

CORRELATED RESPONSE TO SELECTION FOR INTRAMUSCULAR FAT ON THE GUT METAGENOMIC PROFILE

Zubiri-Gaitán A., Martínez-Álvaro, M., Hernández P., Blasco A*.

Institute for Animal Science and Technology, Universitat Politècnica de València, 46022, València, Spain

*Corresponding author: ablasco@dca.upv.es

ABSTRACT

A divergent selection experiment for intramuscular fat content (IMF) in *Longissimus Thoracis et Lumborum* muscle was performed in rabbits during 10 generations at the Universitat Politècnica de València, to study the mechanisms involved in the intramuscular fat deposition. The scope of this experiment was to analyze the correlated response to selection on the gut metagenome, in order to try to elucidate the role of the microbiota and its genes on the mentioned mechanisms.

A total of 89 rabbits from the 10th generation of selection were used to estimate the correlated response to selection on the gut metagenome. Once the metagenome dataset was obtained, the data was transformed using compositional data analysis techniques in order to deal with its compositional nature. Projection to latent structures (PLS) and discriminant PLS (PLS-DA) analyses were used to find the microbial genes affected by selection. The most important variables for both models were those that had a variable importance in projection (VIP) ≥ 0.80 and a confidence interval of the Jack Knife regression coefficient did not include the zero. The PLS-DA model included 240 microbial genes and was able to correctly classify all the samples after a cross-validation procedure. The PLS regression model, on the other hand, included 230 microbial genes with a linear prediction ability of IMF content of 79% after cross-validation (Q^2). Only 122 microbial genes that overlapped between the results of PLS-DA (240) and PLS (230) were considered, helping to differentiate the ones actually related with IMF from those that were fixed due to genetic drift. The sign and magnitude of the correlated response to selection on each microbial gene were estimated as the difference of the relative abundance between the lines, in a linear model that included the line effect. The model was solved using Bayesian inference, and assuming flat priors for all unknowns. The marginal posterior distributions of the mentioned differences were described by the median, highest posterior density interval at 95% probability, and the probability of these differences of being higher or lower than 0. Finally, the metabolic routes affected by those genes were identified thanks to the information provided by the KEGG database. The majority of them were involved in cell wall membrane synthesis, energy production and conversion, and transport and metabolism of lipids, amino acids and coenzymes.

Divergent selection for IMF led to a modification of the gut metagenome, confirming the existence of a link between the host genome and its metagenome. However, a more exhaustive analysis of the genes and the metabolic routes, together with the study of the microbial taxa involved, is necessary to fully understand the role of the microbiome in the intramuscular fat deposition.

Key words: divergent selection, intramuscular fat, metagenome, compositional data, rabbit

INTRODUCTION

Intramuscular fat content (IMF) is one of the main parameters in meat quality, since it affects its juiciness, tenderness and flavor. The genetic and environmental factors involved in intramuscular and carcass fat deposition have been extensively studied and, in the last few years, the role of the gut metagenome has gained importance. Numerous studies suggest that the gut metagenome plays an important role in the host metabolism affecting, among others, the energy harvest and the fat deposition (Krajmalnik-Brown et al., 2012). There is also evidence supporting that the microbiota is determined by the host genome (Bonder et al., 2016). However, the specific role of the gut metagenome in the mechanisms of fat deposition has not been fully elucidated yet.

A divergent selection experiment for IMF content in *Longissimus Thoracis et Lumborum* (LTL) muscle was performed in rabbits during 10 generations at the Universitat Politècnica de València, in order to study the mechanisms of intramuscular fat deposition and its correlation with other relevant traits. After selection, the divergent lines differ in the trait under selection and the correlated ones, and these differences can be directly attributed to the genetic composition of the lines. A correlated response to selection in the gut metagenome would show the existence of a link between those and the host genome.

The scope of this experiment was to analyze the correlated response to selection on the gut metagenome, in order to elucidate its role on the intramuscular fat deposition mechanisms.

MATERIALS AND METHODS

A divergent selection experiment for IMF content was performed in rabbits during 10 generations at the Universitat Politècnica de València. The selection criterion was the average phenotypic value of IMF measured in two full sibs of the candidate (one male and one female) at 9 weeks of age in LTL muscle. The complete selection procedure is described in Martínez-Álvaro et al. (2016). A total of 89 animals from the 10th generation of selection were used for this experiment, 47 from the high-IMF line and 42 from the low-IMF line. Thirty-five of these animals (20 from the high-IMF line and 15 from the low-IMF line) were involved in a related project, and were born after an embryo transfer procedure; the line of the dams that received the embryos was the same as the line of the 35 embryos itself. Animals were slaughtered at 9 weeks of age after 4 hours of fastening period by exsanguination prior electric stunning. Cecum samples were collected immediately after slaughter. The intestinal tract was removed from the abdominal cavity; cecum content was collected in 50 mL sterile Falcon tubes, homogenized and aliquoted in 2 mL cryogenic tubes. The aliquots were immediately submerged in liquid nitrogen and stored at -80°C until the DNA extraction. Bacterial genomic DNA was isolated from frozen cecal samples using the DNeasy PowerSoil kit (QIAGEN Inc., Hilden, Germany).

Samples were sequenced in two different facilities. Thirty-three samples (16 from the high and 17 from low-IMF line) were sequenced at the Fundación para el Fomento de la Investigación Sanitaria y Biomédica (FISABIO, Valencia, Spain). The library was prepared using the Nextera XT DNA library preparation kit (Illumina, Inc., San Diego, CA, USA) according to the manufacturer protocol. Sequencing was carried out on a NextSeq 500 sequencer (Illumina) with 150 bp paired-end chemistry. The remaining 56 samples (31 from the high and 25 from low-IMF line) were sequenced at Sistemas Genómicos (Valencia, Spain). The library was prepared using SureSelect^{QXT} library preparation kit (Agilent Technologies, Inc., Santa Clara, CA, USA) following protocol and recommendations in “SureSelect^{QXT} Whole Genome Library Prep for Illumina Multiplexed Sequencing Featuring Transposase-Based Library Prep Technology”. Sequencing was performed on NextSeq 550 sequencer (Illumina) with 150 bp paired-end chemistry. The data obtained from both sequencing procedures were processed using the SqueezeMeta fully automatic metagenomics analysis pipeline (Tamames and Puente-Sánchez, 2019), using the coassembly mode with default parameters. The microbial genes were identified by homology of the contigs to the Kyoto Encyclopedia of Genes and Genomes (KEGG). Finally, the metabolic routes affected by those genes were identified thanks to the information provided by the KEGG database.

Due to the compositional nature of metagenomics data (Gloor et al., 2017), it was analyzed using compositional data analysis techniques. The data were transformed into an additive log ratio (ALR) (Greenacre, 2018), selecting a reference microbial gene that met the following criteria: being (1) present in all animals, (2) highly abundant, (3) highly correlated with the total number of counts, and (4) not correlated with intramuscular fat. In addition, the batch effect caused by analyzing samples in different laboratories was corrected by fitting a linear model with microbial gene as dependent variable and sequencer effect as fixed effect. The microbial genes showing a correlated response to selection for intramuscular fat content were identified using Projection to Latent Structures (PLS) and discriminant PLS (PLS-DA), computed by SIMCA, P+ 15.0.1, Umetrics (Umea, Sweden). The microbial genes were set as the independent variables; for the PLS regression, IMF content was adjusted as dependent variable, while for the PLS-DA the dependent variable was a categorical vector for high or low-IMF line. The most important variables for both models were those that had a variable importance in projection (VIP) ≥ 0.80 and a confidence interval of the Jack Knife regression coefficient did not include the 0. The models were re-estimated and the variable selection steps were repeated until the best adjustment was achieved, measured

with the Q^2 parameter obtained after the cross-validation procedure. The cross-validation (CV) procedure was performed by dividing the dataset into 7 groups, leaving one group out at each CV round as a test set, and predicting the intramuscular fat content of each sample (PLS), and classifying it as high/low (PLS-DA). A misclassification table was elaborated based on the CV results of the PLS-DA. The microbial genes that were considered as affected by the selection process were only those selected in both PLS-DA and PLS approaches. The sign and magnitude of the correlated response to selection on each microbial gene were estimated as the difference of the relative abundance between the lines, in a linear model that included the line effect. The model was solved using Bayesian inference, and assuming flat priors for all unknowns. The marginal posterior distributions of the mentioned differences were described by the median, highest posterior density interval at 95% probability, and the probability of these differences of being higher or lower than 0 (P_0). The Bayesian analysis was performed with the program Rabbit, developed by the Institute for Animal Science and Technology (Valencia, Spain). Finally, the metabolic routes affected by those genes were identified thanks to the information provided by the KEGG database.

RESULTS AND DISCUSSION

A total of 4726 microbial genes were identified, from which 3937 were present in all samples and were kept for further analysis; the remaining 789 were discarded. The 3937 microbial genes represented 99.66% of the total number of counts. The reference microbial gene selected for the ALR was the RP-S1, involved in ribosomal synthesis. An exploratory analysis of the 89 samples was performed based on the PLS and PLS-DA models, and two samples were discarded based on the distances of the individuals to the model or DMOD criterion.

After the variable selection procedure was completed, the PLS-DA model with the best adjustment included 240 microbial genes and was able to correctly classify all samples after the cross-validation procedure, as shown in table 1. The best PLS regression model, on the other hand, included 230 microbial genes with a linear prediction ability of IMF content of 79% after cross-validation (Q^2). Figure 1 shows the projection of the 87 samples in the space delimited by the first and second latent variables of the PLS-DA model (fig. 1A) and the PLS regression model (fig. 1B). As can be seen in the figures, the PLS-DA model showed a great discrimination between the lines, and the PLS regression also denoted a clear linear relationship between the IMF content and the microbial genes.

Having two contemporary divergently selected lines, raised under identical environmental conditions, allows to confirm that the differences found between the lines are due to their genetic composition. However, it is also important to differentiate the microbial genes actually related with the IMF content from those that were selected during the selection process, due to genetic drift. To do so, a novel approach was used in which only the microbial genes discriminating between lines (identified by PLS-DA) that were also linearly related to the IMF content (identified by PLS regression) were considered. From the 240 and 230 microbial genes identified by PLS-DA and PLS, respectively, 122 overlapped and were considered as a correlated response to selection for IMF content. The sign and magnitude of the correlated response of the former 122 microbial genes were evaluated by fitting a linear model, and the difference between lines was evaluated for each one of them. A total of 77 microbial genes showed high evidence of non-zero differences ($P_0 > 95\%$) between lines and the metabolic routes affected by those genes were identified. Several microbial metabolic routes were found to be genetically associated with the intramuscular fat deposition in rabbits.

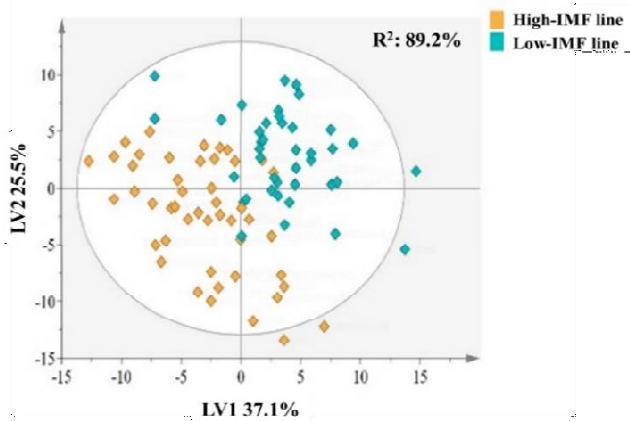
The majority of microbial genes affected by selection, and the metabolic routes they are involved in, act on cell wall membrane synthesis, energy production and conversion, lipid transport and metabolism, amino acid transport and metabolism, and coenzyme transport and metabolism. For instance, 13 genes were involved in the energy production and conversion. Most of them were involved in the oxidative phosphorylation at mitochondria membrane level and in the citrate cycle, and all of them were more abundant in the high-IMF line. The former could evidence a higher amount of energy obtained from the gut microbial metabolism of the rabbits from the mentioned line. A similar scenario can be seen with the genes involved in the lipid transport and metabolism, where 6 out of the 8 differential genes were more abundant in the high-IMF line, a very interesting result and a good starting point for further analyses. The remaining functions did not show a clear tendency towards a line. Overall, a more thorough analysis of the microbial genes, and the functions they are involved in, is still needed.

From these results, we can conclude that a correlated response to selection to IMF content was observed in the gut metagenome, indicating the existence of a link between the host genome and its metagenome. Several metabolic routes were modified after selection. However, a more exhaustive analysis of the genes and the metabolic routes modified by selection, together with the study of the microbial taxa involved, is necessary to understand the mechanisms employed by the microbiome that led to a modification in the intramuscular fat deposition of these lines.

Table 1. Misclassification table obtained after the cross-validation procedure performed in the PLS-DA model

	High-IMF line	Low-IMF line
High-IMF line	47	0
Low-IMF line	0	40
Correctly classified samples (%)	100	100

A.



B.

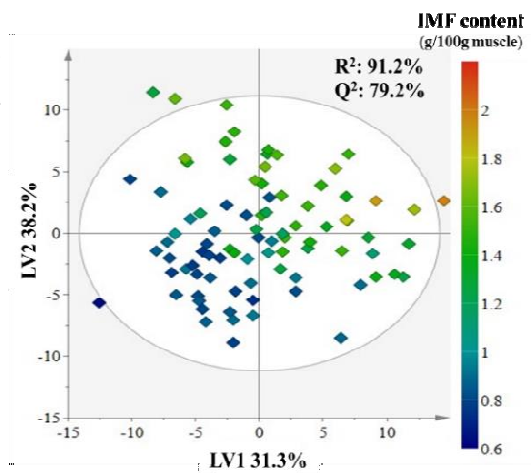


Figure 1. Projection of the 87 samples into the space delimited by the first (LV1) and second (LV2) latent variables of the models (score plot). The ellipse represents the Hotelling T^2 with 95% confidence interval. Each dot represents a sample. **A.** Score plot of the PLS-DA model built with 240 microbial genes. **B.** Score plot of the PLS model built with 230 microbial genes. The model explained 79% of IMF variability after cross-validation (Q^2).

REFERENCES

- Bonder, M. J., A. Kurilshikov, E. F. Tigchelaar, Z. Mujagic, F. Imhann, A. V. Vila, P. Deelen, T. Vatanen, M. Schirmer, S. P. Smekens, D. V Zhernakova, S. A. Jankipersadsing, M. Jaeger, M. Oosting, M. C. Cenit, A. A. M. Masclee, M. A. Swertz, Y. Li, V. Kumar, L. Joosten, H. Harmsen, R. K. Weersma, L. Franke, M. H. Hofker, R. J. Xavier, D. Jonkers, M. G. Netea, C. Wijmenga, J. Fu, and A. Zhernakova. 2016. The effect of host genetics on the gut microbiome. *Nat. Genet.* 48:1407–1412. doi:10.1038/ng.3663.
- Gloor, G. B., J. M. Macklaim, V. Pawlowsky-Glahn, and J. J. Egozcue. 2017. Microbiome datasets are compositional: And this is not optional. *Front. Microbiol.* 8:1–6. doi:10.3389/fmicb.2017.02224.
- Greenacre, M. J. 2018. *Compositional data analysis in practice*. Chapman and Hall/CRC. Available from: <http://public.eblib.com/choice/publicfullrecord.aspx?p=5455461>
- Krajmalnik-Brown, R., Z. E. Ilhan, D. W. Kang, and J. K. DiBaise. 2012. Effects of gut microbes on nutrient absorption and energy regulation. *Nutr. Clin. Pract.* 27:201–214. doi:10.1177/0884533611436116.
- Martínez-Álvaro, M., P. Hernández, and A. Blasco. 2016. Divergent selection on intramuscular fat in rabbits: Responses to selection and genetic parameters. *J. Anim. Sci.* 94:4993–5003. doi:10.2527/jas.2016-0590.
- Tamames, J., and F. Puente-Sánchez. 2019. SqueezeMeta, a highly portable, fully automatic metagenomic analysis pipeline. *Front. Microbiol.* 10:1–10. doi:10.3389/fmicb.2018.03349.