Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range.
Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range

Gemma Henderson¹, Faith Cox¹, Siva Ganesh¹, Arjan Jonker¹, Wayne Young¹, Global Rumen Census Collaborators† & Peter H. Janssen¹

Ruminant livestock are important sources of human food and global greenhouse gas emissions. Feed degradation and methane formation by ruminants rely on metabolic interactions between rumen microbes and affect ruminant productivity. Rumen and camelid foregut microbial community composition was determined in 742 samples from 32 animal species and 35 countries, to estimate if this was influenced by diet, host species, or geography. Similar bacteria and archaea dominated in nearly all samples, while protozoal communities were more variable. The dominant bacteria are poorly characterised, but the methanogenic archaea are better known and highly conserved across the world. This universality and limited diversity could make it possible to mitigate methane emissions by developing strategies that target the few dominant methanogens. Differences in microbial community compositions were predominantly attributable to diet, with the host being less influential. There were few strong co-occurrence patterns between microbes, suggesting that major metabolic interactions are non-selective rather than specific.

Ruminants are one of the most successful groups of herbivorous mammals on the planet, with around 200 species represented by approximately 75 million wild and 3.5 billion domesticated individuals worldwide¹. Ruminants are defined by their mode of plant digestion, and have evolved a forestomach, the rumen, that allows partial microbial digestion of feed before it enters the true stomach. Ruminants themselves do not produce the enzymes needed to degrade most complex plant polysaccharides, and the rumen provides an environment for a rich and dense consortium of anaerobic microbes that fulfil this metabolic role. These rumen microbes ferment feed to form volatile fatty acids that are major nutrient sources for the host animal and contribute significantly to ruminant productivity. The host also uses microbial biomass and some unfermented feed components once these exit the rumen to the remainder of the digestive tract. Ruminants have evolved various rumen anatomies and behaviours to thrive on a range of plant species, and this flexibility has enabled them to occupy many different habitats spanning a wide range of climates². These were also important factors in their domestication, allowing conversion of human-indigestible plant material into readily-accessible animal goods, especially dairy products, meat, and useful fibres. Ruminants have thus played a vital role in sustaining and developing many human cultures, as well as being used as draft animals and having religious and status values.

¹AgResearch Limited, Grasslands Research Centre, Palmerston North 4442, New Zealand. †A comprehensive list of authors and affiliations appear at the end of the paper. Correspondence and requests for materials should be addressed to G.H. (email: gemma.henderson@agresearch.co.nz) or P.H.J. (email: peter.janssen@agresearch.co.nz)
Rumen microbes can be assigned to different functional groups, such as cellulolytics, amylolytics, proteolytics, etc., which degrade the wide variety of feed components or further metabolize some of the products formed by other microbes. For example, methanogens, the methane-forming archaea, are among those that metabolize hydrogen formed by some fermentative microbes to form methane. The methane generated during this fermentation contributes to global anthropogenic greenhouse gas emissions and represents a 2–12% loss of feed energy for the animal. Differences in rumen microbial communities underlie variations in methane formation and the conversion of feed to animal products. Therefore, understanding these communities is key to understanding ruminal transformations of plant material to both undesirable and useful ruminant products.

The aim of this study was to determine the composition of the microbiota in rumen and foregut samples from 742 individual animals from around the world. The resulting dataset allowed us to determine that dietary factors dominate over host species in determining microbial community composition, identify the dominant microbes and their potential associations, and describe the degree of similarity of rumen microbial communities worldwide.

Results and Discussion

This is the largest single study to examine microbial communities across a range of ruminant and camelid species, diets, and geographical regions. A standardised pipeline was used to process samples in order to minimise variation introduced by processing steps such as DNA extraction or PCR amplification. This is important for detecting authentic patterns rather than ones introduced by methodological differences between different studies. The primers chosen amplify, to the best of our current knowledge, the target gene regions from nearly all known bacteria, archaea, and rumen ciliates.

Dominant rumen microbes. Despite the range of ruminants with different feeding strategies and diets, similar rumen bacteria were abundant around the world (Fig. 1). There was some variation in bacterial community compositions in animals from different regions, likely to be caused by differences in diet, climate, and farming practices. The 30 most abundant bacterial groups (Greengenes taxonomy summarised at the genus-level) were all found in over 90% of samples, and together comprised 89.4% of all sequence data (Supplementary Table 1) and were similar to those described in an earlier meta-analysis of rumen microbial communities. All 30 are known rumen-inhabiting bacteria. Because the samples came from a wide range of ruminant species, diets, and geographical locations, these data suggest that new dominant bacteria are not likely to be found in future studies. The seven most abundant bacterial
groups comprised 67.1% of all bacterial sequence data, were detected in all samples (Supplementary Fig. 1), and can be considered the “dominant” rumen bacteria. They were Prevotella, Butyrivibrio, and Ruminococcus, as well as unclassified Lachnospiraceae, Ruminococcaceae, Bacteroidales, and Clostridiales. These might be considered a “core bacterial microbiome” at the genus level or higher, because they are present in a large selection of ruminants, so confirming the suggestion that there is a core rumen microbiome. However, these bacterial groups were not equally abundant in all animal species (P ≤ 0.005; Supplementary Table 2). With the exception of Butyrivibrio, these groups are not adequately represented by characterised cultures, and their functions are not well understood.

Inspection of the most abundant and prevalent bacterial operational taxonomic units (OTUs) in the dataset showed that only 14% fell within a named species, and 70% were not even within a formally recognised genus (Fig. 1). This study shows that, while we appear to recognize the dominant rumen bacteria, considerable microbiological effort is still required to understand them. Some efforts have been made to isolate more cultures and gather more information about these bacteria, for example, the genomes of Prevotella aff. ruminicola Tc2-24, rumen bacterium R-7, and other isolates whose 16S rRNA gene sequences are similar to those of dominant rumen bacterial OTUs (Fig. 2b), have been sequenced as part of the Hungate1000 project.

Because there is a flux of both liquids and solids through the rumen, microbes must actively metabolize to gain energy and multiply to counteract washout and so maintain populations in the rumen. The dominant bacteria found in this study are therefore likely to be responsible for the majority of the

Figure 2. Dominant bacterial and archaeal operational taxonomic units (OTUs). Similarities (Supplementary Tables 8 and 9) of the 50 most abundant and 50 most prevalent bacterial (77 unique OTUs, (a,b) and archaeal (64 unique OTUs, (c,d) OTUs to the most closely related type (a,c) and cultured (b,d) strains are plotted together with prevalence and abundance data. Background shading indicates nominal within-species (dark grey), within-genus (mid grey) and below genus (light grey) similarities. Prevalence indicates the percentage of samples that an OTU occurs in. The size of each circle indicates the mean abundance of each OTU (Supplementary Tables 8 and 9). Bacterial OTU abundances were multiplied by a factor of 15 relative to archaeal OTUs. Mbb. Methanobrevibacter.
transformation of ingested feed in the rumen and cameldid foregut, especially of cellulose, hemicellulose, pectin, starch, fructan, organic acids, and protein, as these are the major energy-yielding substrates used for microbial growth. There is also a convergence of bacterial community structure in the rumen and in the crop of the hoatzin, a bird that relies on a foregut fermentation of ingested leaves. Thus microbial community structure seems to be driven by the similarity of organ function extending across the rumen, the cameldid foregut, and the crop of this unusual bird. More efforts should go into characterizing the metabolism and roles of these bacteria that are the responsible for the majority of feed fermentation, with the aim of enhancing animal productivity and reducing methane emissions.

Nearly all the archaea were identified as methanogens known to be residents of the rumen (Supplementary Table 3, Supplementary Text 1), and their relative abundances were comparable to previous studies. The dominant archaeal groups were remarkably similar in all regions of the world (Fig. 1). This universality and limited diversity was also recently noted in survey of archaea in New Zealand ruminants and could make it possible to successfully mitigate methane emissions by developing strategies, such as vaccines or small-molecule inhibitors, that target the few dominant methanogens.

Members of the Methanobrevibacter gottschalkii and Methanobrevibacter ruminantium clades were found in almost all samples, and were the two largest groups, accounting for 74% of all archaea. Together with a Methanospirillum sp. and two Methanomassiliicoccaceae-affiliated groups, the five dominant methanogen groups comprised 89.2% of the archaeal communities (Supplementary Fig. 1), showing that rumen archaea are much less diverse than rumen bacteria. This likely reflects the narrow range of substrates they use. Methanomicrobiurn has previously been reported as abundant in ruminants in Asia. In our study, they were found to comprise >5% of the archaeal community of some Australian, Brazilian, Chinese, North American, and South African cattle, as well as South African sheep, showing them to be widely distributed, but not universally prevalent. The five dominant methanogen groups were not equally abundant in all animal species groups (P < 0.005; Supplemental Table 4). In contrast to bacteria, the rumen archaea are better represented by cultures, with 58% of the most abundant and prevalent OTUs falling within a named species, and all but 22% within named genera (Fig. 2c). All of the latter were members of Methanomasstillicoccales, which is an order of relatively poorly-characterised methanogens for which representative cultures of as-yet unnamed species and genera are available (Fig. 2d). The 50 most abundant OTUs accounted for 74.5% of the archaeal sequence data, again indicating a much lower diversity than in the bacteria, where the 50 most abundant OTUs made up only 11.0%.

By assigning physiologies (Supplementary Table 5) to the sequence abundance information (Supplementary Table 3), it can be concluded that 77.7% of archaea were hydrogenotrophic methanogens, while 22.1% had the ability to grow with hydrogen plus methyl groups derived from methanol or methylamines. Methanogens able to form methane from acetate (Methanosarcina spp. and Methanosaeta spp.) were extremely rare (<0.015%; Supplemental Data 1), as expected based on their general slow growth rates that would not allow them to be maintained in the rumen under normal conditions.

Almost all protozoal sequence data (>99.9%) were assigned to 12 genus-equivalent protozoal groups (Supplementary Table 6). It was apparent that the variability of protozoa between and within cohorts of co-located animals was much greater than that of bacteria and archaea (Supplementary Fig. 2). It has been reported that there is strong host individuality of rumen protozoal community structure, and this is evident in our study. The genera Entodinium and Epidinium dominated, occurring in more than 90% of samples and representing 54.7% of protozoal sequence data (Supplementary Fig. 1). Many of the protozoal genera were present in greater than 70% of the samples, indicating a wide prevalence. Genera such as Enoploplastron and Ophryoscolex had a wider than expected host distribution. They are considered to be mainly present in sheep and cattle, respectively, but we also found Enoploplastron in cattle, deer, and reindeer samples from twelve countries, and Ophryoscolex in buffalo, goats, deer, sheep, and giraffe samples from 18 countries. Although different rumen protozoa are reported to have limited host and geographical distributions, host specificity has been questioned. It seems likely that further investigation will demonstrate greater ubiquity of the rumen protozoa.

Effects of diet and host on microbial community composition. Because the abundance of microbial groups varied between animal species groups and cohorts (Supplementary Tables 2 and 4; Supplementary Fig. 2), we looked for factors that might underlie this. Rumen and cameldid foregut microbial community structure could be expected to be shaped by morphological, physiological, and even behavioural characteristics that evolved along with the varied feeding strategies in the various ruminal lineages. Indeed, adaptation has resulted in a diversity of rumen sizes and passage rates of rumen contents, allowing ruminant species to exploit a range of feed types. In addition to feed composition effects, these host adaptations might also play a role in regulating rumen microbial community structure. Because our dataset was from ruminants and cameldids from different lineages consuming a range of diets, host and diet effects on rumen microbial community structure could be separated.

To look at diet and host effects, we classified the diets based on forage and browse or concentrate content (Supplementary Table 7) and grouped the animals according to their lineage (Supplementary Data 1). Microbial communities could clearly be discriminated by both host and diet (Fig. 3a), with bacteria being the main drivers behind the observed differences (Fig. 3b). This probably reflects their more diverse metabolic capabilities compared with the less versatile archaea and protozoa. We investigated the patterns of microbial abundances across hosts and diets (Fig. 3c, Supplementary Fig. 3–6).
Figure 3. Effect of host species and dietary forage to concentrate ratios on microbial communities. Diets were grouped (Supplementary Table 7) as forage-dominated (F), mixed forage-concentrate (50–70% forage, FC), mixed concentrate-forage (50–70% concentrate, CF), or concentrate-dominated (C).

(a) Discriminant analysis of microbial communities in samples (represented by points coloured by animal and diet) revealed that both host and diet determined community composition. (b) Bi-plot that shows microbial groups (identified by colours) underlying the separation of samples in panel (a). Several bacterial groups strongly discriminate the samples by host and diet, indicated by their presence towards the outside of the bi-plot. Archaeal and protozoal groups are less discriminatory, and so are clustered nearer the centre. (c) The heatmap shows that bacterial abundances are differentially associated with diet and host (colour key shows the association score; see Supplementary Figs 3–5 for additional data). (d) Unclassified Veillonellaceae, and (e) Fibrobacter are examples of bacteria that caused bovines and caprids to cluster separately from other species in the heat map. The number of samples in each category is given in parentheses in panels (c–e). *indicates unclassified bacteria within an order or family.
**Ruminococcus**, one of the dominant bacteria, was relatively evenly distributed, but this was an exception. For many bacteria, diet was the major factor determining relative abundance. Bacterial communities from forage-fed animals were similar to each other, those from concentrate-fed animals were similar to each other, but distinct from those in forage-fed animals, and those from animals fed mixed diets were intermediate between these. Unclassified Bacteroidales and Ruminococcaceae were more abundant in all animals fed forages. Some-as-yet poorly characterised Bacteroidales are postulated to be able to degrade cellulose, and their genomes encode a broad range of plant polysaccharide degrading capabilities, which could explain their pattern of distribution. In contrast, members of Prevotella and unclassified Saccinivibrionaceae were more abundant in animals fed diets containing concentrate. Based on the physiological capacities of cultured relatives, these are probably major producers of propionate and the propionate-precursor succinate, and so are responsible for the greater levels of propionate formed from concentrate-rich diets. The abundance of only a few other major bacterial groups was associated with host lineage (Fig. 3c). For example, unclassified Veillonellaceae were proportionally more abundant in sheep, deer, and camelids (Fig. 3d). This may be related to differences in rumen and cameldid forestgut sizes, anatomy, and feeding frequencies compared to bovines.

The relative abundances of several major bacterial groups were affected by both host and diet (Fig. 3c). Unclassified Clostridiales were most abundant in bovines fed forage and least abundant in bovines fed high concentrate diets, while in caprines, cervids, and camelids these diet differences were far less pronounced. Butyribrio was most abundant in rumen samples from bovines fed mixes of forage and concentrates. Fibrobacter was most abundant in bovines fed forage. When concentrate was included in bovine diets, the relative abundance of Fibrobacter was decreased, but it was still more abundant than in other animals. To examine its distribution in more detail, we compared Fibrobacter abundances across different rumenomas and found significantly higher levels in bovines compared to deer, sheep, or camelids (Fig. 3e). These data suggest that Fibrobacter is favoured in the bovine rumen and, given that it is cellulose degrader, may play an essential role in the degradation of plant fibre in cattle.

Overall, diet was a major determinant of bacterial community structure. This may be because physical and chemical characteristics of the feed determine the different microbial niches available. In contrast to the post-gastric mammalian digestive tract, and due to the sheer volume of digesta and feed input, there is probably less shaping of the rumen microbial community by local host biological factors such as the immune system, secreted antimicrobial peptides, host-cell glycosylation, and host-derived nutrients.

**Associations between rumen microbes.** The abundance patterns within bacterial, archaeal, and protozoal communities in different hosts fed different diets showed that certain microbes exhibited parallel patterns of relative abundance (Fig. 3c, Supplementary Figs 2–6). We therefore looked for correlations within and between bacteria, archaea, and protozoa (Fig. 4 and Supplementary Fig. 7), reasoning that specific associations should be seen across diets, hosts, and geography. Negative correlations of abundances of groups were observed within the bacteria, archaea, and protozoa, including replacement effects between dominant groups within each of these (Supplementary Text 1 and Supplementary Fig. 7). Few strong positive correlations were found within bacteria, archaea, and protozoa. For example, there was a strong correlation between Veillonellaceae and the TG5 group, driven by their co-occurrence within cervids and caprines. These microbes may cooperate in the rumen, or they may share similar requirements and so certain hosts and diets would offer better opportunities for their growth. This explanation could also underlie the strong positive correlations observed between different groups of methylotrophic methanogens (Supplementary Fig. 7). They may be responding to diets rich in methyl groups, such as feeds with high levels of pectins or osmolesyts such as betaine. The strongest correlation within protozoa was a positive one between Dasytricha and Isotricha. These two genera of holotrichous protozoa display very similar spectra of substrate use, including use of plant soluble sugars and storage carbohydrates, again suggesting that co-occurrence may be due to exploitation of similar opportunities.

We also investigated associations between bacteria, archaea, and protozoa. Strikingly, no strong correlations were detected between archaea and protozoa (Supplementary Fig. 7). Methanogens are known to colonize protozoa, and this mutualistic relationship is believed to enhance methane formation in the rumen. The occurrence of specific symbioses between methanogens and rumen protozoa has been speculated on, but not convincingly demonstrated. The lack of strong co-occurrence patterns within this study indicates that these undoubtedly important associations are probably non-specific, or occur at a strain level. Further investigation is required to corroborate this interesting finding, as mechanisms that mediate the colonization of protozoa by archaea remain to be elucidated. These could have interesting evolutionary aspects if they allow non-specific interactions to form or are mediated by strain-specific mechanisms that confer different partner specificities within archaeal or protozoal species. In contrast, there were some positive associations between bacterial and protozoal groups. Most noticeable were the associations of Isotricha and Dasytricha with Fibrobacter. Fibrobacter were reported to decrease in abundance in animals where protozoa were eliminated, indicating that there may be a mutually beneficial relationship between these protozoa and Fibrobacter, which are surface colonizers of plant material.

No strong associations were found between the most abundant bacteria and archaea (Fig. 4). This was surprising, since rumen bacteria degrade feed and produce the substrates that methanogens use for growth, mainly hydrogen and methyl groups. In contrast, there were distinct positive associations between some less abundant bacteria and archaea. The strongest association was between bacteria such as the
succinate-producing *Succinivibrio* spp., the succinate-using *Dialister*, and the amino-acid-fermenting *Acidaminococcus*, and methanogens belonging to the *Methanomassiliicoccaceae*, *Methanosphaera* sp. A4, and *Methanobrevibacter boviskoreani*. *Succinivibrio* spp. degrade pectin, and methanol is required for growth of *Methanomassiliicoccaceae* and *Methanosphaera*, explaining part of this pattern. Other associations were between the methylo trophic methanogen *Methanosphaera* sp. ISO3-F5 and different bacteria, including members of *Lachnospiraceae*. These associations may be based on the ability of *Lachnospiraceae* to degrade pectin and so provide methanol as a substrate for the methylotrophs.

The associations between other *Methanomassiliicoccaceae* groups and various unclassified members of *Bacteroidales* suggest the possibility of yet further methanol-dependent metabolic interactions. In contrast to archaeal-protozoal interactions, these findings suggest that some archaeal-bacterial interactions are specific, inferring specialised mechanisms for partner recognition or very similar requirements for growth. The basis for these associations remains to be determined. However, the general lack of strong association patterns between protozoa and the major bacteria on the one hand, and the major methanogen groups on the other, suggests that conserved mechanisms may mediate the interactions between hydrogen producing and hydrogen consuming microbes, allowing flexible interactions. This may aid methane mitigation research, since interfering with these potentially universal mechanisms could slow the rate of hydrogen transfer and so slow methane formation. It may also be that the interactions mainly occur via pools of common metabolites, especially where the end products of one group form the substrates of another.

The results of this survey showed that the rumen microbial ecosystem is dominated by a core community composed of poorly-characterised microbes, especially amongst the bacteria. Diet had more influence than animal species on rumen or camelid foregut microbial community composition. Rumen ecosystems are typified by strong metabolic interactions between microbes that facilitate the fermentation of plant material to products useful for both the host and other rumen microbes. The relatively few co-occurrence patterns seen in this study suggest that these microbial interactions do not rely on exclusive associations, and could indicate considerable promiscuity between members of interacting functional groups. Analysis at metagenomic and metatranscriptomic levels could in future uncover whether common functional elements that facilitate interactions are shared among multiple species. It seems plausible that functional redundancy among the microbes means that multiple microbial species can fulfil the same function, with different combinations of microbes being co-selected depending on

---

Figure 4. Associations between bacteria and archaea. The network is based on association scores computed via regularised canonical correlation analysis with an absolute association score greater than 0.15. The colour of the lines indicates the strength of the association. The sizes of the diamonds and circles indicate the mean average abundance and microbial groups are identified by numbers (Supplementary Tables 1 and 3). *Mbb. Methanobrevibacter, Mmc. Methanomassiliicoccales,* *indicates unclassified bacteria within a family.*
the diet. This flexibility of rumen microbial community structure would confer on the ruminant host the ability to exploit a variety of different plant feeds.

Methods

Geographical distribution and diversity of gastrointestinal tract content samples. A total of 742 samples from 32 species or sub-species of ruminants and other foregut fermenters in 35 countries and seven global regions were selected for sequencing of microbial marker genes (Fig. 1, Supplementary Data 1). The samples were from cattle, bison, and buffalo (bovines), sheep and goats (caprids), deer (cervids), and alpacas, llamas, and guanacos (camelids), including diverse breeds of domestic cattle, sheep, and goats, and were largely made up of small cohorts of four or more co-located individuals consuming the same diet. We included foregut samples of camelids in this study, recognizing that these organs have a common function but evolved separately\(^{39}\). The use of animals, including welfare, husbandry, experimental procedures, and the collection of samples used for this study, was, where applicable, approved by named institutional and/or licensing committees and performed in accordance with approved institutional and regulatory guidelines (please refer to Supplementary Data 1 for details of these).

Sample collection, DNA extraction, amplification and processing of samples for high-throughput sequencing. To minimise variation introduced by differing methodologies, such as choice of sampling or DNA extraction method\(^{40}\) and primer-driven gene amplification biases\(^{41}\), we used a standardised pipeline to process samples (unless indicated otherwise in Supplementary Data 1). Briefly, approximately 20 g of whole (i.e., solid and liquid) mid-rumen or camelfid foregut contents were collected via stomach tube, cannula, or post mortem as previously described\(^{33}\). Samples were immediately frozen, freeze-dried, and then couriered to AgResearch. Freeze-dried samples were homogenised in a coffee-blender and DNA was extracted from a representative 30 mg subsample using the PCQI method\(^{40,42}\). We assessed the structure of microbial communities by sequencing regions of bacterial and archaeal 16S rRNA genes and ciliate protozoal 18S rRNA genes in triplicate as described previously\(^{35,37}\) using primers comprised of (5’ to 3’) a sequencing adapter (A or B), a sample-unique 12-base error-correcting Golay barcode on one of each primer pair, a two-base linker, and a group-specific sequence targeting the marker gene. For bacteria, the primers were Ba515Rmod1 (adapter A-barcode GT-CCCGGCGKCTGCGAC) and Ba9F (adapter B-AC-GAGTTTGTGCTGCGGCAGGAC). For archaea, the primers were Ar915aE (adapter A-barcode GT-AGGAATTGGCGGGGGAGCAC) and Ar9136R (adapter B-CA- GCCGTTGTGCTGCGGAGGACC). For protozoa, the primers were Reg1320R (adapter A-barcode TC-AATTGGCAAGAGTCTATCCCC) and RP841F (adapter B-AA-GACTAGGGATTGARGTG) (Supplementary Table 1). Linker A was CCATCTCATCCCTGCGTGTCTCCGACTCAG and linker B was CCAATCCCCTGTGCTGCGTCTGCGACGATCAG. Amplicons were sequenced using 454 GS FLX Titanium chemistry at Eurofins MWG Operon (Ebersberg, Germany). Sample processing and pipeline reproducibility controls were performed to identify variation introduced during sample processing (Supplementary Text 1). Sequence data are available from GenBank [accession numbers PRJNA272135, PRJNA272136, and PRJNA273417].

Phylogenetic analysis of sequencing data. Pyrosequence data were processed and analysed using the QIIME software package version 1.8\(^{43}\). Sequences over 400 bp in length with an average quality score over 25 were assigned to a specific sample via the barcodes. The number of bacterial, archaeal, and ciliate protozoal sequencing reads available for analysis are summarised in Supplementary Data 1. Sequence data were grouped into operational taxonomic units (OTUs) sharing over 97% (bacteria – UCLUST\(^{44}\), 99% (archaea - UCLUST) or 100% (ciliate protozoa – prefix_suffix option in QIIME) sequence similarity. Sequences were assigned to phylogenetic groups by BLAST\(^{45}\). Bacterial 16S RNA genes were assigned using the Greengenes database version 13_5\(^{10}\), archaeal 16S rRNA genes using RIM-DB version 13_11_13\(^{22}\) and ciliate protozoal 18S rRNA genes against an in-house database\(^{46}\). Bacterial and ciliate protozoal data were summarised at the genus level. Archaea were summarised at the species level. Samples for which low read numbers were obtained or that contained high proportions of sequences from “exogenous” bacteria (i.e., likely environmental contaminants such as *Stenotrophomonas*) were excluded from further analyses (Supplementary Text 1).

The identity of the most abundant and prevalent OTUs was determined using BLAST\(^{46}\) against sequences from type material and against all sequences (excluding sequences from model organisms or environmental samples) in the nt database\(^{47}\), Bellerophon (version 3, 200 bp window, Huber-Hugenholtz correction\(^{48}\)) was used to identify chimeric OTU sequences. Sequence similarities greater than 97% and 93% were used as cut-offs to classify OTUs at species- and genus level, respectively. The rationale for these cut-offs was discussed by Kenter\(s\) et al.\(^{49}\).

Simplified classification of dietary information and other factors. The range of diets consumed by the animals from which the samples came was highly diverse and complex. For this reason, and where the information was available, diets were categorised in terms of forage type, forage plant, and forage to concentrate ratio (Supplementary Table 7). Diets likely to contain >5% starch (e.g., whole or grain crops of maize, barley, wheat, rice, as well as pea, potato, sorghum, etc.) or >5% pectin (e.g., beets or legumes such as alfalfa and clover) were also identified. Animals that had been fed their respective diets
for less than a two-week period were noted in Supplementary Data 1. Factors such as gender, age, modifications (e.g., cannulation), treatments (e.g., antibiotics, drench, surgery), farming conditions, season, contact with other animals, and sample processing steps that may affect apparent microbial community compositions (e.g., DNA extraction method, sample fraction used, sample storage, etc.) were also recorded (Supplementary Data 1). Where details were not provided, latitude, longitude, and elevation were estimated using http://www.mapcoordinates.net/en. Climate zones were designated according to the Köppen-Geiger climate classification scheme33.

**Statistical analyses.** The resulting dataset allowed us to establish whether animal or dietary factors relate to rumen and cameldid foregut microbial community composition, identify the dominant microbes and their potential associations, and describe the degree of similarity of rumen and cameldid foregut microbial communities worldwide. Statistical analyses of microbial data were performed using GenStat for Windows31, R software32, and QIIME33. Principal coordinate analysis of Bray-Curtis dissimilarity matrices, analysis of variance, sparse partial least squares discriminant analysis (sPLS-DA, using a sPLS regression approach), and canonical discriminant analyses (CDA) of microbial community composition data in context of the metadata (Supplementary Data 1) were used to identify impacts of factors such as host lineage, diet, etc. on rumen and cameldid foregut microbial communities and to identify the groups associated with these factors. Pearson, Spearman, SparCC33, and regularised canonical correlation analyses (CCA) were used to identify associations within and between archaeal, bacterial, and protozoal groups. Association scores were visualised as relevance networks and clustered image maps (CIM, heatmaps) representing the first two dimensions. González et al. provides a comprehensive overview of sPLS-DA, CCA and the corresponding ‘pairwise associations’, network and CIM techniques and their application34.

**References**


**Acknowledgements**

We thank Ron Ronimus, Paul Newton, and Christina Moon for reading and commenting on the manuscript. We thank all who provided assistance that allowed Global Rumen Census collaborators to supply samples and metadata (Supplemental Text 1). AgResearch was funded by the New Zealand Government as part of its support for the Global Research Alliance on Agricultural Greenhouse Gases. The following funding sources allowed Global Rumen Census collaborators to supply samples and metadata, listed with the primary contact(s) for each funding source: Agencia Nacional de Investigación e Innovación, Martín Fraga; Alberta Livestock and Meat Agency, Canada, Tim A. McAllister; Area de Ciencia y Técnica, Universidad Juan A Maza (Resolución Proy. N° 508/2012), Diego Javier Grilli; Canada British Columbia Ranching Task Force Funding Initiative, John Church; CNPq, Hilário Cuquetto de Ciencia y Técnica, Universidad Juan A Maza; Department of Agriculture, Fisheries and Food, Alexandre B. De Menezes; Meat & Livestock Australia; and Department of Agriculture and Forestry (Australian Government), Chris McSweeney; Ministerio de Agricultura y desarrollo sostenible (Colombia), Olga Lucía Mayorga; Montana Agricultural Experiment Station project (MONB00113), Carl Yeoman; Multistate project W-3177 Enhancing the competitiveness of US beef (MONB00195), Carl Yeoman; NSW Stud Merino Breeders’ Association, Alexandre Vieira Chaves; Queensland Enteric Methane Hub, Diane Ouerkerk; RuminOmics, Jan Kopecny, Ilma Tapio; Rural and Environment Science and Analytical Services Division (RESAS) of the Scottish Government and the Technology Strategy Board, UK, R. John Wallace; Science Foundation Ireland (09/RF/GEN2447), Sinead Waters; Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación, Mario A. Cobos-Peralta; Slovenian Research Agency (project number J1-6732 and P4-0097), Blaz Stres; Strategic Priority Research
Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range.

How to cite this article
et al. Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. Sci. Rep. 5, 14567; doi: 10.1038/srep14567 (2015).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

Consortia

Global Rumen Census Collaborators


2Estacion Experimental del Zaidin, Consejo Superior de Investigaciones Científicas, 18100 Armilla, Granada, Spain.
3Animal Nutrition, Biotechnology and Bioindustry Center, Corporación Colombiana de Investigaciones Agropecuarias, Bogota, Colombia.
4Facultad de Ciencias Veterinarias, Universidad de Concepción, Chillan, Chile.
5Área Microbiología, Departamento de Patología, Facultad de Ciencias Médicas, Universidad Nacional de Cuyo, Centro Universitario, Parque General San Martín S/N, 5500 Mendoza, Argentina.
6University of Cundinamarca, Cundinamarca, Colombia.
8Departamento de Producción Animal, National University of Colombia, AA 2779, Medellín, Colombia.
9Copenhagen Zoo, DK 2000, Frederiksberg, Denmark.
10Instituto de Investigaciones Agropecuarias, INIA Remehue, Osorno, Región de Los Lagos, Chile.
11Department of Animal Science, University of Vermont, Burlington, Vermont 05405, USA.
12Faculty of Veterinary Science, The University of Sydney, New South Wales 2006, Australia.
13Department of Biological Sciences, Thompson Rivers University, Kamloops, B.C., V2C 0C8, Canada.
14Microbiology, School of Biology and Environmental Science, Science Centre West, University College Dublin, Belfield, Dublin 4, Ireland.
15Colegio de Postgraduados, Institución de Enseñanza e Investigación en Ciencias Agrícolas, CP 56230, Montecillo, Mexico.
16Instituto de Biotecnología, Instituto Nacional de Tecnología Agropecuaria, Los Reseros y Repetto, 1686 Hurlingham, Argentina.
18Rowett Institute of Nutrition and Health, University of Aberdeen, Bucksburn, Aberdeen, AB21 9SB, Scotland.
19Cargill Animal Nutrition Innovation Center, Elk River, Minnesota 55330, USA.
20Instituto de Patobiología, Instituto Nacional de Tecnología Agropecuaria, Los Reseros y Repetto, 1686 Hurlingham, Argentina.
21Departamento de Microbiología, Universidad Federal de Viçosa, Campus UFV, 36570-000 Viçosa, Minas Gerais, Brazil.
22Animal Nutrition Group, Wageningen University, 6700 AH Wageningen, The Netherlands.
23Agriculture & Agri-Food Canada, Lacombe, Alberta T4L 1W1, Canada.
24Department of Animal Sciences, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran.
25Faculty of Land and Animal Resources, Agricultural University of Iceland, Keldnaholt, IS-112 Reykjavik, Iceland.
26Department of Agriculture, University of Zululand, KwaDlangezwa, South Africa.
27Department of Microbiología, Instituto de Investigaciones Biológicas Clemente Estable, Av. Italia 3338, CP 11600, Montevideo, Uruguay.
28VITA Marangani, Universidad Nacional Mayor de San Marcos, Lima, Peru.
29 Ministry for Primary Industries Verification Services Hawkes Bay, Silver Fern Farms—Pacific, Whakatu, Hastings, New Zealand.
30Laboratory of Animal Nutrition, Research Faculty of Agriculture, Hokkaido University, N9W9, Kita-ku, Sapporo 060-8589, Japan.
31Institute of Agricultural Sciences, Animal Nutrition, ETH Zürich, CH-8092 Zürich, Switzerland.
33Quality Control Department, Mashhad Feed Mill, Mashhad, Iran.
34Department of Animal Nutrition, Federal University of Agriculture, Abeokuta (FUNAAB), Nigeria.
35Department of Ruminant Sciences, Agricultural Research Center (ARO), Volcani Institute, Bet Dagan 50250, Israel.
36Investigación en Ciencias Animales, Instituto de Investigaciones Agropecuarias Kampenaike, Angamos 1056, Punta Arenas, Región de Magallanes, Chile.
37Natural Resources Institute Finland, FI-31600 Jokioinen, Finland.
38Department of Biology, University of Eastern Finland, FI-70211 Kuopio, Finland.
39Department of Animal Nutrition, Biotechnology and Bioindustry Center, Corporación Colombiana de Investigaciones Agropecuarias, Bogota, Colombia.
Science & Technology, Sunchon National University, Suncheon, Jeonnam 540–743, Korea.  
Animal Studies Building, Gatton Campus and Ecosciences Precinct, The University of Queensland, Dutton Park, Queensland, Australia.  
Institute of Animal Physiology and Genetics, Laboratory of Anaerobic Microbiology, Academy of Sciences of the Czech Republic, Videnska 1083, 142 20, Prague 4, Czech Republic.  
Nordic Genetic Resource Center, NO-1431 As, Norway.  
Swedish University of Agricultural Sciences, Department of Agricultural Research for Northern Sweden, SE-901 83 Umeå, Sweden.  
Department of Animal and Range Sciences, Montana State University, Bozeman, Montana, 59717, USA.  
Animal Science Genetics, University of Missouri-Columbia, Columbia, Missouri 65211, USA.  
Department of Molecular Biology and Genetics, Aarhus University, DK-8830 Tjele, Denmark.  
Buffalo Research Institute, The Chinese Academy of Agricultural Science, Nanning, Guangxi, China.  
Department of Animal Science, Aarhus University, DK-8830 Tjele, Denmark.  
Institut National de la Recherche Agronomique, UMR1213 Herbivores, F-63122 Saint-Genes-Champanelle, France.  
Clermont Université, VetAgro Sup, UMR1213 Herbivores, F-63000 Clermont-Ferrand, France.  
Université de Lyon, VetAgro Sup, UMR1213 Herbivores, F-69280 Marcy l’Etoile, France.  
Lethbridge Research Centre, Agriculture & Agri-Food Canada, Lethbridge, Alberta, Canada T1J 4B1.  
CSIRO Agriculture, Queensland Bioscience Precinct, St. Lucia, Queensland 4067, Australia.  
DairyNZ, Cnr Ruakura and Morrinsville Roads, Newstead, Hamilton 3240, New Zealand.  
NARO Institute of Livestock and Grassland Science, 2 Ikenodai, Tsukuba, Ibaraki, 305-0091 Japan.  
Agroscope, Institute for Livestock Sciences ILS, CH-1725 Posieux, Switzerland.  
University of Ljubljana, Biotechnical Faculty, Department of Animal Science, SI-1230 Domzale, Slovenia.  
Research and Innovation Center, Diamond V, Cedar Rapids, Iowa 52404-5260, USA.  
Forestry Research Institute of Nigeria, Jericho, Ibadan, Oyo State, Nigeria.  
Agri-Science Queensland, Department of Agriculture and Fisheries (DAF), Ecosciences Precinct Dutton Park, Queensland 4012, Australia.  
Embrapa Gado de Leite, Rua Eugenio do Nascimento, EP 36038-330 Juiz de Fora (MG), Brazil.  
Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, As, Norway.  
University of Hohenheim, Institute of Animal Science, Animal Nutrition Group, 70599 Stuttgart, Germany.  
Department of Technology, National Livestock Breeding Center 1, Shibahara, Mafune, Nishigou, Nishishirakawa, Fukushima 961-8511, Japan.  
Natural Resources Institute Finland, FI-31600 Jokioinen, Finland.  
Yamagata Prefectural Animal Industrial Institute, Agricultural Research Center, 1076 Torigoe, Shinjo, Yamagata 996-0041, Japan.  
Department of Biology and Ecology, Faculty of Science and Mathematics, 18000 Niš, Serbia.  
Institute of Subtropical Agriculture, The Chinese Academy of Sciences, Changsha, China.  
CSIRO Animal Food and Health Sciences, James Cook University, Townsville, Queensland, Australia.  
Instituto de Investigaciones Agropecuarias INIA Carillanca, Temuco, Chile.  
Animal and Bioscience Research Department, Animal & Grassland Research and Innovation Centre, Teagasc, Grange, Dunsany, Co. Meath, Ireland.  
Agri-Food and Biosciences Institute, Hillsborough, County Down BT26 6DR, Northern Ireland.  
Present addresses: †AgResearch Ltd. Grasslands Research Centre, Tennent Drive, Palmerston North 4442, New Zealand.  
CSIRO Plant Industry, Canberra, ACT 2601, Australia.  
Instituto de Histología y Embriología de Mendoza, Universidad Nacional de Cuyo, Centro Universitario, Parque General San Martín S/N, 5500 Mendoza, Argentina.  
Department of Chemistry and Bioscience, Aalborg University, DK-9220 Aalborg East, Denmark.  
Research Institute of Organic Agriculture (FiBL), CH-5970 Frick, Switzerland.  
CSIRO, Canberra, ACT 2603, Australia.  
Institute of Biological, Rural and Environmental Sciences, Aberystwyth University, Gogerddan, SY23 3EE, UK.  
School of Animal and Comparative Biomedical Sciences, University of Arizona, Tucson, Arizona, USA.  
Institute of Animal Nutrition and Nutritional Physiology, Justus Liebig University Giessen, 35392 Giessen, Germany.