Invited review

González-Recio, O.; Martínez-Álvaro, M.; Tiezzi, Francesco; Saborío-Montero, A.; Maltecca, C.; Roehe, R.

Document Version
Publisher's PDF, also known as Version of record

Link to publication

Invited Review

Invited review: Novel methods and perspectives for modulating the rumen microbiome through selective breeding as a means to improve complex traits: Implications for methane emissions in cattle
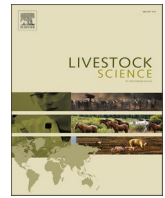
O. González-Recio [a,*], M. Martínez-Álvaro [b], Francesco Tiezzi [c], A. Saborío-Montero [d,e], C. Maltecca [f], R. Roehe [b]

[a] Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (CSIC), Madrid 28040, Spain
[b] Scotland's Rural College, Easter Bush Campus, Edinburgh, Scotland EH25 9RG, UK
[c] Department of Agriculture, Food, Environment and Forestry (DAGRI), University of Florence, Piazzale delle Cascine 18, Firenze 50144 Italy
[d] Escuela de Zootecnia y Centro de Investigación en Nutrición Animal, Universidad de Costa Rica, San José 11501, Costa Rica
[e] Posgrado Regional en Ciencias Veterinarias Tropicales, Universidad Nacional de Costa Rica, Heredia 40104, Costa Rica
[f] Animal Science Department, North Carolina State University, Raleigh, NC 27695, USA

HIGHLIGHTS

- Rumen microbiome composition is partially controlled by the host genomics, both at the taxonomic and microbial function levels, with heritabilities ranging between 0.10 and 0.40.
- Analyzing the microbiome poses some challenges that need to be carefully considered: large complexity, compositionality and lack of standardized bioinformatic procedures.
- Rumen microbiome plays an important role influencing the phenotypic variability of feed digestion related traits, including methane, and it may need to be properly accounted for in the statistical genetic models.
- Microbiota information may be included in the breeding programs, although specific strategies need to be defined in the future.

ARTICLE INFO

ABSTRACT

The rumen microbiome is responsible for methane emission in ruminants. The study of microbes in the rumen has attracted great interest in the last decade. High-throughput sequencing technologies have been key in expanding the knowledge of the microorganisms that populate the rumen through metagenomic studies. There is substantial evidence that the composition of the rumen microbiota is influenced by host genotype. Therefore, modulation of the microbiota poses an important tool for breeding for lower emissions in large and small ruminants. The main challenges of metagenomic studies are addressed and some solutions are proposed when available, including the incorporation of metagenomic information into statistical models regularly used in animal breeding. To incorporate microbiome information into breeding programs, the particularities of the rumen microbiome must be considered, from sampling to inclusion in selection indices. The latest advances in this area are discussed in this review.

1. Contribution of the rumen microbiome to global methane emissions

The first ruminants evolved from the artiodactyls that appeared in the Eocene, when grass began to cover most of the earth's surface. Since then, ruminants have evolved over 50 million years in symbiosis with the microorganisms that populate their rumen, developing a natural bioreactor with very specific characteristics that has contributed significantly to the evolution of the society. The ruminal microbial community plays an important role in ruminants by converting non-

---

edible plant material into nutrients the animal uses to produce high-quality protein that can be utilized by humans. Different microbial species are responsible for the digestion of complex plant carbohydrates, generating nutrients that are absorbed by the host. During enteric fermentation, some microbes generate methane as a by-product which the cow expels through eructation (Crable et al., 2011; Mizrahi et al., 2021). In 2019, methane levels in the atmosphere reached record magnitudes, about two-and-a-half times higher than in the pre-industrial era (NOAA, 2022). Methane has a warming potential between 28 and 34 times greater than CO_2 over a 100-year period (EPA, 2022). Over a 20-year period, it is about 84 times more powerful per unit of mass than carbon dioxide. Ruminants, and cattle in particular, are considered one of the most important sources of global methane (CH_4) emissions, producing about 4623 Mt of CO_2 equivalents (CO_2e) per year, with beef and dairy cattle responsible for 2495 and 2128 Mt of CO_2e per year, respectively (Gerber et al., 2013). About 40% of the globally emitted CH_4 comes from enteric fermentation in ruminants (Moss et al., 2000). Eructated CH_4 not only contributes significantly to global warming, but also represents a loss of dietary energy, estimated to be between 2 and 12% of net energy intake (de Haas et al., 2011; Johnson and Johnson, 1995). Indeed, the level of methane emissions is related to the intake capacity of the animal, and animals that emit less are expected to be more feed efficient (Manzanilla-Pech et al., 2022). Consequently, reducing enteric CH_4 emissions in ruminants has become an important area of research, aligned with the overall commitment to reduce greenhouse gas emissions by 2030. The short methane lifetime in the atmosphere (10–20 yrs) offers an interesting opportunity to reduce contribution of this gas to global warming in a relatively short period of time. Modulating the rumen microbiome to improve feed efficiency and reduce enteric CH_4 emissions from ruminants without altering production and animal health is desirable both as a strategy to decarbonize the livestock industry and to improve its productive efficiency.

The microbial community in the rumen consists mainly of fiber fermenters, predominantly Firmicutes and Bacteroidetes. Bacteroidetes phylum is represented primarily by genera *Prevotella*, Firmicutes phylum is more diverse in terms of genus and includes both primary fermenters *Ruminococcus, Lachnospira, Butyrivibrio* or *Pseudobutyrivibrio* and secondary fermenters *Selenomonas* or *Acetitomaculum. Fibrobacter* genus is another cellulose-degrader highly prevalent in rumen populations (Mizrahi et al., 2021). Proteobacteria and Actinobacteria are also representative rumen phyla. The methanogenic archaea community is less abundant, although different magnitude of relative abundances at genera level have been reported (Martínez-Álvaro et al., 2020; Wallace et al., 2015; Xue et al., 2020). In general, *Methanobrevibacter* (order Methanobacteriales) is the most abundant representative with a relative abundance that can reach up to 70% of all archaea. Protozoa are acquired through direct contact with saliva from other individuals. Still, they are very diverse and usually reach high abundances in the rumen (7–15%), amounting to up half of the rumen biomass (Hungate, 1966; Newbold et al., 2015). Anaerobic fungi inhabit the rumen and have an important role in degrading plant cell walls, mainly unlignified. They are more efficient than bacteria at colonizing and degrade the lignin-containing tissues. There are important interactions between fungi and other rumen microbes, especially with methanogenic microorganisms (Akin and Borneman, 1990).

The core microbiota of domesticated ruminants is, in general terms, similar across species and geographical locations. However, differences in microbial community composition may be primarily due to diet and host (Henderson et al., 2015). Despite the specificity of microorganisms that colonize the rumen, there is a great deal of substantial variability between and within breeds that can be exploited in livestock. Microbiota composition has been reported to be heritable (Wallace et al., 2019a; Saborío-Montero, 2020). There is evidence of genetic control of microbiota composition through changes in physical and physiological conditions in the rumen that promote the growth of specific microbes (Abbas et al., 2020; Gonzalez-Recio et al., 2018; Lassen and Løvendahl,

2016). Thus, modulating the microbiota by genetic or genomic selection on cow genes that promote the proliferation of a healthier and more efficient rumen microbiota is an efficient strategy to improve complex traits. For example, previous studies have predicted a potential reduction in methane emissions of 20 to 25% in 10 to 20 years when applying ad-hoc weights in the selection indices in dairy cattle (de Haas et al., 2011; González-Recio et al., 2020; López-Paredes et al., 2020)

Incorporating metagenomic information into statistical models poses important challenges that must be considered. This review aims to be an updated reference on metagenomic studies in cattle and to serve as a guide for the reader to dive into the specific issues related to the use of the microbiome for methane mitigation. First, we review the evidence on how the cow genome controls the composition of the rumen microbiome. Second, we identify the main challenges in microbiome analysis and propose available strategies to tackle them. Then, we present different statistical models to account for the phenotypic variability explained by the microbiome. Next, different cases of modulation of the rumen microbiome by selective breeding are discussed. Finally, we highlight the most important aspects to consider when incorporating microbiome information to mitigate methane emissions in cattle.

2. Evidence of host genetic control over rumen microbiome

The host influence on the rumen microbiome was first suggested by Weimer et al. (2010) by observing a restoration of the original microbiome features after exchanging ruminant content between cows. Their results showed that the bacterial community re-established itself within days, close to its original profile before the exchange (14d in one of the cows), and in all cases different from the donor cow's composition at the end of the experiment (62 d). Later, the specificity of the host microbiome was also reported for methanogenic Archaea and protozoa populations in beef cattle by Zhou et al. (2009), who studied the variability of ruminal methanogen Archaea and protozoa of four beef heifers fed with two different diets. This study showed that the responses of ruminal methanogenic profiles to distillers dried grain with soluble source were host-dependent, and that these methanogenic profiles differed between individuals (Global $R = 0.87$, $P < 0.01$).

The host genetic control over the microbiota is now widely recognized (Roehe et al., 2016; Li et al., 2019; Martínez-Álvaro et al., 2020; Saborío-Montero et al., 2020; Zhang et al., 2020). Roehe et al. (2016) studied the presence of host genetic control over methane emissions and microbiota composition through sire progeny groups in beef cattle. They identified the host genetic control on the rumen microbiota profile based on significant differences between of sire progeny groups in ruminal archaea : bacteria ratio of Aberdeen Angus and Limousine sired breed types. The statistical model was adjusted for diet, respiration chamber, and randomized block effects. They found that more similar archaea : bacteria ratio was obtained across sire progeny groups, which suggests that the host genetics partially controls relative microbial abundance in the rumen. Another example of host genetic control of the rumen microbiome emerged from comparisons of microbiome composition within and between breeds. Larger diversity between breeds (treatments) than within breeds indicates a breed effect on microbiome composition (in the context of ANOVA). Based on this approach, King et al. (2011) showed that out of a total 55 OTUs at species-level, 20 were common in two breeds (Holstein and Jersey), 23 were only present in Holstein and 12 were only present in Jersey, stressing the greater diversity in Holstein. These authors indicated that on-farm and diet environmental conditions were the same, which suggests that the observed differences were due to host breed genetics. Similar conclusions were obtained from a comparison between Holstein and Brown Swiss raised on the same farm (Gonzalez-Recio et al., 2018). They estimated the percentage of association of taxonomic features with the host genetic background based on the principal components of a genomic relationship matrix of single nucleotide polymorphism markers. In total, 48% of microbes were associated with the host's

genetic background.

In later years, further studies supported the hypothesis of host genetic control over the rumen microbiota by providing h^2 estimates for microbial traits. For example, Difford et al. (2018) estimated significant h^2 values between 0.17 and 0.25 for 8 out of 144 bacteria genera, while only one archaea genus (*Methanobrevibacter* spp.) had a h^2 estimate significantly different from zero (0.22 ± 0.09). This magnitude of h^2 estimate for *Methanobrevibacter* spp. is supported by other studies, such as Saborío-Montero et al. (2020) or Martínez-Álvaro et al. (2022a) which reported estimates ranging from 0.18 to 0.24. For *Prevotella* spp., one of the most abundant complex carbohydrate degraders in the rumen, both Wallace et al. (2019b) and Saborío-Montero et al. (2020) indicated a h^2 estimate larger than 0.4 in dairy cattle. In the latter study, moderate values (~0.30) were estimated for seven ciliate protozoa (*Stentor* spp., *Oxitricha* spp., *Paramecium* spp., *Pseudocohnilembus* spp., *Ichthyphthirius* spp., *Tetrahymena* spp. *and Stylonychia* spp.). In general, broader ranges of h^2 estimates from 0.20 to 0.60 and 0.13 to 0.60 were reported by Wallace et al. (2019b) and Martínez-Álvaro et al. (2022a) respectively, supporting heritable microbial composition at taxonomical level (microbial abundances) and at the functional level (microbial gene abundances)

The host genetic effect on the microbiome implies that members of the same family share a similar functional and taxonomic microbiome due not only to vertical maternal transmission, but also to shared host genomics underlying traits that influence microbial colonization and fitness, transmitted in a classical mendelian sense. Genome-wide association studies searching for causal polymorphisms explaining ruminal microbial abundances have identified SNPs associated with muscle contraction and passage rate (Zhang et al., 2020), absorption of nutrients by the rumen epithelium (Li et al., 2019), immune-related signaling pathways, tight junction of epithelial cells (Fan et al., 2021), mucin-encoding genes critical for gut mucosal health (Fan et al., 2019), cell division and cell cycle (Cardinale and Kadarmideen, 2022), and genes related to appetite, satiety and digestion (Gonzalez-Recio et al., 2023). These findings open the possibility of using genetic selection to modulate the microbiome, treated as another phenotype, and breed animals that favor a specific composition of the rumen metagenome throughout their productive life, which in turn influences complex traits of interest for the sustainability of ruminant productive systems from both economic and environmental perspectives. Since selective breeding for different traits may have a significant impact on the microbiome composition and vice versa, it is paramount to compare various strategies to achieve a healthy and efficient microbiome. Different strategies to include microbiome information in the selective breeding programs, as well as to evaluate current methodological and applicability limitations of each strategy need to be discussed.

3. Challenges in microbiome analysis

Analysis of the microbiome in the context of quantitative genetics presents several methodological challenges that must be considered before conclusions can be drawn about the control of the microbiome over complex traits and before integrating microbiome information into current breeding programs. This review focuses on the three major limitations of metagenomic datasets. It is not intended to be a detailed guide to microbiome analyses, which can be found elsewhere (ten Hoopen et al., 2017), but to explore some solutions to each of the challenges:

(1) A large number of taxa and their functions affect complex traits by dynamically interacting or competing with each other. In other words, it is biologically unlikely that a few taxa explain much of the variability in complex traits of interest; rather, studies show that the additive effect of a large number of taxa and microbial functions and their interactions underlie the biology of a complex trait (Martínez-Álvaro et al., 2020). The interactions

between microbial traits are in some cases very tight, leading to very high correlations between microbial abundances, which are therefore redundant for prediction approaches. To address the high-dimensional challenge, multivariate analysis tools are essential, which include a wide range of approaches such as cluster analysis, principal component analysis (PCA), canonical correspondence analysis (CCA), or redundancy analysis and their variations (Ramette, 2007). Cluster analysis is a descriptive approach to disentangle the correlation structure underlying the microbial community complexity and simplify the system's complexity by grouping variables based on their correlations. Once the microbial components are grouped into a few clusters, it is much easier to resolve the associations between the complex traits and the different clusters, as described in Martínez-Álvaro et al. (2020) when studying the association between rumen microbiome and methane emissions in beef cattle. Principal component analysis (PCA) is an appropriate tool to elucidate the latent variables underlying the main sources of variation within the microbial system. As in cluster analysis, once the latent variables are extracted, a second step is to examine the phenotypic or genetic correlations of the latent variables with the complex trait, as Saborío-Montero et al. (2021b) proposed. Alternatively, the latent variables inherent in the microbiome can be constructed to maximize their covariance with the complex trait of interest using projection onto latent structure linear regression (PLS) or discriminant analysis (PLS-DA), which seems to be a promising strategy for developing selection criteria based on the microbiome because it is supposed to capture the part of the microbiome that best fits the trait.

(2) Microbiome data are compositional because they are generated after a sampling and sequencing process that limits the number of reads. Therefore, abundance data are structured as proportions that sum an irrelevant constant (Gloor et al., 2017) and contain information about the relationship between different microbes rather than the total amount of each microbe (Aitchison, 1982). This implies a lack of *subcompositional coherence* in the analysis of relative abundances (*i.e.*, they change as a function of the total number of microbial traits considered) and the occurrence of *spurious correlations* between them (when variables are forced to sum a constant, this leads to "spurious" correlations between them that are not observed in their original values). To circumvent these challenges, statisticians propose to work with the pairwise log-ratios of the components, which remain constant for the common components when the number of components is reduced or expanded, although the values of their relative abundances change. However, the number of pairwise log-ratio combinations can be overwhelming with thousands of microbial taxa or their functions. Alternatively, a subset of these pairwise log-ratio transformations can be used, always using the same component as the denominator or reference (referred to as an additive log-ratio transformation, or *alr*). To ensure that a reduced subset of pairwise log-ratio transformations can satisfactorily approximate the same multivariate geometric structure of the samples as that of all pairwise log-ratio transformations (referred to as log-ratio geometry), the reference component must be chosen to maximize the Procrustes correlation between the *alr* geometry and the exact log-ratio geometry. Many potential references exist for high-dimensional data such as the microbiome. Recently, Greenacre et al. (2021) showed that in microbiome and other -omics data, there are certain components that, when used as denominators or references, produce *alr* geometries with Procrustes correlations > 0.99 with the log-ratio geometry. These reference components are consistently the less variable components across the samples in the whole composition, greatly simplifying the interpretation of each *alr* as an interpretation of the numerator. Other shortcuts of all pairwise log-ratios that also

respect the exact log-ratio geometry are the centered log-ratio (*clr*) or isometric log-ratio (*ilr*) transformations. Both involve geometric means of all (*clr*) or only a few (*ilr*) parts of the composition in the denominator (*clr*) or both in the numerator and/or denominator (*ilr*). As a result, they are complicated to interpret because they do not have the simplicity of a pairwise logarithm between two components. Moreover, *clr* is not strictly compositionally coherent because its geometric mean contains information about all parts in the composition that change when the composition is reduced or expanded (Eb et al., 2016). Nevertheless, it is most commonly used in microbiome studies (*e. g.*, in the ALDEx2 package) because it does not require selecting a particular component to be chosen as a reference.

(3) Bioinformatics pipelines frequently store microbiome data with massive null values, which can either mean that a component is not present in a sample or is present only at levels below the detection limit, with no opportunity to distinguish between the two (van den Boogaart and Tolosana-Delgado, 2013). The sparse data frameworks of the microbiome make it challenging to evaluate its effects on complex traits, and log transformations do not work with data containing zeros. How best to deal with these zeros remains an open research topic. A common practice is replacing zeros with a number below the detection limit or adding a fixed value (*e.g.*, 1) to all samples and components so that zeros are replaced with 1 s. A more elaborate method for replacing zeros is the Dirichlet sampling method, in which pseudo-counts are added to all components. It should be noted that simply adding a pseudo-count to all components does not preserve the ratios between them, which can be changed by multiplicatively changing the non-zero components (Martín--Fernández and Thió-Henestrosa, 2006). From a breeding perspective, a tempting alternative is to simply discard the missing components that are absent in any of the samples, as it is convenient to be able to measure microbial traits in all selection candidates (*i.e.*, to be part of the core microbiome, see Perlman et al. (2022) for a recent review on this topic). However, it may come at the expenses of some interesting information loss. A Geometric Bayesian Multiplicative procedure can be applied to impute zeros to small values while maintaining compositionality of observed features (Palarea-Albaladejo and Martín-Fernández, 2015).

4. Incorporation of microbiome information into genetic evaluation models

In addition to contributing to the determination of many of the host's traits, the ruminal microbiome composition can itself be considered as a quantitative trait since it is measurable and determined by genetic and environmental components. As discussed in this paper, the microbiome data are both compositional and highly dimensional. For simplicity, we will assume that the microbiome is represented by one or a few variables combining its profiles. These could be its most relevant features, a given number of principal components, some diversity index built on the compositional data, or a dichotomous variable indicating a taxon's presence or absence. While the statistical treatment for these scenarios would be different, here we aim to highlight a working framework that can be applied to any numerical representation of the microbial community composition and utilized to obtain variance components, discover causal variants, or calculate the breeding value of candidate breeders.

Different approaches have been considered for the statistical treatment of microbial information in animal breeding. Christensen et al. (2021) proposed three methods pertinent to this manuscript because the rumen microbial composition could fall under the definition of an 'intermediate omics feature'. Similarly, Tiezzi et al. (2021) conducted a genome-wide association study using information about the gut microbiome as a mediator of host genomic effect on the phenotype (fat deposition). While being conducted on gut microbiome in a monogastric species, the same conceptual framework can be applied to ruminants. We will often refer to these two studies in comparing the proposed models as follows.

4.1. Microbiome as a source of phenotypic variability

First and foremost, given prior knowledge of the impact of microbial composition on the traits of interest, we can use microbial information as an independent variable in the model, as defined below:

$$\mathbf{y} = \mathbf{Xb}_y + \mathbf{Mu}_y + \mathbf{Za}_y + \mathbf{e}_y, \tag{1}$$

where $\mathbf{y}$ is the selection trait of interest, $\mathbf{X}$ and $\mathbf{b}_y$ are respectively the incidence matrix and vector of solutions for the environmental effects, $\mathbf{M}$ is a matrix that contains the information on the microbial features (*e. g.*, species abundance, microbial diversity), $\mathbf{u}_y$ is the vector of microbial effects, $\mathbf{Z}$ and $\mathbf{a}_y$ are the incidence matrix and vector of solutions for the additive genetic effects of the host, $\mathbf{e}_y$ is the residual error. This model would be akin to the 'method 2′ proposed by Christensen et al. (2021) and allows the estimation of both the host and the microbiome effects on the phenotype (the $\beta^{'}$ and $\gamma^{'}$ parameters as per Tiezzi et al., 2021), here defined as $\mathbf{u}_y$ and $\mathbf{a}_y$, respectively. While the genetic effects are fitted as random, the environmental effects could be fitted either as fixed or random. In this paper, we assume these to be fitted as fixed, although we acknowledge that different approaches could be used. The vector of solutions for the genetic effect (also known as estimated breeding values, EBV) can be defined as $\mathbf{a}_y \sim N(0, \mathbf{G}\sigma_{a_y}^2)$, where $\sigma_{a_y}^2$ is the additive genetic variance and $\mathbf{G}$ is a covariance (relationship) matrix built on the pedigree, genomic markers or both (Christensen et al., 2021). Note that, for the additive genetic effect, $\mathbf{Z}$ could also be replaced by a matrix of allele counts and $\mathbf{a}_y$ could be a vector of allele (independent) substitution effects, as in a multiple-regression model (de los Campos et al., 2013). The microbiome effect could be defined in several different ways. The microbiome could be defined by a single descriptor of its richness or diversity (Lu et al., 2018). Therefore $\mathbf{M}$ would be a single-column matrix reporting such measure and $\mathbf{u}_y$ would be a regression coefficient quantifying the change in the phenotype given a unit of change in richness/diversity. The microbiome could be defined by multiple descriptors (*e.g.,* OTU, ASV), $\mathbf{M}$ would be a matrix with number of columns equal to the number of descriptors and $\mathbf{u}_y$ would be a vector reporting the regression coefficients that quantify the change in the phenotype given a unit change in each microbial descriptor. Lastly, $\mathbf{M}$ could simply be an incidence matrix relating the individuals to the observations and $\mathbf{u}_y$ could be a vector of estimated microbial values (EMV), defined as $\mathbf{u}_y \sim N(0, \mathbf{Q}\sigma_{u_y}^2)$, where $\sigma_{u_y}^2$ is the microbial variance and $\mathbf{Q}$ is a covariance matrix among the individuals built on microbial information (Camarinha-Silva et al., 2017; Khanal et al., 2019; Ross et al., 2013; Saborío-Montero et al., 2021a). In this sense, the EMV represents the deviation in performance of that individual from the average of the population given its ruminal microbial composition. Finally, $\mathbf{e}_y$ is assumed to be distributed as $\mathbf{e}_y \sim N(0, \mathbf{I}\sigma_{e_y}^2)$, where $\sigma_{e_y}^2$ is the residual variance and $\mathbf{I}$ is an identity matrix.

Estimates for the variance components for the two random effects $\sigma_{a_y}^2$ and $\sigma_{u_y}^2$ allows to calculate the ratio of each variance component to the total phenotypic variance. Indeed, the influence of microbiome variation on complex traits has often been analogizedparalleled to the influence of genetic variation on complex traits, with the term "microbiability" (m^2), hinting at the widely known heritability, which can be defined in model [1] as $m^2 = \sigma_{u_y}^2 / (\sigma_{a_y}^2 + \sigma_{u_y}^2 + \sigma_{e_y}^2)$. The term m^2 was first used in Difford et al. (2018) for CH_4 emissions in dairy (m^2 was 13%) and has been subsequently broadly used thereafter (Khanal et al., 2021; Ramayo-Caldas et al., 2021). However, a limitation of the

microbiability concept within a breeding perspective is that it considers the microbiome contribution to the phenotypic variance, independently of host genetics. As such, it disregards the potential host genetics contribution to microbiota composition, which might hamper its interpretation. Note that a covariance between $\mathbf{a}_y$ (EBV) and $\mathbf{u}_y$ (EMV), which aims at capturing the dependencies between the genes affecting $\mathbf{y}$ and the part of the microbiota that influences $\mathbf{y}$, could be considered in the model (if all individuals in $\mathbf{G}$ are included in $\mathbf{Q}$ and viceversa). This would be similar, in statistical terms, to the covariance that is usually considered between direct and maternal genetic effects (Eaglen and Bijma, 2009) or between different orders of the polynomial in random regression models (Schaeffer, 2004). However, disentangling the covariance between the two terms might be challenging in practice, due to the different factors determining each vector`s variability. While the EBV of an individual is determined by the genetic architecture of the trait and the known genotype of the individual, the EMV of an individual is determined by the effect of each microbial feature on the trait and the (relative) abundance of microbial features in the individual, which are time-point observations as they may change with time. Genetic and environmental (both permanent and temporary) effects determine such microbial composition. The EMV of an individual could therefore be determined by an environmental component that is not found in the EBV. The discrepancy in the determinants of $\mathbf{a}_y$ and $\mathbf{u}_y$ could hamper the covariance estimation between them, which is, the estimation of such covariance could hold in statistical terms but might lack a biological rationale, an alternative approach will therefore be presented in the flowing.

The host genetic and rumen microbial effects could also be fitted in interaction, as proposed by Khanal et al. (2020). Due to the high dimensionality of both effects, the interaction is usually modeled by creating an *ad hoc* covariance matrix using the Hadamard product of $\mathbf{G}$ and $\mathbf{Q}$. Simon et al., 2019; Margulis and Fester, 1991 used the holobiont concept to describe the system composed by a host and its associated communities of microorganisms. The 'holobiability' (h_o^2) term was first coined by Saborío-Montero et al. (2021a) to describe the proportion of phenotypic variance explained by the joint host-microbiome effect, including the genetic variance, microbiome variance, and the variance generated by the interaction of both. Using an extended version of model 1 including such interaction, h_o^2 can be defined as $h_o^2 = (\sigma_{a_y}^2 + \sigma_{u_y}^2 + \sigma_{axu_y}^2) / (\sigma_{a_y}^2 + \sigma_{u_y}^2 + \sigma_{axu_y}^2 + \sigma_{e_y}^2)$. Notice that again, the success of this method is hampered by the determination of the microbiota composition from both host genetic and environmental effects: the presence of genetic variation in the microbial variables would lead to fit a partial 'genotype by genotype interaction'.

4.2. Microbiome as a trait

Suppose the microbial composition of the rumen of a given individual is determined, at least in part, by the genetic composition of the host. In that case, it could also be considered as a trait itself and added to any phenotypic trait of interest in a bivariate model:

$$\begin{cases} \mathbf{m}_i = \mathbf{X}\mathbf{b}_{m_i} + \mathbf{Z}\mathbf{a}_{m_i} + \mathbf{e}_{m_i} \\ \mathbf{y} = \mathbf{X}\mathbf{b}_y + \mathbf{Z}\mathbf{a}_y + \mathbf{e}_y \end{cases}, \tag{2}$$

where $\mathbf{m}_i$ is the ith microbial feature, $\mathbf{b}_{m_i}$ is the vector of solutions for the environmental effects on the microbial composition, $\mathbf{a}_{m_i}$ is the vector of additive genetic effects that determine the microbial composition and $\mathbf{e}_{m_i}$ is the residual term (determined by uncontrolled random effects that determine the microbiota composition) while $\mathbf{y}$, $\mathbf{X}$, $\mathbf{Z}$, $\mathbf{b}_y$, $\mathbf{a}_y$ and $\mathbf{e}_y$ are as defined above. Note that the microbial feature $\mathbf{m}_i$ (with $i = 1, 2, \ldots, k$) could be a column of the $\mathbf{M}$ matrix in [1] (Aliakbari et al., 2021; Bergamaschi et al., 2020), a principal component of such matrix (Saborío-Montero et al., 2021b) or an ecological measurement of richness and diversity of the microbiota itself (Lu et al., 2018). This model would allow the estimation of the host genetic effects on both the microbiome

and the phenotype (α and γ parameters as per Tiezzi et al. 2021), here defined as $\mathbf{a}_{m_i}$ and $\mathbf{a}_y$.

Here, the covariance between vectors defining residual and host additive genetic effects is straightforward, as in any implementation of the multiple-trait model (Aliakbari et al., 2021; Lu et al., 2018; Saborío-Montero et al., 2020; Martínez-Álvaro et al., 2022a). For the additive genetic effect:

$$\begin{bmatrix} \mathbf{a}_{m_i} \\ \mathbf{a}_y \end{bmatrix} \sim N(0, \ \mathbf{A} \otimes \mathbf{G})$$

And for the residual effect:

$$\begin{bmatrix} \mathbf{e}_{m_i} \\ \mathbf{e}_y \end{bmatrix} \sim N(0, \ \mathbf{R} \otimes \mathbf{I}),$$

where $\mathbf{A}$ is a variance-covariance matrix (**VCV**) between the two additive genetic effects, $\mathbf{G}$ is the additive genetic relationship matrix among the individuals (as defined above), $\mathbf{R}$ is a residual covariance matrix and $\mathbf{I}$ is an identity matrix which makes the residuals values uncorrelated within trait. For the environmental effect, there would be no VCV if this is fitted as fixed, while there could be if this were to be fitted as random. Again, the estimation of the VCV has to be carried out with REML or Gibbs sampling. Note that estimation of the VCV is now feasible since both $\mathbf{a}_{m_i}$ and $\mathbf{a}_y$ only extract the additive genetic component from the dependent variable, removing the noise from the environmental effects, which is not the case present in $\mathbf{u}_y$ from formula [1].

4.3. Microbiome as a trait while accounting for the recursive structure

A structural equation model (SEM) has been proposed combining the approach shown in [1] and [2], a structural equation model (SEM) can be proposed (Christensen et al., 2021; Saborío-Montero et al., 2020; Tiezzi et al., 2021; Varona et al., 2007). Such model would account for the covariance between $\mathbf{a}_{m_i}$ and $\mathbf{a}_y$, as the host genomic composition affects both the microbial composition ($\mathbf{m}_i$) and the phenotypic trait of interest ($\mathbf{y}$), while also considering $\mathbf{M}\mathbf{u}_y$, as the effect that the microbiota composition (as a whole) exerts on the phenotypic trait of interest ($\mathbf{y}$). This model can be written as:

$$\begin{cases} \mathbf{m}_i = \mathbf{X}\mathbf{b}_{m_i} + \mathbf{Z}\mathbf{a}_{m_i} + \mathbf{e}_{m_i} \\ \mathbf{y} = \mathbf{X}\mathbf{b}_y + \mathbf{M}\mathbf{u}_y + \mathbf{Z}\mathbf{a}_y + \mathbf{e}_y \end{cases} \tag{3}$$

While all terms have already been defined, note that their meaning and value could change due to their simultaneous estimation (Tiezzi et al., 2021). This model would mimic the 'method 1′ proposed by Christensen et al. (2021) and would allow the estimation of the full model accounting for the host genetic effect on both the phenotype and the microbiome and the effect of the latter on the phenotype (α', γ' and β' parameters as per Tiezzi et al. 2021), here defined as $\mathbf{a}_y$, $\mathbf{a}_{m_i}$ and $\mathbf{u}_y$.

The SEM would allow disentangling the recursive effect $\mathbf{M}\mathbf{u}_y$ at the phenotypic level from the covariance at the host genetic level, the first describing the putative causal effect that the microbial feature $\mathbf{m}_i$ exerts on $\mathbf{y}$, and the latter describing the shared genetic determination of $\mathbf{m}_i$ and $\mathbf{y}$. In addition, the SEM allows disentangling direct and indirect effects of the host genotype on the phenotypic trait $\mathbf{y}$. Following model [3], $\mathbf{a}_y$ represents the direct effect of the host genotype while the value yielded by the product $\mathbf{a}_{m_i} \times \mathbf{u}_y$ represents the microbiome-mediated effect (Hayes, 2022; Tiezzi et al., 2021). The possibility to disentangle recursive effects from shared covariance and estimate direct and mediated effects has practical implications in inference about the biological processes and animal breeding value estimation (Valente et al., 2010).

4.4. Microbiome as a source of both genetic and environmental variability

Model [3] allows extracting the host genetic covariance from the microbial composition $\mathbf{m}_i$ while still accounting its effect of the pheno-

typic trait **y**. However, some estimability issues could still arise due to the information-dense nature of the microbial composition. This one, in fact, is not only determined by the host but is also affected by controllable environmental conditions (*e.g.,* diet, heat stress) and random environmental variation (*e.g.,* the floral composition of the pasture/ration). In addition, the environmental effects could be permanent of the individual (*e.g.,* stress at weaning) or temporary (*e.g.,* a particular batch of feed). Consequently, microbial composition could change daily, with patterns that may not be fully understood at the moment. Hence again, the need to account for the complex determination of the microbiome even when this one is used as an independent variable.

With this in mind, model [1] could be further expanded, breaking down the $\mathbf{Mu}_y$ component into: 1) a component controlled by the host genotype (*i.e.* $\mathbf{Za}_m$), 2) a component systematically controlled by the environment (*i.e.,* $\mathbf{Xb}_m$) and 3), a component describing random fluctuations (*i.e.* $\mathbf{e}_m$) all taken from a univariate model that resembles the model in [2]. The expanded model could be written as:

$$y = \mathbf{Xb}_y + [\mathbf{Xb}_{m_i}]\vartheta_y + [\mathbf{Za}_{m_i}]\xi_y + [\mathbf{e}_{m_i}]\psi_y + \mathbf{Za}_y + \mathbf{e}_y \qquad (4)$$

In this formula, $\mathbf{Xb}_{m_i}$ could be a vector of environmental Best Linear Unbiased Estimates (BLUE) for $\mathbf{m}_i$, $\mathbf{Za}_{m_i}$ a vector of host additive genetic Best Linear Unbiased Predictions (BLUP) for $\mathbf{m}_i$ and $\mathbf{e}_{m_i}$ a vector of residuals for $\mathbf{m}_i$, while ϑ_y is the regression coefficient for the microbial systematic environmental effect, ξ_y is the regression coefficient for the microbial genetic effect and ψ_y is a regression coefficient for the microbial random environmental effect. All the other terms are defined as in model [1]. Also, consider that the equivalence $\mathbf{m}_i = \mathbf{Xb}_{m_i} + \mathbf{Za}_{m_i} + \mathbf{e}_{m_i}$ should hold, and the factors ϑ_y, ξ_y, and ψ_y have the only function to regulate the contribution of each microbial component into **y**. Similarly, $\mathbf{Xb}_{m_i}$, $\mathbf{Za}_{m_i}$ and $\mathbf{e}_{m_i}$ could be matrices with number of columns equal to the number of microbial features considered (with $i = 1, 2, ..., k$) and number of rows equal to the number of individuals with microbiome data, with ϑ_y, ξ_y, and ψ_y being vectors of length equal to the number of features and containing the (independent) partial regression coefficients. They need to be estimated using auxilliary models, for instances, this model extends the 'method 3′ proposed by Christensen et al. (2021), by accounting for the microbial random environmental predictor ($\mathbf{e}_{m_i}$). The use of the residual deviations (or their variance) has been proposed to model the micro-environmental plasticity of some traits (Rönnegard et al., 2013).

The use of covariates defined as the BLUE for the systematic environmental effect is effective when the model incorporates solutions from another model but has the inconvenience of being a two-step approach, although Christensen et al. (2021) have proposed a method to solve a single system of equation for the estimation of all parameters.

The model in [4] has other advantages as compared, for example, to the model in [1]. First, this model allows the covariance estimation between the terms that regulated by the host genotype. The $\mathbf{a}_y$ component is strictly determined by the host genotype and the genetic architecture of **y**, so is the component $\mathbf{a}_{m_i}$ given the genetic architecture of $\mathbf{m}_i$. Therefore, the shared determination of these components would allow the proper estimation of the covariance between $\mathbf{a}_y$ and $\mathbf{a}_{m_i}$ i.e. the covariance between the additive genetic effect of the host and the microbial effect(s) conditional on the host additive genetic itself. This covariance should be interpreted as the shared genetic architecture between **y** and $\mathbf{m}_i$. Likewise, the same covariance estimation could be carried out between $\mathbf{b}_y$ and $\mathbf{b}_{m_i}$ if the systematic environmental effects were to be fitted as random. This method is proposed as a solution to the issue raised under the model [1], where the estimation of the covariance between the vectors $\mathbf{a}_y$ and $\mathbf{u}_y$ is hampered by the discrepancy in the factors that determine their variation (see model [1]). Second, the model [3] allows to fit the explicit interaction between the host genotype and microbial effects. The main issue of fitting such interaction in [1] was the contemporary presence of host genetic and environmental variation in the microbial taxa or functions (*i.e.* **M**). When [3] is used, fitting the

interaction between $\mathbf{a}_y$ and $\mathbf{b}_{m_i}$ would be equivalent to fitting a proper genotype by environment interaction. Likewise, fitting the interaction between $\mathbf{a}_y$ and $\mathbf{a}_{m_i}$ would be equivalent to fitting a standard additive by additive interaction, with the latter being conditional on the microbial composition. Note that this interaction is not equivalent to the host epistatic interactions that have been previously estimated (Jiang and Reif, 2015): here, the second component describes a host genotype contribution that is 'filtered' by the microbial composition and therefore resembles a marginal mediated effect (Tiezzi et al., 2021).

The preferred model may be context-dependent. For instance, models in [1] and [4] could be implemented to determine the effect of a microbial feature on the phenotype. In contrast, models in [2] and [3] would be better when determining the host genetic effect on microbial composition.

5. Direct selection to modulate the rumen microbiome

The thousands of species and their functions in the rumen microbiome embed large genetic variation and phenotypic plasticity that may have been previously targeted only indirectly in selection for productive traits. This large intrinsic variation provides an opportunity to direct selection towards an optimal ruminal microbial environment that could, potentially, improve efficiency (*e.g.* optimize the fermentation of forages into essential nutrients utilized by the host), minimize methane emissions (*e.g.* reduce substrates used by methanogenesis as excess H_2) and improve host fitness and health.

For a microbial trait to be considered as a selection criterion it must be present in much of the population (*i.e.*, to be part of the core microbiome, see Perlman et al. (2022) for a recent review on this topic), show considerable phenotypic variation across animals, and be heritable and genetically correlated with the productive traits of interest (r_g). Another point to consider is whether the composition of the rumen microbiome, described as the taxonomy of the microbes or their function, best fits these criteria. Even within the same genus, different strains can have very different metabolic pathways (Huttenhower et al., 2012; Martínez-Álvaro et al., 2022a), and some authors suggest that the heritability of a microbial trait is most likely to occur in terms of functional pathways rather than specific strains (Sandoval-Motta et al., 2017). However, the reported h^2 estimates in cattle do not support this hypothesis. Similar h^2 ranges were reported for abundances of core microbial genera (from 0.08 to 0.62 in dairy cattle and from 0.19 to 0.54 in beef cattle) than of core microbial genes (from 0.15 to 0.66 in beef cattle)*,* although with large estimation errors (Cardinale and Kadarmideen, 2022; Martínez-Álvaro et al., 2022a; Saborío-Montero et al., 2020). The magnitude of r_g strongly depends on to what extent the objective trait relies on the effect of the heritable microbiome and its metabolites. In the case of methane emissions, Martínez-Álvaro et al. (2022a) found that, among the heritable part of the rumen microbiome, a larger number of microbial genes had a strong genetic association (r_g from |0.59| to |0.93|) with methane emissions in comparison to microbial genera (115 vs. 29). The same authors proposed a selection index based on the *alr*-transformed abundance of 30 of these microbial genes to mitigate gas emissions, termed as *microbiome-driven breeding strategy*. The expected response to selection exceeded that estimated using direct breeding based on measured methane emissions using respiration chambers, which is explained by a stronger r_g and a larger h^2 of the microbial genes compared to the h^2 of methane emissions. However, the results need to be verified in further studies because EBVs prediction is very sensitive to the variance components, which were estimated based on a small population ($n = 359$ animals). As an alternative to using microbial abundances as a selection criterion, Saborío-Montero et al. (2021b) proposed reducing the dimensionality of the microbiome by singular value decomposition of the *clr*-transformed abundances and using the resulting latent component as a quantitative trait for selection to reduce methane emissions in dairy cattle. This approach could

compromise the full biological interpretation of the selection criterion but simplifies the complexity of the microbiome to a single trait, which could facilitate microbiome breeding in practice. Based on either taxonomical composition at different classification levels or functional composition, the h^2 estimate of the latent component was ~0.3 and its r_g estimate with methane emissions was ~|0.7|, making this strategy appealing for selective breeding.

Some microbial metabolic pathways exert a common influence on several objective traits or are likely to interact with other microbial activities that affect host metabolism. Before one microbial activity is targeted by genetic selection, the expected correlated response on other productive traits and overall animal fitness must be carefully examined. Still, the consequences do not necessarily need to be detrimental. For example, Martínez-Álvaro et al. (2022b) found that a microbiome-driven breeding strategy to increase omega-3 and conjugated linoleic acid (CLA) content in beef can also reduce methane emissions. This is explained by the fact that selected microbial metabolism was involved in diverting H_2 to the synthesis of microbial proteins, resulting in less H_2 available for methanogenesis and biohydrogenation of fat. The existence of host-genomically influenced microbial metabolic pathways that simultaneously affect fatty acid biohydrogenation and methanogenesis is supported by a divergent selection experiment for methane emissions in sheep, in which it was observed that the low-methane yield line had greater levels of fatty acids associated with the early stages of rumen biohydrogenation, such as CLA (Hervás et al., 2022). Another potential advantage of breeding on microbiome profiles is that unfavorable genomic covariances between productive traits, if not extreme, may not necessarily be reflected in the functional microbiome's overall complexity. Some microbial activities may be found with beneficial associations to both productive traits. Favoring these microbial metabolic functions by selection could potentially be used to overcome the unfavorable covariances, as has been shown for different growth rates at different growth stages in beef cattle (Martínez-Álvaro et al., 2022b). However, these assumptions are again based on variance component estimates with large estimation errors and would need to be confirmed using larger databases and independent populations.

An issue yet to be explored in developing microbiome-based selection criteria is the best time to measure the microbiome composition for selection, as h^2 may fluctuate with the age of the host. For example, the rumen microbiome is more diverse in the first weeks of life (O'Hara et al., 2020; Yáñez-Ruiz et al., 2015) and becomes relatively stable in the early stages of growth (after adaptation to a changing diet) until slaughter (Snelling et al., 2019). In this context, the recent study by Fan et al. (2021) has elucidated that the effects of host genetics on the rumen microbiota are not specific to particular growth stages – pre-weaning, postweaning, and fattening – but universal throughout the life of the ruminant, including the early life stages. Besides, similar microbiome composition h^2 have been estimated during the finishing period and at slaughter in beef (Martínez-Álvaro et al., 2022b), and during first lactation (Saborío-Montero et al., 2020) or between 10 and 40 days postpartum in dairy cattle (Cardinale and Kadarmideen, 2022; Wallace et al., 2019b). These results suggests that differences in the microbiota composition may prevail at different life stages, and age and timing of sampling may not be as critical as expected in an animal breeding framework, although further evidences on different ruminant species are needed to confirm this stability. Overall, the above studies suggest that the rumen microbiome is an extremely valuable source of information for selection in ruminants, facing some statistical, and operational challenges, but also offering a wide range of benefits. In some cases, selection based on microbiome profiles may be even more informative than using trait itself, avoiding the costly measurement of, for example, methane emissions (Martínez-Álvaro et al., 2022a).

The potential success of shaping the rumen microbiome by genetic selection is strongly supported by a selection experiment for the caecal microbiome in pigs over 2 generations. The selection criterion was defined based on the previous co-occurrence network study by Ramayo-Caldas et al. (2016), which revealed that two enterotypes defined by either *Prevotella* and *Mitsuokella* or *Ruminococcus* and *Treponema*, were largely associated with divergence in body weight and average daily gain of piglets. Direct selection on the abundance of these 4 microbial genera measured at 60 days of age, showed responses to selection on the abundance of the 4 genera ranging from 0.6 to 1.3 standard deviations in the first generation (h^2 of the four genera ranged from 0.3 to 0.4) and further increases in the second generation, along with correlated responses in other microbial genera and, more importantly, in average daily gain (Rogel-Gaillard et al., 2021). This experiment demonstrates that the fecal microbiota in pigs is sensitive to genetic selection and that the expected responses are consistent with the breeder equation. This is also expected in ruminants, as their fitness is highly dependent on their own microbiome (*e.g.*, ~70% of their energy requirements from volatile fatty acids arise after microbial fermentation, as opposed to 25% in pigs (Bergman, 1990), and exhibit a tight co-evolutionary pattern along history (Perlman et al., 2022). Interestingly, it has been observed in other studies that the correlated response among the microbiome composition and productive traits is bidirectional. For example, in two divergent selection experiments in rabbits, strong (up to 0.7 standard deviations) correlated responses to selection on caecal microbiome function (measured as microbial gene *alr*-transformed abundances) were observed after selection on intramuscular fat over 10 generations (Martínez-Álvaro et al., 2021) or on litter size variability over 13 generations (Casto-Rebollo et al., 2022). In pigs, alpha-diversity and abundance of 52 of the 75 microbial genera analysed were altered after 10 generations of selection for residual feed intake (Aliakbari et al., 2021). In small ruminants, divergent selection on residual feed intake in lambs showed a correlated response in the *clr*-transformed abundance of several genera within phyla *Lachnospiraceae* and *Prevotellaceae* (Marie-Etancelin et al., 2021). A correlated response on the taxonomic composition of the rumen microbiome remained unclear after selection for feeding behavior and for somatic cell score or milk persistence in ewes (Marie-Etancelin et al., 2021; Tortereau et al., 2020). However, the experiments in sheep are based on a single generation (Boggio et al., 2021; Rupp et al., 2009), and the results must be validated. The authors are not aware of similar selection experiments in cattle. The scientific literature may benefit from future specific experiments in these species.

6. The microbiome effect on methane emissions (Case study)

In ruminants, methane emissions are generated by methanogenic Archaea in the digestive tract, reducing the focus of researchers from the entire microbial community to this subset. Ciliate protozoa and fungi have also been associated with increased methane emissions (Guyader et al., 2014; Lopez-Garcia et al., 2022). Selection for a more efficient rumen microbiome requires some considerations along the process, from sampling to genetic evaluations, to inclusion in the selection indices. Sampling the microbiota is the first process to obtain a small portion of the microbial community of the animal for downstream analyses. Sampling can be done *in vivo* or immediately after the animal is slaughtered. The collection of samples at the slaughterhouse allows to obtain a more homogenous sample of the entire rumen content. It may be the preferred option in beef cattle after the fattening period, and also in small ruminants. However, in some breeds, such as dairy cattle, this can be done only late in life. In these cases, *in vivo* extraction through oral or nasopharyngeal tubes performed by oral intubation are recommended. The system usually consists of a mechanical pump unit connected to the other end of the tube and a covered Erlenmeyer trapped in between, in which the sample is collected. The latter technique is less invasive than fistulation and can be performed on larger scale of the populations, accepting the inconvenience of potential contamination with saliva (Shen et al., 2012). This procedure mainly yields liquid samples with some solid components that can be separated by filtration. In most sampling of rumen contents, the animal is partially immobilized

prior to collection to prevent physical harm to the animal, and in some cases, to personnel. Other techniques are possible but are not recommended because of the greater invasiveness. For instance, in fistulated ruminants, sampling consists of extracting the rumen content directly with a canula. Both liquid and solid samples can be collected with this procedure. Rumenocentesis consists of puncturing the abdominal wall with a needle connected to a syringe. These strategies limit the amount of sample that can be collected and are of concern from an animal welfare point of view and are not feasible in commercial herds (Duffield et al., 2004).

After collection of the rumen contents sample, it should be stored in a covered recipient to ensure an anaerobic environment as in the rumen, and to prevent alteration of the rumen community by its intrinsic susceptibility to oxygen (Chaucheyras-Durand and Ossa, 2014). The recipient should be transported in liquid N_2 containers until storage at -80 °C. The sample should be stored under these conditions until DNA extraction protocols can be performed. The extraction of the genetic material (DNA and RNA) is also important. There are several DNA extraction techniques for rumen content samples, either chemical or mechanical. Of these two approaches, mechanical extraction is preferred over chemical extraction, such as the bead beating method (Yu and Morrison, 2004), which provides higher DNA yield, better DNA recovery free of inhibitory substances, and a better representation of bacterial community structure compared to other techniques (Yuan et al., 2012). There are many commercial extraction kits available. Their choice has a great impact on the microbial community composition (Henderson et al., 2013). It is recommended to use kits that are specific for microbial DNA extraction and allow breaking the wall of all cells in the community to obtain a representative metagenome. It is well known that the wall of Archaea are more difficult to break and that they are often underrepresented in the metagenomes. However, there is a trade-off between breaking the cell wall and excessive shearing of DNA material. Researchers are encouraged to optimize this balance in their own labs.

Then, whole metagenome sequencing needs to be performed. Metataxonomic studies using amplicons are discouraged because they may bias the microbial composition due to PCR artifacts, and limited completeness of databases. Shotgun metagenomics with NGS is recommended to obtain a more representative composition of the rumen microbiota. Among the available platforms, Illumina MiSeq and Illumina NextSeq are the most commonly used (Auffret et al., 2018; Martínez-Álvaro et al., 2020; Ramayo-Caldas et al., 2020). Third-generation sequencing with Oxford Nanopore Technologies has also been used satisfactorily (Lopez-Garcia et al., 2022; Stewart et al., 2019a), with lower basecall accuracy but with better taxonomical assignment. Bioinformatics analyses are the next important step along the process. The mapping strategy and pipeline can have an important impact on the resulting community. There are numerous options described in the literature (*e.g.,* Lopez-Garcia et al. 2022, Ramayo-Caldas et al. 2020, Snelling et al. 2019). It is recommended to choose bioinformatics tools that use algorithms capable of correctly mapping a larger number of reads. The non-redundant NCBI database is considered the most comprehensive. Nonetheless, there are still many rumen microbes that do not map to known databases, and therefore it is often recommended to follow a MAG assembly strategy to complement the information in the databases (Stewart et al., 2019b).

Once the microbial composition has been obtained, and metagenomics data have been treated with an appropriate statistical treatment as described in this review, the breeder needs to decide what to retain for selection. This may be the abundance of one or a few microbial taxa or genes of interest, some aggregated variables as described in the previous sections, or the entire whole metagenome information. A sufficiently large number of individuals need to be phenotyped. Metagenomic information is not yet affordable enough for implementing large-scale phenotyping, but a reference population can be maintained. The size of this initial population depends on heritability, desired accuracy, and budget. Different circumstances in a general phenotyping scenario were evaluated in Gonzalez-Recio et al. (2014). A reference population allows all genotyped animals to be genetically evaluated regardless of their own metagenomic phenotype. Finally, genetic and phenotypic correlations between the metagenomic variable and the phenotypes under selection must be estimated, and optimal weight of this trait within a selection index needs to be calculated, either by classical selection index theory or by bioeconomic models, for their final inclusion in the breeding program.

7. Future perspectives

The rumen microbiota plays an important role in the production of methane in the rumen. The genome of the host partially controls the composition of the microbiota. Increasing our knowledge of the rumen microbiota and its interaction with the host is necessary to apply strategies that modulate the microbiota in the right direction. Studying and analyzing the microbiome requires overcoming some methodological challenges, which are highlighted in this review. Metagenomic information is still expensive to obtain, mainly due to the sampling and sequencing costs. It will be necessary to develop new technologies and strategies to incorporate metagenomic information into breeding programs at an affordable cost. This review highlights the lack of current literature on how to incorporate the microbiome information into breeding goals and selective breeding. Important efforts will be needed in this area in the following years.

Declaration of Competing Interest

The authors declare that they have no conflict of interest regarding the topics approached within this study.

References

Abbas, W., Howard, J.T., Paz, H.A., Hales, K.E., Wells, J.E., Kuehn, L.A., Erickson, G.E., Spangler, M.L., Fernando, S.C., 2020. Influence of host genetics in shaping the rumen bacterial community in beef cattle. Sci. Rep. 1 (10), 1–14. https://doi.org/10.1038/s41598-020-72011-9, 2020 10.

Aitchison, J., 1982. The statistical analysis of compositional data. J. R. Stat. Soc. Ser. B (Methodol.) 44, 139–160. https://doi.org/10.1111/J.2517-6161.1982.TB01195.X.

Akin, D.E., Borneman, W.S., 1990. Role of rumen fungi in fiber degradation. J. Dairy Sci. 73 (10), 3023–3032. https://doi.org/10.3168/jds.S0022-0302(90)78989-8.

Aliakbari, A., Zemb, O., Billon, Y., Barilly, C., Ahn, I., Riquet, J., Gilbert, H., 2021. Genetic relationships between feed efficiency and gut microbiome in pig lines selected for residual feed intake. J. Anim. Breed. Genet. 138, 491–507. https://doi.org/10.1111/JBG.12539.

Auffret, M.D., Stewart, R., Dewhurst, R.J., Duthie, C.A., Rooke, J.A., Wallace, R.J., Freeman, T.C., Snelling, T.J., Watson, M., Roehe, R., 2018. Identification, comparison, and validation of robust rumen microbial biomarkers for methane emissions using diverse bos taurus breeds and basal diets. Front. Microbiol. 8, 2642. https://doi.org/10.3389/fmicb.2017.02642.

Bergamaschi, M., Maltecca, C., Schillebeeckx, C., McNulty, N.P., Schwab, C., Shull, C., Fix, J., Tiezzi, F., 2020. Heritability and genome-wide association of swine gut microbiome features with growth and fatness parameters. Sci. Rep. 1 (10), 1–12. https://doi.org/10.1038/s41598-020-66791-3, 2020 10.

Bergman, E.N., 1990. Energy contributions of volatile fatty acids from the gastrointestinal tract in various species. 10.1152/physrev.1990.70.2.567 70, 567–590. 10.1152/PHYSREV.1990.70.2.567.

Boggio, G.M., Meynadier, A., Daunis-I-Estadella, P., Marie-Etancelin, C., 2021. Compositional analysis of ruminal bacteria from ewes selected for somatic cell score and milk persistency. PLoS One 16, e0254874. https://doi.org/10.1371/JOURNAL.PONE.0254874.

Camarinha-Silva, A., Maushammer, M., Wellmann, R., Vital, M., Preuss, S., Bennewitz, J., 2017. Host genome influence on gut microbial composition and microbial prediction of complex traits in pigs. Genetics 206, 1637–1644. https://doi.org/10.1534/genetics.117.200782.

Cardinale, S., Kadarmideen, H.N., 2022. Host genome-metagenome analyses using combinatorial network methods reveal key metagenomic and host genetic features for methane emission and feed efficiency in cattle. Front. Genet. 13 https://doi.org/10.3389/FGENE.2022.795717.

Casto-Rebollo, C., Argente, M.J., García, M.L., Pena, R.N., 2022. Selection for environmental variance shifted the gut microbiome composition driving animal resilience. 10.21203/rs.3.rs-1597523/v1.

Chaucheyras-Durand, F., Ossa, F., 2014. The rumen microbiome: composition, abundance, diversity, and new investigative tools. Prof. Anim. Sci. https://doi.org/10.15232/S1080-7446(15)30076-0.

Christensen, O.F., Börner, V., Varona, L., Legarra, A., 2021. Genetic evaluation including intermediate omics features. Genetics 219. https://doi.org/10.1093/GENETICS/IYAB130.

Crable, B.R., Plugge, C.M., McInerney, M.J., Stams, A.J.M., 2011. Formate formation and formate conversion in biological fuels production. Enzym. Res. 2011 https://doi.org/10.4061/2011/532536.

de Haas, Y., Windig, J.J., Calus, M.P.L., Dijkstra, J., de Haan, M., Bannink, A., Veerkamp, R.F., 2011. Genetic parameters for predicted methane production and potential for reducing enteric emissions through genomic selection. J. Dairy Sci. 94, 6122–6134. https://doi.org/10.3168/jds.2011-4439.

de los Campos, G., Hickey, J.M., Pong-Wong, R., Daetwyler, H.D., Calus, M.P.L., 2013. Whole-genome regression and prediction methods applied to plant and animal breeding. Genetics 193, 327–345. https://doi.org/10.1534/genetics.112.143313.

Difford, G.F., Plichta, D.R., Løvendahl, P., Lassen, J., Noel, S.J., Højberg, O., Wright, A.D. G., Zhu, Z., Kristensen, L., Nielsen, H.B., Guldbrandtsen, B., Sahana, G., 2018. Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows. PLos Genet. 14, e1007580 https://doi.org/10.1371/journal.pgen.1007580.

Duffield, T., Plaizier, J.C., Fairfield, A., Bagg, R., Vessie, G., Dick, P., Wilson, J., Aramini, J., McBride, B., 2004. Comparison of techniques for measurement of rumen pH in lactating dairy cows. J. Dairy Sci. 87, 59–66. https://doi.org/10.3168/jds.S0022-0302(04)73142-2.

Eaglen, S.A.E., Bijma, P., 2009. Genetic parameters of direct and maternal effects for calving ease in Dutch Holstein-Friesian cattle. J. Dairy Sci. 92, 2229–2237. https://doi.org/10.3168/JDS.2008-1654.

EPA, 2022. Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2020. U.S. Environmental Protection Agency. EPA 430-R-22-003. https://www.epa.gov/ghgemissions/draft-inventory-us-greenhouse-gas-emissionsand-sinks-1990-2020.

Fan, P., Bian, B., Teng, L., Nelson, C.D., Driver, J., Elzo, M.A., Jeong, K.C., 2019. Host genetic effects upon the early gut microbiota in a bovine model with graduated spectrum of genetic variation. ISME J. 1 (14), 302–317. https://doi.org/10.1038/s41396-019-0529-2, 2019 14.

Fan, P., Nelson, C.D., Driver, J.D., Elzo, M.A., Peñagaricano, F., Jeong, K.C., 2021. Host genetics exerts lifelong effects upon hindgut microbiota and its association with bovine growth and immunity. ISME J. 8 (15), 2306–2321. https://doi.org/10.1038/s41396-021-00925-x, 2021 15.

Gerber, P.J., Steinfeld, H., Henderson, B., Mottet, A., Opio, C., Dijkman, J., Falcucci, A., Tempio, G., 2013. Tackling Climate Change Through Livestock - A global Assessment of Emissions and Mitigation Opportunities. Food and Agriculture Organization of the United Nations (FAO), Rome.

Gloor, G.B., Macklaim, J.M., Pawlowsky-Glahn, V., Egozcue, J.J., 2017. Microbiome datasets are compositional: and this is not optional. Front. Microbiol. 8 https://doi.org/10.3389/FMICB.2017.02224.

González-Recio, O., López-Paredes, J., Ouatahar, L., Charfeddine, N., Ugarte, E., Alenda, R., Jiménez-Montero, J.A., 2020. Mitigation of greenhouse gases in dairy cattle via genetic selection: 2. Incorporating methane emissions into the breeding goal. J. Dairy Sci. 103, 7210–7221. https://doi.org/10.3168/JDS.2019-17598.

Gonzalez-Recio, O., Coffey, M.P., Pryce, J.E., 2014. On the value of the phenotypes in the genomic era. J. Dairy Sci. 97, 7905–7915. https://doi.org/10.3168/jds.2014-8125.

Gonzalez-Recio, O., Zubiria, I., García-Rodríguez, A., Hurtado, A., Atxaerandio, R., 2018. Short communication: signs of host genetic regulation in the microbiome composition in 2 dairy breeds: holstein and Brown Swiss. J. Dairy Sci. 101, 2285–2292. https://doi.org/10.3168/jds.2017-13179.

Gonzalez-Recio, O., Scrobota, N., López-Paredes, J., Saborío-Montero, A., Fernández, A., López de Maturana, E., Villanueva, B., Goiri, I., Atxaerandio, R., Rodríguez-García, A. Diving into the cow hologenome to reduce methane emissions and increase sustainability. Animal 2023 (under review).

Greenacre, M., Martínez-Álvaro, M., Blasco, A., 2021. Compositional data analysis of microbiome and any-omics datasets: a validation of the additive logratio transformation. Front. Microbiol. 12 https://doi.org/10.3389/FMICB.2021.727398.

Guyader, J., Eugène, M., Nozière, P., Morgavi, D.P., Doreau, M., Martin, C., 2014. Influence of rumen protozoa on methane emission in ruminants: a meta-analysis approach. Animal 8, 1816–1825. https://doi.org/10.1017/S1751731114001852.

Hayes, A.F., 2022. From Guilford Introduction to Mediation, Moderation, and Conditional Process Analysis AF2E 7006, 9–10.

Henderson, G., Cox, F., Kittelmann, S., Miri, V.H., Zethof, M., Noel, S.J., Waghorn, G.C., Janssen, P.H., 2013. Effect of DNA extraction methods and sampling techniques on the apparent structure of cow and sheep rumen microbial communities. PLoS One 8, e74787. https://doi.org/10.1371/journal.pone.0074787.

Henderson, G., Cox, F., Ganesh, S., Jonker, A., Young, W., Janssen, P.H., 2015. Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. Sci. Rep. 5, 14567. https://doi.org/10.1038/srep14567.

Hervás, G., Boussalia, Y., Labbouz, Y., della Badia, A., Toral, P.G., Frutos, P., 2022. Insect oils and chitosan in sheep feeding: effects on *in vitro* ruminal biohydrogenation and fermentation. Anim. Feed Sci. Technol. 285, 115222 https://doi.org/10.1016/J.ANIFEEDSCI.2022.115222.

Hungate, R.E., 1966. Rumen and Its Microbes. Elsevier Science.

Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J.H., Chinwalla, A.T., Creasy, H.H., Earl, A.M., Fitzgerald, M.G., Fulton, R.S., Giglio, M.G., Hallsworth-Pepin, K., Lobos, E.A., Madupu, R., Magrini, V., Martin, J.C., Mitreva, M., Muzny, D. M., Sodergren, E.J., Versalovic, J., Wollam, A.M., Worley, K.C., Wortman, J.R., Young, S.K., Zeng, Q., Aagaard, K.M., Abolude, O.O., Allen-Vercoe, E., Alm, E.J., Alvarado, L., Andersen, G.L., Anderson, S., Appelbaum, E., Arachchi, H.M., Armitage, G., Arze, C.A., Ayvaz, T., Baker, C.C., Begg, L., Belachew, T., Bhonagiri, V., Bihan, M., Blaser, M.J., Bloom, T., Bonazzi, V., Paul Brooks, J., Buck, G.A., Buhay, C.

J., Busam, D.A., Campbell, J.L., Canon, S.R., Cantarel, B.L., Chain, P.S.G., Chen, I.M. A., Chen, L., Chhibba, S., Chu, K., Ciulla, D.M., Clemente, J.C., Clifton, S.W., Conlan, S., Crabtree, J., Cutting, M.A., Davidovics, N.J., Davis, C.C., Desantis, T.Z., Deal, C., Delehaunty, K.D., Dewhirst, F.E., Deych, E., Ding, Y., Dooling, D.J., Dugan, S.P., Michael Dunne, W., Scott Durkin, A., Edgar, R.C., Erlich, R.L., Farmer, C.N., Farrell, R.M., Faust, K., Feldgarden, M., Felix, V.M., Fisher, S., Fodor, A.A., Forney, L.J., Foster, L., di Francesco, V., Friedman, J., Friedrich, D.C., Fronick, C.C., Fulton, L.L., Gao, H., Garcia, N., Giannoukos, G., Giblin, C., Giovanni, M.Y., Goldberg, J.M., Goll, J., Gonzalez, A., Griggs, A., Gujja, S., Kinder Haake, S., Haas, B.J., Hamilton, H.A., Harris, E.L., Hepburn, T.A., Herter, B., Hoffmann, D.E., Holder, M.E., Howarth, C., Huang, K.H., Huse, S.M., Izard, J., Jansson, J.K., Jiang, H., Jordan, C., Joshi, V., Katancik, J.A., Keitel, W.A., Kelley, S. T., Kells, C., King, N.B., Knights, D., Kong, H.H., Koren, O., Koren, S., Kota, K.C., Kovar, C.L., Kyrpides, N.C., la Rosa, P.S., Lee, S.L., Lemon, K.P., Lennon, N., Lewis, C. M., Lewis, L., Ley, R.E., Li, K., Liolios, K., Liu, B., Liu, Y., Lo, C.C., Lozupone, C.A., Dwayne Lunsford, R., Madden, T., Mahurkar, A.A., Mannon, P.J., Mardis, E.R., Markowitz, V.M., Mavromatis, K., McCorrison, J.M., McDonald, D., McEwen, J., McGuire, A.L., McInnes, P., Mehta, T., Mihindukulasuriya, K.A., Miller, J.R., Minx, P. J., Newsham, I., Nusbaum, C., Oglaughlin, M., Orvis, J., Pagani, I., Palaniappan, K., Patel, S.M., Pearson, M., Peterson, J., Podar, M., Pohl, C., Pollard, K.S., Pop, M., Priest, M.E., Proctor, L.M., Qin, X., Raes, J., Ravel, J., Reid, J.G., Rho, M., Rhodes, R., Riehle, K.P., Rivera, M.C., Rodriguez-Mueller, B., Rogers, Y.H., Ross, M.C., Russ, C., Sanka, R.K., Sankar, P., Fah Sathirapongsasuti, J., Schloss, J.A., Schloss, P.D., Schmidt, T.M., Scholz, M., Schriml, L., Schubert, A.M., Segata, N., Segre, J.A., Shannon, W.D., Sharp, R.R., Sharpton, T.J., Shenoy, N., Sheth, N.U., Simone, G.A., Singh, I., Smillie, C.S., Sobel, J.D., Sommer, D.D., Spicer, P., Sutton, G.G., Sykes, S. M., Tabbaa, D.G., Thiagarajan, M., Tomlinson, C.M., Torralba, M., Treangen, T.J., Truty, R.M., Vishnivetskaya, T.A., Walker, J., Wang, L., Wang, Z., Ward, D.v., Warren, W., Watson, M.A., Wellington, C., Wetterstrand, K.A., White, J.R., Wilczek-Boney, K., Wu, Y., Wylie, K.M., Wylie, T., Yandava, C., Ye, L., Ye, Y., Yooseph, S., Youmans, B.P., Zhang, L., Zhou, Y., Zhu, Y., Zoloth, L., Zucker, J.D., Birren, B.W., Gibbs, R.A., Highlander, S.K., Methé, B.A., Nelson, K.E., Petrosino, J.F., Weinstock, G.M., Wilson, R.K., White, O., 2012. Structure, function and diversity of the healthy human microbiome. Nature 7402 (486), 207–214. https://doi.org/10.1038/nature11234, 2012 486.

Jiang, Y., Reif, J.C., 2015. Modeling epistasis in genomic selection. Genetics 201, 759–768. https://doi.org/10.1534/GENETICS.115.177907.

Johnson, K.A., Johnson, D.E., 1995. Methane emissions from cattle. J. Anim. Sci. 73, 2483–2492. https://doi.org/10.2527/1995.7382483X.

Khanal, P., Maltecca, C., Schwab, C., Gray, K., Tiezzi, F., 2019. Genetic parameters of meat quality, carcass composition, and growth traits in commercial swine. J. Anim. Sci. 97, 3669–3683. https://doi.org/10.1093/JAS/SKZ247.

Khanal, P., Maltecca, C., Schwab, C., Fix, J., Bergamaschi, M., Tiezzi, F., 2020. Modeling host-microbiome interactions for the prediction of meat quality and carcass composition traits in swine. Genet. Sel. Evol. 52, 1–13. https://doi.org/10.1186/S12711-020-00561-7/TABLES/1.

Khanal, P., Maltecca, C., Schwab, C., Fix, J., Tiezzi, F., 2021. Microbiability of meat quality and carcass composition traits in swine. J. Anim. Breed. Genet. 138, 223–236. https://doi.org/10.1111/JBG.12504.

King, E.E., Smith, R.P., St-Pierre, B., Wright, A.D.G., 2011. Differences in the rumen methanogen populations of lactating jersey and holstein dairy cows under the same diet regimen. Appl. Environ. Microbiol. 77, 5682. https://doi.org/10.1128/AEM.05130-11.

López-Paredes, J., Goiri, I., Atxaerandio, R., García-Rodríguez, A., Ugarte, E., Jiménez-Montero, J.A., Alenda, R., González-Recio, O., 2020. Mitigation of greenhouse gases in dairy cattle via genetic selection: 1. Genetic parameters of direct methane using noninvasive methods and proxies of methane. J. Dairy Sci. 103, 7199–7209. https://doi.org/10.3168/JDS.2019-17597.

Lassen, J., Løvendahl, P., 2016. Heritability estimates for enteric methane emissions from Holstein cattle measured using noninvasive methods. J. Dairy Sci. 99, 1959–1967. https://doi.org/10.3168/jds.2015-10012.

Li, F., Li, C., Chen, Y., Liu, J., Zhang, C., Irving, B., Fitzsimmons, C., Plastow, G., Guan, L. L., 2019. Host genetics influence the rumen microbiota and heritable rumen microbial features associate with feed efficiency in cattle. Microbiome 7, 92. https://doi.org/10.1186/s40168-019-0699-1.

Lopez-Garcia, A., Saborio-Montero, A., Gutierrez-Rivas, M., Atxaerandio, R., Goiri, I., Garcia-Rodriguez, A., Jimenez-Montero, J.A., Gonzalez, C., Tamames, J., Puente-Sanchez, F., Serrano, M., Carrasco, R., Ovilo, C., Gonzalez-Recio, O., 2022. Fungal and ciliate protozoa are the main rumen microbes associated with methane emissions in dairy cattle. Gigascience 11, 1–14. https://doi.org/10.1093/GIGASCIENCE/GIAB088.

Lu, J., Lu, L., Yu, Y., Cluette-Brown, J., Martin, C.R., Claud, E.C., 2018. Effects of intestinal microbiota on brain development in humanized gnotobiotic mice. Sci. Rep. 1 (8), 1–16. https://doi.org/10.1038/s41598-018-23692-w, 2018 8.

Manzanilla-Pech, C.I.V., Stephansen, R.B., Difford, G.F., Løvendahl, P., Lassen, J., 2022. Selecting for feed efficient cows will help to reduce methane gas emissions. Front. Genet. 13, 885932 https://doi.org/10.3389/fgene.2022.885932.

Margulis, L., & Fester, R. (1991). Bellagio conference and book. Symbiosis as Source of Evolutionary Innovation: Speciation and Morphogenesis. Conference–June 25-30, 1989, Bellagio Conference Center, Italy. Symbiosis (Philadelphia, PA), 11, 93–101.

Marie-Etancelin, C., Tortereau, F., Gabinaud, B., Martinez Boggio, G., le Graverand, Q., Marcon, D., de Almeida, M.L., Pascal, G., Weisbecker, J.L., Meynadier, A., 2021. Apart from the diet, the ruminal microbiota of lambs is modified in relation to their genetic potential for feed efficiency or feeding behavior. Front. Microbiol. 12 https://doi.org/10.3389/FMICB.2021.759432.

Martín-Fernández, J.A., Thió-Henestrosa, S., 2006. Rounded zeros: some practical aspects for compositional data. Geol. Soc. Spec. Publ. 264, 191–201. https://doi.org/10.1144/GSL.SP.2006.264.01.14.

Martínez-Álvaro, M., Auffret, M.D., Stewart, R.D., Dewhurst, R.J., Duthie, C.A., Rooke, J.A., Wallace, R.J., Shih, B., Freeman, T.C., Watson, M., Roehe, R., 2020. Identification of complex rumen microbiome interaction within diverse functional niches as mechanisms affecting the variation of methane emissions in bovine. Front. Microbiol. 11 https://doi.org/10.3389/FMICB.2020.00659.

Martínez-Álvaro, M., Zubiri-Gaitán, A., Hernández, P., Greenacre, M., Ferrer, A., Blasco, A., 2021. Comprehensive functional core microbiome comparison in genetically obese and lean hosts under the same environment. Commun. Biol. 1 (4), 1–10. https://doi.org/10.1038/s42003-021-02784-w, 2021 4.

Martínez-Álvaro, M., Auffret, M.D., Duthie, C.A., Dewhurst, R.J., Cleveland, M.A., Watson, M., Roehe, R., 2022a. Bovine host genome acts on rumen microbiome function linked to methane emissions. Commun. Biol. 5 https://doi.org/10.1038/s42003-022-03293-0.

Martínez-Álvaro, M., Mattock, J., González-Recio, O., Saborío-Montero, A., Weng, Z., Lima, J., Duthie, C.A., Dewhurst, R., Cleveland, J., Watson, M., Roehe, R., 2022 Microbiome-driven breeding strategies to increase growth and performance at different stages in beef. British Society of Animal Science (BSAS) Conference. 12-14 April 2022, Nottingham (UK).

Mizrahi, I., Wallace, R.J., Moraïs, S., 2021. The rumen microbiome: balancing food security and environmental impacts. Nat. Rev. Microbiol. 9 (19), 553–566. https://doi.org/10.1038/s41579-021-00543-6, 2021 19.

Moss, A.R., Jouany, J.P., Newbold, J., 2000. Methane production by ruminants: its contribution to global warming. Ann. Zootech. 49, 231–253. https://doi.org/10.1051/ANIMRES:2000119.

Newbold, C.J., de la Fuente, G., Belanche, A., Ramos-Morales, E., McEwan, N.R., 2015. The role of ciliate protozoa in the rumen. Front. Microbiol. 6 https://doi.org/10.3389/FMICB.2015.01313.

NOAA, 2022. Increase in Atmospheric Methane Set Another Record During 2021. NOAA. https://www.noaa.gov/news-release/increase-in-atmospheric-methane-set-another-record-during-2021. accessed September 23rd 2022.

O'Hara, E., Kenny, D.A., McGovern, E., Byrne, C.J., McCabe, M.S., Guan, L.L., Waters, S. M., 2020. Investigating temporal microbial dynamics in the rumen of beef calves raised on two farms during early life. FEMS Microbiol. Ecol. 96, 203. https://doi.org/10.1093/FEMSEC/FIZ203.

Palarea-Albaladejo, J., Martín-Fernández, J.A., 2015. ZCompositions - R package for multivariate imputation of left-censored data under a compositional approach. Chemom. Intell. Lab. Syst. 143, 85–96. https://doi.org/10.1016/J.CHEMOLAB.2015.02.019.

Perlman, D., Martínez-Álvaro, M., Moraïs, S., Altshuler, I., Hagen, L.H., Jami, E., Roehe, R., Pope, P.B., Mizrahi, I., 2022. Concepts and consequences of a core gut microbiota for animal growth and development. Annu. Rev. Anim. Biosci. 10, 177–201. https://doi.org/10.1146/ANNUREV-ANIMAL-013020-020412.

Rönnegard, L., Felleki, M., Fikse, W.F., Mulder, H.A., Strandberg, E., 2013. Variance component and breeding value estimation for genetic heterogeneity of residual variance in Swedish Holstein dairy cattle. J. Dairy Sci. 96, 2627–2636. https://doi.org/10.3168/JDS.2012-6198.

Ramayo-Caldas, Y., Mach, N., Lepage, P., Levenez, F., Denis, C., Lemonnier, G., Leplat, J. J., Billon, Y., Berri, M., Doré, J., Rogel-Gaillard, C., Estellé, J., 2016. Phylogenetic network analysis applied to pig gut microbiota identifies an ecosystem structure linked with growth traits. ISME J. 10, 2973–2977. https://doi.org/10.1038/ISMEJ.2016.77.

Ramayo-Caldas, Y., Zingaretti, L.M., Pérez-Pascual, D., Alexandre, P.A., Reverter, A., Dalmau, A., Quintanilla, R., Ballester, M., 2021. Leveraging host-genetics and gut microbiota to determine immunocompetence in pigs. Anim. Microbiome 3, 74. https://doi.org/10.1186/s42523-021-00138-9.

Ramayo-Caldas, Y., Zingaretti, L., Popova, M., Estellé, J., Bernard, A., Pons, N., Bellot, P., Mach, N., Rau, A., Roume, H., Perez-Enciso, M., Faverdin, P., Edouard, N., Ehrlich, D., Morgavi, D.P., Renand, G., 2020. Identification of rumen microbial biomarkers linked to methane emission in Holstein dairy cows. J. Anim. Breed. Genet. 137, 49–59. https://doi.org/10.1111/JBG.12427.

Ramette, A., 2007. Multivariate analyses in microbial ecology. FEMS Microbiol. Ecol. 62, 142–160. https://doi.org/10.1111/j.1574-6941.2007.00375.x.

Roehe, R., Dewhurst, R.J., Duthie, C.A., Rooke, J.A., McKain, N., Ross, D.W., Hyslop, J.J., Waterhouse, A., Freeman, T.C., Watson, M., Wallace, R.J., 2016. 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.

Rogel-Gaillard, C., Estellé, J., Borey, M., Larzul, C., 2021. Interplay between host genetics and the dynamics of the gut microbita in pigs. In: Proceedings of the 72nd Annual Meetig of EAAP. Davos. Eauropean Association of Animal Production, 568–568.

Ross, E.M., Moate, P.J., Marett, L.C., Cocks, B.G., Hayes, B.J., 2013. Metagenomic predictions: from microbiome to complex health and environmental phenotypes in humans and cattle. PLoS One 8, e73056. https://doi.org/10.1371/journal.pone.0073056.

Rupp, R., Bergonier, D., Dion, S., Hygonenq, M.C., Aurel, M.R., Robert-Granié, C., Foucras, G., 2009. Response to somatic cell count-based selection for mastitis resistance in a divergent selection experiment in sheep. J. Dairy Sci. 92, 1203–1219. https://doi.org/10.3168/jds.2008-1435.

Saborío-Montero, A., Gutiérrez-Rivas, M., García-Rodríguez, A., Atxaerandio, R., Goiri, I., López de Maturana, E., Jiménez-Montero, J.A., Alenda, R., González-Recio, O., 2020. Structural equation models to disentangle the biological relationship between microbiota and complex traits: methane production in dairy cattle as a case of study. J. Anim. Breed. Genet. 137, 36–48. https://doi.org/10.1111/JBG.12444.

Saborío-Montero, A., Gutiérrez-Rivas, M., López-García, A., García-Rodríguez, A., Atxaerandio, R., Goiri, I., Jiménez-Montero, J.A., González-Recio, O., 2021a. Holobiont effect accounts for more methane emission variance than the additive and microbiome effects on dairy cattle. Livest. Sci. 250, 104538 https://doi.org/10.1016/J.LIVSCI.2021.104538.

Saborío-Montero, A., López-García, A., Gutiérrez-Rivas, M., Atxaerandio, R., Goiri, I., García-Rodriguez, A., Jiménez-Montero, J.A., González, C., Tamames, J., Puente-Sánchez, F., Varona, L., Serrano, M., Ovilo, C., González-Recio, O., 2021b. A dimensional reduction approach to modulate the core ruminal microbiome associated with methane emissions via selective breeding. J. Dairy Sci. 104, 8135–8151. https://doi.org/10.3168/JDS.2020-20005.

Sandoval-Motta, S., Aldana, M., Mart?nez-Romero, E., Frank, A., 2017. The human microbiome and the missing heritability problem. Front. Genet. 8, 80. https://doi.org/10.3389/fgene.2017.00080.

Schaeffer, L.R., 2004. Application of random regression models in animal breeding. Livest. Prod. Sci. 86, 35–45. https://doi.org/10.1016/S0301-6226(03)00151-9.

Shen, J.S., Chai, Z., Song, L.J., Liu, J.X., Wu, Y.M., 2012. Insertion depth of oral stomach tubes may affect the fermentation parameters of ruminal fluid collected in dairy cows1. J. Dairy Sci. 95, 5978–5984. https://doi.org/10.3168/jds.2012-5499.

Simon, JC., Marchesi, J.R., Mougel, C., et al., 2019. Host-microbiota interactions: from holobiont theory to analysis. Microbiome 7, 5. https://doi.org/10.1186/s40168-019-0619-4.

Snelling, T.J., Auffret, M.D., Duthie, C.A., Stewart, R.D., Watson, M., Dewhurst, R.J., Roehe, R., Walker, A.W., 2019. Temporal stability of the rumen microbiota in beef cattle, and response to diet and supplements. Anim. Microbiome 1 (1), 1–14. https://doi.org/10.1186/S42523-019-0018-Y, 2019 1.

Stewart, R.D., Auffret, M.D., Warr, A., Walker, A.W., Roehe, R., Watson, M., 2019. Compendium of 4,941 rumen metagenome-assembled genomes for rumen microbiome biology and enzyme discovery. Nat. Biotechnol. 37, 953–961. https://doi.org/10.1038/S41587-019-0202-3.

ten Hoopen, P., Finn, R.D., Bongo, L.A., Corre, E., Fosso, B., Meyer, F., Mitchell, A., Pelletier, E., Pesole, G., Santamaria, M., Willassen, N.P., Cochrane, G., 2017. The metagenomic data life-cycle: standards and best practices. Gigascience 6, 1–11. https://doi.org/10.1093/GIGASCIENCE/GIX047.

Tiezzi, F., Fix, J., Schwab, C., Shull, C., Maltecca, C., 2021. Gut microbiome mediates host genomic effects on phenotypes: a case study with fat deposition in pigs. Comput. Struct. Biotechnol. J. 19, 530–544. https://doi.org/10.1016/J.CSBJ.2020.12.038.

Tortereau, F., Marie-Etancelin, C., Weisbecker, J.L., Marcon, D., Bouvier, F., Moreno-Romieux, C., François, D., 2020. Genetic parameters for feed efficiency in Romane rams and responses to single-generation selection. Animal 14, 681–687. https://doi.org/10.1017/S1751731119002544.

Valente, B.D., Rosa, G.J.M., de Los Campos, G., Gianola, D., Silva, M.A., 2010. Searching for recursive causal structures in multivariate quantitative genetics mixed models. Genetics 185, 633–644. https://doi.org/10.1534/genetics.109.112979.

van den Boogaart, K.G., Tolosana-Delgado, R., 2013. Analyzing Compositional Data with R. Springer, Berlin, Heidelbergpp, pp. 1–258.

Varona, L., Sorensen, D., Thompson, R., 2007. Analysis of litter size and average litter weight in pigs using a recursive model. Genetics 177, 1791–1799. https://doi.org/10.1534/GENETICS.107.077818.

Wallace, R.J., Rooke, J.A., McKain, N., Duthie, C.A., Hyslop, J.J., Ross, D.W., Waterhouse, A., Watson, M., Roehe, R., 2015. The rumen microbial metagenome associated with high methane production in cattle. BMC Genom. 16 https://doi.org/10.1186/S12864-015-2032-0 [Electronic Resource].

Wallace, J.R., Sasson, G., Garnsworthy, P.C., Tapio, I., Gregson, E., Bani, P., Huhtanen, P., Bayat, A.R., Strozzi, F., Biscarini, F., Snelling, T.J., Saunders, N., Potterton, S.L., Craigon, J., Minuti, A., Trevisi, E., Callegari, M.L., Cappelli, F.P., Cabezas-Garcia, E.H., Vilkki, J., Pinares-Patino, C., Fliegerová, K.O., Mrázek, J., Sechovcová, H., Kopečný, J., Bonin, A., Boyer, F., Taberlet, P., Kokou, F., Halperin, E., Williams, J.L., Shingfield, K.J., Mizrahi, I., 2019a. A heritable subset of the core rumen microbiome dictates dairy cow productivity and emissions. Sci. Adv. 5 https://doi.org/10.1126/SCIADV.AAV8391.

Wallace, J.R., Sasson, G., Garnsworthy, P.C., Tapio, I., Gregson, E., Bani, P., Huhtanen, P., Bayat, A.R., Strozzi, F., Biscarini, F., Snelling, T.J., Saunders, N., Potterton, S.L., Craigon, J., Minuti, A., Trevisi, E., Callegari, M.L., Cappelli, F.P., Cabezas-Garcia, E.H., Vilkki, J., Pinares-Patino, C., Fliegerová, K.O., Mrázek, J., Sechovcová, H., Kopečný, J., Bonin, A., Boyer, F., Taberlet, P., Kokou, F., Halperin, E., Williams, J.L., Shingfield, K.J., Mizrahi, I., 2019b. A heritable subset of the core rumen microbiome dictates dairy cow productivity and emissions. Sci. Adv. 5 https://doi.org/10.1126/sciadv.aav8391.

Weimer, P.J., Stevenson, D.M., Mantovani, H.C., Man, S.L.C., 2010. Host specificity of the ruminal bacterial community in the dairy cow following near-total exchange of ruminal contents. J. Dairy Sci. 93, 5902–5912. https://doi.org/10.3168/jds.2010-3500.

Xue, M.Y., Sun, H.Z., Wu, X.H., Liu, J.X., Guan, L.L., 2020. Multi-omics reveals that the rumen microbiome and its metabolome together with the host metabolome contribute to individualized dairy cow performance. Microbiome 8. https://doi.org/10.1186/S40168-020-00819-8.

Yáñez-Ruiz, D.R., Abecia, L., Newbold, C.J., 2015. Manipulating rumen microbiome and fermentation through interventions during early life: a review. Front. Microbiol. 6, 1133. https://doi.org/10.3389/FMICB.2015.01133/BIBTEX.

Yu, Z., Morrison, M., 2004. Improved extraction of PCR-quality community DNA from digesta and fecal samples. BioTechniques 36, 808–812. https://doi.org/10.2144/04365ST04/ASSET/IMAGES/LARGE/FIGURE2.JPEG.

Yuan, S., Cohen, D.B., Ravel, J., Abdo, Z., Forney, L.J., 2012. Evaluation of methods for the extraction and purification of DNA from the human microbiome. PLoS One 7. https://doi.org/10.1371/JOURNAL.PONE.0033865.

Zhang, Q., Difford, G., Sahana, G., Løvendahl, P., Lassen, J., Lund, M.S., Guldbrandtsen, B., Janss, L., 2020. Bayesian modeling reveals host genetics associated with rumen microbiota jointly influence methane emission in dairy cows. ISME J. 8 (14), 2019–2033. https://doi.org/10.1038/s41396-020-0663-x, 2020 14.

Zhou, M., Hernandez-Sanabria, E., Guan, L.L., 2009. Assessment of the microbial ecology of ruminal methanogens in cattle with different feed efficiencies. Appl. Environ. Microbiol. 75, 6524–6533. https://doi.org/10.1128/AEM.02815-08.