim. Sci.* **26**, 804 (2013).
54. Wang, D. M., Wang, C., Liu, H. Y., Liu, J. X. & Ferguson, J. D. Effects of rumen-protected  $\gamma$ -aminobutyric acid on feed intake, lactation performance, and antioxidative status in early lactating dairy cows. *J. Dairy Sci.* **96**, 3222–3227 (2013).

55. Cheng, J. B. *et al.* Effects of rumen-protected  $\gamma$ -aminobutyric acid on performance and nutrient digestibility in heat-stressed dairy cows. *J. Dairy Sci.* **97**, 5599–5607 (2014).
56. Liu, C. *et al.* Dynamic alterations in yak rumen bacteria community and metabolome characteristics in response to feed type. *Front. Microbiol.* **10**, 1116 (2019).
57. Gui, L. S. *et al.* Altered milk yield and rumen microbial abundance in response to concentrate supplementation during the cold season in Tibetan sheep. *Electron. J. Biotechnol.* **53**, 80–86 (2021).
58. He, J. *et al.* Heat stress during late gestation disrupts maternal microbial transmission with altered offspring's gut microbial colonization and serum metabolites in a pig model. *Environ. Pollut.* **266**, 115111 (2020).
59. Mei, X. *et al.* Insulin sensitivity-enhancing activity of phlorizin is associated with lipopolysaccharide decrease and gut microbiota changes in obese and type 2 diabetes (db/db) mice. *J. Agric. Food Chem.* **64**, 7502–7511 (2016).
60. Zhang, X. *et al.* Modulation of gut microbiota by berberine and metformin during the treatment of high-fat diet-induced obesity in rats. *Sci. Rep.* **5**, 14405 (2015).
61. Zhang, X. *et al.* Structural changes of gut microbiota during berberine-mediated prevention of obesity and insulin resistance in high-fat diet-fed rats. *PLoS One* **7**, e42529 (2012).
62. Ren, Q. *et al.* Bacterial communities in the solid, liquid, dorsal, and ventral epithelium fractions of yak (*Bos grunniens*) rumen. *Microbiologyopen* **9**, e963 (2020).
63. Poudel, P., Froehlich, K., Casper, D. P. & St-Pierre, B. Feeding essential oils to neonatal holstein dairy calves results in increased ruminal prevotellaceae abundance and propionate concentrations. *Microorganisms* **7**, 120 (2019).
64. Quan, J. *et al.* Exploring the fecal microbial composition and metagenomic functional capacities associated with feed efficiency in commercial dly pigs. *Front. Microbiol.* **10**, 52 (2019).
65. Song, X. *et al.* Inulin can alleviate metabolism disorders in ob/ob mice by partially restoring leptin-related pathways mediated by gut microbiota. *Genom. Proteom. Bioinform.* **17**, 64–75 (2019).
66. Suen, G. *et al.* Complete genome of the cellulolytic ruminal bacterium *Ruminococcus albus* 7. *J. Bacteriol.* **193**, 5574–5575. <https://doi.org/10.1128/JB.05621-11> (2011).
67. Bergamaschi, M. *et al.* Gut microbiome composition differences among breeds impact feed efficiency in swine. *Microbiome* **8**, 1–15 (2020).
68. Lv, H. *et al.* Microbial composition in the duodenum and ileum of yellow broilers with high and low feed efficiency. *Front. Microbiol.* **12**, 689653 (2021).
69. Myer, P. R., Smith, T. P. L., Wells, J. E., Kuehn, L. A. & Freetly, H. C. Rumen microbiome from steers differing in feed efficiency. *PLoS One* **10**, e0129174 (2015).
70. Si, J. *et al.* Evaluating the association between feed efficiency and the fecal microbiota of early-life Duroc pigs using 16S rRNA sequencing. *AMB Express* **10**, 1–11 (2020).
71. Jiang, H., Fang, S., Yang, H. & Chen, C. Identification of the relationship between the gut microbiome and feed efficiency in a commercial pig cohort. *J. Anim. Sci.* **99**, skab045 (2021).
72. Metzler-Zebeli, B. U. *et al.* Fecal microbiota transplant from highly feed efficient donors affects cecal physiology and microbiota in low- And high-feed efficient chickens. *Front. Microbiol.* **10**, 1576 (2019).
73. Zeng, H., Guo, C., Sun, D., Seddik, H. E. & Mao, S. The ruminal microbiome and metabolome alterations associated with diet-induced milk fat depression in dairy cows. *Metabolites* **9**, 154 (2019).
74. Clemmons, B. A. *et al.* Rumen fluid metabolomics of beef steers differing in feed efficiency. *Metabolomics* **16**, 1–9. <https://doi.org/10.1007/s11306-020-1643-x> (2020).
75. Zhong, Y. *et al.* Rumen and hindgut bacteria are potential indicators for mastitis of mid-lactating holstein dairy cows. *Microorganisms* **8**, 2042 (2020).
76. Cheng, C. *et al.* Maternal soluble fiber diet during pregnancy changes the intestinal microbiota, improves growth performance, and reduces intestinal permeability in piglets. *Appl. Environ. Microbiol.* **84**, e01047–e1118 (2018).
77. Qiu, X. *et al.* Serum biochemical parameters, rumen fermentation, and rumen bacterial communities are partly driven by the breed and sex of cattle when fed high-grain diet. *Microorganisms* **10**, 323 (2022).
78. Liu, K. *et al.* Ruminal bacterial community is associated with the variations of total milk solid content in Holstein lactating cows. *Anim. Nutr.* **9**, 175–183 (2022).
79. Ogunade, I. M., McCoun, M., Idowu, M. D. & Peters, S. O. Comparative effects of two multispecies directfed microbial products on energy status, nutrient digestibility, and ruminal fermentation, bacterial community, and metabolome of beef steers. *J. Anim. Sci.* **98**, skaa201 (2020).
80. Waters, J. L. & Ley, R. E. The human gut bacteria Christensenellaceae are widespread, heritable, and associated with health. *BMC Biol.* **17**, 1–11. <https://doi.org/10.1186/s12915-019-0699-4> (2019).
81. Yang, H. *et al.* Unraveling the fecal microbiota and metagenomic functional capacity associated with feed efficiency in pigs. *Front. Microbiol.* **8**, 1–11 (2017).
82. Perea, K. *et al.* Feed efficiency phenotypes in lambs involve changes in ruminal, colonic, and small-intestine-located microbiota. *J. Anim. Sci.* **95**, 2585–2592 (2017).
83. Lopes, D. R. G. *et al.* Assessing the relationship between the rumen microbiota and feed efficiency in Nellore steers. *J. Anim. Sci. Biotechnol.* **12**, 1–17 (2021).
84. Lana, R. P., Russell, J. B. & Van Amburgh, M. E. The role of pH in regulating ruminal methane and ammonia production. *J. Anim. Sci.* **76**, 2190–2196. <https://doi.org/10.2527/1998.7682190x> (1998).
85. Lam, S. *et al.* Associations of rumen parameters with feed efficiency and sampling routine in beef cattle. *Animal* **12**, 1442–1450. <https://doi.org/10.1017/S1751731117002750> (2018).
86. Auffret, M. D. *et al.* Identification of microbial genetic capacities and potential mechanisms within the rumen microbiome explaining differences in beef cattle feed efficiency. *Front. Microbiol.* **11**, 1229 (2020).
87. Bagath, M. *et al.* The impact of heat stress on the immune system in dairy cattle: A review. *Res. Vet. Sci.* **126**, 94–102. <https://doi.org/10.1016/j.rvsc.2019.08.011> (2019).
88. Broom, L. J. & Kogut, M. H. The role of the gut microbiome in shaping the immune system of chickens. *Vet. Immunol. Immunopathol.* **204**, 44–51. <https://doi.org/10.1016/j.vetimm.2018.10.002> (2018).
89. Gabler, N. K. & Spurlock, M. E. Integrating the immune system with the regulation of growth and efficiency. *J. Anim. Sci.* **86**, E64–E74. <https://doi.org/10.2527/jas.2007-0466> (2008).
90. Reynolds, J. G., Foote, A. P., Freetly, H. C., Oliver, W. T. & Lindholm-Perry, A. K. Relationships between inflammation- and immunity-related transcript abundance in the rumen and jejunum of beef steers with divergent average daily gain. *Anim. Genet.* **48**, 447–449 (2017).
91. Meuth, S. G. *et al.* TWIK-related acid-sensitive K<sup>+</sup> channel 1 (TASK1) and TASK3 critically influence T lymphocyte effector functions. *J. Biol. Chem.* **283**, 14559–14570 (2008).
92. Hu, Z., Lin, M., Ma, X., Zhao, G. & Zhan, K. Effect of tea tree oil on the expression of genes involved in the innate immune system in goat rumen epithelial cells. *Animals* **11**, 2460 (2021).
93. Zhao, C. *et al.* Inflammatory mechanism of Rumenitis in dairy cows with subacute ruminal acidosis. *BMC Vet. Res.* **14**, 1–8 (2018).

94. Wickham, S., West, M. B., Cook, P. F. & Hanigan, M. H. Gamma-glutamyl compounds: Substrate specificity of gamma-glutamyl transpeptidase enzymes. *Anal. Biochem.* **414**, 208–214 (2011).
95. Osorio, J. S., Ji, P., Drackley, J. K., Luchini, D. & Loo, J. J. Smartamine M and MetaSmart supplementation during the periparturient period alter hepatic expression of gene networks in 1-carbon metabolism, inflammation, oxidative stress, and the growth hormone-insulin-like growth factor 1 axis pathways. *J. Dairy Sci.* **97**, 7451–7464 (2014).
96. Aquilano, K., Baldelli, S. & Ciriolo, M. R. Glutathione: New roles in redox signalling for an old antioxidant. *Front. Pharmacol.* **5**, 196. <https://doi.org/10.3389/fphar.2014.00196> (2014).
97. Gregorio, K. C. R., Laurindo, C. P. & Machado, U. F. Estrogen and glycemic homeostasis: The fundamental role of nuclear estrogen receptors *esr1/esr2* in glucose transporter *glut4* regulation. *Cells* **10**, 99. <https://doi.org/10.3390/cells10010099> (2021).
98. Efstathiadou, Z. A. *et al.* Associations of estrogen receptor alpha and beta gene polymorphisms with lipid levels and insulin resistance in men. *Metabolism* **64**, 611–617 (2015).
99. Nishihara, K. *et al.* Comparative transcriptome analysis of rumen papillae in suckling and weaned Japanese black calves using RNA sequencing. *J. Anim. Sci.* **96**, 2226–2237 (2018).
100. Javurek, A. B. *et al.* Discovery of a novel seminal fluid microbiome and influence of estrogen receptor alpha genetic status. *Sci. Rep.* **6**, 23027 (2016).
101. Lancaster, P. A. *et al.* Phenotypic and genetic relationships of residual feed intake with performance and ultrasound carcass traits in Brangus heifers. *J. Anim. Sci.* **87**, 3887–3896 (2009).
102. Cánovas, A. *et al.* Multi-tissue omics analyses reveal molecular regulatory networks for puberty in composite beef cattle. *PLoS One* **9**, e102551. <https://doi.org/10.1371/journal.pone.0102551> (2014).
103. Yu, Z. & Morrison, M. Improved extraction of PCR-quality community DNA from digesta and fecal samples. *Biotechniques* **36**, 808–812 (2004).
104. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857. <https://doi.org/10.1038/s41587-019-0209-9> (2019).
105. Seedorf, H., Kittelmann, S., Henderson, G. & Janssen, P. H. RIM-DB: A taxonomic framework for community structure analysis of methanogenic archaea from the rumen and other intestinal environments. *PeerJ* **2**, e494 (2014).
106. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596. <https://doi.org/10.1093/nar/gks1219> (2013).
107. Yilmaz, P. *et al.* The SILVA and ‘all-species Living Tree Project (LTP)’ taxonomic frameworks. *Nucleic Acids Res.* **42**, D643–D648 (2014).
108. Glöckner, F. O. *et al.* 25 years of serving the community with ribosomal RNA gene reference databases and tools. *J. Biotechnol.* **261**, 169–176. <https://doi.org/10.1016/j.jbiotec.2017.06.1198> (2017).
109. Lin, H. & Peddada, S. Analysis of compositions of microbiomes with bias correction. *Nat. Commun.* **11**, 3514 (2020).
110. Rohart, F., Gautier, B., Singh, A. & Lê Cao, K. A. mixOmics: An R package for ‘omics feature selection and multiple data integration. *PLoS Comput. Biol.* **13**, e1005752 (2017).
111. R Core Team. R core team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org> Preprint at (2021).
112. Kaimal, V., Bardes, E. E., Tabar, S. C., Jegga, A. G. & Aronow, B. J. ToppCluster: A multiple gene list feature analyzer for comparative enrichment clustering and network-based dissection of biological systems. *Nucleic Acids Res.* **38**, W98–W102 (2010).
113. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. ClusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).

## Acknowledgements

The authors acknowledge financial support from the Ontario Ministry of Agriculture, Food, and Rural Affairs (OMAFRA), Ontario Ministry of Research and Innovation, and the Ontario Agri-Food Innovation Alliance. This study was also supported by the Beef Farmers of Ontario, the Sustainable Beef and Forage Science Cluster funded by the Canadian Beef Cattle Check-Off, Beef Cattle Research Council (BCRC), Alberta Beef Producers, Alberta Cattle Feeders’ Association, Beef Farmers of Ontario, La Fédération des Producteurs de bovins du Québec, and Agriculture and Agri-Food Canada’s Canadian Agricultural Partnership.

## Author contributions

P.A.S.F.: Formal analysis, Methodology, Writing—original draft. S.L.: Formal analysis, Writing—original draft. C.Y.: Investigation, Writing—review and editing. S.M.W.: Investigation, Writing—review and editing. L.L.G.: Conceptualization, Investigation, Writing—review and editing. A.C.: Conceptualization, Methodology, Investigation, Writing—review and editing.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-43097-8>.

**Correspondence** and requests for materials should be addressed to A.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023