nature communications

Article

https://doi.org/10.1038/s41467-024-54025-3

A metagenomic catalogue of the ruminant gut archaeome

Received: 20 December 2023

Accepted: 28 October 2024

Published online: 07 November 2024

Check for updates

Jiandui Mi[1,2] ✉, Xiaoping Jing[3], Chouxian Ma[4], Fuyu Shi[3], Ze Cao[3], Xin Yang[3], Yiwen Yang[5], Apurva Kakade[3], Weiwei Wang[6] & Ruijun Long[3] ✉

While the ruminant gut archaeome regulates the gut microbiota and hydrogen balance, it is also a major producer of the greenhouse gas methane. However, ruminant gut archaeome diversity within the gastrointestinal tract (GIT) of ruminant animals worldwide remains largely underexplored. Here, we construct a catalogue of 998 unique archaeal genomes recovered from the GITs of ruminants, utilizing 2270 metagenomic samples across 10 different ruminant species. Most of the archaeal genomes (669/998 = 67.03%) belong to *Methanobacteriaceae* and *Methanomethylophilaceae* (198/998 = 19.84%). We recover 47/279 previously undescribed archaeal genomes at the strain level with completeness of >80% and contamination of <5%. We also investigate the archaeal gut biogeography across various ruminants and demonstrate that archaeal compositional similarities vary significantly by breed and gut location. The catalogue contains 42,691 protein clusters, and the clustering and methanogenic pathway analysis reveal strain- and host-specific dependencies among ruminant animals. We also find that archaea potentially carry antibiotic and metal resistance genes, mobile genetic elements, virulence factors, quorum sensors, and complex archaeal viromes. Overall, this catalogue is a substantial repository for ruminant archaeal recourses, providing potential for advancing our understanding of archaeal ecology and discovering strategies to regulate methane production in ruminants.

The continuous rise in greenhouse gas (GHG) emissions has triggered rapid changes in global climate patterns, resulting in the frequent occurrence of extreme weather events such as droughts, floods, and heatwaves. These changes profoundly impact the global ecosystem and human livelihoods[1]. Methane (CH_4), which ranks second only to carbon dioxide in terms of greenhouse gases on Earth, has a global warming potential approximately 82.5 times greater than that of carbon dioxide over a 20-year period. Hence, it is recognized as a major contributor to the greenhouse effect[2,3]. Ruminants contribute significantly to methane emissions, making them a focal point of research. CH_4 emissions from ruminants constitute 16% of global greenhouse gas emissions and 30%-32% of global anthropogenic CH_4 emissions[4,5]. In addition, CH_4 emissions represent a substantial energy loss to animals, ranging from 2 to 12% of gross energy intake[6]. Consequently, mitigating CH_4 emissions from ruminants is crucial in the context of sustainable agricultural practices and global climate change mitigation efforts[7]. CH_4 emitted from ruminants is produced primarily by archaea in the gastrointestinal tract (GIT), especially in the rumen[8]. Therefore, a systematic analysis of the composition and function of archaea in the

[1]State Key Laboratory for Animal Disease Control and Prevention, College of Veterinary Medicine, Lanzhou University, Lanzhou 730000, China. [2]Gansu Province Research Center for Basic Disciplines of Pathogen Biology, College of Veterinary Medicine, Lanzhou University, Lanzhou 730000, China. [3]State Key Laboratory of Grassland and Agro-Ecosystems, International Centre for Tibetan Plateau Ecosystem Management, College of Ecology, Lanzhou University, Lanzhou 730000, China. [4]Independent Researcher, Changsha 410023, China. [5]College of Animal Science, South China Agricultural University, Guangzhou 510642, China. [6]College of Animal Science, Guizhou University, Guiyang 550025, China. ✉e-mail: mijiandui@163.com; longrj@lzu.edu.cn

GIT of ruminant animals is crucial for providing background information.

Archaea play a key role in syntrophic metabolism by consuming the end products of biomass fermentation from other microbiota, thus maintaining the hydrogen balance of the GIT[9]. Archaea can utilize compounds such as CO_2, CO, ethanol, formate, acetate or methyl compounds to form CH_4 through three major pathways: hydrogenotrophic, methylotrophic, and aceticlastic. Each trophic type necessitates distinct functions in groups of archaea. Researchers have emphasized that the hydrogenotrophic pathway is responsible for more than 80% of rumen CH_4 production[10–12]. *Methanobrevibacter* spp., which are hydrogenotrophic methanogens, have emerged as the predominant genus involved in rumen methanogenesis[12,13]. Additionally, *Methanosarcinales*, *Methanosphaera*, and *Methanomethylophilaceae* are also involved in reducing methylamine and methanol to CH_4[14–16]. *Methanosarcinales* is notable for its ability to effectively dissolve acetate into carbon dioxide and methane[17]. Yaks, as native ruminants, have been identified as 'low methane' emitters[18,19]. The variations in CH_4 emissions among different ruminants may be attributed to the distinct composition of methanogens in the rumen[19–22]. Although numerous researchers have explored the function and composition of the primary methanogens in the GIT of ruminants[2,9,11,23], comprehensive and global analyses of the systematic abundance of different ruminant gut archaea are lacking.

In this work, we establish a catalogue of the ruminant GIT archaeome by collecting and assembling 2,270 ruminant gut metagenomic samples, including previously assembled and cultured archaeal genomes. The catalogue comprises a total of 998 dereplicated archaeal genomes (99.9% Average Nucleotide Identity-ANI similarity), and 42,691 protein clusters and 216/556 strains (99% ANI similarity) that were previously undescribed. Furthermore, our study extends to the gut biogeography of the archaeome across various ruminants, revealing host- and gut-segment-dependent characteristics of the archaeal composition. This catalogue expands the taxonomic and functional variation of the ruminant gut archaeome, and also holds the potential to advance our understanding of archaeal ecology in ruminants.

Results

A total of 998 unique archaeal genomes were recovered from ruminant gastrointestinal samples

To explore the diversity of archaea in ruminant GIT samples, we recovered genomes from 2270 global metagenomic samples and combined them with publicly available genomes from previous collections of metagenome-assembled genomes (MAGs) and isolates. The 998 obtained nonchimeric and nonredundant genomes spanned a wide range of taxonomic diversity, including those of *Methanobacteriaceae* ($n = 669$; 67.03%), *Methanomassiliicoccales* ($n = 198$; 19.84%), *Methanocorpusculaceae* ($n = 90$; 9.02%), *Methanomicrobiaceae* ($n = 32$; 3.21%), and *Methanosarcinaceae* ($n = 9$; 0.90%) (Supplementary Data 1; Fig. 1). Most of the genomes were taxonomically affiliated with the known genus *Methanobrevibacter* ($n = 616$; 61.72%), which was consistent with previous studies[12,24]. However, this proportion was lower than that reported in archaeal genomes recovered from human GIT samples (998/1167; 85.52%)[25]. Other identified genomes belonged to genera such as *UBA71* of the family *Methanomethylophilaceae* ($n = 105$; 10.52%), *Methanocorpusculum* ($n = 90$; 9.02%), *Methanomethylophilus* ($n = 54$; 5.41%), *Methanosphaera* ($n = 51$; 5.11%), and *Methanomicrobium* ($n = 32$; 3.21%), whereas *SIG5* ($n = 23$; 2.30%) and *MX-02* ($n = 8$; 0.80%) were from *Methanomethylophilaceae*, *Methanimicrococcus* ($n = 7$; 0.70%), and *Methanosarcina* ($n = 2$; 0.20%). *CADBMSO1* of the families *Methanobacteriaceae* and *Methanobacterium* were each represented by a single genome. Among the 998 genomes, 8 belonged to *Methanomethylophilaceae* (0.80%) and could not be assigned to any previously described genus, whereas 140 genomes (14.03%) did not match

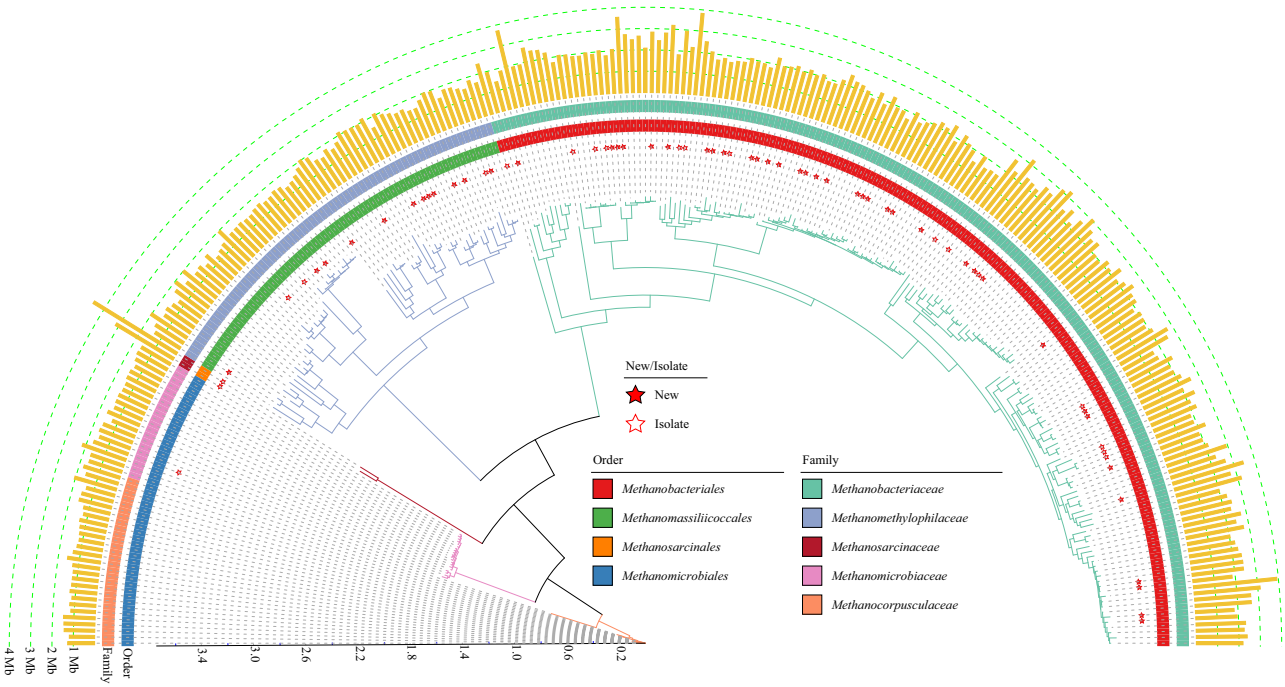

Fig. 1 | Archaeal genomes (279) at the strain level from the ruminant gastrointestinal tract reveal taxonomic expansion of the archaeome. The phylogenetic tree depicts archaeal genomes with a completeness of >80% and contamination of <5%, clustered at 99% similarity (strain level). The characteristics displayed from the center to the outside include the branch of the phylogenetic tree with ultrafast bootstrap values. Newly recovered genomes and isolates are indicated by red solid stars (New) and red stars (Isolate), respectively. The taxonomic affiliations of the MAGs are shown at the order and family levels. The genome size of each MAG is represented by brown bars in megabases (Mb). The different colors of the branches correspond to the colors of the classified families.

https://doi.org/10.1038/s41467-024-54025-3

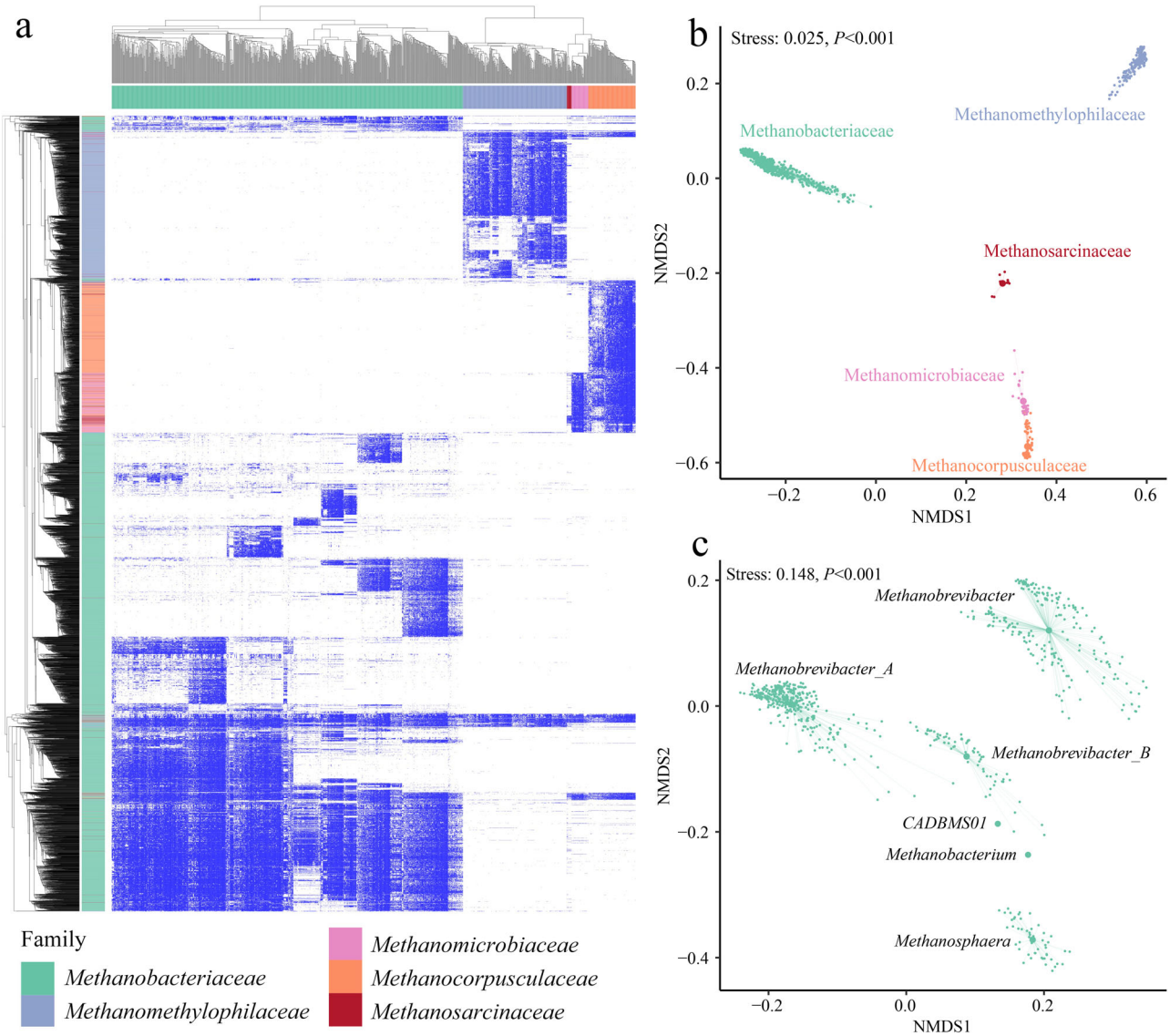

Fig. 2 | Archaeal genomes from the ruminant gastrointestinal tract distribution and the corresponding unified protein catalogue. a Unified ruminant gut archaeal protein catalogue based on protein clustering at 50% sequence identity and 80% coverage via MMseqs2 of all 998 archaeal genomes. The heatmap depicts the presence of 5456 proteins (columns) across the 998 archaeal genomes (rows). Heatmap visualization was performed via the pheatmap package in R. White indicates that no data were available. **b** Nonmetric multidimensional scaling (NMDS) plot of the proteins selected for the heatmap. The NMDS plot shows five distinct clusters corresponding to the archaeal families *Methanobacteriaceae*, *Methanomethylophilaceae*, *Methanocorpusculaceae*, *Methanomicrobiaceae*, and *Methanosarcinaceae*, with the central dots representing the mean NMDS scores for each genome. **c** The NMDS plot shows six distinct clusters corresponding to the archaeal genera *Methanobrevibacter*, *Methanobrevibacter_A*, *Methanobrevibacter_B*, *Methanosphaera*, *Methanobacterium*, and *CADBMS01* of the *Methanobacteriaceae* family with the central dot representing the mean NMDS scores for each genome. The stress value was calculated via metaMDS functions from the vegan package with Bray-Curtis distances (**b**, **c**). The *P* value was calculated by permutational multivariate analysis of variance (PERMANOVA) analysis (**b**, **c**). Source data are provided as a Source Data file.

any known species (Supplementary Data 1). An equally large proportion of the genomes not matching any known species ($n = 70$; 7.01%) were affiliated with the families *Methanobacteriaceae* and *Methanomethylophilaceae*. A phylogenetic tree was constructed for the 279 archaeal genomes at the strain level with a completeness of >80% and contamination of <5% (Fig. 1). Among these genomes, 47 were newly recovered from this study, with 148 genomes derived from the rumen and 105 genomes recovered from mixed assemblies with GIT metagenomic samples (Fig. 1). A phylogenetic tree was also constructed for the 556 archaeal genomes at the strain level with a completeness of >50% and contamination of <5%, of which 216 were previously undescribed (Supplementary Fig. 1).

The archaeal protein profile correlated with different taxonomies

In total, 1,604,531 proteins were identified from the 998 genomes. A protein catalogue of all 998 archaeal genomes was generated by clustering the predicted genes across all the genomes and excluding singleton clustering, resulting in 42,691 cluster representatives (>50% amino acid identity and >80% coverage). A total of 5,456 proteins were shared among >50 genomes in our dataset, revealing the taxonomic distance of the three most abundant families: *Methanobacteriaceae*, *Methanomethylophilaceae*, and *Methanocorpusculaceae* (Fig. 2a). The NMDS plot also demonstrated significant separation among five archaeal families: *Methanobacteriaceae*, *Methanomethylophilaceae*, *Methanocorpusculaceae*, *Methanomicrobiaceae*, and *Methanosarcinaceae*

https://doi.org/10.1038/s41467-024-54025-3

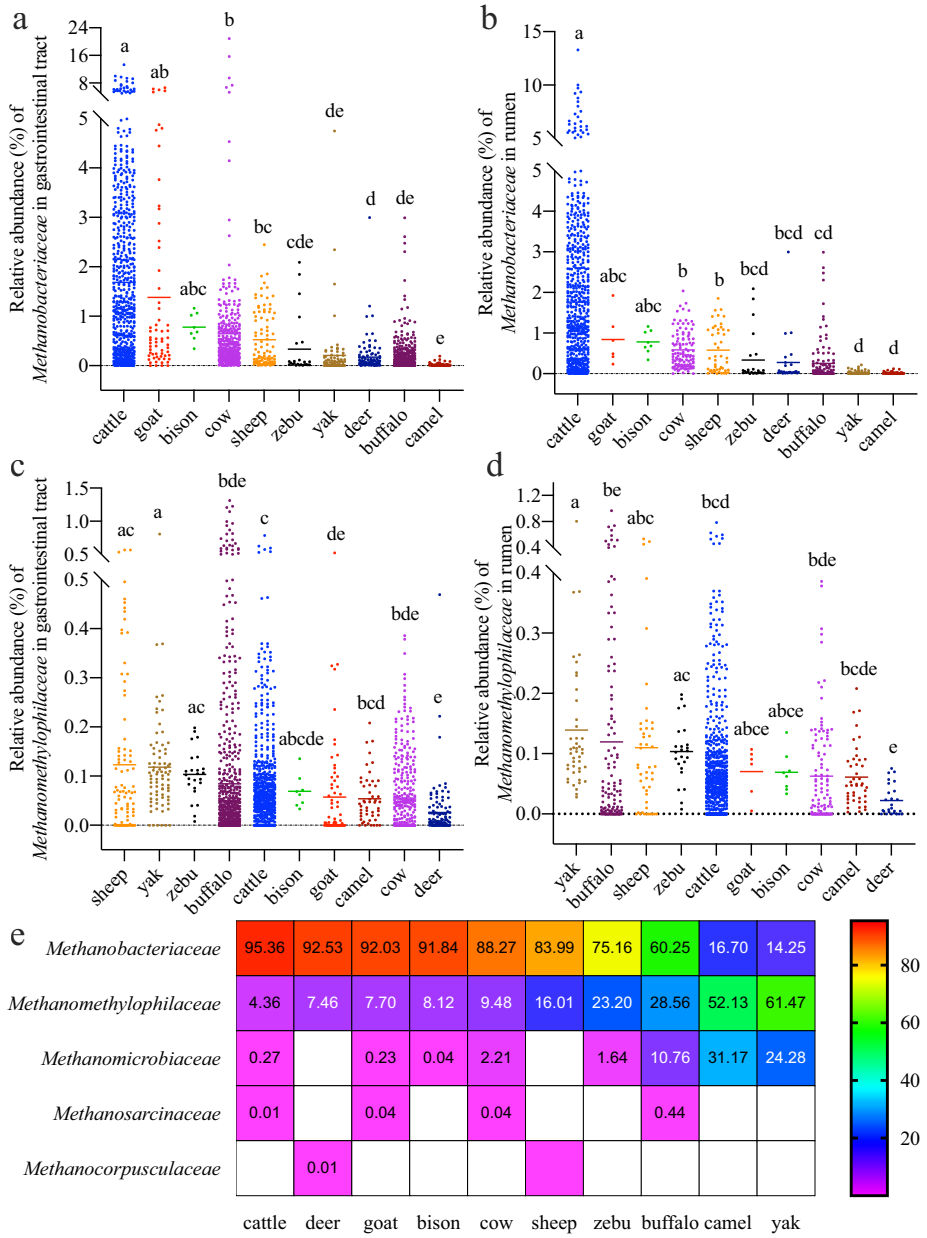

Fig. 3 | Percentage and relative abundance of archaeal genomes at the family level across different ruminant species. a Percentage of *Methanobacteriaceae* in the intestine. **b** Percentage of *Methanobacteriaceae* in the rumen. **c** Percentage of *Methanomethylophilaceae* in the intestine. **d** Percentage of *Methanomethylophilaceae* in the rumen. **e** Heatmap showing the composition of archaeome at the family level in the rumen. From red to purple indicates values from high to low. The number of samples for breeds in the intestine (**a**, **c**) is cattle: $n = 638$, goat: $n = 64$, bison: $n = 8$, cow: $n = 366$, sheep: $n = 92$, zebu: $n = 24$, yak: $n = 79$, deer: $n = 105$, buffalo: $n = 682$, and camel: $n = 50$. The number of samples for breeds in the rumen (**b**, **d**) is cattle: $n = 606$, goat: $n = 6$, bison: $n = 8$, cow: $n = 111$, sheep: $n = 47$, zebu: $n = 24$, yak: $n = 44$, deer: $n = 26$, buffalo: $n = 125$, and camel: $n = 44$. The middle horizontal line represents the mean value (**a**–**d**). The statistical significance of the abundance differences was calculated by the two-sided Kruskal-Wallis test corrected for multiple comparisons via Dunn's test (**a**–**d**). Different letters indicate statistically significant differences ($P < 0.05$), with the specific pairwise comparison *P*-values as Supplementary Data 3 demonstrates (**a**–**d**). Source data are provided as a Source Data file.

(Fig. 2b). To investigate differences at the genus level in the main family, *Methanobacteriaceae*, in the ruminant gut, we created NMDS plots (Fig. 2c). Interestingly, the plots revealed that, according to the GTDB classification, *Methanobrevibacter*, *Methanobrevibacter_A*, and *Methanobrevibacter_B* exhibited significant separation. The results were consistent with the classification rules of the GTDB, where different tails (_A/_B) signify different genera, although they have not yet been named (Fig. 2c). Another significant separation was observed for the genus *Methanosphaera* (Fig. 2c).

Composition of archaeal genomes shaped by breed and gut biogeography

Read-based community profiling revealed variations across breeds and gut biogeography (Supplementary Data 2, Fig. 3, and Fig. 4). For the three major families of archaeal genomes, *Methanobacteriaceae* and *Methanomethylophilaceae*, relative abundance and significance analyses were conducted between breeds in the gastrointestinal and rumen microbiomes (Fig. 3a–d). The relative abundance of *Methanobacteriaceae* was significantly greater in the gastrointestinal and rumen

https://doi.org/10.1038/s41467-024-54025-3

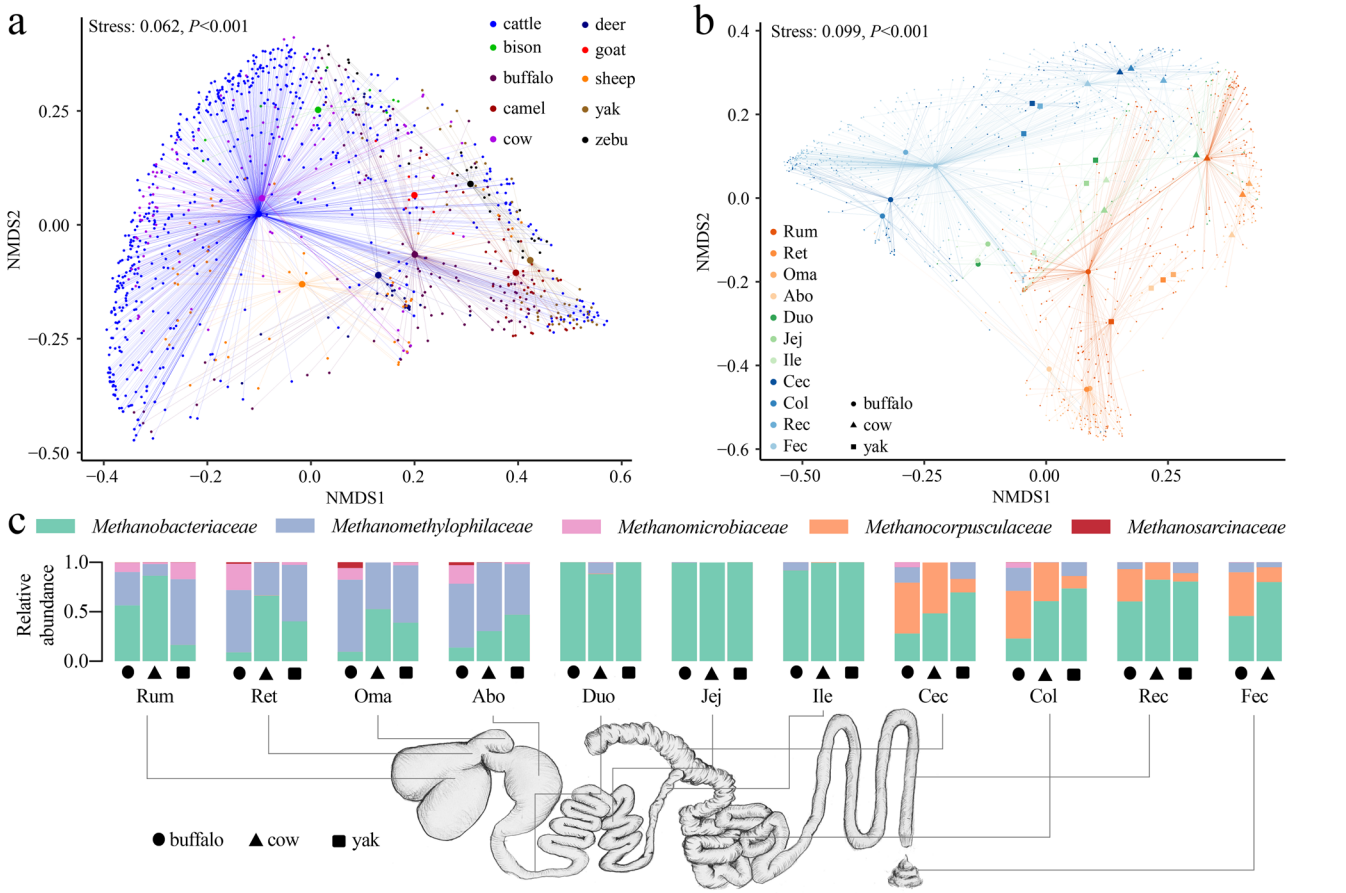

Fig. 4 | Percentages of archaea in the rumen across ruminant species and relative abundances determined through gut biogeography. a Nonmetric multidimensional scaling (NMDS) plot depicting the percentage of archaea in the rumen microbiome across ruminant species, with the central dot representing the mean NMDS score for each sample. **b** NMDS plot illustrating the percentage of archaea in ruminant gut microbiomes across different gut biogeographies of buffalo, cows, and yaks with the central dot representing the mean NMDS scores for each sample. The stress value was calculated via metaMDS functions from the vegan package with Bray-Curtis distances (**a**, **b**). The *P* value was calculated via permutational multivariate analysis of variance (PERMANOVA) analysis (**a**, **b**). The black circle, triangle, and square symbols represent buffalo, dairy cow, and yak, respectively. **c** The relative abundance of archaea in ruminant gut microbiomes across gut biogeographies of buffalo, cows, and yaks. The black circles, triangles, and squares represent buffalo, dairy cows, and yaks, respectively. Rum Rumen, Ret Reticulum, Oma Omasum, Abo Abomasum, Duo Duodenum, Jej Jejunum, Ile Ileum, Cec Cecum, Col Colon, Rec Rectum, and Fec Feces. The schematic diagram of the ruminant gut was hand-drawn and authorized by Jingbo Xia, a friend of J.D. M. (**c**). Source data are provided as a Source Data file.

microbiomes of beef cattle and cows than in those of yak and buffalo ($P < 0.0001$; Supplementary Data 3, Fig. 3a, b). However, there was no difference between the buffaloes and yaks (Fig. 3a, b). In contrast, the relative abundance of *Methanomethylophilaceae* was greater in the entire gastrointestinal and rumen microbiomes of yaks than in those of beef cattle, cows, and buffaloes ($P < 0.0001$; Supplementary Data 3, Fig. 3c, d). A detailed investigation of the relative abundance and composition of families in the rumen was conducted (Supplementary Data 2 and Fig. 3e). The overall relative abundances of the archaeal genomes in the rumens of beef, goat, bison, cow, sheep, zebu, deer, buffalo, yak, and camel were 1.76%, 0.97%, 0.95%, 0.67%, 0.72%, 0.47%, 0.30%, 0.40%, 0.26%, and 0.14%, respectively (Supplementary Data 2). The family *Methanocorpusculaceae*, which includes a special host-associated genus, *Methanocorpusculum*, which is related to 'low methane' emitters, was only read-assigned to the rumen microbiome of sheep, deer, and buffalo and had a very low relative abundance (Fig. 3e)[21]. However, a high relative abundance of the family *Methanocorpusculaceae* was detected in the hindgut, including the cecum, colon, rectum, and feces of buffaloes, cows, and yaks (Fig. 4c). For the families *Methanomethylophilaceae* and *Methanomicrobiaceae*, the relative abundances of the total archaeome were 83.30% and 85.75%, respectively, in the rumens of camels and yaks (Fig. 3e). In summary, the composition of archaeal genomes in the rumen was influenced by breed (Figs. 3e and 4a)[26].

To support previous reports[24,27,28], highlighting that gut biogeography is a crucial factor shaping the diverse composition and functions within animal gut microbiomes, we investigated whether the archaeal genomes of ruminants are determined by gut biogeography. Buffalo, cow, and yak were selected for nonmetric multidimensional scaling (NMDS) analysis, considering the adequate number of representative samples from different gut locations (Fig. 4b). The results demonstrated that the composition of archaeal genomes in buffaloes, cows, and yaks could be differentiated on the basis of gut location (Fig. 4b). Notably, for buffalo and yak, the archaeal genomes could be separated into three distinct gut compartment groups (stomach: rumen, reticulum, omasum, and abomasum; small intestine: duodenum, jejunum, and ileum; and large intestine: cecum, colon, rectum, and feces) at the taxonomic level (Fig. 4b). The analysis also suggested substantial changes in archaeal taxa across the 10 ruminant GITs and feces (Fig. 4c). For example, *Methanobacteriaceae* and *Methanomethylophilaceae* dominated in the stomach region, whereas *Methanobacteriaceae* were more abundant in the small intestine than in the stomach and large intestine, which agreed with the findings of a previous study[29]. *Methanobacteriaceae* and *Methanocorpusculaceae* were

more prevalent in the large intestine and feces. These results underscore that gut biogeography exerts great selective pressures on the ruminant archaeal genome communities.

Comparison of the *Methanomethylophilaceae* clade in different environments

In yaks, we observed a greater relative abundance of *Methanomethylophilaceae* in the rumen, which is the main intestinal site for biomass degradation through interactions between bacteria, archaea, fungi, protozoa, and viruses, than in beef cattle and cows. Therefore, we suggest that host-associated *Methanomethylophilaceae* exhibit taxonomic and functional distinctions from their environmental counterparts due to the unique characteristics of their host environments. According to the phylogenetic tree analysis, *Methanomethylophilaceae* in the ruminant gastrointestinal microbiome separated into three distinct clades, most of which differed from those of strains from host-associated sources (human, termite, and chicken) and environmental sources (wastewater, soil, and sediment) (Supplementary Fig. 2a). Only two human-associated *Methanomethylophilaceae* genomes clustered with ruminant-associated *Methanomethylophilaceae* genomes in the two clades of the phylogenetic tree (Supplementary Fig. 2a). ANI-based analyses of the family *Methanomethylophilaceae* revealed an overall separation between the genomes of different origins (Supplementary Fig. 2b). On basis of their representative microbiota, the *Methanomethylophilaceae* strains in this study could be classified into (1) the exclusively found group of the ruminant gut; (2) host-associated (ruminant, human, and chicken) strains exclusive to the termite gut; and (3) strains with various origins (including ruminant, human, chicken, termite, wastewater, sediment, and soil), highlighting the diversity of *Methanomethylophilaceae* in host-associated and environmental sources[30–32].

Functional and metabolic interactions of the archaeome with different ruminants

We analyzed features that could reveal the advanced functional and metabolic distribution of the ruminant-associated gut archaeome involved in methane metabolism (Fig. 5). In agreement with a previous report[25], in addition to the major components of methanogenesis, such as methyl-coenzyme M reductase (*Mcr*) and heterodisulfide reductase/[NiFe] hydrogenase (Hdr/Mvh) complexes, the major ruminant gut methanogens (*Methanobacteriaceae* and *Methanomethylophilaceae*) possess very distinct methanogenesis pathways (Fig. 5 and Supplementary Data 4). Thus, *Mcr* is commonly used as a target to reduce methane production by inhibiting its activity, as demonstrated by inhibitors such as 3-NOP[33,34]. For example, most of the genetic potential for the H_2/CO_2 pathway comes from *Methanobacteriaceae*. In the methylotrophic pathway of methanogenesis, most of the genes belong to *Methanomethylophilaceae*, with additional genes originating from *Methanosarcinaceae*. The relative abundance of *Methanomethylophilaceae* varies in ruminant breeds, suggesting that the capacity for methyl compound production by the ruminant gut microbiota might be influenced by different populations and diet selections[32,35,36]. However, approximately 10% of the *MtaABC* genes (474 counts), indicating the genetic potential of methanol, were associated with *Methanobacteriaceae*. The presence of *MtaABC* genes in some *Methanobacteriaceae* species strongly indicates that methanol utilization might confer an adaptive advantage in the ruminant guts with high methanol concentrations[25]. However, the conditions under which *Methanobacteriaceae* species use methanol need to be investigated to determine whether they differ between ruminant breeds and/or whether methanol serves as a methanogenic substrate within a broader anabolic pathway for *Methanobacteriaceae*.

A small proportion of genes involved in the methylotrophic pathway of methanogenesis come from *Methanosarcinaceae*, including *MtaABC*, *MtsAB*, *MtmBC*, *MtbABC*, and *MttBC*, which are responsible for the utilization of methanol, methanethiol, methylamin, dimethylamine, and trimethylamine, respectively. This capacity enables *Methanosarcinaceae* to reduce the concentrations of methyl compounds produced by the ruminant gut microbiota, contributing to methane emissions[37,38]. Interestingly, genes involved in the acetoclastic methanogenesis pathway, such as *AckA*, *Pta*, and *CdhA-E*, have also been detected in *Methanosarcinaceae* (Fig. 5a), indicating that in some cases, these *Methanosarcinaceae* might be essential for energy conservation during acetoclastic methanogenesis[39]. The main methylotrophic methanogen *Methanomethylophilaceae*, which also encodes *AcsA*, *AcdAB*, and *CdhA-E*, is involved in the conversion of acetate to acetyl-CoA and methyl-H_4MPT. It has been reported that the acetoclastic methanogenic pathway in *Methanomassiliicoccales* occurs in the monogastric gut[40]. However, the interactions and regulatory mechanisms of methane production in both the methylotrophic and acetoclastic metabolic pathways in *Methanosarcinaceae* and *Methanomethylophilaceae* remain unclear and require further investigation.

Methanocorpusculaceae, the second most important hydrogenotrophic methanogen in the ruminant gut, is involved in all genes in the pathway from CO_2 to CH_4, with the greatest proportion of genes being *FwdA-E* (Fig. 5a). The relative abundance of *Methanocorpusculaceae* is greater in 'low methane' emitters, such as the southern hairy-nosed wombat (*Lasiorhinus latifrons*), which is a foregut fermenter[21]. The relative abundance of *Methanocorpusculaceae* was greater in the large intestine than in the stomach and small intestine (Fig. 4c), indicating that *Methanocorpusculaceae* may be better adapted to the intestinal fermentation environment of nonruminant animals. *Methanocorpusculaceae* also encodes the highest proportion of the membrane-bound complex *EchA-F*, constituting a set of genes involved in electron transfer during methane production. Membrane-bound *Fpo* complexes were more prevalent in *Methanomethylophilaceae* (Fig. 5a).

We also confirmed the enrichment of genes involved in methanogenesis between ruminant breeds (Fig. 5b). These results aligned with our hypothesis that genes involved in methylotrophic methanogenesis, including *Mta*, *Mtb*, *Mtm*, and *Mtt*, were more enriched in the rumens of yaks and camels than in those of zebu, sheep, goats, deer, bison, cows, beef cattle, and buffalo (Fig. 5b). Conversely, the genes involved in hydrogenotrophic methanogenesis, namely, *Fwd*, *Fdh*, *Ftr*, *Mtd*, *Mch*, *Mer*, and *Mtr*, were more enriched in the rumens of camels and yaks. The *Acs* gene involved in acetoclastic methane production was also enriched in the rumens of camels and yaks, following the same trend as that of the genes involved in methylotrophic methanogenesis (Fig. 5b). However, no enrichment was detected in the genes *Mcr*, *Hdr*, and *Mvh*, which are involved in hydrogenotrophic, methylotrophic, and acetoclastic methanogenesis, between the breeds.

Ruminant-associated archaeal genomes potentially carry unique resistance genes and other functions

For antibiotic resistance gene (ARG) detection, the results revealed that 7 *Methanobrevibacter* sp. genomes contained the greatest number of ARGs, primarily *tet* genes that encode proteins responsible for resistance to tetracycline antibiotics, including *tetM*, *tetW*, *tetO*, *tet32*, *tetO/W*, *tet44*, and *tetW/N/W* (Supplementary Fig. 3a). This finding was consistent with a previous report on ARGs in the rumen microbiota, where *tetW* was identified as the dominant ARG[41]. However, these findings differed from those of previous studies[42,43], suggesting potential influences of diet, environment, and management. Two genomes, *Methanobrevibacter B boviskoreani* and *Methanobrevibacter A millerae B*, contained both ARGs and mobile genetic elements (MGEs) (Supplementary Fig. 3a). According to the virulence factor database (VFDB), the two major categories were immune modulation (IM) and stress survival (SS), with a large proportion of the genomes belonging to *Methanobrevibacter, Methanobrevibacter_A, and Methanobrevibacter_B* (Supplementary Fig. 3b). For quorum sensing (QS)

https://doi.org/10.1038/s41467-024-54025-3

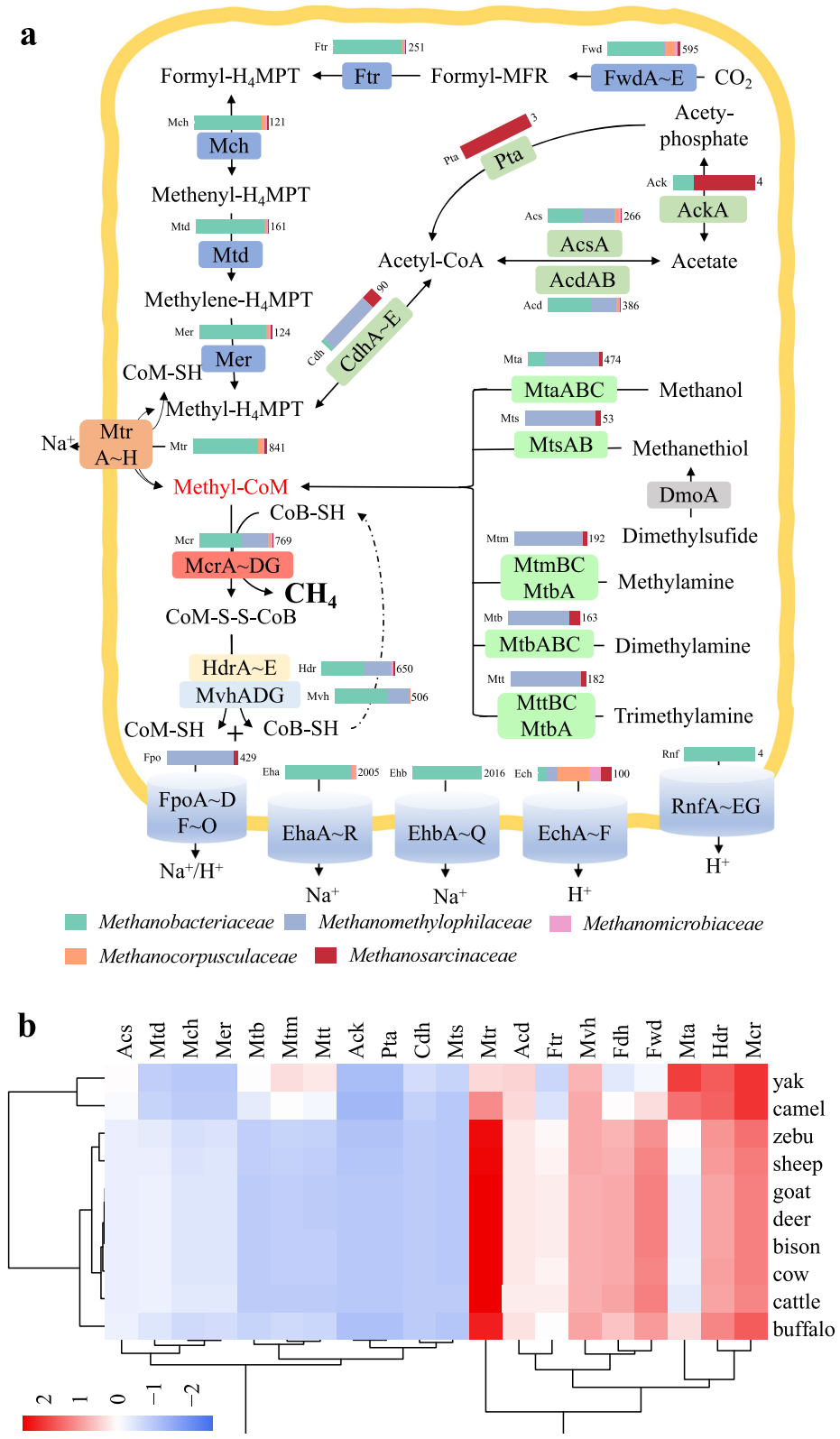

Fig. 5 | Methanogenic pathways in the ruminant gut-associated archaeome at the species level. a The proportion of species with a predicted protein or protein complex is indicated by the bar for total archaeal genomes at the species level ($n = 203$). The number on the right of the bar represents the total number of genes predicted for all archaeal genomes. The blue, light green, and dark green squares on the lines represent genes involved in three different methanogenic pathways: hydrogenotrophic, methylotrophic, and aceticlastic, respectively. **b** Differential enrichment of methanogenesis-associated genes in the rumen across ruminant species. Source data are provided as a Source Data file.

detection, only 4-hydroxyl-2-alkyl-quinolines (HAQs)[44] ($n = 154$) and cholera autoinducer-1 (CAI-1)[45] ($n = 99$) were identified as the predominant languages in the ruminant archaeal genomes (Supplementary Fig. 3c). This differed from a previous report, as AI-2 was suggested as a potentially "universal" signal mediating social behavior in an analysis of 981 rumen bacterial and archaeal genomes[46]. These two dominant QS languages were exclusively found in *Methanobrevibacter* sp. (Supplementary Fig. 3c). We also detected antibacterial biocide and metal resistance genes in ruminant archaeal genomes. The results revealed that the largest category of resistance genes ($n = 13,080$) associated with Cu/Ag resistance (Supplementary Fig. 3d) primarily originated from *Methanobrevibacter* species, including *Methanobrevibacter_A*, *Methanobrevibacter*, and *Methanobrevibacter_B*.

Ruminant lysogenic archaea and proviruses

In total, 164 archaeal genomes in the ruminant archaeome contained at least one viral sequence (mean = 1.99), revealing a lysogeny ratio (LyR) of 16.4% (164/998) among ruminant gut archaea (Supplementary Data 5). This LyR was lower than 28.5% for archaea in marine systems[47]. We identified 327 viral populations in these archaeome datasets, with the virome distributed among 8 complete, 128 high-quality (>90% completeness), and 191 medium-quality (50–90% completeness) categories (Supplementary Fig. 4a and Supplementary Data 6). Among the identified proviruses, 299 viral species were specific to *Methanobrevibacter* (including 148 for *Methanobrevibacter* and 129 and 22 for *Methanobrevibacter A* and *B*, respectively) (Supplementary Fig. 4b). This diversity was greater than that reported for the human gut archaeome[25]. In addition, viromes in extreme environments and the human gut have been extensively explored through metagenomic analysis[25,48–54] because of the importance of the virome in modulating the balance and functions of the ecosystem. The rumen virome database (RVD) was established by mining 975 globally published rumen metagenomes[55]. However, to our knowledge, this is a systematic description of ruminant gut archaeal lysogens and the virome. On the basis of the taxonomic classification released by ICTV 2022[56], 301 vOTUs were assigned to the *Caudoviricetes* class, including the major families *Myoviridae*, *Siphoviridae* and *Podoviridae*, which were removed from the ICTV update (Supplementary Fig. 4c). A total of 1074 AMGs were found in ruminant archaeal viruses. Functional analysis revealed that virome AMGs were involved in miscellaneous (MISC) ($n = 332$), organic nitrogen ($n = 221$), carbon utilization ($n = 221$), energy ($n = 174$), transporters ($n = 119$), and carbon utilization (Woodcroft) ($n = 7$) (Supplementary Fig. 4d). For carbon utilization, the main function was that of glycosyl transferases (Supplementary Fig. 4e). For the energy category, AMG was involved mainly in methanogenesis, including the conversion of CO_2, acetate, methanol, methylamine, dimethylamine, and trimethylamine to form methane (Supplementary Fig. 4e). In summary, ruminant archaeal viruses are potentially ecologically important and may function in the ruminant gut microbiome, particularly in strategies to potentially reduce methane production.

Discussion

Although extensive research has been conducted on the rumen and intestinal microbial communities of ruminants, with a primary focus on bacterial populations[29,57–60], systematic analyses of the ruminant gut archaeome remain insufficient. This study provides insights into the biology of the ruminant gastrointestinal tract (GIT) archaeome by assembling and cataloging 998 nonredundant archaeal genomes. Initial associations between the diversity and function of ruminant gut-associated archaea of different breeds and enterotypes were established. However, some intestinal segments excluding the rumen from many geographic locations, such as regions in America, Africa, Oceania, and Russia, are inadequately sampled. Further efforts should aim to increase the analysis of the gut microbiota in these areas or other environments.

Since a small subset of genomes (21/998 = 2%) has been obtained from cultured archaeal representatives[29,59], it is crucial to obtain more cultured isolates from the ruminant gut to better understand the ecology, evolution, and function of archaeal genomes. Advancements in high-throughput cultivation methods, machine learning, and Raman microspectroscopy technologies can accelerate the isolation of many isolates on demand, overcoming the limitations of traditional labor-intensive methods[61–63]. Moreover, the dataset of archaeal genomes can serve as a reference and a starting point for targeted cultivation of new members of the ruminant gut archaeome. Additionally, the use of long-read metagenomics sequencing techniques, such as Oxford Nanopore Technology (ONT) and PacBio, has the advantage of longer read lengths (capable of obtaining 10 kb to 1 Mb fragments), which can aid in assembling and recovering more complete MAGs[64]. The use of specialized methods to reduce host contamination, increase cell lysis, and improve DNA extraction might also reduce the amount of information available on archaea from different gut biogeographies[65,66]. Devices used to sample regions of the human intestinal tract might also be adapted to easily sample wild ruminant animals that cannot be slaughtered[67]. In summary, the use of adapted technologies might improve and enable profound capture of more diverse ruminant gut archaeomes.

The observed percentage of archaea in the ruminant gut microbiome varied by breed and gut biogeography, ranging from 0.14% to 1.76% in the rumens of different ruminants (Fig. 3g). This finding was similar to the average percentages reported for human gut archaea (~1.2%)[25]. However, similar to that in humans, the abundance of methanogenic archaea in ruminants is highly variable and positively correlated with methane exhalation[25]. Interestingly, low-methane emitters, such as yaks, present different archaeal compositions in the rumen. The *Methanomethylophilaceae* enterotype dominates the landscape of the archaeal community composition according to the definition of enterotypes[26,32]. This finding agrees with studies on methane emissions from yak[18,36] and the comparison of methanogen diversity between yak (*Bos grunniens*) and cattle (*Bos taurus*)[68]. The rumen archaea of camels also contain the *Methanomethylophilaceae* enterotype, which was also observed in camels that produce less methane than in those of ruminants such as cattle, sheep and goats[69]. However, the mechanism underlying the archaeal enterotypes and activity of high- and low-methane emitters remains largely unclear, revealing a potential avenue for reducing methane emissions from ruminants. Further research is essential to elucidate the intricacy of biochemical processes, such as the transfer and utilization of major methyl substrates and dissolved hydrogen, as well as microbial interactions involving bacteria, archaea, viruses, fungi, and protozoa. Advanced techniques such as metagenomics, metatranscriptomics, and stable isotope probing can offer valuable insights into the metabolic activities and functional roles of methylotrophic methanogens in the context of ruminant digestion. Moreover, the rumen, as one of the most efficient anaerobic fiber fermentation systems involving archaea, holds tremendous potential for the development of biomass energy[70].

The presented establishment of ruminant gut-associated archaeal genomes, along with a catalogue of 1.6 million predicted proteins, provides foundational information and serves as a specialized data source for addressing major questions in future research. Understanding the shaping of the diversity and function of the archaeome by ruminant host and archaeome-host interactions, potentially through microRNA secretion by host cells, is crucial for breeding ruminants with low CH_4 emissions and guides microbiome-informed breeding strategies[11]. The ruminant gut microbiome, particularly in the rumen, functions through trophic-like levels to ferment the diet and generate nutrients to meet the host's requirements[11]. Further research is warranted to investigate how the archaeome, as the third sublevel,

interacts with various microorganisms (bacteria, viruses, fungi, and protozoa) at other trophic-like levels, ultimately influencing the host's digestion and utilization efficiency of feed. The ruminant gut archaeome potentially carries virulence and resistance genes, with a high prevalence of *tet*-type category ARGs that confer resistance to tetracycline. Considering archaea, conducting global-scale studies is imperative to evaluate the risks associated with gene transfer, environmental impact, food safety, and human health implications[41,42]. The archaeal virome also encodes auxiliary metabolic genes (AMGs), particularly those related to methane metabolism (Supplementary Fig. 4). However, exploring the process and extent to which the archaeal virome influences methane emissions in ruminants and the potential for targeted reduction of methane emissions by focusing on specific archaea requires further research[1,55].

In summary, we successfully delineated the landscape of the archaeome in the GIT of ruminant animals by assembling 2270 ruminant gastrointestinal metagenomes and incorporating previously assembled and isolated archaeal genomes. We confirmed two distinct enterotypes in ruminant GIT archaea, namely, the *Methanobacteriaceae* enterotype and the *Methanomethylophilaceae* enterotype, with their composition and dominant methanogenesis pathways showing variations on the basis of breed and gut biogeography. The viral community infecting archaea in the GIT of ruminants carries many methanogenesis genes, offering the potential to regulate archaea to reduce methane emissions. These findings provide comprehensive insights into the ecology and functionality of the gut archaeome in ruminant animals, laying a solid foundation for further exploration in mitigating greenhouse gas emissions through precise regulation of archaea. However, the limitations of this study include the lack of additional culture and in-depth mechanistic validation of relevant methanogens[8] and the lack of incorporation of third-generation sequencing data to improve the quality and recovery rate of the assembly. In further studies combining culturomics and the long sequence lengths of third-generation sequencing, complete genomes and MAGs with higher quality could be obtained. This would contribute to gaining deeper insights into higher-resolution strains, allowing for a more precise understanding of their genetic diversity, functional capabilities, and ecological roles in the gut of ruminants.

Methods

Dataset description

A comprehensive catalogue encompassing 998 genomes from the ruminant gut archaeome has been compiled on a global scale. This extensive collection aims to explore the diversity of archaeal genomes and investigate variations in the formation and metabolism of CH_4 across ruminant breeds and gut locations. We gathered 2,270 metagenomic samples from publicly available databases (Supplementary Data 7 and 8; Supplementary Fig. 5)[9,19,23,29,57,58,71–93]. First, we cleaned the metagenomic data using fastp v.0.23.248 (with parameters –detect_adapter_for_pe and –dont_eval_duplication -w 16), followed by the removal of host contamination specific to ruminant breeds using BWA v.0.7.17 (with parameters -k 31 -p -S -K 200000000)[94]. Second, we assembled the cleaned metagenomic data using Megahit v.1.2.9 with a single-sample assembly model (with a -k-list of 39, 59, and 79)[95]. Contigs less than 2 kb in length in the assembly sequences were then removed with seqkit v2.3.0[96]. Next, a sorted bam file for each sample was generated using minimap2 v2.17-r941 (-ax sr)[97] and samtools v1.6[98] with default parameters. The depth file was calculated with the jgi_summarize_bam_contig_depths script in metabat2 v2.15[99]. Metagenome-assembled genomes (MAGs) were subsequently assembled using metabat2 v2.15 with default parameters. Finally, we employed checkM v1.2.2[100] to assess the quality of the MAGs and selected those identified as archaea with an estimated completeness of >50% and contamination of <5%. In total, 815 MAGs were obtained through these assembly procedures. The quality of the collected MAGs ($n = 517$) (details are shown in the Data availability section) from previous studies was assessed, and only those with a completeness of >50% and contamination of <5% ($n = 428$) were retained. Therefore, the total number of archaeal genomes reached 1243 ($n = 815 + 428$).

Genome quality and taxonomic classification

The complete dataset comprising 1243 archaeal genomes underwent dereplication at 99.9% for nonredundant archaeal genomes ($n = 998$), 99% for individual strains ($n = 556$), and 95% for species ($n = 203$) ANI values using DRep v3.4.0[101] (Supplementary Fig. 6). A subset of archaeal genomes at the strain level, totaling 279 and meeting completeness criteria (>80%) and contamination thresholds (<5%), was utilized for constructing the phylogenetic tree. Tree matrices were generated for 279 archaeal genomes with the criteria of >80% completeness and <5% contamination at the strain level (99% ANI similarity) using PhyloPhlAn v3.0.67[102]. Dendrograms, constructed on the basis of the ANI tree matrix, were annotated using the iTOL tool (Interactive Tree of Life)[103]. For cultured archaeal genomes, 19 and 16 out of the total 21 were retained at the individual strain (99%) and species (95%) levels, respectively. Taxonomic annotation of all archaeal genomes was performed using the GTDB Toolkit v2.1.1 (database release207_v2)[104] with default parameters, employing a set of 122 marker genes to identify archaeal MAGs. The taxonomy of the archaeal genomes was summarized via the gtdb_to_ncbi_majority_vote.py script within the GTDB Toolkit.

Genome annotation and protein catalogue

Protein genes were predicted via Prodigal v2.6.3 with default parameters employing a single model[105]. The protein-encoding genes were subsequently annotated via eggNOG-mapper v2.1.9 and eggNOG database v5.0 with default parameters[106,107]. The protein catalogue was generated by combining all the predicted CDSs (totaling 1,604,531) derived from the 998 nonredundant archaeal genomes. MMseqs2 v14.7e284 linclust was utilized to cluster the concatenated protein dataset with the options '–cov-mode 1 -c 0.8 –kmer-per-seq 80 –min-seq-id 0.5'[108]. Proteins were clustered at various percentage identities, and the resulting number of unique proteins per cluster for total taxonomic families was computed and visualized. To minimize the risk of contaminants, nonclustered proteins were filtered out. For clear protein visualization, proteins that clustered only above 2 genes per cluster and were present in more than 50 genomes were displayed via pheatmap and R with NMDS[109].

Relative abundance of archaea in ruminant metagenomes

To assess the relative abundance of archaeal genomes at the species level in the collected metagenomic datasets, we employed CoverM v0.6.1 (https://github.com/wwood/CoverM) and utilized the relative abundance calculation method with genome patterns.

Methane metabolism pathway establishment

A total of 203 genomes at the species level, with completeness exceeding 50% and contamination below 5% in protein sequences, were annotated against the Kyoto Encyclopedia of Genes and Genomes (KEGG) orthologs database using eggNOG-mapper v2.1.9 with default settings[106]. The assigned hits were consolidated for all genomes using KO numbers. The annotated results were further refined, and KOs associated with methane metabolism were retained on the basis of previous studies[13,15,16,25,110–114]. These included genes involved in hydrogenotrophic, methylotrophic, and acetoclastic pathways contributing to methane formation, as well as electronic receptors and transporters on the membrane. The copy numbers of these genes in the archaeal genomes were computed, and the proportion of each gene at the family level was determined on the basis of the summary copy numbers of the total archaeal genomes. The methane metabolism pathways and the proportion of each gene at the family level were

established and visualized for ruminant archaeal genomes. The copy numbers of genes involved in methane metabolism were combined with the relative abundance of archaeal genomes to compare the differences in these genes between different breeds and visualized using pheatmap[109].

Detection of virulence and resistance genes and quorum sensing

To predict potential virulence genes in all 998 archaeal genomes, Diamond v2.0.15[115] was employed to BLASTp against the following databases: SARGs v3.2.3[116], VFDB[117], BacMet[118], MGEs[119], and Quorum sensing[120], with parameters set at an e value of $1e^{-7}$, a query coverage of 85%, and an identity of 60%. Among the 998 genomes, ARGs were identified in only 74, MGEs in 20, QS in 147, VFDB in 779, and BacMet in 635.

Viral identification, quality estimation and AMG detection

To evaluate the presence of viruses, VirSorter2 v.2.2.3[20] was utilized to search all 998 unique archaeal genomes. To mitigate potential contamination issues arising from the binning process, we focused on proviruses flanked within archaeal contigs for this analysis. The identification of viral sequences followed the viral sequence identification SOP as outlined in a previous report (dx.doi.org/10.17504/protocols.io.bwm5pc86). In summary, CheckV v1.0.1 (database v.1.4)[121] was used to assess the quality of VirSorter2-predicted viral sequences, and a quality control check was performed to eliminate false positives. The criteria for virus sequence inclusion were as follows: keep 1 (viral_gene > 0) and keep 2 (viral_gene = 0 and host_gene = 0 or score >= 0.95 or hallmark > 2). However, only viral sequences of medium quality (50-90% completeness), high quality (>90% completeness), and complete quality were selected for further analysis. A total of 327 viral sequences were identified in our collected archaeal genome datasets. This result was obtained by clustering 8 complete, 128 high-quality (>90% completeness), and 191 medium-quality (50–90% completeness) viral sequences at 95% ANI and 85% AF via the supporting code of CheckV. Taxonomy was determined via geNomad v.1.2.0[122], and the auxiliary metabolic genes (AMGs) of the viruses were annotated via DRAM-v v1.4.6[123].

Comparison of different environmental *Methanomethylophilaceae*

To assess the distinction of *Methanomethylophilaceae* between ruminants and other environments, we assembled a dataset comprising 141 archaeal MAGs identified from diverse environmental samples available in the NCBI database. This included but was not limited to the human gut, termite gut, soil, sediment, wastewater, and chicken gut. This dataset served as a reference for comparison with the *Methanomethylophilaceae* set collected and assembled from the ruminant gut microbiome[14–16,111,112,124]. All the genomes from the different environments and ruminants used in the analysis exhibited completeness exceeding 80% and contamination below 5% and were deduplicated at the strain level (99% similarity) using DRep v3.4.0[101]. A total of 124 *Methanomethylophilaceae* genomes, including 64 from ruminant samples, were retained for further analysis (Supplementary Data 9). We constructed a phylogenetic tree incorporating all the *Methanomethylophilaceae* genomes mentioned using PhyloPhlAn v3.0.67[102] with the parameters –diversity low –fast –min_num_markers 1. The tree was annotated using the iTOL tool[103]. To estimate the pairwise ANI distance between the *Methanomethylophilaceae* genomes from the ruminant gut microbiome and other environmental *Methanomethylophilaceae* genome datasets, we employed fastANI v1.33[125]. ANI values ranging from 75% to 100% were selected for plotting using the pheatmap package[109]. The phylogenetic tree was generated using PhyloPhlAn v3.0.67 and visualized via the iTOL website[103].

Statistical analysis

The 'vegan' package v2.6.4 of R v4.2.2 was used to examine the differences in functional and compositional structure, and the 'adonis' function was used to perform permutational multivariate analysis of variance (PERMANOVA) via the Bray–Curtis method with 1000 permutations. Nonmetric multidimensional scaling (NMDS) analyses and stress values were performed via metaMDS functions from the vegan package with Bray–Curtis distances. The rules of thumb for stress values are as follows: <0.05 is excellent, 0.05-0.10 is good, 0.10-0.20 is fair, and >0.20 is poor. NMDS plots were made with the ggplot2 package v3.4.1. The differences in the relative abundance of archaeal composition across different ruminants were analyzed via the Kruskal–Wallis test with Dunn's multiple comparisons test via Prism 9.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

We gathered 2,270 metagenomic samples from publicly available databases (Supplementary Data 6 and 7; Supplementary Fig. 5)[9,19,23,29,57,58,71–93]. We also compiled archaeal MAGs from previous studies, including PRJNA656389 ($n = 74$)[88], PRJNA723218 ($n = 46$)[92], PRJNA631951 ($n = 15$)[86], PRJEB21624 ($n = 28$)[57], PRJNA657473 ($n = 160$)[29], PRJEB31266 ($n = 125$)[57], and PRJNA627251 ($n = 48$)[126], and cultured archaeal genomes from the gut microbiota of ruminants ($n = 21$)[29,59]. The new recovered genomes (597/998) in this study were deposited in the ENA database under Bioproject number PRJEB81441 (BioSamples accession no. SAMEA116318208- SAMEA116318804). Source data are provided with this paper.

Code availability

The R script was deposited in https://github.com/mijiandui/ruminant-archaeome/tree/main (https://doi.org/10.5281/zenodo.13936351).

References

1. Mizrahi, I., Wallace, R. J. & Morais, S. The rumen microbiome: balancing food security and environmental impacts. *Nat. Rev. Microbiol.* **19**, 553–566 (2021).
2. Martínez-Álvaro, M. et al. Bovine host genome acts on rumen microbiome function linked to methane emissions. *Commun. Biol.* **5**, 350 (2022).
3. IPCC. Sections. In: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland, 35-115 (2023).
4. Dangal, S. R. S. et al. Methane emission from global livestock sector during 1890-2014: Magnitude, trends and spatiotemporal patterns. *Glob. Change Biol.* **23**, 4147–4161 (2017).
5. Arndt, C. et al. Full adoption of the most effective strategies to mitigate methane emissions by ruminants can help meet the 1.5 °C target by 2030 but not 2050. *Proc. Natl Acad. Sci. USA* **119**, e2111294119 (2022).
6. Johnson, K. A. & Johnson, D. E. Methane emissions from cattle. *J. Anim. Sci.* **73**, 2483–2492 (1995).
7. Herrero, M. et al. Greenhouse gas mitigation potentials in the livestock sector. *Nat. Clim. Change* **6**, 452–461 (2016).
8. Khairunisa, B. H., Heryakusuma, C., Ike, K., Mukhopadhyay, B. & Susanti, D. Evolving understanding of rumen methanogen ecophysiology. *Front. Microbiol.* **14**, 1296008 (2023).
9. Li, Q. S. et al. Dietary selection of metabolically distinct microorganisms drives hydrogen metabolism in ruminants. *ISME J.* **16**, 2535–2546 (2022).

10. Beauchemin, K. A., Ungerfeld, E. M., Eckard, R. J. & Wang, M. Review: Fifty years of research on rumen methanogenesis: lessons learned and future challenges for mitigation. *Animal* **14**, s2–s16 (2020).
11. Janssen, P. H. & Kirs, M. Structure of the archaeal community of the rumen. *Appl. Environ. Microbiol.* **74**, 3619–3625 (2008).
12. Pitta, D., Indugu, N., Narayan, K. & Hennessy, M. Symposium review: Understanding the role of the rumen microbiome in enteric methane mitigation and productivity in dairy cows. *J. Dairy Sci.* **105**, 8569–8585 (2022).
13. Berghuis, B. A. et al. Hydrogenotrophic methanogenesis in archaeal phylum Verstraetearchaeota reveals the shared ancestry of all methanogens. *Proc. Natl Acad. Sci. USA* **116**, 5037–5044 (2019).
14. Chistoserdova, L., Kalyuzhnaya, M. G. & Lidstrom, M. E. The expanding world of methylotrophic metabolism. *Annu. Rev. Microbiol.* **63**, 477–499 (2009).
15. Sollinger, A. & Urich, T. Methylotrophic methanogens everywhere - physiology and ecology of novel players in global methane cycling. *Biochem. Soc. Trans.* **47**, 1895–1907 (2019).
16. Vanwonterghem, I. et al. Methylotrophic methanogenesis discovered in the archaeal phylum Verstraetearchaeota. *Nat. Microbiol.* **1**, 16170 (2016).
17. Pereira, A. M., de, Lurdes, Nunes Enes Dapkevicius, M. & Borba, A. E. S. Alternative pathways for hydrogen sink originated from the ruminal fermentation of carbohydrates: Which microorganisms are involved in lowering methane emission? *Anim. Microbiome* **4**, 5 (2022).
18. Ding, X. Z., Long, R. J., Kreuzer, M., Mi, J. D. & Yang, B. Methane emissions from yak (*Bos grunniens*) steers grazing or kept indoors and fed diets with varying forage: concentrate ratio during the cold season on the Qinghai-Tibetan Plateau. *Anim. Feed Sci. Tech.* **162**, 91–98 (2010).
19. Zhang, Z. et al. Convergent evolution of rumen microbiomes in high-altitude mammals. *Curr. Biol.* **26**, 1873–1879 (2016).
20. Guo, J. et al. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome* **9**, 37 (2021).
21. Volmer, J. G. et al. Isolation and characterisation of novel *Methanocorpusculum* species indicates the genus is ancestrally host-associated. *BMC Biol.* **21**, 59 (2023).
22. Zhang, Q. et al. Bayesian modeling reveals host genetics associated with rumen microbiota jointly influence methane emission in dairy cows. *ISME J.* **14**, 2019–2033 (2020).
23. Shi, W. et al. Methane yield phenotypes linked to differential gene expression in the sheep rumen microbiome. *Genome Res* **24**, 1517–1525 (2014).
24. Thomas, C. M., Desmond-Le Quéméner, E., Gribaldo, S. & Borrel, G. Factors shaping the abundance and diversity of the gut archaeome across the animal kingdom. *Nat. Commun.* **13**, 3358 (2022).
25. Chibani, C. M. et al. A catalogue of 1,167 genomes from the human gut archaeome. *Nat. Microbiol.* **7**, 48–61 (2021).
26. Costea, P. I. et al. Enterotypes in the landscape of gut microbial community composition. *Nat. Microbiol.* **3**, 8–16 (2018).
27. Donaldson, G. P., Lee, S. M. & Mazmanian, S. K. Gut biogeography of the bacterial microbiota. *Nat. Rev. Micro* **14**, 20–32 (2016).
28. McCallum, G. & Tropini, C. The gut microbiota and its biogeography. *Nat. Rev. Microbiol.* **22**, 105–118 (2024).
29. Xie, F. et al. An integrated gene catalog and over 10,000 metagenome-assembled genomes from the gastrointestinal microbiome of ruminants. *Microbiome* **9**, 137 (2021).
30. Borrel, G., Brugère, J.-F., Gribaldo, S., Schmitz, R. A. & Moissl-Eichinger, C. The host-associated archaeome. *Nat. Rev. Microbiol.* **18**, 622–636 (2020).
31. Moissl-Eichinger, C. et al. Archaea are interactive components of complex microbiomes. *Trends Microbiol* **26**, 70–85 (2018).
32. Xie, F. et al. Unraveling the phylogenomic diversity of *Methanomassiliicoccales* and implications for mitigating ruminant methane emissions. *Genome Biol.* **25**, 32 (2024).
33. Hristov, A. N. et al. An inhibitor persistently decreased enteric methane emission from dairy cows with no negative effect on milk production. *Proc. Natl Acad. Sci. USA* **112**, 10663–10668 (2015).
34. Pitta, D. W. et al. The effect of 3-nitrooxypropanol, a potent methane inhibitor, on ruminal microbial gene expression profiles in dairy cows. *Microbiome* **10**, 146 (2022).
35. Guo, N. et al. Seasonal dynamics of diet-gut microbiota interaction in adaptation of yaks to life at high altitude. *NPJ Biofilms Microbiomes* **7**, 38 (2021).
36. Jing, X. et al. The adaptive strategies of yaks to live in the Asian highlands. *Anim. Nutr.* **9**, 249–258 (2022).
37. Poulsen, M. et al. Methylotrophic methanogenic Thermoplasmata implicated in reduced methane emissions from bovine rumen. *Nat. Commun.* **4**, 1428 (2013).
38. Schorn, S. et al. Diverse methylotrophic methanogenic archaea cause high methane emissions from seagrass meadows. *Proc. Natl Acad. Sci. USA* **119**, e2106628119 (2022).
39. Chadwick, G. L. et al. Comparative genomics reveals electron transfer and syntrophic mechanisms differentiating methanotrophic and methanogenic archaea. *PLoS Biol.* **20**, e3001508 (2022).
40. Feehan, B. et al. Novel complete methanogenic pathways in longitudinal genomic study of monogastric age-associated archaea. *Anim. Microbiome* **5**, 35 (2023).
41. Sabino, Y. N. V. et al. Characterization of antibiotic resistance genes in the species of the rumen microbiota. *Nat. Commun.* **10**, 5252 (2019).
42. Auffret, M. D. et al. The rumen microbiome as a reservoir of antimicrobial resistance and pathogenicity genes is directly affected by diet in beef cattle. *Microbiome* **5**, 159 (2017).
43. Hitch, T. C. A., Thomas, B. J., Friedersdorff, J. C. A., Ougham, H. & Creevey, C. J. Deep sequence analysis reveals the ovine rumen as a reservoir of antibiotic resistance genes. *Environ. Pollut.* **235**, 571–575 (2018).
44. Gophna, U. & Altman-Price, N. Horizontal gene transfer in archaea-from mechanisms to genome evolution. *Annu. Rev. Microbiol.* **76**, 481–502 (2022).
45. Kelly, R. C. et al. The Vibrio cholerae quorum-sensing autoinducer CAI-1: analysis of the biosynthetic enzyme CqsA. *Nat. Chem. Biol.* **5**, 891–895 (2009).
46. Liu, X. et al. Exploring AI-2-mediated interspecies communications within rumen microbial communities. *Microbiome* **10**, 167 (2022).
47. Yi, Y. et al. A systematic analysis of marine lysogens and proviruses. *Nat. Commun.* **14**, 6013 (2023).
48. Li, R., Wang, Y., Hu, H., Tan, Y. & Ma, Y. Metagenomic analysis reveals unexplored diversity of archaeal virome in the human gut. *Nat. Commun.* **13**, 7978 (2022).
49. Medvedeva, S., Borrel, G., Krupovic, M. & Gribaldo, S. A compendium of viruses from methanogenic archaea reveals their diversity and adaptations to the gut environment. *Nat. Microbiol.* **8**, 2170–2182 (2023).
50. Tamarit, D. et al. A closed *Candidatus* Odinarchaeum chromosome exposes Asgard archaeal viruses. *Nat. Microbiol.* **7**, 948–952 (2022).
51. Rambo, I. M., Langwig, M. V., Leão, P., De Anda, V. & Baker, B. J. Genomes of six viruses that infect Asgard archaea from deep-sea sediments. *Nat. Microbiol.* **7**, 953–961 (2022).

52. Medvedeva, S. et al. Three families of Asgard archaeal viruses identified in metagenome-assembled genomes. *Nat. Microbiol.* **7**, 962–973 (2022).
53. Krupovic, M., Dolja, V. V. & Koonin, E. V. The virome of the last eukaryotic common ancestor and eukaryogenesis. *Nat. Microbiol.* **8**, 1008–1017 (2023).
54. Laso-Pérez, R. et al. Evolutionary diversification of methanotrophic ANME-1 archaea and their expansive virome. *Nat. Microbiol.* **8**, 231–245 (2023).
55. Yan, M. et al. Interrogating the viral dark matter of the rumen ecosystem with a global virome database. *Nat. Commun.* **14**, 5254 (2023).
56. Turner, D. et al. Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee. *Arch. Virol.* **168**, 74 (2023).
57. Stewart, R. D. et al. Compendium of 4,941 rumen metagenome-assembled genomes for rumen microbiome biology and enzyme discovery. *Nat. Biotechnol.* **37**, 953–961 (2019).
58. Wilkinson, T. et al. 1200 high-quality metagenome-assembled genomes from the rumen of African cattle and their relevance in the context of sub-optimal feeding. *Genome Biol.* **21**, 229 (2020).
59. Seshadri, R. et al. Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection. *Nat. Biotechnol.* **36**, 359 (2018).
60. Stewart, R. D. et al. Assembly of 913 microbial genomes from metagenomic sequencing of the cow rumen. *Nat. Commun.* **9**, 870 (2018).
61. Liu, S. et al. Opportunities and challenges of using metagenomic data to bring uncultured microbes into cultivation. *Microbiome* **10**, 76 (2022).
62. Huang, Y. et al. High-throughput microbial culturomics using automation and machine learning. *Nat. Biotechnol.* **41**, 1424–1433 (2023).
63. Zhang, J. et al. High-throughput cultivation and identification of bacteria from the plant root microbiota. *Nat. Protoc.* **16**, 988–1012 (2021).
64. Jin, H. et al. A high-quality genome compendium of the human gut microbiome of Inner Mongolians. *Nat. Microbiol.* **8**, 150–161 (2023).
65. Mahnert, A., Blohs, M., Pausan, M.-R. & Moissl-Eichinger, C. The human archaeome: methodological pitfalls and knowledge gaps. *Emerg. Top. Life Sci.* **2**, 469–482 (2018).
66. Wu-Woods, N. J. et al. Microbial-enrichment method enables high-throughput metagenomic characterization from host-rich samples. *Nat. Methods* **20**, 1672–1682 (2023).
67. Shalon, D. et al. Profiling the human intestinal environment under physiological conditions. *Nature* **617**, 581–591 (2023).
68. Huang, X. D., Tan, H. Y., Long, R., Liang, J. B. & Wright, A.-D. G. Comparison of methanogen diversity of yak (*Bos grunniens*) and cattle (*Bos taurus*) from the Qinghai-Tibetan plateau, China. *BMC Microbiol* **12**, 237 (2012).
69. Dittmann, M. T. et al. Methane emission by *Camelids*. *PloS One* **9**, e94363 (2014).
70. Han, Y. et al. Exploring biomimetic potential of ruminant digestion strategies for lignocellulosic biomass utilization: A comprehensive review. *Renew. Sust. Energ. Rev.* **188**, 113887 (2023).
71. Hess, M. et al. Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science* **331**, 463–467 (2011).
72. Wallace, R. J. et al. The rumen microbial metagenome associated with high methane production in cattle. *BMC Genomics* **16**, 839 (2015).
73. Rubino, F. et al. Divergent functional isoforms drive niche specialisation for nutrient acquisition and use in rumen microbiome. *ISME J.* **11**, 932–944 (2017).
74. Svartström, O. et al. Ninety-nine de novo assembled genomes from the moose (Alces alces) rumen microbiome provide new insights into microbial plant biomass degradation. *ISME J.* **11**, 2538–2551 (2017).
75. Solden, L. M. et al. Interspecies cross-feeding orchestrates carbon degradation in the rumen ecosystem. *Nat. Microbiol.* **3**, 1274–1284 (2018).
76. Li, J. et al. A catalog of microbial genes from the bovine rumen unveils a specialized and diverse biomass-degrading environment. *GigaScience* **9**, giaa057 (2020).
77. Malmuthuge, N., Liang, G. & Guan, L. L. Regulation of rumen development in neonatal ruminants through microbial metagenomes and host transcriptomes. *Genome Biol.* **20**, 172 (2019).
78. Lin, L. et al. Ruminal microbiome-host crosstalk stimulates the development of the ruminal epithelium in a lamb model. *Microbiome* **7**, 83 (2019).
79. Wolff, S. M. et al. Diet shifts provoke complex and variable changes in the metabolic networks of the ruminal microbiome. *Microbiome* **5**, 60 (2017).
80. Salaheen, S. et al. Metagenomic analysis of the fecal microbiomes from *Escherichia coli* O157:H7-shedding and non-shedding cows on a single dairy farm. *Food Control* **102**, 76–80 (2019).
81. Al-Masaudi, S. et al. A metagenomics investigation of carbohydrate-active enzymes along the goat and camel intestinal tract. *Int. Microbiol.* **22**, 429–435 (2019).
82. Tanca, A. et al. Diversity and functions of the sheep faecal microbiota: a multi-omic characterization. *Microb. Biotechnol.* **10**, 541–554 (2017).
83. Weinroth, M. D. et al. Effects of ceftiofur and chlortetracycline on the resistomes of feedlot cattle. *Appl. Environ. Microbiol.* **84**, e00610–e00618 (2018).
84. Li, F., Hitch, T. C. A., Chen, Y., Creevey, C. J. & Guan, L. L. Comparative metagenomic and metatranscriptomic analyses reveal the breed effect on the rumen microbiome and its associations with feed efficiency in beef cattle. *Microbiome* **7**, 6 (2019).
85. Mu, Y. Y., Qi, W. P., Zhang, T., Zhang, J. Y. & Mao, S. Y. Gene function adjustment for carbohydrate metabolism and enrichment of rumen microbiota with antibiotic resistance genes during subacute rumen acidosis induced by a high-grain diet in lactating dairy cows. *J. Dairy Sci.* **104**, 2087–2105 (2021).
86. Gharechahi, J. et al. Metagenomic analysis reveals a dynamic microbiome with diversified adaptive functions to utilize high lignocellulosic forages in the cattle rumen. *ISME J.* **15**, 1108–1120 (2021).
87. Teseo, S. et al. A global phylogenomic and metabolic reconstruction of the large intestine bacterial community of domesticated cattle. *Microbiome* **10**, 155 (2022).
88. Tong, F. et al. The microbiome of the buffalo digestive tract. *Nat. Commun.* **13**, 823 (2022).
89. Xue, M.-Y. et al. Integrated meta-omics reveals new ruminal microbial features associated with feed efficiency in dairy cattle. *Microbiome* **10**, 32 (2022).
90. Xue, M.-Y. et al. Investigation of fiber utilization in the rumen of dairy cows based on metagenome-assembled genomes and single-cell RNA sequencing. *Microbiome* **10**, 11 (2022).
91. Bi, Y. et al. Multiomics analysis reveals the presence of a microbiome in the gut of fetal lambs. *Gut* **70**, 853–864 (2021).
92. Lin, L. et al. Genome-centric investigation of bile acid metabolizing microbiota of dairy cows and associated diet-induced functional implications. *ISME J.* **17**, 172–184 (2022).
93. Gharechahi, J., Sarikhan, S., Han, J. L., Ding, X. Z. & Salekdeh, G. H. Functional and phylogenetic analyses of camel rumen microbiota associated with different lignocellulosic substrates. *NPJ Biofilms Microbiomes* **8**, 46 (2022).

94. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
95. Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).
96. Shen, W., Le, S., Li, Y. & Hu, F. SeqKit: a cross-platform and ultrafast toolkit for fasta/q file manipulation. *PloS One* **11**, e0163962 (2016).
97. Li, H. New strategies to improve minimap2 alignment accuracy. *Bioinformatics* **37**, 4572–4574 (2021).
98. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
99. Kang, D. D., Froula, J., Egan, R. & Wang, Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**, e1165 (2015).
100. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**, 1043–1055 (2015).
101. Olm, M. R., Brown, C. T., Brooks, B. & Banfield, J. F. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* **11**, 2864–2868 (2017).
102. Asnicar, F. et al. Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. *Nat. Commun.* **11**, 2500 (2020).
103. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* **49**, W293–w296 (2021).
104. Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* **38**, 5315–5316 (2022).
105. Hyatt, D. et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinforma.* **11**, 119 (2010).
106. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol. Biol. Evol.* **38**, 5825–5829 (2021).
107. Huerta-Cepas, J. et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* **47**, D309–D314 (2018).
108. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).
109. Kolde, R. Pheatmap: pretty heatmaps. R package version 1, 726 (2012).
110. Kurth, J. M., Op den Camp, H. J. M. & Welte, C. U. Several ways one goal-methanogenesis from unconventional substrates. *Appl. Microbiol. Biotechnol.* **104**, 6839–6854 (2020).
111. Adam, P. S., Borrel, G. & Gribaldo, S. An archaeal origin of the Wood-Ljungdahl H(4)MPT branch and the emergence of bacterial methylotrophy. *Nat. Microbiol.* **4**, 2155–2163 (2019).
112. Zhang, C. J., Pan, J., Liu, Y., Duan, C. H. & Li, M. Genomic and transcriptomic insights into methanogenesis potential of novel methanogens from mangrove sediments. *Microbiome* **8**, 94 (2020).
113. Zhou, Z. et al. Non-syntrophic methanogenic hydrocarbon degradation by an archaeal species. *Nature* **601**, 257–262 (2022).
114. Yang, S. et al. Genomic and enzymatic evidence of acetogenesis by anaerobic methanotrophic archaea. *Nat. Commun.* **11**, 3941 (2020).
115. Buchfink, B., Reuter, K. & Drost, H.-G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **18**, 366–368 (2021).
116. Yin, X. et al. ARGs-OAP v3.0: antibiotic-resistance gene database curation and analysis pipeline optimization. *Engineering* **27**, 234–241 (2022).
117. Liu, B., Zheng, D., Zhou, S., Chen, L. & Yang, J. VFDB 2022: a general classification scheme for bacterial virulence factors. *Nucleic Acids Res* **50**, D912–d917 (2022).
118. Pal, C., Bengtsson-Palme, J., Rensing, C., Kristiansson, E. & Larsson, D. G. J. BacMet: antibacterial biocide and metal resistance genes database. *Nucleic Acids Res* **42**, D737–D743 (2013).
119. Pärnänen, K. et al. Maternal gut and breast milk microbiota affect infant gut antibiotic resistome and mobile genetic elements. *Nat. Commun.* **9**, 3891 (2018).
120. Wu, S. et al. Machine learning aided construction of the quorum sensing communication network for human gut microbiota. *Nat. Commun.* **13**, 3079 (2022).
121. Nayfach, S. et al. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nat. Biotechnol.* **39**, 578–585 (2021).
122. Camargo, A. P. et al. Identification of mobile genetic elements with geNomad. *Nat. Biotechnol.* **42**, 1303–1312 (2023).
123. Shaffer, M. et al. DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res* **48**, 8883–8900 (2020).
124. Spang, A., Caceres, E. F. & Ettema, T. J. G. Genomic exploration of the diversity, ecology, and evolution of the archaeal domain of life. *Science* **357**, eaaf3883 (2017).
125. Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114 (2018).
126. Anderson, C. L. & Fernando, S. C. Insights into rumen microbial biosynthetic gene cluster diversity through genome-resolved metagenomics. *Commun. Biol.* **4**, 818 (2021).

Acknowledgements

We thank Professor Pengfei Liu (Lanzhou University) for his help in the guidance on the framework of the manuscript. We thank Teacher Jingbo Xia from the UTT Art Studio for illustrating the schematic diagram of the ruminant gut in Fig. 4c. The research was funded by National Key Research and Development Program of China (Grant No. 2021YFD1200904 to R.J.L., 2023YFD1301704 to J.D.M., and 2023YFD1300900 to J.D.M.), the Joint Funds of the National Natural Science Foundation of Gansu Province, China (Grant No. 23JRRA1514 to J.D.M.), and Program of Coordinated Research Activities from IAEA (CRP D31031 to R.J.L.). This study was also supported by the International Collaboration 111 Programme (BP0719040 to R.J.L.). We would like to express our gratitude to the Supercomputing Center of Lanzhou University for their valuable support in the computation works.

Author contributions

J.D.M., X.P.J. and R.J.L. conceived and designed the study. J.D.M. and C.X.M. provided the methodology and analysis. J.D.M., F.Y.S. and X.Y. collected the data. J.D.M., X.P.J., Z.C., Y.W.Y., A.K., W.W.W. and R.J.L. prepared and revised the manuscript.

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/s41467-024-54025-3.

Correspondence and requests for materials should be addressed to Jiandui Mi or Ruijun Long.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.