Bovine host genome acts on rumen microbiome function linked to methane emissions

Marina Martínez-Álvaro, Marc D. Auffret, Carol-Anne Duthie, Richard J. Dewhurst, Matthew A. Cleveland, Mick Watson & Rainer Roehe

Our study provides substantial evidence that the host genome affects the comprehensive function of the microbiome in the rumen of bovines. Of 1,107/225/1,141 rumen microbial genera/metagenome assembled uncultured genomes (RUGs)/genes identified from whole metagenomics sequencing, 194/14/337 had significant host genomic effects (heritabilities ranging from 0.13 to 0.61), revealing that substantial variation of the microbiome is under host genomic control. We found 29/22/115 microbial genera/RUGs/genes host-genomically correlated (|0.59| to |0.93|) with emissions of the potent greenhouse gas methane (CH₄), highlighting the strength of a common host genomic control of specific microbial processes and CH₄. Only one of these microbial genes was directly involved in methanogenesis (cofG), whereas others were involved in providing substrates for archaea (e.g. bcd and pccB), important microbial interspecies communication mechanisms (ABC.PE.P), host-microbiome interaction (TSTA3) and genetic information processes (RP-L35). In our population, selection based on abundances of the 30 most informative microbial genes provided a mitigation potential of 17% of mean CH₄ emissions per generation, which is higher than for selection based on measured CH₄ using respiration chambers (13%), indicating the high potential of microbiome-driven breeding to cumulatively reduce CH₄ emissions and mitigate climate change.
Ruminants harbour a unique symbiotic gut microbial population that transforms indigestible fibrous feed into high-quality products such as meat and milk for human consumption. Moreover, ruminant livestock is vital to meet global food security and contribute to poverty reduction in an increasing world population. Yet to be solved is the negative environmental impact, as dairy and beef cattle account for 9.5% of all anthropogenic greenhouse gas (GHG) emissions. Of those, ruminal microbial fermentation represents 40–50%; in particular, due to the highly potent GHG methane (CH\(_4\)). Additionally, CH\(_4\) emissions imply a considerable energy loss to the animal, ranging from 2 to 12% of gross energy intake. Therefore, decreasing CH\(_4\) emissions is acknowledged to contribute to the mitigation of climate change and optimize the economic efficiency of cattle production. Ruminal methanogenesis is a complex process dependent on the cooperation of taxonomic communities with different metabolic activities. A diverse community of bacteria, ciliate protozoa, and anaerobic fungi convert complex diet carbohydrates, proteins, and lipids into volatile fatty acids, lactate, microbial proteins, and vitamins while releasing CO\(_2\), H\(_2\), and other compounds. Four orders of ruminal methanogenic archaea use electrons derived from H\(_2\), formate or methyl compounds to reduce carbon dioxide into CH\(_4\) to obtain metabolic energy (Supplementary Table 1). CH\(_4\) emissions are generated from whole metagenomic sequence data of the rumen microbial species, nor the ability of some microbial populations to modulate methanogenesis to reduce CH\(_4\) emissions during ruminal archaea fermentation. This is due to the high phenotypic variability in CH\(_4\) emissions. The core functional microbial KEGG genes were identified by metagenome sequencing of rumen microbial DNA samples from a bovine population designed for a powerful host genomic analysis with high standardization of diets and other husbandry effects. We characterized the core ruminal microbiome by identifying 1,108 cultured microbial genera by mapping our sequences to the Hungate1000 reference genome collection and RefSeq databases (Supplementary Data 1); 225 ruminal uncultured genomes (RUGs) by de novo metagenome-assembly of genomes, 34 of them classified at strain level (Supplementary Data 2), and 1142 functional microbial genes (Supplementary Data 3); present in all (n = 359) and for RUGs in >200 of our animals. Our specific hypothesis is that the host genome influences the abundance of not only functional microbial genes involved in metabolism, but also in interspecies communication, host–microbiome interactions, and genetic information processing. These functions may play a key integrating role in achieving a ruminal balance where fermentation of feed into essential nutrients utilized by the host is optimized and substrates utilized by methanogenesis e.g. H\(_2\) excess are minimized. Our comprehensive description of ruminal microbiome functionalities includes the abundances of 34 microbial genes carried by methanogenic archaea directly implicated in CH\(_4\) metabolism; 511 involved in other metabolic pathways of bacteria, archaea, ciliate protozoa or fungi, indirectly influencing methanogenesis by minimizing required substrates through non-methanogenic routes that yield beneficial nutrients for ruminants (e.g. acetogenesis, propionogenesis, fucose sensing) carried by fungi, bacteria and archaea; 207 in microbial communication processes and host-microbiome interaction (e.g. ABC transporters of different metabolites or fermentation products); 511 involved in genetic information processes (e.g. ribosomal biosynthesis) related to microbial growth; 60 at present not functionally characterized. For each of these 2475 functional and taxonomic characteristics of the rumen microbiome, the host genomic determination and correlation with CH\(_4\) emissions were analysed. After stringent adjustment for multiple testing, we demonstrate that our hypothesis of a common host genomic control is validated by discoveringheritabilities (h\(^2\)) of microbial profiles and host-genomic correlations with CH\(_4\) emissions (r\(_{\text{gCH}_4}\)) significantly deviating from zero, which shows the effectiveness of this strategy. Our results are obtained in bovines, but also provide an indication of potential host genomic effects on functional microbial genes and their biological processes in other species.

Besides providing a better understanding of the complex host genomic effects on the rumen microbiome function, this research provides the basis for a cost-effective microbiome-driven breeding strategy to mitigate CH\(_4\) emissions from cattle without measuring it directly, which is necessary considering the cost-prohibitive limitations of measuring individual animal CH\(_4\) emissions.

**Results**

**Bovine host genomics affected CH\(_4\) emissions produced by ruminal archaea.** CH\(_4\) emissions were accurately measured from individual beef cattle (n = 285) using the gold-standard method of respiration chambers. Animals within the same breed or diet expressed high phenotypic variability in CH\(_4\) emissions with coefficients of variation from 16.3% to 28.5% (Supplementary Fig. 1a, b). Genomic h\(^2\) of CH\(_4\) emissions revealed that 33%...
Host genomics affects the ruminal microbiome composition. We next investigated the proportion of the ruminal microbiome variation at taxonomic and functional levels explained by the host genomic variation among individuals, by estimating $h^2$ of the ruminal abundances of 1107 genera, 225 RUGs, and 1141 microbial genes. Our results demonstrate significant host genomic effects with $h^2$ in a range between 0.13 and 0.61 for the abundances of 194 microbial genera, 14 RUGs, and 337 microbial genes representing cumulatively 58.4, 5.63, and 27.2%, respectively, of the total relative abundance (RA) (Fig. 1 and Supplementary Data 4–6). Among the 194 genera, 20 were highly heritable ($h^2 > 0.40$), which belonged exclusively to bacteria (e.g. Firmicutes *Acidaminococcus* (RA = 0.3%), $h^2 = 0.54$, BF = 8.82 × 10$^{+5}$) and archaea (e.g. hydrogenotrophic methanogen *Methanospirillum* (RA = 0.0005%, $h^2 = 0.40$, BF = 1.04 × 10$^{-2}$). Host genome also shaped the abundance of the hydrogenotrophic/methylo trophic methanogen *Candidatus Methanoplasma* (RA = 0.002%, $h^2 = 0.32$, BF = 6.31 × 10$^{+2}$), and to a lesser extent the abundance of ubiquitous *Methanobrevibacter* (RA = 5.02%, $h^2 = 0.24$, BF = 9.10)—which is coherent with estimates from other studies$^{18–20}$—*Candidatus Methanomethylophilus* (RA = 0.05%, $h^2 = 0.26$, BF = 4.18), *Methanothermus* (RA = 0.002%, $h^2 = 0.25$, BF = 1.26 × 10$^{+1}$) and *Methanoplanus* (RA = 0.0008%, $h^2 = 0.24$, BF = 9.56). Reinforcing the evidence of a host-genomic component in the abundance of methanogenic archaea, 5 RUGs annotated as uncultured *Methanobrevibacter* strains (RA > 0.27%) demonstrated moderate to high $h^2$ estimates (0.39–0.48, BF from 3.5 to 4.65 × 10$^{+3}$), indicating that more specific classification using RUGs provides the opportunity to find highly heritable strains. The most abundant complex carbohydrates degraders in the rumen—*Eubacterium* (RA = 1.02%), *Prevotella* (RA = 39.2%), *Butyrivibrio* (RA = 2.54%), *Bacteroides* (RA = 1.39%) and *Pseudobutyribiviro* (RA = 0.54%)—were highly ($h^2 = 0.51$ for *Eubacterium*, BF = 9.72 × 10$^{+3}$) or moderately ($h^2 = 0.23–0.33$ for the others, BF from 7.42 to 9.73 × 10$^{+3}$) heritable; with 7 highly abundant RUGs (RA > 0.25%) classified as uncultured *Prevotellaceae bacterium* having $h^2$ from 0.32 to 0.45 (BF from 7.48 to 1.67 × 10$^{+2}$). These results support the concepts of a “core heritable microbiome”$^{21,64}$ and stability over time of certain microbial genera abundance such as *Prevotella*.$^{65}$ None of the fungi and protist genera, which are considered to be non-essential for rumen function and highly variable within different host species,$^{66}$ were highly heritable. We elucidated that the specific functional capacity of the ruminal microbiome is heritable by estimating the $h^2$ of a comprehensive set of microbial genes, of which 33 were highly ($h^2 > 0.4$), and 304 were moderately ($0.2 < h^2 < 0.4$) heritable. These microbial genes are involved in a wide variety of metabolic functions (Fig. 1b and Supplementary Data 6), e.g. synthesis of microbial proteins or volatile fatty acids, suggesting that the host genome influences the growth of microbes responsible for the release of nutrients during microbial fermentation.$^{67,68}$ Among 34 microbial genes involved in the CH$_4$ metabolism pathway, 13 showed moderate $h^2$ of 0.22–0.29 (BF from 3.96 to 7.82 × 10$^{+1}$), e.g. mcrA, mcrB, mtrD, mtrE, and cofG. Ribosomal biosynthesis was revealed to be under strong host-genomic
control with 56 heritable microbial genes, representing a cumulative RA of 6.57%, including 9 highly heritable genes ($h^2 = 0.40–0.53$, BF = $1.62 \times 10^{-2}$ to $1.18 \times 10^{-8}$) synthesizing the large ribosomal subunit. Intracellular ribosomal biosynthesis reflects the growth rate of microbial organisms, given that ribosomes can account for up to 40% of their cellular dry mass, and cell fitness and optimal growth are tightly coupled to efficient protein synthesis. Demonstrating that differences among animals in complex microbiome functions are partly due to host genomic variation opens up opportunities to consider a new source of genetic variation not only in ruminants but also in humans, where the $h^2$ of microbial gene abundances was estimated to be even larger ($0.65–0.91$).

**Ruminal microbial mechanisms related to CH$_4$ emissions are influenced by host genomics.** The existence of a common host genomic influence on CH$_4$ emissions and the rumen microbiome was evaluated by estimating host-genomic correlations between CH$_4$ emissions and each microbial genus/RUG/gene abundance ($r_{gCH4}$). Based on the probability of $r_{gCH4}$ being different from 0 ($P_0$) ≥ 0.95, our study revealed 29 microbial genera, 22 RUGs, and 115 functional microbial genes strongly host-genomically correlated with CH$_4$ emissions ($r_{gCH4}$ from $0.59$ to $0.93$, Supplementary Data 7–9). Among the significant microbial communities, most were bacteria (22 genera/17 RUGs) belonging to Bacteroidetes (5/14), Firmicutes (6/2), and Proteobacteria (9/1) phyla. Most microbial genes with strong $r_{gCH4}$ were not directly involved in CH$_4$ metabolism pathways, but rather mechanisms indirectly affecting CH$_4$ production most likely by limiting substrates for methanogenesis, inhibiting methanogens, playing a role in coordinating actions among microbial communities and the host or leading microbial genetic processes. Only H$_2$-oxidizing *Methanobrevibacter* sp. can produce CH$_4$ even under a challenging ruminal environment (e.g., low pH value), however, at a substantially lower level than those adapted to more favourable conditions. To visualize which microbial genus/gene abundances in the rumen are influenced by a common host genomic background, we constructed a co-abundance network based on Pearson correlations among deregressed host genomic effects for each microbial genus/RUG/gene (Supplementary Fig. 2 and Supplementary Data 10). This approach revealed co-abundance clusters of bacterial and fungal genera with strong $r_{gCH4}$ and methanogenic archaea, e.g. fungal *Metschnikowia* ($r_{gCH4} = 0.77$, $P_0 = 0.96$) and archaeal *Methanosarcina* (cluster 9 in Supplementary Fig. 2); and of microbial genes not directly involved in CH$_4$ metabolism but with strong $r_{gCH4}$ (e.g. RP-L6, $r_{gCH4} = 0.71$, $P_0 = 0.96$) and those involved directly in CH$_4$ metabolism (e.g. *flaA*, cluster 1 in Supplementary Fig. 2). The most important host-genomically affected ruminal microbial mechanisms associated with CH$_4$ production (based on $r_{gCH4}$) are as follows:

**Microbial metabolism.** An extensive group of microbial genes involved in amino acid metabolic and transport pathways displayed negative $r_{gCH4}$. Part of this group of microbial genes was involved in the biosynthesis of arginine and branched-chain amino acids via oxocarboxylic acid metabolism ($argF$, $argD$, and *ilvA* with $r_{gCH4} = -0.84$ to $-0.88$, $P_0 = 0.96$; and *argL*, $r_{gCH4} = -0.67$ to $-0.93$, $P_0 = 0.99$). The existence of a common host genomic background, we constructed a co-abundance network based on Pearson correlations among deregressed host genomic effects for each microbial genus/RUG/gene (Supplementary Fig. 2 and Supplementary Data 10). This approach revealed co-abundance clusters of bacterial and fungal genera with strong $r_{gCH4}$ and methanogenic archaea, e.g. fungal *Metschnikowia* ($r_{gCH4} = 0.77$, $P_0 = 0.96$) and archaeal *Methanosarcina* (cluster 9 in Supplementary Fig. 2); and of microbial genes not directly involved in CH$_4$ metabolism but with strong $r_{gCH4}$ (e.g. RP-L6, $r_{gCH4} = 0.71$, $P_0 = 0.96$) and those involved directly in CH$_4$ metabolism (e.g. *flaA*, cluster 1 in Supplementary Fig. 2). The most important host-genomically affected ruminal microbial mechanisms associated with CH$_4$ production (based on $r_{gCH4}$) are as follows:

**Microbial metabolism.** An extensive group of microbial genes involved in amino acid metabolic and transport pathways displayed negative $r_{gCH4}$. Part of this group of microbial genes was involved in the biosynthesis of arginine and branched-chain amino acids via oxocarboxylic acid metabolism ($argF$, $argD$, and *ilvA* with $r_{gCH4} = -0.84$ to $-0.88$, $P_0 = 0.96$; and *argL*,
argC, alaA, ilvH, leuB, and leuD with \( \text{rgCH}_4 = -0.55 \) to \(-0.77\) at lower evidence \( P_0 \geq 0.85\), Figs. 2 and 3). Aconitate hydratase (ACO) catalysing the isomerization of citrate to isocitrate in the early stage of the oxocarboxylic chain extension, and bcd and pccB degrading branched-chain amino acids into branched-chain volatile fatty acids, which have an inhibitory effect on methanogens\(^46\), also expressed negative \( \text{rgCH}_4 = -0.76 \) to \(-0.90\) (\( P_0 \geq 0.95\)). We also estimated negative \( \text{rgCH}_4 \) for microbial genes coding ABC transporters of polar and branched-chain amino acids (\( \text{ABC.PA.A} \), \( \text{ABC.PA.S} \), \( \text{livH} \), \( \text{livG} \), and \( \text{livK} \)) with \( \text{rgCH}_4 = -0.83\) and \(-0.93\), \( P_0 \geq 0.95\). Another group of microbial genes was related to the metabolism of aromatic amino acids tryptophan, tyrosine, and phenylalanine (\( \text{AROA2} \), \( \text{trpA} \), \( \text{trpD} \), \( \text{trpE} \), \( \text{tyrA2} \) and \( \text{paaH} \) with \( \text{rgCH}_4 = -0.74 \) to \(-0.87\), \( P_0 \geq 0.95\) and \( \text{aroC} \), \( \text{aroA} \), \( \text{aroF} \), \( \text{trpG} \), and \( \text{trpB} \), with \( \text{rgCH}_4 = -0.68 \) to \(-0.74\) at lower evidence \( P_0 \geq 0.85\), Fig. 4). More specifically, \( \text{trpE} \), \( \text{trpD} \), and \( \text{trpA} \) take part in the metabolism of L-tryptophan (Fig. 4) whose catabolites (e.g. indole) are important signalling molecules in biofilm formation\(^78\), and activation of the host immune system\(^79\). Moreover, 2-oxocarboxylic acid and tyrosine catabolites are precursors for the biosynthesis of coenzyme \( \text{B} \)\(^77,80\) and methanofuran\(^73\) methanogenic cofactors, and their diversion into the synthesis of other substrates (e.g. arginine, branched-chain amino acids or tryptophan) could explain their negative \( \text{rgCH}_4 \). Lastly, four microbial genes with negative \( \text{rgCH}_4 \) (\(-0.61\) to \(-0.87\), \( P_0 \geq 0.95\)) were associated with methionine metabolism (\( \text{metE} \), \( \text{DNMT1} \)) and transport (\( \text{metQ} \) and \( \text{metN} \)). Methionine is associated with minor methylotrophic methanogenesis pathway\(^81\) in the rumen\(^82,83\) and with an enhancement of microbial long-chain fatty acid production\(^84\), a highly \( \text{H}_2 \) demanding process\(^84\). Our study highlights that the negative association between microbial amino acid metabolism and \( \text{CH}_4 \)\(^85,86\) has a host genomic
component. This could be partly due to host genomic effects on ruminal feed retention times, which have opposite effects on microbial protein synthesis efficiency and CH4 production. We obtained negative \( r_{gCH4} \) (from \(-0.60 \) to \(-0.85\), \( P_0 \geq 0.95\)) for the abundance of several microbial genes responsible for sucrose metabolism (\( sacA, maIZ, bgLB, SPP, \) and \( sucrose phosphorylase, Ras \), Fig. 5), including the highly abundant sucrose fermenter \( \text{Eubacterium} \) (RA = 1.02%), transporters of multiple sugars across the membrane (\( ABC.MS.P1, ABC.MS.S, \) and \( ABC.MS.P \)), and the microbial gene \( PTS-EI \) that catalyses the phosphorylation of incoming sugar substrates concomitantly with their translocation across the cell membrane. Microorganisms capable of fast growth on soluble sugars are suggested to be favoured in hosts with low rumen size and high digesta turnover rate, features also associated with low CH4 emissions. Despite the abundance of easily fermentable carbohydrates, such as sucrose or starch, causes a pH decline which has a strong CH4 reducing effect as a result of pH sensitivity of methanogens or H2-producing microbes. Furthermore, previously mentioned microbial genes \( aroA, trpE \) and \( trpE \) are involved in the shikimate pathway linking sugar metabolism with the synthesis of microbial proteins (aromatic amino acids, tyrosine, phenylalanine, and tryptophan) which are an important source of amino acids for the host. Microbial protein yield from sucrose is suggested to be more persistent over time in comparison to other carbohydrates and partially stored by sucrose utilizers (e.g. \( \text{Eubacterium} \)) for the maintenance of the microbial population.

We also found negative \( r_{gCH4} \) for the abundance of hydrogenotrophic acetogenic bacteria \( \text{Blautila} \) (RA = 0.85%\), together with \( \text{Eubacterium} \) (RA = 0.90\%), and the \( hfs \) microbial gene involved in the reductive Wood-Ljungdahl acetyl-CoA pathway (RA = 0.98%). Acetogens produce volatile fatty acids (mainly acetate but also propionate and butyrate), which serve as host nutrients to improve animal performance and simultaneously compete against methanogens for metabolic \( \text{H}_{2} \) consumption. Despite acetogenesis being thermodynamically less favourable than the reduction of \( \text{CO}_2 \) into \( \text{CH}_4 \), the rumen, this may vary upon microbial interactions and host-genomically influenced ruminal environmental factors. Propionogenesis via acrylate and lactate pathways, and butyrate production from lactate not only reduces \( \text{H}_2 \) availability for methanogenesis but also prevents rumen acidosis and microbiome in redox reactions presented \( r_{gCH4} \) from \(-0.71 \) to \(-0.87\) (\( P_0 \geq 0.96\)). The first two proteins are involved in the synthesis of substrates required for methanogenic cofactors; i.e. \( \text{bioB} \) catalyses the conversion of dethiobiotin to biotin, which competes with coenzyme B for the synthesis of its alkyl portion and \( \text{cobl} \) together with \( \text{hemC} \) (RA = 0.91, \( P_0 = 1.00\)) take part in porphyrin metabolism, required for different processes including the synthesis of porphyrin-based cofactors vitamin \( B_12 \) and F430. Nitrogen fixation protein \( \text{nifU} \) carries out \( N_2 \) reduction into ammonia, which can act as an alternative \( \text{H}_2 \)-consuming sink competing with ruminal methanogenesis. Further negative \( r_{gCH4} \) were obtained for microbial genes in thiamine metabolism (\( iscS, thiD, thiH, \) and \( thiE \)) with \( r_{gCH4} \) from \(-0.88 \) to \(-0.70\) (\( P_0 \geq 0.91\)), hydation of long-chain fatty acid oleate into anti-tumoral hydroxysteriacid (RA = 0.81, \( P_0 = 0.95\), or import of methanogen inhibitors long-chain fatty acids (RA = 0.87, \( P_0 = 0.99\)). Moreover, highly abundant bacteria genera with ruminal fatty acid biohydrogenation activity (\( ABCB \) (RA = 2.54\%, \( r_{gCH4} = -0.37, P_0 = 0.80 \)) were negatively correlated with CH4.

Microbial communication and host–microbiome interaction mechanisms. The majority of methanogens in the rumen are
integrated into the biofilm on the surface of feed particles where H₂ producing bacteria are active. We found strong negative \( r_{gCH_4} \) (−0.78 to −0.92, \( P_0 \geq 0.96 \)) for abundances of microbial genes mediating microbial interactions, involved in ABC transport of cobalt/nickel (cbiO and cbiQ) and quorum sensing–related peptide/nickel ions (ABC.PE.P, ABC.PE.S, ABC.PE.A, ABC.PE.P1) —cobalt and nickel being detrimental for hydrogenotrophic and acetlastic methanogenic activity\(^{48}\) —protein export (secD and secF) and chemotaxis (cheA and mcp); and positive \( r_{gCH_4} \) for transcription protein cbpA (0.85, \( P_0 = 0.97 \)) acting as a microbial response to maintain plasmid replication during amino acid starvation\(^{115}\). CH₄ emissions were also host-genomically correlated with abundances of microbial genes mediating host–microbe interaction; e.g. cbi and balN\(^{116}\) (\( r_{gCH_4} = −0.80, P_0 \geq 0.96 \)) involved in bacterial biosynthesis of secondary bile acids which activate metabolic receptors within gut, host liver, and peripheral tissues\(^{116,117}\) and inhibit CH₄ production in the ruminant by transferring metabolic H₂ into propionate production\(^{118}\). Another interesting finding is that TSTA3, involved in the metabolism of host–microbiome crosstalk mediator fucose\(^{119}\), displays a positive \( r_{gCH_4} \) (0.85, \( P_0 = 0.98 \)). Fucose is a component of mucins present in saliva\(^{120}\), which is produced abundantly by ruminants and acts as a pH buffer during ruminal fermentation due to its phosphate and bicarbonate content\(^{121}\). Cellulolytic Fibrobacter, an indicator of high pH levels in ruminant\(^{122}\), was positively host-genomically correlated to TSTA3 in our data (0.66, \( P_0 = 0.94 \)), while lactic acid producer Kandleria, generally associated with low pH levels and negative \( r_{gCH_4} \) was host-genomically correlated to TSTA3 negatively (−0.70, \( P_0 = 0.90 \)). Thus TSTA3 could be involved in signalling enhanced saliva production, resulting in increased rumen pH that is known to stimulate the growth of methanogenic archaea and CH₄ emissions\(^{123}\).

**Genetic information processes.** Ribosomal biogenesis represented by RP-S10, RP-S12, RP-S17, RP-L2, RP-L3, RP-L6, RP-L23, RP-L28, RP-L34, and RP-L35, was one of the few microbial mechanisms with positive \( r_{gCH_4} \) from 0.71 to 0.84 (\( P_0 \geq 0.95 \)). All of them are universal ribosomal proteins homologous in bacteria, archaea, and eukarya; except for RP-L28, RP-L34, and RP-L35 exclusively found in bacteria\(^{124,125}\). Given that protein synthesis is highly coupled with cellular growth\(^{69}\), these results suggest that the rumen environment provided by low CH₄-emitter host genomes are related to lower growth or activities of specific microbes directly or indirectly involved in methanogenesis.

RUGs enriched with CH₄-related microbial genes are strongly host-genomically correlated to CH₄ emissions. The 20 highly prevalent (present in >200 animals) RUGs containing the highest number of unique proteins from the 115 microbial genes with strong \( r_{gCH_4} \) were all bacterial RUGs carrying between 114 to 180 unique proteins classified into 60 to 84 microbial genes (Fig. 6 and Supplementary Data 11 and 12). Of these 20 highly enriched bacterial RUGs, 18 showed negative \( r_{gCH_4} \) consistently with the majority of the microbial genes; 6 of them with \( r_{gCH_4} < −0.65 \) (\( P_0 \geq 0.85 \)) from which 5 RUGs were classified as uncultured Lachnospiraceae bacterium (RUG10082, RUG13438, RUG13308, RUG13002, RUG12132) and 1 as uncultured Clostridiales bacterium (RUG10940). The abundance of Blautia and Dorea microbial genera within Lachnospiraceae family (identified by alignment to Hungate1000 collection and Refseq databases) also presented negative \( r_{gCH_4} \) (−0.72 (\( P_0 \geq 0.95 \), Supplementary Data 7)). We also investigated the enrichment of these 115 microbial genes in the 6 RUGs with \( r_{gCH_4} \) (\( P_0 \geq 0.95 \)) annotated at the genus level (Supplementary Data 8), and in those RUGs annotated in the same phylogeny level as any of the 29 microbial genera with \( r_{gCH_4} \) (\( P_0 \geq 0.95 \), Supplementary Data 7), which had low occupancies in our cattle population (<200 animals) and therefore were not included in the 225 considered for breeding (see methods). Our findings show that part of the mechanisms identified in this study occurs in the 5 RUGs classified as uncultured Methanobreibacter strains, each carrying at least 45 out of the 115 microbial genes (Supplementary Fig. 3). The uncultured Methanobreibacter strain with positive \( r_{gCH_4} \) (RUG12982) carried fewer unique proteins (67 vs. 75 to 93) and microbial genes (51 vs. 55 to 62) than the other 4 uncultured Methanobreibacter sp. RUGs with negative \( r_{gCH_4} \); lacking, for example, argD in arginine biosynthesis, tyrA2 in tyrosine and tryptophan metabolism, and DNMTI in methionine metabolism, which reinforces the hypothesis of functional versatility amongst different Methanobreibacter strains explaining their different effects and estimated \( r_{gCH_4} \) on CH₄ emissions. Low-occupancy RUGs annotated as Eubacterium ruminatum, Eubacterium pyrrocarbolicus, Kandleria vitulina, and uncultured sp. of Blautia RUGs carried at least 49 out of the 115 microbial genes each (Supplementary Fig. 3). Interestingly, their counterparts identified at the genera level presented negative \( r_{gCH_4} \) (−0.60 (\( P_0 \geq 0.95 \), Supplementary Data 7)).

**Microbiome-driven breeding of the bovine host for mitigation of CH₄ emissions.** The comprehensive findings of the host-genomic associations between microbial genus/RUG/gene abundances and CH₄ emissions enabled us to predict its mitigation potential when applying genomic selection targeting each of them individually (Supplementary Data 13), indirectly informing about the impact of each microbial mechanism on methanogenesis. Considering 30% of our cattle population being selected based on the abundances of each microbial gene, maiZ in sucrose
metabolism, ABC.PE.P in quorum sensing peptide/nickel transport, hemC in porphyrin or upp in pyrimidine metabolism are predicted to result in the highest CH4 mitigation potential (−5.2, −5.3, −5.8 and −6.54% of CH4 emissions mean respectively, \( P_3 \geq 0.99 \)). Subsequently, our study aimed to find a group of heritable (BF > 3) and Deviance Information Criterion difference between models with or without host genomic effects ≤−20 ruminal microbial genera/RUGs/genes (RA > 0.01%) with strong \( r_{gCH4} \) (\( P_2 \geq 0.95 \)) to be used collectively for selecting the host genomes associated with low CH4 emissions. We identified 2 microbial genera (Eubacterium and Blautia), 3 RUGs (two annotated as uncultured Methanobrevibacter sp. and one as uncultured Prevotellaceae bacterium) and 38 microbial genes meeting these requirements (Supplementary Data 14). We selected 30 out of the 38 microbial genes (Fig. 7a) covering several microbial mechanisms, e.g. sugar and nickel transport (ABC.PE.P, ABC.MS.P1 and ABC.MS.S), sucrose sensing (TST3A), chemotaxis (mcp), ribosomal biosynthesis (RP-L6, RP-L23, RP-L28, RP-L35, RP-S12 and RP-S17), reductive acetogenesis (fhs) and metabolism of amino acids (argD), sucrose (SPP), CH4 (cofG), biotin (bioB), propionate (pcfb), porphyrin (hemC), thiamine (thiD) and pyrimidine (upp). A deep study of the host-genomic correlations among these 30 selected microbial genes showed a common host genomic background influencing the abundance of ABC.PE.P, ABC.MS.P1, fhs, cofG, argD, hemC, thiD, upp, tlyC, NTH, and copB with host-genomic correlations among each other ranging from 0.62 (\( P_3 = 0.90 \)) to 0.99 (\( P_2 = 1.00 \)) (Fig. 7b).

Finally, we evaluated the accuracies and response to selection in CH4 emission mitigation in our population based on the prediction of CH4 host genomic effects using three different sources of information: (1) CH4 emissions measured by the “gold-standard” technique of respiration chambers, (2) the 30 microbial gene abundances exhibiting strong \( r_{gCH4} \), and (3) combining both preceding criteria. A single (1) or multiple (2, 3) trait genomic estimation approach was applied in each case. In (2) and (3), CH4 host genomic effects were estimated based on observations of the 30 microbial gene abundances, the genomic relationship matrix amongst individuals, and the estimated host-genomic and residual (co)variance matrix comprising CH4 and the 30 microbial gene abundances; assuming unknown (2) or known (3) CH4 observations (see methods). Using microbiome-driven breeding based on the abundance of 30 specific microbial genes (2) resulted in the mean estimation accuracy of host genomic effects for CH4 emissions to be 34% higher than using measured CH4 emissions (1) (0.70 ± 0.18 vs. 0.52 ± 0.11) and confirmed that functional microbial genes are an extremely valuable source of information to perform host genomic evaluations for CH4 emissions. Using the combined selection criteria (3), the accuracy of estimation was 14% larger than using rumen microbial gene information alone (0.80 ± 0.20). Response to selection in CH4 emissions achieved by selecting animals with low CH4 emission host genomic effects predicted exclusively by microbial gene abundance information resulted in a reduction in emissions of −1.43 ± 0.14 to −3.32 ± 0.77 g CH4/kg DMI per generation, depending on selection intensity (from 1.16 to 2.67 in the analysed population, Fig. 8 and Supplementary Data 16). These results indicate that in our population, microbiome-driven breeding for CH4 emissions reduced its magnitude by 7–17% of its mean per generation, without the necessity for costly measures of CH4 emissions.

Robustness of estimation of genomic parameters of CH4 emissions and microbiome traits from a cross-classified design of breeds and basal diets. The data, from the highly environmentally standardized experiments, comprised of animals from different breeds that were offered different diets, which could be challenging as different breeds might have different genomic backgrounds for the analysed traits, and different diets could have inflated their variances. To consider the difference in means of these effects we fitted a combination of experiment, breed, and diet effects so that an adjustment of each of these effects and their interactions was achieved. To analyse whether after this adjustment the genomic variances of CH4 emissions and the abundances of 30 microbial genes selected for microbiome-driven breeding are homogeneous across breeds and diets, we computed their posterior distributions separately (Supplementary Figs 4 and 5), following the partition of the variance suggested by Sorensen et al.128, and recently used by Lara et al.127. Using this methodology, we found that our genomic parameter estimates for CH4 emissions are based on similar genomic variances across breeds.
(with medians of 3.8, 3.7, 4.0, and 3.9 (g/kg DMI)² for Aberdeen Angus, Limousin, Charolais crosses and pure breed Luing, respectively) and diets (with medians of 3.8 and 3.9 (g/kg DMI)² for Forage and Concentrate based diets) with almost entirely overlapping distributions indicating their homogeneity (Supplementary Fig. 4). We also identified homogeneous genomic variances for the abundances of microbial genes with overlapping distributions across breeds (Supplementary Fig. 5a) and diets (Supplementary Fig. 5b); for example, we estimated genomic variances with medians between 0.024–0.028 (g/kg DMI)² across breeds and 0.028–0.029 (g/kg DMI)² across diets for RP-L35, or between 0.25–0.30 (g/kg DMI)² across breeds and 0.292–0.296 (g/kg DMI)² across diets for cofG (Supplementary Fig. 5). These results indicate that the data recorded under controlled experimental conditions with a cross-classified breed and diet two-way experimental design and progeny groups balanced over diets resulted in reliable and robust genomic parameters estimates.

Discussion

Previous metagenomic studies using 16S rRNA identified microbiota revealed host genomic effects on the rumen microbial community, e.g., Wallace et al. found significant heritabilities for several members of Prevotella and Butyribivrio genera. These species, together with other non-heritable OTUs from the core microbiota, explained up to 40% of the phenotypic variation of CH₄ emissions that is expected to be the result of both genetic and environmental correlations, which we estimated in the present study separately for each taxon and microbial gene. Moreover, the challenge of previous studies is the lack of a comprehensive resolution to sufficiently explain these associations, which we resolved using metagenome-assembled RUGs. For example, we revealed that different Methanobrevibacter strains expressed divergent host-genomic correlations with CH₄ emissions which correspond to the diverse microbial gene content of these strains.

To increase the predictability of CH₄ emissions, Roche et al. suggested genome-resolved metagenomics and identified 20 microbial gene abundances, mainly involved in the methanogenesis pathway, which explained 81% of the phenotypic variation in the emissions. In contrast, the present study uses a host genomic-microbiome analysis strategy and provides a robust and comprehensive insight into joint host-genomic correlations between rumen microbial genes known to affect complex functional mechanisms and CH₄ emissions which further enabled us to predict the expected response of selection in the emissions based on each and a combination of the microbial genes. The findings of this research will be of major importance for the mitigation of the highly potent GHG CH₄ in bovine through genomic selection on the functional microbiome associated with CH₄ referred to as microbiome-driven breeding. The highlights of our research are that the host genome influences CH₄ emissions by favoring the growth of reductive acetogenic microbes limiting the excess of metabolic H₂ substrate (specifically, Blautia, Eubacterium genera and microbial gene fhs found in the genome of uncultured Lachnospiraceae bacterium, Eubacterium pyr- uvatovorans, Eubacterium ruminatum, uncultured Eubacterium and Blautia sp. RUGs); and promoting the shift in the fermentation towards volatile fatty acids (Kandleria genera and microbial genes bed, pccB, fucO, carried by uncultured Lachnospiraceae bacterium and Kandleria vitulina RUGs) and microbial proteins yield including arginine and branched-chain amino acids (argF, argD, ilvA, AROA2), tryptophan, tyrosine and phenylalanine (trpA, trpD, trpE, tyrA2, paaH) and methionine (metE or DNMT1), which are expected to lead to animals with improved efficiency of converting feed into nutrients. Moreover, host genome contributes to lower CH₄ emissions by enhancing the growth of microbes that consume H₂ in alternative pathways (e.g. nitrogen fixation (nifU)); by promoting the pathways that divert specific substrates (e.g. tyrosine and 2-oxocarboxylic acid catalobolites) required to produce methanogenic coenzymes or cofactors (coenzyme B and methanofuran) to other routes; and that inhibit methanogenic organisms (e.g. by the presence of branched-chain amino acids or cobalt/nickel (cbiO, cbiQ)) and maintain a lower optimum ruminal pH (sucroase metabolism (sacA, maIZ, bgLB, SPP, sucrose phosphorylase)) preventing gut disorders (e.g. thiamine metabolism (iscS, thiD, thiH, thiE)). The latter result supports our hypothesis that hosts who are genetically resilient to gut disorders produce less CH₄, which is compatible with nutritional studies demonstrating that blocking methanogenesis has no undesirable effects on cattle health status or feed intake. A further highlight of our study is that the host genome influenced the ABC transport of different metabolites (some of them in quorum sensing processes).
TSTA3, interspecies electron transfers (bioB, cobL, cofG), sensitivity of environmental conditions (cheA, mcp), and host–microbiome interaction mechanisms (ehb, baiN, TSTA3), all host-genomically associated with CH4 emissions. These results shed light on the complex processes of methanogenesis regulated by different microbial mechanisms where communication between microbial communities and their interactions with the host plays an essential role. Genetic information processes in the microbiota (e.g. ribosomal biosynthesis (e.g. RP-S10)) also had a substantial host genomic effect on CH4 emissions, potentially reflecting different microbial community growth profiles.

Our findings on the functional microbial level are complementary to studies investigating the biological mechanisms underlying host genome influence on the colonization and maintenance of specific ruminal microbial groups, such as host genomic effects on rumen size96, muscle contraction associated with passage rate19, or ruminal pH20. Other studies in bovines have elucidated host candidate genes for CH4 emissions involved in similar mechanisms97,98,129,130, fitting into our demonstrated hypothesis that the host genome commonly influences rumen microbiome profile and CH4 emissions. From nutritional studies68, it is well known that the rumen pH has an overarching effect on the rumen microbial community and its metabolism. The rumen pH is intimately related to the production level of buffer-acting saliva98 that is rich in fucose120. We found that the abundance of TSTA3, encoding the sensor for host–microbiome crosstalk mediator fucose119, was host-genomically positively correlated to CH4 emissions and to Fibrobacter genera, an indicator for high pH122, making TSTA3 a highly valuable biomarker for rumen pH, CH4 metabolism and potentially, for host–microbiome mediation to enhance saliva production.

Our results provide comprehensive insight into which communities and functions of the rumen microbiome can be modified by genomic selection to obtain low CH4-emitter animals. We revealed that specific microbiome functionalities (i.e. microbial gene abundances) are more informative for breeding purposes than specific taxonomies, as indicated by a higher number of microbial genes than genera/RUGs being host-genomically correlated to CH4 emissions. This could be due to the closely defined function of those genes, e.g. being involved in producing specific substrates or mediating a specific pathway that interferes with CH4 metabolism; while each microbial genera expresses many substrates or mediating a specific function of those genes, e.g. being involved in producing specific substrates or mediating a specific pathway that interferes with CH4 metabolism.

### Methane emissions data.

Methane emissions were individually measured in 285 of the 363 animals for 48 h within six indirect open-circuit respiration chambers39. One week before entering the respiration chambers, the animals were housed individually in training pens, identical in size and shape to the pens inside the chambers. The animals were allocated to the respiration chambers in a randomized design within breed and diet. Methane emissions were expressed as g of CH4/kg of DMI, by dividing the average CH4 emissions (g/day) by the average DMI (kg/day) recorded both after 48 h.

**Hosts genomic samples.**

For host DNA analysis, 6–10 ml of blood from the 363 steers were collected from the jugular or coccygeal vein in live animals or during slaughter in a commercial abattoir. In addition to the 363 samples, 7 blood and 23 semen samples from sires of the steers were available (n = 393 samples in total). Blood was stored in tubes containing 1.8 mg EDTA/ml blood and immediately frozen to −20 °C. Genomic DNA was isolated from blood samples using the Qiagen QIAamp tissue kit and from semen samples using a Qiagen QIAamp DNA Mini Kit, according to the manufacturer’s instructions. The DNA concentration and integrity were estimated with Nanodrop ND-1000 (NanoDrop Technologies). Genotyping was performed by Neogen Genomics (Ayr, Scotland, UK) using GeneSeek Genomic Profiler (GGP) BovineSNP50k Chip (GeneSeek, Lincoln, NE). Genotypes were filtered for quality control purposes using PLINK version 1.09b1. Single Nucleotide Polymorphisms were removed from further analysis if they met any of the following criteria: no known genomic location, low call rate (<0.95), a single nucleotide polymorphism. The genotypes were then converted to Illumina’s ‘Autosomal call maps’ and analyzed using the ‘linkage’ command of PLINK version 1.09b1. The data were then filtered for minor allele frequency (MAF) >0.05. Seven animals, showing genotypes with a call rate <90%, were removed so that 386 animals and 36,780 autosomal SNPs remained for the analyses.

**Hosts metagenomic samples.**

For microbial DNA analysis, post-mortem digesta samples (approximately 50 ml) from 363 steers were taken at slaughter immediately after the rumen was opened to be emptied. Five ml of strained ruminal fluid was mixed with 10 ml of PBS containing glycerol (87%) and stored at −20 °C. DNA extraction from rumen samples was carried out following the protocol from Yu and Morrison113 based on repeated bead beating with column filtration and DNA concentrations and integrity was evaluated by the same procedure (Nanodrop ND-1000) as for blood samples. Four animals out of 363 did not yield rumen samples of sufficient quality for metagenomics analysis. DNA Illumina TrueSeq libraries were
prepared from genomic DNA and sequenced on Illumina HiSeq systems 2500 (samples from 4 animals from the experiment year 2011), HiSeq systems 4000 (samples from 282 animals from experiment years 2012, 2013 and 2014) or NovaSeq (samples from 76 animals from the experiment year 2017) by Edin-
burgh Genomics (Edinburgh, Scotland, UK). Paired-end reads (2 × 100 bp for HiSeq systems 2500 and 2 × 150 bp for HiSeq systems 4000 and NovaSeq) were generated, resulting in between 7.8 and 47.8 GB per sample (between 26 and 159 million paired reads).

Bioinformatics. For phylogenetic annotation of rumen samples, the sequence reads of 359 samples were aligned to a database including cultured genomes from the Hungate 1000 collection40 and RefSeq genomes41 using Kraken software42. From 1178 cultured microbial genera identified, we used only those present in all the samples with a RA > 0.001% (1108 microbial genera) for downstream analysis, equivalent to 99.99% of the total number of counts. We used the 4941 RUGs generated by Stewart et al.34 with sequences of 282 rumen samples included in this study to identify and quantify the abundance of uncultured species. A detailed description of the metagenomics assembly and binning process and estimation of the depth of each RUG in each sample is described in Stewart et al.34. For breeding purposes, microbial taxa that are present in a large proportion of the animals are required; so we discarded those RUGs present in <200 animals (using a cut-off of 1× coverage) and kept 225 RUGs. RUGs cov- erages <1, which comprised 17.7% of the whole RUGs data set were imputed based on a Geometrical Bayesian multiplicative model (gbm) of replacement by using emulrplftm method in XCompositions package43. This algorithm imputes zero values from a posterior estimate of the multinomial probability assuming a Dirichlet prior distribution with default parameters for GRM method144 and performs a multiplicative read-
justment of non-zero components to respect original proportions in the compo-
sition. The 225-RUGs considered showed a mean relative abundance ±0.15%.
Bioinformatic analysis for the identification of rumen microbial genes was carried out as previously described by Wallace et al.45. Briefly, to measure the abundance of known functional microbial genera whole metagenome sequencing reads were aligned to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (https://www.genome.jp/kegg/coh.html)46. KO (also referred to as microbial genes), common in all animals. As for microbial genera, we used only core microbial genes present in all the samples and with a RA > 0.001% (1142 microbial genes) for downstream analysis, equivalent to 96.25% of the total number of counts. We combined information from KEGG, UniProt and Clusters of Orthologous Groups of protein databases to classify 1141 microbial genes into classes depending on the biological processes they are involved in CH4 metabolism (34), metabolism other than CH4 pathway (511), genetic information processes (329), microbial communication and host-microbiome interaction (207) and other unknown or at present poorly characterized (61).

Statistics and reproducibility

Log-ratio transformation of metagenomic data. To describe the composition of the microbiome at the taxonomic level (cultured microbial genera and RUGs) and functional level (KO or microbial genes) we estimated their RA by dividing each microbial genus/gene in the database selected as a reference. We selected the 16S rRNA gene and **Orb**acterium as reference microbial gene and microbial genus, respectively. These reference variables were selected based on the criteria recom-
med by Greenacre et al.151: (1) present in rumen samples of 359 animals; (2) highly abundant (mean RA 8.56% and 0.35%, respectively); (3) not mentioned to be associated with CH4 emissions in previous literature; (4) low log-ratio variance so the variation mainly proceeds to the numerator (0.09 and 0.24, both located in the first quartile when ordering the microbial variables by log-ratio variance in decreasing order) and (5) reproducing the geometry of the full set of log-ratios in the original data set shown by the estimate of the Procrustes correlation148,152.

The 225 RUGs considered showed a mean relative abundance ±0.15%. To describe the composition of the whole microbial community, we estimated their RA by dividing each microbial genus/gene in the database selected as a reference. We selected the 16S rRNA gene and **Orb**acterium as reference microbial gene and microbial genus, respectively. These reference variables were selected based on the criteria recom-
med by Greenacre et al.151: (1) present in rumen samples of 359 animals; (2) highly abundant (mean RA 8.56% and 0.35%, respectively); (3) not mentioned to be associated with CH4 emissions in previous literature; (4) low log-ratio variance so the variation mainly proceeds to the numerator (0.09 and 0.24, both located in the first quartile when ordering the microbial variables by log-ratio variance in decreasing order) and (5) reproducing the geometry of the full set of log-ratios in the original data set shown by the estimate of the Procrustes correlation148,152.

between the geometrical space defined by all log-ratios and the one defined by the selected additive log-ratios (Procrustes correlation is 0.95 and 0.92). **Orb**acterium is a strictly anaerobic and non-spore-forming bacterial genus from the order Clostridiales and family of Lachnospiraceae; commonly found in the rumen of cattle19,23 and also in the human oral cavity145,150. The abundance of RUGs was centred log-ratio transformed149 as additive log-ratio transformation was here hampered by the difficulty of selecting a reference RUG present in all animals. Assuming f denotes the total number of RUGs (f = 225):

\[
\ln \left( \frac{x_j}{x_{ref}} \right) = \ln(x_j) - \ln(x_{ref}) = 1, \ldots, f, j \neq ref
\]

where \(x_j\) is the RA of each microbial genus/gene in the database selected as a reference. We selected the 16S rRNA gene and **Orb**acterium as reference microbial gene and microbial genus, respectively. These reference variables were selected based on the criteria recom-
med by Greenacre et al.151: (1) present in rumen samples of 359 animals; (2) highly abundant (mean RA 8.56% and 0.35%, respectively); (3) not mentioned to be associated with CH4 emissions in previous literature; (4) low log-ratio variance so the variation mainly proceeds to the numerator (0.09 and 0.24, both located in the first quartile when ordering the microbial variables by log-ratio variance in decreasing order) and (5) reproducing the geometry of the full set of log-ratios in the original data set shown by the estimate of the Procrustes correlation148,152.
and residuals as:

\[ R = e|R_e| \sim N(0, I \otimes |R_e|), \]

where \( G \) and \( R \) are the 2 × 2 host genomic and residual (co)variance matrices between CH₄ emissions and each microbial genus, RUG, or microbial gene, i is an identity matrix of the same order as the number of individuals with data. Bayesian statistics were used\(^{158}\), assuming priors for all unknowns as implemented in THRGIBBSF90 program\(^{159}\). Results were based on Markov chain Monte Carlo chains consisting of 1,000,000 iterations, with a burn-in period of 200,000, and to reduce autocorrelation, the 1 out of every 100 was saved for inference. In all analyses, convergence was tested using the POSTGIBBSF90\(^{159}\) program by calculating the Z criterion of Geweke (varying between −0.05 and 0.05 in univariate and −0.09 and 0.1 in bivariate models). Monte Carlo sampling errors were computed using time-series procedures and checked to be at least 10 times lower than the standard deviation of the marginal posterior distribution. As \( h_j^2 \) estimates, we used the median of its marginal posterior distribution of CH₄, each microbial genus, RUG, or microbial gene, and the highest posterior density interval at 95% probability (HPD\(_{0.95}\)). We considered microbial abundances with \( h_j^2 \) estimates < 0.20 being lowly heritable, 0.20 < \( h_j^2 \) ≤ 0.40 being moderately heritable, and \( h_j^2 \) estimates > 0.40 being highly heritable. To test the significance of host genomic effects we analysed the fitness of the full univariate genomic model vs. the univariate model without host genomic effect by comparing the deviance information criterion (DIC)\(^{160}\) between models and computed the BF using an univariate model without host genomic effect by comparing the deviance. Genomic effects we analysed the software applies Markov Clustering algorithm by a response in CH₄ emissions after host genomic selection for each of these microbial traits using the frequentist approach using AIREMLF90\(^{159}\) and similar results were performed\(^{161}\). For the HPD\(_{0.95}\)%, we considered microbial abundances with \(|R_{CH4}|\) ≤ 0.05 and \( |R_{CH4}| \) ≥ 0.05 in data from 285 animals with CH₄ emissions and metagenomics information. All scenarios were calculated by GBLUP analysis assuming as fixed (tolerance for minimum eigenvalues = 0.001). The difference between original and bent matrices was never higher than the posterior standard error of the corresponding parameters.

Estimation of the selection response of CH₄ emissions based on different sources of information. We analysed three different scenarios to predict host-genomic effects of CH₄ emissions: (1) by using measured CH₄ emissions only, (2) by using the 30 microbial gene abundances only, and (3) by using a combination of both, measured CH₄ emissions and the 30 microbial gene abundances. The three scenarios were computed with data from 285 animals with CH₄ emissions and metagenomics information. All scenarios were calculated by GBLUP analysis assuming as fixed variance components the previously estimated 31 × 31 host genomic and residual variance-covariance matrices of the traits after bending. Scenario (1) was performed using a univariate GBLUP analysis including only measured CH₄ emissions; scenario (2) was computed by fitting a multivariate GBLUP model including the 30 microbial gene abundances host-genomically correlated to CH₄ emissions (using measured CH₄ emissions as missing value\(^{162}\); and scenario (3) considered both microbial gene abundances, the previously estimated 31 × 31 host genomic and residual variance-covariance matrices.

Co-abundance network analysis of host genomic effects on the rumen microbiome. To study the correlation structure among host genomic effects of the log-transformed abundances of 1107 microbial genera, 225 RUGs, and 1141 microbial genes, we built a co-abundance network analysis using deregressed host genomic effects (dGEBVs) for all microbial traits in the 359 samples. Deregressed host genomic effects were calculated from previously described univariate GBLUP models by using ACCF90 and DEPROOF90 programs\(^{159}\). Co-abundance network (Graphia software\(^{163}\)) connected or edged microbial traits (nodes) based on a Pearson correlation ≥ 0.70 among their dGEBVs. The complexity of the graph was reduced by discarding nodes with a minimum number of incident edges (referred to as node degree) of 2, i.e. only those microbial traits Pearson-correlated (≥ 0.70) with at least 2 other microbial traits were kept. The total number of microbial genera, RUGs, and microbial genes included in the network was 2129 out of the 2473 tested. The number of edges of each node was reduced by ranking the edges based on 4-nearest neighbour algorithm and retaining only 80% of them. The software applies Markov Clustering algorithm by a flow simulation model\(^{164}\) to find discrete groups of nodes (clusters) based on their position within the overall topology of the graph. The granularity of the clusters, i.e. the minimum number of nodes that a cluster has to contain, was set to 2 nodes. The network showed 106 clusters, but only those 12 clusters including ≥ 3 methanogenic archaea genera, RUGs and microbial genes involved in CH₄ metabolism pathway according to KEGG\(^{165}\) database or microbial genera/RUGs/gene set-host-genomically correlated with CH₄ emissions (\( P_0 \geq 0.95 \)) were studied in depth.

Data availability

Metagenomic sequence reads for all rumen samples are available under European Nucleotide Archive (ENA) under accession projects PRJEB31266, PRJEB21624, and PRJEB10338. The genotypes of the host animals are readily available from the authors.
54. McInerney, M. J. et al. Physiology, ecology, phylogeny, and genomics of microorganisms capable of syntrophic metabolism. *Ann. NY Acad. Sci. 1125*, 58–59 (2008).
55. Evans, P. N. et al. An evolving view of methane metabolism in the Archaea. *Nat. Rev. Microbiol.* 17, 219–232 (2019).
56. Nomura, M., Gourse, R. & Baughman, G. Regulation of the synthesis of ribosomes and ribosomal components. *Ann. Rev. Biochem.* 53, 75–117 (1984).
57. Liu, Y. & Whitman, W. B. Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. *Ann. NY Acad. Sci.* 1125, 171–189 (2008).
58. McInerney, M. J. et al. Physiology, ecology, phylogeny, and genomics of microorganisms capable of syntrophic metabolism. *Ann. NY Acad. Sci. 1125*, 58–59 (2008).
59. Kamke, J. et al. Rumen metagenome and metatranscriptome analyses of low methane yield sheep reveals a Sharpea-enriched microbiome characterised by lactic acid formation and utilisation. *Microbiome* 4, 1–16 (2016).
60. Yanibada, B. et al. Inhibition of enteric methanogenesis in dairy cows induces changes in plasma metabolites highlighting metabolic shifts and potential markers of emission. *Sci. Rep.* 10, 1–14 (2020).
61. Groopy, J. P. et al. Low-methane yield sheep have smaller rumens and shorter rumen retention time. *Br. J. Nutr.* 111, 578–585 (2014).
62. Steward, C. S., Flint, H. J. & Bryant, M. P. in *The Rumen Microbial Ecosystem* (eds Hobson, P. N. & Steward, C. S.) 10–72 (Blackie Academic and Professional, 1997).
63. Kellermann, S. et al. Two different bacterial community types are linked with the low-methane emission trait in sheep. *PLoS ONE* 9, 1–9 (2014).
64. Strobol, H. J. & Russell, J. B. Effect of pH and energy spilling on bacterial protein synthesis by carbohydrate-limited cultures of mixed rumen bacteria. *J. Dairy Sci.* 69, 2941–2947 (1986).
65. Herrmann, K. M. & Weaver, L. M. The shikimate pathway. *Annu. Rev. Plant Biol.* 50, 473–503 (1999).
66. Hall, M. B. & Herejak, C. Differences in yields of microbial crude protein from in vitro fermentation of carbohydrates. *J. Dairy Sci.* 84, 2486–2493 (2001).
67. Volop, L. & Verstraete, W. Gastro-enteric methane versus sulphate and hydrogen. *Nature* 363, 465–466 (1993).
68. Demeyer, D., De Graeve, K., Durand, M. & Stevani, J. Acetate: a hydrogen sink in hindgut fermentation as opposed to rumen fermentation. *Acta Vet. Scand Suppl.* 86, 68–75 (1989).
69. Lopez, M. Cintos, F. M., Wallace, R. J. & Newbold, C. J. Effect of adding aceticogenic bacteria on methane production by mixed rumen microorganisms. *Annu. Rev. Food Sci. Technol.* 10, 1–22 (2019).
70. Ovens, F. N., Secrist, D. S., Hill, W. J. & Gill, D. R. Acidosis in cattle: a review. *J. Anim. Sci.* 76, 275–286 (1998).
71. Doyle, N. et al. Use of lactic acid bacteria to reduce methane production in ruminants, a critical review. *Front. Microbiol.* 10, 2207 (2019).
72. Kruger Ben Shabat, S. et al. Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants. *ISME J.* 10, 2958–2972 (2016).
73. Ugalvova, N. B., Sacanell, C. J. & Jarrett, J. T. Spectroscopic changes during a single turnover of biontin synthase: destruction of a [2Fe-2S] cluster accompanies sulfur insertion. *Biochemistry* 40, 8352–8358 (2001).
74. Howells, D. M., Harich, K., Xu, H. & White, R. J. K. Acid chain elongation reactions involved in the biosynthesis of coenzyme B (7-mercaptoheptanoyl threonine phosphate) in methanogenic archaea. *Biochemical J.* 379, 10108–10117 (2004).
75. Widdel, F. Growth of methanogenic bacteria in pure culture with 2-propanol and other alcohols as hydrogen donors. *Appl. Environ. Microbiol.* 51, 1056–1062 (1986).
76. Moore, S. J. et al. Elucidation of the biosynthesis of the methane catalyst coenzyme F430. *Nature* 543, 78–82 (2017).
77. Bulen, W. A. & LeComte, J. R. The nitrosenase system from Azotobacter: two-enzyme requirement for N2 reduction, ATP-dependent H2 evolution, and ATP hydrolysis. *Proc. Natl Acad. Sci. USA* 56, 979–986 (1966).
78. Wang, M., Wang, H., Zheng, H., Dewhurst, R. J. & Roche, R. A heat diffusion multilayer network approach for the identification of functional biomarkers in rumen methane emissions. *Methods* https://doi.org/10.1016/j.jymeth.2020.09.014 (2020).
79. Jenkins, T. C., Abughaizaleh, A. A., Freeman, S. & Thies, E. J. The production of ARG genes: the effects of diet and supplements. *Ann. Vet. Anim. Sci.* 84, 43–131 (1999).
80. Herrmann, K. M. & Weaver, L. M. The shikimate pathway. *Annu. Rev. Plant Biol.* 50, 473–503 (1999).
81. Hall, M. B. & Herejak, C. Differences in yields of microbial crude protein from in vitro fermentation of carbohydrates. *J. Dairy Sci.* 84, 2486–2493 (2001).
82. Volop, L. & Verstraete, W. Gastro-enteric methane versus sulphate and hydrogen. *Nature* 363, 465–466 (1993).
83. Demeyer, D., De Graeve, K., Durand, M. & Stevani, J. Acetate: a hydrogen sink in hindgut fermentation as opposed to rumen fermentation. *Acta Vet. Scand Suppl.* 86, 68–75 (1989).
84. Baldwin, L. R., Wood, W. A. & Emery, R. S. Conversion of lactate-4-c to propionate by the rumen microflora. *J. Bacteriol.* 83, 907–913 (1961).
85. Janssen, P. H. Influence of hydrogen on rumen methane formation and fermentation balances through microbial growth kinetics and fermentation thermodynamics. *Annu. Rev. Food Sci. Technol.* 160, 1–22 (2010).
86. Ovens, F. N., Secrist, D. S., Hill, W. J. & Gill, D. R. Acidosis in cattle: a review. *J. Anim. Sci.* 76, 275–286 (1998).
87. Doyle, N. et al. Use of lactic acid bacteria to reduce methane production in ruminants, a critical review. *Front. Microbiol.* 10, 2207 (2019).
88. Kruger Ben Shabat, S. et al. Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants. *ISME J.* 10, 2958–2972 (2016).
89. Ugalvova, N. B., Sacanell, C. J. & Jarrett, J. T. Spectroscopic changes during a single turnover of biontin synthase: destruction of a [2Fe-2S] cluster accompanies sulfur insertion. *Biochemistry* 40, 8352–8358 (2001).
90. Howells, D. M., Harich, K., Xu, H. & White, R. J. K. Acid chain elongation reactions involved in the biosynthesis of coenzyme B (7-mercaptoheptanoyl threonine phosphate) in methanogenic archaea. *Biochemical J.* 379, 10108–10117 (2004).
91. Widdel, F. Growth of methanogenic bacteria in pure culture with 2-propanol and other alcohols as hydrogen donors. *Appl. Environ. Microbiol.* 51, 1056–1062 (1986).
92. Moore, S. J. et al. Elucidation of the biosynthesis of the methane catalyst coenzyme F430. *Nature* 543, 78–82 (2017).
93. Bulen, W. A. & LeComte, J. R. The nitrosenase system from Azotobacter: two-enzyme requirement for N2 reduction, ATP-dependent H2 evolution, and ATP hydrolysis. *Proc. Natl Acad. Sci. USA* 56, 979–986 (1966).
94. Wang, M., Wang, H., Zheng, H., Dewhurst, R. J. & Roche, R. A heat diffusion multilayer network approach for the identification of functional biomarkers in rumen methane emissions. *Methods* https://doi.org/10.1016/j.jymeth.2020.09.014 (2020).
95. Jenkins, T. C., Abughaizaleh, A. A., Freeman, S. & Thies, E. J. The production of ARG genes: the effects of diet and supplements. *Ann. Vet. Anim. Sci.* 84, 43–131 (1999).
96. Herrmann, K. M. & Weaver, L. M. The shikimate pathway. *Annu. Rev. Plant Biol.* 50, 473–503 (1999).
112. Leng, R. A. Interactions between microbial consortia in biofilms: a paradigm shift in rumen microbial ecology and enteric methane mitigation. *Anim. Prod. Sci.* 54, 519–543 (2014).

113. Won, M. Y., Oyama, L. B., Courtney, S. J., Creevey, C. J. & Huws, S. A. Can rumen bacteria communicate to each other? *Microbiome* 8, 1–8 (2020).

114. Patra, A., Park, T., Kim, M. & Yu, Z. Rumen methanogens and mitigation of methane emission by anti-methanogenic compounds and substances. *J. Anim. Sci. Biotechnol.* 8, 1 (2017).

115. Wahlström, A., Sayin, S. I., Marschall, H. U. & Bäckhed, F. Intestinal crosstalk: a new approach to host-microbial interactions. *Microbiome* 15, 1–12 (2014).

116. Lara, L. A. d. C., Pocrnic, I., Gaynor, R. C. & Gorjanc, G. Temporal and variances in the course of artificial selection. *Heredity* 77, 287–294 (2001).

117. Hess, M. K. et al. A restriction enzyme reduced representation sequencing two contrasting diets on blood methaemoglobin and performance of two finishing beef steers. *Anim. Sci. Rep.* 30, 1–9 (2011).

118. Wächter, A., Sayam, S. I., Marschall, H. U. & Bäckhed, F. Intestinal crosstalk: a new approach to host-microbial interactions. *Microbiome* 15, 1–12 (2014).

119. Poehlein, A., Schneider, D., Soh, M., Daniel, R. & Seedorf, H. Comparative genomic analysis of members of the genera *methanosphaera* and *methanobrevibacter* reveals distinct clades with specific potential metabolic functions. *Arch. Microbiol.* 198, 541–550 (2016).

120. Hoorens, P. R. et al. Genome wide analysis of the bovine mucin genes and their expression in response to diet and parabiotic association. *BMC Genomics* 16, 839 (2015).

121. Maekawa, M., Beauchemin, K. A. & Christensen, D. A. Effect of concentrate level and feeding management on chewing activities, saliva production, and ruminal pH of lactating dairy cows. *J. Dairy Sci.* 91, 1448–1463 (2008).

122. Lee, M., Jeong, S., Seo, J. & Seo, S. Changes in the ruminal fermentation and membrane vesicles in cellulolytic *Ruminococcus* spp. of ruminal origin. *Appl. Environ. Microbiol.* 76, 1351–1358 (2010).

123. Kellems, R. E., Miller, D. H. & Somarriba, M. A. The statistical analysis of quantitative data in microbiology. *Microbiol. Mol. Biol. Rev.* 58, 349–383 (1994).

124. Lecompte, O., Ripp, R., Thierry, J. C., Moras, D. & Poch, O. Comparative annotation of *Oribacterium* sp. strain C9, isolated from a cattle rumen. *PLoS ONE* 13, e0201143 (2018).

125. Lee, M., Jeong, S., Seo, J. & Seo, S. Changes in the ruminal fermentation and bacterial community structure by a sudden change to a high-concentrate diet in finishing beef steers. *Front. Microbiol.* 8, 2224 (2017).

126. Maekawa, M., Beauchemin, K. A. & Christensen, D. A. Effect of concentrate level and feeding management on chewing activities, saliva production, and ruminal pH of lactating dairy cows. *J. Dairy Sci.* 91, 1448–1463 (2008).

127. Lara, L. A. d. C., Pocrnic, I., Gaynor, R. C. & Gorjanc, G. Temporal and variances in the course of artificial selection. *Heredity* 77, 287–294 (2001).

128. Greenacre, M. Variable selection in compositional data analysis using pairwise logratios. *Math. Geosci.* 51, 649–682 (2018).

129. Greenacre, M. *Compositional Data Analysis in Practice* (CRC Press, 2019).

130. Iwasawa, K. et al. Dysbiosis of the salivary microbiota in pediatric-onset primary sclerosing cholangitis and its potential as a biomarker. *Sci. Rep.* 8, 1–10 (2018).

131. Oksanen, J. et al. vegan: Community Ecology Package. *R* package version 2.5-7. (2021).

132. Zhang, X. & Pusztai, A. Genome wide analysis of the bovine mucin genes and their expression in response to diet and parabiotic association. *BMC Genomics* 16, 839 (2015).

133. Greenacre, M. Comparative data analysis. *Annu. Rev. Stat. Appl.* 8, 271–299 (2021).

134. Bach, J. C. & Bäckhed, F. Intestinal crosstalk: a new approach to host-microbial interactions. *Microbiome* 15, 1–12 (2014).

135. Blasco, A. *Bayesian Data Analysis for Animal Scientists: The Basics*. https://doi.org/10.1007/978-3-319-54274-4 (2017).

136. Misztal, I. et al. Manual for BLUPF90 Family of Programs (Univ. Georgia Athens, 2018).

137. Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & Van Der Linde, A. *Bayesian Data Analysis* (Chapman & Hall/CRC, 2008).

138. Newton, M. A. & Raftery, Adrian E. Approximate Bayesian Inference with the Bayesian-frequentist approach. *J. Anim. Breed. Genet.* 120, 583–616 (2002).

139. van Doorn, J. et al. The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychon. Bull. Rev.* https://doi.org/10.3758/s13423-020-01798-5 (2020).

140. Freeman, T. C. & Mackay, T. F. C. *Introduction to Quantitative Genetics* (Pearson, 1981).

141. Freeman, T. C. Graphia: a platform for the graph-based visualisation and analysis of complex networks from microarray expression data. *PLoS Comput. Biol.* 3, 2032–2042 (2007).

142. Schneeberger, M., Barbwick, S. A., Crow, G. H. & Hammond, K. Economic indices using breeding values predicted by BLUP. *J. Anim. Breed. Genet.* 109, 180–187 (1992).

Acknowledgements

The authors thank Professor Ignacy Misztal and Dr. Shogo Tsuruta for making software available to us, Professor Agustín Blasco and Professor Chris Haley for their statistic advice, and Professor Michael Greenacre for his advice on compositional data analysis. We also thank Bin Zhao for his contribution to the identification and biological description of metagenomics data and Dr. Larissa Zetouni for her comments on the manuscript.
Author contributions
M.M.-A., R.R., and M.W. conceived and designed the overall study, and M.M-A., M.W.,
and R.R. conceived, designed, and executed the bioinformatics analysis. M.D.A., C.-A.D.,
R.J.D., and M.A.C. provided essential insight into microbiology, rumen metabolism,
nutrition, methane emissions and animal breeding. M.M.-A. and R.R. wrote the initial
draft, and subsequently, all authors contributed intellectually to the interpretation and
presentation of the results in the manuscript, which was edited and approved by all
authors.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material
available at https://doi.org/10.1038/s42003-022-03293-0.

Correspondence and requests for materials should be addressed to Rainer Roehe.

Peer review information Communications Biology thanks Grum Gebreyesus and the other
anonymous reviewer for their contribution to the peer review of this work. Primary handling
editors: Anna Heintz-Buschart and Caitlin Karniski. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in
published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons
Attribution 4.0 International License, which permits use, sharing,
adaptation, distribution and reproduction in any medium or format, as long as you give
appropriate credit to the original author(s) and the source, provide a link to the Creative
Commons license, and indicate if changes were made. The images or other third party
material in this article are included in the article’s Creative Commons license, unless
indicated otherwise in a credit line to the material. If material is not included in the
article’s Creative Commons license and your intended use is not permitted by statutory
regulation or exceeds the permitted use, you will need to obtain permission directly from
the copyright holder. To view a copy of this license, visit http://creativecommons.org/
licenses/by/4.0/.

© The Author(s) 2022