The Value of Gut Microbiota to Predict Feed Efficiency and Growth of Rabbits Under Different Feeding Regimes

María Velasco-Galilea (maria.velasco@irta.es)
Institute of Agrifood Research and Technology (IRTA) - Animal Breeding and Genetics
Miriam Piles
Institute of Agrifood Research and Technology (IRTA) - Animal Breeding and Genetics
Yuliaxis Ramayo-Caldas
Institute of Agrifood Research and Technology (IRTA) - Animal Breeding and Genetics
Juan P. Sánchez
Institute of Agrifood Research and Technology (IRTA) - Animal Breeding and Genetics

Research Article

Keywords: rabbits, Animal, feeding regimes, nutrient, feed efficiency, breeding

Posted Date: May 24th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-441480/v2

The value of gut microbiota to predict feed efficiency and growth of rabbits under different feeding regimes

María Velasco-Galilea[1*], Miriam Piles[1], Yuliaxis Ramayo-Caldas[1] and Juan P. Sánchez[1]

[1]Institute of Agrifood Research and Technology (IRTA) - Animal Breeding and Genetics, E08140 Caldes de Montbui, Barcelona, Spain

***Corresponding author:**

María Velasco-Galilea

maria.velasco@irta.es

Abstract

Gut microbiota plays an important role in nutrient absorption and could impact rabbit feed efficiency. This study aims at investigating such impact by evaluating the value added by microbial information for predicting individual growth and cage phenotypes related to feed efficiency. The dataset comprised individual average daily gain and cage-average daily feed intake from 425 meat rabbits, in which cecal microbiota was assessed, and their cage mates. Despite microbiota was not measured in all animals, consideration of pedigree relationships with mixed models allowed the study of cage-average traits. The inclusion of microbial information into certain mixed models increased their predictive ability up to 20% and 46% for cage-average feed efficiency and individual growth traits, respectively. These gains were associated with large microbiability estimates and with reductions in the heritability estimates. However, large microbiabililty estimates were also obtained with certain models but without any improvement in their predictive ability. A large proportion of OTUs seems to be responsible for the prediction improvement in growth and feed efficiency traits, although specific OTUs have a higher weight. Rabbit growth and feed efficiency are influenced by host cecal microbiota and considering microbial information in models improves the prediction of these complex phenotypes.

Introduction

Feed efficiency (FE) is a fundamental trait in rabbit breeding since food expenses often represent up to 70% of the production costs [1-Cartuche et al., 2014]. The difficulties entailed in measuring the individual animals' feed intake (FI) are the main responsible for most programs do not perform a direct selection for FE. An alternative commonly used to improve

FE is the indirect selection for average daily gain (ADG) or body weight (BW) at the end of the growing period [2-Estany et al., 1992]. Nevertheless, the genetic correlation between these growth traits and FE may be not high enough to result in an optimal selection response [3-Piles et al., 2004]. Therefore, it would be worth exploring new traits allowing alternative selection strategies such as FE definitions based on cage-average FI records. In this regard, the present study uses cage-average records of FI and individual records of BW collected from animals raised in groups, thus reflecting the reality of commercial farms where animals are raised in groups.

The cecum is the main organ harboring the microbial fermentation processes in the domestic meat rabbit, *Oryctolagus cuniculus*. This organ hosts a complex microbial ecosystem dominated by bacterial phyla *Firmicutes*, *Tenericutes*, and *Bacteroidetes* [4-Velasco-Galilea et al., 2018]. The interactions that are continuously taking place between bacteria and their host ensure the homeostatic balance maintenance of the cecum ecosystem. Previous studies revealed that relative abundances of these, and other less abundant taxa, vary between individuals and are affected by external factors such as the breeding farm, the level of feeding, or the administration of antibiotics 5-Velasco-Galilea et al., 2020].

In the field of livestock production, certain studies have hypothesized that the rabbit gut microbiota could be associated with growth [6-Zeng et al., 2014] and FE [7-Drouilhet et al., 2016]. Furthermore, a recent study has identified several operational taxonomic units (OTUs) and KEEG pathways associated with ADG in commercial meat rabbits [8-Fang et al., 2020]. Nonetheless, a fact that should not be overlooked is the strong impact on the animals' growth and FE exerted by the breeding environment or common rabbit breeding strategies such as

feed restriction [9-Gidenne et al., 2012], thus when considering the role of gut microbiota on performance traits these management and environmental effects must not be ignored. To our knowledge, no published study has so far investigated the connection between the gut microbiota and animal performance together with these external factors that also affect growth and FE while shaping microbial communities [5-Velasco-Galilea et al., 2020]. Moreover, the existing collinearity between microbiota and management effects difficult the finding of real associations of the animal growth with specific taxa abundances.

This study aims at understanding the role of microbial communities inhabiting the cecum on the FE and the growth of rabbits raised in collective cages under different feeding regimes. The use of sparse partial least squares regression (sPLSR) and mixed models in cross-validation schema will allow unraveling the value of cecal microbiota to predict cage FE and individual growth performances in a rabbit line selected for post-weaning growth.

Results

Influence of genetics and cecal microbiota on rabbit growth and FE

Table 1 includes statistics of marginal posterior distributions for heritabilities (h^2), microbiabilities (m^2), and phenotypic variances for individually recorded traits (ADG_{AL} and ADG_R) obtained with the dataset including only records of animals in which microbiota was assessed (mDataset). Similarly, Table 2 and Table 3 include estimates for the same parameters referring both to individual growth and cage-average traits ($\overline{ADFI}_{AL}$, $\overline{ADRFI}_{AL}$ and $\overline{ADFCR}_{AL}$). In these latter two cases, the estimates were computed with the dataset including records of animals in which microbiota was assessed as well as of their cage mates

(fullDataset). Statistics were obtained with the model not including the microbial effect (M1) and with the models fitting the microbial effect (M2) by considering different prior assumptions. Trace plots and histograms of Markov chains from the posterior distribution of the parameters of these models using different prior assumptions and datasets are included as Additional file 4.

The heritabilities (h^2) obtained with M1 and the mDataset were 0.21 and 0.29 for ADG_{AL} and ADG_R, respectively (Table 1). The posterior means of h^2 obtained with M1 and the fullDataset were markedly lower, 0.15 and 0.09 for ADG_{AL} and ADG_R, respectively (Table 2 and Table 3). However, estimates cannot be considered significantly different between datasets. The h^2 estimates with M2 models including the microbial effect ranged, depending on the prior assumption for the microbial effects and the dataset used for the analysis, from 0.05 to 0.15 for ADG_{AL} and from 0.07 to 0.09 for ADG_R. These ranges for m^2 varied from 0.00 to 0.79 for ADG_{AL} and from 0.00 to 0.77 for ADG_R. In general, it was observed that the higher the magnitude of m^2, the higher the changes in the h^2 estimates from M1 to M2. It is important to note that the lowest estimates of m^2 for both traits were obtained in the analyses in which all the individual records were considered for the study and the elements of the covariance matrices for animals without microbial composition were generated considering cage-average CSS OTU counts ($\mathbf{M}_{\bar{\mathbf{O}}}$, $\mathbf{M}_{\bar{\mathbf{B}}}$ or $\mathbf{M}_{\bar{\mathbf{U}}}$) (Table 3). The posterior means of m^2 for both traits were almost null for nearly all the cases studied with these covariance matrices, except for ADG_{AL} when the covariance matrix was defined from the Bray-Curtis distance matrix ($\mathbf{M}_{\bar{\mathbf{B}}}$) and for ADG_R when the covariance matrix was defined from the weighted Unifrac distance matrix ($\mathbf{M}_{\bar{\mathbf{U}}}$) . Note that large estimation errors were observed in both cases.

These errors can also be linked with the poor mixing of the sampling processes that are evidenced in the trace plots provided in the Additional file 4.

Regarding cage-average traits, the posterior means of h^2 obtained with M1 were medium-high ranging from 0.26 ($\overline{\text{ADFI}}_{\text{AL}}$) to 0.49 ($\overline{\text{ADRFI}}_{\text{AL}}$) (Tables 2 and 3). When the microbial effect was included, these posterior means tended to decrease. The h^2 obtained with M2 models ranged, depending on the prior assumption for the microbial effects, from 0.11 to 0.24 for $\overline{\text{ADFI}}_{\text{AL}}$, from 0.12 to 0.44 for $\overline{\text{ADRFI}}_{\text{AL}}$, and from 0.08 to 0.30 for $\overline{\text{ADFCR}}_{\text{AL}}$. The posterior means of m^2 ranged from 0.03 to 0.58 for $\overline{\text{ADFI}}_{\text{AL}}$, from 0.10 to 0.76 for $\overline{\text{ADRFI}}_{\text{AL}}$, and from 0.16 to 0.78 for $\overline{\text{ADFCR}}_{\text{AL}}$. Note that for all cage-average traits the highest posterior mean of h^2 and the lowest posterior mean of m^2 were obtained when the microbial covariance matrix was expanded using cage-average CSS OTU counts and then computing their cross-product ($\mathbf{M}_{\bar{\mathbf{O}}}$). The lowest posterior means of h^2 and the highest posterior means of m^2 were obtained with the microbial covariance matrix $\mathbf{M}_{\bar{\mathbf{U}}}$ (i.e., expanding the OTU table using cage-average CSS OTU counts and then computing the weighted Unifrac distance matrix). It is worth mentioning that, similarly to growth traits, the posterior means of the parameters obtained with M2 models based on expanding the CSS OTU table by cage-average before computing the respective distance matrices ($\mathbf{M}_{\bar{\mathbf{O}}}$, $\mathbf{M}_{\bar{\mathbf{B}}}$ or $\mathbf{M}_{\bar{\mathbf{U}}}$) (Table 3) are associated with large posterior standard errors. For these analyses, poor mixing was also observed (Additional file 4). Given our dataset size, the covariance structure generated with this expansion procedure seems not suitable to properly identify the covariance between animals due to sharing cecal microbial composition. The posterior means of h^2 and m^2 for these traits seem to be more consistent when they were obtained with the M2 models based on the

expansion of the microbial relationship matrices that just included ones in the diagonal and zeros outside the diagonal for the animals without microbial information (Table 2). In this case, a similar pattern was obtained with $\mathbf{M_{O,0}}$, $\mathbf{M_{B,0}}$ and $\mathbf{M_{U,0}}$: h^2 decrease from 0.26 (M1) to 0.19 for $\overline{\text{ADFI}}_{\text{AL}}$, from 0.49 (M1) to 0.32 for $\overline{\text{ADRFI}}_{\text{AL}}$, and from 0.34 (M1) to 0.21 for $\overline{\text{ADFCR}}_{\text{AL}}$ while m^2 ranged from 0.45 to 0.49 for $\overline{\text{ADFI}}_{\text{AL}}$, from 0.38 to 0.42 for $\overline{\text{ADRFI}}_{\text{AL}}$, and from 0.45 to 0.49 for $\overline{\text{ADFCR}}_{\text{AL}}$.

Table 1 Means (SD) of marginal posterior distributions of the heritability (h^2), microbiability (m^2) and phenotypic variance (Phe. Var.) for ADG_{AL} and ADG_R obtained with the mDataset.

Parameter	Model	Microbial matrix	ADG_{AL}	ADG_R
h^2	M1	--	0.21 (0.14)	0.29 (0.19)
Phe. Var.	M1	--	41.20 (4.37)	32.80 (3.93)
h^2	M2	$\mathbf{M_O}$	0.07 (0.07)	0.13 (0.09)
m^2	M2	$\mathbf{M_O}$	0.67 (0.15)	0.56 (0.12)
Phe. Var.	M2	$\mathbf{M_O}$	93.08 (26.03)	57.90 (12.51)
h^2	M2	$\mathbf{M_B}$	0.05 (0.05)	0.07 (0.06)
m^2	M2	$\mathbf{M_B}$	0.79 (0.12)	0.77 (0.10)
Phe. Var.	M2	$\mathbf{M_B}$	193.85 (83.54)	129.08 (46.78)
h^2	M2	$\mathbf{M_U}$	0.08 (0.09)	0.14 (0.13)
m^2	M2	$\mathbf{M_U}$	0.60 (0.26)	0.49 (0.26)
Phe. Var.	M2	$\mathbf{M_U}$	174.85 (168.52)	91.03 (72.38)

ADG_{AL}: average daily gain in rabbits fed *ad libitum*; ADG_R: average daily gain in rabbits fed under restriction; SD: standard deviation; M1: model without microbial effects; M2: model fitting the microbial effects; $\mathbf{M_O}$: microbial relationship covariance matrix defined from CSS normalized OTU counts, $\mathbf{M_B}$: microbial relationship covariance matrix defined from Bray-Curtis distance matrix; $\mathbf{M_U}$: microbial relationship covariance matrix defined from weighted Unifrac distance matrix.

Table 2 Means (SD) of marginal posterior distributions of the heritability (h^2), microbiability (m^2) and phenotypic variance (Phe. Var.) for individual traits (ADG_{AL} and ADG_R) and cage-average traits ($\overline{ADFI}_{AL}$, $\overline{ADRFI}_{AL}$ and $\overline{ADFCRI}_{AL}$) obtained with the fullDataset by expanding the corresponding microbial relationship matrix with ones in the diagonal and zeros outside.

Parameter	Model	Microbial matrix[1]	ADG_{AL}	ADG_R	$\overline{ADFI}_{AL}$	$\overline{ADRFI}_{AL}$	$\overline{ADFCR}_{AL}$
h^2	M1	--	0.15 (0.09)	0.09 (0.07)	0.26 (0.18)	0.49 (0.20)	0.34 (0.20)
Phe. Var.	M1	--	79.79 (4.67)	57.02 (3.40)	635.14 (102.99)	206.59 (33.06)	0.20 (0.03)
h^2	M2	$\mathbf{M_{O,0}}$	0.11 (0.06)	0.08 (0.05)	0.19 (0.13)	0.33 (0.15)	0.22 (0.14)
m^2	M2	$\mathbf{M_{O,0}}$	0.63 (0.06)	0.66 (0.05)	0.48 (0.18)	0.38 (0.17)	0.47 (0.18)
Phe. Var.	M2	$\mathbf{M_{O,0}}$	90.54 (5.47)	66.50 (4.13)	676.55 (118.29)	219.47 (37.77)	0.21 (0.04)
h^2	M2	$\mathbf{M_{B,0}}$	0.12 (0.07)	0.07 (0.06)	0.19 (0.13)	0.31 (0.15)	0.22 (0.14)
m^2	M2	$\mathbf{M_{B,0}}$	0.56 (0.06)	0.61 (0.05)	0.49 (0.18)	0.42 (0.17)	0.49 (0.17)
Phe. Var.	M2	$\mathbf{M_{B,0}}$	92.04 (5.67)	68.13 (4.38)	711.55 (128.31)	227.88 (40.04)	0.22 (0.04)
h^2	M2	$\mathbf{M_{U,0}}$	0.13 (0.07)	0.07 (0.06)	0.19 (0.13)	0.32 (0.15)	0.22 (0.15)
m^2	M2	$\mathbf{M_{U,0}}$	0.52 (0.06)	0.58 (0.05)	0.45 (0.19)	0.40 (0.17)	0.45 (0.18)
Phe. Var.	M2	$\mathbf{M_{U,0}}$	92.11 (5.78)	68.26 (4.43)	711.42 (128.01)	226.68 (39.58)	0.22 (0.04)

ADG_{AL}: average daily gain in rabbits fed *ad libitum*; ADG_R: average daily gain in rabbits fed under restriction; $\overline{ADFI}_{AL}$: average daily feed intake in rabbits fed *ad libitum*; $\overline{ADRFI}_{AL}$: average daily residual feed intake in rabbits fed *ad libitum*; $\overline{ADFCR}_{AL}$: average daily feed conversion ratio in rabbits fed *ad libitum*; SD: standard deviation; M1: model without microbial effects; M2: model fitting the microbial effects.

[1]The expansion of the microbial relationship matrix ($\mathbf{M_O}$, $\mathbf{M_B}$ or $\mathbf{M_U}$) was done by including ones in the diagonal and zeros outside the diagonal for the animals without microbial information.

Table 3 Means (SD) of marginal posterior distributions of the heritability (h^2), microbiability (m^2) and phenotypic variance (Phe. Var.) for individual traits (ADG_{AL} and ADG_R) and cage-average traits ($\overline{ADFI}_{AL}$, $\overline{ADRFI}_{AL}$ and $\overline{ADFCRI}_{AL}$) obtained with the fullDataset by expanding the OTU matrix with the cage-average counts.

Parameter	Model	Microbial matrix[1]	ADG_{AL}	ADG_R	$\overline{ADFI}_{AL}$	$\overline{ADRFI}_{AL}$	$\overline{ADFCR}_{AL}$
h^2	M1	--	0.15 (0.09)	0.09 (0.07)	0.26 (0.18)	0.49 (0.20)	0.34 (0.20)
Phe. Var.	M1	--	79.79 (4.67)	57.02 (3.40)	635.14 (102.99)	206.59 (33.06)	0.20 (0.03)
h^2	M2	$\mathbf{M}_{\bar{O}}$	0.14 (0.09)	0.09 (0.07)	0.24 (0.17)	0.44 (0.19)	0.30 (0.18)
m^2	M2	$\mathbf{M}_{\bar{O}}$	0.08 (0.05)	0.00 (0.00)	0.03 (0.06)	0.10 (0.12)	0.16 (0.09)
Phe. Var.	M2	$\mathbf{M}_{\bar{O}}$	85.71 (6.42)	57.08 (3.40)	635.52 (102.28)	209.30 (34.46)	0.21 (0.03)
h^2	M2	$\mathbf{M}_{\bar{B}}$	0.09 (0.06)	0.09 (0.07)	0.16 (0.12)	0.23 (0.13)	0.20 (0.14)
m^2	M2	$\mathbf{M}_{\bar{B}}$	0.39 (0.13)	0.06 (0.03)	0.44 (0.19)	0.56 (0.17)	0.44 (0.16)
Phe. Var.	M2	$\mathbf{M}_{\bar{B}}$	133.31 (32.36)	61.00 (6.57)	1059.88 (359.15)	407.68 (135.59)	0.32 (0.09)
h^2	M2	$\mathbf{M}_{\bar{U}}$	0.15 (0.09)	0.07 (0.06)	0.11 (0.10)	0.12 (0.12)	0.08 (0.08)
m^2	M2	$\mathbf{M}_{\bar{U}}$	0.00 (0.00)	0.25 (0.23)	0.58 (0.24)	0.76 (0.20)	0.78 (0.17)
Phe. Var.	M2	$\mathbf{M}_{\bar{U}}$	79.83 (4.67)	88.33 (43.15)	2106.33 (1622.31)	1284.29 (948.14)	1.20 (0.80)

ADG_{AL}: average daily gain in rabbits fed *ad libitum*; ADG_R: average daily gain in rabbits fed under restriction; $\overline{ADFI}_{AL}$: average daily feed intake in rabbits fed *ad libitum*; $\overline{ADRFI}_{AL}$: average daily residual feed intake in rabbits fed *ad libitum*; $\overline{ADFCR}_{AL}$: average daily feed conversion ratio in rabbits fed *ad libitum*; SD: standard deviation; M1: model without microbial effects; M2: model fitting the microbial effects.

[1]The expansion of the microbial relationship matrix ($\mathbf{M_O}$, $\mathbf{M_B}$ or $\mathbf{M_U}$) was done before computing the respective distance matrices, assigning to the animals without microbial information the cage-average of the CSS normalized OTU counts.

Predictive ability of individual growth and cage FE from microbial information

Table 4 shows the correlation coefficient between observed and predicted records of individual traits (ADG_{AL} and ADG_R) in the validation set reached with the different tested models and the mDataset. It was observed that the consideration of microbial information resulted in a significant prediction improvement of the individually measured growth traits only when $\mathbf{M_O}$ or $\mathbf{M_B}$ were used as covariance matrix between individual microbial effects. The consideration of microbial information in M2 models improved the predictive capacity of ADG_{AL} and ADG_R by 25% and 46%, respectively.

Table 4 Across 100 replicates average (SD) correlation coefficient between observed and predicted ADG_{AL} and ADG_R records with sPLSR and mixed models using the mDataset.

Model	Microbial matrix	ADG_{AL}	ADG_R
M1	--	0.30(0.15)	0.39(013)
M2	$\mathbf{M_O}$	0.36(0.13)*[a]	0.56(0.11)*[a]
M2	$\mathbf{M_B}$	0.38(0.13)*[a]	0.57(0.12)*[a]
M2	$\mathbf{M_U}$	0.30(0.14)	0.39(0.13)
sPLSR1	--	0.50 (0.11)	0.28 (0.14)
sPLSR2	--	0.51 (0.11)	0.19 (0.16)

ADG_{AL}: average daily gain in rabbits fed *ad libitum*; ADG_R: average daily gain in rabbits fed under restriction; SD: standard deviation; M1: mixed model without microbial effects; M2: mixed model fitting the microbial effects; $\mathbf{M_O}$: microbial relationship covariance matrix defined from CSS normalized OTU counts, $\mathbf{M_B}$: microbial relationship covariance matrix defined from Bray-Curtis distance matrix; $\mathbf{M_U}$: microbial relationship covariance matrix defined from weighted Unifrac distance matrix; sPLS1: sparse Partial Least Squares Regression model with systematic effects as predictors; sPLS2: sparse Partial Least Squares Regression model with systematic effects and CSS OTU counts as predictors.

*M2 or sPLSR2 correlation between observed and predicted records significantly higher (bootstrapped paired t test) than M1 or sPLSR1 correlation after Bonferroni correction for multiple testing at the $P < 0.05$ level.

[a]M2 or sPLSR2 correlation between observed and predicted records higher than M1 or sPLSR1 correlation in at least 80% of the replicates.

When $\mathbf{M_U}$ was used as covariance matrix between individual microbial effects no improvement of the predictive capacity was observed for any trait. The same was observed when microbial information was included in sPLSR2 models fitting systematic effects and CSS OTU counts. sPLSR2 models did not exhibit better predictive ability than those models just fitting the systematic effects (sPLSR1).

Table 5 shows the correlation coefficient between observed and predicted records of individual growth traits (ADG_{AL} and ADG_R) in the validation set when different mixed models and microbial covariance matrices were used. In this case, the analyses were conducted using the fullDataset. Here the correlation coefficient between observed and predicted records of each trait in the validation set was computed separately for the animals with microbial information and for the animals without this information. The only consistent improvement in the predictive ability was observed on animals in which cecal microbiota was assessed for ADG_R using M2 models based on the expansion of the microbial relationship matrices including ones in the diagonal and zeros outside the diagonal. The predictive capacity of ADG_R with these M2 models increased by 17% with respect to M1.

Finally, Table 6 shows the correlation coefficient between observed and predicted records of cage-average traits ($\overline{ADFI}_{AL}$, $\overline{ADRFI}_{AL}$ and $\overline{ADFCR}_{AL}$) in the validation set reached with the different mixed and sPLSR models under study using the fullDataset.

Table 5 Across 100 replicates average (SD) correlation coefficient between observed and mixed model predicted ADG_{AL} and ADG_R records using the fullDataset by expanding the microbial relationship covariance matrix in different ways.

		Animals with microbial information		Animals without microbial information	
Model	**Microbial matrix**	**ADG_{AL}**	**ADG_R**	**ADG_{AL}**	**ADG_R**
M1	--	0.46 (0.15)	0.48 (0.15)	0.39 (0.11)	0.42 (0.14)
M2	$\mathbf{M_{O,0}}^1$	0.47 (0.14)	0.56 (0.14)*[a]	0.37 (0.10)	0.42 (0.14)
M2	$\mathbf{M_{B,0}}^1$	0.46 (0.15)	0.57 (0.15)*[a]	0.37 (0.10)	0.43 (0.14)
M2	$\mathbf{M_{U,0}}^1$	0.45 (0.15)	0.55 (0.14)*[a]	0.37 (0.10)	0.43 (0.14)
M2	$\mathbf{M_{\bar{O}}}^2$	0.47 (0.14)*	0.48 (0.15)	0.39 (0.10)	0.42 (0.14)
M2	$\mathbf{M_{\bar{B}}}^2$	0.47 (0.15)*	0.48 (0.15)	0.39 (0.10)*	0.42 (0.14)
M2	$\mathbf{M_{\bar{U}}}^2$	0.45 (0.15)	0.48 (0.15)	0.39 (0.10)	0.42 (0.14)

ADG_{AL}: average daily gain in rabbits fed *ad libitum*; ADG_R: average daily gain in rabbits fed under restriction; SD: standard deviation; M1: mixed model without microbial effects; M2: mixed model fitting the microbial effects; $\mathbf{M_O}$: microbial relationship covariance matrix defined from CSS normalized OTU counts, $\mathbf{M_B}$: microbial relationship covariance matrix defined from Bray-Curtis distance matrix; $\mathbf{M_U}$: microbial relationship covariance matrix defined from weighted Unifrac distance matrix.

[1]The expansion of the microbial relationship matrix ($\mathbf{M_O}$, $\mathbf{M_B}$ or $\mathbf{M_U}$) was done by including ones in the diagonal and zeros outside the diagonal for the animals without microbial information.

[2]The expansion of the microbial relationship matrix ($\mathbf{M_O}$, $\mathbf{M_B}$ or $\mathbf{M_U}$) was done before computing the respective distance matrices, assigning to the animals without microbial information the cage-average of the CSS normalized OTU counts.

*M2 correlation between observed and predicted records significantly higher (bootstrapped paired t test) than M1 correlation after false discovery rate correction for multiple testing at the $P < 0.05$ level.

[a]M2 correlation between observed and predicted records higher than M1 correlation in at least 80% of the replicates.

Table 6 Across 100 replicates average (SD) correlation coefficient between observed and predicted individual cage-average $\overline{\mathbf{ADFI}}_{\mathbf{AL}}$, $\overline{\mathbf{ADRFI}}_{\mathbf{AL}}$ and $\overline{\mathbf{ADFCRI}}_{\mathbf{AL}}$ records with sPLSR and mixed models using the fullDataset.

Model	Microbial matrix	$\overline{\mathbf{ADFI}}_{\mathbf{AL}}$	$\overline{\mathbf{ADRFI}}_{\mathbf{AL}}$	$\overline{\mathbf{ADFCR}}_{\mathbf{AL}}$
M1	--	0.79 (0.11)	0.42 (0.21)	0.61 (0.16)
M2	$\mathbf{M}_{\mathbf{O},\mathbf{0}}$[1]	0.83 (0.08)*[a]	0.50 (0.19)*[a]	0.69 (0.12)*[a]
M2	$\mathbf{M}_{\mathbf{B},\mathbf{0}}$[1]	0.83 (0.08)*[a]	0.50 (0.19)*[a]	0.69 (0.12)*[a]
M2	$\mathbf{M}_{\mathbf{U},\mathbf{0}}$[1]	0.82 (0.08)*[a]	0.50 (0.18)*[a]	0.69 (0.12)*[a]
M2	$\mathbf{M}_{\bar{\mathbf{O}}}$[2]	0.79 (0.11)	0.41 (0.21)	0.61 (0.16)
M2	$\mathbf{M}_{\bar{\mathbf{B}}}$[2]	0.79 (0.11)	0.41 (0.21)	0.61 (0.16)
M2	$\mathbf{M}_{\bar{\mathbf{U}}}$[2]	0.79 (0.11)	0.42 (0.21)	0.61 (0.15)
sPLSR1	--	0.79 (0.08)	-0.31 (0.14)	0.65 (0.15)
sPLSR2	--	0.73 (0.09)	0.17 (0.21)*[a]	0.39 (0.18)

$\overline{\text{ADFI}}_{\text{AL}}$: average daily feed intake in rabbits fed *ad libitum*; $\overline{\text{ADRFI}}_{\text{AL}}$: average daily residual feed intake in rabbits fed *ad libitum*; $\overline{\text{ADFCR}}_{\text{AL}}$: average daily feed conversion ratio in rabbits fed *ad libitum*; SD: standard deviation; M1: mixed model without microbial effects; M2: mixed model fitting the microbial effects; $\mathbf{M}_{\mathbf{O}}$: microbial relationship covariance matrix defined from CSS normalized OTU counts, $\mathbf{M}_{\mathbf{B}}$: microbial relationship covariance matrix defined from Bray-Curtis distance matrix; $\mathbf{M}_{\mathbf{U}}$: microbial relationship covariance matrix defined from weighted Unifrac distance matrix; sPLS1: sparse Partial Least Squares Regression model with systematic effects as predictors; sPLS2: sparse Partial Least Squares Regression model with systematic effects and cage-average CSS OTU counts as predictors.

[1]The expansion of the microbial relationship matrix ($\mathbf{M}_{\mathbf{O}}$, $\mathbf{M}_{\mathbf{B}}$ or $\mathbf{M}_{\mathbf{U}}$) was done by including ones in the diagonal and zeros outside the diagonal for the animals without microbial information.

[2]The expansion of the microbial relationship matrix ($\mathbf{M}_{\mathbf{O}}$, $\mathbf{M}_{\mathbf{B}}$ or $\mathbf{M}_{\mathbf{U}}$) was done before computing the respective distance matrices, assigning to the animals without microbial information the cage-average of the CSS normalized OTU counts.

*M2 or sPLSR2 correlation between observed and predicted records significantly higher (bootstrapped paired t test) than M1 or sPLSR1 correlation after false discovery rate correction for multiple testing at the $P < 0.05$ level.

[a]M2 or sPLSR2 correlation between observed and predicted records higher than M1 or sPLSR1 correlation in at least 80% of the replicates.

The M2 mixed models in which the elements of the covariance matrices for animals without microbial information were generated from cage-average CSS OTU counts did not add any predictive value for any trait. On the contrary, the consideration of microbial information resulted in a significant improvement of the predictive ability of all traits with all M2 mixed

models based on microbial relationship matrices expanded with ones in the diagonal and zeros outside the diagonal for the animals without microbial information. When these models are used, the predictive ability increased by 5%, 20% and 14% for $\overline{\text{ADFI}}_{\text{AL}}$, $\overline{\text{ADRFI}}_{\text{AL}}$ and $\overline{\text{ADFCR}}_{\text{AL}}$, respectively, over M1. These improvements were nearly the same irrespectively the covariance matrix considered: $\mathbf{M_{O,0}}$, $\mathbf{M_{B,0}}$ or $\mathbf{M_{U,0}}$.

Regarding the sPLSR multivariate approach, the correlation coefficient between observed and predicted records reached in the validation set with the model that only included the systematic effects as predictors (sPLSR1) was pretty high and in most cases better than that achieved with the sPLSR2 models (i.e., also including the cage-average CSS OTU counts as predictors). The only exception was observed for $\overline{\text{ADRFI}}_{\text{AL}}$ what could be said to be expected since a correction by batch effect is implicit in its definition. Thus, the systematic effects considered do not play any role in the prediction of the observations, indeed, an average negative correlation associated with large dispersion was observed. This average correlation turned positive (although of low magnitude: 0.17) when CSS OTU counts were considered, resulting in a significant improvement of the predictive capacity of the model for this cage-average phenotype.

Identification of relevant OTUs for the prediction of rabbit growth and FE

The observed improvement in the predictive ability of the sPLSR2 model for $\overline{\text{ADRFI}}_{\text{AL}}$ could be explained by the systematic selection of 7 OTUs in more than 80 out of the 100 replicates conducted. Table 7 shows the taxonomic assignment with the RDP classifier of the selected OTUs, and their representative sequences can be found in Additional file 5. Out of these

OTUs, 5 belong to family *Lachnospiraceae* and 2 are unclassified bacteria. The Pearson's correlations between these OTUs and $\overline{\text{ADRFI}}_{\text{AL}}$ were computed to quantify the degree of association. These correlations ranged from -0.33 to 0.31 (Table 7).

Table 7 Taxonomic assignment of the OTUs selected in the sPLSR analysis for $\overline{\text{ADRFI}}_{\text{AL}}$.

OTU ID and taxonomical assignment	Pearson's correlation
874627 Unclassified *Bacteria*	0.31
NR1922 Unclassified *Lachnospiraceae*	-0.27
NR153 Unclassified *Lachnospiraceae*	0.31
NR3628 Unclassified *Lachnospiraceae*	-0.33
NR381 Unclassified *Lachnospiraceae*	-0.31
NR4083 Unclassified *Lachnospiraceae*	0.32
NR768 Unclassified *Bacteria*	-0.27

$\overline{\text{ADRFI}}_{\text{AL}}$: average daily residual feed intake in rabbits fed *ad libitum*.

On the other hand, sPLSR models were used to fit the posterior means of the individual microbial effects predicted for growth and FE traits with M2 models and microbial covariance matrices $\mathbf{M_{O,0}}$, $\mathbf{M_{B,0}}$ or $\mathbf{M_{U,0}}$ to identify the most relevant OTUs for the prediction of such phenotypes. Table 8 shows, for each trait and covariance matrix, the number of OTUs selected from a total of 946 in at least 80 out of the 100 replicates conducted.

Table 8 Number of OTUs selected in at least 80 out of the 100 sPLSR replicates conducted for microbial effects predicted with covariance matrices $\mathbf{M_{O,0}}$, $\mathbf{M_{B,0}}$ and $\mathbf{M_{U,0}}$ for growth and FE traits.

Trait	$\mathbf{M_{O,0}}$	$\mathbf{M_{B,0}}$	$\mathbf{M_{U,0}}$	Most relevant[1]
$\mathbf{ADG_{AL}}$	911	931	673	16
$\mathbf{ADG_{R}}$	887	874	621	13
$\overline{\mathbf{ADFI}}_{\mathbf{AL}}$	850	785	490	25
$\overline{\mathbf{ADRFI}}_{\mathbf{AL}}$	600	793	480	18
$\overline{\mathbf{ADFCR}}_{\mathbf{AL}}$	824	832	877	13

ADG_{AL}: average daily gain in rabbits fed *ad libitum*; ADG_R: average daily gain in rabbits fed under restriction; $\overline{ADFI}_{AL}$: average daily feed intake in rabbits fed *ad libitum*; $\overline{ADRFI}_{AL}$: average daily residual feed intake in rabbits fed *ad libitum*; $\overline{ADFCR}_{AL}$: average daily feed conversion ratio in rabbits fed *ad libitum*; $\mathbf{M_{O,0}}$: microbial relationship covariance matrix defined from CSS normalized OTU counts and expanded by including ones in the diagonal and zeros outside the diagonal for the animals without microbial information, $\mathbf{M_{B,0}}$: microbial relationship covariance matrix defined from Bray-Curtis distance matrix and expanded by including ones in the diagonal and zeros outside the diagonal for the animals without microbial information; $\mathbf{M_{U,0}}$: microbial relationship covariance matrix defined from weighted Unifrac distance matrix and expanded by including ones in the diagonal and zeros outside the diagonal for the animals without microbial information.

[1]The most relevant OTUs were those with the greatest loading weights and that were selected with $\mathbf{M_{O,0}}$, $\mathbf{M_{B,0}}$ and $\mathbf{M_{U,0}}$.

Additionally, Table S1 shows the taxonomy of the most relevant OTUs (i.e., those having the greatest loading weights and selected with the three M2 models) for the prediction of growth and FE traits based on the individual microbial effects predicted with the linear mixed models. The Pearson's correlations between each OTU and the traits are also shown in Table S1 while their representative sequences can be found in Additional file 7. Sixteen OTUs seemed to have an important weight for the prediction improvement f ADG_{AL}. Ten of them belong to phylum *Firmicutes*, 2 to phylum *Euryarchaeota*, and 4 OTUs are unclassified *Bacteria*. Thirteen OTUs were found to be relevant to improve the predictive ability of mixed models for ADG_R. Of these OTUs, 10 belong to phylum *Firmicutes*, 2 to phylum *Euryarchaeota* and 1 to phylum *Bacteroidetes*. Twenty-five OTUs were found to be involved

298 in the improvement of the predictive ability of mixed models for $\overline{\text{ADFI}}_{\text{AL}}$. Most of them (20
299 OTUs) belong to phylum *Firmicutes*, 1 to phylum *Bacteroidetes*, 1 to phylum *Actinobacteria*,
300 1 to phylum *Proteobacteria*, and 2 OTUs are unclassified *Bacteria*. Eighteen OTUs were
301 found to be relevant to improve the predictive ability of mixed models for $\overline{\text{ADRFI}}_{\text{AL}}$. Out of
302 these OTUs, 9 belong to phylum *Firmicutes*, 3 to phylum *Bacteroidetes*, 2 to phylum
303 *Actinobacteria*, 1 to phylum *Proteobacteria*, and 3 OTUs are unclassified *Bacteria*. Finally,
304 13 OTUs were responsible for the prediction improvement of $\overline{\text{ADFCR}}_{\text{AL}}$ when microbial
305 information was fitted in the proposed mixed models. Most of them (8 OTUs) belong to
306 phylum *Firmicutes*, 2 to phylum *Bacteroidetes*, and 3 OTUs are unclassified *Bacteria*. It is
307 worth mentioning that some OTUs were found to be relevant for the prediction of more than
308 one trait. In this regard, two OTUs belonging to genus *Methanobrevibacter* and one to order
309 *Clostridiales* were found to be relevant for the prediction of both growth traits, i.e., ADG_{R}
310 and ADG_{AL}. One OTU taxonomically assigned to family *Lachnospiraceae* was found to be
311 relevant for the prediction of both ADG_{AL} and $\overline{\text{ADFI}}_{\text{AL}}$. Seven OTUs (2 belonging to genus
312 *Eisenbergiella*, 1 to class *Alphaproteobacteria*, 1 to genus *Longibaculum*, 1 to family
313 *Erysipelotrichaceae*, 1 to family *Lachnospiraceae*, and 1 unclassified *Bacteria*) were found
314 to be relevant for the prediction of both $\overline{\text{ADFI}}_{\text{AL}}$ and $\overline{\text{ADRFI}}_{\text{AL}}$. Two OTUs (1 belonging to
315 genus *Ruminococcus*, and 1 to family *Lachnospiraceae*) were found to be relevant for the
316 prediction of both ADG_{R} and $\overline{\text{ADFI}}_{\text{AL}}$. Two OTUs (1 belonging to genus *Butyricimonas*, and
317 1 unclassified *Bacteria*) were found to be relevant for the prediction of both $\overline{\text{ADRFI}}_{\text{AL}}$ and
318 $\overline{\text{ADFCR}}_{\text{AL}}$. One OTU belonging to genus *Butyricicoccus* was found to be relevant for the
319 prediction of ADG_{R}, ADG_{AL} and $\overline{\text{ADFI}}_{\text{AL}}$. Finally, 2 OTUs (1 belonging to family

320 *Lachnospiraceae*, and 1 to genus *Blautia*) were found to be relevant for the prediction of
321 ADG_R, $\overline{ADFI}_{AL}$ and $\overline{ADRFI}_{AL}$ (Table S1).
322
323 ## Discussion
324 The role of microbial communities inhabiting the rabbit cecum on key breeding traits related
325 to FE remains unknown. To shed light on this matter, we have reported heritabilities and
326 microbiabilities of ADG under different feeding regimes commonly used in meat rabbit
327 commercial farms. We have also computed such ratios for cage-average traits related to FI
328 and FE in animals fed AL. Dealing with such cage-average performances, while having only
329 measured cecal microbial information in a few animals per cage, is a statistical modeling
330 challenge. We have faced it using different approaches, with the final objective of evaluating
331 the predictive value of microbial information for both individual growth and cage-average
332 FE phenotypes.
333
334 The study of ADG has particular significance for rabbit breeding programs since this trait is
335 commonly selected to indirectly improve FE. Apart from that, the commercial application of
336 feed restriction (i.e., a reduction in the amount of the feed provided to the animal) is common
337 since it improves FE and reduces mortality and morbidity caused by enteric disorders [10-
338 Gidenne et al., 2009]. Piles and Sánchez (2019) [11] estimated a low genetic correlation
339 between ADG_{AL} and ADG_R, and the genome-wide association study conducted by Sánchez
340 et al. (2020) [12] identified different QTL regions for both traits. Such findings support the
341 existence of different genetic backgrounds for these traits. Thus, in this study, we reported
342 the posterior means of the heritability (h^2) for ADG_{AL} and ADG_R separately. In line with

343 previous results [11], we have found a lower h^2 for ADG_R, which implies difficulties to
344 achieve a response to selection for growth or indirectly for FE.
345
346 In this context, one can understand the relevance of exploring whether microbiota explains a
347 significant percentage of the phenotypic variance of these traits as well as the value of
348 microbial information to predict such complex traits as tools to define the degree of influence
349 of microbial information on the traits of interest. A clear effect of microbial composition on
350 the traits of interest would open the door to search and select for taxa positively associated
351 with them. Ross et al. (2013) [13], motivated by the existence of numerous exploratory
352 studies in humans and other animals aiming at relating the microbiome to a complex trait,
353 tested a method to predict body mass index in humans and methane production phenotypes
354 in cattle. Their results showed that microbial information could be useful to predict complex
355 host phenotypes, and even suggested that it could exceed prediction accuracies based on the
356 host genome for traits largely influenced by the gut microbiota. Following that study, others
357 have been conducted in an attempt to evaluate the utility of microbial information to predict
358 complex phenotypes in different livestock species. However, to date, there is a lack of
359 knowledge about the value of microbial information to predict phenotypes related to growth
360 in rabbits. This is the first study to assess the value of cecal microbiota to predict individual
361 growth traits in meat rabbits using different modeling approaches. What is more, this is the
362 first time that the predictive value of microbial information is evaluated when this
363 information has not been measured in all the individuals contributing to the phenotype. The
364 first challenge we faced was to properly define a between-animals relationship matrix due to
365 microbial effects (**M**). Thus, we replicated each analysis with three alternative definitions of
366 **M**: one defined from CSS normalized OTU counts ($\mathbf{M_0}$) and two defined from two classical

367 measures of distance; Bray-Curtis ($\mathbf{M_B}$) and weighted Unifrac ($\mathbf{M_U}$). A second challenge was
368 to define an appropriate way to expand **M** for those animals in which cecal microbiota was
369 not assessed. These developments are strongly linked with several prediction tools based on
370 kernel methods already proposed [14-Ramon et al, 2021]. In our study, we have derived
371 kernel matrices by implementing an ad-hoc solution to transform distance matrices into
372 proper covariance matrices, while Ramon et al. (2021) [14] directly derived the kernel
373 matrices associated with distance metrics from raw information. Not having microbial
374 information for all the animals under study would request, anyhow, some heuristics to
375 generate valid covariance matrices to be included in the mixed models.
376
377 Despite the difficulties mentioned above and the fact that, in general, a low predictive ability
378 for growth traits was observed (the correlation coefficient between observed and predicted
379 records in the validation set with M1 was not higher than 0.4), we have been able to detect a
380 certain predictive ability improvement by considering microbial information. Such
381 consideration improved the predictive capacity of mixed models for ADG_{AL} and ADG_R by
382 25% and 46%, respectively, in the dataset comprised of only the rabbits in which cecal
383 microbiota was assessed (mDataset). When the role of the microbial information was
384 assessed by inspecting the percentage of phenotypic variance explained by the bacterial
385 effect, a large proportion was attributed to the bacterial effect, being this large proportion of
386 the phenotypic variance accompanied by a sharp reduction of the h^2 which is probably related
387 to a certain degree of association between cecal microbiota and host genotype. This was even
388 observed for the case in which the definition of the **M** covariance matrix was based on the
389 weighted Unifrac distance matrix. However, for this particular case, we did not see any
390 improvement when considering microbial information for predicting ADG_{AL} or ADG_R. This

result highlights the need to accompany any assessment of the proportion of the phenotypic variance attributed to the microbial effect (i.e., microbiability) by validation of its actual predictive value.

The predictive value of models not including the microbial effect for growth traits was slightly higher (up to 0.46-0.48) with the fullDataset (i.e., that comprised of records from rabbits in which cecal microbiota was assessed and from their cage mates without such microbial information) than with the mDataset. In this case, however, the predictive value added by microbial information was more limited, being only significant for ADG_R of animals in which microbiota was assessed, and exclusively when the expansion of **M** for those animals without microbial information was based on the identity matrix. Despite this limited predictive value of the microbial information when the fullDataset was studied, and similar to that observed in some cases when the mDataset was considered, a very large percentage of the phenotypic variation of ADG_{AL} was estimated to be explained by cecal microbiota when the covariance matrix **M** was expanded using the identity matrix. The large estimates of m^2 for this trait can be said to be artifacts given that they are not accompanied by an improvement in the predictive capacity of the model, and they seem to be associated with an increase of the phenotypic variance estimates regarding M1. Such increase could be associated with an increment of the residual variance in the model, probably linked with the existence of a certain degree of collinearity between the covariance matrices of the different factors in the model. In this regard, the results obtained using covariance matrixes **M** expanded with cage-average CSS OTU counts could be said to be more coherent, since the null microbiability estimates are associated with a null improvement of the prediction of both growth traits (ADG_{AL} or ADG_R).

Fang et al. (2020) [15] found that only 10% of the phenotypic variance of finishing weight in commercial meat rabbits was explained by the gut microbiome. Besides that, previous studies in Japanese quails [16-Vollmar et al., 2020] and pigs [17-Camarinha-Silva et al., 2017] estimated m^2 for body weight gain of 0.18 and 0.28, respectively. These large differences between our current results for growth traits and the previous ones could be simply due to the study of different definitions of these traits in different species or to the use of different approaches and definitions of **M** to estimate m^2. We report a predictive value of cecal microbiota for ADG_{AL}, in line with that reported for daily gain in pigs by Camarinha-Silva et al. (2017) [17] applying microbial best linear unbiased prediction (M-BLUP) and by Maltecca et al. (2019) [18] using Bayesian models, machine learning approaches and semi-parametric kernel model. In our study, another important point to note is that the predictive value of cecal microbiota was higher for ADG_R than for ADG_{AL}. This result suggests that ADG_R is more strongly influenced by gut microbial composition than ADG_{AL}, which is more affected by host genetics as Piles and Sánchez (2019) [11] previously evidenced.

Regarding the study of cage-average phenotypes, the current difficulties in individually recording FI of rabbits bred in group suppose the major limitation to conduct a direct selection for FE. Therefore, definitions of FE in this study rely on group records of FI and individual records of growth. In addition to this constraint, in the current study, we have also faced the challenge that supposes not having microbial information for all the individuals of a cage. Our modeling approaches allow including the phenotypic information of cage mates on which cecal microbiota was not assessed. Thus, we present the first study to predict cage-average FI and FE traits in a rabbit sire line with a mixed model approach using microbial

information although it was only measured in approximately 30% of the individuals within cage. To deal with this limitation, we tested two different expansions of three microbial covariance matrices for the animals in which microbiota was not assessed to be able to consider the contributions of all individuals to the cage performance traits.

Our modeling approaches exhibited moderate predictive abilities for the cage-average phenotypes, higher than those obtained for the individually measured growth traits. This result was not surprising since the prediction of individual measures is more challenging than averages. Moreover, the inclusion of microbial information increased the predictive ability of mixed models by 5%, for $\overline{\text{ADFI}}_{\text{AL}}$, 20% for $\overline{\text{ADRFI}}_{\text{AL}}$ and 14% $\overline{\text{ADFCR}}_{\text{AL}}$ over the model not considering a microbial effect. It is worth mentioning that this improvement was only achieved when the expansion of the microbial relationship matrix for those animals without microbial information was based on the identity matrix (i.e., for those animals without microbial information the diagonal elements of the covariance matrix were set to one and elements outside the diagonal were fixed to zero). These improvements in the prediction were accompanied by large microbiability estimates, which in turn were associated with a reduction of heritability estimates. Clear evidence of ill-conditioned models was observed for those cases in which the expansion of the covariance matrices was based on cage-average CSS OTU counts given that large microbiabilities were estimated but they were not associated with improvements in the prediction, but with increased phenotypic variance estimates. The consideration of cage-average CSS counts to expand the covariance matrix could increased the collinearity between the individual microbial and the cage effects, deteriorated the parameters identification, and altered convergence properties (Additional file 4).

Previous studies have evaluated the value of gut microbiota to predict complex traits related to FE in other livestock species. In cattle, Delgado et al. (2019) [19] found a set of microbial contigs obtained from a *de novo* metagenome assembly that allowed high classification power for samples with extreme values of FE and FI traits. They found that these microbial contigs had a certain predictive ability for such traits in an independent cattle population. In pigs, Camarinha-Silva et al. (2017) [17] achieved higher prediction accuracies for FI and feed conversion with microbial best linear unbiased prediction (M-BLUP) method than with the same method but employing the genomic relationship matrix (G-BLUP). They quantified that 21% of the phenotypic variance of feed conversion in pigs is explained by the gut microbiome. In Japanese quails [16] and pigs [17], 9% and 16% of the phenotypic variance of FI, respectively, seem to be explained by the gut microbiome. In line with these studies estimating microbiabilities of traits related to FI and FE, we have also reported that a large percentage of the phenotypic variance of these phenotypes can be explained by the cecal microbiota. Such percentage was, in most cases, larger than that explained by the additive genetic effects. Nonetheless, as we have previously indicated, large microbiability estimates are not always associated with improvements in the predictive capacity of the models. Thus, such estimates should be interpreted with caution.

What seems clear from our results is that in those cases in which an improvement in the predictive ability of the model was evidenced, the estimated high microbiability was accompanied by a reduction in the heritability estimates with respect to those obtained in models not fitting the microbial effect. We interpret this as indirect evidence of certain host genetic control over the gut microbial composition. Several studies have already reported the

existence of moderate heritability for certain microbial taxa and diversity indexes in humans [20-Goodrich et al., 2014; 21-Goodrich et al., 2016], pigs [22-Lu et al., 2018; 23-Cheng et al., 2018; 24-Crespo-Piazuelo et al., 2018; 25-Reverter et al., 2021] or cattle [26-Sasson et al., 2017]. A preliminary study in the same meat rabbit population used in the current study has also directly shown that cecal microbiota is under genetic control [27-Velasco-Galilea et al., 2018]. These results are relevant from a biological perspective to better understand the symbiotic relationship between host and gut microbial communities, but also from an applied perspective. In the case we confirm that relevant OTUs (i.e., associated with performance traits of interest) have a clear host genetic control, selective breeding could be considered as an additional tool to promote the presence of such favorable microbial taxa in the gut of a given livestock population.

The predictive ability of multivariate sPLSR models for the traits under study did not improve by considering microbial information, except for $\overline{\text{ADRFI}}_{\text{AL}}$. This result was discouraging since with this approach we had hoped to identify the group of OTUs responsible of an improvement in the predictive ability. The unique case in which we identified a group of OTUs that appears to confer a predictive value was for $\overline{\text{ADRFI}}_{\text{AL}}$. We detected some unclassified OTUs belonging to family *Lachnospiraceae* moderately correlated with this trait, some of them positively and others negatively. This is not surprising given this is a big family encompassing numerous different genera. Siegerstetter et al. (2017) [28] found different *Lachnospiraceae* genera enriched in both low or high residual feed intake chickens and suggested that these bacteria could promote the host FE by stimulating fatty acid, amino acid, and vitamin synthesis. In short, with sPLSR we have not been able to detect the

improvement in the predictive ability observed with mixed models, suggesting the existence of an added value of microbial information that cannot be captured by all predictive machineries when the amount of data and microbial information are limited.

Our implemented mixed models approach integrates all the available pedigree information in the analysis. Such information is particularly relevant for the analysis of cage-average traits since it allows to share information between cages according to the additive genetic relationships. This way, predictions of individual phenotypes include variability between cage mates. However, the same cage-average measurement was assigned to all cage mates in the sPLSR model approach.

We have thus tried an alternative application of sPLSR models by fitting the posterior means of individual microbial effects estimated with M2 mixed models for each trait to identify the most relevant OTUs contributing to the improvement of the model predictive ability. This approximation has allowed us to identify for each trait a number of OTUs that are systematically chosen by the sPLSR models fitted with the three different matrices based on the identity matrix (i.e., those that we have found associated with gains in the predictive ability of the model) having the greatest loading weights.

We have detected four unclassified OTUs belonging to family *Lachnospiraceae* moderately correlated with growth traits: one positively and other negatively with ADG_R, and two positively with ADG_{AL}. This is not surprising given this is a big family encompassing numerous different genera. Fang et al. (2020) [15] identified a positive association between members of this family and ADG of commercial meat rabbits. Another study in the same

534 population of rabbits reported a positive association between members of family
535 *Lachnospiraceae* and finishing BW [8-Fang et al., 2020]. Interestingly, we have found two
536 different OTUs belonging to genus *Methanobrevibacter* positively associated with ADG_{AL}
537 and negatively with ADG_R. Kušar and Avguštin (2010) [29] suggested that methanogenic
538 microorganisms inhabiting the rabbit cecum are predominantly *Methanobrevibacter* species.
539 This result was supported by the study conducted by Velasco-Galilea et al. (2018) [4] in
540 which all archaeal species identified in the rabbit cecum and feces belonged to such
541 methanogenic genus that encompasses different hydrogenotrophic methane-producing
542 species. Conversely, McGovern et al. (2017) [30] and McCabe et al. (2015) [31] reported a
543 negative correlation between the abundance of this genus and body mass index, as well as an
544 overrepresentation of this genus in cattle under fed restriction.
545
546 We have identified a positive association between an unclassified member of family
547 *Ruminococcaceae* and ADG_R. This result is in agreement with the above-mentioned studies
548 in meat rabbits that also identified a positive association of this family with ADG and
549 finishing BW [15-Fang et al., 2020; 8-Fang et al., 2020]. Interestingly, we have found a
550 negative association between genus *Bacteroides* and ADG_R and $\overline{ADFI}_{AL}$, as well as between
551 genus *Butyricicoccus* and ADG_R. Genus *Bacteroides* has been associated with obesity in
552 humans [32- de la Cuesta-Zuluaga et al., 2018]. However, it is worth mentioning that this
553 genus encompasses pathogenic species, such as *Bacteroides fragilis* [33-Yekani et al., 2020],
554 that could lead to a diversion of nutrients from growth towards immune response. Previous
555 studies have hypothetized that an overgrowth of *Bacteroides* species in the rabbit gut could
556 lead to a decrease of butyrate yield and, consequently, to the incidence of epizootic rabbit
557 enteropathy [34-Jin et al., 2018]. Several studies have demonstrated that the application of

feed restriction after weaning reduces the risk of enteric disorders in rabbits [10-Gidenne et al., 2009; 35-Romero et al., 2010; 9-Gidenne et al., 2012]. In this regard, a lighter presence of genus *Bacteroides* in restricted animals could be associated with the benefits conferred by this feeding strategy. Previous studies, indeed, have found a negative correlation between this genus and pig BW [36-Mach et al., 2015; 37-Yang et al., 2016].

It is also noteworthy that we have identified three different OTUs taxonomically assigned to genus *Neglecta* that are negatively associated with $\overline{\text{ADFI}}_{\text{AL}}$. This genus encompasses pathogenic bacterial species, and it has been associated positively with pig ADG in a previous study conducted by Tran et al. (2018) [38]. On the other hand, we have identified two and five unclassified OTUs belonging to family *Lachnospiraceae* positively correlated with $\overline{\text{ADRFI}}_{\text{AL}}$ and $\overline{\text{ADFI}}_{\text{AL}}$, respectively. In cattle, in accordance with our results, Li and Guan (2017) [39] and Shabat et al. (2016) [40] found an overrepresentation of family *Lachnospiraceae* in less efficient animals (greater RFI). High relative abundances of members belonging to this family could suggest a more active cecum fermentation, which leads to increased butyrate short-chain fatty acid that is a nutrient for the gut of the animal. Besides that, we have found one OTU taxonomically assigned to genus *Olsenella* that seems to be relevant for the prediction of $\overline{\text{ADRFI}}_{\text{AL}}$, and that is positively associated with this trait. Members of this genus ferment starch and glycogen substrates to produce lactic, acetic, and formic acid [41-Göker et al., 2010]. In line with our results, Ellison et al. (2017) [42] and Kubasova et al. (2018) [43] reported higher abundances of *Olsenella* in the rumen of low feed efficient lambs and piglets, respectively.

581 On another note, we have found several OTUs relevant for the prediction of traits related to
582 FE analyzed in this study, i.e., $\overline{\mathrm{ADRFI}}_{\mathrm{AL}}$ and $\overline{\mathrm{ADFCR}}_{\mathrm{AL}}$. Two OTUs taxonomically assigned
583 to genus *Paramuribaculum* were found negatively correlated with $\overline{\mathrm{ADRFI}}_{\mathrm{AL}}$. Members of
584 this genus are involved in the metabolism of carbohydrates, lipids, vitamins, and amino acids
585 as well as in glycan biosynthesis [44- Lagkouvardos et al., 2019]. On the other hand, we have
586 identified OTUs belonging to class *Acidaminococcaceae* and genus *Negativibacillus*
587 positively correlated with $\overline{\mathrm{ADFCR}}_{\mathrm{AL}}$. Zhang et al. (2021) [45] suggested a role of genus
588 *Negativibacillus* in sheep feed efficiency throughout the fermentation of complex
589 carbohydrates. Conversely, Elolimy et al. (2020) [46] identified an enrichment of class
590 *Acidaminococcaceae* and genus *Negativibacillus* in the most efficient Holstein heifer calves.
591
592 Finally, we want to highlight that, in line with previous studies, we have observed that
593 bacterial members assigned to the same taxonomic group can either be positively or
594 negatively associated with a given phenotype. The observed heterogeneity in this study
595 includes members of family *Lachnospiraceae* and genera *Rumminoccocus*, *Butyricicoccus*,
596 and *Bacteroides*. This suggests that these OTUs belong to functionally and/or physiologically
597 different species encompassed within the same taxa.
598

599 ## Conclusions

600 Significant improvements in the prediction of individual growth and cage-average traits
601 related to FE were observed when cecal microbial information was fitted into the models.
602 However, these improvements are not general and depend to a large extent on the prediction
603 method used as well as on the prior information considered to define the covariance matrix

604 between animals due to their cecal microbial effect. We have introduced a novel modeling
605 approach based on the traditional mixed animal models that, relying on the pedigree
606 information, enables the estimation of variance components and the evaluation of the
607 predictive value of microbial information for cage-average performances even when
608 microbiota was not assessed in all individuals of the cage. Caution must be taken, however,
609 to interpret the magnitude of the proportion of the phenotypic variance explained by the
610 individual gut microbial effect since large microbiabilities estimates are not necessarily
611 associated with gains in the predictive ability of the model. Part of the effect associated with
612 the prediction improvement by considering cecal microbial information could be said to
613 partially have a genetic origin. In general, a certain drop in heritability estimates was
614 observed when both additive genetic and individual microbial effects were fitted at the time.
615 Cecal microbiota seems to have a polibacterial role in growth and FE traits since, although
616 we have identified certain OTUs with a relevant weight, a large proportion of OTUs are
617 responsible for the prediction improvement achieved with mixed models.
618

619 # Methods

620 Animals

621 All animals involved in the study were raised at the rabbit facilities of the Institute of
622 Agrifood Research and Technology (IRTA) in two different periods. The animals come from
623 the Caldes line [47-Gómez et al., 2002] that has been selected for post-weaning growth since
624 1983, and it is commonly used as a terminal sire line within the three-ways crossbreeding
625 schema for rabbit meat production in Spain. The animals used in this study were randomly
626 selected from 5 batches of a larger experiment conducted to estimate the effect of the

627 interaction between the genotype and the feeding regime on growth, feed efficiency, carcass
628 characteristics, and health status of the animals [11-Piles and Sánchez, 2019].
629
630 Most of the animals were produced in 4 batches in a semi-open-air facility during the first
631 semester of 2014, and the remaining were produced in a single batch in another facility under
632 better controlled environmental conditions in spring 2016. The animals bred in the first
633 facility were housed in collective cages containing 8 kits each one from weaning (32 days of
634 age) until the end of the fattening period (66 days of age). On the other hand, the kits raised
635 in the second facility were housed in cages of 6 kits each one and their growing period was
636 slightly shorter (32 - 60 days of age).
637
638 Beyond these differences, all animals received the same management and were fed with a
639 standard pelleted diet. Water was provided *ad libitum* and feed was supplied once per day in
640 a feeder with three places for the 4 - 5 weeks that the fattening lasted. At weaning, the animals
641 were randomly assigned to one of the two different feeding regimes under assessment: (1)
642 *ad libitum* (AL) or (2) restricted (R) to 75% of the AL FI. The amount of feed supplied to the
643 animals under R in each week for each batch was computed as 0.75 times the average FI of
644 kits on AL from the same batch during the previous week, plus 10% to account for a FI
645 increase as the animals grow. Kits under both feeding regimes were categorized into two
646 groups according to their BW at weaning (big if their BW was greater than 700 g or small
647 otherwise) to generate homogeneous groups regarding animal size within feeding regime. A
648 maximum of two kits from the same litter were assigned to a single cage to avoid confounding
649 between cage and maternal effects.
650

651 The individual BW was weekly recorded for all animals in both feeding regimes, and the
652 cage FI was also weekly recorded in AL cages. From BW raw records, individual ADG was
653 computed as the slope of the within animal regression of all BW measurements on their
654 respective ages in days. This trait was individually computed for each feeding regime, thus
655 obtaining ADG on AL (ADG_{AL}) or under R (ADG_{R}). For the AL cages, three additional traits
656 were computed. The individual average daily feed intake ($\overline{ADFI}_{AL}$) was computed as the total
657 FI of the cage during the whole growing period divided by the number of days and the number
658 of kits that each cage contained. The individual average daily residual feed intake ($\overline{ADRFI}_{AL}$)
659 was obtained as the residual of a batch-nested multiple regression of $\overline{ADFI}_{AL}$ on the $\overline{ADG}_{AL}$
660 and the cage-average mid-growing-period day metabolic weight ($\overline{MW}_{AL}$). Finally, the
661 individual average daily feed conversion ratio ($\overline{ADFCR}_{AL}$) was computed as the ratio between
662 $\overline{ADFI}_{AL}$ and the ADG_{AL} cage-average ($\overline{ADG}_{AL}$).
663
664 Two different datasets were considered for the analyses performed in this study. The
665 mDataset was represented by the 425 kits from which cecal samples were collected at the
666 end of their growing period for microbiota assessment, and the fullDataset included these
667 425 kits and their cage mates. On average, cecal microbiota was assessed in 2 kits by cage.
668 The number of animals and cages within feeding regime and batch are shown in Table 9, and
669 the descriptive statistics of the traits under study are presented in Table 10.
670
671 **Table 9 Number of individual and cages within feeding regime and batch. Animals with**
672 **microbiota assessed and non-assessed are distinguished for the individual records.**

	Individuals		Cages
	With microbiota	W/o microbiota	

Batch	R	AL	R	AL	R	AL
1	45	44	51	52	16	16
2	30	27	66	61	12	11
3	41	35	103	84	18	15
4	53	61	195	211	31	34
5	32	57	96	126	16	23

R: Animals under restriction; AL: animals fed *ad libitum*.

Table 10 Descriptive statistics of growth and FE traits.

Trait	Dataset	N	Mean	SD	IQR
ADG_{AL} (g/day)[1]	mDataset	224	55.12	6.52	7.30
ADG_{AL} (g/day)[1]	fullDataset	758	53.21	9.42	8.49
ADG_{R} (g/day)[1]	mDataset	201	36.35	5.85	7.56
ADG_{R} (g/day)[1]	fullDataset	712	35.35	7.99	8.27
$ADFI_{AL}$ (g/day)[2]	fullDataset	99	151.37	17.01	20.93
$ADRFI_{AL}$ (g/day)[2]	fullDataset	99	0.00	5.92	6.66
$ADFCR_{AL}$ (g/day)[2]	fullDataset	99	2.84	0.24	0.33

ADG_{AL}: average daily gain in rabbits fed *ad libitum*; ADG_{R}: average daily gain in rabbits fed under restriction; $\overline{ADFI}_{AL}$: average daily feed intake in rabbits fed *ad libitum*; $\overline{ADRFI}_{AL}$: average daily residual feed intake in rabbits fed *ad libitum*; $\overline{ADFCR}_{AL}$: average daily feed conversion ratio in rabbits fed *ad libitum*; SD: standard deviation; IQR: interquartile range; mDataset: dataset including only records of animals in which microbiota was assessed; fullDataset: dataset including records of animals in which microbiota was assessed as well as of their cage mates.
[1]Refers to individual traits.
[2]Refers to cage traits.

Sample processing, DNA extraction and sequencing

Animals were slaughtered (at 66 and 60 days of age in first and second facility, respectively) and cecal samples of 425 rabbits were collected in a sterile tube, kept cold in the laboratory (4ºC), and stored at -80ºC. DNA extraction, amplification, Illumina library preparation and sequencing followed methods described in previous studies [4-Velasco-Galilea et al., 2018; 5-Velasco-Galilea et al., 2020]. Whole genomic DNA was extracted from 250 mg of each

691 biological sample according to manufacturer's instructions of kit ZR Soil Microbe DNA
692 MiniPrep™ kit (ZymoResearch, Freiburg, Germany). Cecal samples were mechanically
693 lysed in a FastPrep-24™ Homogenizer (MP Biomedicals, LLC, Santa Ana, CA, United
694 States) at a speed of 6 m/s for 60 s, thus facilitating an efficient lysis of archaeal and bacterial
695 species. Integrity and purity of DNA extracts were measured with Nanodrop ND-1000
696 spectrophotometer equipment (NanoDrop products; Wilmington, DE, United States)
697 following Desjardins and Conklin's protocol [48- Desjardins and Conklin, 2010]. All DNA
698 extracts showed adequate integrity and purity (absorbance ratio 260 nm/280 nm > 1.6) to
699 avoid PCR inhibition issues. A fragment of the 16S rRNA gene that included the V4-V5
700 hypervariable regions was amplified with the F515Y/R926 pair of primers (5'-
701 GTGYCAGCMGCCGCGGTAA-3', 5'-CCGYCAATTYMTTTRAGTTT-3') [49-Parada et
702 al., 2016]. The initial polymerase chain reaction (PCR) was conducted for each sample using
703 12.5 µl 2x KAPA HiFi HotStart Ready Mix, 5 µl forward primer, 5 µl reverse primer and 2.5
704 µl template DNA (5 ng/ µl). The PCR conditions were the following: initial denaturation for
705 3 minutes at 95 ºC, 25 cycles of 30 seconds at 95 ºC, 30 seconds at 55 ºC and 30 seconds at
706 72 ºC; and final extension for 2 minutes at 72 ºC. The fragment was then re-amplified in a
707 limited-cycle PCR reaction to add sequencing adaptors and 8 nucleotide dual-indexed
708 barcodes of the multiplex Nextera® XT kit (Illumina, Inc., San Diego CA, United States)
709 according to manufacturer's instructions. The adaptors and barcodes were added to both ends
710 of the fragment in a second PCR by using 25 µl 2x KAPA HiFi HotStart Ready Mix, 5 µl
711 index i7, 5 µl index i5, 10 µl PCR Grade water and 5 µl concentrated amplicons of the initial
712 PCR. The second PCR conditions were the following: initial denaturation for 3 minutes at 95
713 ºC, 8 cycles of 30 seconds at 95 ºC, 30 seconds at 55 ºC and 30 seconds at 72 ºC; and final
714 extension for 5 minutes at 72 ºC. Final libraries were cleaned up with AMPure XP beads,

715 validated by running 1 µl of a 1:50 dilution on a Bioanalyzer DNA 1000 chip (Agilent
716 Technologies, Inc., Santa Clara, CA, United States) to verify their size, quantified by
717 fluorometry with PicoGreen dsDNA quantification kit (Invitrogen, Life Technologies,
718 Carlsbad, CA, United States), pooled at equimolar concentrations and paired-end sequenced
719 in 5 parallel plates in a Illumina MiSeq 2 x 250 platform at the Genomics and Bioinformatics
720 Service (SGB) of the Autonomous University of Barcelona (UAB).
721
722 Bioinformatic pipeline for OTU calling
723 Sequence processing was performed using QIIME software (version 1.9.0) [50- Caporaso et
724 al., 2010] as described in 5-Velasco-Galilea et al. 2020. The first step consists of assembling
725 the paired-ended V4-V5 16S rRNA gene reads into contigs with the python script
726 *multiple_join_paired_ends.py*. The resulting contigs were filtered (those with a quality score
727 smaller than Q19 were discarded) and assigned to samples using the python script
728 *split_libraries.py* with default parameters. Chimeric sequences generated in the PCR were
729 detected with UCHIME algorithm [51- Edgar et al., 2011] and removed. The filtered contigs
730 were clustered into operational taxonomic units (OTUs) with a 97% similarity threshold
731 using the script *pick_open_reference_otus.py* with default parameters [52- Rideout et al.,
732 2014]. This script uses the UCLUST algorithm [53- Edgar, 2010], to first align the sequences
733 against Greengenes reference database (version gg_13_5_otus) [54- McDonald et al., 2012],
734 and then to make a *de novo* clustering of those contigs that did not match the database. After
735 doubletons removal, the filtered OTU table contained the sequence counts of 963 OTUs for
736 425 samples. Finally, the OTU table was normalized with the cumulative sum scaling (CSS)
737 method [55- Paulson et al., 2013]. Taxonomic assignment of representative sequences of

each OTU was conducted with the QIIME default parameters of the UCLUST consensus taxonomy assigner by mapping the sequences against the Greengenes reference database gg_13_5_otus. The raw sequence data were deposited in the sequence read archive of NCBI under the BioProject accession number PRJNA524130. Metadata, OTU table, and corresponding taxonomic assignments are also included as Additional files 1, 2 and 3, respectively. In summary, after executing the bioinformatic processing, 14,928,203 filtered sequences clustered into 963 different OTUs were obtained for 425 cecal rabbit samples. Most of these OTUs were assigned to phyla *Firmicutes* (76.74%), *Tenericutes* (7.22%) and *Bacteroidetes* (6.26%). Details on the taxonomic assignment can be found at Velasco-Galilea et al. (2020) [5].

Statistical analyses: mixed models

(I) Parameter estimation:

The following univariate microbial mixed linear model was fitted to estimate the marginal posterior distributions of additive, litter, cage, and microbial effects of the individual growth traits ADG_{AL} and ADG_R with the mDataset:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Z_A a} + \mathbf{Z_L l} + \mathbf{Z_C c} + \mathbf{Z_M m} + \mathbf{e},$$

where **y** is a vector containing the phenotypes (ADG_{AL} or ADG_R); **β** is a vector of the systematic effects of batch (5 levels) and of BW at weaning (2 levels: big and small) with its corresponding incidence matrix **X**; **a** is a vector including the additive genetic effects with the corresponding incidence matrix $\mathbf{Z_A}$; **l** is a vector with litter birth effects with the corresponding incidence matrix $\mathbf{Z_L}$; **c** is a vector including cage effects with the corresponding incidence matrix $\mathbf{Z_C}$; **m** is a vector having the animal microbial effects with

761 the corresponding incidence matrix $\mathbf{Z_M}$; finally **e** is a vector of residuals. The mDataset used
762 in these analyses included phenotypic information of 425 rabbits born from 318 litters and
763 housed in 192 cages, while the pedigree included relationships of 2,547 individuals.
764
765 The fullDataset was used to estimate the marginal posterior distributions of additive, litter,
766 and microbial effects of $\overline{\text{ADFCR}}_{\text{AL}}$, $\overline{\text{ADFI}}_{\text{AL}}$ and $\overline{\text{ADRFI}}_{\text{AL}}$ from records on the 99 AL cages
767 available. The following univariate microbial mixed linear was fitted:

768 $$\mathbf{y = X\beta + Z_A a + Z_L l + Z_M m + e,}$$

769 where **y** is a vector containing cage trait phenotypes ($\overline{\text{ADFCR}}_{\text{AL}}$, $\overline{\text{ADFI}}_{\text{AL}}$ or $\overline{\text{ADRFI}}_{\text{AL}}$); **β** is a
770 vector including the systematic effects of batch (5 levels) and of BW at weaning (2 levels:
771 big and small) with its corresponding incidence matrix **X**. As described above, vectors **a**, **l**,
772 **m** and **e** correspond to additive genetic, litter birth, animal microbial and residual effects,
773 respectively. However, the corresponding incidence matrices $\mathbf{Z_A}$, $\mathbf{Z_L}$ and $\mathbf{Z_M}$ are not
774 composed by zeros and ones but by real numbers representing the proportions of the different
775 levels of the factor contributing to the cage-average.
776
777 In both models, the same sets of prior distributions were considered for the different factors.
778 The systematic effects (**β**) were *a priori* assumed to follow uniform distributions. The
779 assumed prior distribution for the additive genetic effects was $\mathbf{a} \sim NMV\ (\mathbf{0}, \mathbf{A}\sigma^2_{\text{A}})$, with **A**
780 being the numerator relationship matrix [56-Henderson, 1973] and σ^2_{A} being the additive
781 genetic variance. The prior distribution assumed for the litter effects was $\mathbf{l} \sim NMV\ (\mathbf{0}, \mathbf{I}\sigma^2_{\text{L}})$,
782 with **I** being an identity matrix of appropriate dimension, and σ^2_{L} being the litter birth
783 variance. The prior distribution for the cage effects was $\mathbf{c} \sim NMV\ (\mathbf{0}, \mathbf{I}\sigma^2_{\text{C}})$, with **I** also being

784 an identity matrix of appropriate dimension, and σ_C^2 being the cage variance. In different
785 analyses, alternative prior distributions were assumed for the vector of animal-specific
786 microbial effects, being its general form $\mathbf{m} \sim NMV\ (\mathbf{0}, \mathbf{M}\sigma_M^2)$, with $\mathbf{M}$ being a between-
787 animals relationship matrix due to microbial effects, and σ_M^2 being the animal microbial
788 variance. Three alternative definitions of $\mathbf{M}$ were considered in three separate analyses: i)
789 $\mathbf{M_O} = \mathbf{OO'}$, with $\mathbf{O}$ being the row-normalized CSS OTU count matrix, [n (animals) x m
790 (OTUs)]; the $\mathbf{O}$ matrix was row-wise normalized by dividing the row vector elements by the
791 row norms ensuring that $\mathbf{M_O}$ had ones in its diagonal (this definition is fairly similar to that
792 previously proposed by Difford et al. (2018) [57]; ii) $\mathbf{M_B} = 1 - \frac{\mathbf{B^2}}{2}$; with $\mathbf{B}$ being the Bray-
793 Curtis distance matrix [58-Bray and Curtis, 1957] computed from the CSS OTU count
794 matrix; and iii) $\mathbf{M_U} = 1 - \frac{\mathbf{U^2}}{2}$; with $\mathbf{U}$ being the weighted Unifrac distance matrix [59-
795 Lozupone and Knight, 2005] computed from the CSS OTU count matrix. Both distance
796 matrices ($\mathbf{B}$ and $\mathbf{U}$) were computed using the "phyloseq" R package [60-McMurdie and
797 Holmes, 2013].
798
799 To deal with the fact that microbial information was only available for some of the rabbits
800 within a cage, it was necessary to generate the rows and columns of the between-animal
801 covariance matrices due to the cecal microbial content for the animals not having microbial
802 information assessed. This approach allows to consider the contributions of all individuals to
803 the cage-average performance traits. Two different expansion strategies were adopted: i)
804 assigning to the animals without microbial information the within cage-average of each CSS
805 OTU count, and then computing $\mathbf{M_{\bar{O}}}$, $\mathbf{M_{\bar{B}}}$ and $\mathbf{M_{\bar{U}}}$ between the 1,470 animals under study (425
806 having microbial information plus their cage mates without microbial information); ii) first

807 computing $\mathbf{M_O}$, $\mathbf{M_B}$ and $\mathbf{M_U}$ from the 425 animals with microbial information and then
808 expanding with ones in the diagonal and zeros out of the diagonal the rows and columns
809 corresponding to animals not having microbial information, thus obtaining $\mathbf{M_{O,0}}$, $\mathbf{M_{B,0}}$ and
810 $\mathbf{M_{U,0}}$. The resulting covariance matrices were forced to be positive definite by conducting an
811 eigen-value decomposition, saving all the positive eigen-values and their associated eigen-
812 vectors, and finally reconstructing the covariance matrices from these elements. Note that the
813 original (obtained between the 425 animals having microbial composition) Bray-Curtis or
814 unweighted Unifrac distance matrices could be undefined matrices, i.e., mixing positive and
815 negative eigen values, since distance matrices are pairwise constructed. Thus, certain
816 incongruities could exist when the distances are studied beyond pairs of individuals, which
817 translate into non-positive definition of the whole distance matrix. These incongruities must
818 be corrected if the distance matrix is going to be used as a covariance matrix.
819
820 The MCMC Bayesian estimation procedure was conducted using gibbsf90test program [61-
821 Misztal et al., 2015]. Chains of 2,000,000 samples were run discarding the first 500,000 to
822 allow the algorithm to reach convergence to the marginal posterior distributions. Finally, one
823 in every 10 samples was saved. Trace plots and histograms of Markov chains from the
824 posterior distribution of the parameters of Bayesian models fitted for the individual growth
825 traits and for the cage FE traits are included as Additional file 4.
826
827 The fractions of the phenotypic variance of ADG_{AL} and ADG_R explained by σ^2_A (heritability),
828 σ^2_L (litter variance ratio), σ^2_C (cage variance ratio), and σ^2_M (microbiability; [57-Difford et al.,
829 2018]) were calculated as:

$$h^2 = \frac{\sigma_A^2}{\sigma_P^2};\ l^2 = \frac{\sigma_L^2}{\sigma_P^2};\ c^2 = \frac{\sigma_C^2}{\sigma_P^2};\ m^2 = \frac{\sigma_M^2}{\sigma_P^2},$$

where $\sigma_P^2 = \sigma_A^2 + \sigma_L^2 + \sigma_C^2 + \sigma_M^2 + \sigma_e^2$ is the phenotypic variance.

Similarly, for the cage traits ($\overline{ADFCR}_{AL}$, $\overline{ADFI}_{AL}$ and $\overline{ADRFI}_{AL}$), the fractions of the phenotypic variance explained by σ_A^2 (heritability), σ_L^2 (litter variance ratio), and σ_M^2 (microbiability) were calculated as:

$$h^2 = \frac{\sigma_A^2}{\sigma_P^2};\ l^2 = \frac{\sigma_L^2}{\sigma_P^2};\ \ m^2 = \frac{\sigma_M^2}{\sigma_P^2},$$

where $\sigma_P^2 = \sigma_A^2 + \sigma_L^2 + \sigma_M^2 + 7\sigma_e^2$ is the phenotypic variance. Given that σ_e^2 represents the cage residual mean, it is necessary to multiply it by 7 (the average number of animals within cage in this study), thus obtaining an individual residual variance estimate referred to individual records. Note that l^2 and c^2 were defined but related results are not presented in this study.

(II) Predictive ability assessment:

For each trait, two cross-validations assessments were conducted to evaluate whether including microbial information in the model improves its predictive ability. The first one was based on the above-described mixed model whose predictive performance was compared with that of the same model but without considering the microbial effect. Cross-validations were replicated 100 times. In each of them, the dataset for the individually measured traits (ADG_{AL} and ADG_R) was randomly split into training and validation sets with probabilities 0.9 and 0.1, respectively. This partition was done in a manner that ensured all litters and cages of the animals in the validation set were also represented in the training set. For the

cage traits ($\overline{\text{ADFCR}}_{AL}$, $\overline{\text{ADFI}}_{AL}$ and $\overline{\text{ADRFI}}_{AL}$), the dataset was randomly split in a way that cages within a given batch were assigned to the training or the testing set with probabilities 0.8 and 0.2, respectively. The predictive ability of each model was defined as the average, across 100 replicates, correlation coefficient between predicted and observed phenotypes in the validation set. In this cross-validation assessment, the training step of the model was conducted using the expectation–maximization residual maximum likelihood (EM-REML) algorithm as implemented in the program remlf90 [61-Misztal et al., 2015]. Paired t test [62-R] was applied to compare the across replicates mean correlations obtained with the model considering microbial effect to that from the model that ignored this information. The tests were assumed paired because the same dataset was used in each replicate of both analyses (i.e., with and without bacterial effect). Empirical bootstrap p-values for the paired t test were computed after generating 1,000 bootstrap samples under the null hypothesis of the correlation coefficients from both models across the 100 replicates. The bootstrap p-value was defined as the proportion of bootstrap rounds having an estimated difference equal to or greater than that obtained with the original dataset. A p-value lower than 0.05, after Bonferroni correction [63-Bonferroni, 1936], was considered to support the rejection of the hypothesis of both models having the same predictive ability. In those cases where the null hypothesis was rejected, the percentage of times across the 100 replicates that the correlation coefficient obtained with the model considering microbial information was higher than that obtained with the model that ignored such information was computed.

Statistical analyses: multivariate models

(I) Predictive ability assessment:

875 Another predictive performance assessment was conducted using a multivariate approach.
876 Individual (ADG_{AL} and ADG_{R}) and cage traits ($\overline{ADFCR}_{AL}$, $\overline{ADFI}_{AL}$ and $\overline{ADRFI}_{AL}$) were fitted
877 with sparse Partial Least Squares Regression (sPLSR) models. The predictors of the first
878 sPLSR model where the columns of the design matrix obtained with the *model.matrix()* R
879 function [62-R] after fitting for each trait a linear model defined by the same systematic
880 effects as those used in the mixed model approach (i.e., batch and body size at weaning). The
881 second sPLSR model fitted for each trait include as predictors the abovementioned
882 systematic effects together with the 946 CSS OTU counts which were detected in at least 5%
883 of the samples and had a sum of its counts resulting in a frequency greater than 0.01% of the
884 total sum of all OTUs counts across all samples. CSS OTU counts on the 425 rabbits having
885 measures of gut microbial composition were directly used for the analysis of the individual
886 growth records. For the cage-average traits, it was needed to associate these cage-average
887 performances to the cage-average CSS OTU counts. For each trait, the corresponding dataset
888 was randomly divided into 5 folds, 4 of which constituted the learning dataset, and the
889 remaining was used as the validation dataset. Before fitting the sPLSR on the learning
890 dataset, optimal tuning parameters sparsity and number of latent components were chosen by
891 an internal 5-fold cross-validation using *cv.spls()* function of the "spls" R package [64-Chung
892 et al., 2019] within ranges (0.01-0.99) and (1-20) for sparsity and number of latent
893 components, respectively. With the tuning parameters returned by the *cv.spls()* function, the
894 combination that resulted in the minimum mean squared prediction error (MSPE) was used
895 to finally fit the sPLSR to the learning dataset by the function *spls()*. Then, the fitted sPLSR
896 model was used to predict the host trait performances of the validation dataset. This process
897 was replicated 20 times with different seeds, thus obtaining 100 replicates for each trait and
898 model tested. The predictive ability of each model was defined as the average, across 100

899 replicates, correlation coefficient between predicted and observed host trait phenotypes in
900 the validation dataset. The significance of the differences in the correlation coefficient
901 between observed and predicted records across these 100 replicates was tested using the
902 bootstrap paired t tests previously described for the mixed model analysis. In this case, the
903 comparison involved the correlations between observed and predicted records obtained with
904 a model just fitting the systematic effects and with other model fitting both systematic effects
905 and CSS OTU counts. Additionally, when the predictive ability of the model including the
906 microbial information was declared as better than that obtained with that of the model only
907 including the systematic effects as predictors, the taxonomy of those OTUs selected in more
908 than 80% of the sPLSR replicates was studied with the reference taxonomic database RDP
909 [65-Wang et al., 2007]. Finally, the Pearson's correlation was computed to quantify the
910 degree of association between selected OTUs and the trait of interest.
911
912 *(II) Identification of relevant OTUs:*
913 Multivariate sPLSR models were also used to fit the posterior means of the individual
914 microbial effects predicted with the univariate microbial mixed linear models that led to a
915 significant prediction improvement of growth and FE traits. This approach was conducted in
916 an attempt to identify the most relevant OTUs for the prediction of such phenotypes. In each
917 case, the microbial composition records associated with the animals that conformed the
918 mDataset were randomly divided into 5 folds (1 and 4 folds constituted the validation and
919 the learning dataset, respectively). Before fitting the sPLSR on the learning dataset, optimal
920 tuning parameters sparsity and number of latent components were chosen by an internal 5-
921 fold cross-validation using *cv.spls()* function of the "spls" R package as described above. A

922 sPLSR model was then fitted to the learning dataset by the function *spls()* with the tuning
923 parameters returned by the *cv.spls()* function using the 946 CSS OTU counts as predictors.
924 This process was replicated 20 times with different seeds for each trait and model tested to
925 select those OTUs chosen in at least 80 out of the 100 replicates conducted. The OTUs
926 considered as relevant for the prediction of a given trait were those having the greatest
927 loading weights (i.e., below 5^{th} and above 95^{th} percentile values) and that were selected with
928 all the models tested. The taxonomy of the relevant OTUs was studied with the reference
929 taxonomic database RDP and the Pearson's correlation was computed to quantify the degree
930 of association between each OTU and the trait of interest.
931

932 ## References

933 [1] Cartuche, L., Pascual, M., Gómez, E. A. & Blasco, A. Economic weights in rabbit meat
934 production. *World Rabbit Science*, **22(3)**, 165-177 (2014).
935

936 [2] Estany, J., Camacho, J., Baselga, M. & Blasco, A. Selection response of growth rate in
937 rabbits for meat production. *Genetics Selection Evolution*, **24(6)**, 527-537 (1992).
938

939 [3] Piles, M., Gomez, E. A., Rafel, O., Ramon, J. & Blasco, A. Elliptical selection experiment
940 for the estimation of genetic parameters of the growth rate and feed conversion ratio in
941 rabbits. *Journal of animal science*, **82(3)**, 654-660 (2004).
942

943 [4] Velasco-Galilea, M. *et al*. Rabbit microbiota changes throughout the intestinal
944 tract. *Frontiers in microbiology*, **9**, 2144 (2018).

945

946 [5] Velasco-Galilea, M. *et al*. Breeding farm, level of feeding and presence of antibiotics in
947 the feed influence rabbit cecal microbiota. *Animal Microbiome*, **2(1)**, 1-16 (2020).
948
949 [6] Zeng, B. *et al*. The bacterial communities associated with fecal types and body weight of
950 rex rabbits. *Scientific reports*, **5(1)**, 1-8 (2015).
951
952 [7] Drouilhet, L. *et al*. Direct and correlated responses to selection in two lines of rabbits
953 selected for feed efficiency under ad libitum and restricted feeding: I. Production traits and
954 gut microbiota characteristics. *Journal of animal science*, **94(1)**, 38-48 (2016).
955
956 [8] Fang, S. *et al*. Dynamic distribution of gut microbiota in meat rabbits at different growth
957 stages and relationship with average daily gain (ADG). *BMC microbiology*, **20**, 1-13 (2020).
958
959 [9] Gidenne, T., Combes, S. & Fortun-Lamothe, L. Feed intake limitation strategies for the
960 growing rabbit: effect on feeding behaviour, welfare, performance, digestive physiology and
961 health: a review. *Animal*, **6(9)**, 1407-1419 (2012).
962
963 [10] Gidenne T. *et al*. Feed restriction strategy in the growing rabbit. 2. Impact on digestive
964 health, growth and carcass characteristics. *Animal*, **3(4)**, 509-515 (2009).
965
966 [11] Piles, M. & Sánchez, J. P. Use of group records of feed intake to select for feed efficiency
967 in rabbit. *Journal of animal breeding and genetics*, **136(6)**, 474-483 (2019).
968

969 [12] Sánchez, J. P. *et al*. Genome-wide association study for feed efficiency in collective
970 cage-raised rabbits under full and restricted feeding. *Animal Genetics*, **51(5)**, 799-810 (2020).
971
972 [13] Ross, E. M., Moate, P. J., Marett, L. C., Cocks, B. G. & Hayes, B. J. Metagenomic
973 predictions: from microbiome to complex health and environmental phenotypes in humans
974 and cattle. *PloS one*, **8(9)**, e73056 (2013).
975
976 [14] Ramon, E. *et al*. kernInt: A Kernel Framework for Integrating Supervised and
977 Unsupervised Analyses in Spatio-Temporal Metagenomic Datasets. *Frontiers in*
978 *microbiology*, **12**, 60 (2021).
979
980 [15] Fang, S. *et al*. Effects of gut microbiome and short-chain fatty acids (SCFAs) on
981 finishing weight of meat rabbits. *Frontiers in Microbiology*, **11**, 1835 (2020).
982
983 [16] Vollmar, S. *et al*. The gut microbial architecture of efficiency traits in the domestic
984 poultry model species Japanese quail (Coturnix japonica) assessed by mixed linear models.
985 *G3: Genes, Genomes, Genetic*, **10(7)**, 2553-2562 (2020).
986
987 [17] Camarinha-Silva, A. *et al*. Host genome influence on gut microbial composition and
988 microbial prediction of complex traits in pigs. *Genetics*, **206(3)**, 1637-1644 (2017).
989
990 [18] Maltecca, C. *et al*. Predicting growth and carcass traits in swine using microbiome data
991 and machine learning algorithms. *Scientific reports*, **9(1)**, 1-15 (2019).
992

[19] Delgado, B. *et al*. Whole rumen metagenome sequencing allows classifying and predicting feed efficiency and intake levels in cattle. *Scientific reports*, **9(1)**, 1-13 (2019).

[20] Goodrich, J. K. *et al*. Human genetics shape the gut microbiome. *Cell*, **159(4)**, 789-799 (2014).

[21] Goodrich, J. K. *et al*. Genetic determinants of the gut microbiome in UK twins. *Cell host & microbe*, **19(5)**, 731-743 (2016).

[22] Lu, D. *et al*. Host contributes to longitudinal diversity of fecal microbiota in swine selected for lean growth. *Microbiome*, **6(1)**, 1-15 (2018).

[23] Cheng, P. *et al*. Exploratory analysis of the microbiological potential for efficient utilization of Fiber between Lantang and Duroc pigs. *Frontiers in Microbiology*, **9**, 1342 (2018).

[24] Crespo-Piazuelo, D. *et al*. Characterization of bacterial microbiota compositions along the intestinal tract in pigs and their interactions and functions. *Scientific reports*, **8(1)**, 1-12 (2018).

[25] Reverter, A. *et al*. A gene co-association network regulating gut microbial communities in a Duroc pig population. *Microbiome*, **9(1)**, 1-16 (2021).

[26] Sasson, G. *et al*. Heritable bovine rumen bacteria are phylogenetically related and correlated with the cow's capacity to harvest energy from its feed. *MBio*, **8(4)**, (2017).

[27] Velasco-Galilea, M. *et al*. Determinismo genético de la microbiota intestinal del conejo. In: XIX Reunión Nacional de Mejora Genética Animal; 2018 June 14-15; León, Spain. www.acteon.webs.upv.es/CONGRESOS/Z-XIX_Reunion_MG_LEON_2018/043_VelascoGalilea.pdf

[28] Siegerstetter, S. C. *et al*. Intestinal microbiota profiles associated with low and high residual feed intake in chickens across two geographical locations. *PloS one*, **12(11)**, e0187766 (2017).

[29] Kušar, D. & Avguštin, G. Molecular profiling and identification of methanogenic archaeal species from rabbit caecum. *FEMS Microbiology Ecology*, **4(3)**, 623-630 (2010).

[30] McGovern, E. *et al*. Plane of nutrition affects the phylogenetic diversity and relative abundance of transcriptionally active methanogens in the bovine rumen. *Scientific reports*, **7(1)**, 1-10 (2017).

[31] McCabe, M. S. *et al*. Illumina MiSeq phylogenetic amplicon sequencing shows a large reduction of an uncharacterised Succinivibrionaceae and an increase of the Methanobrevibacter gottschalkii clade in feed restricted cattle. *PloS one*, **10(7)**, e0133234 (2015).

[32] De La Cuesta-Zuluaga, J. *et al*. Body size phenotypes comprehensively assess cardiometabolic risk and refine the association between obesity and gut microbiota. *International Journal of Obesity*, **42(3)**, 424-432, (2018).

[33] Yekani, M. *et al*. Y. To resist and persist: Important factors in the pathogenesis of Bacteroides fragilis. *Microbial Pathogenesis*, **149**, 104506 (2020).

[34] Jin, D. X. *et al*. The underlying microbial mechanism of epizootic rabbit enteropathy triggered by a low fiber diet. *Scientific reports*, **8(1)**, 1-15 (2018).

[35] Romero, C., Cuesta, S., Astillero, J. R., Nicodemus, N. & De Blas, C. Effect of early feed restriction on performance and health status in growing rabbits slaughtered at 2 kg live-weight. *World Rabbit Science*, **18(4)**, 211-218 (2010).

[36] Mach, N. *et al*. Early-life establishment of the swine gut microbiome and impact on host phenotypes. *Environmental microbiology reports*, **7(3)**, 554-569 (2015).

[37] Yang, H. *et al*. Uncovering the composition of microbial community structure and metagenomics among three gut locations in pigs with distinct fatness. *Scientific reports*, **6(1)**, 1-11 (2016).

[38] Tran, H. *et al*. Effects of spray-dried porcine plasma on fecal microbiota in nursery pigs. *Journal of animal science*, **96(3)**, 1017-1031 (2018).

[39] Li, F. & Guan, L. Metatranscriptomic profiling reveals linkages between the active rumen microbiome and feed efficiency in beef cattle. *Applied and environmental microbiology*, **83(9)**, e00061-17 (2017).

[40] Shabat, S. K. B. *et al*. Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants. *The ISME journal*, **10(12)**, 2958-2972 (2016).

[41] Göker, M. *et al*. Complete genome sequence of Olsenella uli type strain (VPI D76D-27C T). Standards in Genomic Sciences, **3(1)**, 76-84 (2010).

[42] Ellison, M. J. *et al*. Diet and feed efficiency status affect rumen microbial profiles of sheep. *Small Ruminant Research*, **156**, 12-19 (2017).

[43] Kubasova, T. *et al*. Effects of host genetics and environmental conditions on fecal microbiota composition of pigs. *PLoS One*, **13(8)**, e0201901 (2018).

[44] Lagkouvardos, I. *et al*. Sequence and cultivation study of Muribaculaceae reveals novel species, host preference, and functional potential of this yet undescribed family. *Microbiome*, **7(1)**, 1-15 (2019).

[45] Zhang, Y. K. *et al*. Characterization of the rumen microbiota and its relationship with residual feed intake in sheep. *Animal*, **15(3)**, 100161 (2021).

[46] Elolimy, A., Alharthi, A., Zeineldin, M., Parys, C. & Loor, J. J. Residual feed intake divergence during the preweaning period is associated with unique hindgut microbiome and metabolome profiles in neonatal Holstein heifer calves. *Journal of animal science and biotechnology*, **11(1)**, 1-15 (2020).

[47] Gómez, E. A., Rafel, O. & Ramon, J. The Caldes strain (Spain). *Options Méditerranéennes. Série B: Etudes et Recherches (CIHEAM)*. (2002).

[48] Desjardins, P. & Conklin, D. NanoDrop microvolume quantitation of nucleic acids. Journal of visualized experiments: *JoVE*, **45**, (2010).

[49] Parada, A. E., Needham, D. M. & Fuhrman, J. A. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental microbiology*, **18(5)**, 1403-1414 (2016).

[50] Caporaso, J. G. *et al*. QIIME allows analysis of high-throughput community sequencing data. *Nature methods*, **7**, 335-336 (2010).

[51] Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, **27**, 2194-2200 (2011).

[52] Rideout, J. R. *et al*. Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. *PeerJ*, **2**, e545 (2014).

[53] Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, **26**, 2460-2461 (2010).

[54] McDonald, D. *et al*. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *The ISME journal*, **6(3)**, 610-618 (2012).

[55] Paulson, J. N., Stine, O. C., Bravo, H. C. & Pop, M. Differential abundance analysis for microbial marker-gene surveys. *Nature methods*, **10(12)**, 1200 (2013).

[56] Henderson, C. R. Sire evaluation and genetic trends. *Journal of Animal Science*; Symposium: 0-41 (1973).

[57] Difford, G. F. *et al*. Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows. *PLoS genetics*, **14(10)**, e1007580 (2018).

[58] Bray, J. R. & Curtis, J. T. An ordination of upland forest communities of southern Wisconsin. *Ecological monographs*, **27**, 325-349 (1957).

[59] Lozupone, C. & Knight, R. UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and environmental microbiology*, **71**, 8228–8235 (2005).

[60] McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one*, **8(4)**, e61217 (2013).

[61] Manual for BLUPF90 family of programs [Internet] Athens, GA, USA: University of Georgia; 2015. Available from: http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all2.pdf.

[62] R Development Core Team. R: A Language and Environment for Statistical Computing. (2010) [http://cran.r-project.org].

[63] Bonferroni, C. Teoria statistica delle classi e calcolo delle probabilita. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze, **8**, 3-62 (1936).

[64] Chung, D., Chun, H. & Todorov, M. V. Package 'spls' (2019).

[65] Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology*, **73(16)**, 5261-5267 (2007).

Acknowledgements

We would like to thank Oscar Perucho, Josep Ramon and Carmen Requena (staff of Unitat de Cunicultura, IRTA) for their contribution to data recording and animal care during the experiment. We also want to thank Oriol Rafel, Marc Viñas, Miriam Guivernau and Olga González for their help collecting and processing the biological samples. We acknowledge

Armand Sánchez, Nicolas Boulanger and Joana Ribes (Genomics and NGS Unit, CRAG) for their assistance in massive libraries preparation.

Author contributions

JS and MP conceived the experimental design. JS, MP and MVG collected biological samples. MVG and MP processed the samples in the laboratory. MVG processed and analyzed the sequencing data, interpreted data, prepared figures and tables, and wrote the manuscript. JS and YRC helped analyzing the sequencing data. JS, MP and YRC helped interpreting the data, and revised the manuscript. All authors read and approved the final manuscript.

Additional information

Additional file 1: metadata.txt. Metadata associated with the 425 rabbit cecal samples analyzed in this study.

Additional file 2: otu_table.txt. Filtered and CSS-normalized OTU table used for statistical analyses in this study.

Additional file 3: OTUs_tax_assignments.txt. Taxonomic assignments for all OTUs in Additional file 2.

Additional file 4: Trace plots and histograms of Markov chains from the posterior distribution of the parameters of Bayesian models.

Additional file 5: Rep_seqs_sPLSR_selected_OTUs.fna. Representative sequences of the OTUs selected in the sPLSR analysis for $\overline{\mathrm{ADRFI}}_{\mathrm{AL}}$.

Additional file 6: Table S1. Relevant OTUs for the prediction of individual traits (ADG_{AL} and ADG_R) and cage-average traits ($\overline{ADFI}_{AL}$, $\overline{ADRFI}_{AL}$ and $\overline{ADFCRI}_{AL}$).

Additional file 7: Relevant_OTUs.fna. Representative sequences of the OTUs relevant OTUs for the prediction of individual traits (ADG_{AL} and ADG_R) and cage-average traits ($\overline{ADFI}_{AL}$, $\overline{ADRFI}_{AL}$ and $\overline{ADFCRI}_{AL}$) in Additional file 6.

List of abbreviations

$\overline{ADFCR}_{AL}$: average daily feed conversion ratio on *ad libitum* feeding regime

$\overline{ADFI}_{AL}$: average daily feed intake on *ad libitum* feeding regime

ADG: average daily gain

ADG_{AL}: average daily gain on *ad libitum* feeding regime

$\overline{ADG}_{AL}$: cage-average daily gain on *ad libitum* feeding regime

ADG_R: average daily gain on restricted feeding regime

$\overline{ADRFI}_{AL}$: average daily residual feed intake on *ad libitum* feeding regime

AL: *ad libitum* feeding regime

BW: body weight

CSS: cumulative sum scaling

FE: feed efficiency

FI: feed intake

fullDataset: dataset including records of animals in which microbiota was assessed as well as of their cage mates

mDataset: dataset including only records of animals in which microbiota was assessed

MSPE: mean squared prediction error

$\overline{\mathbf{MW}}_{\mathbf{AL}}$: cage-average mid growing period day metabolic weight ($BW^{0.75}$)

OTU: operational taxonomic unit

PCR: polymerase chain reaction

R: restricted feeding regime

EM-REML: expectation–maximization residual maximum likelihood

sPLSR: sparse partial least squares regression

Declarations

Ethics approval and consent to participate

This study was carried out in compliance with the ARRIVE guidelines. This study was carried out in accordance with the relevant guidelines and regulations of the animal care and use committee of the Institute for Food and Agriculture Research and Technology (IRTA). The protocol was approved by the committee of the Institute for Food and Agriculture Research and Technology (IRTA).

Availability of data and materials

The raw sequence data were deposited in the sequence read archive of NCBI under the accession number SRP186982 (BioProject PRJNA524130). Metadata, the filtered and CSS-normalized OTU table and corresponding taxonomic assignments have all been included as Additional files 1, 2 and 3, respectively.

Competing interests

The authors declare that they have no competing interests.

Funding

The experimental design of this work was conducted thanks to funding from INIA project RTA2011-00064-00-00. This study was part of the Feed-a-Gene project that received funding from the European Union's H2020 program under grant agreement no. 633531, and the Spanish project RTI2018-097610R-I00. MVG is a recipient of a "Formación de Personal Investigador (FPI)" pre-doctoral fellowship from INIA, associated with the research project RTA2014-00015-C2-01. YRC is recipient of a Ramon y Cajal post-doctoral fellowship (RYC2019-027244-I) from the Spanish Ministry of Science and Innovation.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- metadata.txt.txt
- otutable.txt.txt
- OTUstaxassignments.txt.txt
- Additionalfile4.rar
- RepseqssPLSRselectedOTUs.fna.txt
- TableS1.docx
- RelevantOTUs.fna.txt