

Upload this completed form to website with submission

ARTICLE INFORMATION	Fill in information in each box below
Article Type	Research article
Article Title (within 20 words without abbreviations)	Comparative pangenome analysis of methanogenic archaea from diverse ecosystems reveals potential targets for methane mitigation in rumen microbiome
Running Title (within 10 words)	Pangenome for methane mitigation in rumen microbiome
Author	Jeongkuk Park ^{1,a} , Ki Beom Jang ^{1,a} , Min-Geun Kang ¹ , Junhyeok Kyung ¹ , Jihyun Yoon ¹ , Sangdon Ryu ² , and Younghoon Kim ^{1,*}
Affiliation	1Department of Agricultural Biotechnology, Research Institute of Agriculture and Life Science, Seoul National University, Seoul 08826, South Korea 2Honam National Institute of Biological Resources, Mokpo 58762, Korea
ORCID (for more information, please visit https://orcid.org)	Jeongkuk Park (https://orcid.org/0000-0003-0743-4795) Ki Beom Jang (https://orcid.org/0000-0001-6794-9569) Min-Geun Kang (https://orcid.org/0000-0002-2204-6443) Junhyeok Kyung (https://orcid.org/0009-0004-3596-6769) Jihyun Yoon (https://orcid.org/0009-0007-9928-2480) Sangdon Ryu (https://orcid.org/0000-0001-5338-8385) Younghoon Kim (https://orcid.org/0000-0001-6769-0657)
Competing interests	No potential conflict of interest relevant to this article was reported.
Funding sources State funding sources (grants, funding sources, equipment, and supplies). Include name and number of grant if available.	This work was supported by the National Research Foundation of Korea (NRF) grants funded by the Korean government (MEST; NRF-2021R1A2C3011051) and by Basic Science Research Program through the NRF funded by the Ministry of Education (RS-2025-25418389).
Acknowledgements	Not applicable.
Availability of data and material	Upon reasonable request, the datasets of this study can be available from the corresponding author.
Authors' contributions Please specify the authors' role using this form.	Conceptualization: Kim Y, Park J. Data curation: Jang KB, Park J. Formal analysis: Jang KB, Park J. Methodology: Jang KB, Park J. Software: Jang KB, Park J. Validation: Jang KB, Park J, Kang MG, Kyung J, Yoon J, Ryu S. Investigation: Jang KB, Park J, Kang MG, Kyung J, Yoon J, Ryu S. Writing - original draft: Jang KB, Park J, Kang MG, Kyung J, Yoon J, Ryu S.. Writing - review & editing: Jang KB, Park J, Kang MG, Kyung J, Yoon J, Ryu S.
Ethics approval and consent to participate	This manuscript does not require IRB/IACUC approval because there are no human and animal participants.

CORRESPONDING AUTHOR CONTACT INFORMATION

For the corresponding author (responsible for correspondence, proofreading, and reprints)	Fill in information in each box below
First name, middle initial, last name	Younghoon Kim
Email address – this is where your proofs will be sent	ykeys2584@snu.ac.kr
Secondary Email address	ykeys2584@snu.ac.kr
Address	Department of Agricultural Biotechnology and Research Institute of Agriculture and Life Science, Seoul National University, Seoul 08826, South Korea
Cell phone number	+82-10-4135-2584
Office phone number	+82-02-880-4808
Fax number	+82-02-880-4821

1

ACCEPTED

Running head: Pangenome for methane mitigation in rumen microbiome

**Comparative pangenome analysis of methanogenic archaea from diverse
ecosystems reveals potential targets for methane mitigation in rumen
microbiome**

Jeongkuk Park^{1,a}, Ki Beom Jang^{1,a}, Min-Geun Kang¹, Junhyeok Kyung¹, Jihyun Yoon¹, Sangdon Ryu²,
and Younghoon Kim^{1,*}

¹Department of Agricultural Biotechnology, Research Institute of Agriculture and Life Science, Seoul
National University, Seoul 08826, South Korea

²Honam National Institute of Biological Resources, Mokpo 58762, Korea

^aThese authors contributed equally to this work

*To whom correspondence should be addressed: ykeys2584@snu.ac.kr

ABSTRACT

Rumen methanogenesis is a major biological contributor to methane emissions in ruminants, yet the extent to which functional markers align with taxonomic relationships and how genome content varies across habitats, remains poorly resolved. In this study, we integrated broad phylogenetic frameworks with pangenome-resolved analysis to characterize methanogenic archaea from diverse ecosystems, including seawater, freshwater, sewage, rumen, human gut, soil, and cockroach sources. By combining these insights with pangenome reconstruction and KEGG-based pathway mapping of methanogenesis, we reveal key evolutionary and functional patterns. Notably, phylogenies based on 16S rRNA and *mcrA* genes showed limited concordance: only two clades exhibited overlap between trees, with most clustering patterns lacking environmental specificity. This discrepancy reflects the deep conservation of 16S rRNA compared with the evolutionary plasticity of *mcr* genes, shaped by lateral gene transfer, gene loss, and pathway modularity. The pangenome comprised of 8,695 orthogroups across 71 genomes, with core and soft-core genes enriched in translation, amino acid metabolism, and coenzyme biosynthesis, while the shell contained many poorly annotated orthogroups, highlighting annotation gaps in archaeal genomes. KEGG analysis revealed habitat-specific signatures: rumen methanogens were notably depleted in genes of the acetyl-CoA pathway, whereas human gut methanogens lacked key cofactor biosynthesis modules, including those for coenzymes M, B, F₄₂₀, and methanofuran. From rumen-derived shotgun metagenomes, we identified 53 methane-producing, 4 canonical methanogenic, 10 potential competitor, and 1 methanotrophic metagenome-assembled genomes (MAGs) based on functional gene content. Competitor candidates included nitrate-reducing and Wood–Ljungdahl pathway (WLP)-utilizing acetogens, suggesting hydrogen redirection under high-hydrogen or inhibitor conditions. These findings support a functional marker strategy that integrates 16S rRNA with pathway-specific genes and a pangenome framework to enhance ecological interpretations of methanogens and to prioritize potential targets for methane mitigation in ruminants.

Keywords: Comparative pangenome, rumen microbiome, MAGs, methanogen, methane mitigation

INTRODUCTION

Comprehensive efforts have been directed toward elucidating the ecological and metabolic features of methanogenic archaea to inform effective strategies for mitigating methane emissions in livestock production systems. Methanogenic archaea inhabit a range of anoxic environments including the rumen, marine sediments, wetlands, and deep-sea hydrothermal vents, where their dominant methanogenic pathways, primarily hydrogenotrophic and methylotrophic, vary according to ecological context [1, 2]. However, many methanogens remain difficult to isolate and cultivate under laboratory conditions, resulting in a knowledge base that is skewed toward *in vitro* data from non-rumen habitats and remains fragmentary with respect to rumen-specific biology [3]. Within the rumen, methane production is strongly influenced by methanogen diversity, hydrogen flux, and syntrophic interactions with bacterial guilds, yet the phylogenetic and functional relationships remain poorly defined [4, 5].

Comparative pangenome analysis provides a genome-wide framework for dissecting these relationships, enabling gene-level resolution of phylogenetic, functional, and metabolic divergence across diverse taxa. Notably, the Zoonomia Consortium's alignment of 240 mammalian genomes has established a methodological benchmark for large-scale comparative analysis, integrating large-scale genome alignment, normalization, and statistical inference in a unified analytical pipeline [6]. Building upon such advances in large-scale genome comparative analysis, similar efforts in the field of animal microbiology have followed suit. The Hungate1000 project provided foundational resources by sequencing 410 cultured rumen microbial genomes, which were later expanded by the assembly of 913 metagenome-assembled genomes (MAGs) from bovine rumen metagenomes and the construction of a 4,941-member Rumen Uncultured Genome catalogue, expanding the landscape of species- and function-level variation and strengthening quantitative links to methanogenesis pathways [7, 8]. More recently, genome-resolved surveys have expanded archaeal resources and environmental breadth, including a catalogue of 998 unique ruminant-gut archaeal genomes across 10 host species and large MAG datasets from Nelore cattle, which together sharpen genome-level resolution of rumen methanogenesis [9, 10]. Despite these advances,

several critical questions remain unanswered. It still lacks a genome-resolved understanding of how methanogen gene repertoires vary across habitats and how these differences map onto phylogenetic lineages, functional capacities, and methanogenic strategies within the rumen. Furthermore, there is no integrated framework that connects methanogen genomic features with co-occurring microbial guilds that either supply reductants or act as metabolic competitors within the methanogenesis network.

To address these knowledge gaps, we employed a two-stage comparative genomics approach integrating phylogenetic reconstruction, pangenome analysis, and metagenomic profiling to systematically characterize rumen methanogens and their ecological interactions. Initially, we constructed phylogenetic trees using 16S rRNA and the key methanogenesis marker gene *mcrA* from methanogens isolated across diverse ecosystems, and performed habitat-stratified pangenomic comparisons to identify gene-level signatures and clustering patterns of each habitat. And then, we investigated rumen shotgun metagenomic data to delineate candidate methane substrate producers, consumers, and competitor lineages within the methanogenic network, thereby nominating functionally relevant targets for methane mitigation in rumen ecosystems [11].

MATERIALS AND METHODS

Genome collection and phylogenetic distribution

Genomic sequences were retrieved primarily from GenBank, with curated versions cross-validated via RefSeq. Accession numbers, version identifiers, BioProject, BioSample, TaxID, and download dates are detailed in Supplementary Table S1. The final dataset included 71 methanogen reference genomes, spanning 10 from the rumen, 4 from the human gut, 17 freshwater, 14 sewage, 20 seawater, 4 soil, and 2 cockroach-derived strains. For 16S rRNA-based phylogeny, rRNA loci were predicted and extracted using Barrnap v0.9, followed by coordinate-based trimming to remove incomplete ends. Multiple sequence alignment was performed using MAFFT v7 [12]. For functional phylogeny, the *mcrA* gene (encoding methyl-coenzyme M reductase subunit A) was detected via HMMER v3 searches against the corresponding Pfam domain, retaining only full-length hits in line with established use of *mcrA* as a methanogen marker. The resulting amino-acid sequences were aligned with MAFFT v7 [13-15]. Maximum-likelihood phylogenies were inferred using IQ-TREE v2, and annotated phylogenetic trees were visualized with iTOL v6 [16, 17].

Pangenome analysis

Pangenome analysis from diverse ecosystems was conducted following established protocols with minor modifications [11]. The analytical workflow is summarized in Figure 1A. Protein-coding genes were predicted and functionally annotated from assembled genomes with Prokka v1.14.6 [18]. The resulting translated proteomes were used for orthogroup inference with OrthoFinder v3.1.0, and orthogroups (COG) presence-absence matrices were used to delineate core, shell and cloud gene sets following standard pangenome definitions [19]. Functional annotation of protein sequences was assigned using eggNOG-mapper v2.1.13 against the eggNOG v5 orthology database. KEGG Orthology (KO) and Gene Ontology (GO) assignments derived through the eggNOG framework were used to functionally characterize orthogroups [20, 21]. The combined outputs from OrthoFinder and eggNOG-mapper

provided gene clusters with consolidated functions, which formed the basis for downstream interpretation of methanogen pangenome structure across environments [19]. Data visualization was carried out using Seaborn v0.13.2 and Matplotlib v3.10.6 [22, 23]. The COG were scored as present or absent per strain and classified by prevalence as follows: core, $\geq 99\%$ of genomes; soft-core, $\geq 95\%$ and $< 99\%$; shell, $\geq 15\%$ and $< 95\%$; cloud, $< 15\%$ [24]. Results for the cloud set were omitted.

KEGG mapping of methanogenesis pathway

From eggNOG-mapper v2.1.13 annotations of the methanogen reference genomes used in the pangenome analysis, KEGG Orthology assignments linked to the methanogenesis pathway ko00680 were retrieved and compiled. A KO was deemed present if at least one encoded protein was annotated with the respective identifier. Environment-specific KO frequency was calculated as the proportion of genomes in each habitat containing the KO. The resulting KO-by-environment matrix was visualized both as pathway tile maps and matrix plots using Matplotlib v3.10.6.

Functional analysis of rumen MAGs

Schematic workflows of rumen MAGs functional analysis were illustrated in Fig. 1B. Shotgun metagenome reads were obtained from dairy cow rumen fluid sample sourced from previously published datasets by Kang et al [25]. Long-read metagenome assembly was performed with metaMDBG v1.2 [26]. Read mapping and coverage profiling were performed using minimap2 v2.30. Binning was conducted with SemiBin2 v2.2.0 [27]. Genome quality was evaluated with CheckM2 v1.1.0 and contamination was filtered using GUNC v1.0.6. Bins with $\geq 50\%$ completeness and $< 10\%$ contamination were retained [28-30]. Nonredundant representatives were obtained by dereplication with dRep v3.6.2 [31]. Taxonomic assignment used GTDB-Tk v2.4.1 with GTDB reference data release r226 [32]. Functional annotation of predicted protein sequences was performed using eggNOG-mapper v2.1.13. MAGs were categorized into four ecological roles using curated marker genes.

Methanogens were identified as archaeal MAGs encoding the methyl-coenzyme M reductase operon (*mcrA/B/G*), optionally supported by *mtrA/B/E*, which are canonical to methanogenic energy metabolism [33, 34]. Substrate supply producers were defined as MAGs encoding one or more routes that provide key methanogenic substrates: H₂ production via [FeFe]-hydrogenase (*hydA*) together with its maturation genes (*hydE/F/G*); formate production via pyruvate-formate lyase (*pflA/pflB*); acetate production via the phosphotransacetylase–acetate kinase pair (*pta/ackA*) [35, 36]. Competitive sinks-competitors captured respiratory pathways that divert the same reductants (H₂/electrons) away from methanogenesis. Dissimilatory nitrate reduction to ammonium (DNRA) was identified by *nrfA/H*. The reductive Wood–Ljungdahl pathway (WLP) evidenced by *cooS* (CODH), with *fhs* and *metF* treated as supportive folate-branch markers rather than strict requirements [37-39]. Methanotrophs were identified by the presence of the particulate methane monooxygenase gene *pmoA*, a widely used functional and phylogenetic marker [40, 41]. Out of 903 initial bins from SemiBin2, 151 passed CheckM2 filtering, 116 passed GUNC quality control, and 106 dereplicated MAGs remained. Ultimately, 67 MAGs with functional roles were used in downstream analysis.

Statistical analysis

Differential distribution of orthogroups by clade and habitat was assessed using two-sided Fisher's exact tests (2×2 contingency tables) with Benjamini–Hochberg FDR correction ($q \leq 0.05$). Enrichment was interpreted via odds ratios (OR): OR > 1 denoting enrichment, OR < 1 indicating depletion. Clade-specific orthogroups were defined as those present in $\geq 80\%$ of strains within a clade and $\leq 20\%$ outside it [42, 43].

RESULTS

16S rRNA and *mcrA* phylogeny reveal limited clade concordance

Comparative phylogenetic analysis based on 16S rRNA and *mcrA* sequences identified two clades exhibiting substantial overlap, with Jaccard indices exceeding 0.7. The first clade, comprising strains from freshwater and seawater habitats, showed a Jaccard index of 0.75 and included NC_009051.1_Methanoculleus_marisnigri_JR1, NC_009712.1_Methanoregula_boonei_6A8, NZ_AP019781.1_Methanoculleus_chikugoensis_strain_MG62, NZ_CP109831.1_Methanoculleus_submarinus_strain_DSM_15122, NZ_CP113361.1_Methanogenium_organophilum_strain_DSM_3596, and NZ_JOMF01000012.1_Methanomicrobium_mobile_DSM_1539. A second clade, spanning freshwater and sewage-derived isolates, yielded a Jaccard index of 0.714 and consisted of NC_007796.1_Methanospirillum_hungatei_JF-1, NC_018227.2_Methanoculleus_bourgensis_MS2, NC_019943.1_Methanoregula_formicica_SMSP, NZ_CP036172.1_Methanofollis_aquaemaris_strain_N2F9704, and NZ_CP091092.1_Methanomicrobium_antiquum_strain_DSM_21220.

No additional clades showed concordance above the 0.7 threshold. The lack of clade concordance between the relatively conserved 16S rRNA tree and the *mcrA* tree indicates accelerated sequence divergence in *mcrA*. Moreover, the failure of environment-based clustering to persist on the *mcrA* phylogeny points to frequent horizontal gene transfer events and inherent diversification of this functional marker (Fig. 2).

Pangenome architecture and functional classification of COG

From the 71 methanogen genomes analyzed, a total of 8,695 orthogroups were identified and classified into core ($\geq 99\%$ genomes), soft-core ($\geq 95\%$ and $< 99\%$), and shell ($\geq 15\%$ and $< 95\%$) components. This yielded 385 core, 94 soft-core, and 2,573 shell orthogroups. The most abundant COG

categories within each group were as follows: J (translation, ribosomal structure and biogenesis; 35.4%), E (amino acid transport and metabolism; 12.8%), and H (coenzyme transport and metabolism; 12.2%) for core; H (22.1%), E (19.5%), and J (10.4%) for soft-core; and C (energy production and conversion; 16.9%), K (transcription; 8.9%), and P (inorganic ion transport and metabolism; 8.8%) for shell (Fig. 3B). These distributions align with prior archaeal pangenomic surveys, emphasizing conserved functions in translational machinery and cofactor metabolism across core genomes [44]. In contrast to typical bacterial patterns, where J, E, and H frequently appear alongside C and F in the core at comparable proportions, the methanogen set here places energy metabolism and nucleotide transport and metabolism predominantly outside the core. The core was enriched for COG J, E, and H, contrasting with the bacterial tendency to balance J, E, and H with C and F at comparable proportions. Therefore, these results support strong conservation of the information-processing machinery in archaea and suggest an elevated fraction of poorly annotated proteins reflecting limited study coverage [44-46]. Additionally, rare or composite COG, including IQ, FG, NU, DJ, DZ, EGP, BQ, and A, were uniquely found within the shell COG. Within the rumen-enriched clade including NZ_CP118753.2_Methanosphaera_sp._ISO3-F5; NC_013790.1_Methanobrevibacter_ruminantium_M1; NZ_CP014265.1_Methanobrevibacter_olleyae_strain_YLM1; NZ_FMXB01000001.1_Methanobrevibacter_millerae_strain_DSM_16643; NC_009515.1_Methanobrevibacter_smithii_ATCC_35061; NZ_BAGX02000054.1_Methanobrevibacter_boviskoreani_JH1, enrichment testing at $q \leq 0.05$ showed COG J to be clade-specifically depleted (OR = 0.08), whereas COG M (Cell wall/membrane/envelope biogenesis) was relatively enriched (OR = 2.46) (Fig. 3C). The enrichment of cell wall, membrane and envelope functions in the rumen methanogen clade, relative to other clades, supports the interpretation that these taxa engage in potential ecological and physical interactions with coexisting rumen microbiome [47].

Environment-specific patterns in methanogenesis gene retention

KEGG pathway ko00680 was used to assess methanogenesis gene distribution. The methyl-coenzyme M reductase complex, a hallmark of methanogenic metabolism, was universally present across all genomes. Other core methanogenesis enzymes such as heterodisulfide reductase and ferredoxin reductase, also showed complete conservation across environments. However, human gut-derived methanogens exhibited substantial depletion of genes involved in the biosynthesis of coenzyme M, coenzyme B, coenzyme F₄₂₀, and methanofuran, including K00200, K00201, K00202, K00203, K00205, K00319, K00320, K00441, K00577, K00578, K00579, K00580, K00581, K00584, K00672, K01499, K06914, K07072, K07144, K09733, K11212, K11780, K11781, K14941, K16792, K16793, and K18933. This pattern is an environment-specific signature of human gut methanogens and indicates a strong dependence on community-supplied intermediates. In contrast, rumen-derived methanogens were notably depleted in genes involved in the acetyl-CoA pathway genes, including K00192 (OR = 0.1), K00193 (OR = 0.08), and K00194 (OR = 0.06). The depletion of acetyl-CoA pathway genes in rumen methanogens indicates a reduced role for acetoclastic methanogenesis. This suggests that hydrogen is competitively diverted by coexisting rumen microbiome (e.g. acetogen) via the WLP, potentially limiting methane yield from acetate oxidation (Fig. 4).

Functional categorization of methanogenesis-related genes in rumen MAGs.

From long-read metagenomic assemblies of rumen fluid, 67 high-quality MAGs were assigned to ecological roles based on methanogenesis-related gene content (Fig. 5). Genomes carrying genes that can supply substrates for methanogenesis were designated producers; archaeal genomes encoding the core methanogenesis machinery were designated methanogens; genomes encoding pathways that competitively consume methanogenesis substrates were designated competitors; genomes encoding methane oxidation were designated methanotrophs. These included 53 producers, 4 methanogens, 10 competitors, and 1 methanotroph.

Among competitor candidates, DNRA MAGs were identified as Bin.31__s_Aristaeella_sp900315675, Bin.35__s_UBA3792_sp002369195, Bin.38__s_UBA3792_sp902792295,

Bin.39__s_Faecousia_sp900315595, Bin.41__s_Bulleidia_intestinalis,
Bin.75__s_Denitrobacterium_detoxificans, Bin.122__s_Aristaeella_sp900322155, and
Bin.126__s_Chordicoccus_sp900320575. The thermodynamically favorable nature of DNRA over
methanogenesis positions these taxa as key electron sinks in methane-suppressing communities. DNRA
offers a more favorable free-energy change than the reduction of CO₂ to methane, enabling more efficient
withdrawal of electrons. Rumen methane-mitigation strategies that supplement nitrate operate on this
principle. The DNRA competitor candidates identified here support the feasibility of suppressing
methanogenesis through electron competition [48]. The WLP was detected in Bin.4__g_FB2012 and
Bin.73__s_Ruminococcus_sp002394695, suggesting these taxa divert hydrogen away from
methanogenesis (Fig. 5). Given the depletion of acetoclastic methanogenesis in rumen methanogens, we
prioritized taxa that can divert hydrogen as mitigation candidates, and the genus *Ruminococcus* appears to
be a promising candidate for hydrogen competition in the rumen ecosystem.

DISCUSSION

This study includes two major components, an environment-based analysis of methanogenic archaea and a rumen shotgun MAG analysis. For the methanogen component, we combined established elements from prior workflows. The phylogenetic reconstruction followed the analysis flow of Ou et al. [49], and the pangenome analysis followed the framework of Prondzinsky et al. [11]. Environmental classification was assigned from the isolation context of each reference genome, allowing us to examine genetic features of methanogens alongside habitat-specific patterns. For the rumen MAG component, the assembly, mapping, and binning steps were guided by the Oxford Nanopore long-read shotgun workflow, which is well-suited for processing Oxford Nanopore Technologies data. Quality filtering and curation followed the methods of Richy et al. [50], and functional annotation was performed using eggNOG-mapper [20], consistent with the methanogen analysis. Although the rumen MAG pipeline is not directly derived from a single paper, once MAGs are recovered by binning, the subsequent interpretation is tool-independent, following the curation and QC sequence proposed by Richy et al. [50]. Similarly, eggNOG-mapper provides a complementary route to functional annotation in lineages, such as methanogens, where curated annotations remain limited. Therefore, these choices support the reproducibility of the workflow employed in this study.

Methanogenic archaea play a central role in regulating hydrogen flux and methane output across anaerobic ecosystems [1]. However, conventional reliance on cultivation-dependent methods and single-marker analyses has left substantial gaps in our understanding of their ecological differentiation and evolutionary dynamics [7]. To address this, we integrated broad phylogenetic analyses with genome-resolved comparisons to explore how environmental context shapes methanogen lineage structure and pathway composition, and how these patterns could inform methane-mitigation strategies [51]. The pangenome-based genomic analysis examines lineage-specific and niche-linked gene clusters that are involved in microbial metabolism, particularly methanogenesis, thereby identifying candidate genomic and metabolic targets for methane mitigation [52]. In line with this perspective, a methanogen pangenome

analysis resolved a large conserved core together with habitat relevant accessory variability, providing a rational map to distinguish conserved enzymatic nodes from context-specific accessory modules that are pertinent to mitigation [11]. Advancing high-quality genome references and refining functional annotations will connect genotype to mechanism in rumen systems, supporting the design of targeted, microbiome-informed strategies to reduce methane emissions in livestock production [8].

The *mcrA* is a canonical component of methanogenesis and has served as a defining molecular marker for methanogens. Earlier studies suggested that *mcrA* sequences could substitute for 16S rRNA in taxonomic classification because both markers produced largely congruent phylogenetic trees [15, 53]. However, more recent work has revealed substantial discordance between *mcrA*- and 16S-based phylogenies, influenced by environmental niche, taxonomic scope, primer design, and analytical framework [54]. These findings underscore the limitations of using any single functional gene as a universal proxy for evolutionary inference [55, 56]. In other words, using a single functional gene as a representative marker to describe evolutionary relationships or to conduct phylogenetic analysis among different organisms can be inappropriate. Our results strongly reinforce this point: except for two minor clades (one comprising 5 methanogens from sewage and freshwater, and another comprising 6 methanogens from freshwater and seawater), no meaningful congruence was observed between 16S rRNA and *mcrA* trees (Fig. 2). This limited overlap suggests that phylogenetic resolution in methanogens operates primarily at the genus level and that evolutionary patterns are not clearly partitioned by habitat. The divergence reflects the conservative evolutionary trajectory of 16S rRNA compared with the more dynamic *mcrA*, which is subject to lateral gene transfer, modular pathway organization, and selective gene loss [57, 58]. Supporting this, functional analyses revealed that methanogens exhibited low retention of acetoclastic methanogenesis genes, whereas human gut methanogens showed marked depletion in cofactor biosynthesis modules for coenzymes M, B, F₄₂₀, and methanofuran (Fig. 4). These pathway differences argue against a single canonical methanogenesis route across habitats and help explain why *mcrA*-based phylogeny alone does not recapitulate 16S rRNA relationships. Therefore, the phylogenetic and genome content results support a marker strategy that integrates 16S rRNA with multiple pathway

genes and genome-resolved context, rather than relying on *mcrA* alone, for robust taxonomic inference and for ecological interpretation across environments.

Because *mcrA* did not reliably capture environment-specific structure in our phylogenies, we integrated pangenome-based clade definitions with environmental metadata to track niche-linked genomic adaptations. The methanogen pangenome revealed a clear hierarchical organization, with core and soft-core gene sets enriched in housekeeping functions such as translation (COG J), amino acid metabolism (COG E), and coenzyme metabolism (COG H) (Fig. 3A). This pattern is in accordance with previous archaeal pangenomic studies that reported similar enrichment of informational processes in conserved gene sets [59]. The archaeal distribution contrasts with typical bacteria, where J, E and H tend to occur alongside C and F for energy metabolism and nucleotide transport and metabolism at comparable proportions [45]. These results indicate strong conservation of transcription, translation, and replication in archaea, and also suggest a higher fraction of proteins lacking confident annotation due to limited study coverage [44, 46]. The shell gene set further contained numerous poorly annotated or lineage-specific orthogroups, highlighting persistent annotation gaps within archaeal genomics. Within the rumen-derived methanogen clade, the shell showed enrichment for cell wall, membrane, and envelope biogenesis, consistent with tight attachment to symbiotic bacteria in the rumen and with sustained requirements for maintenance and remodeling of archaeal cell envelopes based on pseudomurein and S-layers [60] (Fig. 3B). This interpretation aligns with prior observations of *Methanobrevibacter ruminantium* M1, which encodes multiple adhesin-like proteins such as Mru_1499 capable of binding directly to protozoa and hydrogen-donating bacteria [61]. The ether lipid membrane typical of archaea also requires compositional maintenance under volatile fatty acids, long-chain fatty acids, and osmotic stress prevalent in the rumen, which helps explain the relative expansion of membrane and envelope biosynthesis pathways [47]. The observed depletion of COG J outside the core aligns with the broader archaeal trend, which preserves translation and ribosome biogenesis functions as deeply conserved core features [62].

Methanogenesis is generally categorized into hydrogenotrophic, acetoclastic, and methylotrophic pathways [63]. Our KEGG-based screening revealed that genes supporting methylotrophic

methanogenesis were retained at relatively low frequencies (mean ratio 0.32) across environments (Fig. 4), consistent with previous reports indicating that hydrogenotrophic methanogenesis is the dominant route, with methylotrophy operating as a niche-specific, facultative strategy [64]. The result reflects gene presence only, and previous studies show that methylotrophic methanogenesis can become prominent in settings with abundant methylated substrates [1, 65]. Among environment-specific distinctions, human gut methanogens exhibited pronounced depletion of cofactor biosynthetic genes associated with coenzymes M, B, F₄₂₀, and methanofuran. This is consistent with metabolic streamlining during host-associated adaptation [66]. Human gut-associated *Methanomassiliicoccales* exemplify this trend, with loss of canonical coenzyme M biosynthesis genes and atypical energy conservation modules, as well as loss of the Wood–Ljungdahl methyl branch, supporting a hydrogen-dependent methyl-reducing lifestyle that increases reliance on community-supplied intermediates [67, 68]. Alternative or yet-unresolved routes for coenzyme M formation and exogenous supply have been proposed, which together suggest that human gut methanogens may depend more on community interdependence and pathway substitution than on strict *de novo* cofactor synthesis [66, 69].

Interestingly, depletion of acetyl-CoA pathway genes was clearly observed in rumen-derived methanogens (Fig. 4). This pattern supports that hydrogenotrophic methanogenesis predominates in the rumen and that acetoclastic activity is low [70]. It consists of the dominance of *Methanobrevibacter*, which relies primarily on the hydrogenotrophic route, and the low prevalence of *Methanosarcinales*, which can perform acetoclastic methanogenesis. Together, these features indicate limited acetate use in rumen methanogenesis. The deficit of acetyl-CoA pathway genes in rumen methanogens is therefore best interpreted as an outcome of ecological selection pressures [71, 72]. Acetate has been considered a substrate for acetoclastic methanogenesis, which led to the view that acetate-forming bacteria hinder methane mitigation [73]. Recent studies instead highlight competition for metabolic hydrogen as the main control [74, 75]. When methanogenesis is suppressed, electron flow is redirected toward alternative hydrogen sinks such as propionate formation and reductive acetogenesis, with accompanying increases in volatile fatty acid yields. Consistent with this mechanism, recent genome-resolved and metabolite-

profiling work shows that 3-nitrooxypropanol (3-NOP) cuts methane by ~60% while stimulating reductive acetogenesis and shifting SCFA and H₂ dynamics without depressing intake [76]. A previous meta-analysis also indicated consistent shifts toward propionate and activation of acetogenic lineages, while acetoclastic methanogenesis appears uncommon in the rumen [77]. Taken together, these results support positioning acetate production as part of a reallocation of hydrogen sinks among the competitive routes that can antagonize methanogenesis under inhibition regimes [78].

In the genome-resolved functional classification based on rumen shotgun sequencing data, competitors were divided into a DNRA group that draws electrons away from methanogenesis and a WLP group that competes for H₂. Thermodynamically, DNRA offers a larger free-energy gain than CO₂ reduction. Under standard conditions, the reduction of CO₂ to CH₄ is on the order of -131 kJ mol⁻¹, whereas the reduction of NO₃⁻ to NH₃ is on the order of -600 kJ mol⁻¹, making DNRA a stronger electron sink [79]. This thermodynamic advantage intensifies competition for the reductants used in methanogenesis, and nitrate has been widely evaluated as a rumen methane mitigation agent [80]. Recent *in vitro* and applied studies corroborate nitrate-driven methane suppression with concomitant microbiome and fermentation shifts, including dose-dependent responses and cation-specific effects [81]. Prior studies reported methane reductions with *Denitrobacterium* supplementation alone, in combination with nitrate, and with the combined treatment showing a greater effect than either alone, consistent with electron diversion via DNRA by *Denitrobacterium* [48, 82]. Consistent with this, our analysis detected DNRA genes in genera such as *Aristaeella*, *Faecousia*, *Bulleidia*, and *Chordicoccus* (Fig. 5). Although direct evidence linking these genera to DNRA remains limited and precludes immediate designation as nitrate reducers, their gene content highlights them as competitor candidates for future methane mitigation assays. In the rumen, the WLP is thermodynamically and kinetically disadvantaged while methanogenesis dominates, yet it can emerge as an alternative H₂ sink when methanogenesis is suppressed or when H₂ partial pressure increases [83]. Classical incubations and experiments with methanogenesis inhibitors such as BES repeatedly showed activation of indigenous acetogens and increased acetate production under methanogenesis suppression, supporting an inhibition-induced shift toward WLP [84, 85]. More

recent meta-omics studies indicate the presence of acetogenic lineages not only within Lachnospiraceae but also within Ruminococcaceae, with enrichment of WLP marker genes and acetate formation when methanogenic pressure is reduced [86, 87]. In the present study, the detection of Ruminococcaceae MAGs including *fhs* and *cooS* as WLP markers suggests potential facultative switching to an H₂-sink role, even though these taxa are generally recognized as primary fermenters that produce H₂ (Fig. 5). Unlike DNRA, WLP does not depend on an external electron acceptors, but under localized H₂ accumulation or methanogenesis inhibition, it can operate as a competing electron sink and reroute H₂ flow in ways that contribute to methane mitigation [88].

Our findings confirm that the comparative genome-resolved approach is a powerful tool for identifying a targeted strategy to mitigate ruminant methane production. Integrating phylogenetic and pangenome information allows the approach to focus on a practical set of organisms and modules for experimental validation. The WLP, identified as a novel candidate, provides a concrete establishment for targeted assays and follow-up experiments. These findings enhance our genome-resolved understanding of methanogens and reveal how their metabolic pathways vary in response to ecological niche differentiation. However, functional annotation of archaeal genomes remains incomplete because there are still uncharacterized proteins compared to bacteria, which limits pathway-level inference and experimental validation due to the limited cultivability of methanogens. Coordinated multiomics workflows, integrating expanded archaeal genome references with culturomics, metatranscriptomics, and metabolomics, will be essential to confirm whether the predicted roles are expressed under actual rumen conditions. Systematic curation of archaeal functional databases and targeted validation of key modules highlighted by pangenome analyses may enhance annotation precision and strengthen the mechanistic relation between methanogen genomics and methane-mitigation strategies, thereby supporting the development of more effective and sustainable interventions in livestock production.

Summary and Conclusion

In this study, we integrated broad phylogenetic reconstruction with pangenome-resolved analyses to interrogate the diversity and functional attributes of methanogens across diverse ecosystems. By constructing both 16S rRNA- and *mcrA*-based phylogenies and coupling them with pangenome comparisons, we assessed habitat-specific clustering and identified gene-level signatures associated with environmental adaptation. Using rumen shotgun sequencing metagenomes, we determined putative methanogenic producers, hydrogen competitors, and methanotrophs within the methane network and prioritized candidate taxa for *in vitro* methane mitigation strategies. In contrast, rumen methanogens are enriched for genes linked to cell envelope biogenesis, consistent with sustained physical interactions in the rumen. Furthermore, the observed depletion of acetyl-CoA pathway genes in rumen methanogens suggests limited acetoclastic activity and raises the potential for hydrogen redirection via the WLP in non-archaeal partners. These findings support a functional marker strategy that integrates 16S rRNA with pathway-specific genes and a pangenome framework to enhance ecological interpretations of methanogens and to prioritize potential targets for methane mitigation in ruminants.

426 **Acknowledgments**

427 This work was supported by the National Research Foundation of Korea (NRF) grants funded by the
428 Korean government (MEST; NRF-2021R1A2C3011051) and by Basic Science Research Program
429 through the NRF funded by the Ministry of Education (RS-2025-25418389).

ACCEPTED

REFERENCES

1. Bueno de Mesquita CP, Wu D, Tringe SG. Methyl-based methanogenesis: an ecological and genomic review. *Microbiology and Molecular Biology Reviews*. 2023;87(1):e00024-22.
2. Wen X, Yang S, Horn F, Winkel M, Wagner D, Liebner S. Global biogeographic analysis of methanogenic archaea identifies community-shaping environmental factors of natural environments. *Frontiers in microbiology*. 2017;8:1339.
3. Costa KC, Whitman WB. Model organisms to study methanogenesis, a uniquely archaeal metabolism. *Journal of bacteriology*. 2023;205(8):e00115-23.
4. 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. *Journal of Dairy Science*. 2022;105(10):8569-85.
5. Beauchemin KA, Ungerfeld EM, Eckard RJ, Wang M. Fifty years of research on rumen methanogenesis: Lessons learned and future challenges for mitigation. *Animal*. 2020;14(S1):s2-s16.
6. Diane P. Genereux AS, Joel Armstrong, Jeremy Johnson, Voichita D. Marinescu, Eva Murén, David Juan, Gill Bejerano, Nicholas R. Casewell, Leona G. Chemnick, Joana Damas, Federica Di Palma, Mark Diekhans, Ian T. Fiddes, Manuel Garber, Vadim N. Gladyshev, Linda Goodman, Wilfried Haerty, Marlys L. Houck, Robert Hubley, Teemu Kivioja, Klaus-Peter Koepfli, Lukas F. K. Kuderna, Eric S. Lander, Jennifer R. S. Meadows, William J. Murphy, Will Nash, Hyun Ji Noh, Martin Nweeia, Andreas R. Pfennig, Katherine S. Pollard, David A. Ray, Beth Shapiro, Arian F. A. Smit, Mark S. Springer, Cynthia C. Steiner, Ross Swofford, Jussi Taipale, Emma C. Teeling, Jason Turner-Maier, Jessica Alfoldi, Bruce Birren, Oliver A. Ryder, Harris A. Lewin, Benedict Paten, Tomas Marques-Bonet, Kerstin Lindblad-Toh & Elinor K. Karlsson. A comparative genomics multitool for scientific discovery and conservation. *Nature*. 2020;587(7833):240-5.
7. Seshadri R, Leahy SC, Attwood GT, Teh KH, Lambie SC, Cookson AL, et al. Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection. *Nature biotechnology*. 2018;36(4):359-67.
8. Stewart RD, Auffret MD, Warr A, Wiser AH, Press MO, Langford KW, et al. Assembly of 913 microbial genomes from metagenomic sequencing of the cow rumen. *Nature communications*. 2018;9(1):870.
9. Contevelle LC, Silva Jvd, Andrade BGN, Coutinho LL, Palhares JCP, Regitano LCdA. Recovery of metagenome-assembled genomes from the rumen and fecal microbiomes of *Bos indicus* beef cattle. *Scientific Data*. 2024;11(1):1385.
10. Mi J, Jing X, Ma C, Shi F, Cao Z, Yang X, et al. A metagenomic catalogue of the ruminant gut archaeome. *Nature Communications*. 2024;15(1):9609.
11. Prondzinsky P, Toyoda S, McGlynn SE. The methanogen core and pangenome: conservation and variability across biology's growth temperature extremes. *DNA Research*. 2023;30(1):dsac048.
12. Alneberg J, Karlsson CM, Divne A-M, Bergin C, Homa F, Lindh MV, et al. Genomes from uncultivated prokaryotes: a comparison of metagenome-assembled and single-amplified genomes. *Microbiome*. 2018;6(1):173.
13. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer EL, et al. Pfam: The protein families database in 2021. *Nucleic acids research*. 2021;49(D1):D412-D9.
14. Eddy SR. Accelerated profile HMM searches. *PLoS computational biology*. 2011;7(10):e1002195.
15. Luton PE, Wayne JM, Sharp RJ, Riley PW. The *mcrA* gene as an alternative to 16S rRNA in the phylogenetic analysis of methanogen populations in landfill. *Microbiology*. 2002;148(11):3521-30.
16. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular biology and evolution*. 2020;37(5):1530-4.

17. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic acids research*. 2024;52(W1):W78-W82.
18. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068-9.
19. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome biology*. 2019;20(1):238.
20. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Molecular biology and evolution*. 2021;38(12):5825-9.
21. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. *Nucleic acids research*. 2021;49(D1):D545-D51.
22. Waskom ML. Seaborn: statistical data visualization. *Journal of open source software*. 2021;6(60):3021.
23. Hunter JD. Matplotlib: A 2D graphics environment. *Computing in science & engineering*. 2007;9(03):90-5.
24. Livingstone PG, Morpew RM, Whitworth DE. Genome sequencing and pan-genome analysis of 23 *Corallococcus* spp. strains reveal unexpected diversity, with particular plasticity of predatory gene sets. *Frontiers in microbiology*. 2018;9:3187.
25. Kang M-G, Kwak M-J, Kang A, Park J, Lee DJ, Mun J, et al. Metagenome-based microbial metabolic strategies to mitigate ruminal methane emissions using *Komagataeibacter*-based symbiotics. *Science of The Total Environment*. 2025;987:179793.
26. Benoit G, Raguideau S, James R, Phillippy AM, Chikhi R, Quince C. High-quality metagenome assembly from long accurate reads with metaMDBG. *Nature Biotechnology*. 2024;42(9):1378-83.
27. Pan S, Zhao X-M, Coelho LP. SemiBin2: self-supervised contrastive learning leads to better MAGs for short-and long-read sequencing. *Bioinformatics*. 2023;39(Supplement_1):i21-i9.
28. Orakov A, Fullam A, Coelho LP, Khedkar S, Szklarczyk D, Mende DR, et al. GUNC: detection of chimerism and contamination in prokaryotic genomes. *Genome biology*. 2021;22(1):178.
29. Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nature methods*. 2023;20(8):1203-12.
30. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy T, et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nature biotechnology*. 2017;35(8):725-31.
31. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *The ISME journal*. 2017;11(12):2864-8.
32. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics*. 2022;38(23):5315-6.
33. Lyu X, Yu H, Lu Y. Diversity and function of soluble heterodisulfide reductases in methane-metabolizing archaea. *Microbiology Spectrum*. 2025;13(5):e03238-24.
34. Chen H, Gan Q, Fan C. Methyl-coenzyme M reductase and its post-translational modifications. *Frontiers in microbiology*. 2020;11:578356.
35. Schütze A, Benndorf D, Püttker S, Kohrs F, Bettenbrock K. The impact of *ackA*, *pta*, and *ackA-pta* mutations on growth, gene expression and protein acetylation in *Escherichia coli* K-12. *Frontiers in Microbiology*. 2020;11:233.
36. Britt RD, Rao G, Tao L. Biosynthesis of the catalytic H-cluster of [FeFe] hydrogenase: the roles of the Fe-S maturase proteins HydE, HydF, and HydG. *Chemical Science*. 2020;11(38):10313-23.
37. Pandey C, Kumar U, Kaviraj M, Minick K, Mishra A, Singh J. DNRA: a short-circuit in biological N-cycling to conserve nitrogen in terrestrial ecosystems. *Science of the Total Environment*. 2020;738:139710.
38. Ragsdale SW, Pierce E. Acetogenesis and the Wood–Ljungdahl pathway of CO₂ fixation. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*. 2008;1784(12):1873-98.

39. Diao M, Dykema S, Koeksoy E, Ngugi DK, Anantharaman K, Loy A, et al. Global diversity and inferred ecophysiology of microorganisms with the potential for dissimilatory sulfate/sulfite reduction. *FEMS Microbiology Reviews*. 2023;47(5):fuad058.
40. Horz H-P, Yimga MT, Liesack W. Detection of methanotroph diversity on roots of submerged rice plants by molecular retrieval of *pmoA*, *mmoX*, *mxoF*, and 16S rRNA and ribosomal DNA, including *pmoA*-based terminal restriction fragment length polymorphism profiling. *Applied and Environmental Microbiology*. 2001;67(9):4177-85.
41. Tentori EF, Richardson RE. Methane monooxygenase gene transcripts as quantitative biomarkers of methanotrophic activity in *Methylosinus trichosporium* OB3b. *Applied and environmental microbiology*. 2020;86(23):e01048-20.
42. Rivals I, Personnaz L, Taing L, Potier M-C. Enrichment or depletion of a GO category within a class of genes: which test? *Bioinformatics*. 2007;23(4):401-7.
43. Falcon S, Gentleman R. Using GOSTats to test gene lists for GO term association. *Bioinformatics*. 2007;23(2):257-8.
44. Makarova KS, Wolf YI, Koonin EV. Archaeal clusters of orthologous genes (arCOGs): an update and application for analysis of shared features between Thermococcales, Methanococcales, and Methanobacteriales. *Life*. 2015;5(1):818-40.
45. Hyun JC, Monk JM, Palsson BO. Comparative pangenomics: analysis of 12 microbial pathogen pangenomes reveals conserved global structures of genetic and functional diversity. *BMC genomics*. 2022;23(1):7.
46. Makarova KS, Wolf YI, Koonin EV. Towards functional characterization of archaeal genomic dark matter. *Biochemical Society Transactions*. 2019;47(1):389-98.
47. Siliakus MF, van der Oost J, Kengen SW. Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure. *Extremophiles*. 2017;21(4):651-70.
48. Anderson RC, Ripley LH, Bowman JG, Callaway TR, Genovese KJ, Beier RC, et al. Ruminal fermentation of anti-methanogenic nitrate-and nitro-containing forages in vitro. *Frontiers in veterinary science*. 2016;3:62.
49. Ou Y-F, Dong H-P, McIlroy SJ, Crowe SA, Hallam SJ, Han P, et al. Expanding the phylogenetic distribution of cytochrome b-containing methanogenic archaea sheds light on the evolution of methanogenesis. *The ISME Journal*. 2022;16(10):2373-87.
50. Richy E, Thiago Dobbler P, Tlaskal V, López-Mondéjar R, Baldrian P, Kyselková M. Long-read sequencing sheds light on key bacteria contributing to deadwood decomposition processes. *Environmental Microbiome*. 2024;19(1):99.
51. Stewart RD, Auffret MD, Warr A, Walker AW, Roehe R, Watson M. Compendium of 4,941 rumen metagenome-assembled genomes for rumen microbiome biology and enzyme discovery. *Nature biotechnology*. 2019;37(8):953-61.
52. Waters S, Roskam E, Smith P, Kenny D, Popova M, Eugène M, et al. The role of rumen microbiome in the development of methane mitigation strategies for ruminant livestock. *Journal of Dairy Science*. 2025.
53. Steinberg LM, Regan JM. *mcrA*-targeted real-time quantitative PCR method to examine methanogen communities. *Applied and Environmental Microbiology*. 2009;75(13):4435-42.
54. Fischer MA, Güllert S, Neulinger SC, Streit WR, Schmitz RA. Evaluation of 16S rRNA gene primer pairs for monitoring microbial community structures showed high reproducibility within and low comparability between datasets generated with multiple archaeal and bacterial primer pairs. *Frontiers in microbiology*. 2016;7:1297.
55. Wilkins D, Lu X-Y, Shen Z, Chen J, Lee PK. Pyrosequencing of *mcrA* and archaeal 16S rRNA genes reveals diversity and substrate preferences of methanogen communities in anaerobic digesters. *Applied and environmental microbiology*. 2015;81(2):604-13.
56. Wang Y, Wegener G, Williams TA, Xie R, Hou J, Tian C, et al. A methylotrophic origin of methanogenesis and early divergence of anaerobic multicarbon alkane metabolism. *Science advances*. 2021;7(27):eabj1453.

57. Boyd JA, Jungbluth SP, Leu AO, Evans PN, Woodcroft BJ, Chadwick GL, et al. Divergent methyl-coenzyme M reductase genes in a deep-subseafloor Archaeoglobi. *The ISME journal*. 2019;13(5):1269-79.
58. Wolf YI, Makarova KS, Yutin N, Koonin EV. Updated clusters of orthologous genes for Archaea: a complex ancestor of the Archaea and the byways of horizontal gene transfer. *Biology direct*. 2012;7(1):46.
59. Gaba S, Kumari A, Medema M, Kaushik R. Pan-genome analysis and ancestral state reconstruction of class halobacteria: probability of a new super-order. *Scientific reports*. 2020;10(1):21205.
60. Rodrigues-Oliveira T, Belmok A, Vasconcellos D, Schuster B, Kyaw CM. Archaeal S-layers: overview and current state of the art. *Frontiers in microbiology*. 2017;8:2597.
61. Ng F, Kittelmann S, Patchett ML, Attwood GT, Janssen PH, Rakonjac J, et al. An adhesin from hydrogen-utilizing rumen methanogen *Methanobrevibacter ruminantium* M 1 binds a broad range of hydrogen-producing microorganisms. *Environmental microbiology*. 2016;18(9):3010-21.
62. Hook SE, Wright A-DG, McBride BW. Methanogens: methane producers of the rumen and mitigation strategies. *Archaea*. 2010;2010(1):945785.
63. Liu Y, Whitman WB. Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. *Annals of the New York Academy of Sciences*. 2008;1125(1):171-89.
64. Conrad R. Importance of hydrogenotrophic, acetoclastic and methylotrophic methanogenesis for methane production in terrestrial, aquatic and other anoxic environments: a mini review. *Pedosphere*. 2020;30(1):25-39.
65. Zhuang G-C, Heuer VB, Lazar CS, Goldhammer T, Wendt J, Samarkin VA, et al. Relative importance of methylotrophic methanogenesis in sediments of the Western Mediterranean Sea. *Geochimica et Cosmochimica Acta*. 2018;224:171-86.
66. Thomas CM, Taib N, Gribaldo S, Borrel G. Comparative genomic analysis of *Methanimitococcus blatticola* provides insights into host adaptation in archaea and the evolution of methanogenesis. *ISME communications*. 2021;1(1):47.
67. Lang K, Schuldes J, Klingl A, Poehlein A, Daniel R, Brune A. New mode of energy metabolism in the seventh order of methanogens as revealed by comparative genome analysis of “*Candidatus Methanoplasma termitum*”. *Applied and environmental microbiology*. 2015;81(4):1338-52.
68. Zinke LA, Evans PN, Santos-Medellín C, Schroeder AL, Parks DH, Varner RK, et al. Evidence for non-methanogenic metabolisms in globally distributed archaeal clades basal to the *Methanomassiliicoccales*. *Environmental Microbiology*. 2021;23(1):340-57.
69. Wu H-H, Pun MD, Wise CE, Streit BR, Mus F, Berim A, et al. The pathway for coenzyme M biosynthesis in bacteria. *Proceedings of the National Academy of Sciences*. 2022;119(36):e2207190119.
70. Costa KC, Leigh JA. Metabolic versatility in methanogens. *Current opinion in biotechnology*. 2014;29:70-5.
71. Pereira AM, de Lurdes Nunes Enes Dapkevicius M, Borba AE. Alternative pathways for hydrogen sink originated from the ruminal fermentation of carbohydrates: Which microorganisms are involved in lowering methane emission? *Animal Microbiome*. 2022;4(1):5.
72. Khairunisa BH, Heryakusuma C, Ike K, Mukhopadhyay B, Susanti D. Evolving understanding of rumen methanogen ecophysiology. *Frontiers in Microbiology*. 2023;14:1296008.
73. Ungerfeld EM. Metabolic hydrogen flows in rumen fermentation: principles and possibilities of interventions. *Frontiers in microbiology*. 2020;11:589.
74. Wang K, Xiong B, Zhao X. Could propionate formation be used to reduce enteric methane emission in ruminants? *Science of the Total Environment*. 2023;855:158867.
75. Janssen PH. Influence of hydrogen on rumen methane formation and fermentation balances through microbial growth kinetics and fermentation thermodynamics. *Animal feed science and technology*. 2010;160(1-2):1-22.

76. Ni G, Wang M, Walker N, Muetzel S, Schmidt O, Fischer A, et al. Methanogenesis inhibition remodels microbial fermentation and stimulates acetogenesis in ruminants. *Proceedings of the National Academy of Sciences*. 2025;122(41):e2514823122.
77. Ungerfeld EM. Shifts in metabolic hydrogen sinks in the methanogenesis-inhibited ruminal fermentation: a meta-analysis. *Frontiers in microbiology*. 2015;6:37.
78. Ungerfeld E, Pitta D. Biological consequences of the inhibition of rumen methanogenesis. *animal*. 2024;101170.
79. Latham EA, Anderson RC, Pinchak WE, Nisbet DJ. Insights on alterations to the rumen ecosystem by nitrate and nitrocompounds. *Frontiers in microbiology*. 2016;7:228.
80. Hristov AN, Oh J, Firkins J, Dijkstra J, Kebreab E, Waghorn G, et al. Special topics—Mitigation of methane and nitrous oxide emissions from animal operations: I. A review of enteric methane mitigation options. *Journal of animal science*. 2013;91(11):5045-69.
81. Tanaka K, Collins S, Polkoff K, Fellner V. Inhibiting methanogenesis by targeting thermodynamics and enzymatic reactions in mixed cultures of rumen microbes in vitro. *Frontiers in Microbiology*. 2024;15:1322207.
82. Granja-Salcedo YT, Fernandes RM, Araujo RCd, Kishi LT, Berchielli TT, Resende FDd, et al. Long-term encapsulated nitrate supplementation modulates rumen microbial diversity and rumen fermentation to reduce methane emission in grazing steers. *Frontiers in Microbiology*. 2019;10:614.
83. Ungerfeld EM. A theoretical comparison between two ruminal electron sinks. *Frontiers in microbiology*. 2013;4:319.
84. Le Van TD, Robinson JA, Ralph J, Greening RC, Smolenski WJ, Leedle JA, et al. Assessment of reductive acetogenesis with indigenous ruminal bacterium populations and *Acetitomaculum ruminis*. *Applied and environmental microbiology*. 1998;64(9):3429-36.
85. Bae B, Pope L, Park K, Lee C. Effects of nitrate in combination with lactate on rumen fermentation and methane production in an in vitro batch culture. *Animal Bioscience*. 2025;38(8):1672.
86. Greening C, Geier R, Wang C, Woods LC, Morales SE, McDonald MJ, et al. Diverse hydrogen production and consumption pathways influence methane production in ruminants. *The ISME journal*. 2019;13(10):2617-32.
87. Gagen EJ, Padmanabha J, Denman SE, McSweeney CS. Hydrogenotrophic culture enrichment reveals rumen Lachnospiraceae and Ruminococcaceae acetogens and hydrogen-responsive Bacteroidetes from pasture-fed cattle. *FEMS microbiology letters*. 2015;362(14):fnv104.
88. Belanche A, Bannink A, Dijkstra J, Durmic Z, Garcia F, Santos FG, et al. Feed additives for methane mitigation: A guideline to uncover the mode of action of antimethanogenic feed additives for ruminants. *Journal of Dairy Science*. 2025;108(1):375-94.

FIGURE LEGENDS

Fig. 1. Comparative pangenomic workflow for methanogens across diverse ecosystems. (A) Pipeline for reference methanogen genomes. Assembled genomes were annotated with Prokka, translated proteomes were functionally annotated with eggNOG-mapper, and orthogroups were inferred with OrthoFinder to yield gene clusters with consolidated annotations. (B) Pipeline for analyzing rumen shotgun sequencing metagenomes. Long-read assemblies were generated with metaMDBG, reads were mapped with minimap2, bins were reconstructed with SemiBin2, quality was screened with CheckM2 and GUNC, nonredundant representatives were obtained with dRep, taxonomy was assigned with GTDB-Tk, and proteins were annotated with eggNOG-mapper to enable downstream functional clustering. Both pipelines produce orthogroups with consolidated functional labels that feed into prevalence analysis and pathway-level summaries.

Fig. 2. Phylogenetic relationships of methanogens from diverse ecosystems using 16S rRNA and *mcrA* gene sequences. (A) Phylogenetic tree constructed from 16S rRNA gene sequences extracted from 71 methanogen genomes. Source environments are color-coded as follows: seawater (cyan), freshwater (pink-violet), sewage (orange), rumen (red), human gut (gray), ground (brown), and cockroach (yellow). (B) Phylogenetic tree based on *mcrA* gene sequences from the same 71 genomes. Red boxes highlight clades showing concordance between the 16S rRNA and *mcrA* phylogenies from rumen-derived methanogen genomes. Trees were inferred under a maximum-likelihood framework and include N = 71 genomes in both panels. Both trees are midpoint-rooted; the scale bar denotes expected substitutions per site. Branch-support values were computed and are shown at major nodes (see Methods for model settings; if not displayed, they were omitted for clarity). Tip labels were truncated to ensure legibility at print size; full strain names are provided in Supplementary Table S1.

Fig. 3. Comparative pangenome analysis of methanogens across diverse ecosystems. (A)

Orthogroups were classified by prevalence across all genomes: $\geq 99\%$ (core), $\geq 95\%$ and $< 99\%$ (soft-core), and $\geq 15\%$ and $< 95\%$ (shell). Core, soft-core, and shell genes are represented in dark, medium, and light green, respectively. Source environments are color-coded as follows: seawater (cyan), freshwater (pink-violet), sewage (orange), rumen (red), human gut (gray), ground (brown), and cockroach (yellow). The “Pressure” scale denotes the per-orthogroup presence ratio across strains (0 - 1). (B) COG functional category distributions for methanogen pangenome partitions. Bar plots show the percentage of orthogroups assigned to each COG functional category, based on eggNOG-mapper annotations mapped onto OrthoFinder-derived orthogroup. X-axis: COG category; y-axis: percentage of orthogroups. Correct “COG DISTRIBUTION” in all panel titles.

Fig. 4. Identification of methanogenesis-associated genes in methanogens from diverse ecosystems.

Methanogenesis pathway map overlaid with KEGG Orthology (KO)-level, environment-specific gene retention values. Each KO box corresponds to a methanogenesis-related complex within KEGG pathway ko00680. Gene retention is defined as the proportion of genomes within each environment that encode the KO, ranging from 0 (absent) to 1 (present in all genomes). Color intensity reflects the per-environment gene retention ratio (see right-side gradient scale). Source environments are color-coded as follows: seawater (cyan), freshwater (pink-violet), sewage (orange), rumen (red), human gut (gray), ground (brown), and cockroach (yellow). Between environment differences were tested with two-sided Fisher’s exact tests and adjusted by the Benjamini–Hochberg procedure; unless stated otherwise, significance is reported at $q < 0.05$. Significant between-environment differences included K00192 rumen OR = 0.10, K00193 rumen OR = 0.08, K00194 rumen OR = 0.06, K22480 rumen OR = 11.375, K00918 seawater OR = 8, K01985 seawater OR = 0.06. No human gut-specific retention was observed for K00200, K00201, K00202, K00203, K00205, K00319, K00320, K00441, K00577, K00578, K00579, K00580, K00581, K00584, K00672, K01499, K06914, K07072, K07144, K09733, K11212, K11780,

K11781, K14941, K16792, K16793, K18933.

Fig. 5. Functional classification of methanogenesis-associated roles in rumen MAGs. Metagenome-assembled genomes (MAGs) derived from rumen shotgun datasets were functionally categorized based on the presence of methanogenesis-related genes. Substrate supply genes included: hydrogen production (*hydA/E/F/G*), formate production (*pflA/pflB*), and acetate production (*pta/ackA*). Methanogens were marked by the presence of *mcrA/B/G* (methyl-coenzyme M reductase) and *mtrA/B/E* (coenzyme M methyltransferase). Competing pathways included: Wood–Ljungdahl pathway (*cooS/fhs/metF*) and nitrate reduction to ammonium (*nrfA/H*). Methanotrophs were identified via *pmoA* (particulate methane monooxygenase). Black tiles indicate gene presence; blank tiles denote absence. Rows are grouped by functional modules: substrate supply (teal), methanogenesis (orange), competitive sinks (red), and methanotrophy (light salmon). The role of MAGs are color-coded in the top bar: producers (blue), methanogens (red), competitors (orange), and methanotrophs (purple).

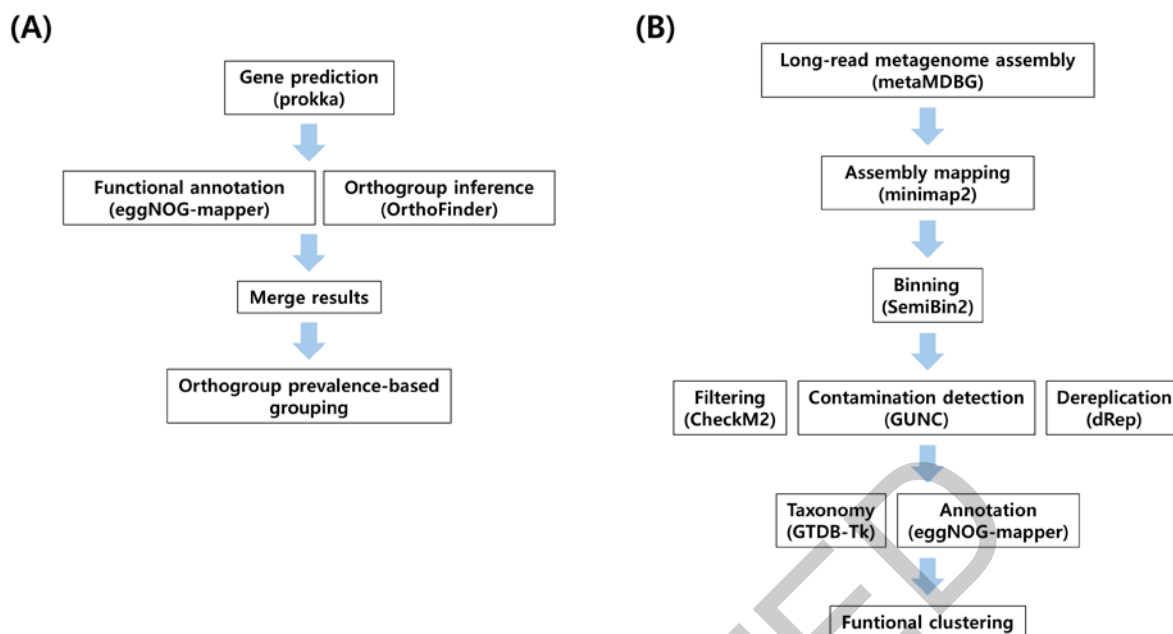


Figure 1.

(A)



(B)



Figure 2.

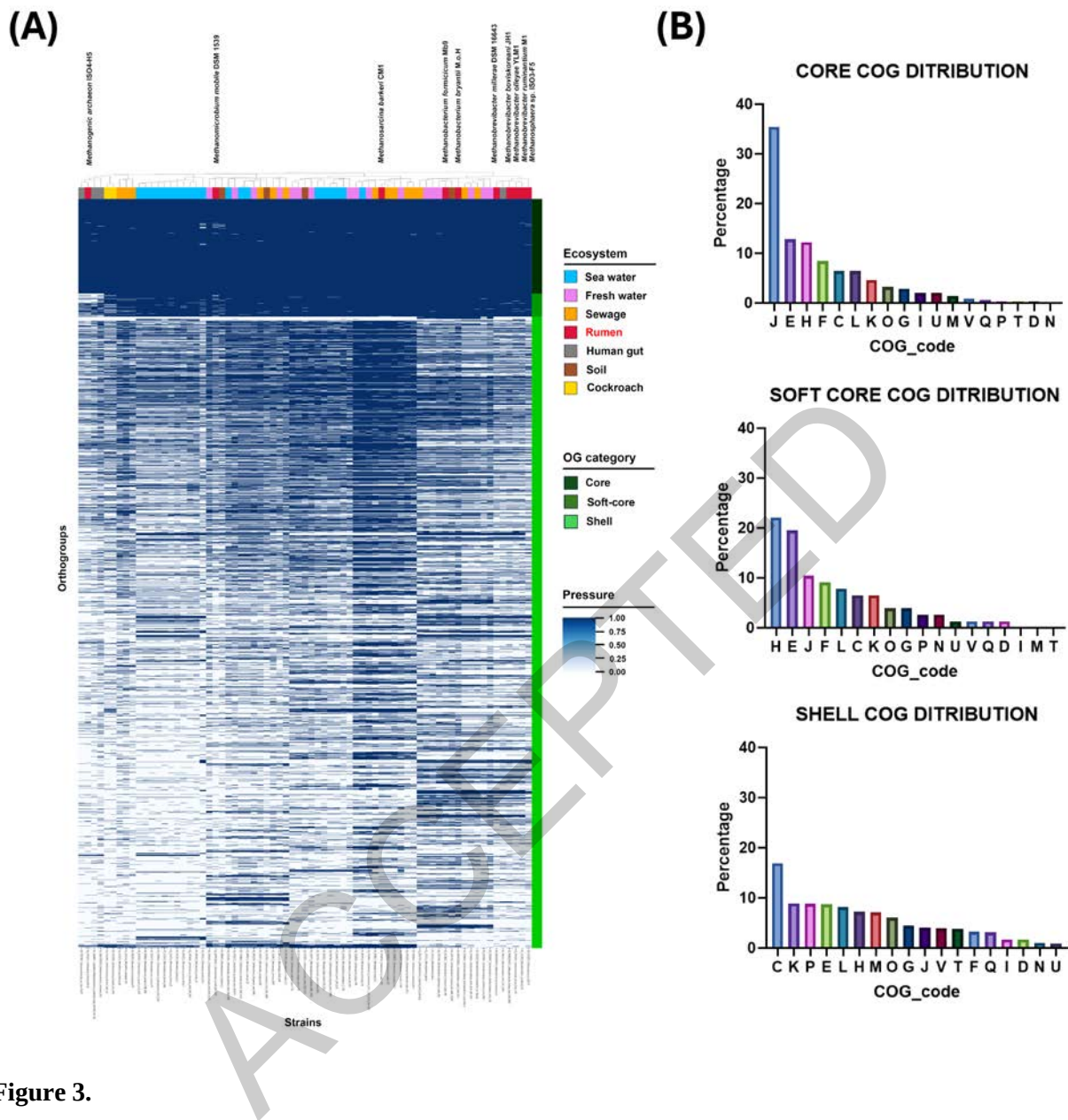


Figure 3.

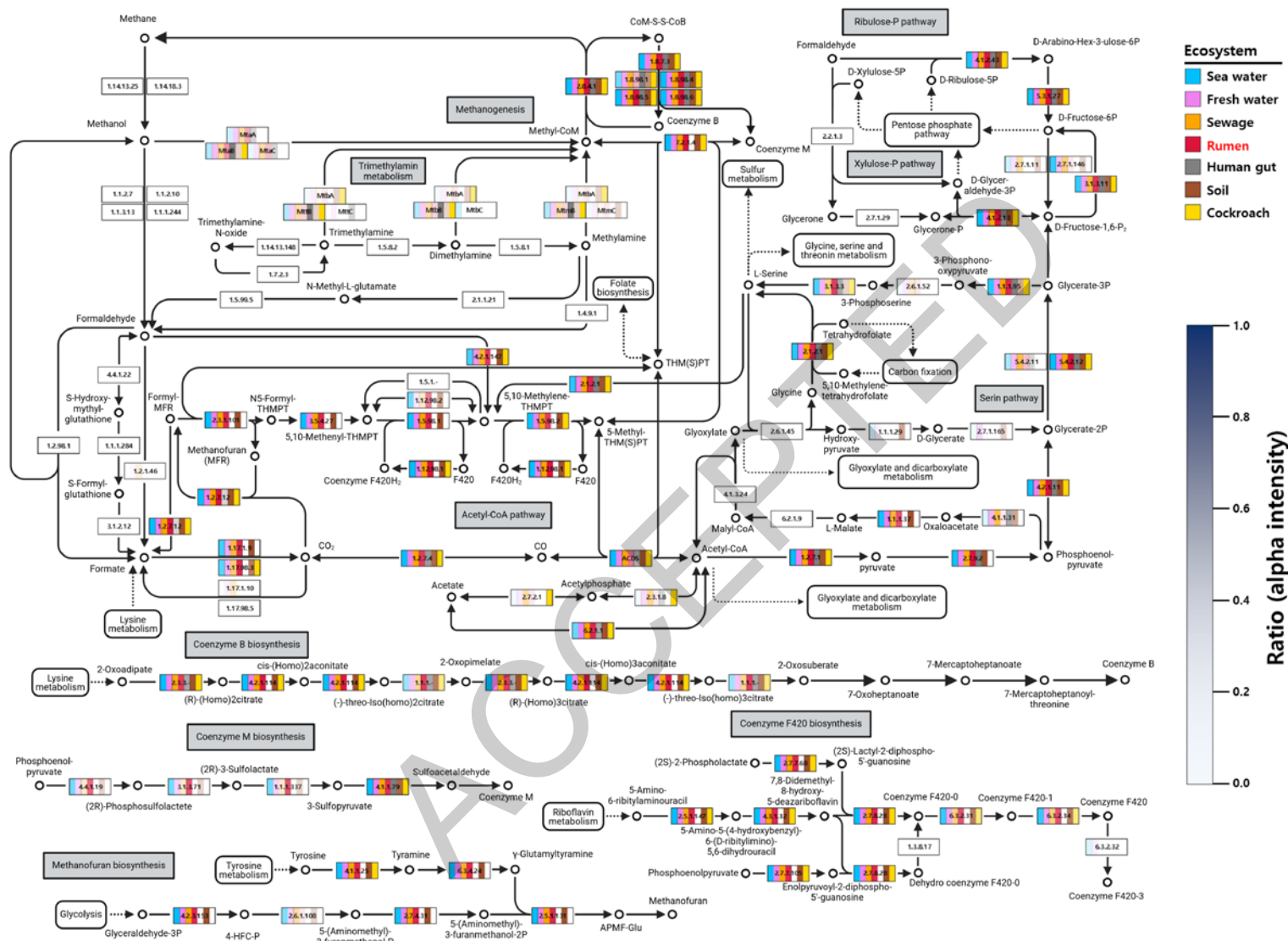
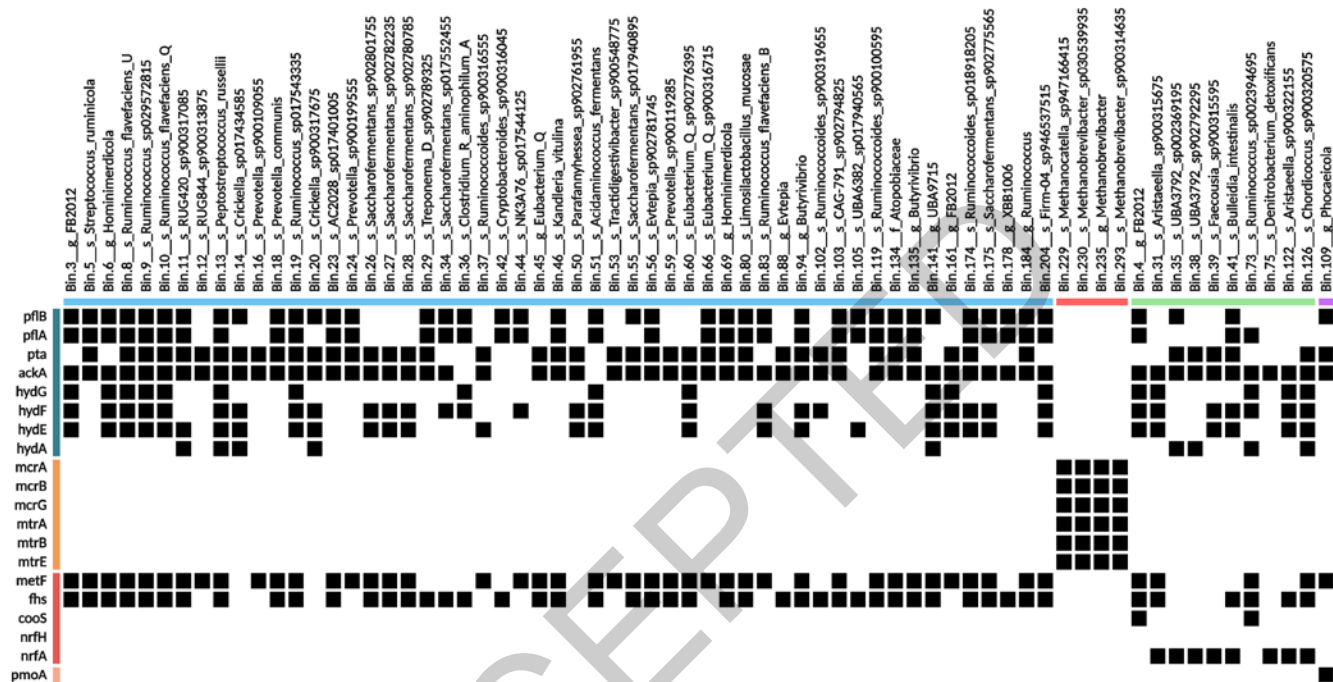


Figure 4.

Substrate supply
Methanogenesis
Competing sink
Methanotrophy



Producer
Methanogen
Competitor
Methanotroph

Figure 5.