RUNNING HEADER: HOST GENETIC COMPONENT OF THE MICROBIOME

SHORT REPORT: Signs of host genetic regulation in the microbiome composition in cattle

O. Gonzalez-Recio[*], I. Zubiria[†], A. García-Rodríguez[†], A. Hurtado[††], R. Atxaerandio[†]

[*] Departamento de Mejora Genética Animal. Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria. 28040 Madrid, Spain

[†] Departamento de Producción Animal. NEIKER-Tecnalia. Granja Modelo de Arkaute Apartado 46, 01080 Vitoria-Gasteiz, Spain

[††] Departamento de Sanidad Animal. NEIKER-Tecnalia. Berreaga 1, 48160 Derio, Spain

Corresponding author: Oscar González-Recio. E-mail: gonzalez.oscar@inia.es

ABSTRACT

Previous studies have revealed certain genetic control by the host over the microbiome composition, although in many species the host genetic link controlling microbial composition is yet unknown. This potential association is important in livestock to study all factors and interactions that rule the effect of the microbiome in complex traits. This report aims to study whether the host genotype exerts any genetic control on the microbiome composition of the rumen in cattle. Data on 16S and 18S rRNA gene-based analysis of the rumen microbiome in 18 dairy cows from two different breeds (Holstein and Brown Swiss) were used. The effect of the genetic background of the animal (through the breed and Single Nucleotide Polymorphisms; SNP) on the relative abundance (RA) of archaea, bacteria and ciliates (with average relative abundance per breed >0.1%) was analysed using Bayesian statistics. In total, 13 genera were analysed for bacteria (5), archaea (1), and ciliates (7). All these bacteria and archaea genera showed association to the host genetic background both for breed and SNP markers, except RA for the genera *Butyrivibrio* and *Ruminococcus* that showed association with the SNP markers but not with the breed composition. Relative abundance of 57% (4/7) of ciliate analysed showed to be associated to the genetic background of the host. This host genetic link was observed in some genus of *Trichostomatia* family. For instance, the breed had a significant effect on *Isotricha*, *Ophryoscolex* and *Polyplastron*, and the SNP markers on *Entodinium*, *Ophryoscolex* and *Polyplastron*. In total, 77% (10/13) of microbes analysed showed to be associated to the host genetic background (either by breed or SNP genotypes). Further, the results showed a significant association between DGAT1, ACSF3, AGPAT3 and STC2 genes with the relative abundance *Prevotella* genus with a false discovery rate lower than 15%. The results in this study support the

41 hypothesis and provide some evidence that there exist a host genetic component in cattle that
42 can partially regulate the composition of the microbiome.
43
44
45 **Keywords**: genomic, breed, SNP, Holstein, Brown Swiss, microbiome, NGS
46 **Abbreviations**: FDR: False Discovery Rate; NGS: Next Generation Sequencing; MAF:
47 Minor Allele Frequency; OTU: operational taxonomic units; PC: Principal components; RA:
48 Relative abundance; SNP: Single Nucleotide Polymorphisms.
49
50
51

The research interest on the microbiome and its effects on the host, both in humans [1,2] and livestock [3,4], is raising in the last years. The microbiome plays an important role in the phenotypic expression of many phenotypes such as feed efficiency, disease status, or methane emission. Traditionally, microbes have been studied in the lab, without considering their effect on complex features and their interaction with the host. In the particular case of livestock, the traits of interest are usually related to productive, health or environmental factors. In the last decade, more attention has been focused on the interactions between microbes and diet [5–8], methane emissions [9–13] and the microbiome compositions across hosts, environment and age [4,7,14]. It has also been proposed as a predictor of complex traits [13,15].

Therefore, there is an increasing interest on determining whether a host genetic control on the microbiome composition exists. Recent studies show some evidences that support the hypothesis that there is some sort of host control over the composition of the microbiome in mammals. For instance, Weimer et al. (2010) reported that after a near-total exchange of ruminal contents, the ruminal bacterial composition returned to a similar status to that prior the exchange. More recently, [17] showed differences between sire progeny groups on the archaea:bacteria ratio in Aberdeen Angus and Limousin cattle breeds, and [18] reported heritabilities above 0.20 for the relative abundance of several microbes in a twin human study.

It is of interest to provide more evidences on the host genetic control of the microbiome composition because some selection intensity could be applied to select individuals with a favourable microbiome for a given breeding goal, as the reduction of methane yield or the improvement of the feed efficiency, for example.

This trial was carried out in accordance with Spanish Royal Decree 53/2013 for the protection of animals used for experimental and other scientific purposes. In this study,

77 ruminal content was sampled from 18 dairy cows (10 Holstein and 8 Brown Swiss) from
78 Fraisoro Farm School (Zizurkil, Gipuzkoa, Spain). These cows were undergoing a nutrition
79 experiment. They were randomly assigned to one of two experimental concentrate
80 supplements. Concentrates were formulated to contain cold-pressed rapeseed cake or palm as
81 fat sources, and to provide equal amounts of crude protein, energy and fat. Both breeds were
82 fed both diets. The effect of the treatment was adjusted as a 2-levels factor in the statistical
83 analyses, but results are not reported here as this is not the objective of this study.
84 Rumen samples were taken 4 times over two consecutive days. Sampling began at 00:00 and
85 12:00 h on d 1, and 06:00 and 18:00 h on d 2; each sampling taking approximately 2 h.
86 Ruminal samples were collected from each dairy cow using a stomach tube connected to a
87 mechanical pumping unit. About 100 ml of each ruminal extraction were placed into a
88 container and were frozen immediately after the extraction and then stored at -20±5ºC until
89 analysis.
90 Then, samples were gradually thawed overnight at refrigeration (5±3ºC) and squeezed
91 through four layers of sterile cheesecloth to separate solid (solids with a particle size smaller
92 than the diameter of the tube) from liquid digesta phases. This latter phase was subsequenty
93 separated into planktonic organisms and bacteria associated with the liquid fraction. The solid
94 phase was separated in associated and adherent fractions. Fractionation procedures were
95 carried out following the methodology described in [19]. The four fractions were lyophilized
96 and composited to obtain a unique sample with the four fractions represented proportionally
97 (on dry matter basis).
98 After composition, DNA extraction was performed using the commercial Power Soil DNA
99 Isolation kit (Mo Bio Laboratories Inc, Carlsbad, CA, USA) following manufacturer's
100 instructions. The extracted DNA was subjected to paired-end Illumina sequencing of the V4
101 hypervariable region of the 16S rRNA [20] and of the V7 region of the 18S rRNA genes. The

libraries were generated by means of Nextera kit. The 250 bp paired-end sequencing reactions were performed on a MiSeq platform (Illumina, San Diego, CA, USA).

Sequence data were processed using the QIIME software package version 1.9.1 [21]. Sequences below 220 bp in length and Phred score below 20 were discarded. In total, 3,261,168 and 3,431,242 reads from the 16S and 18S rRNA regions respectively, were analysed. Sequence data were grouped into operational taxonomic units (OTU) sharing 97% sequence similarity, and assigned to phylogenetic groups by BLAST [22].

Bacterial and archaeal 16S rRNA genes were assigned using the GreenGenes database (May 2013 version) and ciliate protozoal 18S rRNA genes against SILVA database (March 2015 version). Data were summarised at the genus level. Relative abundance (RA) of genera in each animal was calculated after excluding those genera that appeared in <0.1% proportion in the previous step. Only genera showing average RA>0.1% in both breeds were kept for subsequent analyses.

Genotypes from animals under study were also obtained with the Illumina 9K chip (Illumina, Inc, San Diego, CA, USA). A total of 9,146 SNPs with minor allele frequency (MAF) >0.05 in the whole genotyped Spanish population were kept (data provided by the Spanish Holstein association www.conafe.com from more than 3,000 individuals).

We used two strategies to analyse the host genetic effect on the microbiome composition. Our response variable was the RA of the most common ruminal microbes previously found, and the model adjusted by diet treatment (2 groups, with or without cold-pressed rapeseed cake) and age (primiparous *vs* multiparous) groups and days in milk as a covariate. In the first strategy, differences at the breed level (Holstein *vs* Brown Swiss) were estimated (Model 1).

Model 1)

$$RA_{ijklm} = \mu + diet_j + age\_group_k + lactation\_group_l + breed_m + e_{ijklm}$$

The second strategy included the first two principal components (PC) of a genomic relationship matrix instead of the breed effect as

Model 2)

$$RA_{ijklmn} = \mu + diet_j + age\_group_k + lactation\_group_l + PC\_1_m + PC\_2_n + e_{ijklmn}$$

This genomic relationship matrix was calculated as in [23], where the genome relationship between individuals *i* and *j* can be calculated as

$$G_{ij} = \frac{1}{L}\sum_{k=1}^{L}\frac{(g_{ik} - \hat{p}_k)(g_{jk} - \hat{p}_k)}{\hat{p}_k(1-\hat{p}_k)}$$

where g_{ik} refers to the gene frequency value genotypes *AA*, *Aa* and *aa*, coded as 1, 0.5 and 0, respectively, of individual *i* or *j* at locus *k* (*k* = 1, 9146). Gene frequency is half the number of copies of the reference allele *A*. Then, $\hat{p}_k$ was the estimated allele frequency in the whole genotyped population as provided by CONAFE. The first PC of this matrix aims to detect stratification at the breed level (Holstein *vs* Brown Swiss), whereas the subsequent PC are expected to capture genomic differences between individuals.

Bayesian analyses were performed to estimate the breed and principal component effects [24] using an in-home suite of programs written in R software [25]. Evidence of a host genetic effect was considered when the 80% of the posterior distribution for the breed or the PC had the same sign (either positive or negative). This is, 80% of the posterior probability for the respective effect fell either above or below zero. Those microorganisms that showed evidence of a host genetic control were selected to implement genome wide association analyses. Here, the RA of those microorganisms was used as a dependent variable, and the SNP markers were selected as explicative variables in a single marker linear regression model, including

151 breed and diet as environmental factors. The p-values were adjusted on false discovery rate
152 (FDR).
153 The gene content of the significant SNP was examined using the bovine genome annotation
154 in BioMart tool of Emsembl (ensembl.org/biomart) using Ensembl Genes 75 database. The
155 National Center for Biotechnology Information (NCBI) database and PubMed were
156 employed to investigate the potential biological relation of the genes that contained the SNP
157 and the microbes in order to propose candidate genes that underlie the detected associations.

158

159 **RESULTS AND DISCUSSION**

160 The results from the 16S rRNA region showed a 98:2 for the bacteria:archaea ratio. The more
161 abundant bacterial phyla were *Bacteroidetes* (58%), *Firmicutes* (33%) and TM7 (*Candidatus*
162 *Saccharibacteria*) (4%). Methanobacteria was the most abundant clade among the archaeas.
163 Taxa composition was similar to those reported before in other ruminal microbiome
164 communities [7,13], being mainly microbes related to peptide and cellulose degradation or to
165 the synthesis of microbial protein and volatile fatty acids.

166

167 *Bacteria and archaea*

168 The RA of genera analysed are shown in Figure 1. *Prevotella* was the most abundant
169 bacteria-archaea genus in both breeds, followed by *Butyrivibrio* and *Succiniclasticum.* The
170 archaea *Methanobrevibacter* was more abundant than the rest of the archaea genera detected
171 in the samples.
172 Table 1 shows the results from the statistical analyses of the host genetic component on the
173 different RA. The analyses showed differences between breeds for 4 (*Methanobrevibacter*,
174 *Succiniclasticum*, *Prevotella* and *YRC22*) out of the 6 archaea and bacteria genera analysed

175 from 16S rRNA region. However, either the first or second genomic PC were significant for
176 all other genus analysed (Table 1).
177
178 *Ciliate*
179 Figure 2 shows the relative abundance of the analysed ciliate in both breeds. The genus
180 *Entodinium* was the most abundant among the ciliate protozoal, followed by *Isotricha.*
181 Phenotypically, *Ophryoscolex*, *Diplodinium* and *Polypastron* were more abundant in
182 Holstein, whereas *Dasytricha* showed larger RA in Brown Swiss. The breed effects showed
183 differences in 3 (*Isotricha*, *Ophryoscolex* and *Polyplastron*) out of 7 ciliate genus analysed.
184 The genomic PCs were also statistically significant for these genera, except for *Isotricha*,
185 where the posterior distribution did not show a significant effect for the PCs (Table 2).
186
187 Despite the small sample size, RA for 77% (10/13) of the genera analysed were found to be
188 regulated by some host genetic factor (breed, SNP marker, or both), which suggest that the
189 microbiome composition is regulated by some genetic mechanisms in the host. The host
190 genetic background showed to have effect on a larger proportion of bacteria-archaea, in
191 comparison to ciliates. We did not find a host genetic effect on the relative abundance of
192 genera *Trichostomatia*, *Dasytricha*, and *Diplodinium*. These microbes might be more
193 influenced by diet than by the host genetic effect, and larger sample sizes might be neccesary
194 to detect differences between breeds or host genetic effects.
195 [18] also showed host genetic effect on the RA of different genera and families of *Firmicutes*
196 and *Euryarchaeota* (e.g. *Turicibacter*, *Blautia Clostridiaceae*, *Ruminococcaceae* or
197 *Methanobrevibacter*) in humans. Their study also showed a host genetic effect on some
198 *Tenericutes*, *Proteobacteria* (Family *Oxalobacteraceae*) and *Actinobacteria* (Genera
199 *Bifidobacterium* and *Actinomyces*). Our study also shows a host genetic effect on some

genera of Firmicutes but also on some *Bacteroidetes* differently to [18] and ciliate which were not analysed in the human study as they are not abundant in the human gut. Roehe et al. [17] showed differences in the microbial community of progeny daughters from different cattle breeds and sires, suggesting that even under the same diet and environmental circumstances, individuals can differ in their microbial communities depending on their progenitors.

Microbial networks for 16S and 18S-gene rRNA regions were constructed using the algorithm described by [26] and their graphical representations are shown in Figures S1 and S2. The microorganisms that showed to be related to the host genetics are relevant in the composition of the ruminal environment and the degradation of feed. For instance, bacteria from the *Prevotella*, the most abundant group, and *Paraprevotella* genera are involved in the metabolism of proteins and peptides in the rumen. They break down protein and carbohydrates in feed [27], synthesize *de novo* peptides and use products of cellulose degradation from other cellulolytic bacteria [28,29]. Further, bacteria from the genus *Ruminococcus* break down cellulose, hemicellulose and produce succinic acid as a major fermentation product together with acetic and formic acids, hydrogen and CO_2. These products are then used by other bacteria, some from the *Succiniclasticum* genus, which convert succinate to propionate as an energy-yielding mechanism. *Butyrivibrio* bacteria are proteolytic bacteria and are involved in the degradation of hemicellulose walls, and lipid hydrogenation. They produce mainly butyrate, that is metabolized through the rumen wall to produce energy. Further, archaeas from the *Methanobrevibacter* genus use hydrogen and CO_2 products and by-products from other microorganism (e.g. *Ruminococcus*) to synthesise methane. The archaea, mainly organisms related with genera *Methanobrevibacter* and *Methanosphaera*, are highly associated with methane emission in ruminants [27].

225 *Methanobrevibacter* has been associated to methane emissions in many previous studies, e.g.
226 [13,30,31].
227 *Entodinium* ciliate are able to engulf small plant particles and degrade cellulose [27,32].They
228 are considered as cellulolytic microorganisms. *Isotricha* and *Dasytricha* use soluble sugar,
229 and many carbohydrates enzymatic activities have been detected. *Polypastron* ciliates can
230 actively ingest large cellulosic fibres of the rumen fluid [27,33]. The products of rumen
231 ciliates are more or less similar and include acetate, butyrate, and lactate. They also produce
232 CO_2 and hydrogen during the synthesis that can be converted to methane by methanogenic
233 archaea and protozoa. Ciliates interact with other rumen microorganism as they can ingest
234 bacteria as protein source. A host genetic effect on the RA of these microorganisms explain
235 the heritability found in related traits such as feed efficiency or methane yield [34–36].
236
237 Genome-wide association analyses was performed for the RA the four microorganisms that
238 showed significant effect on both breed and PC1 effect (*Methanobrevibacter*,
239 *Succiniclasticum*, *Prevotella*, and *Polyplastron* genera). The generalized linear model
240 implemented included the breed, diet and the bovine SNP marker effects, and p-values were
241 adjusted on false discovery rate (FDR). As expected, the small sample size caused that most
242 of the markers with significant P-values (<0.01) presented a large FDR. We chose the
243 threshold of FDR<0.15 (equivalent P-value of 1.81×10^{-4}) to report significant SNP markers.
244 After this adjustment, significant bovine SNP markers were found for *Prevotella* genus RA
245 (Table 3). Most of these markers were within known genes with functions related mainly to
246 metabolic pathways and signalling on the central neural system. The role of the microbiome
247 in the metabolic status and the development of several central system disorders have been
248 well establish in humans [37,38], and our results suggest that there are also associations
249 between genes involved in metabolic and neural processes and the rumen microbiota

compositions. It must be highlighted that we found association between the DGAT1 gene and the RA of P. Prevotella. The DGAT1 gene is a major gene with a large effect on the fat composition in milk [39–41]. The association found in this study shows that the effect of the DGAT1 on the milk fat composition may be partially regulated by some effect on the microbiome composition, where individuals carrying the A (A/G) allele of the ARS-BFGL-NGS-4939 SNP tend to host a larger proportion of *Prevotella* microorganisms which are also involved in the protein and peptide degradation in the rumen, in the production of saturated fatty acids as well as in saccharolytic pathways. Other genes with significant association to the *Prevotella* RA were the ACSF3, AGPAT3 and STC2, all of them previously associated to fatty acids or cell metabolism.

The results in this study provide some evidence that support the hypothesis of a host genetic component that can partially regulate the composition of the microbiome, and indirectly some metabolic pathways. In this sense, it seems that there is a genetic component in the regulation of some groups of H_2- producing microorganisms included in the *Firmicutes* phylum and ciliate protozoa and H_2-utilizers bacteria associated to *Bacteroidetes*. This is relevant because diets and management practices can be specifically designed to compensate those genotypes that are more susceptible to harbour less efficient microorganisms from a nutritional and energetic point of view. Results from this study must be considered carefully due to the reduced sample size. Future studies should allow to better estimate heritability of the microbiome composition in cattle, as well as covariance components with other traits of interest (e.g. feed efficiency, productivity, or methane emissions). Still, if these results were confirmed, breeding strategies could be developed to select future livestock generations prone to harbor a favourable microbiome composition that improves feed digestion and utilization, while precluding presence of harmful microbes or composition thereof.

275

276

DECLARATIONS

Ethics approval and consent to participate

It was not considered to ask for an authorization because the procedures used in animals were those used under a common clinical veterinary procedure, therefore not subject to regulation by the Spanish and European Legislation related with the protection of animals used for scientific purposes. Nevertheless, the animals were manipulated according to the Spanish Policy for Animal Protection RD 53/2013, which meets the European Union Directive 86/609 about the protection of animals used in experimentation.

Consent for publication

Not applicable

Availability of data and material

The datasets during and/or analysed during the current study available from the corresponding author on reasonable request, and the authors plan to upload them to a data repository soon.

Competing interests

The authors declare that they have no competing interests.

Funding

298 Funding from grants INIA RTA2012-00065-C04 and LIFE SEEDCAPITAL-12 ENV/ES/590
299 is acknowledged.

300 **Authors' contributions**

301 OGR performed the analyses of sequence data, managed the genotyping of the animals,
302 implemented the statistical analyses, discuss the results and wrote the first draft of the
303 manuscript. AGR, IZU and RAT made the experimental design and executed the
304 experiments, collected and analyzed the samples, discussed the results and helped to write the
305 manuscript. AHU performed the preparatory actions for the sequencing analysis and helped
306 writing the manuscript. All authors read and approved the final manuscript.

307 **Acknowledgments**

308 Authors wish to thank Iurancha González for assistance during sampling and Beatriz Oporto
309 and Medelin Ocejo for DNA preparations, as well as the staff from Escuela Agraria de
310 Fraisoro (Basque Department of Environment, Rural Development, Agriculture and Fishery).

311

312

313

314 **REFERENCES**

315 1. Huttenhower C, Knight R, Brown CT, Caporaso JG, Clemente JC, Gevers D, et al.
316 Advancing the microbiome research community. Cell [Internet]. Elsevier Inc.;
317 2014;159:227–30. Available from: http://dx.doi.org/10.1016/j.cell.2014.09.022
318 2. Waldor MK, Tyson G, Borenstein E, Ochman H, Moeller A, Finlay BB, et al. Where next
319 for microbiome research? PLoS Biol. [Internet]. Public Library of Science; 2015 [cited 2015
320 Oct 25];13:e1002050. Available from:
321 http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002050
322 3. Malmuthuge N, Guan LL. Gut microbiome and omics: a new definition to ruminant
323 production and health. Anim. Front. [Internet]. American Society of Animal Science; 2016
324 [cited 2016 Apr 12];6:8. Available from:
325 https://www.animalsciencepublications.org/publications/af/articles/6/2/8
326 4. Jewell KA, McCormick CA, Odt CL, Weimer PJ, Suen G. Ruminal Bacterial Community
327 Composition in Dairy Cows Is Dynamic over the Course of Two Lactations and Correlates
328 with Feed Efficiency. Appl. Environ. Microbiol. [Internet]. 2015 [cited 2016 Jan
329 19];81:4697–710. Available from: http://aem.asm.org/content/early/2015/04/27/AEM.00720-
330 15
331 5. Saro C, Ranilla MJ, Tejido ML, Carro MD. Influence of forage type in the diet of sheep on
332 rumen microbiota and fermentation characteristics. Livest. Sci. [Internet]. Elsevier; 2014
333 [cited 2016 Feb 15];160:52–9. Available from:
334 http://www.livestockscience.com/article/S1871141313005295/fulltext
335 6. Saro C, Ranilla MJ, Carro MD. Postprandial changes of fiber-degrading microbes in the
336 rumen of sheep fed diets varying in type of forage as monitored by real-time PCR and
337 automated ribosomal intergenic spacer analysis. J. Anim. Sci. [Internet]. 2012 [cited 2016

338 Feb 15];90:4487–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23100580
339 7. Henderson G, Cox F, Ganesh S, Jonker A, Young W, Janssen PH. Rumen microbial
340 community composition varies with diet and host, but a core microbiome is found across a
341 wide geographical range. Sci. Rep. [Internet]. Nature Publishing Group; 2015 [cited 2015 Oct
342 17];5:14567. Available from:
343 http://www.nature.com/srep/2015/151009/srep14567/full/srep14567.html
344 8. Mohammed R, Brink GE, Stevenson DM, Neumann AP, Beauchemin KA, Suen G, et al.
345 Bacterial communities in the rumen of Holstein heifers differ when fed orchardgrass as
346 pasture vs. hay. Front. Microbiol. [Internet]. 2014 [cited 2016 Mar 16];5:689. Available from:
347 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4260508&tool=pmcentrez&rende
348 rtype=abstract
349 9. Deng W, Xi D, Mao H, Wanapat M. The use of molecular techniques based on ribosomal
350 RNA and DNA for rumen microbial ecosystem studies: a review. Mol. Biol. Rep. [Internet].
351 2008 [cited 2016 Feb 12];35:265–74. Available from:
352 http://www.ncbi.nlm.nih.gov/pubmed/17484038
353 10. Wadhwa M, Bakshi MPS, Makkar HPS. Modifying gut microbiomes in large ruminants:
354 Opportunities in non-intensive husbandry systems. Anim. Front. [Internet]. 2016;6:27.
355 Available from: https://dl.sciencesocieties.org/publications/af/abstracts/6/2/27
356 11. Yáñez-Ruiz DR, Macías B, Pinloche E, Newbold CJ. The persistence of bacterial and
357 methanogenic archaeal communities residing in the rumen of young lambs. FEMS Microbiol.
358 Ecol. [Internet]. 2010 [cited 2016 Feb 15];72:272–8. Available from:
359 http://www.ncbi.nlm.nih.gov/pubmed/20236326
360 12. Hayes BJ, Lewin HA, Goddard ME. The future of livestock breeding: Genomic selection
361 for efficiency, reduced emissions intensity, and adaptation. Trends Genet. [Internet]. Elsevier
362 Ltd; 2013;29:206–14. Available from: http://dx.doi.org/10.1016/j.tig.2012.11.009

363 13. Wallace RJ, Rooke JA, McKain N, Duthie C-A, Hyslop JJ, Ross DW, et al. The rumen
364 microbial metagenome associated with high methane production in cattle. BMC Genomics
365 [Internet]. BioMed Central; 2015 [cited 2016 Feb 15];16:839. Available from:
366 http://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-015-2032-0
367 14. Wang L, Xu Q, Kong F, Yang Y, Wu D, Mishra S, et al. Exploring the Goat Rumen
368 Microbiome from Seven Days to Two Years. PLoS One [Internet]. Public Library of Science;
369 2016 [cited 2016 May 19];11:e0154354. Available from:
370 http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0154354
371 15. Ross EM, Moate PJ, Marett LC, Cocks BG, Hayes BJ. Metagenomic predictions: from
372 microbiome to complex health and environmental phenotypes in humans and cattle. PLoS
373 One [Internet]. Public Library of Science; 2013 [cited 2016 Feb 15];8:e73056. Available
374 from: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0073056
375 16. Weimer PJ, Stevenson DM, Mantovani HC, Man SLC. Host specificity of the ruminal
376 bacterial community in the dairy cow following near-total exchange of ruminal contents. J.
377 Dairy Sci. [Internet]. Elsevier; 2010;93:5902–12. Available from:
378 http://dx.doi.org/10.3168/jds.2010-3500
379 17. Roehe R, Dewhurst RJ, Duthie C-A, Rooke JA, McKain N, Ross DW, et al. Bovine Host
380 Genetic Variation Influences Rumen Microbial Methane Production with Best Selection
381 Criterion for Low Methane Emitting and Efficiently Feed Converting Hosts Based on
382 Metagenomic Gene Abundance. PLoS Genet. [Internet]. Public Library of Science; 2016
383 [cited 2016 Mar 21];12:e1005846. Available from:
384 http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1005846
385 18. Goodrich JK, Davenport ER, Beaumont M, Jackson MA, Knight R, Ober C, et al. Genetic
386 Determinants of the Gut Microbiome in UK Twins. Cell Host Microbe [Internet]. Elsevier;
387 2016 [cited 2016 May 11];19:731–43. Available from:

388 http://www.cell.com/article/S1931312816301536/fulltext

389 19. Yu Z, Foster R. Methods in gut microbial ecology for ruminants. The Netherlands:

390 Springer Netherlands; 2005.

391 20. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et

392 al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample.

393 Proc. Natl. Acad. Sci. U. S. A. [Internet]. National Academy of Sciences; 2011 [cited 2016

394 Oct 17];108 Suppl:4516–22. Available from:

395 http://www.ncbi.nlm.nih.gov/pubmed/20534432

396 21. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al.

397 QIIME allows analysis of high-throughput community sequencing data. Nat. Methods

398 [Internet]. Nature Publishing Group; 2010 [cited 2014 Jul 10];7:335–6. Available from:

399 http://dx.doi.org/10.1038/nmeth.f.303

400 22. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search

401 tool. J. Mol. Biol. [Internet]. 1990 [cited 2016 Jun 22];215:403–10. Available from:

402 http://www.ncbi.nlm.nih.gov/pubmed/2231712

403 23. Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N, Cunningham JM, et al.

404 Genome partitioning of genetic variation for complex traits using common SNPs. Nat. Genet.

405 [Internet]. 2011;43:519–25. Available from:

406 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4295936&tool=pmcentrez&rende

407 rtype=abstract

408 24. Sorensen D, Gianola D. Likelihood, Bayesian and MCMC methods in quantitative

409 genetics. Springer-Verlag; 2002.

410 25. R Core Team. R: A language and environment for statistical computing [Internet].

411 Vienna, Austria: R Foundation for Statistical Computing; 2015. Available from:

412 https://www.r-project.org/

413 26. Cuscó A, Casellas J, Francino O. CoMINO v.1.2: Network analysis program for
414 directional interactions in microbiome. Barcelona Debates Hum. microbiome from microbes
415 to Med. 2016.
416 27. Jouany JP. Rumen Microbial Metabolism and Ruminant Digestion [Internet]. Editions
417 Quae; 1991 [cited 2016 Jul 11]. Available from: https://books.google.fr/books?id=Bcsa8Z4u-
418 E4C
419 28. Atasoglu C, Valdés C, Walker ND, Newbold CJ, Wallace RJ. De novo synthesis of amino
420 acids by the ruminal bacteria Prevotella bryantii B14, Selenomonas ruminantium HD4, and
421 Streptococcus bovis ES1. Appl. Environ. Microbiol. [Internet]. American Society for
422 Microbiology (ASM); 1998 [cited 2016 Oct 17];64:2836–43. Available from:
423 http://www.ncbi.nlm.nih.gov/pubmed/9687438
424 29. Lou J, Dawson KA, Strobel HJ. Glycogen Formation by the Ruminal Bacterium
425 Prevotella ruminicola. Appl. Environ. Microbiol. 1997;63:1483–8.
426 30. Zhou M, Hernandez-Sanabria E, Guan LL. Characterization of variation in rumen
427 methanogenic communities under different dietary and host feed efficiency conditions, as
428 determined by PCR-denaturing gradient gel electrophoresis analysis. Appl. Environ.
429 Microbiol. [Internet]. 2010 [cited 2016 Feb 15];76:3776–86. Available from:
430 http://aem.asm.org/content/76/12/3776.full
431 31. Zhou M, Hernandez-Sanabria E, Guan LL. Assessment of the microbial ecology of
432 ruminal methanogens in cattle with different feed efficiencies. Appl. Environ. Microbiol.
433 [Internet]. 2009 [cited 2016 Feb 15];75:6524–33. Available from:
434 http://aem.asm.org/content/75/20/6524.abstract
435 32. Dauvrin T. La caractérisation de l'invertase du Cilié du rumen Isotricha prostoma, révèle
436 certaines propriétés originales. Université Catholique de Louvain; 1988.
437 33. Belanche A, de la Fuente G, Moorby JM, Newbold CJ. Bacterial protein degradation by

438 different rumen protozoal groups. J. Anim. Sci. [Internet]. American Society of Animal
439 Science; 2012 [cited 2016 Jul 11];90:4495–504. Available from:
440 http://www.animalsciencepublications.org/publications/jas/abstracts/90/12/4495
441 34. Basarab JA, Beauchemin KA, Baron VS, Ominski KH, Guan LL, Miller SP, et al.
442 Reducing GHG emissions through genetic improvement for feed efficiency: effects on
443 economically important traits and enteric methane production. Animal [Internet]. 2013 [cited
444 2016 Feb 15];7 Suppl 2:303–15. Available from:
445 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3691002&tool=pmcentrez&rende
446 rtype=abstract
447 35. Pryce JE, Wales WJ, de Haas Y, Veerkamp RF, Hayes BJ. Genomic selection for feed
448 efficiency in dairy cattle. Animal [Internet]. 2014;8:1–10. Available from:
449 http://www.journals.cambridge.org/abstract_S1751731113001687\nhttp://journals.cambridge
450 .org/download.php?file=/ANM/ANM8_01/S1751731113001687a.pdf&code=5f1942a3f4fe81
451 0a3a840b894cced959\nhttp://www.ncbi.nlm.nih.gov/pubmed/24128704
452 36. Pryce JE, Gonzalez-Recio O, Nieuwhof G, Wales WJ, Coffey MP, Hayes BJ, et al. Hot
453 topic: Definition and implementation of a breeding value for feed efficiency in dairy cows. J.
454 Dairy Sci. [Internet]. Elsevier Inc.; 2015 [cited 2016 Feb 8];98:7340–50. Available from:
455 http://www.scopus.com/inward/record.url?eid=2-s2.0-84942553085&partnerID=tZOtx3y1
456 37. Foster JA, Lyte M, Meyer E, Cryan JF. Gut microbiota and brain function: An evolving
457 field in neuroscience. Int. J. Neuropsychopharmacol. 2015;
458 38. Nicholson JK, Holmes E, Kinross J, Burcelin R, Gibson G, Jia W, et al. Host-Gut
459 Microbiota Metabolic Interactions. Science (80-.). 2012;336.
460 39. Mohammed S, Mohammed SA, Rahamtalla SA, Ahmed SS, elamin khalid mohammed,
461 Dousa BM, et al. DGAT1 Gene in Dairy Cattle: A Review. Glob. J. Anim. Sci. Res.
462 [Internet]. 2014 [cited 2016 Oct 11];3:191–8. Available from:

463 http://www.gjasr.com/index.php/GJASR/article/view/142

464 40. Grisart B, Farnir F, Karim L, Cambisano N, Kim J-J, Kvasz A, et al. Genetic and

465 functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in

466 affecting milk yield and composition. Proc. Natl. Acad. Sci. U. S. A. [Internet]. National

467 Academy of Sciences; 2004 [cited 2016 Oct 11];101:2398–403. Available from:

468 http://www.ncbi.nlm.nih.gov/pubmed/14983021

469 41. Coppieters W, Riquet J, Arranz J-J, Berzi P, Cambisano N, Grisart B, et al. A QTL with

470 major effect on milk yield and composition maps to bovine Chromosome 14. Mamm.

471 Genome [Internet]. Springer-Verlag; 1998 [cited 2016 Oct 11];9:540–4. Available from:

472 http://link.springer.com/10.1007/s003359900815

473

474

475

476 ***Table 1***. *Effect*[1] *of the breed (Holstein* vs *Brown Swiss) and the two first principal component*
477 *of a genomic relationship matrix based on genotypes on the relative abundance of different*
478 *bacteria and archaea genera. Only genera that are present with average relative abundance*
479 *larger than 0.1% in both breeds are shown.*

Domain	Phylum	(Family) Genus	Breed	PC1	PC2
Archaea	*Euryarchaeota*	*Methanobacteriaceae Methanobrevibacter*	*	*	*
Bacteria	*Firmicutes*	*Lachnospiraceae Butyrivibrio*	N.S.	N.S.	*
Bacteria	*Firmicutes*	*Veillonellaceae Succiniclasticum*	*	*	N.S.
Bacteria	*Firmicutes*	*Ruminococcaceae Ruminococcus*	N.S.	N.S.	*
Bacteria	*Bacteroidetes*	*Prevotellaceae Prevotella*	*	*	*
Bacteria	*Bacteroidetes*	*Paraprevotellaceae YRC22*	*	N.S.	*

480 [1] * states that >80% of the posterior distribution of the effect was either larger or lower than
481 zero, suggesting a significant effect of the breed or of the principal component on the relative
482 abundance. N.S. states otherwise.

483

484

Table 2. *Effect*[1] *of the breed (Holstein vs Brown Swiss) and the two first principal component of a genomic relationship matrix based on genotypes on the relative abundance of different ciliate genera. Only genera that are present with average relative abundance larger than 0.1% in both breeds are shown.*

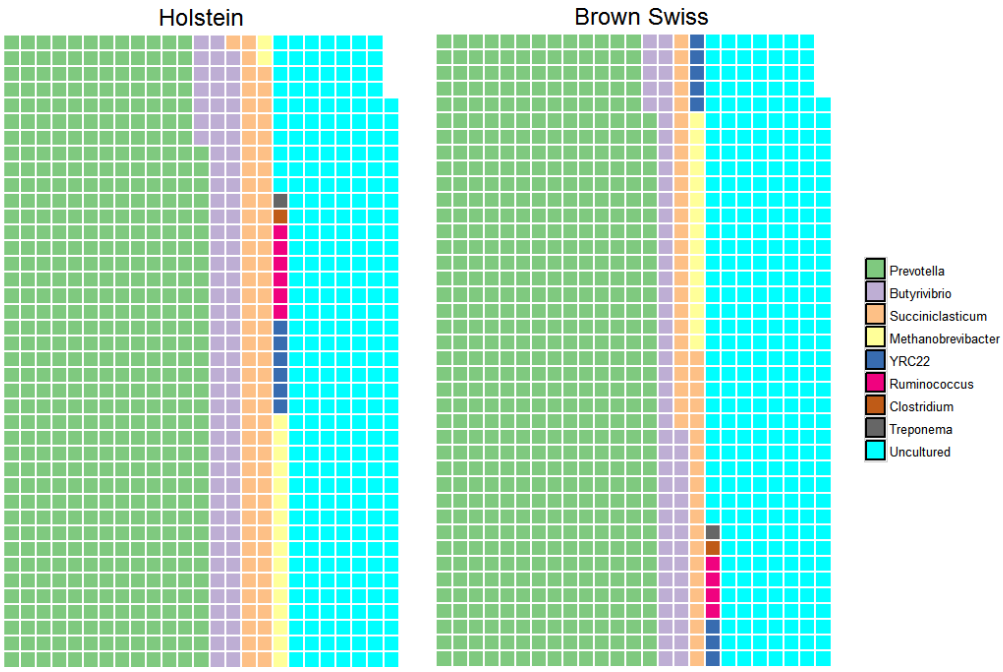
Domain	Order	(Family) *Genus*	Breed	PC1	PC2
Eukaryota	*Ciliophora*	*Litostomatea Trichostomatia*	N.S.	N.S.	N.S.
Eukaryota	*Ciliophora*	*Trichostomatia Dasytricha*	N.S.	N.S.	N.S.
Eukaryota	*Ciliophora*	*Trichostomatia Entodinium*	N.S.	N.S.	*
Eukaryota	*Ciliophora*	*Trichostomatia Isotricha*	*	N.S.	N.S.
Eukaryota	*Ciliophora*	*Trichostomatia Diplodinium*	N.S.	N.S.	N.S.
Eukaryota	*Ciliophora*	*Trichostomatia Ophryoscolex*	*	N.S.	*
Eukaryota	Ciliophora	Trichostomatia *Polyplastron*	*	*	*

[1] * states that >80% of the posterior distribution of the effect was either larger or lower than zero, suggesting a significant effect of the breed or of the principal component on the relative abundance. N.S. states otherwise.

***Table 3**. Genes contained within significant bovine SNP markers for the relative abundance of P. Prevotella, and their position.*

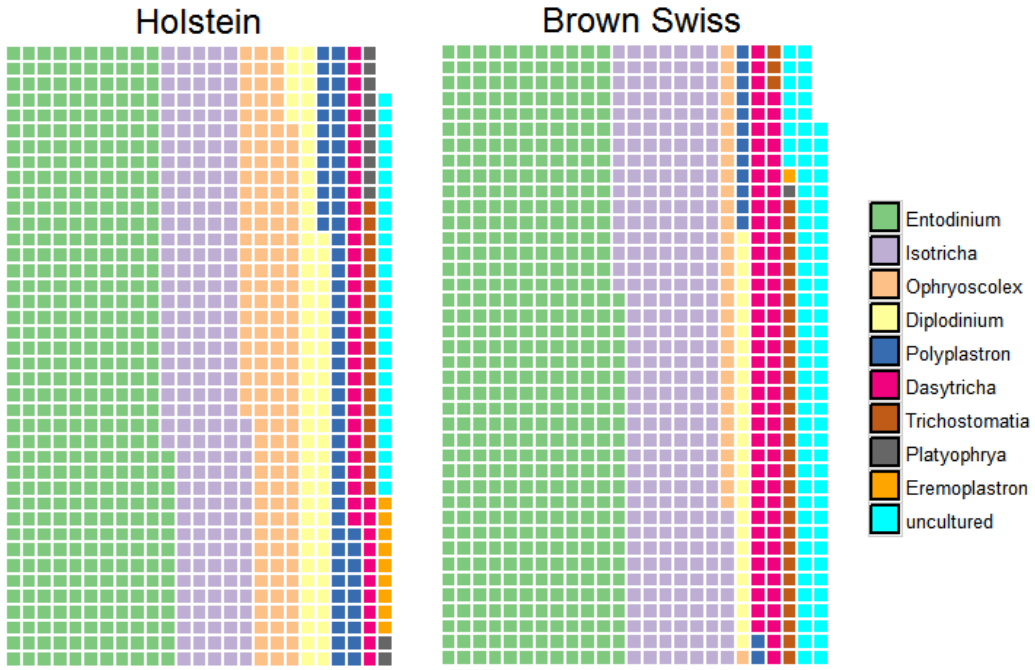
SNP name	SNP position	Gene	Related function	P-value
ARS-BFGL-NGS-13121	1:146833973	AGPAT3	Metabolic pathways and glycerolipid metabolism	$5.02x10^{-5}$
ARS-BFGL-NGS-106490	3:13635591	Unknown	-	$1.60\ x10^{-4}$
ARS-BFGL-NGS-28573	3:24081964	Unknown	-	$2.05\ x10^{-5}$
BTB-01512420	8:72495155	ADAMDEC1	Dendritic cell maturation and functions	$1.56\ x10^{-4}$
ARS-BFGL-NGS-32158	12:90983897	RASA3	Ras signalling pathway; control of intracellular signaling networks	$1.81\ x10^{-4}$
ARS-BFGL-NGS-4939	14:1801116	DGAT1	Conversion of diacylglycerol and fatty acyl CoA to triacylglycerol; metabolic status	$1.81\ x10^{-4}$
ARS-BFGL-NGS-31386	18:14208633	ACSF3	Fatty acid, triacylglycerol, and ketone body metabolism	$1.81\ x10^{-4}$
ARS-BFGL-NGS-112014	18:34794005	CES3	Fatty acyl and cholesterol ester metabolism	$1.72\ x10^{-4}$
ARS-BFGL-NGS-31292	20:4907906	STC2	Autocrine or paracrine functions and cell metabolism	$1.81\ x10^{-4}$
ARS-BFGL-NGS-31656	26:51426365	Unknown	-	$1.81\ x10^{-4}$
BTA-122892-no-rs	X:81638519	SLC16A2	Transporter of thyroid hormone and development of the central nervous system	$1.81\ x10^{-4}$

Figure 1. *Relative abundance of Bacteria and Euryarchaea with average relative abundance larger than 0.1% in both breeds.*



513 ***Figure 2***. *Relative abundance of genera of ciliate with average relative abundance larger*

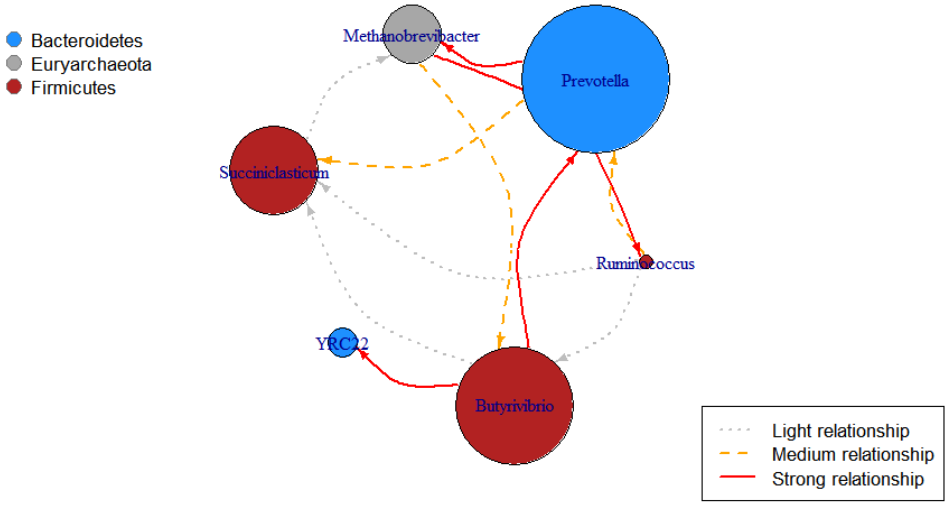
514 *than 0.1% in both breeds.*



515

516

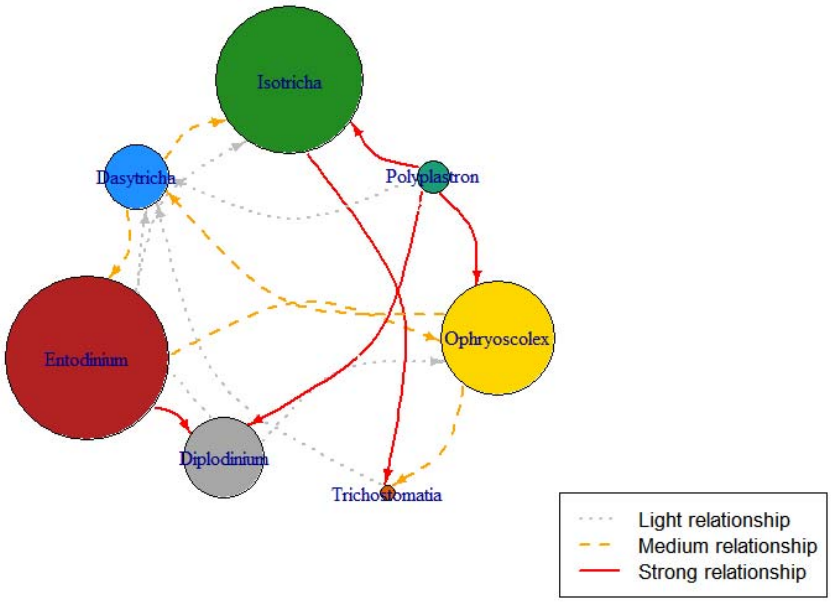
517 ***Figure S1***. *Microbial network based on 16S rRNA-gene based region for microorganism*
518 *with relative abundance larger than 0.1% in both breeds. The size of the nodes represents the*
519 *relative abundance of the genera.*



520

521

522 ***Figure S2***. *Microbial network based on 18S rRNA-gene based region for ciliates with*
523 *relative abundance larger than 0.1% in both breeds. The size of the nodes represents the*
524 *relative abundance of the genera.*



525

526