Taxonomic and functional analysis of metagenome data from bovine rumen samples and its association with feed efficiency

Delgado[1,2], B., Bach[3], A., Guasch[4], I., Elcoso[4], G., González[1], C., Gonzalez-Recio[1,5], O.
[1]Ctra. La Coruña, km. 7.5, 28040, Madrid, Spain (INIA).
[1]Departamento de Mejora Genética Animal. Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria O.A., M.P. 28040 Madrid, Spain. [2]Ctra. Colmenar, km. 15, 28049, Madrid, Spain (UAM). [3]Department of Ruminant Production, ICREA-IRTA, 08140 Caldes de Montbui, Spain. [4]Blanca from the Pyreness, 25795 Hostalets the Tost, Spain. [5]Departamento de Producción Agraria. Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas. UPM. Ciudad Universitaria s/n 28040 Madrid, Spain.
E-mail: beatriz.delgado01@estudiante.uam.es

INTRODUCTION

The rumen is a key organ in cattle because of its role in fiber decomposition and production of different substances that can be used by the animal for its nutrition. This is made thanks to resident microorganisms belonging to Archaea, Bacteria and Eukarya kingdoms, which allow ruminants to use the lignocellulose and other molecules from plants as their main source of energy (Creevey *et al*., 2014). Microorganisms depend on each other to develop its functions as some of them produce molecules that serve as a substrate for the others (Wallace *et al*., 2017) - e.g. protozoa produce H_2 that methanogens use to synthesize CH_4. Microbial interactions among rumen microorganisms are responsible of the nutrition and well-being of cattle and have an enormous impact in the environment as they are one of the main sources of CH_4.
In addition to its role in food digestion, the rumen microbiome has been studied for years in an attempt to discover its association with health and production in cattle. Recent studies show correlation between the host genetics and the microbiome composition (Gonzalez-Recio *et al*. 2018), which could be used in breeding programs by selecting animals with a favorable microbiome.
Metagenomics are one of the most powerful tools to unravel new functions and microorganisms in diverse ecosystems. Although rumen microbiome databases are relatively weak (Tapio *et al*., 2017) it is interesting to use this approach to discover which microorganisms and functions are enhanced or present in animals with high feed efficiency.
The aim of this study is to describe bovine rumen microbiome, its composition and functions, as well as to analyze its relationship with feed efficiency

METHODS

Data. Rumen samples from 30 animals selected among 96 for extreme phenotype of feed efficiency (High or Low) were first grinded until solid and liquid phases were homogenized using a blender. DNA was extracted using "Power Soil DNA Isolation Kit" (Mo Bio Laboratories Inc., Carlsbad, CA, USA) and "DNeasy Power Soil Kit" (QIAGEN, Valencia, CA, USA). All samples were sequenced using Illumina MiSeq in an external sequencing service (FISABIO, Valencia).
Data analysis. We used MEGAN Community Edition (CE) (Huson *et al*., 2016) software for the taxonomic binning and functional analysis, avoiding a previous assembly step. For that, we aligned the sequences contained in the fastq.gz files against the NCBI-nr protein database using DIAMOND (Buchfink *et al*., 2015), a command-line software which generates DAA (Diamond Alignment Archive) files. The diamond files were later "meganized" by using the *daa2rma* tool of MEGAN, obtaining an RMA file for each animal. To compare all 30 samples, we introduced the RMA files in a comparison document and we later grouped the samples by high and low efficiency.
The taxonomic assignment was made by MEGAN using the naïve LCA algorithm, which assigns each read to the lowest common ancestor. For the functional analysis we performed a search against Gene Ontology and mapped the reads to InterPro families, resulting in an InterPro2GO classification with three different nodes corresponding to cellular component, molecular function and biological processes.

Statistical method. Functional analysis was performed using the GLM package from R using the following model:

$$\mathbf{y} = \mu + \mathbf{P} + \mathbf{fe} + \mathbf{e}$$

Where **y** = relative gene abundance transformed to $\log(1 + x)$; **P** = Parity (primiparous vs. multiparous); **fe** = feed efficiency (classes: High and Low); **e** = error.
P values for the feed efficiency effect were adjusted for multiple comparison using the Benjamini & Hochberg (1995) correction.

RESULTS AND DISCUSSION

Taxonomic classification resulted in a taxonomic tree containing bacteria, archaea, eukaryotic microorganisms and viruses present in the 30 different rumen samples. Then, the samples were classified in MEGAN by feed efficiency in two groups (15 animals per group), high and low, calculated as milk production (kg/d) divided by feed consumption (kg/d).
Comparison of the two groups showed differences between abundances of bacteria, archaea and eukaryotic microorganisms. The phylum Bacteroidetes was more abundant in efficient cows (Figure 1) and Firmicutes was less abundant, whilst Archaea and Eukarya kingdoms were more abundant and diverse in less efficient cows. More efficient individuals presented a larger relative abundance of *Bacteroidetes* (*P*=0.041), and a lower, but not significant, relative abundance of *Firmicutes* (P=0.119). The most abundant genus in the Bacteroidetes group was *Prevotella*, which was also more abundant in the cows within the high efficiency group (P=0.003). Cows within the group of less efficient individuals presented microbiota with a larger abundance of Methanobacteria (P=0.004) and *Methanobrevibacter* (P=0.003) (Figure 1).
Methanobrevibacter species such as *M. ruminantium* and *M. olleyae* were not present in efficient cows, whose archaea population was composed almost exclusively by *M. millerae*, being this species the most abundant in less efficient cows. Moreover, in efficient cows we observed a lower number of reads associated to the eukaryotic phylum Ciliophora as well as an overall decrease in the other phyla in this kingdom. Microorganisms among phylum Ciliophora include ciliates from genera such as *Entodinium* or *Isotricha*, whose association with rumen methanogens have been known for decades (Chagan et al., 1999). Different studies show that *Methanobrevibacter* spp. and other genera of methanogens live inside ciliates and on their membrane surface, improving the utilization of H_2, a fermentation product of ciliates, to produce CH_4 by methanogenic archaea (Morgavi et al., 2010; Belanche et al., 2014). A significant reduction in the number of ciliates might be accompanied by a decrease in methanogens due to the close relationship between them, meaning that the H_2 produced by protozoa can be used to produce volatile fatty acids which are energy sources for the animal (Knapp et al., 2014). The overall result is an increase in feed efficiency and production, as food is not wasted in producing methane but used by the animal to nourish.
Functional annotation and analysis of the metagenomic data showed that functions and pathways such as hydrolase activity (GO:0016787), biosynthetic processes (GO:0009058), carbohydrate metabolism (GO:0005975) and amino acid metabolic processes (GO:0006520) are increased in the rumen microbiome when compared to other pathways (Table 1). These functions are related to food processing and are essential for the animal by using all food components that microorganisms can decompose and transform into molecules of high nutritional value. Other common pathways such as nucleotide binding (GO:0000166), transport (GO:0006810) and DNA metabolism (GO:0006259) were also abundant in the samples. In addition to Gene Ontology pathways, more than 4000 proteins were detected in the samples, being glycoside hydrolases (IPR026892, IPR000322) the most abundant, as well as other enzymes related to carbohydrate metabolism such as pyruvate oxidoreductases and phosphate dikinases (IPR011895, IPR010121). Other important proteins in nutrition had a higher number of reads, such as carbamoyl-phosphate synthase (IPR006275), which intervene in the metabolism of N_2; and different peptidases take part in protein decomposition.
High and low efficiency groups showed differences in homeostatic and cofactor metabolic processes as well as in cellular respiration pathways and pilus components but they were not significant after the BH correction (P>0.05) (Table 2). Genes related to carbohydrate metabolism and amino acid transport also had differences between the groups, although they

were not significant after the correction either. More rumen microbiome samples must be analyzed in order to establish relationships between metabolic functions and enzyme activity with feed efficiency in cows.

BIBLIOGRAPHY

Belanche, A., de la Fuente, G., & Newbold, C. J. (2014). Study of methanogen communities associated with different rumen protozoal populations. *FEMS microbiology ecology*, *90*(3), 663-677.

Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature methods*, *12*(1), 59.

Chagan, I., Tokura, M., Jouany, J. P., & Ushida, K. (1999). Detection of methanogenic archaea associated with rumen ciliate protozoa. *The Journal of general and applied microbiology*, *45*(6), 305-308.

Creevey, C. J., Kelly, W. J., Henderson, G., & Leahy, S. C. (2014). Determining the culturability of the rumen bacterial microbiome. *Microbial biotechnology*, *7*(5), 467-479.

Gonzalez-Recio, O., Zubiria, I., García-Rodríguez, A., Hurtado, A., & Atxaerandio, R. (2018). Signs of host genetic regulation in the microbiome composition in 2 dairy breeds: Holstein and Brown Swiss. *Journal of dairy science*, *101*(3), 2285-2292.

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., ... & Tappu, R. (2016). MEGAN community edition-interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS computational biology*, *12*(6), e1004957.

Knapp, J. R., Laur, G. L., Vadas, P. A., Weiss, W. P., & Tricarico, J. M. (2014). Invited review: Enteric methane in dairy cattle production: Quantifying the opportunities and impact of reducing emissions. *Journal of Dairy Science*, *97*(6), 3231-3261.

Morgavi, D. P., Forano, E., Martin, C., & Newbold, C. J. (2010). Microbial ecosystem and methanogenesis in ruminants. *Animal*, *4*(7), 1024-1036.

Tapio, I., Snelling, T. J., Strozzi, F., & Wallace, R. J. (2017). The ruminal microbiome associated with methane emissions from ruminant livestock. *Journal of animal science and biotechnology*, *8*(1), 7.

Wallace, R. J., Snelling, T. J., McCartney, C. A., Tapio, I., & Strozzi, F. (2017). Application of meta-omics techniques to understand greenhouse gas emissions originating from ruminal metabolism. *Genetics Selection Evolution*, *49*(1), 9.

Table 1. *Relative abundance (%) of the 15 most abundant metabolic pathways and functions from Gene Ontology in the 30 rumen samples.*

	% Relative abundance
Unclassified molecular processes	8.418
GO:0016787 hydrolase activity	7.658
GO:0009058 biosynthetic process	5.917
GO:0000166 nucleotide binding	5.798
GO:0016740 transferase activity	5.636
Unclassified biological processes	4.330
GO:0003676 nucleic acid binding	4.119
GO:0006810 transport	3.974
Unclassified cellular processes	3.601
GO:0006259 DNA metabolic process	3.419
GO:0016020 membrane	3.110
GO:0005975 carbohydrate metabolic process	2.992
GO:0006520 cellular amino acid process	2.792
GO:0006807 nitrogen compound process	2.720

Table 2. *Metabolic pathways and functions from Gene Ontology analyzed with GLM with the lowest p-values (<0.05) and their corrected p-values.*

	p-value (<0.05)	Corrected p-value
GO:0042592 homeostatic process	0.0147	0.8026
GO:0051186 cofactor metabolic process	0.0291	0.8026
GO:0045333 cellular respiration	0.0418	0.8026
GO:0009289 pilus	0.0433	0.8026

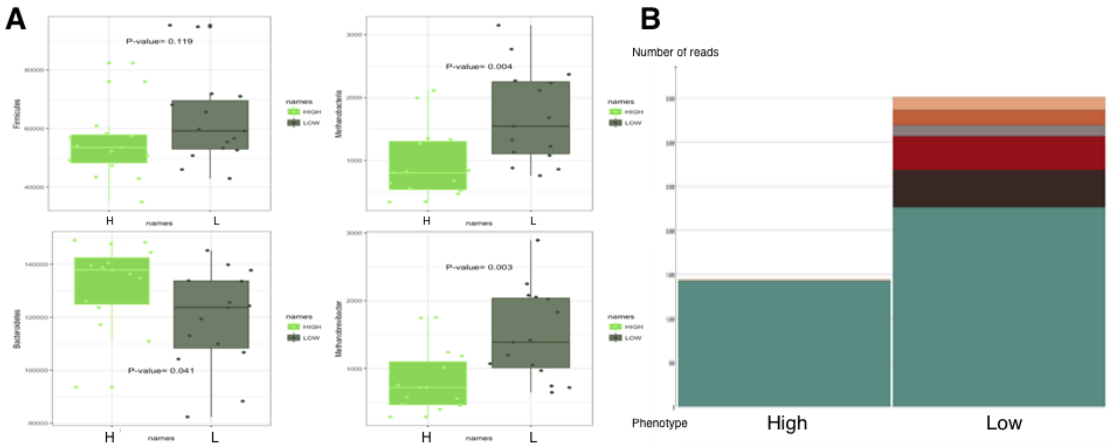

Figure 1. *A) Relative abundance of* Firmicutes*, Methanobacteria, Bacteroidetes and Methanobrevibacter in high (green) and low (dark green) groups. B) Differences in the abundance of Archaea between high (left) and low (right) feed efficiency groups.*

Taxonomic and functional analysis of metagenome data from bovine rumen samples and its association with feed efficiency

ABSTRACT: Rumen microbiota composition is a key factor for healthy and productive livestock and it has an essential role in climate change due to the large amount of methane produced by cattle worldwide. Its relationship with production deserves further study, for example which microbial groups are associated with a higher feed efficiency and how host genetics regulate its composition. Metagenomics is one of the most powerful tools to reveal microorganisms present in the rumen microbiome and their functions. Analysis of metagenomic data from 30 different rumen samples showed that some taxonomic groups, such as Bacteroidetes and *Methanobrevibacter* were significantly different in high and low feed efficiency cows. Bacteroidetes group was more abundant in the group with a higher feed efficiency and *Methanobrevibacter* in low feed efficiency cows, which could mean that lower amounts of methanogenic archaea and thus methane production are related with a better use of food to increase livestock production. Moreover, a functional analysis revealed that the principal functions of the microbiome are those related to carbohydrate and amino acid metabolism, which are essential for the animal to nourish. There were differences in metabolic pathways and genes between high and low feed efficiency groups, although they were not significant after multiple testing correction, and it is interesting to analyze more rumen microbiome data in the future to establish relationships between feed efficiency and enhanced metabolic pathways.

Keywords*:* metagenomics, microbiome, livestock, feed efficiency