World Journal of Microbiology and Biotechnology (2025) 41:267
https://doi.org/10.1007/s11274-025-04504-6

REVIEW

The crosstalk between host and rumen microbiome in cattle: insights from multi-omics approaches and genome-wide association studies

Xiaoxing Ye[1] · Goutam Sahana[1] · Mogens Sandø Lund[1] · Zexi Cai[1]

Received: 6 May 2025 / Accepted: 18 July 2025 / Published online: 28 July 2025

Abstract

The rumen microbiome plays a pivotal role in cattle health, productivity, and environmental sustainability by driving key processes such as nutrient metabolism, feed efficiency, and methane emissions. The intricate interactions among host genome, microbiome composition, and dietary factors shape a dynamic ecosystem that underpins animal performance and sustainability outcomes. Leveraging recent advancements in sequencing technologies and multi-omics approaches, this review synthesizes emerging insights into the genetic and functional profile of the rumen microbiome. Meanwhile, association studies with microbiome and genome-wide markers have identified key microbial and genetic markers linked to methane emissions, paving the way for microbiome-driven breeding programs aimed at reducing environmental impacts. This review also addresses challenges in integrating microbiome datasets, bioinformatics workflows, and reference databases, offering strategies to overcome these obstacles through multi-omics and advanced computational tools. By highlighting actionable pathways for microbiome-informed breeding and management strategies, this work provides a novel framework for improving livestock productivity and mitigating environmental footprints.

Keywords Cattle · Rumen microbiome · Microbiome GWAS · Multi-omics · Host-microbiome interactions · Methane mitigation

Abbreviations

mGWAS	Microbiome-wide association studies
MAGs	Metagenome-assembled genomes
VFAs	Volatile fatty acids
NGS	Next generation sequencing
OTUs	Operational taxonomic units
ASVs	Amplicon sequence variants
SGNs	Species-level genome bins
PacBio	Pacific Bioscience
SMRT	Single-Molecule Real-Time
ONT	Oxford Nanopore Technologie
RMG	Rumen Microbial Gene Catalogue
SCFAs	short-chain fatty acids
NMR	Nuclear magnetic resonance
ML	Machine learning
LD	Linkage disequilibrium
MAS	Marker-assisted selection
GBS	Genotyping-by-sequencing
PCA	Principal component analysis

✉ Xiaoxing Ye
xiaoxing.ye@qgg.au.dk

[1] Center for Quantitative Genetics and Genomics, Aarhus University, CF Møllers Allé 3, Aarhus 8000, Denmark

Introduction

The interactions among host genetics (Bubier et al. 2021), the rumen microbiome (Matthews et al. 2019), and dietary factors (Gruninger et al. 2019; Leeming et al. 2019) create a complex, dynamic system pivotal for health (O'Hara et al. 2020), growth performance (Daghio et al. 2021; Boggio et al. 2023), and methane emissions (Matthews et al. 2019; Liu et al. 2021) in cattle. These relationships drive key phenotypes such as feed efficiency and methane emissions, which are vital for achieving sustainable livestock production. In this ecosystem, host genetics significantly influence the composition and functionality of the rumen microbiome, a trait with considerable heritability (Li et al. 2019b; Wallace et al. 2019). Conversely, the microbiome directly shapes host phenotypes by modulating metabolic pathways, as evidenced by breed-specific differences (Hernandez-Sanabria

et al. 2013). This bidirectional interplay, often referred to as "co-metabolism", highlights the need for integrative studies that encompass both host and microbial contributions to phenotypic outcomes (Wang et al. 2024b).

Recent advances in sequencing technologies and multi-omics approaches have transformed our understanding of the rumen microbiome. These developments have enabled the transition from traditional culture-based techniques to high-throughput metagenomics, metatranscriptomics, and other omics platforms, uncovering microbial diversity, functionality, and ecological dynamics with unprecedented precision (Uruburu 2003; Wassan et al. 2023). Multi-omics integration has further illuminated the rumen microbiome's critical role in cattle productivity, revealing key metabolic pathways and microbial interactions associated with health, feed efficiency, and methane mitigation (Xue et al. 2020; Fonseca et al. 2023). Additionally, microbiome-wide association studies (mGWAS) have begun to elucidate the genetic underpinnings of host-microbiome interactions, uncovering loci associated with microbiome composition and functionality (Fregulia et al. 2022). Integrating rumen microbiome data into predictive models for methane emissions offers a transformative opportunity to incorporate methane reduction as a novel breeding objective. These findings hold great promise for developing methane mitigation strategies for cattle, especially opening the possibility for targeted breeding programs that optimize cattle performance while addressing environmental challenges (Ismail et al. 2023). This aligns with global sustainability goals by enabling the identification of microbial and genetic markers that drive low-emission phenotypes. Such insights offer new ways to select multi-trait phenotypes including microbial metagenomes.

This review aims to (1) provide a comprehensive overview of recent bioinformatics and multi-omics advances in understanding the cattle rumen microbiome; (2) discuss the challenges and opportunities of integrating multi-omics data into this research domain; and (3) explore the potential of mGWAS and other genomic tools in advancing livestock breeding strategies for improved health, productivity, and environmental sustainability.

Rumen microbiome: composition, diversity and function

Rumen is a highly active microbial ecosystem, essential for the digestion of fibrous feed and the production of energy metabolites critical for ruminant productivity (Xue et al. 2020; Ismail et al. 2023; Lawther et al. 2024). Despite extensive studies using advanced technologies, only a fraction of ruminal microbes has been fully characterized (Gruninger et al. 2019). Major initiatives like the Hungate1000 Project have significantly expanded our knowledge by sequencing 501 genomes (480 bacterial and 21 archaeal) from rumen microbial communities (Seshadri R, 2018). Subsequent efforts have further enriched genomic resources, including 913 metagenome-assembled genomes (MAGs) (Stewart et al. 2018), 4941 MAGs (Stewart et al. 2019), over 10,000 MAGs from ruminants (Xie et al. 2021), and 1200 high-quality MAGs from African cattle (Wilkinson et al. 2020) (Fig. 1).

Bacteria

Bacteria represent the most abundant and taxonomically diverse group of microorganisms within the rumen, with their population estimated at 10^{10}-10^{11} cells per milliliter of rumen fluid. Bacteria are classified based on their association with specific rumen factions, namely solid, liquid, and epithelial compartments, which are crucial roles in fiber degradation, metabolism of soluble nutrients, and regulation of epithelial cell turnover. Among bacteria, *Prevotella*, *Ruminococcus* and *Butyrivibrio* are dominant genera, playing crucial roles in nitrogen cycling and fiber degradation. These bacteria derive nitrogen primarily from ammonia and depend on low concentrations of C4-C5 branched-chain fatty acids for growth (Henderson et al. 2015; Puniya et al. 2015; Zhao et al. 2022; Kupczyński et al. 2024). For instance, *Butyrivibrio* is noted for its role in hemicellulose degradation, and thrives in forage-based diets. Similarly, species like *Butyrivibrio fibrisolvens*, *Prevotella ruminicola*, *Lachnospira multiparus*, *Succinivibrio dextrinosolvens*, and *Fibrobacter succinogenes* contribute to the fermentation of cellulose, hemicellulose, and pectin - key substrates in fiber-based diets (Henderson et al. 2015). These bacterial activities result in the production of volatile fatty acids (VFAs) and fermentable sugars, driving energy metabolism and nutrient absorption while reinforcing host-microbiome interactions (Paz et al. 2016; Dao et al. 2021; Wu et al. 2021; Betancur-Murillo et al. 2023).

Archaea

Although comprising only 0.3–3% of the total microbiome, archaea play a pivotal role in rumen dynamics and methane emissions (Janssen and Kirs 2008). Similar to bacteria, archaea are associated with distinct fractions, including solid, liquid, and epithelial compartments. Despite their morphological resemblance to bacteria, archaea exhibit distinct metabolic characteristics. The predominant functional group among rumen archaea comprises methanogens, which generate methane as a metabolic byproduct. Methanogens, such as *Methanobrevibacter*, are predominant in cattle rumen. In Fig. 2, a global rumen census identified *Methanobrevibacter*

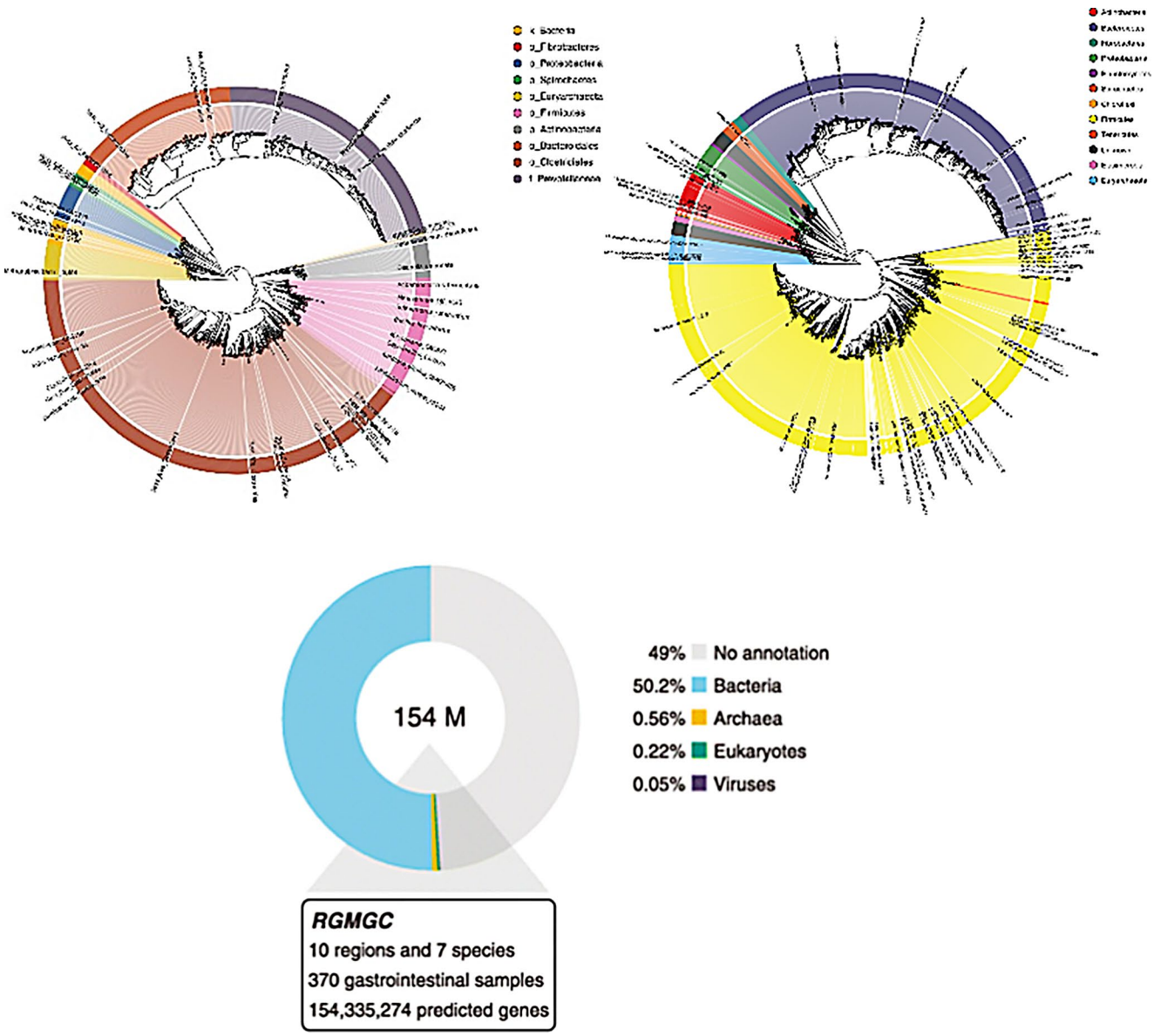

Fig. 1 An overview of metagenome-assembled genomes for ruminants based on recent research adapted from Stewart et al. (2018), Stewart et al. (2019), and Xie et al. (2021)

gottschalkii as the most abundant methanogen, accounting for approximately 54% of the total archaeal community in sheep, followed by *Methanobrevibacter ruminantium* (20%) and *Methanosphaera sp. ISO3-F5* (Henderson et al. 2015). *Methanobrevibacter* can utilize hydrogen produced by bacteria, protozoa, and fungi to generate methane and ATP, influencing energy balance within the rumen ecosystem (Kim et al. 2011; Wallace et al. 2014; Henderson et al. 2015; McAllister et al. 2025). This bidirectional interplay, often referred to as 'co-metabolism', highlights the need for integrative studies that encompass both host and microbial contributions to phenotypic outcomes Metabolomics provides critical insights into the shared metabolic outputs of fermentation processes in the rumen, where microbial degradation of fibrous feed generates VFAs that supply up to 70% of the host's energy requirements for maintenance and production, with some chemical energy inevitably lost as methane (Harirchi et al. 2022).

Protozoa, fungi, and viruses

While the roles of bacteria and archaea are relatively well-studied, the functions of protozoa, fungi, and viruses remain less clear. Among protozoa, genera *Entodinium* and *Epidinium* are predominant (Park et al. 2021; Li et al. 2022; Andersen et al. 2023). Recent advances in the genomic

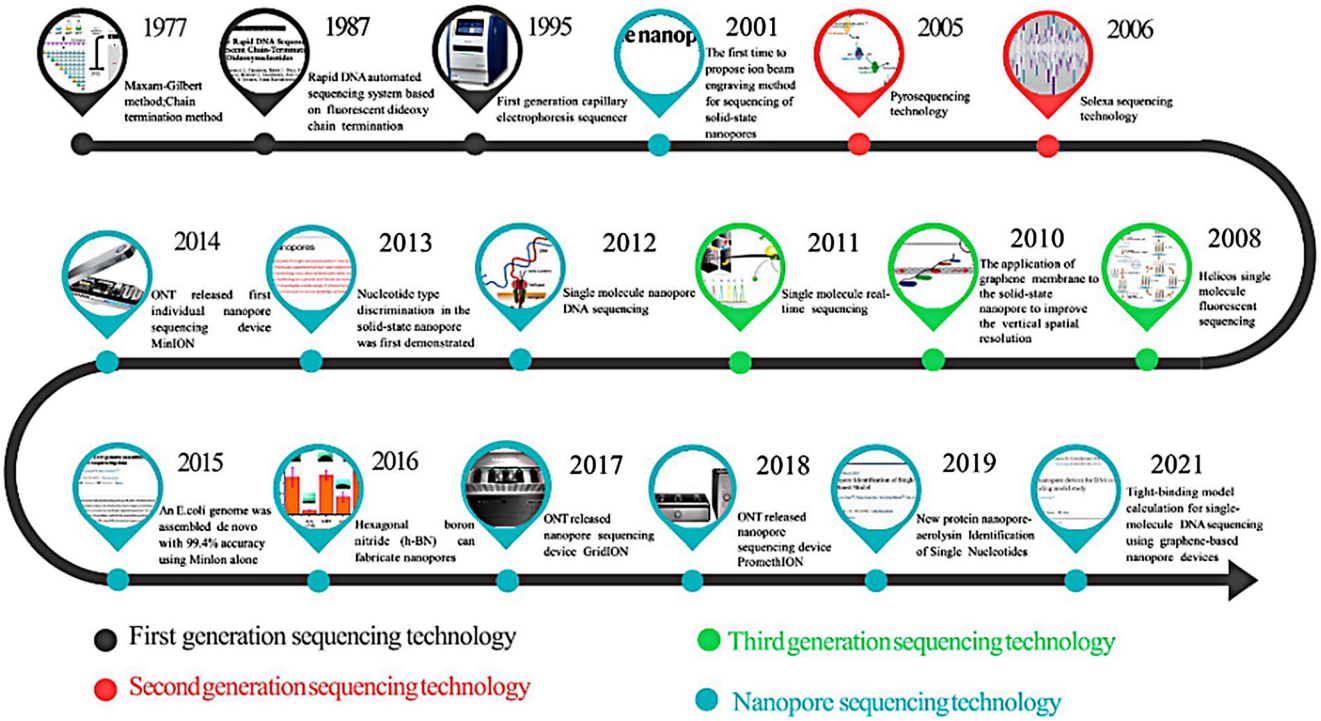

Fig. 2 Timeline of the introduction of sequencing technologies and platforms. Adapted from Chen et al. (2023).

characterization of rumen ciliate protozoa, such as the sequencing of *Entodinium caudatum* (Park et al. 2021), and single cell approaches to reconstruct ciliate phylogeny (Li et al. 2022), have expanded our understanding of protozoa. Meanwhile, Andersen et al. (2023) discussed the metabolic influence of core ciliates protozoa in rumen microbiome. Therefore, recent research could help us to understand protozoa well.

Protozoa contribute to the mechanical breakdown of plant feed and nutrient cycling by predating bacteria and other protozoa, helping stabilize microbial communities (Yu et al. 2024). Fungi specialize in degrading plant cell walls, synergizing with bacteria and protozoa to enhance feed digestibility and improve energy metabolism (Lee et al. 2000). Viruses play a significant role in shaping microbial community dynamics and are abundant in rumen (Yan and Yu 2024). Viruses, particularly bacteriophages, regulate microbial dynamics by infecting and lysing specific bacterial populations, controlling their abundance and transferring genetic information through transduction (Yan and Yu 2024). These interactions among eukaryotic microbes, viruses, and their hosts create a dynamic ecosystem that influences productivity and microbial resilience (Vargas-Albores et al. 2023). While understanding the microbial players is essential, advancements in sequencing technologies have revolutionized how we study these ecosystems, as discussed below.

Advancements in sequencing technologies: unlocking the rumen microbiome

The rapid development of sequencing technologies (Chen et al. 2023) (Fig. 2) has greatly enhanced rumen microbiome research, revealing microbial diversity, functionality, and ecological dynamics. These developments have deepened our understanding of microbe-host interactions and their impact on ruminant health and productivity (McCann et al. 2014). This section overviews short-read (Next generation sequencing [NGS]) and long-read (third generation) sequencing technologies for microbiome research, including amplicon sequencing, shotgun metagenomics, and metatranscriptomics (Fig. 3). The strengths and limitations of each approach, along with bioinformatics workflows and reference databases, are discussed to enhance sequencing-based microbiome analyses.

Amplicon sequencing: establishing taxonomic profiles

Amplicon sequencing, particularly targeting conserved regions of the 16 S rRNA (Abellan-Schneyder et al. 2021), and 23 S rRNA genes (García-Martínez et al. 1999) for bacteria and archaea, 18 S rRNA (Choi and Park 2020) for eukaryotes such as protozoa and algae, and the ITS for fungi (Schoch et al. 2012), has long been used microbial

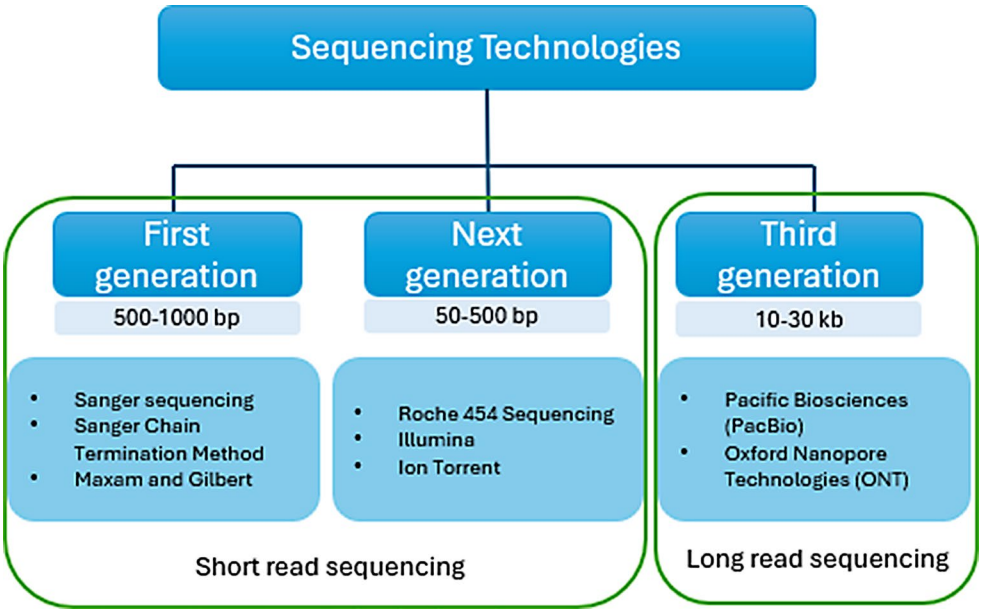

Fig. 3 An overview of short-read and long-read sequencing

diversity study. This cost-effective approach provides detailed taxonomic profiles, enabling the quantification of microbial diversity and community structure in the rumen microbiome. Its cost-effectiveness and accessibility make it a widely used method for studying microbial communities (Větrovský and Baldrian 2013), especially when combined with reference databases like RDP (Maidak et al. 1996), SILVA (Quast et al. 2013), and Greengenes (Quast et al. 2013), which provide reliable taxonomic frameworks. However, it suffers from several limitations, including PCR biases, limited species-level resolution, and the inability to infer functional potential (Acinas et al. 2005; Bharti and Grimm 2021). Additionally, amplicon sequencing suffers from limited species-level resolution, however, it lacks functional information (Rodino et al. 2021; Zaiko et al. 2022). Consequently, using whole-genome approaches can help to offer more comprehensive resolution and functional insights.

Shotgun metagenomic sequencing: unveiling functional potential

Shotgun metagenomic sequencing offers a more comprehensive analysis by sequencing the entire DNA content of microbial communities. This technique provides detailed insights into microbial functional genes, metabolic pathways, and host-microbe interactions (Handelsman 2004; Denman et al. 2018). It avoids amplification biases and has become the gold standard for rumen microbiome studies aiming to capture metabolic diversity and host-microbe interactions.

A major advance stemming from shotgun metagenomics is the genome-resolved metagenomics approach, which enables the reconstruction of MAGs. Several large-scale studies have significantly expanded rumen microbial gene catalogs and genome databases. For instance, Seshadri R (2018) and Stewart et al. (2018) expanded public rumen datasets fivefold, increasing mapping rates of rumen metagenomes to 50–70%. Additionally, Xie et al. (2021) explored the spatial associations between 8,745 uncultured genomes and ruminant physiological adaptations, further enriching the genomic landscape of the rumen microbiota. More recently, Mi et al. (2024) constructed the ruminant gut archaeome with 998 unique archaeal genomes using metagenomic sequencing. Xu et al. (2024) ecovered over 131,000 MAGs and 11,000 high-quality species-level genome bins (SGBs) across 17 ungulate species. These efforts demonstrate the power of metagenomics in expanding the known functional landscape of the rumen.

Importantly, genome-resolved metagenomics have also been extended to underrepresented microbial domains such as viruses and eukaryotes. For example, Yan et al. (2023) developed an approach to reconstruct > 397,180 species-level viral operational taxonomic units (vOTUs), rumen virome database substantially increases the detection rate of rumen viruses from metagenomes compared with IMG/VR V3. Another study was shown to facilitate comprehensive analyses of protozoa and fungi within more than 1000 rumen metagenomes, revealing a greater genomic diversity among protozoa than previously documented (Yan 2025).

Despite the strengths of short-read metagenomics, its limitations include fragmented assemblies, lack of strain-level resolution, and high computational demands. These challenges can be addressed by integrating long-read sequencing platforms (e.g., Oxford Nanopore, PacBio HiFi), which improve assembly contiguity and enable

recovery of full-length genes and operons (Yu et al. 2025). Hybrid assemblies combining short and long reads represent a promising strategy to generate near-complete MAGs and explore complex genomic architectures such as mobile elements and gene clusters involved in fiber degradation or methanogenesis.

While metatranscriptomics can help address this by analyzing the complete set of RNA transcripts to indicate which genes are being transcribed (Terrón-Camero et al. 2022), it is important to recognize that transcript abundance does not necessarily reflect protein levels or functional enzyme activities (Edfors et al. 2016), as study have shown a low and variable correlation between mRNA and protein concentrations in complex biological samples (Maier et al. 2009).

Long-read sequencing: advancing microbial genomics and functional studies

Long-read sequencing technologies, such as Pacific Biosciences' (PacBio) Single-Molecule Real-Time (SMRT) sequencing and Oxford Nanopore Technologies' (ONT) MinION and PromethION, have transformed microbial genomics by addressing limitations of short-read platforms. These technologies generate significantly longer reads (10–50 kbp for PacBio; >10 kbp for ONT), enabling the resolution of repetitive regions, complex genomic structures, and full-length RNA transcripts that short-read sequencing cannot easily resolve (Steijger et al. 2013; Gao et al. 2016; Lang et al. 2020). Recent advancements in sequencing chemistry and base-calling algorithms—such as ONT Q20 + and PacBio HiFi circular consensus sequencing—have significantly improved read accuracy (> 99%), making long-read sequencing more reliable and applicable to metagenomics and other omics studies. However, computational demands and higher costs remain challenges for widespread adoption.

Long-read sequencing such as ONT R10.4 and PacBio HiFi facilitates the generation of near-complete bacterial genomes without requiring short-read or reference-based polishing (Sereika et al. 2022; Greenman et al. 2024). Besides, long-read sequencing allows for full-length 16 S rRNA gene retrieval, compared with short-read sequencing, which fragments the 16 S rRNA gene into hypervariable regions (e.g., V3–V4 or V4–V5). This advancement enhances taxonomic resolution from genus to species and even strain level, reducing classification biases and improving phylogenetic analyses (Buetas et al. 2024). For instance, PacBio and ONT platforms have been successfully applied to generate full-length 16 S rRNA gene sequences, enabling researchers to differentiate closely related microbial taxa that would otherwise be indistinguishable using short-read approaches (Buetas et al. 2024).

Long-read metatranscriptomics improves detection of low-abundance transcripts and enhanced resolution of full-length transcripts, which are critical for linking transcriptional activity to individual genomes or MAGs. This approach provides a real-time snapshot of gene expression, revealing the metabolic activity of microbial communities and their responses to dietary changes (Zhang et al. 2022c; Pitta et al. 2022), animal productivity (Xue et al. 2022), and breed differences (Li et al. 2019a; Zhang et al. 2020b). For instance, Yan et al. (2024) combined long-read sequencing with metatranscriptomic data to provide a genome-wide perspective of carbon mineralization pathways in an anaerobic bioreactor. Similarly, ONT's MinION has been utilized for real-time monitoring of microbial communities, achieving species-level resolution and supporting dynamic analyses in diverse environments (Kerkhof et al. 2017). These capabilities make long-read metatranscriptomics a powerful tool for investigating microbial processes and host-microbiome interactions.

Despite its advantages, long-read sequencing faces challenges, including high sequencing costs and computational demands. However, integrating long-read sequencing with hybrid approaches and complementary omics methods is expected to enhance its utility in unraveling microbial ecosystem dynamics. As these technologies continue to advance, their application in microbial genomics, taxonomy, and functional analysis will further expand, solidifying their role as essential tools in microbiome research.

Bioinformatics tools and reference databases

The application of sequencing technologies relies heavily on robust bioinformatics workflows and comprehensive reference databases. These tools are essential for processing, analyzing, and interpreting the vast amounts of data generated, ensuring accurate taxonomic classification and functional annotation.

Tools and database for amplicon sequencing

The 16 S rRNA gene sequencing, tools like QIIME2 (Bolyen et al. 2019) and DADA2 (Callahan et al. 2016) have become standard for processing and analyzing microbial diversity. These tools perform key functions such as filtering raw reads, denoising sequences, and generating ASVs, which offer higher resolution compared to traditional OTUs (Bharti and Grimm 2021). QIIME2 is particularly versatile, supporting plugin-based workflows that can integrate phylogenetic analysis and alpha/beta diversity metrics, while DADA2 excels in error correction and precise delineation of microbial populations.

The biological interpretation of the 16 S rRNA gene sequencing depends on the availability of accurate and comprehensive reference databases. For 16 S rRNA studies, general-purpose databases such as SILVA, RDP, and Greengenes2 provide robust frameworks for taxonomic classification (Quast et al. 2013; Cole et al. 2014). SILVA 138.2 (https://www.arb-silva.de/documentation/release-1382/) for example, offers expanded annotations and a comprehensive phylogenetic framework, while the newly revitalized Greengenes2 (McDonald et al. 2024) improves taxonomic resolution through updated algorithms and datasets. Despite these improvements, discrepancies in taxonomic assignments across databases remain a concern, underscoring the need for validation across multiple resources.

Tools and database for short-read metagenomics

Short-read sequencing technologies (read lengths < 250–300 base pairs) present several bioinformatics challenges, including fragmented assemblies due to difficulty in spanning repetitive or complex genomic regions, and the computational burden of managing massive data volumes (Das and Vikalo 2013). Andrew (2018) found that achieving high-quality de novo assemblies with short reads typically requires significantly more sequencing depth than long-read technologies. Ayling et al. (2020) have highlighted both the challenges and advancements in short-read metagenomics. The limited length of short reads can hinder the assembly of complex genomes, especially in metagenomic samples with high diversity.

Tools like Kraken2 (Wood et al. 2019), Sourmash (Irber et al. 2024), Skani (Shaw and Yu 2023), and MetaPhlAn4 (Blanco-Míguez et al. 2023) are widely used for taxonomic profiling. Kraken2 and Sourmash employ k-mer matching for rapid taxonomic classification, while MetaPhlAn4 focuses on clade-specific marker genes for high-resolution profiling of microbial communities. And Skani is a hashing-based tool for taxonomic profiling. Additionally, tools such as MEGAHIT (Li et al. 2015) and metaSPAdes (Nurk et al. 2017) are optimized for metagenomic assembly, enhancing the reconstruction of microbial genomes from short-read data. These bioinformatics workflows enable researchers to link microbial composition with functional potential, a key objective in rumen microbiome research (Bharti and Grimm 2021).

Accurate sequence annotation is another crucial component of metagenomic analysis. Three core strategies are commonly employed: (1) Similarity-based methods, such as BLAST (Altschul et al. 1990) and DIAMOND (Buchfink et al. 2015), which align sequences to known genes in reference databases; (2) Profile Hidden Markov Models (HMMs), used in tools like HMMER (Finn et al. 2011), which detect conserved protein domains; and (3) Structure-based approaches, such as AlphaFold (Jumper et al. 2021), which predict protein tertiary structures to infer potential biological functions. Combining these approaches helps to improve annotation accuracy, particularly when working with unknown or poorly characterized microbial genomes.

For genome-resolved metagenomics, binning methods are essential for reconstructing MAGs. Common binning tools include MetaBAT2 (Kang et al. 2019), MaxBin2 (Wu et al. 2016), and CONCOCT (Alneberg et al. 2014), which group assembled contigs based on sequence composition, coverage depth, and other genomic features. These MAGs provide high-resolution insights into the metabolic potential of uncultured or novel rumen taxa. For downstream functional annotation of MAGs, tools like DRAM (Shaffer et al. 2020) and METABOLIC (Zhou et al. 2022) are increasingly used to predict gene functions, reconstruct metabolic pathways, and assess ecological roles, including those related to methane production and fiber degradation.

The choice of reference databases in microbiome studies and human health can influence the outcomes (Smith et al. 2022; Szopinska-Tokov et al. 2023). This limitation can result in incomplete or inaccurate identification of microbial taxa and their functional genes. In metagenomics studies they rely on comprehensive reference databases for both taxonomic and functional annotation. Widely used metagenome databases, such as KEGG (Kanehisa et al. 2021), EggNOG (Cantalapiedra et al. 2021), NCBI RefSeq (Goldfarb et al. 2025), and GTDB (Parks et al. 2022), are invaluable for analyzing microbial genomes. However, these databases often fall short in representing rumen-specific microbes and their unique metabolic pathways (Smith et al. 2022). For instance, KEGG is effective for metabolic pathway analysis but may not capture novel enzymes or pathways unique to the rumen environment. Similarly, GTDB provides standardized phylogenetic classifications but offers limited taxonomic annotations. Resources such as Hungate1000 and the CAZy functional annotation-based Rumen Microbial Gene Catalogue (RMG) have been developed to address this gap, offering curated genomic data specific to rumen environments (Li et al. 2020). These specialized databases are critical for studying key processes like fiber degradation and methane production.

The bioinformatics tools for long-read sequencing

Long-read sequencing relies on the same reference databases as short-read metagenomics, including KEGG, GTDB, and NCBI RefSeq, for exploring taxonomy and functional properties. However, the key distinction lies in the bioinformatics tools required for analysis. Long-read sequencing demands specialized tools to handle higher

error rates, longer read lengths, and greater structural variation insights. Unlike short-read sequencing, which relies on tools optimized for fragmented assemblies, long-read workflows prioritize raw read correction, genome assembly, and functional annotation (Amarasinghe et al. 2020). For base-calling and error correction, ONT relies on tools such as Guppy, Dorado, and Bonito, which leverage neural network models to improve read accuracy (Wick et al. 2019). Similarly, PacBio utilizes DeepConsensus (Baid et al. 2023) to enhance base-calling accuracy for HiFi sequencing system, while SequalTools (Hufnagel et al. 2020) can offer quality control and read subsampling functionalities. These base-calling tools and computational resources are integral to processing PacBio sequencing data, ensuring high accuracy and facilitating downstream analyses.

For assembly, tools like metaFlye (Kolmogorov et al. 2020) and Canu (Koren et al. 2017) are widely used to construct high-contiguity genomes from long reads. Hybrid assembly tools such as Unicycler (Wick et al. 2017) integrate short- and long-read data to maximize accuracy and completeness. For structural variation detection, tools such as Sniffles2 (Smolka et al. 2024) and SVIM (Heller and Vingron 2019) help identify large genomic rearrangements, which are often missed by short-read sequencing. However, as long-read sequencing continues to improve, there is a growing need for databases optimized for high-contiguity genome assemblies and structural variant annotations.

Future advancements should focus on expanding niche-specific databases to include underrepresented taxa and improving hybrid workflows that integrate short- and long-read sequencing data. Leveraging machine learning models to predict functional capabilities from taxonomic profiles can further enhance the interpretability of microbiome data. By addressing current gaps in bioinformatics and database coverage, researchers can achieve a more detailed and accurate understanding of microbial dynamics, driving progress in livestock productivity and sustainability research.

Omics and multi-omics approaches: applications and integration in Microbiome research

Understanding the structural and functional properties of microbial communities requires comprehensive and integrative approaches. Multi-omics analyses have emerged as powerful tools, combining data from various omics technologies to investigate microbiomes in diverse environments. These approaches would enable us to figure out the fundamental biological questions such as which metabolites are presented under specific conditions (metabolomics), and which genes are expressed into RNA (metatranscriptomics) and translated into proteins (metaproteomics). Combining these omics methods enables a better understanding of the functional processes occurring in the system under study and provides much richer information for predictive models of the research aims (Jansson and Baker 2016) (Fig. 4).

Metabolomic: decoding microbial metabolic outputs

Metabolomics focuses on identifying and quantifying metabolites produced by the microbiome, providing a direct measure of metabolic outputs and linking microbial activity to functional outcomes. For instance, the production of short-chain fatty acids (SCFAs), such as acetate, propionate, and butyrate, plays a critical role in host energy metabolism and productivity (Wang et al. 2024b). These SCFAs are key indicators of microbial interactions with the host, reflecting the efficiency of fiber degradation and fermentation processes.

The diversity and dynamic range of metabolites in the rumen pose challenges for comprehensive profiling. Advanced analytical techniques, such as nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry are essential for accurate identification and quantification of metabolites. For example, Xue et al. (2022) identified six carbohydrate-related metabolites that differentiated efficiently from inefficient microbiomes with high predictive accuracy, demonstrating how metabolomics can uncover markers of feed efficiency. Integrating metabolomic data with other omics layers enhances our understanding of the metabolic interactions between the host and its microbiome.

Metatranscriptomic sequencing: Understanding Real-Time functionality

Metatranscriptomics enables the profiling of actively transcribed genes in microbial communities, offering a dynamic view of metabolic activity and host-microbiome interactions at a specific time point (Aguiar-Pulido et al. 2016; Ojala et al. 2023). This approach is instrumental in uncovering how microbial functions respond to environmental and dietary stimuli and help identify the expressed functional capacity within the rumen ecosystem (Wang et al. 2024b).

It is important to note that transcript abundance may not directly correlate with protein expression or functional activity due to post-transcriptional regulation, making it essential to interpret metatranscriptomic data in the context of complementary proteomic or metabolomic evidence (Maier et al. 2009; Edfors et al. 2016).

Metatranscriptomics has revealed key metabolic pathways within the rumen microbiome. For instance, Mann et al. (2018) identified 1,607 functional features, including central metabolism, galactose metabolism, and starch and sucrose metabolism pathways in rumen wall microbial

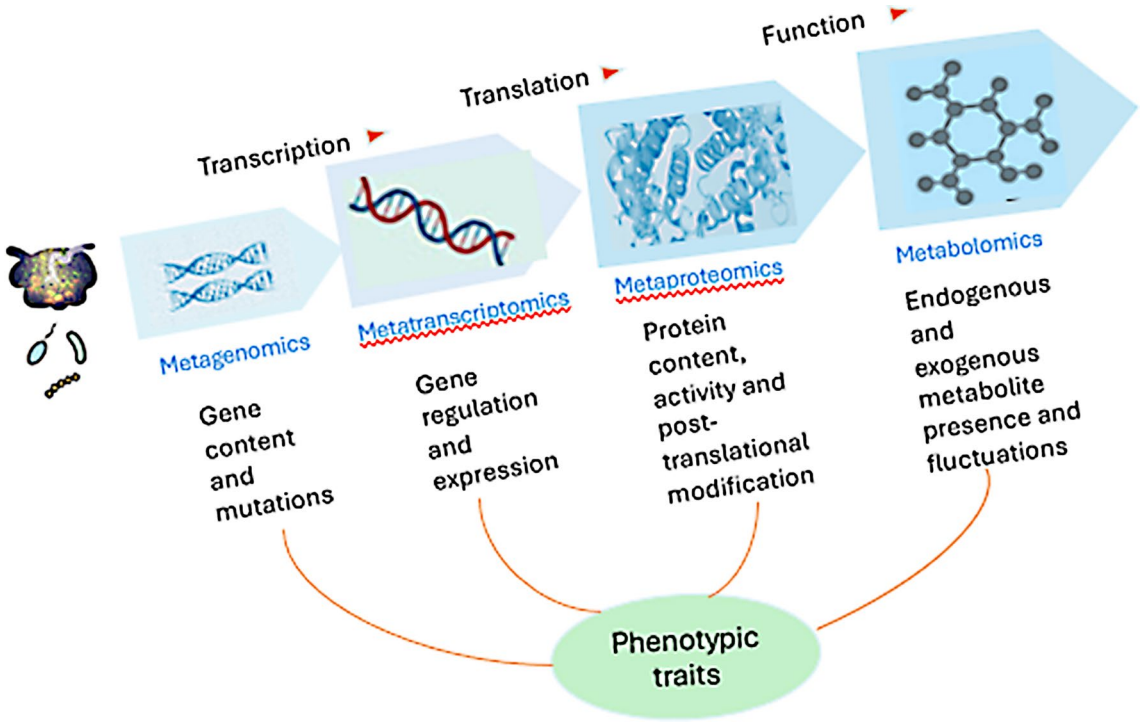

Fig. 4 A schematic illustrating the information from multi-omics techniques and how their integration is essential to understanding the phenotypic traits

communities. Similarly, Xie et al. (2022) demonstrated the upregulation of pathways related to fiber degradation, glycolysis, and the pentose phosphate pathway, highlighting the method's ability to characterize active microbial processes and nutrient cross-sharing within the rumen ecosystem. These findings illustrate how metatranscriptomics can investigate diet-driven changes, fiber degradation efficiency, and microbial metabolic interactions.

Moreover, this method enables the investigation of specific microbial roles. Fiber-degrading bacteria, such as *Ruminococcus flavefaciens*, *Ruminococcus albus*, and *Fibrobacter succinogenes* play critical roles in breaking down cellulose and hemicellulose. Similarly, amylolytic bacteria like *Prevotella ruminicola* and *Streptococcus bovis* metabolize starch into simpler sugars, while protein-degrading bacteria including *Ruminobacter amylophilus*, *Butyrivibrio fibrisolvens*, and *Clostridium spp*. synthesize microbial proteins essential for the host's nutrition (Dolan and Chang 2017; Cerqueira et al. 2020; Wang et al. 2020; Wei et al. 2022; Zhang et al. 2022b). For functional annotation of carbohydrate-active enzymes involved in fiber degradation, specialized resources such as the CAZy database (Cantarel et al. 2009) and rumen-specific reference collections like RUS-refDB are increasingly important for improving the accuracy and depth of metagenomic and metatranscriptomic analyses. This approach provides valuable insights into microbial functions relevant to rumen processes—such as methane emissions, feed efficiency, and fiber degradation—as demonstrated in several informative publications about the rumen metatranscriptome (Li et al. 2019a; Martínez-Álvaro et al. 2022; Xie et al. 2022; Malik et al. 2023). For instance, Xue et al. (2022) identified *Selenomonas* as a keystone taxon in carbohydrate metabolism, emphasizing its ecological importance within the rumen. Consequently, metatranscriptomics is a powerful tool for understanding the active functional dynamics of the rumen microbiome.

Despite its advantages, metatranscriptomics cannot on their own confirm protein expression or enzymatic activity, and its insights are most robust when integrated with metaproteomics and metabolomics. When integrated with metagenomic and metabolomic datasets, it provides a comprehensive framework for exploring the interplay between microbial activity and host productivity, enabling a deeper understanding of the rumen's ecological and functional complexity.

Metaproteomic: functional insights through proteins

Metaproteomic identifies and quantifies proteins in the rumen microbiome, revealing insights into the functional proteins produced and their roles in microbial activity. By directly measuring proteins involved in metabolic

processes, it provides a more immediate view of realized microbial function and host-microbiome interactions, such as enzymes involved in plant fiber degradation or proteins influencing host immunity. For instance, the identification of enzymes related to carbohydrate metabolism can help elucidate how the microbiome processes dietary component (Hagen et al. 2021; Andersen et al. 2023).

Recent advances in mass spectrometry and bioinformatics have greatly expanded the scope and sensitivity of metaproteomic analyses, enabling more comprehensive profiling of complex microbial communities (Van Den Bossche et al. 2021; Stamboulian et al. 2022). This technique is critical for linking microbial activity to biochemical pathways and understanding host-microbiome interactions at a mechanistic level. However, the complex nature of proteins and their post-translational modifications presents analytical challenges. Additionally, low-abundance proteins may evade detection, potentially missing important functional insights. Advanced analytical workflows and expanding protein databases specific to the rumen microbiome will further enhance the accuracy and depth of metaproteomic studies.

Integrating multi-omics data: A holistic perspective

Advancements in sequencing and multi-omics technologies have further facilitated the investigation of host-microbiome interactions. By integrating genome, transcriptome, and metabolome, researchers can uncover specific microbial pathways associated with breed-specific traits such as feed efficiency or methane mitigation (Nyholm et al. 2020; Zhu et al. 2022). Multi-omics approaches also enable the identification of microbial populations that drive productivity in different breeds, offering actionable insights for breeding and management strategies. For example, identifying breeds with microbial communities favoring reduced hydrogen availability to methanogens could naturally lower methane emissions while enhancing feed efficiency (Zhang et al. 2022a). These approaches offer actionable insights for selective breeding, enabling the development of livestock with optimized microbiota for enhanced productivity and reduced environmental impact.

Metagenomics uncovers the genetic blueprint of microbial communities, while metatranscriptomics identifies actively expressed genes under specific conditions (Aguiar-Pulido et al. 2016; Ojala et al. 2023). Proteomics quantifies functional proteins, linking gene expression to biochemical pathways and metabolomics measures metabolic products, revealing the outcomes of microbial activity (Reel et al. 2021). Xue et al. (2020) demonstrated the power of multi-omics by linking the rumen metagenome and metabolome to the serum metabolome in high-milk-yield cows. Their findings indicated that *Prevotella* species contributed to amino acid biosynthesis, correlating with improved productivity. Similarly, Kolli et al. (2023) mapped host-microbiome networks using integrated metagenomic and metabolomic data, elucidating key microbial interactions driving nutrient utilization.

One of the most promising applications of multi-omics could be its use in selective breeding for sustainability. Identifying breeds with microbial communities favoring reduced hydrogen availability to methanogens could naturally lower methane emissions while improving feed efficiency (Zhang et al. 2022a). These insights enable the development of livestock with optimized microbiota, enhancing both productivity and environmental sustainability. Future research should explore underrepresented microbial groups, such as viruses and fungi, which may also influence host metabolism and performance. Expanding reference databases like Hungate1000 and incorporating diverse livestock breeds into studies will further strengthen our ability to tackle global challenges in ruminant production.

Despite its advantages, multi-omics integration presents challenges, including the vast complexity of data and the need for robust computational frameworks. High-performance bioinformatics tools are essential to manage, process, and interpret the large datasets generated, enabling researchers to map complex microbial interactions with greater precision.

Machine learning in multi-omics integration

Machine learning (ML) has become an indispensable tool for integrating and analyzing multi-omics data (Reel et al. 2021; Kang et al. 2022). By uncovering patterns and relationships within complex datasets, ML algorithms enhance the predictive power of multi-omics approaches. For example, Rabaglino et al. (2023) demonstrated that transcriptomic data with ML algorithms could accurately predict embryonic competence in cattle, improving reproductive efficiency. Similarly, Peng et al. (2023), used random forest regression to predict methane emissions and identify significant microbial contributors in Holstein cows.

Xue et al. (2022) employed random forest models to predict ruminal metabolites associated with feed efficiency, achieving a prediction accuracy of 95.06%. These studies illustrate how ML can optimize data integration and uncover actionable insights for improving livestock health and productivity. However, the successful application of ML requires well-curated training datasets and robust validation strategies to ensure reliable and reproducible results.

Moving forward, researchers should focus on developing computational frameworks to streamline multi-omics integration and expanding rumen-specific reference databases for proteomics and metabolomics. By combining

contemporary sequencing technologies with advanced analytics, multi-omics approaches can drive innovations in ruminant health and productivity, paving the way for more sustainable livestock systems.

Host and microbiota interactions

The intricate interactions between the hosts and their microbiomes are crucial for various biological processes, including metabolism, immunity, and overall health. These interactions occur through direct and indirect mechanisms, including microbial metabolism of dietary components, production of bioactive molecules (Nicholson et al. 2012), immune system modulation (Belkaid and Hand 2014), and signaling pathways that influence host gene expression (Thaiss et al. 2016). The host, in turn, shapes the microbiome through factors such as immune responses, gut barrier function, and dietary influences, creating a dynamic and reciprocal relationship (Hooper et al. 2012).

Breakthroughs in sequencing and multi-omics technologies have significantly expanded our understanding of host-microbiome interactions, revealing key functional and ecological relationships across humans, model organisms, and livestock species, particularly cattle (Nyholm et al. 2020; Xue et al. 2020; Zhu et al. 2022). However, despite the recognized importance of these interactions, their mechanistic underpinnings remain largely understudied (Wang et al. 2024b). Understanding how host genetics, microbiome composition, and environmental factors interact is essential for developing targeted interventions that optimize animal health, feed efficiency, and methane mitigation strategies. The section will introduce the advance of the understanding of host and microbiota interactions, many of which are derived from the integration of different layers of information.

Host genome's influence on the Microbiome composition

Studies comparing cattle breeds have revealed distinct microbial compositions that reflect breed-specific characteristics. For instance, Li et al. (2019a) compared three cattle breeds (Angus, Charolais, and Kinsella composite hybrid), finding distinct differences in their rumen microbial communities, even when breeds were fed the same diet (Fig. 5). This suggests that host genomes have an intrinsic influence on microbial populations, independent of external factors like diet. These microbial distinctions are critical because they directly contribute to breed-specific traits such as milk production and fat composition. Holsteins, known for their high milk yield, exhibit higher abundances of genera like *Ethanoligenens* and *Succinivibrio*, which are associated with efficient fermentation pathways and enhanced VFA production. In contrast, Jersey cows, which excel in milkfat production, harbor microbial profiles better suited for lipid metabolism, highlighting how microbiota are tailored to breed-specific needs (Wang et al. 2024a).

Such breed-specific differences extend beyond bacterial communities to the activity of archaea, particularly methanogens like *Methanobrevibacter*. Methanogens utilize hydrogen produced by other microbes to generate methane; a process influenced by the host's breed. Research has shown that breeds such as Angus and Charolais tend to have microbial communities characterized by higher relative

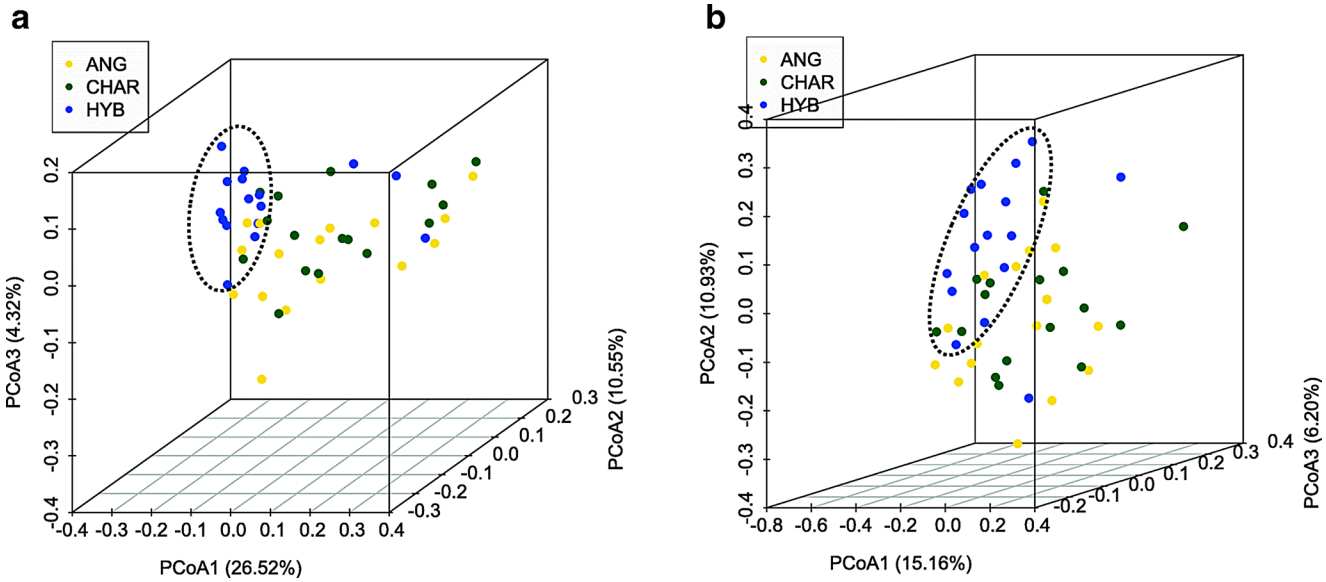

Fig. 5 Host breeds influence rumen microbial composition. The top three PCoAs were plotted for bacteria (**a**) and archaea (**b**). Adapted from Li et al. (2019b).

abundances of bacteria, archaea, and protozoa, whereas the Kinsella Composite hybrid exhibits microbiota that supports reduced hydrogen availability to methanogens, potentially lowering methane emissions (Zhang et al. 2022a). This demonstrates that breed-related microbial differences can be leveraged to address environmental concerns, providing opportunities to select breeds with microbiota that align with sustainability goals.

Another critical aspect is the influence of viruses and phages within the microbial ecosystem. Recent research demonstrated that bacteriophages play a role in shaping microbial community dynamics through lysis and horizontal gene transfer, processes that may differ between breeds. Understanding these microbial and viral interactions could reveal additional layers of breed-specific microbial regulation (Fan et al. 2020). Meanwhile, core microbial taxa, including *Unclassified Clostridiales*, *Butyrivibrio*, and *Fibrobacter*, vary by host genetics, affecting rumen function. Differences in microbial diversity between breeds like Holstein and Jersey are linked to traits such as feed efficiency and methane emissions (Paz et al. 2016; Noel et al. 2019; Islam et al. 2021; Olijhoek et al. 2022). These findings highlight the potential for selective breeding programs to optimize microbial traits for improved productivity and environmental sustainability.

Microbiome genome-wide association studies (mGWAS)

MGWAS link host genetic markers to variation in microbial composition, providing valuable insights into microbiota contributions to traits like feed efficiency and methane emissions (Gilbert et al. 2016; Uffelmann et al. 2021). This has been facilitated by extensive scientific collaboration and the exponential increase of technical and methodological advancements illustrated in Fig. 6 (Awany D, 2019). While extensively used in human studies to explore microbiota-disease relationships, its application in livestock remains underexplored but highly promising. For example, the *LCT* gene locus in humans, associated with lactose tolerance, is linked to an increased abundance of *Bifidobacterium*, demonstrating a direct host-microbiome interaction (López-García et al. 2022). Similarly, mGWAS in pigs has revealed associations between fecal microbiota and feed efficiency (Aliakbari et al. 2022).

In ruminants, applications of mGWAS remained limited so far. Boggio et al. (2023) conducted studies in sheep,

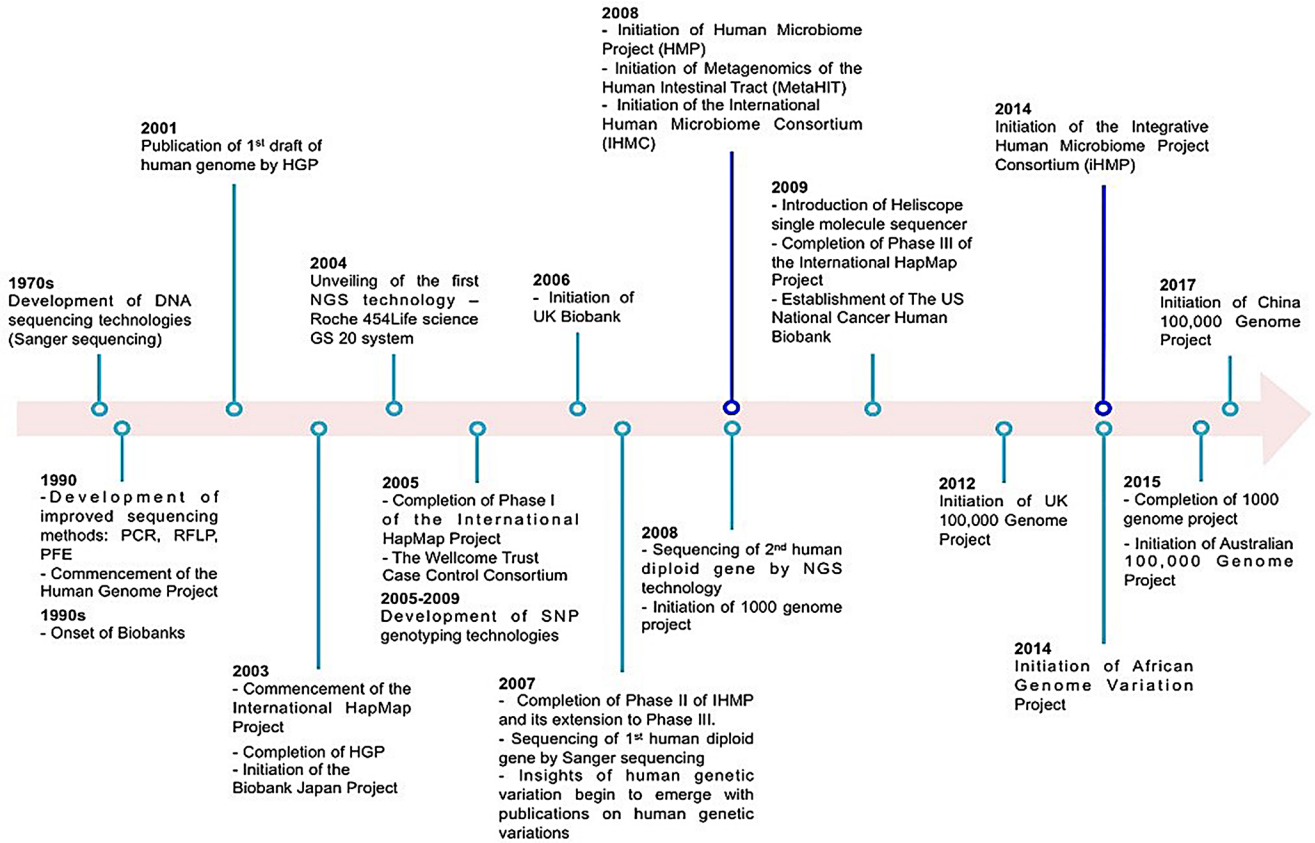

Fig. 6 Timetable of partial major technological and large-scale collaborative projects (excluding data repositories) on host and mGWAS adapted from Awany D (2019)

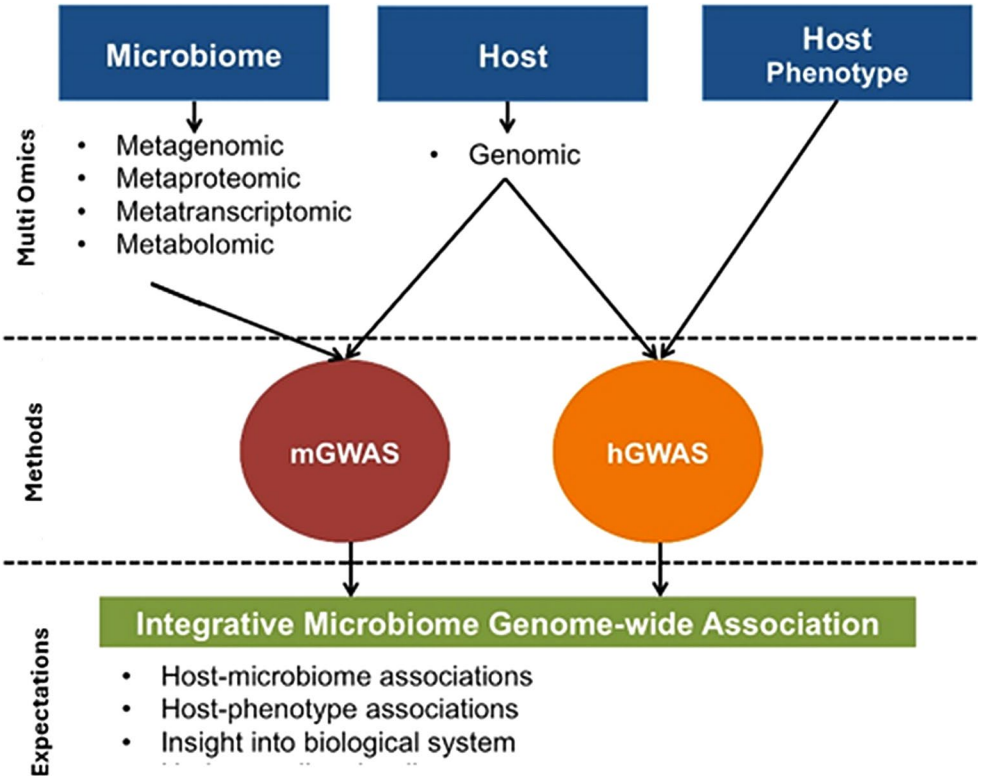

Fig. 7 Illustrative representation of future integrating mGWAS with multi-omics data to accelerate selective breeding for host-microbiome associations, adapted from Awany D (2019)

showing that microbial contributions significantly influence dairy traits. These analyses, conducted across taxonomic levels (phylum, genus, species) or using ASVs, reveal the profound impact of host genetics on the microbiome's structure and functionality (Abdul-Aziz et al. 2016). Additionally, advanced statistical tools like microbiome-specific pipelines (Hua et al. 2022) and models tailored to microbiome data, such as those used by Rothschild D (2018) have further refined these insights.

mGWAS challenges and future opportunities

Despite its potential, mGWAS in ruminants faces challenges, including the complexity of the rumen microbiome and the need for large, well-characterized datasets. Kurilshikov et al. (2021) demonstrated this by analyzing 18,340 human samples across 24 cohorts to identify 31 genetic loci affecting microbiome composition, highlighting the scale of data required for robust associations. Compared to human studies, long-range linkage disequilibrium (LD) in cattle require specialized analytical approaches (Sahana et al. 2023). Similarly, marker-assisted selection (MAS) provides an additional framework to select animals with favorable microbiome profiles that contribute to health, productivity, and sustainability. For ruminants, integrating mGWAS with MAS offers an opportunity to select animals with favorable microbiome traits. For example, MAS has identified beneficial genetic variants in humans linked to obesity and diabetes (Cuevas-Sierra et al. 2019; Smoczek et al. 2020), an approach that could be adapted for livestock to enhance efficiency and reduce methane emissions.

Future research should prioritize large-scale mGWAS in cattle, leveraging multi-omics data integration and advanced computational frameworks for more accurate host-microbiota interaction mapping. Figure 7 illustrates how statistical modeling can be used to bridge the gap between host genetics and microbiome dynamics, enabling data-driven selective breeding strategies for sustainable livestock production, as discussed in detail in the following section.

Continued advancements in sequencing technology, machine learning-based association mapping, and large-scale data aggregation will be essential for realizing the full potential of mGWAS in ruminants. Standardized pipelines, improved reference databases, and enhanced statistical frameworks will further strengthen genetic selection programs, making microbiome-driven breeding a feasible approach for improving livestock health and productivity.

Prediction of the methane emission incorporating microbiome information

Recent advances in microbiome research and predictive modeling offer novel insights into host-microbiota interactions that influence methane production. Here, we explore the potential of microbiome-informed models for methane prediction, discussing key methodologies, statistical

frameworks, and future applications. The integration of host genetic data, microbiome composition, and environmental factors could pave the way for precision livestock breeding, optimizing microbial communities to reduce methane emissions while maintaining productivity.

Advances in predictive models

Methane emission from ruminant production is a major environmental concern and incorporating microbiome information into predictive models offers a promising solution. The rumen microbiome, particularly methanogenic archaea, plays a central role in methane production. By leveraging microbiome features such as diversity indices, taxonomic composition, and functional gene abundance alongside host genomic data, researchers have developed predictive models with improved accuracy.

Martínez-Álvaro et al. (2022) demonstrated that the bovine genome influences rumen microbiome functions linked to methane emissions, emphasizing the value of integrating host and microbiome data. Difford et al. (2018) demonstrated that rumen microbiome composition is a heritable trait and can be leveraged for genomic prediction of methane emissions in dairy cattle. Zhang et al. (2020a) further quantified this relationship, showing that host genetics account for 24% of methane variance, with an additional 7% explained by microbiota. Rowe et al. (2019) introduced a genotyping-by-sequencing (GBS) approach, linking microbial profiles to methane emissions and highlighting the heritability of microbial communities as a key factor in methane production. Recent studies have further refined microbiome-based predictive models. González-Recio et al. (2023) reviewed that the potential of incorporating microbiome data into selective breeding programs, improving trait selection for methane reduction. López-García et al. (2022) identified fungi and ciliate protozoa as key microbial groups associated with methane emissions, suggesting potential microbial targets for mitigation strategies.

These studies showed that incorporating microbiome profiles into genomic selection models improved the accuracy of methane emission predictions, suggesting that microbiome-informed breeding could be an effective strategy for reducing methane output in livestock.

Innovative techniques

Predictive frameworks now incorporate advanced statistical and machine learning methods to identify microbial biomarkers. For instance, Saborío-Montero et al. (2021) used principal component analysis (PCA) to identify microbial traits associated with methane emissions, achieving heritability estimates of up to 30%, which is notably higher than those typically reported in ruminant methane studies, such as 0-0.13 in sheep (Pinares-Patiño et al. 2013; Robinson et al. 2015; Jonker et al. 2018) and approximately 0.15–0.20 in beef and dairy cattle (Hayes et al. 2015; de Haas et al. 2016; Donoghue et al. 2016; Lassen and Løvendahl 2016). The relatively low heritability of the methane emission trait in sheep and cattle is thought to be partly due to variation in host rumen volume, liveweight, CO_2 production, feed intake and microbiome composition (Løvendahl et al. 2018; Martínez-Álvaro et al. 2022; Wang et al. 2023). Ross et al. (2020) reported that the accuracy of genomic predictions of methane production improved when metabolomic and microbiome data were integrated. Furthermore, machine learning approaches, such as random forest regression, enhance prediction accuracy by identifying microbial taxa and pathways that significantly contribute to methane variance (Wang et al. 2021).

Predictive models have practical implications for methane mitigation. For example, Wallace et al. (2019) identified a heritable core microbiome in cattle, providing biomarkers for selective breeding programs targeting reduced methane emissions. Pérez-Enciso et al. (2021) demonstrated that combining metagenomic and host genomic data improves predictive accuracy for methane-related traits, outperforming single-data approaches. Beyond genetics, network-based analyses have revealed associations between microbial communities and VFA production, offering new avenues for methane reduction strategies.

The integration of microbiome and host genomic data represents a transformative approach to sustainable livestock production. Future efforts should focus on expanding predictive models to include environmental factors and microbiome manipulation strategies, such as targeted dietary interventions or microbial engineering. These innovations align with global sustainability goals, enabling environmentally friendly livestock breeding while maintaining productivity.

Conclusion and perspectives

Integrating multi-omics approaches with host genetic data offers new opportunities for optimizing key traits in cattle, such as feed efficiency, methane emissions, and overall productivity. This review highlights the growing role of microbiome research in livestock breeding, emphasizing the potential of incorporating microbial data into genomic selection and predictive modeling. Advances in sequencing technologies, statistical modeling, and computational tools have enabled more precise characterization of host-microbiome interactions, paving the way for microbiome-informed breeding strategies.

Future research should prioritize machine learning frameworks to analyze the vast datasets generated by multi-omics studies, enhancing the accuracy of predictive models that link microbial traits with host-genomic markers. Additionally, genome-resolved metagenomic approaches will be crucial for assembling high quality microbial genomes directly from metagenomic databases, allowing for a more precise characterization of microbial functions and interactions. Meanwhile, the niche-specific reference databases will further improve the accuracy of taxonomic and functional annotations, reducing the reliance on general databases that may lack coverage for specific microbial communities.

Longitudinal studies considering dietary and environmental variability will be essential for capturing the dynamic nature of host-microbiome interactions over time. While mGWAS have already demonstrated the feasibility of identifying host genetic loci associated with microbiome composition, further refinement of statistical approaches and multi-omics integration will be crucial for practical applications in selective breeding programs.

By addressing these challenges, the integration of microbiome data with host-genetic and phenotypic information can drive more sustainable livestock management, contributing to climate change mitigation efforts while maintaining animal-based food production.

Acknowledgements We are grateful to our reviewers and editor for the helpful comments to improve the manuscript.

Author contributions XY wrote the manuscript. ZC and GS modified the manuscript. All authors reviewed the manuscript and approved to submit the mauscript.

Funding Not applicable.

Data availability No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

Abdul-Aziz MA, Cooper A, Weyrich LS (2016) Exploring relationships between host genome and microbiome: new insights from genome-Wide association studies. Front Microbiol 7. https://doi.org/10.3389/fmicb.2016.01611

Abellan-Schneyder I, Matchado MS, Reitmeier S et al (2021) Primer, pipelines, parameters: issues in 16S rRNA gene sequencing. mSphere 6:101128msphere01202–101128msphere01220. https://doi.org/10.1128/msphere.01202-20

Acinas SG, Sarma-Rupavtarm R, Klepac-Ceraj V, Polz MF (2005) Appl Environ Microbiol 71:8966–8969. https://doi.org/10.1128/AEM.71.12.8966-8969.2005. PCR-Induced Sequence Artifacts and Bias: Insights from Comparison of Two 16S rRNA Clone Libraries Constructed from the Same Sample

Aguiar-Pulido V, Huang W, Suarez-Ulloa V et al (2016) Metagenomics, metatranscriptomics, and metabolomics approaches for Microbiome analysis: supplementary issue: bioinformatics methods and applications for big metagenomics data. Evolutionary Bioinf 12. https://doi.org/10.4137/EBO.S36436. s1, EBO.S36436

Aliakbari A, Zemb O, Cauquil L et al (2022) Microbiability and microbiome-wide association analyses of feed efficiency and performance traits in pigs. Genet Selection Evol 54:29. https://doi.org/10.1186/s12711-022-00717-7

Alneberg J, Bjarnason BS, de Bruijn I et al (2014) Binning metagenomic contigs by coverage and composition. Nat Methods 11:1144–1146. https://doi.org/10.1038/nmeth.3103

Altschul SF, Gish W, Miller W et al (1990) Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Amarasinghe SL, Su S, Dong X et al (2020) Opportunities and challenges in long-read sequencing data analysis. Genome Biol 21:30. https://doi.org/10.1186/s13059-020-1935-5

Andersen TO, Altshuler I, Vera-Ponce de León A et al (2023) Metabolic influence of core ciliates within the rumen Microbiome. ISME J 17:1128–1140. https://doi.org/10.1038/s41396-023-01407-y

Andrew C (2018) The challenge of genome sequence assembly. Open Bioinforma J. https://doi.org/10.2174/1875036201811010231

Awany DAI, Dalvie S, Hemmings S et al (2019) Host and Microbiome Genome-Wide association studies: current state and challenges. Front Genet 9. https://doi.org/10.3389/fgene.2018.00637

Ayling M, Clark MD, Leggett RM (2020) New approaches for metagenome assembly with short reads. Brief Bioinform 21:584–594. https://doi.org/10.1093/bib/bbz020

Baid G, Cook DE, Shafin K et al (2023) DeepConsensus improves the accuracy of sequences with a gap-aware sequence transformer. Nat Biotechnol 41:232–238. https://doi.org/10.1038/s41587-022-01435-7

Belkaid Y, Hand TW (2014) Role of the microbiota in immunity and inflammation. Cell 157:121–141. https://doi.org/10.1016/j.cell.2014.03.011

Betancur-Murillo CL, Aguilar-Marín SB, Jovel J (2023) Prevotella: A key player in ruminal metabolism. https://doi.org/10.3390/microorganisms11010001. Microorganisms 11

Bharti R, Grimm DG (2021) Current challenges and best-practice protocols for Microbiome analysis. Brief Bioinform 22:178–193. https://doi.org/10.1093/bib/bbz155

Blanco-Míguez A, Beghini F, Cumbo F et al (2023) Extending and improving metagenomic taxonomic profiling with uncharacterized species using metaphlan 4. Nat Biotechnol 41:1633–1644. https://doi.org/10.1038/s41587-023-01688-w

Boggio GM, Monteiro HF, Lima FS et al (2023) Host and rumen Microbiome contributions to feed efficiency traits in Holstein cows. J Dairy Sci. https://doi.org/10.3168/jds.2023-23869

Bolyen E, Rideout JR, Dillon MR et al (2019) Reproducible, interactive, scalable and extensible Microbiome data science using QIIME 2. Nat Biotechnol 37:852–857. https://doi.org/10.1038/s41587-019-0209-9

Bubier JA, Chesler EI, Weinstock GM (2021) Host genetic control of gut Microbiome composition. Mamm Genome 32:263–281. https://doi.org/10.1007/s00335-021-09884-2

Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59–60. https://doi.org/10.1038/nmeth.3176

Buetas E, Jordán-López M, López-Roldán A et al (2024) Full-length 16S rRNA gene sequencing by PacBio improves taxonomic resolution in human Microbiome samples. BMC Genomics 25:310. https://doi.org/10.1186/s12864-024-10213-5

Callahan BJ, McMurdie PJ, Rosen MJ et al (2016) DADA2: High-resolution sample inference from illumina amplicon data. Nat Methods 13:581–583. https://doi.org/10.1038/nmeth.3869

Cantalapiedra CP, Hernández-Plaza A, Letunic I et al (2021) eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol Biol Evol 38:5825–5829. https://doi.org/10.1093/molbev/msab293

Cantarel BL, Coutinho PM, Rancurel C et al (2009) The Carbohydrate-Active enzymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res 37:D233–D238. https://doi.org/10.1093/nar/gkn663

Cerqueira FM, Photenhauer AL, Pollet RM et al (2020) Starch digestion by gut bacteria: crowdsourcing for carbs. Trends Microbiol 28:95–108. https://doi.org/10.1016/j.tim.2019.09.004

Chen P, Sun Z, Wang J et al (2023) Portable nanopore-sequencing technology: trends in development and applications. Front Microbiol Volume 14–2023. https://doi.org/10.3389/fmicb.2023.1043967

Choi J, Park JS (2020) Comparative analyses of the V4 and V9 regions of 18S rDNA for the extant eukaryotic community using the illumina platform. Sci Rep 10:6519. https://doi.org/10.1038/s41598-020-63561-z

Cole JR, Wang Q, Fish JA et al (2014) Ribosomal database project: data and tools for high throughput rRNA analysis. Nucleic Acids Res 42:D633–D642. https://doi.org/10.1093/nar/gkt1244

Cuevas-Sierra A, Ramos-Lopez O, Riezu-Boj JI et al (2019) Diet, gut microbiota, and obesity: links with host genetics and epigenetics and potential applications. Adv Nutr 10:S17–S30. https://doi.org/10.1093/advances/nmy078

Daghio M, Ciucci F, Buccioni A et al (2021) Correlation of breed, growth performance, and rumen microbiota in two rustic cattle breeds reared under different conditions. Front Microbiol 12. https://doi.org/10.3389/fmicb.2021.652031

Dao TK, Dao TK, Le NG et al (2021) Understanding the Role of Prevotella Genus in the Digestion of Lignocellulose and Other Substrates in Vietnamese Native Goats' Rumen by Metagenomic Deep Sequencing, Animals. https://doi.org/10.3390/ani11113257

Das S, Vikalo H (2013) Base calling for high-throughput short-read sequencing: dynamic programming solutions. BMC Bioinformatics 14:129. https://doi.org/10.1186/1471-2105-14-129

de Haas H, Natter K, Vezzoli S (2016) Growing restrictiveness or changing selection?? The nature and evolution of migration policies. Int Migrat Rev. https://doi.org/10.1111/imre.12288

Denman SE, Morgavi DP, McSweeney CS (2018) Review: the application of omics to rumen microbiota function. Animal 12:s233–s245. https://doi.org/10.1017/S175173111800229X

Difford GF, Plichta DR, Løvendahl P et al (2018) Host genetics and the rumen Microbiome jointly associate with methane emissions in dairy cows. PLoS Genet 14:e1007580. https://doi.org/10.1371/journal.pgen.1007580

Dolan KT, Chang EB (2017) Diet, gut microbes, and the pathogenesis of inflammatory bowel diseases. Mol Nutr Food Res 61:1600129. https://doi.org/10.1002/mnfr.201600129

Donoghue KA, Bird-Gardiner T, Arthur PF et al (2016) Genetic and phenotypic variance and covariance components for methane emission and postweaning traits in Angus cattle12. J Anim Sci 94:1438–1445. https://doi.org/10.2527/jas.2015-0065

Edfors F, Danielsson F, Hallström BM et al (2016) Gene-specific correlation of RNA and protein levels in human cells and tissues. Mol Syst Biol 12:883. https://doi.org/10.15252/msb.20167144

Fan P, Bian B, Teng L et al (2020) Host genetic effects upon the early gut microbiota in a bovine model with graduated spectrum of genetic variation. ISME J 14:302–317. https://doi.org/10.1038/s41396-019-0529-2

Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. Nucleic Acids Res 39:W29–W37. https://doi.org/10.1093/nar/gkr367

Fonseca PAS, Lam S, Chen Y et al (2023) Multi-breed host rumen epithelium transcriptome and Microbiome associations and their relationship with beef cattle feed efficiency. Sci Rep 13:16209. https://doi.org/10.1038/s41598-023-43097-8

Fregulia P, Campos MM, Dias RJP et al (2022) Taxonomic and predicted functional signatures reveal linkages between the rumen microbiota and feed efficiency in dairy cattle Raised in tropical areas. Front Microbiol 13. https://doi.org/10.3389/fmicb.2022.1025173

Gao S, Ren Y, Sun Y et al (2016) PacBio full-length transcriptome profiling of insect mitochondrial gene expression. RNA Biol 13(9):820–825. https://doi.org/10.1080/15476286.2016.1197481

García-Martínez J, Acinas SG, Antón AI, Rodríguez-Valera F (1999) Use of the 16S–23S ribosomal genes spacer region in studies of prokaryotic diversity. J Microbiol Methods 36(1):55–64. https://doi.org/10.1016/S0167-7012(99)00011-1

Gilbert JA, Quinn RA, Debelius J et al (2016) Microbiome-wide association studies link dynamic microbial consortia to disease. Nature 535:94–103. https://doi.org/10.1038/nature18850

Goldfarb T, Kodali VK, Pujar S et al (2025) NCBI refseq: reference sequence standards through 25 years of curation and annotation. Nucleic Acids Res 53:D243–D257. https://doi.org/10.1093/nar/gkae1038

González-Recio O, Martínez-Álvaro M, Tiezzi F et al (2023) Invited review: novel methods and perspectives for modulating the rumen Microbiome through selective breeding as a means to improve complex traits: implications for methane emissions in cattle. Livest Sci 269:105171. https://doi.org/10.1016/j.livsci.2023.105171

Greenman N, Hassouneh SA, Abdelli LS et al (2024) Improving Bacterial Metagenomic Research through Long-Read Sequencing, Microorganisms. https://doi.org/10.3390/microorganisms12050935

Gruninger RJ, Ribeiro GO, Cameron A, McAllister TA (2019) Invited review: application of meta-omics to understand the dynamic nature of the rumen Microbiome and how it responds to diet in ruminants. Animal 13:1843–1854. https://doi.org/10.1017/S1751731119000752

Hagen LH, Brooke CG, Shaw CA et al (2021) Proteome specialization of anaerobic fungi during ruminal degradation of recalcitrant plant fiber. ISME J 15:421–434. https://doi.org/10.1038/s41396-020-00769-x

Handelsman J (2004) Metagenomics: application of genomics to uncultured microorganisms. Microbiol Mol Biol Rev 68:669–685. https://doi.org/10.1128/mmbr.68.4.669-685.2004

Harirchi S, Wainaina S, Sar T et al (2022) Microbiological insights into anaerobic digestion for biogas, hydrogen or volatile fatty acids (VFAs): a review. Bioengineered 13:6521–6557. https://doi.org/10.1080/21655979.2022.2035986

Hayes BJ, Donoghue KA, Reich C et al (2015) Genomic estimated breeding values for methane production in Australian beef cattle, Feed efficiency, methane emission and fertility, pp. 118–121

Heller D, Vingron M (2019) SVIM: structural variant identification using mapped long reads. Bioinformatics 35:2907–2915. https://doi.org/10.1093/bioinformatics/btz041

Henderson G, Cox F, Ganes, S et al (2015) Rumen microbial community composition varies with diet and host, but a core Microbiome is found across a wide geographical range. Sci Rep 5:14567. https://doi.org/10.1038/srep14567

Hernandez-Sanabria E, Goonewardene LA, Wang Z et al (2013) Influence of sire breed on the interplay among rumen microbial populations inhabiting the rumen liquid of the progeny in beef cattle. PLoS ONE 8:e58461. https://doi.org/10.1371/journal.pone.0058461

Hooper LV, Littman DR, Macpherson AJ (2012) Interactions between the microbiota and the immune system. Science 336:1268–1273. https://doi.org/10.1126/science.1223490

Hua X, Song L, Yu G et al (2022) MicrobiomeGWAS: A tool for identifying host genetic variants associated with Microbiome composition. Genes 13(7):1224. https://doi.org/10.3390/genes13071224

Hufnagel DE, Hufford MB, Seetharam AS (2020) SequelTools: a suite of tools for working with PacBio sequel Raw sequence data. BMC Bioinformatics 21:429. https://doi.org/10.1186/s12859-020-03751-8

Irber L, Pierce-Ward NT, Abuelanin M et al (2024) Sourmash v4: A multitool to quickly search, compare, and analyze genomic and metagenomic data sets. J Open Source Softw 9:6830. https://doi.org/10.21105/joss.06830

Islam M, Kim SH, Ramos SC et al (2021) Holstein and Jersey steers differ in rumen microbiota and enteric methane emissions even fed the same total mixed ration. Front Microbiol 12. https://doi.org/10.3389/fmicb.2021.601061

Ismail MA, Jiang H, Mudasir N et al (2023) Identification of candidate genes and functional pathways associated with body size traits in Chinese Holstein cattle based on GWAS analysis. https://doi.org/10.3390/ani13060992. Animals 13

Janssen PH, Kirs M (2008) Structure of the archaeal community of the rumen. Appl Environ Microbiol 74:3619–3625. https://journals.asm.org/doi/abs/ https://doi.org/10.1128/aem.02812-07

Jansson JK, Baker ES (2016) A multi-omic future for Microbiome studies. Nat Microbiol 1:16049. https://doi.org/10.1038/nmicrobiol.2016.49

Jonker A, Hickey SM, Rowe SJ et al (2018) Genetic parameters of methane emissions determined using portable accumulation chambers in lambs and Ewes grazing pasture and genetic correlations with emissions determined in respiration chambers1. J Anim Sci 96:3031–3042. https://doi.org/10.1093/jas/sky187

Jumper J, Evans R, Pritzel A et al (2021) Highly accurate protein structure prediction with alphafold. Nature 596:583–589. https://doi.org/10.1038/s41586-021-03819-2

Kanehisa M, Furumichi M, Sato Y (2021) KEGG: integrating viruses and cellular organisms. Nucleic Acids Res 49:D545–D551. https://doi.org/10.1093/nar/gkaa970

Kang DD, Li F, Kirton E et al (2019) MetaBAT 2: an adaptive Binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 7:e7359. https://doi.org/10.7717/peerj.7359

Kang M, Ko E, Mersha TB (2022) A roadmap for multi-omics data integration using deep learning. Brief Bioinform 23:bbab454. https://doi.org/10.1093/bib/bbab454

Kerkhof LJ, Dillon KP, Häggblom MM, McGuinness LR (2017) Profiling bacterial communities by minion sequencing of ribosomal operons. Microbiome 5:116. https://doi.org/10.1186/s40168-017-0336-9

Kim M, Morrison M, Yu Z (2011) Status of the phylogenetic diversity census of ruminal microbiomes. FEMS Microbiol Ecol 76:49–63. https://doi.org/10.1111/j.1574-6941.2010.01029.x

Kolli U, Jalodia R, Moidunny S et al (2023) Multi-omics analysis revealing the interplay between gut Microbiome and the host following opioid use. Gut Microbes 15:2246184. https://doi.org/10.1080/19490976.2023.2246184

Kolmogorov M, Bickhart DM, Behsaz B et al (2020) MetaFlye: scalable long-read metagenome assembly using repeat graphs. Nat Methods 17:1103–1110. https://doi.org/10.1038/s41592-020-00971-x

Koren S, Walenz BP, Berlin K et al (2017) Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 27:722–736. https://doi.org/10.1101/gr.215087.116

Kupczyński R, Pacyga K, Lewandowska K et al (2024) Milk Odd- and Branched-Chain fatty acids as biomarkers of rumen fermentation. Animals 14(11):1706. https://doi.org/10.3390/ani14111706

Kurilshikov A, Medina-Gomez C, Bacigalupe R et al (2021) Large-scale association analyses identify host factors influencing human gut Microbiome composition. Nat Genet 53:156–165. https://doi.org/10.1038/s41588-020-00763-1

Lang D, Zhang S, Ren P et al (2020) Comparison of the two up-to-date sequencing technologies for genome assembly: HiFi reads of Pacific biosciences sequel II system and ultralong reads of Oxford nanopore. Gigascience 9(12):giaa123. https://doi.org/10.1093/gigascience/giaa123

Lassen J, Løvendahl P (2016) Heritability estimates for enteric methane emissions from Holstein cattle measured using noninvasive methods. J Dairy Sci 99:1959–1967. https://doi.org/10.3168/jds.2015-10012

Lawther K, Santos FG, Oyama LB, Huws SA (2024) Invited review: chemical signalling within the rumen Microbiome. Anim Bioscience 37:337–345. https://doi.org/10.5713/ab.23.0374

Lee SS, Ha JK, Cheng K (2000) Relative contributions of bacteria, protozoa, and fungi to in vitro degradation of orchard grass cell walls and their interactions. Appl Environ Microbiol 66:3807–3813. https://doi.org/10.1128/aem.66.9.3807-3813.2000

Leeming ER, Johnson AJ, Spector TD, Le Roy CI (2019) Effect of diet on the gut microbiota: rethinking intervention duration. Nutrients 11(12):2862. https://doi.org/10.3390/nu11122862

Li D, Liu CM, Luo R et al (2015) MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31:1674–1676. https://doi.org/10.1093/bioinformatics/btv033

Li F, Hitch TCA, Chen Y et al (2019a) Comparative metagenomic and metatranscriptomic analyses reveal the breed effect on the rumen Microbiome and its associations with feed efficiency in beef cattle. Microbiome 7:6. https://doi.org/10.1186/s40168-019-0618-5

Li F, Li C, Chen Y et al (2019b) Host genetics influence the rumen microbiota and heritable rumen microbial features associate with feed efficiency in cattle. Microbiome 7:92. https://doi.org/10.1186/s40168-019-0699-1

Li J, Zhong H, Ramayo-Caldas Y et al (2020) A catalog of microbial genes from the bovine rumen unveils a specialized and diverse biomass-degrading environment. GigaScience 9, giaa057. https://doi.org/10.1093/gigascience/giaa057

Li Z, Wang X, Zhang Y et al (2022) Genomic insights into the phylogeny and biomass-degrading enzymes of rumen ciliates. ISME J 16:2775–2787. https://doi.org/10.1038/s41396-022-01306-8

Liu K, Zhang Y, Yu Z et al (2021) Ruminal microbiota–host interaction and its effect on nutrient metabolism. Anim Nutr 7:49–55. https://doi.org/10.1016/j.aninu.2020.12.001

López-García A, Saborío-Montero A, Gutiérrez-Rivas M et al (2022) Fungal and ciliate protozoa are the main rumen microbes associated with methane emissions in dairy cattle. GigaScience 11:giab088. https://doi.org/10.1093/gigascience/giab088

Løvendahl P, Difford GF, Li B et al (2018) Review: selecting for improved feed efficiency and reduced methane emissions in dairy cattle. Animal 12:s336–s349. https://doi.org/10.1017/S1751731118002276

Maidak BL, Olsen GJ, Larsen N et al (1996) The ribosomal database project (RDP). Nucleic Acids Res 24:82–85. https://doi.org/10.1093/nar/24.1.82

Maier T, Güell M, Serrano L (2009) Correlation of mRNA and protein in complex biological samples. FEBS Lett 583:3966–3973. https://doi.org/10.1016/j.febslet.2009.10.036

Malik PK, Trivedi S, Kolte AP et al (2023) Comparative analysis of rumen metagenome, metatranscriptome, fermentation and methane yield in cattle and buffaloes fed on the same diet. Front Microbiol 14. https://doi.org/10.3389/fmicb.2023.1266025

Mann E, Wetzels SU, Wagner M et al (2018) Metatranscriptome sequencing reveals insights into the gene expression and functional potential of rumen wall Bacteria. Front Microbiol. 9https://doi.org/10.3389/fmicb.2018.00043

Martínez-Álvaro M, Auffret MD, Duthie CA et al (2022) Bovine host genome acts on rumen Microbiome function linked to methane emissions. Commun Biology 5:350. https://doi.org/10.1038/s42003-022-03293-0

Matthews C, Crispie F, Lewis E et al (2019) The rumen microbiome: a crucial consideration when optimising milk and meat production and nitrogen utilisation efficiency. Gut Microbes 10:115–132. https://doi.org/10.1080/19490976.2018.1505176

McAllister TA, Thomas KD, Gruninger RJ et al (2025) International Symposium on Ruminant Physiology: Rumen fungi, archaea, and their interactions. Journal of Dairy Science 108, 7545–7566. https://doi.org/10.3168/jds.2024-25713

McCann JC, Wickersham TA, Loor JJ (2014) High-throughput methods redefine the rumen Microbiome and its relationship with nutrition and metabolism. Bioinform Biol Insights 8:109–125. https://doi.org/10.4137/bbi.S15389

McDonald D, Jiang Y, Balaban M (2024) Greengenes2 unifies microbial data in a single reference tree. Nat Biotechnol 42:715–718. https://doi.org/10.1038/s41587-023-01845-1

Mi J, Jing X, Ma C et al (2024) A metagenomic catalogue of the ruminant gut archaeome. Nat Commun 15:9609. https://doi.org/10.1038/s41467-024-54025-3

Nicholson JK, Holmes E, Kinross J et al (2012) Host-Gut microbiota metabolic interactions. Science 336:1262–1267. https://doi.org/10.1126/science.1223813

Noel SJ, Olijhoek DW, Mclean F et al (2019) Rumen and fecal microbial community structure of Holstein and Jersey dairy cows as affected by breed, diet, and residual feed intake. Animal 9(8):498. https://doi.org/10.3390/ani9080498

Nurk S, Meleshko D, Korobeynikov A, Pevzner PA (2017) MetaSPAdes: a new versatile metagenomic assembler. Genome Res 27:824–834. https://doi.org/10.1101/gr.213959.116

Nyholm L, Koziol A, Marcos S (2020) Holo-Omics: integrated Host-Microbiota Multi-omics for basic and applied biological research. iScience 23:101414. https://doi.org/10.1016/j.isci.2020.101414

O'Hara E, Neves ALA, Song Y, Guan LL (2020) The role of the gut Microbiome in cattle production and health: driver or passenger? Annu Rev Anim Biosci 8:199–220. https://doi.org/10.1146/annurev-animal-021419-083952

Ojala T, Kankuri E, Kankainen M (2023) Understanding human health through metatranscriptomics. Trends Mol Med 29:376–389. https://doi.org/10.1016/j.molmed.2023.02.002

Olijhoek DW, Hellwing ALF, Noel SJ et al (2022) Feeding up to 91% concentrate to Holstein and Jersey dairy cows: effects on enteric methane emission, rumen fermentation and bacterial community, digestibility, production, and feeding behavior. J Dairy Sci 105:9523–9541. https://doi.org/10.3168/jds.2021-21676

Park T, Wijeratne S, Meulia T et al (2021) The macronuclear genome of anaerobic ciliate entodinium caudatum reveals its biological features adapted to the distinct rumen environment. Genomics 113:1416–1427. https://doi.org/10.1016/j.ygeno.2021.03.014

Parks DH, Chuvochina M, Rinke C et al (2022) GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. Nucleic Acids Res 50:D785–D794. https://doi.org/10.1093/nar/gkab776

Paz HA, Anderson CL, Muller MJ et al (2016) Rumen bacterial community composition in Holstein and Jersey cows is different under same dietary condition and is not affected by sampling method. Front Microbiol 7. https://doi.org/10.3389/fmicb.2016.01206

Peng C, May A, Abeel T (2023) Unveiling microbial biomarkers of ruminant methane emission through machine learning. Front Microbiol 14. https://doi.org/10.3389/fmicb.2023.1308363

Pérez-Enciso M, Zingaretti LM, Ramayo-Caldas Y, de los Campos G (2021) Opportunities and limits of combining Microbiome and genome data for complex trait prediction. Genet Selection Evol 53:65. https://doi.org/10.1186/s12711-021-00658-7

Pinares-Patiño CS, Hickey SM, Young EA et al (2013) Heritability estimates of methane emissions from sheep. Animal 7:316–321. https://doi.org/10.1017/S1751731113000864

Pitta DW, Indugu N, Melgar A et al (2022) The effect of 3-nitrooxypropanol, a potent methane inhibitor, on ruminal microbial gene expression profiles in dairy cows. Microbiome 10:146. https://doi.org/10.1186/s40168-022-01341-9

Puniya AK, Singh R, Kamra DN (2015) Rumen microbiology: from evolution to revolution. Springer New Delhi. https://doi.org/10.1007/978-81-322-2401-3

Quast C, Pruesse E, Yilmaz P et al (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596. https://doi.org/10.1093/nar/gks1219

Rabaglino MB, Salilew-Wondim D, Zolini A et al (2023) Machine-learning methods applied to integrated transcriptomic data from bovine blastocysts and elongating conceptuses to identify genes predictive of embryonic competence. FASEB J 37:e22809. https://doi.org/10.1096/fj.202201977R

Reel PS, Reel S, Pearson E et al (2021) Using machine learning approaches for multi-omics data analysis: A review. Biotechnol Adv 49:107739. https://doi.org/10.1016/j.biotechadv.2021.107739

Robinson DL, Goopy JP, Hegarty RS, Oddy VH (2015) Comparison of repeated measurements of methane production in sheep over 5 years and a range of measurement protocols1. J Anim Sci 93:4637–4650. https://doi.org/10.2527/jas.2015-9092

Rodino KG, Wolf MJ, Sheldon S et al (2021) Detection of Tick-Borne Bacteria from whole blood using 16S ribosomal RNA gene PCR followed by Next-Generation sequencing. J Clin Microbiol 59. https://doi.org/10.1128/jcm.03129-03120

Ross EM, Hayes BJ, Tucker D (2020) Genomic predictions for enteric methane production are improved by metabolome and Microbiome data in sheep (Ovis aries). J Anim Sci 98:skaa262. https://doi.org/10.1093/jas/skaa262

Rothschild D, Weissbrod O, Barkan E et al (2018) Environment dominates over host genetics in shaping human gut microbiota. Nature 555:210–215. https://doi.org/10.1038/nature25973

Rowe SJ, Hess M, Zetouni L et al (2019) Breeding low emitting ruminants: predicting methane from microbes. Proceedings 36(1):177. https://doi.org/10.3390/proceedings2019036177

Saborío-Montero A, López-García A, Gutiérrez-Rivas M et al (2021) A dimensional reduction approach to modulate the core ruminal Microbiome associated with methane emissions via selective breeding. J Dairy Sci 104:8135–8151. https://doi.org/10.3168/jds.2020-20005

Sahana G, Cai Z, Sanchez MP et al (2023) Invited review: good practices in genome-wide association studies to identify candidate sequence variants in dairy cattle. J Dairy Sci 106:5218–5241. https://doi.org/10.3168/jds.2022-22694

Schoch CL, Seifert KA, Huhndorf S et al (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc Natl Acad Sci 109:6241–6246. https://doi.org/10.1073/pnas.1117018109

Sereika M, Kirkegaard RH, Karst SM (2022) Oxford nanopore R10.4 long-read sequencing enables the generation of near-finished bacterial genomes from pure cultures and metagenomes without short-read or reference Polishing. Nat Methods 19:823–826. https://doi.org/10.1038/s41592-022-01539-7

Seshadri R, Leahy S, Attwood G et al (2018) Cultivation and sequencing of rumen Microbiome members from the Hungate1000 collection. Nat Biotechnol 36:359–367. https://doi.org/10.1038/nbt.4110

Shaffer M, Borton MA, McGivern BB et al (2020) DRAM for distilling microbial metabolism to automate the curation of Microbiome function. Nucleic Acids Res 48:8883–8900. https://doi.org/10.1093/nar/gkaa621

Shaw J, Yu YW (2023) Fast and robust metagenomic sequence comparison through sparse Chaining with Skani. Nat Methods 20:1661–1665. https://doi.org/10.1038/s41592-023-02018-3

Smith RH, Glendinning L, Walker AW, Watson M (2022) Investigating the impact of database choice on the accuracy of metagenomic read classification for the rumen Microbiome. Anim Microbiome 4:57. https://doi.org/10.1186/s42523-022-00207-7

Smoczek M, Vital M, Wedekind D et al (2020) A combination of genetics and microbiota influences the severity of the obesity phenotype in diet-induced obesity. Sci Rep 10:6118. https://doi.org/10.1038/s41598-020-63340-w

Smolka M, Paulin LF, Grochowski CM et al (2024) Detection of mosaic and population-level structural variants with Sniffles2. Nat Biotechnol 42:1571–1580. https://doi.org/10.1038/s41587-023-02024-y

Stamboulian M, Canderan J, Ye Y (2022) Metaproteomics as a tool for studying the protein landscape of human-gut bacterial species. PLoS Comput Biol 18:e1009397. https://doi.org/10.1371/journal.pcbi.1009397

Steijger T, Abril JF, Engström PG et al (2013) Assessment of transcript reconstruction methods for RNA-seq. Nat Methods 10:1177–1184. https://doi.org/10.1038/nmeth.2714

Stewart RD, Auffret MD, Warr A et al (2018) Assembly of 913 microbial genomes from metagenomic sequencing of the cow rumen. Nat Commun 9:870. https://doi.org/10.1038/s41467-018-03317-6

Stewart RD, Auffret MD, Warr A et al (2019) Compendium of 4,941 rumen metagenome-assembled genomes for rumen Microbiome biology and enzyme discovery. Nat Biotechnol 37:953–961. https://doi.org/10.1038/s41587-019-0202-3

Szopinska-Tokov J, Bloemendaal M, Boekhorst J et al (2023) A comparison of bioinformatics pipelines for compositional analysis of the human gut Microbiome. BioRxiv 202320022013528280. https://doi.org/10.1101/2023.02.13.528280

Terrón-Camero LC, Gordillo-González F, Salas-Espejo E, Andrés-León E (2022) Comparison of metagenomics and metatranscriptomics tools: A guide to making the right choice. Genes 13(12):2280. https://doi.org/10.3390/genes13122280

Thaiss CA, Zmora N, Levy M, Elinav E (2016) The Microbiome and innate immunity. Nature 535:65–74. https://doi.org/10.1038/nature18847

Uffelmann E, Huang QQ, Munung NS et al (2021) Genome-wide association studies. Nat Reviews Methods Primers 1:59. https://doi.org/10.1038/s43586-021-00056-9

Uruburu F (2003) History and services of culture collections. Int Microbiol 6:101–103. https://doi.org/10.1007/s10123-003-0115-2

Van Den Bossche T, Arntzen MØ, Becher D et al (2021) The metaproteomics initiative: a coordinated approach for propelling the functional characterization of microbiomes. Microbiome 9:243. https://doi.org/10.1186/s40168-021-01176-w

Vargas-Albores F, Garibay-Valdez E, Medina-Félix D, Martínez-Porchas M (2023) The micro-eukaryotic community: an underrated component of the mammalian gut microbiota? Front Microbiol 14. https://doi.org/10.3389/fmicb.2023.1123513

Větrovský T, Baldrian P (2013) The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. PLoS ONE 8:e57923. https://doi.org/10.1371/journal.pone.0057923

Wallace RJ, Rooke JA, Duthie CA et al (2014) Archaeal abundance in post-mortem ruminal digesta May help predict methane emissions from beef cattle. Sci Rep 4:5892. https://doi.org/10.1038/srep05892

Wallace RJ, Sasson G, Garnsworthy PC et al (2019) A heritable subset of the core rumen Microbiome dictates dairy cow productivity and emissions. Sci Adv 5:eaav8391. https://doi.org/10.1126/sciadv.aav8391

Wang Y, Shen Q, Zhong S et al (2020) Comparison of rumen microbiota and serum biochemical indices in white cashmere goats fed ensiled or sun-dried mulberry leaves. Microorganisms 8:981. https://doi.org/10.3390/microorganisms8070981

Wang M, Wang H, Zheng H et al (2021) A heat diffusion multilayer network approach for the identification of functional biomarkers in rumen methane emissions. Methods 192:57–66. https://doi.org/10.1016/j.ymeth.2020.09.014

Wang W, Zhang Y, Zhang X et al (2023) Heritability and recursive influence of host genetics on the rumen microbiota drive body weight variance in male Hu sheep lambs. Microbiome 11:197. https://doi.org/10.1186/s40168-023-01642-7

Wang L, Wang K, Hu L et al (2024a) Microbiological characteristics of the Gastrointestinal tracts of Jersey and Holstein cows. Animals 14(21):3137. https://doi.org/10.3390/ani14213137

Wang W, Dong Y, Guo W et al (2024b) Linkages between rumen microbiome, host, and environment in yaks, and their implications for Understanding animal production and management. Front Microbiol 15. https://doi.org/10.3389/fmicb.2024.1301258

Wassan TJ, Wang H, Zheng JH (2023) Recent advances in the phylogenetic analysis to study rumen Microbiome. Curr Bioinform 18:1–14. https://doi.org/10.2174/1574893618666230605120615

Wei W, Zhen Y, Wang Y et al (2022) Advances of rumen functional Bacteria and the application of Micro-Encapsulation fermentation technology in ruminants: A review. Fermentation 8(10):564. https://doi.org/10.3390/fermentation8100564

Wick RR, Judd LM, Gorrie CL, Holt KE (2017) Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595

Wick RR, Judd LM, Holt KE (2019) Performance of neural network basecalling tools for Oxford nanopore sequencing. Genome Biol 20:129. https://doi.org/10.1186/s13059-019-1727-y

Wilkinson T, Korir D, Ogugo M et al (2020) 1200 high-quality metagenome-assembled genomes from the rumen of African cattle and

their relevance in the context of sub-optimal feeding. Genome Biol 21:229. https://doi.org/10.1186/s13059-020-02144-7

Wood DE, Lu J, Langmead B (2019) Improved metagenomic analysis with kraken 2. Genome Biol 20:257. https://doi.org/10.1186/s13059-019-1891-0

Wu YW, Simmons BA, Singer SW (2016) MaxBin 2.0: an automated Binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 32:605–607. https://doi.org/10.1093/bioinformatics/btv638

Wu X, Huang S, Huang J et al (2021) Identification of the potential role of the rumen Microbiome in milk protein and fat synthesis in dairy cows using metagenomic sequencing. Animals 11(5):1247. https://doi.org/10.3390/ani11051247

Xie F, Jin W, Si H et al (2021) An integrated gene catalog and over 10,000 metagenome-assembled genomes from the Gastrointestinal Microbiome of ruminants. Microbiome 9:137. https://doi.org/10.1186/s40168-021-01078-x

Xie Y, Sun H, Xue M, Liu J (2022) Metagenomics reveals differences in microbial composition and metabolic functions in the rumen of dairy cows with different residual feed intake. Anim Microbiome 4:19. https://doi.org/10.1186/s42523-022-00170-3

Xu B, Song P, Jiang F et al (2024) Large-scale metagenomic assembly provide new insights into the genetic evolution of gut microbiomes in plateau ungulates. Npj Biofilms Microbiomes 10:120. https://doi.org/10.1038/s41522-024-00597-3

Xue MY, Sun HZ, Wu XH et al (2020) Multi-omics reveals that the rumen Microbiome and its metabolome together with the host metabolome contribute to individualized dairy cow performance. Microbiome 8:64. https://doi.org/10.1186/s40168-020-00819-8

Xue MY, Xie YY, Zhong Y et al (2022) Integrated meta-omics reveals new ruminal microbial features associated with feed efficiency in dairy cattle. Microbiome 10:32. https://doi.org/10.1186/s40168-022-01228-9

Yan M, Yu Z (2024) Viruses contribute to microbial diversification in the rumen ecosystem and are associated with certain animal production traits. Microbiome 12:82. https://doi.org/10.1186/s40168-024-01791-3

Yan M, Pratama AA, Somasundaram S et al (2023) Interrogating the viral dark matter of the rumen ecosystem with a global Virome database. Nat Commun 14:5254. https://doi.org/10.1038/s41467-023-41075-2

Yan W, Wang D, Wang Y et al (2024) Metatranscriptomics-guided genome-scale metabolic reconstruction reveals the carbon flux and trophic interaction in methanogenic communities. Microbiome 12:121. https://doi.org/10.1186/s40168-024-01830-z

Yan M, Andersen TO, Pope PB, Yu Z (2025) Probing the eukaryotic microbes of ruminants with a deep-learning classifier and comprehensive protein databases. Genome Res 35:368–378. https://doi.org/10.1101/gr.279825.124

Yu Z, Yan M, Somasundaram S (2024) Rumen protozoa and viruses: the predators within and their functions—A mini-review. JDS Commun 5:236–240. https://doi.org/10.3168/jdsc.2023-0433

Yu Z, Somasundaram S, Yan M (2025) International symposium on ruminant physiology: rumen protozoa and viruses—New insights into their diversity and potential roles through omics lenses; A review. J Dairy Sci 108:7530–7544. https://doi.org/10.3168/jds.2024-25780

Zaiko A, Greenfield P, Abbott C (2022) Towards reproducible metabarcoding data: lessons from an international cross-laboratory experiment. Mol Ecol Resour 22:519–538. https://doi.org/10.1111/1755-0998.13485

Zhaa Y, Zhang Y, Khas E et al (2022) Effects of Allium Mongolicum Regel ethanol extract on three flavor-related rumen branched-chain fatty acids, rumen fermentation and rumen bacteria in lambs. Front Microbiol 13. https://doi.org/10.3389/fmicb.2022.978057

Zhang Q, Difford G, Sahana G et al (2020a) Bayesian modeling reveals host genetics associated with rumen microbiota jointly influence methane emission in dairy cows. ISME J 14:2019–2033. https://doi.org/10.1038/s41396-020-0663-x

Zhang Y, Li F, Chen Y et al (2020b) Metatranscriptomic profiling reveals the effect of breed on active rumen eukaryotic composition in beef cattle with varied feed efficiency. Front Microbiol 11. https://doi.org/10.3389/fmicb.2020.00367

Zhang Y, Li F, Chen Y, Guan LL (2022a) The Effects of Breed and Residual Feed Intake Divergence on the Abundance and Active Population of Rumen Microbiota in Beef Cattle, Animals 12 (15), 1966. https://doi.org/10.3390/ani12151966

Zhang Z, Wei W, Yang S et al (2022b) Regulation of dietary protein solubility improves ruminal nitrogen metabolism in vitro: role of bacteria–protozoa interactions. Nutrients 14:2972. https://doi.org/10.3390/nu14142972

Zhang J, Zhuoga D, Zhang X et al (2022c) Different feeding strategies can affect growth performance and rumen functions in Gangba sheep as revealed by integrated transcriptome and Microbiome analyses. Front Microbiol 13. https://doi.org/10.1038/nmicrobiol.2016.49

Zhou Z, Tran PQ, Breister AM et al (2022) METABOLIC: high-throughput profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and community-scale functional networks. Microbiome 10:33. https://doi.org/10.1186/s40168-021-01213-8

Zhu Y, Bu D, Ma L (2022) Integration of multiplied omics, a step forward in systematic dairy research. Metabolites 12(3):225. https://doi.org/10.3390/metabo12030225