Microbiability of milk composition and genetic control of microbiota effects in sheep

Guillermo Martinez Boggio, O.F. Christensen, Andres Legarra, Annabelle Meynadier, Christel Marie-Etancelin

J. Dairy Sci. 106
https://doi.org/10.3168/jds.2022-22948

Microbiability of milk composition and genetic control of microbiota effects in sheep

G. Martinez Boggio,[1]* O. F. Christensen,[2] A. Legarra,[1] A. Meynadier,[1] and C. Marie-Etancelin[1]*
[1]GenPhySE - Université de Toulouse - INRAE - ENVT, 31326, Castanet-Tolosan, France
[2]Center for Quantitative Genetics and Genomics, Aarhus University, DK-8000 Aarhus C, Denmark

ABSTRACT

Recently, high-dimensional omics data are becoming available in larger quantities, and models have been developed that integrate them with genomics to understand in finer detail the relationship between genotype and phenotype, and thus improve the performance of genetic evaluations. Our objectives are to quantify the effect of the inclusion of microbiome data in the genetic evaluation for dairy traits in sheep, through the estimation of the heritability, microbiability, and how the microbiome effect on dairy traits decomposes into genetic and nongenetic parts. In this study we analyzed milk and rumen samples of 795 Lacaune dairy ewes. We included, as phenotype, dairy traits and milk fatty acids and proteins composition; as omics measurements, 16S rRNA rumen bacterial abundances; and as genotyping, 54K SNP chip for all ewes. Two nested genomic models were used: a first model to predict the individual contributions of the genetic and microbial abundances to phenotypes, and a second model to predict the additive genetic effect of the microbial community. In addition, microbiome-wide association studies for all dairy traits were applied using the 2,059 rumen bacterial abundances, and the genetic correlations between microbiome principal components and dairy traits were estimated. Results showed that in general the inclusion of both genetic and microbiome effect did not improve the fit of the model compared with the model with the genetic effect only. In addition, for all dairy traits the total heritability was equal to the direct heritability after fitting microbiota effects, due to a microbiability being almost zero for most dairy traits and heritability of the microbial community was very close to zero. Microbiome-wide association studies did not show operational taxonomic units with major effect for any of the dairy traits evaluated, and the genetic correlations between the first 5 principal components and dairy traits were low to moderate. So far, we can conclude that, using a substantial data set of 795 Lacaune dairy ewes, rumen bacterial abundances do not provide improved genetic evaluation for dairy traits in sheep.

Key words: rumen microbiota, heritability, microbiability, microbiome-wide association studies, dairy sheep

INTRODUCTION

In plant and animal breeding, it has recently become possible to obtain high-dimensional omics data, such as metabolites, gene expression, proteins, microbial genes, in larger quantities, in addition to the already available genotypes. This has enabled the integration of different types of data to uncover the genotype-phenotype relationship (Morgante et al., 2020), which will be beneficial for the development of an optimized breeding strategy to improve complex traits (Fernie and Schauer, 2009; Guo et al., 2016).

However, to use omics information in genomic evaluations, there must be methods and models available that can account for these high-dimensional data. To this end, first, Hayes et al. (2017) used a multitrait genetic model including as correlated traits near infrared and nuclear magnetic resonance-derived phenotypes. Then, Weishaar et al. (2020) used 2 models: a microbial model to describe the microbiome effect on phenotypes (i.e., the microbiability defined by Difford et al. (2018) for rumen microbiome in cattle), and a genetic model for microbial abundances, which quantify the host genetic control on the microbiota. The genetic effect on the trait, not mediated by the microbiome, is included in a final step in a selection index. However, in this regard, Christensen et al. (2021) proposed a joint model including the genetic effect not mediated by the microbiome in the first model. This joint model was proposed for intermediate traits, which are those traits between the DNA action and phenotype expression. Although microbiome data are not strictly an intermediate trait

Received October 26, 2022.
Accepted February 28, 2023.
*Corresponding authors: guillermo.martinezboggio@gmail.com and christel.marie-etancelin@inrae.fr

from a molecular point of view, in practice, it is a source of information between genotype and phenotype.

In ruminants, the rumen microbiome is crucial to digest plant fiber and is closely associated to productive traits (Difford et al., 2018; Xue et al., 2018; Matthews et al., 2019). Particularly in dairy ruminants, the rumen fermentation products, such as VFA, and the microbial AA are used by the mammary gland to produce milk with a high solid content which allows the production of high-quality cheese. Our previous results showed that rumen bacterial abundances are controlled by host genetics in Lacaune dairy ewes (Martinez Boggio et al., 2022) and that fine milk composition is linked with particular rumen bacteria (Martinez Boggio et al., 2021). So, we hypothesized that there is an indirect genetic effect on dairy traits mediated by the bacterial abundances in the ewe's rumen. Therefore, we propose to apply the method defined by Christensen et al. (2021) on a Lacaune dairy sheep dataset including 16S rRNA bacterial abundances, and fine milk composition traits as phenotypes, aiming at quantifying the effect of the inclusion of the rumen microbial abundance in the genetic evaluation of dairy traits, estimating model parameters such as the direct heritability (not mediated by the microbiome; h_d^2), microbiability $\left(c_m^2\right)$, and the heritability of the predicted microbiome effect $\left(h_m^2\right)$, which amounts to a general heritability of the microbial community (Christensen et al., 2021). In addition, to evaluate the link between rumen microbial structure and milk composition traits, we propose to aggregate the rumen microbiota into a few principal components (**PrC**), and for each of these to estimate the heritability (h^2) and genetic correlations with dairy traits.

MATERIALS AND METHODS

The study was conducted at the INRAE Experimental Unit of La Fage (UE 321 agreement A312031, Roquefort, France) with a protocol for rumen sampling approved by the French Ministry of Higher Education, Research and Innovation – Animal Ethics Committee (approval number: APAFIS#6292–2016080214271984 v8).

Animals' Phenotyping

Data of dairy traits and rumen bacterial abundance from 795 multiparous Lacaune dairy ewes were collected from 2015 to 2019. For details, see Martinez Boggio et al. (2023). The data consisted of milk yield (**MY**) records and midinfrared (**MIR**) spectra predictions of the milk fat content (**FC**) and protein content (**PC**) for 795 ewes, and the fine profile of daily milk fatty acids (**FA**) and proteins for 563 ewes. The accuracies of the predictive equations of ewe milk FA and proteins were retrieved from Ferrand-Calmels et al. (2014) and Ferrand et al. (2012), respectively. The coefficients of determination for the FA retained for this study, were higher than 0.91, except for *c*9*t*11 C18:2 and C18:3n-3 with values of 0.74. Meanwhile, for caseins, the coefficients of determination were higher than 0.82, and equal to 0.77 and 0.26, for β-LG and α-LA, respectively. The FA predicted by MIR were SFA, such as butyric acid (C4:0), caproic acid (C6:0), caprylic acid (C8:0), capric acid (C10:0), lauric acid (C12:0), and palmitic acid (C16:0), and UFA, such as oleic acid (*c*9 C18:1), rumenic acid (*c*9*t*11 C18:2) and α-linolenic acid (C18:3n-3). The proteins predicted by MIR were caseins, namely α_{S1}-CN, α_{S2}-CN, β-CN, and κ-CN, and 2 whey proteins, namely α-LA and β-LG.

Rumen Sampling

Rumen sampling was performed in morning or afternoon within 3 d around the official milk recording of the flock. Ruminal contents were sampled from each ewe using a vacuum pump and a medical gastric tube. After extraction, the DNA strands of the 795 samples were sequenced using V3-V4 region of 16S rRNA gene with Illumina MiSeq technology at the Genomic and Transcriptomic Platform (INRAE, Toulouse, France). More details on rumen sampling, DNA extraction, amplicon sequencing and bioinformatic process are provided in Martinez Boggio et al. (2023). We obtained an abundance table with 2,059 operational taxonomic units (**OTU**) as result of the bioinformatic process of DNA sequences using FROGS 3.0 pipeline (Escudié et al., 2018). Note that, given the Swarm clustering (Mahé et al., 2015), chimera removal and abundance filtering applied, OTU obtained from FROGS can be considered as amplicon sequencing variants.

Genotyping

Among the 795 ewes, 743 were genotyped using Illumina Ovine SNP50 BeadChip (54,241 SNPs), and 52 ewes were genotyped with Illumina Ovine SNP15 (16,681 SNPs) followed by imputation to a medium-density SNP chip as part of the Lacaune dairy sheep genomic selection program (Larroque et al., 2017). Genotypes were subjected to quality control, including minimum call rates of 90% for SNPs and 95% for individuals and exclusion of SNPs with a minor allele frequency lower than 5%. The final data set included 773 genotyped individuals and 35,492 autosomal SNPs.

Statistical Analyses

The methodology applied for inference with complete omics data was the one proposed by Christensen et al. (2021), where a joint model for omics and phenotypes was proposed, and a method for prediction of breeding values developed. The method consisted of the following 2 steps: (1) prediction of the individual contributions of the genetic and microbiome to phenotypes, and (2) prediction of the additive genetic effect of the microbial community. Christensen et al. (2021) named the first genetic effect as the "residual additive genetic effect," but here we prefer to name it the "direct genetic effect," as it is the genetic effect directly influencing the phenotype (i.e., not mediated by the microbiome). Similarly, we here name the additive genetic effect of the microbial community as "indirect genetic effect," consistent with terminology in Bijma (2014) in the sense that it is the sum of indirect genetic effects of bacteria in the rumen. As an additional analysis, in the first step, we compared several mixed models to assess the relevance of including the microbiome effect to the genetic one, as well as the additive genetic and microbiome interaction effect. The individual dairy traits were analyzed separately using single trait models.

Step 1. The first step consisted on fitting a single trait model including the direct genetic effect not mediated by the microbiome and the microbiome effect (Equation [1]):

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Z}_1\boldsymbol{a}_d + \boldsymbol{Z}_2\boldsymbol{m} + \boldsymbol{e}, \qquad [1]$$

where $\boldsymbol{y}$ is the vector of observations, $\boldsymbol{b}$ is the vector of fixed effects, $\boldsymbol{a}_d$ is the vector of random direct (not mediated by the microbiome) additive genetic effects, $\boldsymbol{m}$ is the vector of individual effects due to microbiome, and $\boldsymbol{e}$ is the vector of random residual effects, $\boldsymbol{X}$ is the incidence matrix for $\boldsymbol{b}$, and $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ are incidence matrices for $\boldsymbol{a}_d$ and $\boldsymbol{m}$, respectively. The distributional assumptions are $\boldsymbol{a}_d \sim N\left(0, \boldsymbol{G}\sigma^2_{a_d}\right)$, $\boldsymbol{m} \sim N\left(0, \boldsymbol{O}\sigma^2_m\right)$, $\boldsymbol{e} \sim N\left(0, \boldsymbol{I}\sigma^2_e\right)$, where $\boldsymbol{G}$ is a genomic relationship matrix computed based on the first method proposed by VanRaden (2008), $\boldsymbol{O}$ is the microbial similarity matrix, and $\boldsymbol{I}$ is an identity matrix; $\sigma^2_{a_d}$ is the direct additive genetic variance, σ^2_m is the microbiome effect variance, and σ^2_e is the residual variance. The microbiability was estimated as $c^2_m = \sigma^2_m / \sigma^2_y$, and the direct heritability as $h^2_d = \sigma^2_{a_d} / \sigma^2_y$ (h^2_d was called h^2_r by Christensen et al., 2021), where σ^2_y is the phenotypic variance.

The fixed effects considered in $\boldsymbol{b}$ (Equation [1]) were defined by ANOVA test with $P < 0.05$. For all dairy traits, we included the effects of DIM (28–133 DIM) as a covariate and the sampling year (with 5 levels: 2015 to 2019). For FC and milk FA, we included the number of lactations nested in the sampling year (with 7 levels: in 2015, ewes in second, third, or fourth and more lactations, and in 2016 to 2019 ewes in second lactation only), and for PC, milk proteins and FA, we also included litter size, with 2 levels (1, or 2 and more lambs).

The microbial similarity matrix ($\boldsymbol{O}$) for the 795 ewes was computed based on the rumen bacterial abundance similarities between animals using the method proposed by Ross et al. (2013):

$$\boldsymbol{O} = \frac{\boldsymbol{M}\boldsymbol{M}'}{n}.$$

The matrix was computed as a variance-covariance matrix from rumen bacterial abundances as where $\boldsymbol{M}$ is the abundance matrix with $n = 2{,}059$ OTU, with zeros corrected with the geometric Bayesian-multiplicative method (Martín-Fernández et al., 2015) and centered log-ratio (**CLR**) transformed. Other corrections and transformations, such as adding one to all values, logarithmic transformation, or use different number of OTU, did not greatly affect the results (not shown). The OTU abundances were precorrected for the significant fixed effects (for more than 10% of OTU; $P < 0.05$) not included in $\boldsymbol{b}$ (Equation [1]), such as total number of sequences per rumen sample, sequencing run, and sampling time and order, all of them nested on the sampling year. Note that the microbial matrix $\boldsymbol{O}$, is called $\boldsymbol{M}$ by other authors (Difford et al., 2018; Weishaar et al., 2020), but here we use $\boldsymbol{M}$ for the abundance matrix to have the same notation as Christensen et al. (2021). Similarly, microbiability, that we call c^2_m, was named m^2 in Difford et al. (2018). Matrix $\boldsymbol{O}$ had an overall mean of 0 and a mean of the diagonal of 1.

To assess the relevance of including the microbiome effect and its interaction with the genetic effect in the prediction model, we used the model with genetic and microbiome effects (Equation [1]), and a model with the genetic effect only, a model with the microbiome effect only, and the model in Equation [1] plus the genome-by-microbiome interaction as follows:

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Z}_1\boldsymbol{g} + \boldsymbol{e}, \qquad [2]$$

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Z}_2\boldsymbol{m} + \boldsymbol{e}, \qquad [3]$$

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Z}_1\boldsymbol{a}_d + \boldsymbol{Z}_2\boldsymbol{m} + \boldsymbol{Z}_3\boldsymbol{am} + \boldsymbol{e}, \qquad [4]$$

where $\boldsymbol{g}$ is the vector of random additive genetic effects, $\boldsymbol{am}$ is the vector of additive genetic and microbiome

interaction effect, $\boldsymbol{Z}_3$ is the incidence matrix for $\boldsymbol{am}$, and $\boldsymbol{y}$, $\boldsymbol{b}$, $\boldsymbol{a}_d$, $\boldsymbol{m}$, $\boldsymbol{e}$, $\boldsymbol{X}$, $\boldsymbol{Z}_1$, and $\boldsymbol{Z}_2$ are described above. The distributional assumptions are $\boldsymbol{g} \sim N\left(0, \boldsymbol{G}\sigma_g^2\right)$, $\boldsymbol{m} \sim N\left(0, \boldsymbol{O}\sigma_m^2\right)$, and $\boldsymbol{am} \sim N\left(0, \left[\boldsymbol{G}^{\circ}\boldsymbol{O}\right]\sigma_{am}^2\right)$, where $\boldsymbol{G}$ is a genomic relationship matrix, $\boldsymbol{O}$ is the microbial similarity matrix, and $\boldsymbol{G}^{\circ}\boldsymbol{O}$ is the Hadamard product of $\boldsymbol{G}$ and $\boldsymbol{O}$ (as implemented by Khanal et al. 2020); σ_g^2 is the additive genetic variance, σ_m^2 is the microbiome effect variance, and σ_{am}^2 is the variance for the interaction effect. In Equation [2] the heritability was estimated as $h^2 = \sigma_g^2 / \sigma_y^2$.

The model with the genetic effect only (Equation [2]), defined as the null hypothesis, was compared with models in Equations [1] and [4] on the basis of a likelihood ratio test. The P-values were computed considering that the distribution of likelihood ratio test asymptotically is a mixture of chi-squared distributions (Visscher, 2006). The significance for chi-squared test with one degree of freedom was defined at $P < 0.05$.

Step 2. The second step consisted on fitting a single trait model using the estimates of microbiome effect (obtained from Equation [1]) as phenotype,

$$\hat{\boldsymbol{m}} = \boldsymbol{W}\boldsymbol{\beta} + \boldsymbol{Z}_1\boldsymbol{g} + \boldsymbol{\varepsilon}, \qquad [5]$$

where $\hat{\boldsymbol{m}}$ is the predicted effects $\boldsymbol{m}$ in Equation [1] for each individual, $\boldsymbol{\beta}$ is the vector of fixed effects (same effects as in $\boldsymbol{b}$); $\boldsymbol{g}$ is the vector of indirect random (mediated by the microbiome) additive genetic effects; and $\boldsymbol{\varepsilon}$ is the vector of random residual effects; $\boldsymbol{W}$ is the incidence matrix for $\boldsymbol{\beta}$. The distributional assumptions are $\boldsymbol{g} \sim N\left(0, \boldsymbol{G}\sigma_g^2\right)$ and $\boldsymbol{\varepsilon} \sim N\left(0, \boldsymbol{I}\sigma_\varepsilon^2\right)$, where $\boldsymbol{G}$ is a genomic relationship matrix, and $\boldsymbol{I}$ is an identity matrix; σ_g^2 is additive genetic variance and σ_ε^2 is residual variance. The heritability of $\hat{\boldsymbol{m}}$ was calculated as $h_{\hat{m}}^2 = \sigma_g^2 / \left(\sigma_g^2 + \sigma_\varepsilon^2\right)$. Here, contrary to our previous work (Martinez Boggio et al., 2022) we assume (or force) the same heritability for all OTU, which is a simplifying assumption in Christensen et al. (2021). In theory, whatever the trait used in Equation [5] and based on the assumption of equal heritability for all OTU, the appropriate heritability to use for such a phenotype $\hat{\boldsymbol{m}}$ is equal to the heritability of the microbial community $\left(h_m^2\right)$. However, first, when $\sigma_m^2 \to 0$, there is no variation in $\hat{\boldsymbol{m}}$, and h_m^2 (although well defined) is not estimable. Second, because the variance components of Equation [1] are themselves estimates, the estimate of variance components in Equation [5], which uses inferred $\hat{\boldsymbol{m}}$, will suffer from incorrect estimation of variance components in Equation [1].

Thus, the total heritability $\left(h_t^2\right)$ of a phenotype is decomposed into an indirect microbiome mediated heritability $\left(c_m^2 h_m^2\right)$, and a direct heritability $\left(h_d^2\right)$ according to the formula $h_t^2 = c_m^2 h_m^2 + h_d^2$ presented by Christensen et al. (2021). Note that when the model includes only the genetic effect, h_t^2 is equal to h^2. All variance components were estimated using BLUPF90+ (Misztal et al., 2002) through a mixture of EM-REML and AI-REML.

Microbiome-Wide Association Study

Single-OTU regression analyses were applied to test the effect of the 2,059 OTUs, one at a time, and obtain the associated P-value. The model used was the Equation [2] plus the specific OTU as a fixed covariate. The P-value of the estimate of the regression coefficient for the fitted OTU-covariate was obtained by converting the estimate and its standard error to Z-score and applying a chi-squared test. The model was fitted using BLUPF90+ (Misztal et al., 2002) with OPTION method VCE.

Genetic Parameters of Rumen Microbial Structure

Principal component analysis (**PCA**) was used to reduce the dimensionality and aggregate the rumen microbiota variance into few variables. We performed PCA on the CLR-transformed abundance matrix $\boldsymbol{M}$, using the centered and scaled option of the prcomp function from the stats package in R (R Core Team, 2021).

Principal components were used as phenotypic variables describing the rumen microbiota structure. The host genetic effect over microbiota and dairy traits was estimated using a 2-traits models including one of each PrC and one of each dairy trait:

$$\begin{bmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\end{bmatrix} = \begin{bmatrix}\boldsymbol{X}_1 & 0\\ 0 & \boldsymbol{X}_2\end{bmatrix}\begin{bmatrix}\boldsymbol{b}_1\\ \boldsymbol{b}_2\end{bmatrix} + \begin{bmatrix}\boldsymbol{W}_1 & 0\\ 0 & \boldsymbol{W}_2\end{bmatrix}\begin{bmatrix}\boldsymbol{g}_1\\ \boldsymbol{g}_2\end{bmatrix} + \begin{bmatrix}\boldsymbol{\varepsilon}_1\\ \boldsymbol{\varepsilon}_2\end{bmatrix}, \qquad [6]$$

where $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ are the vectors of observations for PrC and dairy traits (1–18) for each individual, respectively, $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ are the vectors of fixed effects for each trait, $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$ are the vectors of random additive genetic effects, and $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ are the vectors of random residual effects; $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, $\boldsymbol{W}_1$, and $\boldsymbol{W}_2$ are the incidence matrices for $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{g}_1$, and $\boldsymbol{g}_2$, respectively. The distributional assumptions are $\boldsymbol{g} = N\left(0, \boldsymbol{G} \otimes \boldsymbol{K}\right)$ and $\boldsymbol{\varepsilon} = N\left(0, \boldsymbol{I} \otimes \boldsymbol{R}\right)$, where $\otimes$ denotes the Kronecker product between 2 matrices, $\boldsymbol{G}$ is a genomic relationship matrix

computed based on the first method proposed by VanRaden (2008), $\boldsymbol{I}$ is an identity matrix, and $\boldsymbol{K}$ and $\boldsymbol{R}$ are the genetic and residual variance-covariance matrices for the random additive genetic and residual effects, respectively.

The fixed effects considered in $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ (Equation [6]) were defined by ANOVA test with $P < 0.05$. For all PrC, we included the effects of DIM as a covariate, the sampling year, and the number of lactations nested in the sampling year, and for dairy traits the same effects as defined in Equation [1].

The heritability for each PrC was calculated as $h^2 = \sigma_g^2 / \sigma_y^2$. The genetic correlations between ith PrC and jth dairy trait were computed as follows:

$$corr_{g_{PrC_iDT_j}} = \sigma_{g_{PrC_iDT_j}} \Big/ \sqrt{\sigma^2_{g_{PrC_i}} \times \sigma^2_{g_{DT_j}}},$$

where $\sigma_{g_{PrC_iDT_j}}$ is the additive genetic covariance between PrC_i and DT_j, $\sigma^2_{g_{PrC_i}}$ is the additive genetic variance for PrC_i, and $\sigma^2_{g_{DT_j}}$ is the additive genetic variance for DT_j. Genetic correlations with absolute values higher than twice the standard error were considered to differ from zero.

RESULTS

Description of Phenotypes

A data summary on dairy traits and rumen bacterial composition is available in Martinez Boggio et al. (2022). Briefly, MY averaged 1.95 ± 0.59 L, with an FC of 7.37 ± 1.14 g/100 mL and a PC of 5.71 ± 0.52 g/100 mL. Among the fine milk components measured by MIR spectra, β-CN was the most abundant protein (2.10 ± 0.23 g/100 mL) and palmitic acid (C16:0) the most abundant FA (1.96 ± 0.37 g/100 mL).

Comparison Between Genetic, Microbiome and Interaction Models

Results of the comparison between mixed models are reported in Table 1. The models that included both microbiome and genomic information (Equation [1]), and microbiome, genomic information, and their interaction (Equation [4]) were better than the model that included only genomic information (Equation [2]) for only 3 of the 18 dairy traits. More specifically, the inclusion of the microbiome effect to the genetic one was significant for β-LG ($P < 0.01$), α_{S2}-CN and κ-CN ($P < 0.05$). However, we obtained nonsignificant effects of the inclusion of the microbiome effect in all FA (Table 1).

Microbiability and Heritability of the Microbial Community

The microbiability estimated from the full model was almost zero for most dairy traits (Table 2), with exception of β-LG (0.06 ± 0.05), α_{S2}-CN (0.07 ± 0.05), κ-CN and α_{S1}-CN (0.04 ± 0.04) and PC (0.03 ± 0.03), where values were barely higher than zero. We also run models without fitting a genetic effect (Equation [3]), this analysis was called "microbial mixed model" by Weishaar et al. (2020) and the estimated microbiabilities were also very low (even for β-LG c^2_m, it was estimated as 0.07 instead of 0.06). So, we conclude that in our dataset, the microbiome effect is not "captured" by the genetic effect.

For all the dairy traits with a microbiability almost equal to zero, the model in step 2 did not run because the predicted phenotypes $\hat{\boldsymbol{m}}$ have zero variance and h^2_m is therefore not estimable. For the remaining traits having a nonzero microbiability (albeit small in all cases), h^2_m was very close to zero. So, in practice, h^2_t of dairy traits was equal to the direct heritability (from step 1) as shown in Table 2. Furthermore, we tested the contribution of each specific OTU on dairy traits using a microbiome-wise association study, but no OTU for any dairy trait showed a significant effect (results not shown).

Genetic Correlations Between Microbiome PrC and Dairy Traits

The rumen microbiota with 2,059 OTU was aggregated in few variables by applying PCA to the CLR abundances. We kept the first 5 PrC that explain up to 25% of the microbial variance, and the variance explained by each PrC is shown in Table 3. We obtained low to moderate genetic correlations (from −0.44 ± 0.39 to 0.18 ± 0.44), mainly with PrC2 to PrC4, however the genetic correlations had large standard errors that were not significantly different from zero (Table 3).

DISCUSSION

In this study, we evaluated the impact of rumen bacterial abundance on milk composition traits in dairy ewes, and its genetic control by the host. Rumen 16S rRNA gene sequencing were considered as a source of information between genotype and phenotype in the model of Christensen et al. (2021), because rumen bacteria are closely associated with milk composition traits (Matthews et al., 2019), even though they are not part of the molecular network which operates at various omics levels (e.g., transcriptomics, proteomics,

Table 1. Estimates of variance components [additive genetic $\left(\sigma^2_g \text{ or } \sigma^2_{a_d}\right)$, microbiome $\left(\sigma^2_m\right)$ and its interaction $\left(\sigma^2_{am}\right)$, and residual $\left(\sigma^2_r\right)$ variances], and heritability (h^2) and direct heritability $\left(h^2_d\right)$, microbiability $\left(c^2_m\right)$, and heritability of the interaction between additive genetic and microbiome effect $\left(h^2_{am}\right)$, for model with genetic only (Model G), model with microbiome effect only (Model M), model with genetic and microbiome effects (Model G+M), and model with genetic, microbiome and its interaction (Model G+M+GM); values shown ± SE (given in parentheses if exceptionally low)

Model	Trait[1]	*P*-value[2]	σ^2_g or $\sigma^2_{a_d}$	σ^2_m	σ^2_{am}	σ^2_r	h^2 or h^2_d	c^2_m	h^2_{am}
G	MY		0.13 ± 0.04			0.39 ± 0.02	0.25 ± 0.07		
M	MY			0.003 ± 0.01		0.52 ± 0.03		0.007 ± 0.02	
G+M	MY	0.50	0.13 ± 0.04	0.003 ± 0.01		0.39 ± 0.02	0.25 ± 0.07	0.005 ± 0.02	
G+M+GM	MY	0.50	0.13 ± 0.04	0.003 ± 0.01	<10e-4 (<10e-3)	0.39 ± 0.02	0.25 ± 0.07	0.005 ± 0.02	<10e-3 (<10e-3)
G	PC		0.40 ± 0.06			0.37 ± 0.05	0.52 ± 0.07		
M	PC			0.05 ± 0.04		0.70 ± 0.05		0.07 ± 0.05	
G+M	PC	0.06	0.39 ± 0.06	0.03 ± 0.03		0.34 ± 0.05	0.51 ± 0.07	0.03 ± 0.04	
G+M+GM	PC	0.06	0.39 ± 0.06	0.03 ± 0.03	<10e-4 (<10e-3)	0.34 ± 0.05	0.51 ± 0.07	0.03 ± 0.03	<10e-4 (<10e-3)
G	FC		0.33 ± 0.05			0.28 ± 0.04	0.54 ± 0.07		
M	FC			<10e-4 (<10e-3)		0.60 ± 0.03		<10e-4 (<10e-3)	
G+M	FC	0.50	0.33 ± 0.05	<10e-4 (<10e-3)		0.28 ± 0.04	0.54 ± 0.07	<10e-4 (<10e-3)	
G+M+GM	FC	0.50	0.33 ± 0.05	<10e-4 (<10e-3)	<10e-4 (<10e-3)	0.28 ± 0.04	0.54 ± 0.07	<10e-4 (<10e-3)	<10e-4 (<10e-3)
G	α_{S1}-CN		0.30 ± 0.06			0.30 ± 0.05	0.50 ± 0.09		
M	α_{S1}-CN			0.03 ± 0.03		0.56 ± 0.04		0.06 ± 0.05	
G+M	α_{S1}-CN	0.05	0.29 ± 0.06	0.02 ± 0.02		0.28 ± 0.05	0.49 ± 0.09	0.04 ± 0.04	
G+M+GM	α_{S1}-CN	0.05	0.29 ± 0.06	0.02 ± 0.02	<10e-4 (<10e-3)	0.28 ± 0.05	0.49 ± 0.09	0.04 ± 0.04	<10e-4 (<10e-3)
G	α_{S2}-CN		0.09 ± 0.01			0.04 ± 0.01	0.69 ± 0.08		
M	α_{S2}-CN			0.01 ± 0.01		0.13 ± 0.01		0.04 ± 0.05	
G+M	α_{S2}-CN	0.02	0.10 ± 0.01	0.01 ± 0.01		0.03 ± 0.01	0.70 ± 0.08	0.07 ± 0.05	
G+M+GM	α_{S2}-CN	0.02	0.10 ± 0.01	0.01 ± 0.01	<10e-4 (<10e-3)	0.03 ± 0.01	0.70 ± 0.08	0.07 ± 0.05	<10e-4 (<10e-3)
G	β-CN		0.32 ± 0.08			0.44 ± 0.06	0.42 ± 0.09		
M	β-CN			0.01 ± 0.02		0.74 ± 0.06		0.01 ± 0.03	
G+M	β-CN	0.50	0.32 ± 0.08	<10e-4 (<10e-3)		0.44 ± 0.06	0.42 ± 0.09	<10e-4 (<10e-3)	
G+M+GM	β-CN	0.50	0.32 ± 0.08	<10e-4 (<10e-3)	<10e-4 (<10e-3)	0.44 ± 0.06	0.42 ± 0.09	<10e-4 (<10e-3)	<10e-4 (<10e-3)
G	κ-CN		0.34 ± 0.07			0.36 ± 0.06	0.48 ± 0.09		
M	κ-CN			0.04 ± 0.03		0.65 ± 0.05		0.06 ± 0.05	
G+M	κ-CN	0.03	0.33 ± 0.07	0.03 ± 0.03		0.34 ± 0.06	0.47 ± 0.09	0.04 ± 0.04	
G+M+GM	κ-CN	0.03	0.33 ± 0.07	0.03 ± 0.03	<10e-4 (<10e-3)	0.34 ± 0.06	0.47 ± 0.09	0.04 ± 0.04	<10e-4 (<10e-3)
G	α-LA		0.21 ± 0.08			0.55 ± 0.07	0.28 ± 0.10		
M	α-LA			<10e-4 (<10e-3)		0.75 ± 0.04		<10e-4 (<10e-3)	
G+M	α-LA	0.50	0.21 ± 0.08	<10e-4 (<10e-3)		0.55 ± 0.07	0.28 ± 0.10	<10e-4 (<10e-3)	
G+M+GM	α-LA	0.50	0.21 ± 0.08	<10e-4 (<10e-3)	<10e-4 (<10e-3)	0.55 ± 0.07	0.28 ± 0.10	<10e-4 (<10e-3)	<10e-4 (<10e-3)
G	β-LG		0.38 ± 0.09			0.44 ± 0.07	0.46 ± 0.09		
M	β-LG			0.06 ± 0.04		0.74 ± 0.06		0.08 ± 0.05	
G+M	β-LG	0.005	0.36 ± 0.09	0.05 ± 0.04		0.40 ± 0.07	0.45 ± 0.09	0.06 ± 0.05	
G+M+GM	β-LG	0.005	0.36 ± 0.09	0.05 ± 0.04	<10e-4 (<10e-3)	0.40 ± 0.07	0.45 ± 0.09	0.06 ± 0.05	<10e-4 (<10e-3)
G	C4:0		0.26 ± 0.06			0.30 ± 0.05	0.46 ± 0.09		
M	C4:0			<10e-4 (<10e-3)		0.56 ± 0.03		<10e-4 (<10e-3)	
G+M	C4:0	0.50	0.26 ± 0.06	<10e-4 (<10e-3)		0.30 ± 0.05	0.46 ± 0.09	<10e-4 (<10e-3)	
G+M+GM	C4:0	0.50	0.26 ± 0.06	<10e-4 (<10e-3)	<10e-4 (<10e-3)	0.30 ± 0.05	0.46 ± 0.09	<10e-4 (<10e-3)	<10e-4 (<10e-3)
G	C6:0		0.34 ± 0.08			0.43 ± 0.06	0.45 ± 0.09		
M	C6:0			<10e-4 (<10e-3)		0.76 ± 0.05		<10e-4 (<10e-3)	
G+M	C6:0	0.50	0.34 ± 0.08	<10e-4 (<10e-3)		0.43 ± 0.06	0.45 ± 0.09	<10e-4 (<10e-3)	
G+M+GM	C6:0	0.50	0.34 ± 0.08	<10e-4 (<10e-3)	<10e-4 (<10e-3)	0.43 ± 0.06	0.45 ± 0.09	<10e-4 (<10e-3)	<10e-4 (<10e-3)
G	C8:0		0.32 ± 0.07			0.36 ± 0.06	0.47 ± 0.09		
M	C8:0			0.001 ± 0.02		0.67 ± 0.04		0.002 ± 0.03	
G+M	C8:0	0.50	0.32 ± 0.07	<10e-4 (<10e-3)		0.36 ± 0.06	0.47 ± 0.09	<10e-4 (<10e-3)	
G+M+GM	C8:0	0.50	0.32 ± 0.07	<10e-4 (<10e-3)	<10e-4 (<10e-3)	0.36 ± 0.06	0.47 ± 0.09	<10e-4 (<10e-3)	<10e-4 (<10e-3)

Continued

Table 1 (Continued). Estimates of variance components [additive genetic $\left(\sigma^2_g \text{ or } \sigma^2_{a_d}\right)$, microbiome $\left(\sigma^2_m\right)$ and its interaction $\left(\sigma^2_{am}\right)$, and residual $\left(\sigma^2_r\right)$ variances], and heritability (h^2) and direct heritability $\left(h^2_d\right)$, microbiability $\left(c^2_m\right)$, and heritability of the interaction between additive genetic and microbiome effect $\left(h^2_{am}\right)$, for model with genetic only (Model G), model with microbiome effect only (Model M), model with genetic and microbiome effects (Model G+M), and model with genetic, microbiome and its interaction (Model G+M+GM); values shown ± SE (given in parentheses if exceptionally low)

Model	Trait[1]	*P*-value[2]	σ^2_g or $\sigma^2_{a_d}$	σ^2_m	σ^2_{am}	σ^2_r	h^2 or h^2_d	c^2_m	h^2_{am}
G	C10:0		0.35 ± 0.07			0.32 ± 0.05	0.52 ± 0.09		
M	C10:0			0.004 ± 0.02		0.67 ± 0.05		0.01 ± 0.03	
G+M	C10:0	0.50	0.35 ± 0.07	<10e-4 (<10e-3)		0.32 ± 0.05	0.52 ± 0.09	<10e-4 (<10e-3)	
G+M+GM	C10:0	0.50	0.35 ± 0.07	<10e-4 (<10e-3)	<10e-4 (<10e-3)	0.32 ± 0.05	0.52 ± 0.09	<10e-4 (<10e-3)	<10e-4 (<10e-3)
G	C12:0		0.35 ± 0.07			0.28 ± 0.05	0.55 ± 0.09		
M	C12:0			0.02 ± 0.03		0.62 ± 0.05		0.03 ± 0.04	
G+M	C12:0	0.50	0.35 ± 0.07	<10e-4 (<10e-3)		0.28 ± 0.05	0.55 ± 0.08	<10e-4 (<10e-3)	
G+M+GM	C12:0	0.50	0.35 ± 0.07	<10e-4 (<10e-3)	<10e-4 (<10e-3)	0.28 ± 0.05	0.55 ± 0.09	<10e-4 (<10e-3)	<10e-4 (<10e-3)
G	C16:0		0.21 ± 0.04			0.24 ± 0.04	0.47 ± 0.08		
M	C16:0			<10e-4 (<10e-3)		0.45 ± 0.03		<10e-4 (<10e-3)	
G+M	C16:0	0.41	0.21 ± 0.04	0.004 ± 0.01		0.24 ± 0.04	0.47 ± 0.08	0.008 ± 0.03	
G+M+GM	C16:0	0.41	0.21 ± 0.04	0.004 ± 0.01	<10e-4 (<10e-3)	0.24 ± 0.04	0.47 ± 0.08	0.008 ± 0.03	<10e-4 (<10e-3)
G	*c*9 C18:1		0.23 ± 0.06			0.30 ± 0.05	0.43 ± 0.10		
M	*c*9 C18:1			<10e-4 (<10e-3)		0.52 ± 0.03		<10e-4 (<10e-3)	
G+M	*c*9 C18:1	0.50	0.23 ± 0.06	<10e-4 (<10e-3)		0.30 ± 0.05	0.43 ± 0.10	<10e-4 (<10e-3)	
G+M+GM	*c*9 C18:1	0.50	0.23 ± 0.06	<10e-4 (<10e-3)	<10e-4 (<10e-3)	0.30 ± 0.05	0.43 ± 0.10	<10e-4 (<10e-3)	<10e-4 (<10e-3)
G	*c*9*t*11 C18:2		0.20 ± 0.06			0.38 ± 0.05	0.35 ± 0.09		
M	*c*9*t*11 C18:2			<10e-4 (<10e-3)		0.57 ± 0.03		<10e-4 (<10e-3)	
G+M	*c*9*t*11 C18:2	0.50	0.20 ± 0.06	<10e-4 (<10e-3)		0.38 ± 0.05	0.35 ± 0.09	<10e-4 (<10e-3)	
G+M+GM	*c*9*t*11 C18:2	0.50	0.20 ± 0.06	<10e-4 (<10e-3)	<10e-4 (<10e-3)	0.38 ± 0.05	0.35 ± 0.09	<10e-4 (<10e-3)	<10e-4 (<10e-3)
G	C18:3n-3		0.17 ± 0.05			0.28 ± 0.04	0.36 ± 0.09		
M	C18:3n-3			<10e-4 (<10e-3)		0.44 ± 0.03		<10e-4 (<10e-3)	
G+M	C18:3n-3	0.50	0.17 ± 0.05	<10e-4 (<10e-3)		0.28 ± 0.04	0.36 ± 0.09	<10e-4 (<10e-3)	
G+M+GM	C18:3n-3	0.50	0.17 ± 0.05	<10e-4 (<10e-3)	<10e-4 (<10e-3)	0.28 ± 0.04	0.36 ± 0.09	<10e-4 (<10e-3)	<10e-4 (<10e-3)

[1]MY= milk yield; PC = milk protein content; FC = milk fat content.

[2]*P*-values from the comparison between models G+M and G+M+GM with model G.

Table 2. Estimates of direct heritability (not mediated by the microbiome; h_d^2), microbiability $\left(c_m^2\right)$, heritability of the microbial community $\left(h_m^2\right)$, and total heritability $\left(h_t^2\right)$ as product of $\left(c_m^2 h_m^2\right)$ plus $\left(h_d^2\right)$; values shown ± SE (given in parentheses if exceptionally low)

Trait[1]	h_d^2	c_m^2	h_m^2	h_t^2
MY	0.25 ± 0.07	0.005 ± 0.02	—[2]	0.25
FC	0.54 ± 0.07	<10e-4 (<10e-3)	—[2]	0.54
PC	0.51 ± 0.07	0.03 ± 0.03	0.004 ± 0.003	0.51
α_{S1}-CN	0.49 ± 0.09	0.04 ± 0.04	0.006 ± 0.004	0.49
α_{S2}-CN	0.70 ± 0.08	0.07 ± 0.05	0.03 ± 0.007	0.70
β-CN	0.42 ± 0.09	<10e-4 (<10e-3)	—[2]	0.42
κ-CN	0.47 ± 0.08	0.04 ± 0.04	0.006 ± 0.003	0.47
α-LA	0.28 ± 0.10	<10e-4 (<10e-3)	—[2]	0.28
β-LG	0.44 ± 0.09	0.06 ± 0.05	0.003 ± 0.002	0.44
C4:0	0.46 ± 0.09	<10e-4 (<10e-3)	—[2]	0.46
C6:0	0.44 ± 0.09	<10e-4 (<10e-3)	—[2]	0.44
C8:0	0.47 ± 0.09	<10e-4 (<10e-3)	—[2]	0.47
C10:0	0.52 ± 0.09	<10e-4 (<10e-3)	—[2]	0.52
C12:0	0.55 ± 0.08	<10e-4 (<10e-3)	—[2]	0.55
C16:0	0.47 ± 0.08	0.01 ± 0.03	—[2]	0.47
*c*9 C18:1	0.43 ± 0.09	<10e-4 (<10e-3)	—[2]	0.43
*c*9*t*11 C18:2	0.35 ± 0.09	<10e-4 (<10e-3)	—[2]	0.35
C18:3n-3	0.36 ± 0.09	<10e-4 (<10e-3)	—[2]	0.36

[1]MY= milk yield; FC = milk fat content; PC = milk protein content.

[2]Model in Equation [5] not run.

metabolomics). Moreover, a genetic control of bacterial abundances by the host, and genetic associations with the fine milk composition, had been shown Martinez Boggio et al. (2022). This close link of some rumen microbiota with milk FA is due to rumen fermentation products, such as VFA, which influence the milk FA profile through de novo synthesis of FA in the mammary gland and biohydrogenation of UFA in the rumen (Lourenço et al., 2010; Osorio et al., 2016). Furthermore, the link between rumen microbiota and milk proteins is mainly due to the fact that the passage of microbial proteins to the intestine supplies the mammary gland with essential AA used for protein production (Osorio et al., 2016). Although rumen bacteria provide precur-

Table 3. Proportion of the variance explained by the first 5 principal components (PrC), heritability, and genetic correlations between the first 5 principal components and dairy traits; values shown ± SE (given in parentheses if exceptionally low)

Item[1]	PrC1	PrC2	PrC3	PrC4	PrC5
Proportion of variance (%)	9.58	7.39	6.10	1.58	1.32
Heritability	0.01 ± 0.04	0.21 ± 0.06	0.23 ± 0.06	0.10 ± 0.06	0.11 ± 0.06
Genetic correlations					
MY	0.24 ± 1.09	−0.003 ± 0.25	0.02 ± 0.24	−0.52 ± 0.78	0.18 ± 0.48
FC	−0.07 ± 0.65	−0.25 ± 0.20	−0.14 ± 0.18	0.14 ± 0.37	0.16 ± 0.34
PC	0.25 ± 0.53	−0.27 ± 0.20	−0.24 ± 0.17	−0.04 ± 0.36	−0.11 ± 0.29
α_{S1}-CN	0.52 ± 0.72	−0.42 ± 0.22	−0.23 ± 0.20	0.18 ± 0.44	−0.01 ± 0.38
α_{S2}-CN	−0.41 ± 0.72	−0.42 ± 0.28	−0.18 ± 0.17	0.17 ± 0.51	−0.07 ± 0.33
β-CN	0.75 ± 2.07	−0.21 ± 0.23	−0.15 ± 0.22	0.14 ± 1.26	−0.01 ± 0.46
κ-CN	0.07 ± 0.55	−0.35 ± 0.22	−0.21 ± 0.20	0.17 ± 0.51	−0.11 ± 0.45
α-LAC	0.90 ± 0.56	0.08 ± 0.31	−0.06 ± 0.29	−0.18 ± 0.64	−0.22 ± 0.63
β-LG	0.00 ± 0.70	−0.32 ± 0.23	−0.19 ± 0.21	0.14 ± 0.47	−0.13 ± 0.38
C4:0	−0.46 ± 0.64	−0.20 ± 0.24	0.10 ± 0.20	0.09 ± 0.60	−0.23 ± 0.40
C6:0	−0.77 ± 0.96	−0.25 ± 0.24	−0.06 ± 0.21	0.05 ± 0.40	−0.23 ± 0.38
C8:0	−0.90 ± 0.49	−0.18 ± 0.22	−0.15 ± 0.21	0.03 ± 0.49	−0.06 ± 0.42
C10:0	−0.85 ± 0.39	−0.16 ± 0.23	−0.19 ± 0.19	0.13 ± 0.44	−0.06 ± 0.37
C12:0	−0.60 ± 0.57	−0.05 ± 0.20	−0.18 ± 0.18	0.16 ± 0.65	0.02 ± 0.42
C16:0	0.59 ± 0.67	−0.19 ± 0.23	−0.02 ± 0.21	0.11 ± 0.44	0.01 ± 0.38
*c*9 C18:1	0.99 ± 1.01	−0.21 ± 0.24	−0.05 ± 0.22	−0.21 ± 0.48	−0.13 ± 0.55
*c*9*t*11 C18:2	0.95 ± 0.28	−0.12 ± 0.30	0.01 ± 0.25	−0.04 ± 0.48	0.11 ± 0.61
C18:3n-3	−0.16 ± 1.04	−0.44 ± 0.39	−0.28 ± 0.25	0.07 ± 0.51	0.10 ± 0.45

[1]MY= milk yield; FC = milk fat content; PC = milk protein content.

sors to produce milk FA and proteins, the microbiota effect obtained in this study was only significant for some milk proteins, such as β-LG, α_{S2}-CN, and κ-CN.

For all dairy traits included in this study, total heritability was moderate to high (0.25–0.70). However, the contribution to total heritability was almost due to the direct genetic effect on the phenotype [i.e., total heritability $\left(c_m^2 h_m^2 + h_d^2\right)$ was almost equal to the direct heritability $\left(h_d^2\right)$]. This was due, first, to the weak effect of the microbiota on most dairy traits, as shown by low values of microbiability $\left(c_m^2\right)$ estimated with each of the models evaluated. In the literature, a zero c_m^2 was reported for most milk FA in dairy cows, except for C15:0 (0.42 ± 0.18) and C18:3n-3 (0.31 ± 0.14; Buitenhuis et al., 2019), and there are no references for milk proteins, for which we obtained a very low impact of microbiota (e.g., α_{S2}-CN 0.07 ± 0.05 and β-LG 0.06 ± 0.05). Second, the heritability of the microbial community $\left(h_m^2\right)$ as a whole estimated in step 2 was close to zero, and similar to the average heritability of 0.04 ± 0.03 obtained for the 2,059 OTU that made up this community (Martinez Boggio et al., 2022). Thus, those elements evidence a noncontribution of the genetic effect on the phenotype mediated by the rumen bacterial abundance $\left(c_m^2 h_m^2\right)$.

Despite the results obtained with milk composition traits, the methodology proposed by Christensen et al. (2021) allowed to separate the direct genetic effect on phenotypes from the indirect genetic effect on phenotypes mediated by the microbiome. This discrimination may be useful as Weishaar et al. (2020) proposed to improve a trait by changing the rumen microbiota composition, selecting for mediated breeding values, or by other metabolic pathways selecting for nonmediated breeding values. The method builds on some assumptions that are debatable.

In step 2, the model assumed a constant heritability for all OTU to simplify the calculations, even if across OTU heritability estimated varied between 0 and 0.29 (Martinez Boggio et al., 2022). The method also assumes (or imposes) that the direct and indirect genetic effects are uncorrelated. This means that a gene A that contributes to the genetic variation of the indirect genetic effect on phenotype via the microbiota does not contribute to the genetic variation of the direct genetic effect and vice versa. This assumption is convenient, but probably also robust for microbiome data, because it has been seen that only a small percentage of the microbiota have genetic correlations with phenotypes different from zero, and these correlations have alternating signs for different microbiota (Martinez Boggio et al., 2022; Martínez-Álvaro et al., 2022), resulting in an average genetic correlation between indirect and direct genetic effect of zero, at least for traits where the proportion of variance explained by microbiota is small. Eventually, the model implies that the covariance between the (total) genetics background for the trait of interest and the genetics of the microbiome (each OTU weighted by its role in the trait) is a function of $c_m^2 h_m^2$. The squared genetic correlation between the trait of interest and the trait explained by microbiome is in fact $r_g^2 = \frac{c_m^2 h_m^2}{h_t^2} = 1 - \frac{h_d^2}{h_t^2}$ (Legarra and Christensen, 2022). Note that for each OTU, there will be a different correlation (which may be positive or negative), so genetic correlations between each trait and OTU may not be very informative. Moreover, in a previous study (Martinez Boggio et al., 2022) we estimated these correlations to be weak (only 4% were higher than 0.50), and here most of the genetic correlations with the PrC were not significant. Finally, the use of a joint model may limit the quantification of host genetic control of the microbial community, since at the extreme when the predicted microbiota effects $\hat{\boldsymbol{m}}$ are zero, the parameters of the model in step 2 could not be estimated.

We evaluated the inclusion of the rumen microbiota as a correlated trait with the phenotype of interest through a 2-trait model, which was previously reported by (Saborío-Montero et al., 2021). In our study, we obtained a genetic control of the host over the microbiota represented by 5 PrC, but most of the genetic correlations although moderate have large standard errors, implying that the accuracy of genomic prediction may not be improved by the use of a multitrait model with omics traits as proposed Hayes et al. (2017).

The results here showed that including 16S rRNA gene sequencing as an additional information source of data will not improve accuracy of EBV. It would be interesting to evaluate some elements that may improve this accuracy. Firstly, the use of traits more closely associated with rumen microbiota, such as rumen FA or methane emissions, which are direct products of the rumen microbial community (Hurtaud et al., 1993; Dehority, 2003). In the literature for those traits, moderate but relevant effects of the rumen microbiota were obtained. For metabolic traits, such as milk acetone and β-hydroxybutyric acid, a c_m^2 of 0.15 ± 0.09 were estimated for both traits (Gebreyesus et al., 2020). In addition, for methane emissions Difford et al. (2018) and Ramayo-Caldas et al. (2020) in dairy cows and Hess et al. (2020) in meat sheep, reported c_m^2 of 0.13 ± 0.09, 0.16 ± 0.09, and 0.19 ± 0.07, respectively. Secondly, replacing the use of microbiota abundances by microbial functions to construct the microbial similarity matrix would avoid functional redundancy at the ru-

men microbiome (Weimer, 2015). Furthermore, we suggest much further work is needed to evaluate more complex models that fit the biological nature of the microbial community in the rumen, because the methodology used in this study was designed for intermediate -omics traits between DNA and the phenotype of interest, which is not really the case with microbiome.

CONCLUSIONS

Rumen microbial abundance does not contribute to the phenotypic variance of most of the fine milk composition traits and no causative OTU were detected. In turn, there was no genetic control of the microbial community. So far, we can conclude that, using a substantial data set of 795 Lacaune dairy ewes, rumen bacterial abundances do not provide improved genetic evaluation for dairy traits in sheep. Further work will be needed to evaluate traits that are a direct product of the microbiota (e.g., rumen VFA or methane emissions) as well as the substitution of microbial abundances for microbial functions.

ACKNOWLEDGMENTS

Financial support was provided by INRAE (Toulouse, France), La Region Occitanie (Toulouse, France) and the CRB-Anim infrastructure project (grant no. ANR-11-INBS-0003). We are grateful to the technical staff at INRAE Experimental Unit of La Fage for animal care and help with rumen sampling. We thank Béatrice Gabinaud and Marie-Luce Chemit (UMR GenPhySE, INRAE, Toulouse) for preparing microbiota samples, Get-Plage platform (INRAE, Toulouse) for DNA sequencing, Géraldine Pascal (UMR GenPhySE, INRAE, Toulouse) for designed and coordinated the bioinformatic analysis, Hélène Larroque and Rachel Rupp (UMR GenPhySE, INRAE, Toulouse) for access to the genetic resources produced at INRAE La Fage. We are grateful to the genotoul bioinformatics platform Toulouse Occitanie (Bioinfo Genotoul, Toulouse) for providing help and computing and storage resources. The authors have not stated any conflicts of interest.

REFERENCES

Bijma, P. 2014. The quantitative genetics of indirect genetic effects: A selective review of modelling issues. Heredity 112:61–69. https://doi.org/10.1038/hdy.2013.15.

Buitenhuis, B., J. Lassen, S. J. Noel, D. R. Plichta, P. Sørensen, G. F. Difford, and N. A. Poulsen. 2019. Impact of the rumen microbiome on milk fatty acid composition of Holstein cattle. Genet. Sel. Evol. 51:23. https://doi.org/10.1186/s12711-019-0464-8.

Christensen, O. F., V. Börner, L. Varona, and A. Legarra. 2021. Genetic evaluation including intermediate omics features. Genetics 219:iyab130. https://doi.org/10.1093/genetics/iyab130.

Dehority, B. A. 2003. Gross anatomy, physiology and environment of the ruminant stomach. Pages 19–42 in Rumen Microbiology. Nottingham Univ. Press.

Difford, G. F., D. R. Plichta, P. Løvendahl, J. Lassen, S. J. Noel, O. Højberg, A.-D. G. Wright, Z. Zhu, L. Kristensen, H. B. Nielsen, B. Guldbrandtsen, and G. Sahana. 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.

Escudié, F., L. Auer, M. Bernard, M. Mariadassou, L. Cauquil, K. Vidal, S. Maman, G. Hernandez-Raquet, S. Combes, and G. Pascal. 2018. FROGS: Find, Rapidly, OTUs with Galaxy Solution. Bioinformatics 34:1287–1294. https://doi.org/10.1093/bioinformatics/btx791.

Fernie, A. R., and N. Schauer. 2009. Metabolomics-assisted breeding: A viable option for crop improvement? Trends Genet. 25:39–48. https://doi.org/10.1016/j.tig.2008.10.010.

Ferrand, M., G. Miranda, H. Larroque, O. Leray, S. Guisnel, F. Lahalle, M. Brochard, and P. Martin. 2012. Determination of protein composition in milk by mid-infrared spectrometry. In ICAR 38 Annual Meeting, Cork, Ireland.

Ferrand-Calmels, M., I. Palhière, M. Brochard, O. Leray, J. M. Astruc, M. R. Aurel, S. Barbey, F. Bouvier, P. Brunschwig, H. Caillat, M. Douguet, F. Faucon-Lahalle, M. Gelé, G. Thomas, J. M. Trommenschlager, and H. Larroque. 2014. Prediction of fatty acid profiles in cow, ewe, and goat milk by mid-infrared spectrometry. J. Dairy Sci. 97:17–35. https://doi.org/10.3168/jds.2013-6648.

Gebreyesus, G., G. F. Difford, B. Buitenhuis, J. Lassen, S. J. Noel, O. Højberg, D. R. Plichta, Z. Zhu, N. A. Poulsen, U. K. Sundekilde, P. Løvendahl, and G. Sahana. 2020. Predictive ability of host genetics and rumen microbiome for subclinical ketosis. J. Dairy Sci. 103:4557–4569. https://doi.org/10.3168/jds.2019-17824.

Guo, Z., M. M. Magwire, C. J. Basten, Z. Xu, and D. Wang. 2016. Evaluation of the utility of gene expression and metabolic information for genomic prediction in maize. Theor. Appl. Genet. 129:2413–2427. https://doi.org/10.1007/s00122-016-2780-5.

Hayes, B. J., J. Panozzo, C. K. Walker, A. L. Choy, S. Kant, D. Wong, J. Tibbits, H. D. Daetwyler, S. Rochfort, M. J. Hayden, and G. C. Spangenberg. 2017. Accelerating wheat breeding for end-use quality with multi-trait genomic predictions incorporating near infrared and nuclear magnetic resonance-derived phenotypes. Theor. Appl. Genet. 130:2505–2519. https://doi.org/10.1007/s00122-017-2972-7.

Hess, M. K., S. J. Rowe, T. C. Van Stijn, H. M. Henry, S. M. Hickey, R. Brauning, A. F. McCulloch, A. S. Hess, M. R. Kirk, S. Kumar, C. Pinares-Patiño, S. Kittelmann, G. R. Wood, P. H. Janssen, and J. C. McEwan. 2020. A restriction enzyme reduced representation sequencing approach for low-cost, high-throughput metagenome profiling. PLoS One 15:e0219882. https://doi.org/10.1371/journal.pone.0219882.

Hurtaud, C., H. Rulquin, and R. Verite. 1993. Effect of infused volatile fatty acids and caseinate on milk composition and coagulation in dairy cows. J. Dairy Sci. 76:3011–3020. https://doi.org/10.3168/jds.S0022-0302(93)77640-7.

Khanal, P., C. Maltecca, C. Schwab, J. Fix, M. Bergamaschi, and F. Tiezzi. 2020. Modeling host-microbiome interactions for the prediction of meat quality and carcass composition traits in swine. Genet. Sel. Evol. 52:41. https://doi.org/10.1186/s12711-020-00561-7.

Larroque, H., M. Chassier, R. Saintilan, and J. M. Astruc. 2017. Imputation accuracy from a low density SNP panel in 5 dairy sheep breeds in France. Page 513 in Annual Meeting of the EEAP. Tallinn, Estonia. Wageningen Academic Publishers.

Legarra, A., and O. F. Christensen. 2022. Genomic evaluation methods to include intermediate correlated features such as high-throughput or omics phenotypes. JDS Commun. 4:55–60. https://doi.org/10.3168/jdsc.2022-0276.

Lourenço, M., E. Ramos-Morales, and R. J. Wallace. 2010. The role of microbes in rumen lipolysis and biohydrogenation and their manipulation. Animal 4:1008–1023. https://doi.org/10.1017/S175173111000042X.

Mahé, F., T. Rognes, C. Quince, C. de Vargas, and M. Dunthorn. 2015. Swarm v2: Highly-scalable and high-resolution amplicon clustering. PeerJ 3:e1420. https://doi.org/10.7717/peerj.1420.

Martín-Fernández, J.-A., K. Hron, M. Templ, P. Filzmoser, and J. Palarea-Albaladejo. 2015. Bayesian-multiplicative treatment of count zeros in compositional data sets. Stat. Model. 15:134–158. https://doi.org/10.1177/1471082X14535524.

Martínez-Álvaro, M., J. Mattock, M. Auffret, Z. Weng, C.-A. Duthie, R. J. Dewhurst, M. A. Cleveland, M. Watson, and R. Roehe. 2022. Microbiome-driven breeding strategy potentially improves beef fatty acid profile benefiting human health and reduces methane emissions. Microbiome 10:166. https://doi.org/10.1186/s40168-022-01352-6.

Martinez Boggio, G., C. Marie-Etancelin, J.-M. Menras, R. Tomas, M.-L. Chemit, B. Gabinaud, G. Pascal, and A. Meynadier. 2023. A large database linking the rumen bacterial composition and milk traits in Lacaune sheep. Sci. Data 10:17. https://doi.org/10.1038/s41597-022-01912-3.

Martinez Boggio, G., A. Meynadier, A. J. Buitenhuis, and C. Marie-Etancelin. 2022. Host genetic control on rumen microbiota and its impact on dairy traits in sheep. Genet. Sel. Evol. 54:77. https://doi.org/10.1186/s12711-022-00769-9.

Martinez Boggio, G., A. Meynadier, P. Daunis-i-Estadella, and C. Marie-Etancelin. 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.

Matthews, C., F. Crispie, E. Lewis, M. Reid, P. W. O'Toole, and P. D. Cotter. 2019. The rumen microbiome: A crucial consideration when optimising milk and meat production and nitrogen utilisation efficiency. Gut Microbes 10:115–132. https://doi.org/10.1080/19490976.2018.1505176.

Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. H. Lee. 2002. BLUPF90 and related programs (BGF90). Pages 743–744 in Proceedings of the 7th WCGALP. Montpellier, France.

Morgante, F., W. Huang, P. Sørensen, C. Maltecca, and T. F. C. Mackay. 2020. Leveraging multiple layers of data to predict *Drosophila* complex traits. G3 (Bethesda) 10:4599–4613. https://doi.org/10.1534/g3.120.401847.

Osorio, J. S., J. Lohakare, and M. Bionaz. 2016. Biosynthesis of milk fat, protein, and lactose: Roles of transcriptional and posttranscriptional regulation. Physiol. Genomics 48:231–256. https://doi.org/10.1152/physiolgenomics.00016.2015.

R Core Team. 2021. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Accessed Sep. 22, 2022. http://www.R-project.org/.

Ramayo-Caldas, Y., L. Zingaretti, M. Popova, J. Estellé, A. Bernard, N. Pons, P. Bellot, N. Mach, A. Rau, H. Roume, M. Perez-Enciso, P. Faverdin, N. Edouard, D. Ehrlich, D. P. Morgavi, and G. Renand. 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.

Ross, E. M., P. J. Moate, L. Marett, B. G. Cocks, and B. J. Hayes. 2013. Investigating the effect of two methane-mitigating diets on the rumen microbiome using massively parallel sequencing. J. Dairy Sci. 96:6030–6046. https://doi.org/10.3168/jds.2013-6766.

Saborío-Montero, A., A. López-García, M. Gutiérrez-Rivas, R. Atxaerandio, I. Goiri, A. García-Rodriguez, J. A. Jiménez-Montero, C. González, J. Tamames, F. Puente-Sánchez, L. Varona, M. Serrano, C. Ovilo, and O. González-Recio. 2021. 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.

VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423. https://doi.org/10.3168/jds.2007-0980.

Visscher, P. M. 2006. A note on the asymptotic distribution of likelihood ratio tests to test variance components. Twin Res. Hum. Genet. 9:490–495. https://doi.org/10.1375/twin.9.4.490.

Weimer, P. J. 2015. Redundancy, resilience, and host specificity of the ruminal microbiota: Implications for engineering improved ruminal fermentations. Front. Microbiol. 6. https://doi.org/10.3389/fmicb.2015.00296.

Weishaar, R., R. Wellmann, A. Camarinha-Silva, M. Rodehutscord, and J. Bennewitz. 2020. Selecting the hologenome to breed for an improved feed efficiency in pigs—A novel selection index. J. Anim. Breed. Genet. 137:14–22. https://doi.org/10.1111/jbg.12447.

Xue, M., H. Sun, X. Wu, L. L. Guan, and J. Liu. 2018. Assessment of rumen microbiota from a large dairy cattle cohort reveals the pan and core bacteriomes contributing to varied phenotypes. Appl. Environ. Microbiol. 84:e00970-e18. https://doi.org/10.1128/AEM.00970-18.