

Review

Metagenomic and Computational Perspectives on the Rumen Microbiome: A Mini Review

Vishakha Uttam^{1*}, Anil Kumar Mishra¹, Raja Kolandanoor Nachiappan¹, Aditya D. Deshpande²,
Sayed Nabil Abedin³, Sonam Dwivedi⁴, Garima Choudhary⁵, Harinder Deep Singh¹, Prasoon Nayak¹,
Tamanna¹ and Rafiul¹

¹ICAR-National Bureau of Animal Genetic Resources, Karnal - 132001, Haryana, India

²ICAR-Central Institute of Fisheries Education, Mumbai - 400061, India

³ICAR-Indian Veterinary Research Institute, Mukteswar - 263138, Nainital, Uttarakhand, India

⁴ICAR-Central Institute for Research on Cattle, Meerut - 250001, Uttar Pradesh, India

⁵ICAR-Central Sheep & Wool Research Institute, Avikanagar - 304501, Rajasthan, India

ABSTRACT

The rumen microbiome is a diverse and functionally complex microbial community that plays a pivotal role in feed digestion, nutrient metabolism, animal productivity, and environmental sustainability. Conventional culture-based approaches have provided valuable insights into rumen microbial diversity but remain limited by the inability to cultivate majority of the rumen microorganisms. Advances in metagenomic sequencing have transformed rumen microbiome research by allowing culture-independent characterization of microbial communities, revealing wide taxonomic diversity, functional genes, metabolic pathways, and novel microbial resources with significant biotechnological and agricultural potential. Parallel developments in computational metagenomics, including progressive bioinformatics pipelines, long-read sequencing technologies, artificial intelligence (AI), and machine learning (ML), have considerably improved genome reconstruction, functional annotation, and prediction of microbial functions from complex sequencing datasets. These methods have applications in enzyme discovery, identification of fibre degrading microorganisms, feed efficiency, methane mitigation, and microbiome-informed livestock management. Despite these advances, challenges such as large-scale data processing, incomplete reference databases, computational complexity, and lack of standardized analytical workflows to limit the full exploitation of metagenomic data. This review summarizes recent advances in metagenomic and computational approaches for rumen microbiome research, highlighting sequencing technologies, bioinformatics tools, AI and ML based analytical methods, major applications, current challenges, and future perspectives.

Keywords: Computational metagenomics, Machine learning, Metagenomics, Rumen microbiome

*Corresponding author: vishakhauttam13@gmail.com

Received: July 28, 2026

Accepted: August 07, 2026

INTRODUCTION

The rumen is a complex and dynamic ecosystem primarily composed of anaerobic bacteria, protozoa, anaerobic fungi, methanogenic archaea, and bacteriophages (Huws *et al.*, 2018). These microorganisms interact with both the host and one another, allowing the breakdown of plant cell wall carbohydrates that are largely indigestible to humans (Mizrahi *et al.*, 2013). Through the fermentation of polysaccharides, rumen microorganisms produce short-chain fatty acids (SCFAs), mainly acetate, propionate, and butyrate, which are absorbed through the rumen epithelium and utilized by the host for maintenance and growth (Seshadri *et al.*, 2018). In addition to their essential role in ruminant nutrition, rumen microorganisms signify a valuable resource for the biofuel and biotechnology

industries, with metagenomic methods facilitating the discovery of novel proteins with potential industrial applications (Stewart *et al.*, 2018). During polysaccharide degradation, methanogenic archaea utilize hydrogen (H₂) as an energy source and combine it with carbon dioxide (CO₂) to produce methane (CH₄), which is expelled through eructation and released into the atmosphere (Leahy *et al.*, 2013). Therefore, understanding the functional dynamics of the rumen microbiome is essential for developing technologies and management strategies that improve the efficiency of ruminant production while mitigating greenhouse gas emissions (Seshadri *et al.*, 2018). However, studying rumen microorganisms remains challenging because the majority of this specialized microbial community, particularly those involved in degrading cellulosic plant material, are resistant to cultivation (Rinke *et al.*, 2013).

Recent advances in sequencing technologies have changed traditional microbiology into a more advanced and comprehensive discipline. Major revolutions in molecular microbiology have led to the advent of metagenomics, which includes the collective genomes and genetic material present within living systems, including humans, as well as those recovered from environmental samples such as soil and water (Athanasopoulou *et al.*, 2023). Metagenomics is a scientific discipline that focuses on the direct analysis of the genetic material of microorganisms within their natural environments to provide both taxonomic and functional insights (Ramazzotti *et al.*, 2018). The National Institutes of Health (NIH) defines metagenomics as the analysis of the collective genetic material recovered from all organisms, mainly microbes, within a bulk environmental sample to understand their composition and function (Kim *et al.*, 2024). This approach enables a wide range of applications, including estimation of relative abundance, taxonomic profiling, assessment of microbial diversity and richness, functional characterization, metabolic pathway analysis, phylogenetic studies, and detection of pathogenic organisms (Poussin *et al.*, 2018). Furthermore, predictive models based on AI and ML have become increasingly valuable for extracting meaningful insights from the large volumes of metagenomic data generated in laboratories. These computational approaches also facilitate a better understanding of microbial ecology and behaviour, thereby supporting informed decision-making, effective management strategies, and conservation planning (Kumar *et al.*, 2024).

Composition of the rumen microbiome

The rumen microbiome comprises a diverse assemblage of bacteria, protozoa, archaea, and fungi, with bacteria representing the most abundant and diverse group, followed by protozoa, archaea, and fungi (Mizrahi *et al.*, 2021). Unlike the eukaryotic members of the rumen microbiota, bacterial and archaeal communities are indispensable for the survival and efficient functioning of the ruminant host (Mizrahi *et al.*, 2013). Among the ruminal bacterial community, the phyla Firmicutes and Bacteroidetes are the most dominant at both the genomic and proteomic levels. For example, key fibre-degrading microorganisms such as *Fibrobacter succinogenes*, *Ruminococcus flavefaciens*, and *Ruminococcus albus* are involved in cellulose degradation, while members of the genera *Prevotella*, *Butyrivibrio*, and *Pseudobutyrvibrio* play important roles in hemicellulose degradation. In addition, bacteria such as *Treponema bryantii*, which utilize the products released during cellulose degradation, contribute to the

rumen fibre degradation network (Hart *et al.*, 2018; Morais *et al.*, 2019).

Methanogenic archaea, predominantly belonging to the phylum *Euryarchaeota*, utilize fermentation-derived hydrogen to reduce carbon dioxide and are the principal contributors to enteric methane production (Tapio *et al.*, 2017). Rumen protozoa, mainly ciliate species, constitute a substantial proportion of the microbial biomass and participate in fibre degradation, starch utilization, and hydrogen transfer during rumen fermentation (Newbold *et al.*, 2015). Anaerobic fungi facilitate the colonization and degradation of lignocellulosic plant material through mechanical penetration and the secretion of fibrolytic enzymes, thereby improving fibre digestion (Solomon *et al.*, 2016). In addition, bacteriophages regulate bacterial populations through lytic interactions, influencing the composition, diversity, and functional stability of the rumen microbial ecosystem (Gilbert *et al.*, 2017). Therefore, rumen microbiome functions through a complex network of microbial interactions that collectively facilitate the breakdown of feed components and the production of diverse metabolites. For example, SCFAs serve as the primary energy source for the host, while microbial biomass provides an important source of protein (Mizrahi *et al.*, 2013). Therefore, rumen microorganisms and their metabolic products play a central role in animal nutrition, productivity, and methane emissions, thereby influencing livestock sustainability and global food security.

Metagenomics as a tool for microbial bioprospecting

Metagenomics has emerged as a powerful culture-independent approach for microbial bioprospecting, enabling comprehensive characterization of both culturable and unculturable microorganisms and their functional potential (Athanasopoulou *et al.*, 2023). The term “metagenomics” was coined by Handelsman *et al.* (1998), and the field expanded rapidly with the development of whole-genome shotgun sequencing, which enabled direct analysis of microbial communities from environmental samples (Giani *et al.*, 2020).

Although techniques such as culturomics, PCR, 16S rRNA sequencing, FISH, and electrophoresis improved microbial characterization, their limitations in cultivation, taxonomic resolution, and functional assessment remain significant (Liu *et al.*, 2021; Zhang *et al.*, 2021). NGS-based amplicon sequencing, particularly 16S rRNA and ITS sequencing, enables high-throughput microbial profiling, while long-read platforms such as PacBio and Oxford Nanopore provide improved taxonomic resolution (Pei *et al.*, 2023). However, PCR

bias and limited functional information restrict its applications.

Shotgun metagenomics overcomes many of these limitations by sequencing total DNA, enabling assessment of microbial diversity, genetic composition, and functional potential. It includes taxonomic metagenomics for microbial profiling and functional metagenomics for identifying genes and metabolic pathways (Lam *et al.*, 2015). Nevertheless, high sequencing costs, large datasets, computational demands, host-DNA contamination, and inability to distinguish live from dead microorganisms remain challenges. Consequently, computational metagenomics has become essential for efficient analysis and interpretation of complex metagenomic datasets.

Computational metagenomics

Computational metagenomics involves the use of bioinformatics tools and algorithms to process and interpret metagenomic sequencing data. Key steps include quality control, genome assembly, metagenomic binning, taxonomic classification, gene prediction, and functional annotation (Lema *et al.*, 2023). The workflow generally involves sample collection, DNA extraction, library preparation, sequencing, quality filtering, assembly of reads into contigs, and reconstruction of metagenome-assembled genomes (MAGs). Binning

groups contigs based on sequence composition and abundance profiles, enabling taxonomic and functional characterization of microbial communities (Lapidus *et al.*, 2021; Setubal *et al.*, 2021).

Sequencing technology strongly influences downstream analyses. Illumina short-read sequencing provides high accuracy and cost-effectiveness, whereas long-read platforms such as PacBio and Oxford Nanopore improve genome assembly, resolution of repetitive regions, recovery of near-complete genomes, and detection of low-abundance taxa (Amarasinghe *et al.*, 2020). Hybrid approaches combining short- and long-read data can further enhance assembly and functional characterization (Kim *et al.*, 2024).

Taxonomic and functional annotation utilizes specialized algorithms and databases such as KEGG, eggNOG, COG, UniProt, MetaCyc, and Gene Ontology (Kanehisa *et al.*, 2017; Caspi *et al.*, 2020; Cantalapiedra *et al.*, 2021). Integrated pipelines and workflow systems, including MOCAT2, SnakeMAGs, SnakeWRAP, and Snakemake, improve automation, reproducibility, scalability, and cross-study comparability (Kultima *et al.*, 2016; Tadrent *et al.*, 2022; Krapohl *et al.*, 2022). Collectively, these approaches enable comprehensive assessment of rumen microbial diversity, metabolic potential, and associations with host traits (Wani *et al.*, 2022; Roy *et al.*, 2024).

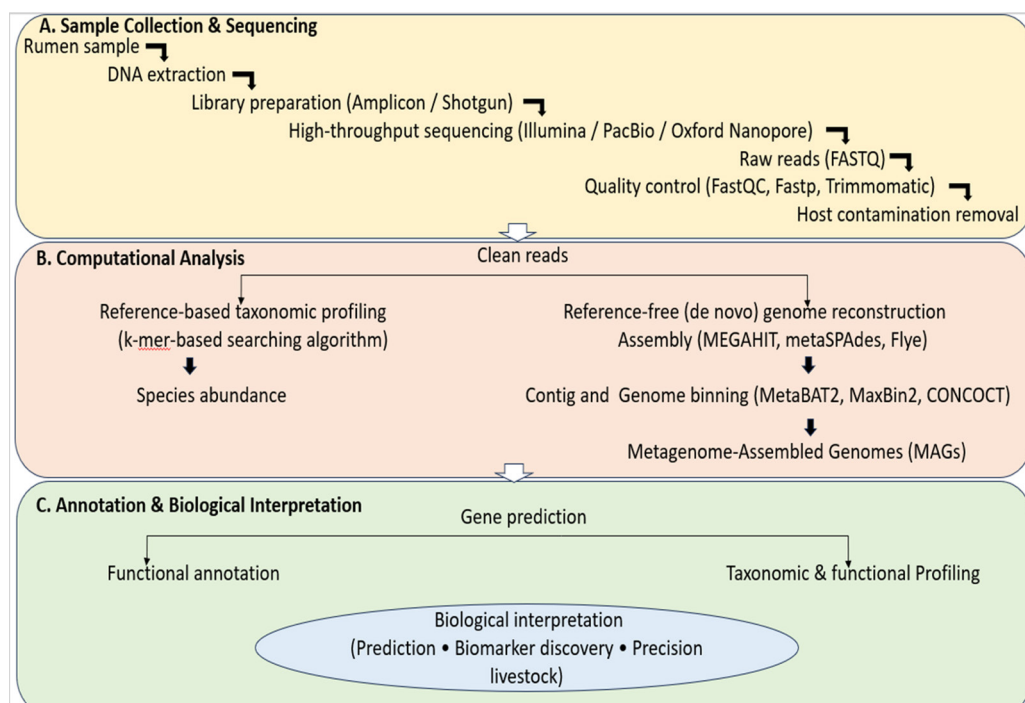


Fig. 1: Schematic illustration of the computational metagenomics workflow for rumen microbiome analysis. The workflow begins with rumen sample collection, DNA extraction, library preparation, high-throughput sequencing, and quality control to generate clean sequencing reads. The processed reads are subsequently analyzed using either reference-based taxonomic profiling or reference-free (de novo) genome reconstruction involving assembly and genome binning to recover metagenome-assembled genomes (MAGs). The resulting data are subjected to gene prediction, functional annotation, and taxonomic and functional profiling, enabling biological interpretation for biomarker discovery, prediction of host-associated traits, and precision livestock applications.

Advances in computational metagenomics

Recent advances in computational metagenomics, bioinformatics, artificial intelligence (AI), and machine learning (ML) have enabled more predictive and functional analysis of complex microbial communities (Yu *et al.*, 2024). Genome-centric metagenomics using metagenome-assembled genomes (MAGs) links genes to their source microorganisms, facilitating the identification of uncultured microbes and their roles in feed degradation, nutrient utilization, fermentation, and methane production.

Machine learning enables analysis of complex microbiome datasets for phenotype prediction, biomarker discovery, and functional trait identification. Algorithms such as Random Forest, Support Vector Machine, and regression models, along with PCA and PCoA, are widely used to identify and visualize microbial patterns (Pasolli *et al.*, 2016). More recently, deep learning approaches, including neural networks, CNNs, and RNNs, have improved the detection of complex nonlinear patterns and prediction of microbial functions and host phenotypes (Monteiro *et al.*, 2024).

Table 1: Overview of sequencing technologies and computational analysis in metagenomic studies

Step	Common approaches/ tools	Primary purpose	Advantages	Limitations
Amplicon sequencing (16S rRNA/ITS)	Illumina	Taxonomic profiling of bacterial, archaeal, or fungal communities	Cost-effective, standardized workflows, suitable for community composition analysis	Limited taxonomic resolution; provides little or no functional information
Shotgun metagenomic sequencing	Illumina	Sequencing of total microbial DNA for taxonomic and functional analysis	High accuracy, functional profiling, detection of diverse microorganisms	Requires higher sequencing depth and computational resources
Long-read metagenomic sequencing	PacBio HiFi, Oxford Nanopore Technologies	Genome reconstruction, strain-level characterization, recovery of metagenome-assembled genomes (MAGs)	Improved genome assembly, structural variant detection, enhanced strain resolution	Higher DNA input requirements, increased cost or platform-specific error profiles
Quality control and preprocessing	FASTQC, Fastp, Trimmomatic, Cutadapt	Assessment and filtering of sequencing reads	Improves accuracy of downstream analyses	Additional computational processing required
Genome assembly and binning	MEGAHIT, metaSPAdes, Flye; MetaBAT2, MaxBin2, CONCOCT	Reconstruction of microbial genomes	Enables recovery of genomes from uncultured microorganisms	Assembly is challenging in complex communities with closely related or low-abundance species
Taxonomic and functional analysis	Kraken2, MetaPhlAn, HUMAnN, eggNOG-mapper, Prokka	Identification of microbial taxa and prediction of gene functions	Comprehensive characterization of microbial communities	Accuracy depends on reference databases and annotation quality
Computational workflows	MetaWRAP, ATLAS, IMP, MOCAT2, SnakeMAGs; Snakemake, Nextflow	Automated and reproducible metagenomic analysis pipelines	Standardized, scalable and reproducible analyses	Some pipelines require substantial computational resources and bioinformatics expertise
Machine learning and AI	Random Forest, Support Vector Machine (SVM), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Feed-forward Neural Networks (FFN), PCA, PCoA	Biomarker discovery, phenotype prediction, microbial classification, dimensionality reduction and pattern recognition	Identifies complex biological patterns, improves prediction, handles high-dimensional datasets	Requires large, high-quality datasets; limited interpretability; risk of overfitting and batch effects

Applications of metagenomics in ruminant research

Metagenomic approaches have significantly advanced our understanding of the rumen microbiome by enabling comprehensive characterization of microbial diversity, functional genes, and metabolic pathways. Beyond taxonomic profiling, metagenomics has facilitated the discovery of novel microbial resources and provided valuable insights into microbial functions associated with feed digestion, animal productivity, and environmental sustainability. These advances have expanded the application of metagenomics in ruminant research, ranging from enzyme discovery and microbial functional characterization to precision livestock management and methane mitigation.

Identification of novel enzymes through functional metagenomics

Functional metagenomics has emerged as an important approach for the discovery of novel enzymes from the rumen microbiome by screening cloned metagenomic libraries for specific functional activities. This approach has enabled the identification of numerous carbohydrate-active enzymes (CAZymes), including cellulases, endoglucanases, esterases, and other polysaccharide-degrading enzymes involved in the digestion of plant biomass (Li, 2015). Although challenges such as low expression efficiency of foreign genes and limitations in functional screening remain, recent advances, including habitat-biased enrichment strategies and fluorescence-activated cell sorting, have substantially improved enzyme discovery (Ekkers *et al.*, 2012). These findings have expanded opportunities for identifying novel biocatalysts with applications in livestock nutrition, biotechnology, and biofuel production.

Identification of fibre-degrading microorganisms and functional pathways

Sequencing-based metagenomics has significantly improved the characterization of rumen microbial communities by providing comprehensive insights into their taxonomic composition and functional potential. High-throughput sequencing has facilitated the identification of microorganisms and carbohydrate-active enzyme (CAZyme) families involved in cellulose, hemicellulose, and starch degradation, thereby improving our understanding of rumen fibre digestion. Furthermore, metagenomic analyses have revealed variations in microbial functional pathways associated with dietary substrates and different physiological stages of animals, providing valuable information on microbial adaptation and nutrient utilization (Gharechahi *et al.*, 2021).

Improving feed utilization and industrial applications

The knowledge generated through rumen metagenomics has created opportunities for manipulating the rumen microbiome to improve feed utilization and animal productivity. The identification of novel fibrolytic enzymes and microbial species has facilitated the development of enzyme-based feed additives, improved microbial cultivation strategies, and enhanced methods for optimizing rumen fermentation. In addition, rumen-derived enzymes possess considerable commercial potential in industries such as biomass conversion and biofuel production, highlighting the broader biotechnological significance of rumen metagenomic research (Hua *et al.*, 2022).

Improving feed efficiency and prediction of production traits

Metagenomic studies have demonstrated that the rumen microbiome contributes significantly to economically important production traits, including methane emissions, residual feed intake (RFI), milk composition, feed efficiency, and body weight. The microbial contribution to these traits, often referred to as microbiability, highlights the importance of the rumen microbiome in livestock performance. Furthermore, microbiome-wide association studies (MWAS) and machine learning approaches have enabled the identification of microbial taxa associated with production traits and improved predictive models by integrating microbial and host information. These findings demonstrate that the rumen microbiome can provide not only candidate microbial or functional biomarker but also as an important predictor of animal performance, supporting the development of microbiome-informed breeding and precision livestock management strategies (Zhang *et al.*, 2023a). Recent evidence further suggests that heritable rumen microorganisms contribute more strongly to production traits than non-heritable microbial taxa. Based on this concept, holobiability has been proposed as a novel metric that integrates host genetic heritability with microbiability to quantify the combined influence of the host genome and rumen microbiome on economically important traits (Saborío-Montero *et al.*, 2021). This integrated framework has the potential to improve genomic prediction and accelerate genetic improvement through combined host-microbiome selection strategies.

Methane mitigation and sustainable livestock production

Metagenomics has become an important tool for understanding enteric methane production and identifying potential targets for methane-mitigation strategies. From a biological perspective, metagenomics

identifies the genetic repertoire and functional potential of microbial communities but cannot determine which genes are actively expressed under specific environmental conditions (Wang *et al.*, 2015). Integration of metagenomics with complementary meta-omics approaches, including metatranscriptomics, metaproteomics, and metabolomics, together with host genomic information, provides a systems-level understanding of the interactions among diet, the rumen microbiome, and host genetics that regulate methane production (Huws *et al.*, 2018). These integrated datasets facilitate the identification of candidate microbial taxa, genes, metabolic pathways, and host genetic regions associated with methane emissions, while machine learning and systems biology approaches further improve the prediction of methane-related traits and support the development of microbiome-informed breeding, precision feeding, and targeted microbial interventions for environmentally sustainable ruminant production (Roehe *et al.*, 2016).

Major challenges

Despite its potential for comprehensive microbial characterization, metagenomics faces several technical, computational, and biological limitations. Biases can arise during sample processing, DNA extraction, library preparation, and sequencing, while metagenomic studies require greater sequencing depth, high-quality DNA, higher costs, and substantial computational resources compared with 16S rRNA-based approaches (Wang *et al.*, 2015). Long-read technologies improve genome reconstruction but may still be affected by platform-specific errors (Lema *et al.*, 2023).

Computational challenges include incomplete reference databases, large and complex datasets, and difficulties in genome assembly, binning, and functional annotation, particularly for low-abundance or closely related microorganisms (Lema *et al.*, 2023). The high dimensionality and sparsity of metagenomic data can also complicate statistical and machine learning analyses, increasing the risk of overfitting and reducing model generalizability (Yu *et al.*, 2024). Furthermore, sequence-similarity-based functional annotation may produce uncertain or inaccurate predictions when closely related reference genomes are unavailable (Wang *et al.*, 2015).

Future perspectives

Future advances in rumen microbiome research are expected to be driven by the integration of genome-centric metagenomics with complementary multi-omics approaches, including metatranscriptomics, metaproteomics, and metabolomics, to achieve a comprehensive understanding of microbial

composition, gene expression, protein activity, metabolic functions, and their interactions with the host (Yu *et al.*, 2024). Genome-centric multi-omics approaches based on MAGs will enable the assignment of transcripts, proteins, and metabolites to individual microorganisms, thereby providing organism-level insights into key rumen processes such as feed digestion, fermentation, methane production, and animal productivity (Yu *et al.*, 2024). Furthermore, integrating differential gene expression, pathway enrichment, and metabolite profiling will facilitate the identification of functional biomarkers associated with feed efficiency, health, and production traits, supporting precision manipulation of the rumen microbiome for sustainable livestock production (Yu *et al.*, 2024).

Realizing these advances will require continued improvements in sequencing technologies together with the development of comprehensive rumen-specific reference genome, protein, and metabolite databases, expansion of high-quality MAG collections, standardized analytical workflows, and advanced bioinformatics, artificial intelligence, and machine learning-based tools for accurate genome assembly, functional annotation, pathway reconstruction, and multi-omics data integration (Yu *et al.*, 2024). Collaborative initiatives to develop and continuously update rumen microbial reference resources will further strengthen genome-centric analyses and improve the interpretation of microbial functions. Collectively, these developments are expected to accelerate biomarker discovery, enhance prediction of microbiome functions, and enable targeted microbiome engineering strategies to improve feed efficiency, animal health, productivity, and environmental sustainability while mitigating methane emissions in ruminants.

CONCLUSION

Conventional culture-based approaches have provided important perceptions into the rumen microbiome but remain limited in their ability to characterize the vast diversity and functional potential of unculturable microorganisms. Advances in metagenomics have changed rumen microbiome research by allowing inclusive characterization of microbial diversity, taxonomic composition, functional genes, and metabolic pathways, thereby providing deeper understandings into the microbial processes underlying feed digestion, fermentation, methane production, and animal productivity. Similar developments in computational metagenomics, supported by high-throughput sequencing, bioinformatics, artificial intelligence, and machine learning, have greatly improved genome reconstruction, functional annotation, and biological interpretation of increasingly complex microbiome

datasets. Also, the integration of genome-centric metagenomics with complementary multi-omics methods is expected to provide a more comprehensive understanding of host microbiome interactions and microbial functions. Collectively, these advances will facilitate the development of microbiome informed strategies for improving feed efficiency, animal health and productivity, methane mitigation, and the long-term sustainability of ruminant production.

REFERENCES

- Amarasinghe SL, Su S, Dong X, Zappia L, Ritchie ME, Gouil Q. 2020. Opportunities and challenges in long-read sequencing data analysis. *Genome Biology*, 21(1): 30.
- Athanasopoulou K, Adamopoulos PG, Scorilas A. 2023. Unveiling the human gastrointestinal tract microbiome: the past, present, and future of metagenomics. *Biomedicines*, 11(3): 827.
- Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. 2021. eggNOG-mapper v2: Functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Molecular Biology and Evolution*, 38(12): 5825–5829.
- Caspi R, Billington R, Keseler IM, Kothari A, Krummenacker M, Midford PE, Ong WK, Paley S, Subhraveti P, Karp PD. 2020. The MetaCyc database of metabolic pathways and enzymes: A 2019 update. *Nucleic Acids Research*, 48(D1): D445–D453.
- Ekkers DM, Cretoiu MS, Kielak AM, van Elsas JD. 2012. The great screen anomaly—a new frontier in product discovery through functional metagenomics. *Applied Microbiology and Biotechnology*, 93(3): 1005–1020.
- Gharechahi J, Vahidi MF, Bahram M, Han JL, Ding XZ, Salekdeh GH. 2021. Metagenomic analysis reveals a dynamic microbiome with diversified adaptive functions to utilize high lignocellulosic forages in the cattle rumen. *The ISME Journal*, 15: 1108–1120.
- Giani AM, Gallo GR, Gianfranceschi L, Formenti G. 2020. Long walk to genomics: History and current approaches to genome sequencing and assembly. *Computational and Structural Biotechnology Journal*, 18: 9–19.
- Gilbert RA, Kelly WJ, Altermann E, Leahy SC, Minchin C, Ouwerkerk D, *et al.* 2017. Toward understanding phage–host interactions in the rumen: Complete genome sequences of lytic phages infecting rumen bacteria. *Frontiers in Microbiology*, 8: 2340.
- Handelsman J. 2004. Metagenomics: Application of genomics to uncultured microorganisms. *Microbiology and Molecular Biology Reviews*, 68(4): 669–685.
- Hart EH, Creevey CJ, Hitch T, Kingston-Smith AH. 2018. Meta-proteomics of rumen microbiota indicates niche compartmentalisation and functional dominance in a limited number of metabolic pathways between abundant bacteria. *Scientific Reports*, 8: 10501.
- Hua D, Hendriks WH, Xiong B, Pellikaan WF. 2022. Starch and cellulose degradation in the rumen and applications of metagenomics on ruminal microorganisms. *Animals*, 12(21): 3020.
- Huws SA, Creevey CJ, Oyama LB, Mizrahi I, Denman SE, Popova M, *et al.* 2018. Addressing global ruminant agricultural challenges through understanding the rumen microbiome: Past, present, and future. *Frontiers in Microbiology*, 9: 2161.
- Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. 2017. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research*, 45(D1): D353–D361.
- Kim C, Pongpanich M, Porntaveetus T. 2024. Unraveling metagenomics through long-read sequencing: A comprehensive review. *Journal of Translational Medicine*, 22(1): 111.
- Krapohl J, Pickett BE. 2022. SnakeWRAP: A Snakemake workflow to facilitate automated processing of metagenomic data through the metaWRAP pipeline. *F1000Research*, 11: 265.
- Kultima JR, Coelho LP, Forslund K, Huerta-Cepas J, Li SS, Driessen M, Voigt AY, Zeller G, Sunagawa S, Bork P. 2016. MOCAT2: A metagenomic assembly, annotation and profiling framework. *Bioinformatics*, 32: 2520–2523.
- Kumar B, Lorusso E, Fosso B, Pesole G. 2024. A comprehensive overview of microbiome data in the light of machine learning applications: categorization, accessibility, and future directions. *Frontiers in Microbiology*, 15: 1343572.
- Lam KN, Cheng J, Engel K, Neufeld JD, Charles TC. 2015. Current and future resources for functional metagenomics. *Frontiers in Microbiology*, 6: 1196.
- Lapidus AL, Korobeynikov AI. 2021. Metagenomic data assembly—The way of decoding unknown microorganisms. *Frontiers in Microbiology*, 12: 61379.
- Leahy SC, Kelly WJ, Ronimus RS, Wedlock N, Altermann E, Attwood GT. 2013. Genome sequencing of rumen bacteria and archaea and its application to methane mitigation strategies. *Animal*, 7(Suppl. 2): 235–243.
- Lema NK, Gameda MT, Woldeamayat AA. 2023. Recent advances in metagenomic approaches, applications, and challenges. *Current Microbiology*, 80: 347.
- Li R. 2015. Rumen microbiology. In: Puniya AK, Singh R, Kamra DN (eds). *Rumen Microbiology: From Evolution to Revolution*. Springer, p. 223.
- Liu YX, Qin Y, Chen T, Lu M, Qian X, Guo X, Bai Y. 2021. A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein & Cell*, 12: 315–330.

- Mizrahi I. 2013. Rumen symbioses. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (eds). *The Prokaryotes: Prokaryotic Biology and Symbiotic Associations*. Springer, pp. 533–544.
- Mizrahi I, Wallace RJ, Morais S. 2021. The rumen microbiome: Balancing food security and environmental impacts. *Nature Reviews Microbiology*, 19(9): 553–566.
- Monteiro HF, Figueiredo CC, Mion B, Santos JEP, Bisinotto RS, Peñagaricano F, *et al.* 2024. An artificial intelligence approach of feature engineering and ensemble methods depicts the rumen microbiome contribution to feed efficiency in dairy cows. *Animal Microbiome*, 6(1): 5.
- Morais S, Mizrahi I. 2019. Islands in the stream: From individual to communal fibre degradation in the rumen ecosystem. *FEMS Microbiology Reviews*, 43: 362–379.
- Newbold CJ, De La Fuente G, Belanche A, Ramos-Morales E, McEwan NR. 2015. The role of ciliate protozoa in the rumen. *Frontiers in Microbiology*, 6: 1313.
- Pasolli E, Truong DT, Malik F, Waldron L, Segata N. 2016. Machine learning meta-analysis of large metagenomic datasets: Tools and biological insights. *PLoS Computational Biology*, 12(7): e1004977.
- Pei XM, Yeung MHY, Wong ANN, Tsang HF, Yu ACS, Yim AKY, Wong SCC. 2023. Targeted sequencing approach and its clinical applications for the molecular diagnosis of human diseases. *Cells*, 12: 493.
- Poussin C, Sierro N, Boue S, Battey J, Scotti E, Belcastro V, *et al.* 2018. Interrogating the microbiome: Experimental and computational considerations in support of study reproducibility. *Drug Discovery Today*, 23(9): 1644–1657.
- Ramazzotti M, Bacci G. 2018. 16S rRNA-based taxonomy profiling in the metagenomics era. In: Nagarajan M (ed). *Metagenomics: Perspectives, Methods, and Applications*. Academic Press, pp. 103–119.
- Rinke C, Schwientek P, Sczyrba A, *et al.* 2013. Insights into the phylogeny and coding potential of microbial dark matter. *Nature*, 499(7459): 431–437.
- Roehe R, Dewhurst RJ, Duthie CA, Rooke JA, McKain N, Ross DW, *et al.* 2016. Bovine host genetic variation influences rumen microbial methane production with best selection criterion for low methane-emitting and efficiently feed-converting hosts based on metagenomic gene abundance. *PLoS Genetics*, 12(2): e1005846.
- Roy G, Prifti E, Belda E, Zucker JD. 2024. Deep learning methods in metagenomics: A review. *Microbial Genomics*, 10(4): 001231.
- Saborío-Montero A, Gutiérrez-Rivas M, López-García A, García-Rodríguez A, Atxaerandio R, Goiri I, *et al.* 2021. Holobiont effect accounts for more methane emission variance than the additive and microbiome effects on dairy cattle. *Livestock Science*, 250: 104538.
- Seshadri R, Leahy SC, Attwood GT, Teh KH, Lambie SC, Cookson AL, Elae-Fadrosch EA, Pavlopoulos GA, Hadjithomas M, Varghese NJ, Paez-Espino D, Perry R, Henderson G, Creevey CJ, Terrapon N, Lapebie P, Drula E, Lombard V, Rubin E, *et al.* 2018. Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection. *Nature Biotechnology*, 36(4): 359–367.
- Setubal JC. 2021. Metagenome-assembled genomes: Concepts, analogies, and challenges. *Biophysical Reviews*, 13(6): 905–909.
- Solomon KV, Haitjema CH, Henske JK, Gilmore SP, Borges-Rivera D, Lipzen A, *et al.* 2016. Early-branching gut fungi possess a large, comprehensive array of biomass-degrading enzymes. *Science*, 351: 1192–1195.
- Stewart RD, Auffret MD, Warr A, Wiser AH, Press MO, Langford KW, Liachko I, Snelling TJ, Dewhurst RJ, Walker AW, Roehe R, Watson M. 2018. Assembly of 913 microbial genomes from metagenomic sequencing of the cow rumen. *Nature Communications*, 9(1): 870.
- Tadrent N, Dedeine F, Hervé V. 2022. SnakeMAGs: A simple, efficient, flexible and scalable workflow to reconstruct prokaryotic genomes from metagenomes. *F1000Research*, 11: 1522.
- Tapio I, Snelling TJ, Strozzi F, Wallace RJ. 2017. The ruminal microbiome associated with methane emissions from ruminant livestock. *Journal of Animal Science and Biotechnology*, 8: 11.
- Wang WL, Xu SY, Ren ZG, Tao L, Jiang JW, Zheng SS. 2015. Application of metagenomics in the human gut microbiome. *World Journal of Gastroenterology*, 21(3): 803–814.
- Wani AK, Roy P, Kumar V, ul Gani Mir T. 2022. Metagenomics and artificial intelligence in the context of human health. *Infection, Genetics and Evolution*, 100: 105267.
- Yu Z, Yan M, Wang J. 2024. Rumen microbiome nutriomics: Harnessing omics technologies for enhanced understanding of rumen microbiome functions and ruminant nutrition. *Animal Nutriomics*, 1: e10.
- Zhang B, Lin S, Moraes L, *et al.* 2023. Methane prediction equations including genera of rumen bacteria as predictor variables improve prediction accuracy. *Scientific Reports*, 13(1): 21305.
- Zhang J, Liu YX, Guo X, Qin Y, Garrido-Oter R, Schulze-Lefert P, Bai Y. 2021. High-throughput cultivation and identification of bacteria from the plant root microbiota. *Nature Protocols*, 16: 988–1012.