Italian Journal of Animal Science

ISSN: 1828-051X (Online) Journal homepage: www.tandfonline.com/journals/tjas20

Milk production efficiency: candidate genes, digestibility and methane emission for late lactation in a sample of Italian Holstein cows

Francesca Bernini, Andrea Delledonne, Michela Frigerio , Silvia Grossi, Matteo Dell'Anno, Chiara Punturiero, Matteo Santoru, Luciana Rossi, C. A. Sgoifo Rossi, Alessandro Bagnato & Maria G. Strillacci

To cite this article: Francesca Bernini, Andrea Delledonne, Michela Frigerio , Silvia Grossi, Matteo Dell'Anno, Chiara Punturiero, Matteo Santoru, Luciana Rossi, C. A. Sgoifo Rossi, Alessandro Bagnato & Maria G. Strillacci (2026) Milk production efficiency: candidate genes, digestibility and methane emission for late lactation in a sample of Italian Holstein cows, Italian Journal of Animal Science, 25:1, 509-524, DOI: 10.1080/1828051X.2026.2647475

To link to this article: https://doi.org/10.1080/1828051X.2026.2647475

ITALIAN JOURNAL OF ANIMAL SCIENCE
2026, VOL. 25, NO. 1, 509–524
https://doi.org/10.1080/1828051X.2026.2647475

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Milk production efficiency: candidate genes, digestibility and methane emission for late lactation in a sample of Italian Holstein cows

Francesca Bernini[a], Andrea Delledonne[a], Michela Frigerio[a], Silvia Grossi[a], Matteo Dell'Anno[a,b], Chiara Punturiero[a], Matteo Santoru[a], Luciana Rossi[a], C. A. Sgoifo Rossi[a], Alessandro Bagnato[a] and Maria G. Strillacci[a]

[a]Dipartimento di Medicina Veterinaria e Scienze Animali (DIVAS), Università degli Studi di Milano, Lodi, Italia; [b]Dipartimento di Scienze Veterinarie, Università degli Studi di Messina, Messina, Italia

ABSTRACT

The aim of this study was to investigate physiological and genomic differences between high- (HEFF) and low-efficiency (LEFF) Holstein cows, using a residual-based definition of productive efficiency derived from automatic milking systems (AMS) data. Individual efficiency was quantified as the deviation from expected daily milk yield, estimated using a linear mixed-effects model applied to AMS records. This approach enabled the identification of 68 HEFF and 68 LEFF cows within a single herd managed under uniform feeding conditions. A genome-wide association study (GWAS) based on selective genotyping using the GGP Bovine 100k SNP array identified 28 significant SNPs across 13 autosomes, mapping to 51 candidate genes. These genes included loci associated with lipid metabolism and energy utilisation (*CD36, LPGAT1*), methane-related efficiency (*VPS13B*), and immune regulation (*TLR9*), suggesting a complex genetic architecture underlying productive efficiency. To support the biological relevance of the residual-based classification, zootechnical parameters and apparent total tract digestibility (aTTD) were evaluated over three consecutive days in a subset of 15 HEFF and 15 LEFF cows during late lactation. HEFF cows produced more milk than LEFF (33.76 vs. 28.19 kg/d; $p < 0.05$) and emitted significantly less methane per litre of milk (14.40 vs. 17.20 g/L; $p = 0.02$), despite similar total methane output. Furthermore, HEFF cows exhibited higher fibre digestibility, particularly for neutral detergent fibre (aTTD NDF: 57.62 vs. 48.09%; $p = 0.003$). Overall, efficiency defined by model residuals corresponds to measurable biological differences in digestibility and methane emission intensity, supporting the use of residual-based phenotypes as practical indicators of productive efficiency in dairy cattle.

HIGHLIGHTS

- The residual-based model identified low and high-efficiency dairy cows using AMS data.
- GWAS revealed 28 significant SNPs and 51 genes related to nutrient metabolism, CH_4 efficiency, immune response, and reproduction.
- Efficient cows showed higher fibre and protein digestibility and lower methane emissions intensity.

ARTICLE HISTORY
Received 6 November 2025
Revised 11 March 2026
Accepted 15 March 2026

KEYWORDS
Production efficiency; GWAS; selective genotyping; methane emissions

Introduction

Improving the productive efficiency of dairy cows is a critical goal for the global livestock sector, driven by the dual need to enhance food security and reduce the environmental footprint of animal production (Gerber et al. 2013; National Academies of Sciences and Medicine 2021).

Holstein cows are globally recognised for their high milk yield (MY) (Brito et al. 2021) and widespread distribution, and are the leading breed in the U.S., Europe, and several other regions. According to data extracted from the 2023 World Annual Statistics Report (World Holstein Friesian Federation – https://whff.info/annual-statistics/), the global population of milk-recorded cows was 13.96 million, producing 145 million tons of milk, with an average yield of 10,460 kg per cow. In terms of milk production, the most important countries were the USA (2.8 million cows with

CONTACT Francesca Bernini francesca.bernini@unimi.it

Supplemental data for this article can be accessed online at https://doi.org/10.1080/1828051X.2026.2647475.

12,775 kg per cow), Germany (2 million cows with 10,056 kg per cow), the Netherlands (1.2 million cows with 9505 kg per cow), France (1.2 million cows with 9563 kg per cow), and Italy (1.1 million cows with 10,460 kg per cow).

Productive efficiency reflects the ability to convert feed into milk efficiently, and by definition, can be described as the level of an animal's output relative to its input over its productive life or cycle (Berry and Crowley 2013). An efficient animal is capable of producing the same amount of milk with less feed. This characteristic is influenced by digestibility, which plays a pivotal role in the process and affects long-term feed efficiency, milk production, and nutrient excretion (Lee and Hristov 2013; Min and Solaiman 2015). However, conventional digestibility measurements based on total faecal collection, although accurate, are laborious and unsuitable for routine farm application (Titgemeyer et al. 2001). Rapid alternatives such as near-infrared spectroscopy (NIRs) offer timely on-farm estimates but require validation against laboratory standards (Benedeti et al. 2019).

Production efficiency is also linked to environmental sustainability, as more efficient animals generally exhibit lower methane emission intensity (g of CH_4/L of milk) (Capper and Cady 2020). Although direct CH_4 quantification under farm conditions is difficult, predictive models based on production and diet characteristics remain valuable for exploring efficiency–emission relationships (Niu et al. 2018). Indicators such as residual feed intake (RFI) and residual methane emissions (RME) have been widely used to study these aspects (Connor et al. 2013; Salamone et al. 2024). Interestingly, substantial individual variation in production efficiency exists within the breed (Pryce et al. 2014), thus effective genetic programs can be implemented.

The diffusion of automatic milking systems (AMS) has enabled continuous recording of productive and behavioural parameters, generating large datasets that support the identification of individual efficiency differences (Halachmi et al. 2019). Several methods exist for modelling expected individual daily milk production, thereby enabling the identification of individual efficiency. These include parametric approaches such as classical linear and nonlinear functions (e.g. Wilmink, Wood, Ali lactation curve models) and random regression models fitted with Legendre polynomials. Additionally, non-parametric approaches such as various spline-based methods offer flexible alternatives that do not assume a predetermined functional form. Mixed models are well suited for analysing these data, accounting for repeated measures, animal-level random effects, and the genetic relationship among individuals (Strabel and Misztal 1999; Tullo et al. 2014).

The use of deviations from expected milk production has been increasingly adopted to characterise individual differences in efficiency, stability, or resilience in dairy cows. For example, Poppe et al. (2021) derived resilience indicators from daily milk yield records by modelling expected lactation curves and analysing the variability and temporal structure of residuals, highlighting the potential of high-frequency AMS data to capture individual responses to environmental perturbations.

Related approaches have also been used within genetic evaluation frameworks, where model-based predictions and residual components contribute to the assessment of genetic merit and long-term production trends (Guinan et al. 2023). In this context, although the phenotypes considered in the present study are based on proxies rather than direct measurements, they provide a pragmatic framework to explore the integration of high-frequency AMS data with genomic information and targeted physiological indicators of efficiency.

Simultaneously, the availability of dense genotype data for individuals has enabled genomic selection: a revolution in dairy breeding that has accelerated genetic progress due to more precise estimation of breeding values, especially because they can be estimated for young animals that lack phenotypic information (VanRaden et al. 2009; Meuwissen et al. 2016; Brito et al. 2020). Dense genotype data have allowed even finer QTL detection through genome-wide association analysis (GWAS), which in combination with AMS-derived phenotypic data offers a powerful approach to identify quantitative trait loci (QTL) associated with productive efficiency and CH_4 emissions (Liu et al. 2020). GWAS for QTL identification requires both phenotypic data and genomic markers. Statistical power depends primarily on three factors: the number of individuals analysed, the effect size of variants, and the significance threshold applied (Sahana et al. 2023). When working with small populations, whole-population analysis may lack adequate statistical power; in such cases, selective genotyping offers a cost-effective alternative (Darvasi and Soller 1994). Selective genotyping effectively detects associations with quantitative traits when appropriately parameterised (Van Gestel et al. 2000). However, its application in GWAS is most suitable when paired with candidate gene-level significance thresholds ($p < 0.01$) rather than stringent

genome-wide thresholds ($p < 5 \times 10^{-8}$) (Xing and Xing 2009).

In this context, the primary objectives were twofold: 1) to evaluate a simplified approach for deriving production-efficiency phenotypes; 2) to use these derived phenotypes to identify associated QTL in a small cohort of Holstein cows by integrating genomic data. These primary objectives were complemented by secondary goals. In particular, digestibility estimates were used to validate the production-efficiency indicators derived from AMS data and to estimate CH_4 emissions. Field-based instrumentation was combined with laboratory analysis to validate the digestibility estimates. Although this study relies on phenotypic proxies rather than direct measurements, it contributes to improving our understanding of the relationships among the investigated measures.

Material and methods

Ethic statement

The study was approved by the Animal Welfare Organisation (OPBA, University of Milan; Protocol number 67_2023) and conducted according to EU and Italian animal welfare regulations (Directive 2010/63/EU; D. Lgs. 26/2014).

Data processing and filtering criteria

This study was based on a dataset comprising 657,429 daily milk yield (MY) records collected between 2017 and 2024 from 1,086 Holstein cows, recorded using nine AMS units and monitoring collars on a single commercial farm located in Lombardy, Italy. Along with daily milk yield, the variables recorded by the AMS included days in milk (DIM), date, somatic cell count (SCC), and fat, protein, and lactose contents (%). Additionally, rumination time (minutes) and feeding time (minutes) were recorded by accelerometers attached to the collars; these data were subsequently uploaded to the farm's digital system as daily aggregated values and were downloaded together with the other variables.

For the analysis, we used information about DIM, daily milk yield, daily rumination and feeding time. To ensure data consistency and robustness of the analysis, a multi-step quality control and filtering procedure was applied. As a first step, data editing was performed to remove records with missing values in key variables or with a daily milk yield of zero. After this initial preprocessing step, 575,759 daily records remained. Subsequently, biological and analytical filtering criteria were applied:

- Lactation number: only cows in second to fifth lactation were included to reduce variability associated with physiological immaturity and adaptation to AMS. First-lactation cows were excluded due to their transitional physiological status and learning phase associated with AMS, which could bias efficiency assessments (Connor et al. 2019; Siewert et al. 2019).
- Lactation completeness: only complete lactations were retained, defined as those with a duration between 200 and 305 DIM, consistent with the farm's average lactation length. The cut-off value was also set at 305 DIM since it is considered a traditional and defined length for standardising lactation records in dairy cows (Santschi et al. 2011; Kok et al. 2016).

After applying all filtering criteria, the final dataset used for model fitting consisted of 206,334 daily MY records from 453 Holstein cows.

Statistical model for efficiency classification

Data were analysed using a mixed-effects model implemented in JMP® 19 software (2025 JMP® JMP Statistical Discovery LLC, Cary, NC) to evaluate MY:

$$Y_{i,t,k,l,m,n,o} = \beta_0 + \gamma_{k,l} + \delta_m + \beta_1 \cdot timefeed_{i,t} + \theta_n + \iota_o + \beta_2 \cdot timerum_{i,t} + \beta_3 \cdot AVGmilkings_i + u_i + \varepsilon_{i,t,k,l,m,n,o}$$

where $Y_{i,t,k,l,m,n,o}$: daily MY for cow i recorded at the level of factors specified by the subscripts, β_0: overall intercept, $\gamma_{k,l}$: fixed effect of the year-month interaction (level k for year, level l for month), δ_m: fixed effect of DIM category (level m), β_1: regression coefficient for feeding time, $timefeed_{i,t}$: feeding time (min/day) for cow i on day t, θ_n: fixed effect of parity order (level n), ι_o: fixed effect of age at calving in months (level o), β_2: regression coefficient for rumination time, $timerum_{i,t}$: rumination time (min/day) for cow i on day t, β_3: regression coefficient for average number of daily milkings, $AVGmilkings_i$: average number of daily milkings for cow i, u_i: random effect associated with cow i, assumed $u_i \sim \mathbb{N}(0, \sigma^2_u)$ and $\varepsilon_{i,t,k,l,m,n,o}$: residual error term, assumed $\varepsilon_{i,t,k,l,m,n,o} \sim \mathbb{N}(0, \sigma\_e\text{^}2)$.

DIM was included as a categorical fixed effect to describe the average lactation trajectory, using thirty classes defined as one initial class covering DIM 5–14

followed by consecutive 10–day intervals. This approach allowed a flexible representation of lactation dynamics without imposing a predefined parametric curve. The distribution of observations across DIM classes was evaluated to ensure adequate support for model estimation. Across the 30 DIM classes, the number of observations per class averaged 6878 (median = 7284), ranging from 4613 to 7352 records, indicating a generally balanced representation across the lactation trajectory. A detailed summary of the number of observations per DIM class is reported in Table S1.

Parity order and age at calving were included as categorical fixed effects. Feeding time (min/day), rumination time (min/day), and average number of daily milkings were included as continuous covariates. Feeding and rumination times were incorporated to account for short-term behavioural and intake-related variation affecting daily milk yield, thereby improving the estimation of marginal residuals used for cow efficiency classification (Paudyal 2021; Nascimento et al. 2024).

Daily milk yield residuals were computed as the difference between observed values and model predictions based on fixed effects only (marginal predictions). Residuals were then averaged within each lactation. To ensure robustness of the efficiency phenotype, only cows showing consistently positive or negative mean residuals across all available lactations were retained and classified as high-efficiency (HEFF) or low-efficiency (LEFF). The cow random effect was included in the model to account for repeated measurements, but it was not used to derive the efficiency phenotype, which was based exclusively on marginal residuals in order to preserve persistent between-cow differences. Daily milk yield residuals were evaluated using standard diagnostic procedures, including inspection of their distribution, Q–Q plots, and plots of residuals against predicted values, days in milk, and parity.

Genotype-phenotype association study

To investigate the genetic basis underlying productive efficiency, a selective genotyping strategy combined with a DNA pooling–like approach was applied (Darvasi and Soller 1994). This method enables the detection of QTL by focusing genotyping efforts on animals with extreme phenotypic values and relies on contrasts in allele frequencies between extreme groups, thereby reducing individual-level noise and increasing statistical power and sensitivity to weak genetic signals, particularly in studies with limited sample size (Darvasi and Soller 1994; Huang et al. 2010; Strillacci et al. 2014, 2023; Lipkin et al. 2016, 2024).

In detail, a total of 68 HEFF and 68 LEFF cows, corresponding to the top and bottom 15% of the average residual distribution of 453 cows (resulting after editing), were included in the genotypic analysis. Genotypes of all selected cows, obtained using the GGP Bovine 100K SNP array (GeneSeek®), were available (SNP call rate $\geq$ 95% and minor allele frequency $\geq$ 1%). To implement the experimental design, each group (HEFF and LEFF) was randomly divided into two biological replicates of 34 sample each (HEFF1, HEFF2; LEFF1, and LEFF2), using a random assignment procedure implemented in R. Allele frequencies for each SNP were calculated within each group using the 'genotype statistics by marker' function in SNP & Variation Suite software (SVS v8.9, Golden Helix Inc., Bozeman, MT, USA). The statistical comparison between HEFF and LEFF groups was performed using an in-house R script (R version 4.0.5). Monomorphic markers, which are non-informative, were removed from the dataset, along with SNPs showing the top 5% of absolute allele frequency differences between biological replicates (i.e. HEFF1 vs. HEFF2 and LEFF1 vs. LEFF2). After this filtering, 65,845 autosomal SNPs remained for analysis.

Marker-trait associations were assessed using a single-marker test, in which the difference in mean A allele frequencies between the phenotypic tails was calculated while accounting for the variation in allele frequencies within each tail. This adjustment corrects for within-group variability, providing a more accurate assessment of the significance of allele frequency differences between the extremes. GWAS results were visualised by a Manhattan plot generated with the *qqman* R package (Turner 2014), using both the Bonferroni and FDR multiple testing correction thresholds (set at 0.05).

Gene annotation

All SNPs above the FDR threshold were used for the SNP annotation, performed using the NCBI Genome Data Viewer tool, based on the genomic position of each significant SNP mapped to the ARS-UCD1.2 assembly. When the SNP was located within a gene, that gene was considered a candidate gene. If the SNP did not fall within a gene, a genomic window of ±250kb of the SNP position was explored. The QTL associated with each gene were identified using the Cattle QTL database of AnimalQTLdb (https://www.animalgenome.org/cgi-bin/QTLdb/BT/index).

Evaluation of production performance and CH_4 emissions

To investigate potential physiological mechanisms underlying differences in productive efficiency, faecal and ration analyses were conducted on a subset of 30 animals used in the GWAS (15 HEFF and 15 LEFF), due to the high cost of these laboratory measurements. The selected animals were balanced for lactation number and stage of lactation (DIM 250–290) and were chosen from the high and low phenotypic tails to evaluate whether cows with contrasting production levels, under the same lactation stage and feeding conditions, exhibit differences in digestion and faecal characteristics.

Zootechnical, digestive, and CH_4-related parameters were measured over three consecutive days in a subset of cows during late lactation. This short monitoring period was designed to allow a controlled comparison between efficiency groups at the same physiological stage, rather than to characterise performance across the entire lactation. All animals were maintained under identical environmental conditions and fed the same total mixed ration (TMR), reported in Table S2. In addition, all cows received a concentrate through the AMS, with the quantity tailored to their individual MY. This ensured that the additional energy intake was proportional to the production level, minimising its confounding effect on the production data.

All zootechnical parameters detailed below were collected from the 30 cows over three measurement time points during late lactation. AMS data represented the primary source of information for daily milk yield and behavioural traits. One of the three days coincided with the official milk recording carried out by ARAL (Breeders Association of the Lombardy Region). These official records were used exclusively to obtain milk composition values required for dry matter intake (DMI) and CH_4 estimation, while all production data used in the analyses were derived from the AMS. All statistical comparisons between HEFF and LEFF groups were performed using Student's t-tests. Variables were reported as mean ± standard deviation, and significance was considered at $p < 0.05$. Analyses were conducted in R (v.4.5.1).

Production performance: Milk parameters, body condition scores, body weight and dry matter intake (DMI)

Individual daily MY, rumination time, and feeding time were obtained from the AMS and collar-based sensors for each of the three consecutive measurement days. Milk composition parameters (fat, protein, and lactose percentages) were obtained from both the AMS and the official milk recording carried out by ARAL (Breeders Association of the Lombardy Region), depending on data availability. Specifically, milk composition values from the official milk recording were used for the measurement day coinciding with the control test and were employed in the calculation of apparent total tract digestibility (aTTD), estimation of dry matter intake (DMI) and CH_4 emission. For the remaining two measurement days, milk composition data recorded by the AMS were used for aTTD estimation. For each cow, daily values were averaged across the three measurement days prior to subsequent analyses.

During each measurement period, body condition score (BCS) was assessed using a 5-point scale by two independent scorers and averaged. Body weight (BW) was measured once per measurement period using a weight tape (Animeter tape, Kerbl Austria Handels GmbH, Austria). Although this method provides lower precision than direct weighing with a scale, it represents a practical and commonly used approach under commercial farm conditions. The potential reduction in accuracy associated with this measurement method was therefore considered when interpreting DMI estimates. For the analysis the following equation of De Souza et al. (2019) was used:

$$\begin{aligned}\text{DMI (kg)} = [(3.7 + \text{parity} \times 5.7) + 0.305 \\ \times \text{milkenergy (Mcal)} + 0.022 \times \text{BW(kg)} \\ + (-0.689 + \text{parity} \times -1.87) \times \text{BCS}] \\ \times [1 - (0.212 + \text{parity} \times 0.136) \times e^{(-0.053 \times \text{DIM})}]\end{aligned}$$

In this equation, milk energy (Mcal) represents the net energy output in milk and was calculated based on daily milk yield and milk composition (fat, protein, and lactose), according to standard energy-corrected milk formulations. BW (kg) corresponds to the estimated live weight of the cow, BCS is the body condition score, DIM represents days in milk, and parity indicates lactation number. All input variables required by the equation were derived from either official milk recording data or on-farm measurements collected during the same measurement period.

Feed and faecal sample collection and NIRs analysis

The chemical composition of feed and faecal samples was evaluated on field using a portable Near-infrared spectroscopy (NIRs, Polispec, IT Photonics, Italy). The NIRs was used to measure the content of dry matter

(DM), crude protein (CP), crude fats (CF), neutral detergent fibre (NDF), acid detergent fibre (ADF), acid detergent lignin (ADL), acid insoluble ash (AIA) and starch, expressed as a percentage of DM.

The characteristics of the TMR were evaluated on the fresh feed daily at each time point, measuring ten different sections along the length of both feed bunks and averaging the results. Individual faecal samples were collected from the rectal ampulla of each cow. With the same sampling criteria samples were also collected and stored for the subsequent laboratory analysis.

Bromatological analyses

For bromatological analyses of the same nutrients described for the NIRs procedure, samples were pooled into a composite sample per animal (representing the average of three consecutive days), dried in a forced-air oven at 65 °C, and ground using a centrifugal mill equipped with a 1 mm mesh screen before being processed for laboratory analysis.

The analysis covered both the TMR and faecal samples following the official methodologies established by the Association of Official Analytical Chemists (AOAC 2023) for the following parameters:

- Dry matter (DM): determined by drying at 65 °C until reaching a constant weight, according to AOAC method 930.15.
- Crude protein (CP): quantified using the Kjeldahl method (AOAC 2001.11), which includes digestion, ammonia distillation, and titration for total nitrogen and protein content determination using the nitrogen-to-protein conversion coefficient of 6.25.
- Ether extract (EE): determined by Randall-Soxtec extraction (AOAC 2003.5), involving petroleum ether immersion, solvent evaporation, and lipid residue weighing.
- Starch: analysed using a commercial colorimetric kit following the rapid total starch method, which includes enzymatic hydrolysis and a colorimetric reaction registered spectrophotometrically, performed according to the manufacturer's instructions (Total Starch Assay Kit, Megazyme Ltd, Bray, Ireland).
- Fibre fractions: Including neutral detergent fibre (NDF), acid detergent fibre (ADF), and acid detergent lignin (ADL), analysed according to the method developed by van Soest et al. (1991).
- Crude ash: determined by incinerating samples in a muffle furnace at 550 °C for 3 hours, according to AOAC method 942.05, to quantify the total inorganic fraction.
- Acid-Insoluble ash (AIA): determined by incinerating the ash and treating it with 2N HCl, following the method described by Liu (2022).

All analyses were performed in technical duplicate to ensure accuracy and reproducibility of results.

Apparent total tract digestibility (aTTD) evaluation

The aTTD coefficients were calculated for each nutrient (DM, CP, NDF, ADF, starch and EE) using the following formula, as introduced by Merchen (1988) and applied by Jancewicz et al. (2016):

$$\mathrm{aTTD} = \left\{\left[\mathrm{DC}_D/(\mathrm{ADL}_D - \mathrm{AIA}_D)\right] - \left[\mathrm{CD}_F/(\mathrm{ADL}_F - \mathrm{AIA}_F)\right]\right\}/\left[\left(\mathrm{DC}_D/(\mathrm{ADL}_D - \mathrm{AIA}_D)\right] * 100\right.$$

where DC_D and DC_F represent the target nutrient concentration in the diet and faeces respectively. ADL_D and ADL_F represent the ADL concentration in the diet and faeces. AIA_D and AIA_F represent the AIA concentration in the diet and faeces.

CH_4 emissions and CH_4 emission intensity

Similarly, the daily CH_4 production was then estimated using the equation proposed by Niu et al. (2018):

$$\mathrm{CH_4(g/d)} = -72.4 + 3.15 \times \mathrm{NDF(\%DM)} + 2.65 \times \mathrm{ECM} + 23.9 \times \mathrm{milk\ fat\ (\%)} + 0.290 \times \mathrm{BW\ (kg)}$$

where NDF is the neutral detergent fibre of the feed as analysed by the NIRs system, DM is the dry matter of the feed as analysed by the NIRs system, ECM is the energy corrected milk calculated using the formula of Tyrrell and Reid (1965), and BW is the average body weight collected during the 3 days period. All variables required for the equation were sourced from official milk recording data or on-farm collected measurements. CH_4 intensity (g/kg of milk) was also computed by relating the daily CH_4 production to daily milk production.

Results

Efficiency classification of the animals

The robustness of the residual-based classification was further supported by diagnostic analyses of the mixed model. The distribution of daily milk yield residuals was approximately symmetric around zero and the

Table 1. Descriptive statistics for the two efficiency groups obtained from Automatic Milking Sistems (AMS) data, N° is the number of animals considered in the group.

Group	N°	RES (kg) ***	MY (kg) ***	Time_feed (min)	Time_rum (min)*
HEFF	68	7.26 ± 2.18	51.30 ± 3.79	237.51 ± 42.70	560.11 ± 27.41
LEFF	68	−7.66 ± 1.13	39.11 ± 4.39	245.10 ± 35.50	546.49 ± 36.46
p-value	–	9.548895E-68	1.131746E-35	0.2619109	0.01517791

Note: Statistical significance levels: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Average values and standard deviation of residuals (res), average production (MY), feeding time (time_feed) and rumination time (time_rum) of the two efficiency groups across all lactations. HEFF - high efficiency and LEFF - low efficiency.

Q–Q plot showed that residuals closely followed the expected theoretical normal distribution (Figure S1A). When residuals were plotted against days in milk, no clear systematic trend was observed across the lactation curve, indicating that the model adequately captured the main lactation-stage effects. Similarly, the distribution of residuals across parity classes was comparable, suggesting that parity-related variation was appropriately accounted for in the model (Figure S1B). Finally, the plot of residuals against predicted milk yield did not reveal any major structure or funnel-shaped pattern (Figure S1C), indicating no strong evidence of heteroscedasticity. Although the Spearman correlation between residuals and days in milk was statistically significant because of the very large number of observations, its magnitude was negligible ($\rho = 0.025$; $p < 2.2 \times 10^{-16}$), indicating the absence of a biologically meaningful association. These results support the use of average residuals as a proxy for persistent production efficiency and confirm that the identification of HEFF and LEFF cows was not driven by lactation stage or parity-related artefacts.

Descriptive statistics for the two efficiency groups (68 HEFF and 68 LEFF) are presented in Table 1. As expected, HEFF cows exhibited higher average MY across all available lactations (51.30 kg) and residual values (7.26 ± 2.18) compared to LEFF cows, confirming their superior productive performance. No significant differences were observed in feeding time, whereas for the rumination time, HEFF animals ruminated more (560.11 min/d) compared to the LEFF (546.49 min/d; $p = 0.015$). Dry matter intake (DMI) was not included in this section because individual intake estimates were available only for the subset of animals involved in the physiological validation experiment (see Table 2), where additional measurements related to digestibility and methane emissions were collected.

Genotype-phenotype association study

The results of the GWAS comparing HEFF and LEFF cows are shown in Figure 1. The Manhattan plot highlights the distribution of $-\log10(p)$ across all bovine autosomes. A total of 28 significant SNPs were identified across 13 autosomes. These SNPs lead to the identification of 51 positional candidate genes (Table 3). SNPs located within coding or intronic regions are reported under the 'intragenic' column, while nearby genes within ±250 kb of intergenic SNPs are listed as 'upstream' or 'downstream'. The full list of associated QTL for each candidate gene, based on the Animal QTLdb, is provided in Table S3.

Table 2. Average residuals of the sampled cows (res) milk yield (MY), milk quality parameters (fat, protein, lactose), rumination (time_rum), feeding time (time_feed), estimated DMI (kg) and CH_4 emissions (g/d) in HEFF and LEFF cows.

Class	HEFF	LEFF	*p*-value
Res	5.79 ± 2.67	−4.96 ± 2.74	0.0000000000146
MY	33.76 ± 6.08	28.19 ± 5.16	0.0115
fat	1.37 ± 0.16	1.18 ± 0.18	0.00391
protein	3.59 ± 0.14	3.54 ± 0.19	0.412
lactose	4.96 ± 0.11	4.98 ± 0.08	0.571
Time_rum	528.27 ± 65.16	503.21 ± 57.93	0.350
Time_feed	184.00 ± 56.14	214.90 ± 54.93	0.200
DMI	29.13 ± 2.08	28.71 ± 2.54	0.623
CH4_tot (NIRs)	470.11 ± 28.04	470.65 ± 32.03	0.965
CH4_liter_milk (NIRs)	14.40 ± 2.96	17.20 ± 3.23	0.0196
CH4_tot (Lab)	495.86 ± 30.49	504.82 ± 40.07	0.496
CH4_liter_milk (Lab)	15.19 ± 3.14	18.46 ± 3.61	0.0134

Values are reported as mean ± SD. residuals (res); average production (MY); feeding time (time_feed); rumination time (time_rum) and dry matter intake (DMI).

TMR and faecal composition

Table 4 reports the compositional analysis of the TMR and faeces collected from HEFF ($n = 15$) and LEFF ($n = 15$) cows, based on both NIRs and laboratory data. Faecal composition estimated by NIRs differed markedly between efficiency groups. In general, HEFF cows exhibited a significantly lower concentration of fibrous components in faeces compared with LEFF cows. In particular, NDF and ADF were reduced in HEFF animals (NDF: 39.76 ± 5.59 vs. 47.40% ± 6.83%; $p = 0.002$; ADF: 23.18 ± 3.70 vs. 28.25% ± 4.72%; $p = 0.003$), together with ADL (9.33 ± 0.39 vs. 9.81% ± 0.41%; $p = 0.003$) and AIA (1.64 ± 0.13 vs. 1.86% ± 0.19%; $p = 0.001$). These results indicate a lower faecal excretion of structural carbohydrates in HEFF cows, consistent with improved fibre utilisation.

Conversely, DM was significantly higher in HEFF cows compared with LEFF cows (14.68 ± 1.00 vs.

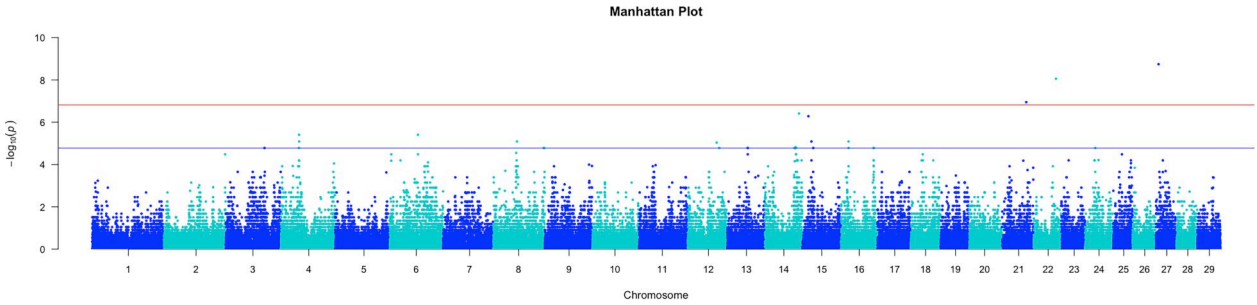

Figure 1. Manhattan plot of the genome-wide association results comparing HEFF (high efficiency) and LEFF (low efficiency) cows. The red line shows the Bonferroni-corrected significance threshold, while the blue is the false discovery rate (FDR) threshold, both set at a significance level of 0.05.

13.53% ± 1.10%; $p = 0.006$). Faecal starch concentration was also higher in HEFF cows (1.50 ± 0.67 vs. 0.69% ± 0.71%; $p = 0.003$), whereas no significant differences were detected for CP or EE. Ash content was lower in HEFF animals (14.03 ± 0.75 vs. 14.78% ± 0.69%; $p = 0.008$). These results could indicate a slightly lower digestibility of fibrous fractions in the LEFF group. However, these patterns were not confirmed by the laboratory data, which showed no significant group-level differences for any of the measured parameters.

Production performance and CH_4 emissions

Table 2 compares the production performances, DMI and CH_4 emissions of the two sampled groups. In terms of production performance, groups are comparable with the larger sample used in the GWAS, exhibiting statistically significant differences in both residuals and MY. Some production performance traits and certain emission-related traits differed between efficiency groups. As expected by construction, HEFF cows showed significantly higher mean residual values than LEFF cows (5.79 ± 2.67 vs. −4.96 ± 2.74; $p < 0.001$), confirming the effectiveness of the residual-based classification. HEFF cows also exhibited higher average daily MY (33.76 ± 6.08 vs. 28.19 ± 5.16 kg/d; $p = 0.011$) and greater fat yield (1.37 ± 0.16 vs. 1.18 ± 0.18 kg/d; $p = 0.004$).

No significant differences were observed between groups for milk protein and lactose content, nor for total CH_4 emissions. However, CH_4 intensity expressed per litre of milk was significantly lower in HEFF cows (14.40 ± 2.96 vs. 17.20 ± 3.23 g CH_4/L; $p = 0.020$), indicating a more favourable conversion of feed into milk relative to methane output. The use of body weight estimated *via* weight tape rather than direct weighing may have slightly reduced the precision of CH_4 estimates; however, this limitation is expected to affect absolute values rather than relative differences between efficiency groups.

aTTD analysis

To investigate potential physiological differences in nutrient utilisation between HEFF and LEFF cows, the aTTD was estimated using both NIRs and laboratory analyses. The assessed feed components included DM, CP, EE, NDF, ADF and starch. The results for both techniques are summarised in Table 5, aTTD values differed between efficiency groups and between analytical approaches. Based on NIRs-derived estimates, HEFF cows exhibited significantly higher digestibility of CP (62.41 ± 3.33 vs. 59.53% ± 3.63%; $p = 0.032$), aNDF (57.62 ± 6.53 vs. 48.09% ± 8.93%; $p = 0.003$), and ADF (65.88% ± 6.70 vs. 56.12% ± 9.08%; $p = 0.003$) compared with LEFF cows. Digestibility of starch was slightly lower in HEFF cows (98.04 ± 0.88 vs. 99.05% ± 0.95%; $p = 0.005$), whereas no significant differences were observed for DM digestibility or lipids.

Laboratory-based aTTD estimates confirmed higher digestibility of several components in HEFF cows, albeit with greater variability. In particular, HEFF cows showed higher digestibility of DM (86.25 ± 2.21 vs. 84.19% ± 2.79%; $p = 0.034$), CP (61.41 ± 6.50 vs. 56.12% ± 7.05%; $p = 0.042$), and ADF (31.52 ± 8.34 vs. 19.12% ± 13.56%; $p = 0.006$). Differences in aNDF digestibility were not significant when estimated from laboratory data, while starch and lipids digestibility showed only numerical trends favouring HEFF cows.

Discussion

The classification of cows into HEFF and LEFF groups was based on individual residuals from a linear mixed-effects model that accounted for key factors influencing daily MY, including DIM, parity, feeding and rumination times, age, and milking frequency. The lack of systematic patterns in daily residuals across days in milk and parity supports the robustness of the mixed model and suggests that the residual-based classification of cows was not driven by temporal or parity-

Table 3. Summary of significant SNPs from the GWAS and gene annotation.

SNP	Chr	bp	p-value	Intragenic	Upstream	Downstream
Hapmap41493-BTA-28649	3	85077748	1.66E-05			
BovineHD0400011183	4	40115485	1.66E-05	SEMA3C		
BovineHD0400011267	4	40474946	8.17E-06		CD36	GNAT3
BovineHD0400011271	4	40481319	3.94E-06		CD36	GNAT3
BTA-111059–no-rs	6	62338657	3.94E-06	BNIP3		
ARS-BFGL-NGS-72492	8	51919977	8.17E-06	PCSK5		
ARS-BFGL-NGS-102553	8	111057264	1.66E-05	TRAPPC12		
ARS-BFGL-NGS-41309	8	111672008	1.66E-05	MYT1L		
BovineHD1200017519	12	63497662	9.16E-06			
BovineHD1200019122	12	69207619	1.66E-05		DCT, TGDS, GPR180, SOX21	
BTB-01837541	13	45320699	1.66E-05		PITRM1, PFKP	
BovineHD1400018787	14	64990713	1.66E-05	VPS13B		
BovineHD1400018792	14	65004906	1.66E-05	VPS13B		
BovineHD1400019527	14	67276670	1.66E-05	CPQ		
BovineHD1400019537	14	67309823	1.66E-05	CPQ		
BovineHD1400019548	14	67341621	1.49E-05	CPQ		
BovineHD1400021634	14	74788465	3.86E-07		MMP16	
BovineHD1500003342	15	13163097	5.21E-07			
BovineHD1500005099	15	19885618	8.17E-06			
BovineHD1500006313	15	23874724	1.66E-05	TTCL12		
BovineHD1600004670	16	16603900	8.17E-06			
Hapmap51075-BTA-86900	16	16665275	1.66E-05			
ARS-BFGL-NGS-36880	16	71839497	1.66E-05		LPGAT1, NEK2,	SLC30A1, RD3, TRAF5, RCOR3
ARS-BFGL-NGS-106233	16	72025137	1.66E-05	TRAF5		
BovineHD2100015369	21	53223744	1.13E-07			
BovineHD2200014147	22	48873180	8.78E-09		WDR82, MIRLET7G, PPM1M, TWF2, TLR9, ALAS1, POC1A, DUSP7	RPL29, ACY1, ABHD14A, ABHD14B, PCBP4, GPR62, PARP3, RRP9, IQCF1, IQCF5, IQCF2, IQCF6, GRM2
BovineHD2400006157	24	22352444	1.66E-05	DTNA		
ARS-BFGL-NGS-55443	27	6126003	1.81E-09		XKR5, TAP, SPAG11B, ZNF705A	LAP, DEFB13

Note: The 'intragenic' column indicates genes where the SNP falls within a gene boundary; the 'upstream' and 'downstream' columns list genes within ±250 kb of the non-intragenic SNPs position.

related artefacts but rather reflects persistent deviations from expected production. This residual-based approach follows the same rationale as RFI, which quantifies efficiency as the deviation between observed and expected feed intake based on energy sinks like MY and metabolic body weight (Berry and Crowley 2013; Olijhoek et al. 2020). In our case, the "residual yield" metric identifies animals that consistently produce more or less milk than expected (given their physiological and environmental context), thus offering a practical proxy for productive efficiency. While RFI requires direct feed intake data, our approach exploits high-frequency data collected *via* AMS and sensors, that are more feasible in commercial settings and aligned with the goals of precision livestock farming. Although this residual-based phenotype provides a practical proxy for productive efficiency, it does not explicitly integrate resource use and productive output into a single composite metric. In particular, direct measures of feed intake, energy balance, or environmental outputs were not available for the entire population and therefore could not be included in the efficiency definition. Future studies should aim to develop composite efficiency indices that jointly account for production, resource utilisation, and environmental emissions, allowing a more comprehensive evaluation of productive efficiency in dairy cattle.

The use of model residuals as efficiency indicators has gained momentum, particularly in systems equipped with sensor-based phenotyping (Salamone et al. 2024). Integrated phenotyping platforms combining automated milking systems and sensor technologies allow continuous and high-frequency recording of behavioural and productive traits, generating large longitudinal datasets that improve the estimation of expected performance and the reliability of residual-based phenotypes. Residuals allow the consistent identification of outperforming or underperforming cows across lactations, increasing the reliability of subsequent genomic analyses. To minimise bias from transient physiological states, we excluded animals in early lactation or in first parity, when energy imbalance, growth, and adaptation to AMS could distort performance metrics (Connor et al. 2019; Siewert et al. 2019). This ensures that the observed deviations reflect true differences in efficiency rather than temporary effects.

Genotype-phenotype association study

The selective genotyping approach used, proved to be effective for identifying genomic regions underlying productive and health traits, in line with previous studies in cattle (Lipkin et al. 1998; Huang et al. 2010; Strillacci et al. 2014; Kurz et al. 2019). In this study, 28 significant SNPs distributed across 13 BTAs, encompassing 51 genes. These genes represent a diverse set of functional candidates that converge on key biological processes, which can be broadly grouped into categories related to digestive physiology, metabolic and endocrine regulation, immune function, and reproductive biology.

Nutrient metabolism, CH_4 and feed efficiency

Comparison of HEFF and LEFF dairy cows revealed genomic regions containing genes involved in energy extraction, nutrient digestion, and utilisation, processes closely linked to DMI and CH_4 emissions. Genes such as *LPGAT1* and *CD36* play central roles in lipid metabolism, potentially contributing to differences in feed efficiency. *LPGAT1* influences lipid metabolism, energy partitioning, body fat deposition, metabolic efficiency, and longevity (Getaneh and Alemayehu 2022). *CD36*, a transmembrane receptor in mammary epithelial cells, mediates long-chain fatty acid uptake and delivery to intracellular pathways for milk fat synthesis, and its high expression in lactating cows suggests a role in improved lipid utilisation, milk yield, and production efficiency in HEFF cows (Tian et al. 2022). *ALAS1* has been identified as one of the most significantly differentially expressed genes between high- and low-residual feed intake (RFI) Holstein–Friesian cows, suggesting a role in metabolic efficiency (Higgins et al. 2019). *PFKP*, previously identified in association with ruminal bacterial family phenotypes (Golder et al. 2023), highlights a potential role in modulating carbohydrate metabolism in the rumen, as it is a key regulatory enzyme of glycolysis. Variation at this locus may influence substrate availability for microbial fermentation, thereby indirectly affecting feed efficiency and nutrient utilisation in dairy cattle, consistent with Golder et al., who showed that the mammalian genome interacts with the rumen metabolome and microbial communities, ultimately impacting milk protein composition.

Among the positional candidate genes identified in previous studies on predicted methane emissions in dairy cattle, *VPS13B* was associated with all six distinct CH_4-related traits and has been repeatedly reported in the literature as being linked to milk production and composition traits (Rezvannejad et al. 2022; Salgado Pardo et al. 2022; Fresco et al. 2025). In the present study, we identified two significant SNPs, BovineHD1400018787 and BovineHD1400018792, mapping within *VPS13B*. These results are consistent with prior findings in cattle, goats, and sheep, indicating a conserved role of *VPS13B* in both methane-related traits and milk production efficiency.

The region flanking SNP BovineHD1200019122 overlaps one of the regions (containing *DCT*, *TGDS*, *GPR180*, and *SOX21*) reported by Soares et al. to explain the highest proportion of genetic variance for clinical ketosis (Soares et al. 2021). Given that ketosis is associated with negative energy balance in early lactation, compromising dry matter intake and nutrient partitioning, genetic variants in this region may may be associated with differences in milk yield and overall production efficiency among dairy cows (Yang et al. 2019).

Immune response and production efficiency

Health status and immune competence are increasingly recognised as key components of feed efficiency and productive longevity in dairy cattle (Aleri et al. 2019). Several genes identified in this study are involved in innate immunity and inflammatory regulation, suggesting a link between efficiency and resilience. The lingual antimicrobial peptide (*LAP*), a member of the bovine β-defensin family, has been detected in bovine milk in biologically active forms with antimicrobial activity comparable to human β-defensins. *LAP* expression increases in response to mammary infection and is positively associated with somatic cell count, supporting its role in innate mammary immune defence (Isobe et al. 2009). Tracheal antimicrobial peptide (TAP) exhibits bactericidal activity against pathogens responsible for bovine respiratory disease, and its expression in tracheal epithelial cells increases in response to bacterial components, such as lipopolysaccharide. In a study on bovine β-defensin, the downregulation of LAP and TAP expression in the rumen epithelium following butyrate infusion highlights a close interaction between diet-derived metabolites and host defense peptide expression, supporting a link between nutrition, immune regulation, and microbial homeostasis (Meade et al. 2014). Similarly, *DEFB13*, another β-defensin expressed in the bovine mammary gland, contributes to epithelial antimicrobial protection against a broad range of pathogens (Gurgul et al. 2019). *TLR9* plays a key role in innate immune recognition by sensing bacterial DNA and activating inflammatory responses in the mammary gland. Genetic variation or altered

Table 4. Comparison of TMR (total mixed ration) and faecal composition from NIRs and laboratory data, values are reported as mean ± SD.

	Comp	NIRs	Laboratory		Comp	HEFF	LEFF	*p*-value		HEFF	LEFF	*p*-value
TMR	DM	47.03 ± 0.78	48.93 ± 1.06	Faeces NIRs data	DM	14.68 ± 1.00	13.53 ± 1.10	0.006	Laboratory data	15.87 ± 0.91	16.50 ± 1.19	0.114
	CP	16.43 ± 0.22	14.28 ± 0.19		CP	17.82 ± 0.67	18.01 ± 0.58	0.408		13.34 ± 0.70	13.38 ± 1.09	0.9
	NDF	33.66 ± 0.77	44.27 ± 3.97		NDF	39.76 ± 5.59	47.40 ± 6.83	0.002		64.62 ± 3.25	64.83 ± 4.19	0.884
	ADF	23.87 ± 0.40	22.65 ± 0.42		ADF	23.18 ± 3.70	28.25 ± 4.72	0.003		37.20 ± 3.03	39.11 ± 2.69	0.078
	ADLD	3.45 ± 0.20	5.37 ± 1.01		ADL	9.33 ± 0.39	9.81 ± 0.41	0.003		12.50 ± 2.04	13.30 ± 2.35	0.329
	AIAD	0.54 ± 0.04	0.57 ± 0.12		AIA	1.64 ± 0.13	1.86 ± 0.19	0.001		3.21 ± 0.92	2.63 ± 1.66	0.252
	Starch	25.33 ± 0.97	21.28 ± 1.30		Starch	1.50 ± 0.67	0.69 ± 0.71	0.003		0.82 ± 0.43	0.93 ± 0.32	0.414
	EE	4.29 ± 0.08	4.09 ± 0.15		EE	4.96 ± 0.64	5.23 ± 0.60	0.248		1.90 ± 0.33	1.98 ± 0.43	0.588
	ASH	7.41 ± 0.07	4.72 ± 0.93		ASH	14.03 ± 0.75	14.78 ± 0.69	0.008		10.92 ± 0.94	11.25 ± 1.65	0.519

Note: Comp = Component; DM = dry matter; CP = crude protein; NDF = neutral detergent fibre; ADF; acid detergent fibre; ADLD: acid detergent lignin concentration; AIAD: acid insoluble ash concentration; EE: ether extract; ASH: crude ash; HEFF: high efficiency and LEFF: low efficiency.

Table 5. Estimated apparent total tract digestibility (aTTD) values (%) from NIRs and laboratory data for HEFF and LEFF groups for: dry matter (DM), crude protein (CP), neutral detergent fibre (NDF), acid detergent fibre (ADF), starch and ether extract (EE); values are reported as mean ± SD.

NIRs data						
Group	DM	CP	aNDF	ADF	Starch	EE
HEFF	88.95 ± 1.09	62.41 ± 3.33	57.62 ± 6.53	65.88 ± 6.70	98.04 ± 0.88	59.97 ± 6.44
LEFF	89.44 ± 0.77	59.53 ± 3.63	48.09 ± 8.93	56.12 ± 9.08	99.05 ± 0.95	54.83 ± 7.38
p-value	0.170	0.032	0.003	0.003	0.005	0.052
Laboratory data						
Group	DM	CP	NDF	ADF	Starch	EE
HEFF	86.25 ± 2.21	61.41 ± 6.50	34.39 ± 10.96	31.52 ± 8.34	98.45 ± 0.86	80.90 ± 3.44
LEFF	84.19 ± 2.79	56.12 ± 7.05	31.49 ± 11.49	19.12 ± 13.56	97.91 ± 0.84	77.08 ± 7.18
p-value	0.034	0.042	0.486	0.006	0.091	0.078

regulation of *TLR9* has been associated with differences in immune responsiveness and susceptibility to mastitis, influencing udder health and production efficiency (Badami et al. 2019; Lesta et al. 2025).

The QTL region identified by Badia-Bringué et al. annotated with *RD3, TRAF5,* and *RCOR3* and associated with high IFN-γ levels, overlaps with the region identified in the present study between the significant SNPs ARS-BFGL-NGS-36880 and ARS-BFGL-NGS-106233, with the latter mapping within the *TRAF5* gene. *TRAF5* is an adaptor protein of the TNF receptor-associated factor family involved in TNFα-mediated signalling. It contributes to the activation of NF-κB and cell death pathways and plays a regulatory role in immune homeostasis by modulating TLR responsiveness and limiting excessive inflammation (Badia-Bringué et al. 2023).

Reproduction and other traits

Although the primary focus of this study was feed efficiency, several identified genomic regions are associated with additional traits, indicating pleiotropic genetic effects (Table S3). Notably, some regions overlap with QTL for fertility-related traits, such as first-service conception and number of inseminations per conception, involving genes like *PITRM1*, as well as with QTL for body size, stature, and body depth, linked to genes including *CPQ*. Structural traits such as udder attachment, udder depth, and udder height, associated with QTL overlapping *GPR180* and *CPQ*, are known to influence milking efficiency, productive longevity, and susceptibility to mastitis (Amma et al. 2024; Medeiros et al. 2024). Together with the fertility-related QTL, these findings suggest that metabolic efficiency and feed efficiency are genetically interconnected with reproductive performance and structural soundness, likely through shared mechanisms related to energy balance and resource allocation.

Digestibility, DMI, and CH_4 emissions

This study also explored physiological differences between HEFF and LEFF cows in terms of nutrient utilisation and environmental impact, to evaluate whether selection for higher efficiency could promote sustainability through reduced CH_4 emissions and improved feed conversion. The analyses were conducted on a limited sample of cows (15 HEFF and 15 LEFF) and during a three-day period of time, restricting the generalisability of our results. Furthermore, CH_4 emissions were estimated using predictive equations and should therefore be interpreted as approximate values. Accordingly, the differences observed between the

groups should not be extrapolated to other herds or management conditions but should be considered as a promising basis and direction for further research or expanded studies. Lastly, because these measurements were collected over a limited number of days and during a single lactation phase, the observed differences should be interpreted as phase-specific. As expected, HEFF cows produced more milk than LEFF cows despite similar dry matter intake, suggesting a more favourable conversion of feed into milk, likely related to enhanced ruminal functionality. They also showed a significantly longer rumination time, and improved nutrient utilisation. This was supported by the higher fibre and protein digestibility observed in HEFF cows using NIRs.

Importantly, NIRs determines aNDF (which includes the ash correction for NDF), and this methodological difference may explain the discrepancies between the two approaches used to determine aTTD. Furthermore, the absence of available urine samples prevented the calculation of a complete nitrogen balance, limiting the accuracy of protein utilisation estimates, even though a significant difference emerged in CP digestibility between the groups.

The present results indicate that HEFF cows exhibited higher apparent total tract digestibility of crude protein and structural carbohydrates compared with LEFF cows, consistently observed using both NIRs- and laboratory-based approaches. This agreement between methods, despite differences in analytical variability, supports the robustness of the association between digestive efficiency and productive efficiency, as previously reported in dairy cattle (Schalla et al. 2012).

Higher crude protein digestibility in HEFF cows suggests a more efficient utilisation of dietary nitrogen, which is known to support milk yield and protein synthesis through enhanced microbial protein production and improved nitrogen use efficiency (Christensen et al. 2015; Xie et al. 2019). In addition, HEFF cows showed greater digestibility of fibre fractions, particularly NDF and ADF, which are widely recognised as key drivers of feed efficiency and energy availability in dairy cows (Huhtanen et al. 2009; Pino et al. 2018). These differences may reflect improved rumen function in HEFF cows, such as enhanced fibrolytic microbial activity or a more stable ruminal environment favouring fibre degradation (Xue et al. 2022; Tapio et al. 2023).

Although starch digestibility differed slightly between groups, the magnitude of this difference was small and likely of limited biological relevance. Overall, the improved digestibility of protein and fibre fractions observed in HEFF cows provides a plausible mechanistic explanation for their higher productive efficiency and lower methane intensity per unit of milk produced.

Lastly, although differences in feeding time, registered by collars, were not significant, HEFF cows showed a longer rumination time, a behavioural trend consistent with more efficient feed processing. Prolonged rumination enhances fibre breakdown and microbial access to cellulose (Souza et al. 2022), which aligns with the higher fibre digestibility found in HEFF cows.

These findings suggest that HEFF cows achieve greater feed conversion efficiency, which also translates into a lower CH_4 intensity per litre of milk (Capper 2021). Although CH_4 values were predicted rather than directly measured, similar patterns have been reported in studies using direct measurements, supporting the notion that more efficient animals are also less environmentally impactful. This evidence underpins the development of combined indices such as RFI and RME.

The residual-based classification captures meaningful phenotypic differences extending beyond production to digestive capacity and sustainability-related traits, highlighting the importance of integrating digestive and environmental phenotyping into multi-trait efficiency frameworks.

Conclusions

This study provides an integrative and targeted evaluation of productive efficiency in Holstein dairy cows by combining high-frequency phenotypic data, genomic information, and short-term nutritional efficiency measures. Using a residual-based classification approach, we identified cows showing persistent deviations from the expected milk yield under relatively stable physiological conditions, enabling the definition of contrasting efficiency profiles within a single herd. Under the specific conditions of late lactation and short-term monitoring, efficient cows produced more milk and showed improved fibre digestibility together with lower CH_4 emissions per litre of milk, suggesting a more favourable combination of nutrient utilisation and productive efficiency.

The selective genotyping approach effectively identified genomic regions previously associated with efficiency-related traits. Twenty-eight significant SNPs mapped to 52 candidate genes with known or plausible roles in metabolism, digestion, immunity, and reproduction. Although further validation is required, these findings provide a solid foundation for future

research and for developing genomic selection strategies that incorporate efficiency as a target trait.

From a methodological standpoint, some inconsistencies emerged between NIRs- and laboratory-based digestibility estimates, with the latter showing higher variability. These discrepancies likely highlight the need for caution when interpreting short-term physiological measurements and may reflect differences in analytical principles and individual biological variability that should be considered when interpreting physiological results.

Overall, our findings demonstrate that the residual-based classification captures biologically meaningful differences among animals and represents a promising framework for evaluating both digestive and environmental efficiency. Future studies incorporating full-lactation modelling, direct measurements of feed intake and methane emissions, and larger multi-herd populations will be essential to validate these preliminary findings and to further elucidate the genetic basis of productive and environmental efficiency in dairy cattle.

Acknowledgement

The authors acknowledge support from the University of Milan through the APC initiative.

Disclosure statement

No potential conflict of interest was reported by the authors.

University of Milan's departmental fund 'Line 2 Research Support Plan: Title of research: BREES: Breeding for estimated efficiency and sustainability'; European Union Next-GenerationEU, CN00000022;

Authors contributions

CRediT: **Francesca Bernini**: Conceptualization, Formal analysis, Visualization, Writing – original draft, Writing – review & editing; **Andrea Delledonne**: Formal analysis, Resources, Writing – original draft, Writing – review & editing; **Michela Frigerio**: Formal analysis, Resources; **Silvia Grossi**: Investigation, Writing – original draft; **Matteo Dell'Anno**: Investigation, Writing – original draft; **Chiara Punturiero**: Formal analysis, Resources; **Matteo Santoru**: Investigation, Resources; **Luciana Rossi**: Data curation; **C. A. Sgoifo Rossi**: Data curation; **Alessandro Bagnato**: Funding acquisition, Supervision; **Maria G. Strillacci**: Formal analysis, Funding acquisition, Supervision, Writing – original draft, Writing – review & editing.

Funding

This study was carried out within the Agritech National Research Centre and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) -MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4–D.D. 1032 17/06/2022, CN00000022). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them. This research was also supported by the University of Milan's departmental fund 'Line 2 Research Support Plan: Title of research: BREES: Breeding for estimated efficiency and sustainability'.

ORCID

Francesca Bernini http://orcid.org/0000-0001-8430-6813
Andrea Delledonne http://orcid.org/0000-0003-3034-3517
Silvia Grossi http://orcid.org/0000-0003-4588-0171
Matteo Dell'Anno http://orcid.org/0000-0002-2171-864X
Chiara Punturiero http://orcid.org/0000-0001-7778-7882
Matteo Santoru http://orcid.org/0009-0008-9582-8519
Luciana Rossi http://orcid.org/0000-0003-1178-4683
C. A. Sgoifo Rossi http://orcid.org/0000-0002-3602-9660
Alessandro Bagnato http://orcid.org/0000-0003-2087-1287
Maria G. Strillacci http://orcid.org/0000-0002-8972-1730

Data availability statement

The data supporting the conclusions of this manuscript are included in the Supplementary Materials.

References

Aleri JW et al. 2019. Associations between immune competence, stress responsiveness, and production in Holstein-Friesian and Holstein-Friesian × Jersey heifers reared in a pasture-based production system in Australia. J Dairy Sci. 102(4):3282–3294. https://doi.org/10.3168/jds.2018–14578

AOAC. 2023. Official Methods of Analysis. AOAC: Washington, D. C.

Amma Z, Reiczigel J, Fébel H, Solti L. 2024. Relationship between milk yield and reproductive parameters on three Hungarian dairy farms. Vet Sci. 11(5):218. https://doi.org/10.3390/vetsci11050218

Badami S et al. 2019. Identification of single nucleotide variations in the Toll-like receptor 9 (TLR9) gene and its association to mastitis susceptibility in dairy cattle. Biol Rhythm Res. 50(6):887–896. https://doi.org/10.1080/09291016.2018.1502233

Badia-Bringué G et al. 2023. Association between High Interferon-Gamma Production in Avian Tuberculin-Stimulated Blood from Mycobacterium avium subsp. paratuberculosis-Infected Cattle and Candidate Genes Implicated in Necroptosis. Microorganisms. 11(7):1817. https://doi.org/10.3390/microorganisms11071817

Benedeti PDB et al. 2019. Prediction of in vivo organic matter digestibility of beef cattle diets from degradation parameters estimated from in situ and in vitro incubations. J Agric Sci. 157(9-10):711–720. https://doi.org/10.1017/S0021859620000180

Berry DP, Crowley JJ. 2013. Cell biology symposium: genetics of feed efficiency in dairy and beef cattle. J Anim Sci. 91(4):1594–1613. https://doi.org/10.2527/jas.2012-5862

Brito LF et al. 2021. Review: genetic selection of high-yielding dairy cattle toward sustainable farming systems in a rapidly changing world. Animal. 15 Suppl 1:100292. https://doi.org/10.1016/J.ANIMAL.2021.100292

Brito LF et al. 2020. Genetic mechanisms underlying feed utilization and implementation of genomic selection for improved feed efficiency in dairy cattle. Can J Anim Sci. 100(4):587–604. https://doi.org/10.1139/cjas-2019-0193

Capper JL, Cady RA. 2020. The effects of improved performance in the US dairy cattle industry on environmental impacts between 2007 and 2017. J Anim Sci. 98(1):1–14. https://doi.org/10.1093/jas/skz291

Capper JL. 2021. Current issues and controversies in assessing the environmental impacts of livestock production. CABI Reviews. (2021) 16, No. 044. https://doi.org/10.1079/PAVSNNR202116044

Christensen RG et al. 2015. Effects of feeding birdsfoot trefoil hay on neutral detergent fiber digestion, nitrogen utilization efficiency, and lactational performance by dairy cows. J Dairy Sci. 98(11):7982–7992. https://doi.org/10.3168/jds.2015-9348

Connor EE et al. 2013. Use of residual feed intake in Holsteins during early lactation shows potential to improve feed efficiency through genetic selection. J Anim Sci. 91(8):3978–3988. https://doi.org/10.2527/jas.2012-5977

Connor EE, Hutchison JL, Van Tassell CP, Cole JB. 2019. Defining the optimal period length and stage of growth or lactation to estimate residual feed intake in dairy cows. J Dairy Sci. 102(7):6131–6143. https://doi.org/10.3168/jds.2018-15407

Darvasi A, Soller M. 1994. Selective DNA pooling for determination of linkage between a molecular marker and a quantitative trait locus. Genetics. 138(4):1365–1373. https://doi.org/10.1093/genetics/138.4.1365

De Souza RAd, Tempelman RJ, Allen MS, VandeHaar MJ. 2019. Updating predictions of dry matter intake of lactating dairy cows. J Dairy Sci. 102(9):7948–7960. https://doi.org/10.3168/jds.2018-16176

Fresco S, Sanchez M-P, Boichard D, Fritz S, Martin P. 2025. Sequence-based GWAS reveals genes and variants associated with predicted methane emissions in French dairy cows. Genet Sel Evol. 57(1):32. https://doi.org/10.1186/s12711-025-00977-z

Getaneh M, Alemayehu K. 2022. Candidate genes associated with economically important traits in dairy goats. Cogent Food & Agriculture. 8(1):2149131. https://doi.org/10.1080/23311932.2022.2149131

Gerber PJ et al. 2013. *Tackling climate change through livestock: a global assessment of emissions and mitigation opportunities*. Rome: Food and Agriculture Organization of the United Nations (FAO).

Golder HM et al. 2023. Associations among the genome, rumen metabolome, ruminal bacteria, and milk production in early-lactation Holsteins. J Dairy Sci. 106(5):3176–3191. https://doi.org/10.3168/jds.2022-22573

Guinan FL et al. 2023. Changes in genetic trends in US dairy cattle since the implementation of genomic selection. J Dairy Sci. 106(2):1110–1129. https://doi.org/10.3168/jds.2022-22205

Gurgul A et al. 2019. Diversifying selection signatures among divergently selected subpopulations of Polish Red cattle. J Appl Genet. 60(1):87–95. https://doi.org/10.1007/s13353-019-00484-0

Halachmi I, Guarino M, Bewley J, Pastell M. 2019. Smart animal agriculture: application of real-time sensors to improve animal well-being and production. Annu Rev Anim Biosci. 7(1):403–425. https://doi.org/10.1146/annurev-animal-020518-114851

Higgins MG et al. 2019. The effect of breed and diet type on the global transcriptome of hepatic tissue in beef cattle divergent for feed efficiency. BMC Genomics. 20(1):525. https://doi.org/10.1186/s12864-019-5906-8

Huang W, Kirkpatrick BW, Rosa GJM, Khatib H. 2010. A genome-wide association study using selective DNA pooling identifies candidate markers for fertility in Holstein cattle. Anim Genet. 41(6):570–578. https://doi.org/10.1111/j.1365-2052.2010.02046.x

Huhtanen P, Rinne M, Nousiainen J. 2009. A meta-analysis of feed digestion in dairy cows. 2. The effects of feeding level and diet composition on digestibility. J Dairy Sci. 92(10):5031–5042. https://doi.org/10.3168/jds.2008-1834

Isobe N, Nakamura J, Nakano H, Yoshimura Y. 2009. Existence of functional lingual antimicrobial peptide in bovine milk. J Dairy Sci. 92(6):2691–2695. https://doi.org/10.3168/jds.2008-1940

Jancewicz LJ et al. 2016. Characterization of the variation in the daily excretion of faecal constituents and digestibility predictions in beef cattle fed feedlot diets using near-infrared spectroscopy. Can J Anim Sci. 96(4):532–549. https://doi.org/10.1139/cjas-2015-0193

Kok A et al. 2016. Effective lactation yield: a measure to compare milk yield between cows with different dry period lengths. J Dairy Sci. 99(4):2956–2966. https://doi.org/10.3168/jds.2015-10559

Kurz JP et al. 2019. A genome-wide association study for mastitis resistance in phenotypically well-characterized Holstein dairy cattle using a selective genotyping approach. Immunogenetics. 71(1):35–47. https://doi.org/10.1007/s00251-018-1088-9

Lee C, Hristov AN. 2013. Evaluation of acid-insoluble ash and indigestible neutral detergent fiber as total-tract digestibility markers in dairy cows fed corn silage-based diets. J Dairy Sci. 96(8):5295–5299. https://doi.org/10.3168/jds.2012-6442

Lesta A, Marín-García PJ, Llobat L. 2025. A Comprehensive Review of the Bovine Immune Response to Pathogens. Int J Mol Sci. 26(17):8461. https://doi.org/10.3390/ijms26178461

Lipkin E et al. 1998. Quantitative trait locus mapping in dairy cattle by means of selective milk DNA pooling using dinucleotide microsatellite markers: analysis of milk protein percentage. Genetics. 149(3):1557–1567. https://doi.org/10.1093/genetics/149.3.1557

Lipkin E, Smith J, Soller M, Burt DW, Fulton JE. 2024. Mapping quantitative trait loci regions associated with Marek's disease on chicken autosomes by means of selective DNA pooling. Sci Rep. 14(1):31896. https://doi.org/10.1038/s41598-024-83356-w

Lipkin E et al. 2016. The use of Kosher phenotyping for mapping QTL affecting susceptibility to Bovine respiratory disease. PLoS One. 11(4):e0153423. https://doi.org/10.1371/journal.pone.0153423

Liu K. 2022. New and improved methods for measuring acid insoluble ash. Anim Feed Sci Technol. 288:115282. https://doi.org/10.1016/j.anifeedsci.2022.115282

Liu L et al. 2020. GWAS-based identification of new loci for milk yield, fat, and protein in Holstein cattle. Animals. 10(11):2048. https://doi.org/10.3390/ani10112048

Meade KG, Cormican P, Narciandi F, Lloyd A, O'Farrelly C. 2014. Bovine β-defensin gene family: opportunities to improve animal health? Physiol Genomics. 46(1):17–28. https://doi.org/10.1152/physiolgenomics.00085.2013

Medeiros GC et al. 2024. Genetic parameters for udder conformation traits derived from Cartesian coordinates generated by robotic milking systems in North American Holstein cattle. J Dairy Sci. 107(9):7038–7051. https://doi.org/10.3168/jds.2023-24208

Merchen NR. 1988. Digestion absorption and excretion in ruminants. The ruminant animal digestive physiology and nutrition. 172–201.

Meuwissen T, Hayes B, Goddard M. 2016. Genomic selection: a paradigm shift in animal breeding. Animal Frontiers. 6(1):6–14. https://doi.org/10.2527/af.2016-0002

Min BR, Solaiman S. 2015. Prediction of feed intake and its relationships with chemical composition of diets in goats consuming concentrate, Bahiagrass pasture and mimosa browse. OJAS. 05(03):283–293. https://doi.org/10.4236/ojas.2015.53033

Nascimento BM et al. 2024. Genetic relationships between behavioral traits and feed efficiency traits in lactating Holstein cows. J Dairy Sci. 107(10):8141–8149. https://doi.org/10.3168/jds.2023-24526

National Academies of Sciences, Engineering, and Medicine. 2021. Nutrient requirements of dairy cattle. National Academies Press (US).

Niu M et al. 2018. Prediction of enteric methane production, yield, and intensity in dairy cattle using an intercontinental database. Glob Chang Biol. 24(8):3368–3389. https://doi.org/10.1111/gcb.14094

Olijhoek DW, Difford GF, Lund P, Løvendahl P. 2020. Phenotypic modeling of residual feed intake using physical activity and methane production as energy sinks. J Dairy Sci. 103(8):6967–6981. https://doi.org/10.3168/jds.2019-17489

Paudyal S. 2021. Using rumination time to manage health and reproduction in dairy cattle: a review. Vet Q. 41(1):292–300. https://doi.org/10.1080/01652176.2021.1987581

Pino F, Mitchell LK, Jones CM, Heinrichs AJ. 2018. Comparison of diet digestibility, rumen fermentation, rumen rate of passage, and feed efficiency in dairy heifers fed ad-libitum versus precision diets with low- and high-quality forages. J Appl Anim Res. 46(1):1296–1306. https://doi.org/10.1080/09712119.2018.1498788

Poppe M, Mulder HA, Veerkamp RF. 2021. Validation of resilience indicators by estimating genetic correlations among daughter groups and with yield responses to a heat wave and disturbances at herd level. J Dairy Sci. 104(7):8094–8106. https://doi.org/10.3168/jds.2020-19817

Pryce JE, Wales WJd, De Haas Y, Veerkamp RF, Hayes BJ. 2014. Genomic selection for feed efficiency in dairy cattle. Animal. 8(1):1–10. https://doi.org/10.1017/S1751731113001687

Rezvannejad E, Asadollahpour Nanaei H, Esmailizadeh A. 2022. Detection of candidate genes affecting milk production traits in sheep using whole-genome sequencing analysis. Vet Med Sci. 8(3):1197–1204. https://doi.org/10.1002/vms3.731

Sahana G, Cai Z, Sanchez MP, Bouwman AC, Boichard D. 2023. Invited review: good practices in genome-wide association studies to identify candidate sequence variants in dairy cattle. J Dairy Sci. 106(8):5218–5241. https://doi.org/10.3168/jds.2022-22694

Salamone M et al. 2024. Milk yield residuals and their link with the metabolic status of dairy cows in the transition period. J Dairy Sci. 107(1):317–330. https://doi.org/10.3168/jds.2023-23641

Salgado Pardo JI et al. 2022. Candidate genes and their expressions involved in the regulation of milk and meat production and quality in goats (Capra hircus). Animals. 12(8):988. https://doi.org/10.3390/ani12080988

Santschi DE, Lefebvre DM, Cue RI, Girard CL, Pellerin D. 2011. Complete-lactation milk and component yields following a short (35-d) or a conventional (60-d) dry period management strategy in commercial Holstein herds. J Dairy Sci. 94(5):2302–2311. https://doi.org/10.3168/jds.2010-3594

Schalla A et al. 2012. Hot topic: apparent total-tract nutrient digestibilities measured commercially using 120-hour in vitro indigestible neutral detergent fiber as a marker are related to commercial dairy cattle performance. J Dairy Sci. 95(9):5109–5114. https://doi.org/10.3168/jds.2012-5650

Siewert JM, Salfer JA, Endres MI. 2019. Milk yield and milking station visits of primiparous versus multiparous cows on automatic milking system farms in the Upper Midwest United States. J Dairy Sci. 102(4):3523–3530. https://doi.org/10.3168/jds.2018-15382

Soares RAN, Vargas G, Duffield T, Schenkel F, Squires EJ. 2021. Genome-wide association study and functional analyses for clinical and subclinical ketosis in Holstein cattle. J Dairy Sci. 104(9):10076–10089. https://doi.org/10.3168/jds.2020-20101

Souza JG, Ribeiro CV, Harvatine KJ. 2022. Meta-analysis of rumination behavior and its relationship with milk and milk fat production, rumen pH, and total-tract digestibility in lactating dairy cows. J Dairy Sci. 105(1):188–200. https://doi.org/10.3168/jds.2021-20535

Strabel T, Misztal I. 1999. Genetic parameters for first and second lactation milk yields of Polish Black and White cattle with random regression test-day models. J Dairy Sci. 82(12):2805–2810. https://doi.org/10.3168/jds.S0022-0302(99)75538-4

Strillacci MG et al. 2014. Quantitative trait loci mapping for conjugated linoleic acid, vaccenic acid andΔ 9-desaturase in Italian Brown Swiss dairy cattle using selective DNA pooling. Anim Genet. 45(4):485–499. https://doi.org/10.1111/age.12174

Strillacci MG et al. 2023. Antibiotic treatments and somatic cell count as phenotype to map QTL for mastitis susceptibility in Holstein cattle breed. Ital J Anim Sci. 22(1):190–199. https://doi.org/10.1080/1828051X.2023.2175498

Tapio M, Fischer D, Mäntysaari P, Tapio I. 2023. Rumen microbiota predicts feed efficiency of primiparous nordic red dairy cows. Microorganisms. 11(5):1116. https://doi.org/10.3390/microorganisms11051116

Tian Z et al. 2022. Transcriptional regulation of milk fat synthesis in dairy cattle. J Funct Foods. 96:105208. https://doi.org/10.1016/j.jff.2022.105208

Titgemeyer EC, Armendariz CK, Bindel DJ, Greenwood RH, Löest CA. 2001. Evaluation of titanium dioxide as a digestibility marker for cattle. J Anim Sci. 79(4):1059–1063. https://doi.org/10.2527/2001.7941059x

Tullo E, Biffani S, Maltecca C, Rizzi R. 2014. Genetic parameters for milk yield and persistency in Carora dairy cattle breed using random regression model. Ital J Anim Sci. 13(4):3484. https://doi.org/10.4081/ijas.2014.3484

Turner SD. 2018. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. Journal of Open Source Software. 3(25):731.

Tyrrell HF, Reid JT. 1965. Prediction of the energy value of cow's milk. J Dairy Sci. 48(9):1215–1223. https://doi.org/10.3168/jds.S0022-0302(65)88430-2

Van Gestel S, Houwing-Duistermaat JJ, Adolfsson R, van Duijn CM, Van Broeckhoven C. 2000. Power of selective genotyping in genetic association analyses of quantitative traits. Behav Genet. 30(2):141–146. https://doi.org/10.1023/A:1001907321955

Van Soest PJ, Robertson JB, Lewis BA. 1991. Symposium: carbohydrate methodology, metabolism, and nutritional implications in dairy cattle. J Dairy Sci. 74(10):3583–3597.

VanRaden PM et al. 2009. Invited Review: reliability of genomic predictions for North American Holstein bulls. J Dairy Sci. 92(1):16–24. https://doi.org/10.3168/jds.2008-1514

Xie Y, Wu Z, Wang D, Liu J. 2019. Nitrogen partitioning and microbial protein synthesis in lactating dairy cows with different phenotypic residual feed intake. J Anim Sci Biotechnol. 10(1):54. https://doi.org/10.1186/s40104-019-0356-3

Xing C, Xing G. 2009. Power of selective genotyping in genome-wide association studies of quantitative traits. BMC Proc. 3 Suppl 7(Suppl 7):S23. https://doi.org/10.1186/1753-6561-3-s7-s23

Xue MY et al. 2022. Integrated meta-omics reveals new ruminal microbial features associated with feed efficiency in dairy cattle. Microbiome. 10(1):32. https://doi.org/10.1186/s40168-022-01228-9

Yang W, Zhang B, Xu C, Zhang H, Xia C. 2019. Effects of ketosis in dairy cows on blood biochemical parameters, milk yield and composition, and digestive capacity. J Vet Res. 63(4):555–560. https://doi.org/10.2478/jvetres-2019-0059