Recent Advances in the Phylogenetic Analysis to Study Rumen Microbiome

[Link to publication record in Ulster University Research Portal](#)

Document Version
Author Accepted version

REVIEW ARTICLE

Title: Recent Advances in the Phylogenetic Analysis to Study Rumen Microbiome

Jyotsna Talreja Wassan[1], Haiying Wang[2], Huiru (Jane) Zheng[*2]

1. *Department of Computer Science, Maitreyi College, University of Delhi, INDIA*
2. *School of Computing, Ulster University, Belfast, United Kingdom*

Recent rumen microbiome studies are progressive due to the advent of next-generation sequencing technologies, computational models, and gene referencing databases. Rumen metagenomics enables the linking of the genetic structure and composition of the rumen microbial community to the functional role it plays in the ecosystem.

- **Background** Systematic investigations of the rumen microbiome including its composition in cattle, have revealed the importance of microbiota in rumen functions. Various research studies identified different types of microbiome species and their relationships that reside within the rumen leading to a greater understanding of their functional contribution
- **Objective** The objective of this scoping review is to highlight the role of the phylogenetic and functional composition of the microbiome in cattle functions. It is driven by a natural assumption that closely related microbial genes/ operational taxonomical unit (OTUs)/Amplicon Sequence Variants (ASVs) by phylogeny are highly correlated and tend to have similar functional traits.
- **Methods** PRISMA approach has been used to conduct the current scoping review providing state-of-the-art studies for a comprehensive understanding of microbial genes' phylogeny in the rumen microbiome and their functional capacity.
- **Results** 44 studies were included in the review facilitating phylogenetic advancement in studying important cattle functions and identifying key microbiota. Microbial genes and their inter-relations have the potential to accurately predict the phenotypes linked to ruminants such as feed efficiency, milk production, and high/low methane emissions. In this review, a variety of cattle were considered ranging from cows, buffaloes, lambs, and Angus Bulls. etc. Also, results from the reviewed literature indicate that metabolic pathways in microbiome genomic groupings result in better carbon channelling thereby affecting the methane production by ruminants.
- **Conclusion** The mechanistic understanding of the phylogeny of the rumen microbiome could lead to a better understanding of ruminant functions. The composition of the rumen microbiome is crucial for the understanding of dynamics within the rumen environment. The integration of biological domain knowledge with functional gene activity, metabolic pathways, and rumen metabolites could lead to a better understanding of the rumen system.

1. INTRODUCTION

The rumen is a four-chambered stomach that is found in specialized herbivorous ruminants such as cows, sheep, goats, deer, and buffalo. It is important to study microbial dynamics present in the rumen. The most widely used analysis for rumen environments involves the 16S rRNA gene [1]. The prevalent curated databases for 16S analysis are Greengenes[2], RDP [3], SILVA [4], and RIM-DB[5]. Nonetheless, the taxonomy of such databases is derived from methods such as the Bayesian approach, curation of phylogenetic trees, and alignment of 16S sequences with genus or species as the lowermost level of analysis. The studies [6]–[9] relating to rumen metagenomics primarily investigate the rumen microbial community in the rumen through the relative abundance of microbial genes or their relationships on a phylogenetic tree identified in a metagenomics analysis. To date, a variety of microbial phylotypes have been identified in the rumen using metagenomic approaches[10]. The interactions and relationships among different microorganisms residing in the rumen can influence the cattle's energy-harvesting capability and feed efficiency[11]. Ruminants support efficient fermentation that converts plant materials to

human edible food such as high-protein milk, meat, and other dairy products. The rumen ecosystem is diverse, complex, and dynamic, with microorganisms interacting for the host's digestion and metabolism. The relationships and interactions between microorganisms can drive changes in microbial ecology [12]. Those changes can affect fermentation and subsequently animal health and performance [13]. The relationships and interactions between microorganisms can drive changes in rumen microbial ecology [12]. Those changes can affect fermentation and subsequently animal health and performance. Hence, it is important to study how relationships between microorganisms present in cattle are important in determining cattle health and production. Different cattle diets can also influence microbial composition leading to variations in microbial ecology [14]. The rumen contains a variety of prokaryotic and eukaryotic microorganisms utilizing plant material (like cellulose) and converting them to microbial protein to obtain energy by emitting greenhouse gases [15]. Excessive nitrogen in manure may also adversely affect ruminant products. Diet is one of the most influential functional parameters affecting the rumen microbiome and fermentation [2], [16]. It has also been observed that host animals may also influence the rumen microbial population through the effect of nutrition. Recent efforts have allowed phylogenetic information to be linked with the metabolic information in cattle rumen [17]. In this paper, we reviewed the recent literature of ten years (2012-2022) studying relationships between different microbiota and how these affect associations between the rumen microbiome and host phenotypes e.g., rumen feed efficiency, conversion of plant-based feed to useful nutrients, methane emissions, etc. This review will highlight the interplay of the host animal, its environment, and constituent microbial composition and phylogeny, allowing for deeper insights into rumen metagenomics. The review will focus on understanding the rumen microbiome in terms of: "Who is there?" "How they are related to microbial lineages "and "What do they do?". Rapid advances in computational tools and methods aid in studying the phylogenetic structure and function of the rumen microbiome. Systematic computational analysis of the rumen microbiome, virome, and plasmidome, has revealed unknown compositions and functions in the rumen.

2. BACKGROUND

Phylogeny covers evolutionary relationships from different taxonomical levels from domain to species. The abundance counts of microbiota and relationships between microbes by their evolutionary taxonomy (i.e., phylogeny) play an important role in analysing functional roles. Phylogeny is also important for the development of new microbial therapeutics. It is assumed that closely related genes share their evolutionary history and tend to show similar responses to biological functions [18]. Therefore, integrating biological domain knowledge of phylogeny is important in microbiome analysis. The association of microbial genes determines its metabolic behaviour. The researchers need to determine the pathways that coexist to

determine the behaviour of the system. Ranging from methane emission relating to archaeal lineages [19], specialist degradation of plant feed by Ruminococcus [20], to the dietary manipulation of the rumen microbiome [21]. Therefore, linking functionality to phylogeny allows the function to associate linked genes with host functions. Without knowing which microbial genes are responsible for cattle functional activities, it becomes impossible to target them. Phylogeny could play important role in (a) identification of rumen microbial genes linked to ruminant production systems, (b) understanding variations of rumen microbial communities in response to functions such as diets, disease, etc., (c) producing novel methods for mitigating the effects of methane production by ruminants, and (d) devising algorithms for integrated analyses of multi-comics datasets and co-occurrence networks [22]. The rumen-reticulum in cattle is a habitat for bacterial and protozoal species supporting digestion, and absorption of nutrients. Kim et al. [23] identified 13,478 bacterial and 3516 archaeal sequences of which the Firmicutes, Bacteroidetes, and Proteobacteria were found to be the most predominant in cattle rumen. Stewart et al.[24] studied the phylogeny of microbial rumens of 43 cattle raised in Scotland and assembled 913 rumen-uncultured genomes (RUGs) from them. Seshadri et al. [25] reported 410 reference archaeal and bacterial genomes from the Hungate collection [25]. Stewart et al. [26] further found more Hungate genomes adding large numbers of undiscovered microbes in the rumen. The study by Stewart et al. [26] indicated that the phylogenetic tree consists of large numbers of genomes from the Firmicutes and Bacteroidetes phyla (dominated by Clostridiales and Bacteroidales, respectively), but also contains many new genomes from the Actinobacteria, Fibrobacteres, and Proteobacteria phyla. In 2017, Parks et al. [27], reported the reconstruction of 7,903 bacterial and archaeal genomes. FibRumBa database (www.jcvi.org/rumenomics/) also proved useful for researchers interested in rumen microbiology and related genomics [28].

3. METHODS

In this paper, we intended to review studies highlighting the use of phylogeny and biological lineages of the rumen microbiome in the prediction of ruminant functions. However, one of the key challenges in these studies is the high degree of variability in the rumen microbiome responsible for cattle functions. This review is guided by the preferred conceptual boundaries, as highlighted by Peters et al.[29] in their publication. A Google Scholar, Institute of Electrical and Electronics Engineers (IEEE Xplore), and PubMed search was carried out to identify papers on the role of microbial phylogeny in cattle rumen functional metagenomics, published between January 2012 and May 2022. These papers were reviewed, evaluated, and classified by different kinds of cattle functions, and the findings were summarized. The summary of factors considered in this review is noted in Table 1.

Table 1. The details of Scoping Review

Title	Role of phylogeny in Determining Cattle Functions
Review Objective	To investigate the current state of methods using phylogeny and composition of microbial genes in detecting ruminant functions.
Question	How can phylogeny of rumen microbiota help in predicting ruminant functions to understand the fundamental mechanism of the ecosystem?
Population	Ruminants (cattle) population (e.g., cows, buffaloes, deer, etc.)
Concept	The current review studies the trend in the last ten years (2012–2022) of studying phylogeny (various genus and species) of microbiome harboured in the rumen of cattle to predict functions such as diet, productivity, and methane emissions.
Context	Use of genomic sequences with abundance counts and their relationships on a phylogenetic tree
Evidence	Bacterial, archaeal, and fungal microbial lineages exert the most significant role in shaping rumen microbial diversity and community composition affecting rumen functions.

We conducted a review of studies that used rumen microbial composition involving phylogeny in the classification of ruminant functions. The search strategy of the literature is outlined using the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA)[30] flow diagram, shown in Fig.(1), and is discussed below.

- **Identification**

To thoroughly review the literature, the following method was adopted. A keyword search was conducted to find relevant articles from Google Scholar, PubMed, and IEEE. The review intended to study ruminant populations, but not human populations to address the functional analysis of bacterial and archaeal populations. Human population formed the "exclusion criteria" for the current review. 2380 results were obtained by searching literature with the primary keywords of - "Phylogeny", "cattle rumen", "metagenomics", "microbiome". Thereafter, next-level keyword search was conducted by adding more specific keywords - "cattle diet", "methane emission", "disease", "co-occurrence", and "analysis", respectively to include studies involving the major role of the rumen microbial community in cattle functions. 1910 out of 2380 articles were retained in the process.

The publications before January 2012 were excluded. The duplicate articles obtained from the three platforms were removed, leaving behind 751 articles. The articles were then sorted by relevance via recommendations available at the Google Scholar and PubMed search platforms. The keyword search terms have a relevancy score based on how often they appear in text (either title or body) of research articles, and how many research articles contain the keyword term.
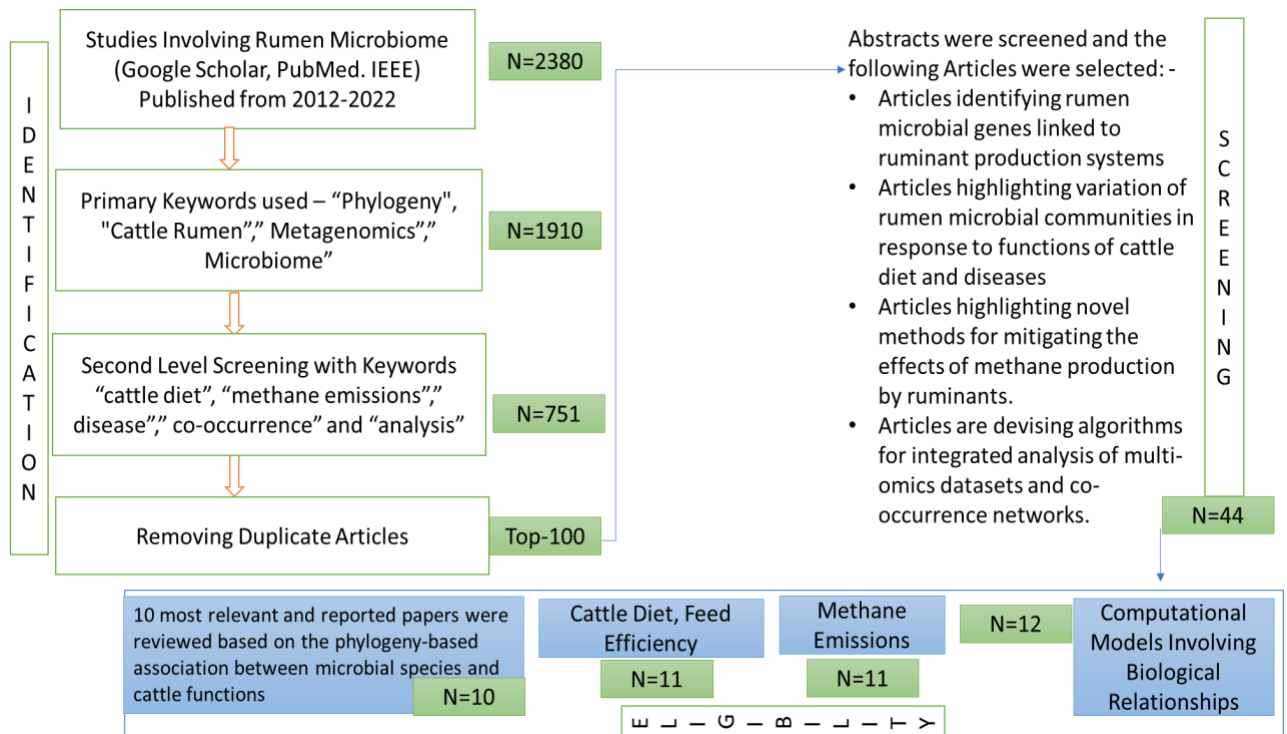
Fig. (1). PRISMA Approach for the current Review

The intersection between keyword search and the matching document on the web then results in an overall score, and the results are sorted by that score. This is known as "Sort by Relevance" of the articles [30]. At IEEE Xplore platform, results were further filtered by content type, publication year, abstract, keyword search, and citation score. Thereafter, in total of top-100 recommended articles were considered for the review from all three platforms.

- **Screening**

We intended to exclude survey or review papers, theses, book chapters, and patents. To the best of our knowledge, previous studies have not conducted a scoping review of phylogenetic applications in detecting metagenomic functional phenotypes`. This scoping review is focused on the applicability of phylogeny-based models in two primary functional phenotypes – "Diet" and "Methane emissions" associated with cattle functioning and performance. The remaining articles were screened based on the primary focus of the current review. Abstracts were screened to understand the application of phylogeny in predicting cattle functions. The articles intending to include informative biological lineages of the microbiome affecting primary cattle functions were included.

Post abstract screening, full articles were screened. Articles focusing on the following factors were included forming the inclusion criteria.

- Articles identifying rumen microbial genes linked to ruminant production systems.
- Articles highlighting variations of rumen microbial communities in response to functions such as diets, disease, etc.
- Articles highlighting novel methods for mitigating the effects of methane production by ruminants.
- Articles devising algorithms for integrated analyses of multi-omics datasets and co-occurrence networks.

Preference was given to highly cited articles. In total, 44 studies were included after screening as detailed below.

- Role of Phylogeny of rumen microbiome in Cattle Diet and Feed Efficiency – 11 papers [keywords: - "Phylogeny", "rumen", "microbiome", "diet" and/or "feed efficiency"].
- Phylogeny-bases analyses dictating Methane Emissions by Ruminants – 11 papers [keywords: - "Phylogeny", "rumen", "microbiome", "methane", "emissions"].
- Phylogeny-bases analyses using Varied Computational Models – 12 papers [keywords: - "Phylogeny", "rumen", "microbiome", "computational", "analysis", "and/or "cooccurrence networks" and/or "interactions" and/or "integrative analysis"]. Any studies dealing with only descriptive or raw data were excluded from this assessment.
- The 10 most relevant and reported papers were reviewed based on phylogeny-based associations between microbial species and cattle functions.

- **Eligibility**

All selected 44 papers led to a better understanding of how the cattle rumen microbial system responds to perturbations. Different microbial lineages from phylogeny can share functions; henceforth, combining phylogeny into the functional analysis of cattle microbiome enables better interpretation and understanding of the rumen ecosystem.

4. RESULTS

A total of 44 papers were included in this review. The review deals with phylogeny-based approaches that could be used to predict varied cattle performance with integrative analysis of abundances and relationships of microbial genes present in cattle rumen. The review highlights how host genetics and functions such as diet, productivity, and methane emissions are connected to rumen microbiome composition and relationships. The rumen-reticulum in cattle is a habitat for bacterial and protozoal species supporting digestion, and absorption of nutrients. A cattle carbohydrate-rich diet consists of predisposed decomposed to acidosis including feed grains; feedstuffs such as molasses and potatoes; by-products such as brewer's grains; and bakery products [31], [32]. Biochemically, ingestion of large amounts of the carbohydrate-rich diet causes an increase in lineages of *Streptococcus* and *Lactobacillus* microbial species [33]. These species intend to convert the starches to lactic acid leading to inflammation of rumen mucosal inflammation, which may further lead the animal to ulcers, liver abscesses, or laminitis from absorbed toxins. Methanogenesis in the rumen of cattle and other ruminants is a major source of atmospheric methane and involves Methanomicrobium and Methylotrophic Methanosarcina microbial species in the phenomena [34]. The production efficiency of cattle is also linked to the phylogenetic composition of the ruminal microbiome, as shown in popular studies linking it to residual feed intake [35],[36]. The study by Rojas et al. [37], surveyed the gut microbiotas of African buffalo and domestic cattle, along with other mammals. The study concluded that host phylogenetic relatedness and diet are strong drivers of microbiota structure present in rumen. Their findings suggest that hosts related by phylogeny may share similar niches for colonization, but these are influenced by the host diet, and other ecological conditions. Ryu et al.[38] further emphasized phylogenetic relationship, diet, physiology of host, and the environment -all of these affect the microbiome composition. Cattle, and poultry, inhabitant heritable microbes involving host genes and pathways actually help in shaping the microbiome.

- **Role of phylogeny of rumen microbiome in cattle diet and feed efficiency**

Sasson et al. [39] elucidated the relationship between the phylogenetic and functional composition of the rumen microbiome. The study [39] considered the population of 78 Holstein-Friesian dairy cows, using a combination of rumen microbiome abundance and phylogeny linking them to phenotypes. The studies [35], [39] indicated that the genetic composition and functionality are linked to the cow's capacity

to harvest energy from its feed, as well as to other physiological traits. Interestingly the studies [35], [39] also highlighted, that the variation in abundance profiles is also attributed to heritable genetic factors. The order Bacteroidales were found to be dominant in heritable OTUs. Considering physiological attributes, the majority of heritable OTUs were positively correlated with residual feed intake, dry matter intake, and milk protein. Understanding the natural evolutionary progress of the rumen microbiome and its function is necessary to evaluate changes in the microbial environment caused by diet. In a study by McCann et al. [40], the milk replacer diet showed great phylogenetic fluctuation with significant decreases observed for the Firmicutes and Proteobacteria phyla. At 14 days of age, Prevotella, Bacteroides, Oscillibacter, Paraprevotella, Butyricimonas, and Pelistega were found to be dominant; however, by 42 days of age, the main genera composition had shifted to Bacteroides, Prevotella, Porphyromonas, Butyricimonas, and Coprococcus. The study indicated that phylogenetic variation was present among animals consuming a 30% roughage and 70% concentrated diet. Prevotella was the most abundant genus observed in both diets. Rikenella was detected in greater abundance in the hay diet relative to wheat. Firmicutes phylum was found to be a core bacterial component of the rumen in all kinds of diets [40].Zehavi et al. [9] found that abundance and phylogeny are the main factors determining the cultivability of rumen microbes. selecting different clades that are scattered across the branches of the phylogenetic tree rather than being confined to a specific clade [9]. Das et al. [41] emphasized that the microbiome and their relationships are responsible for the bioconversion of nutrients in the given diet into a source of energy for the ruminants. The study identified three species of microbes i.e., Butyrivibrio fibrisolvens, Streptococcus species, and Clostridium aminophilus playing important role in diet conversion to feed efficiency. In an interesting recent study by Zhou et al. [42], the rumen microbiome in Angus bulls fed with either a backgrounding diet (BCK) or a finishing diet (HG) was compared. The dietary differences were observed at various taxonomic levels of the phylogeny. The phyla Firmicutes and Bacteroidetes, and the genus Selenomonas showcased opposite patterns in their response to different diets. Complex rumen enriched with microorganisms helps to degrade biomass during the animal's digestion process. Some new enzymes from naturally evolved biomass-degrading microbial communities are good candidates for organic matter degradability (OMD) in ruminants. The bacterial community in the rumen of four sheep using phylogenetic profiling, depicted a positive correlation between the Prevotellaceae (Bacteroidetes) and OMD [43]. McGovern et al [44] suggested that rumen microbial features are heritable from phylogeny and could be influenced by host genetics, further improving feed efficiency and optimizing rumen fermentation by targeting a variety of cattle and their rumen microbiota emphasizing the idea of Sasson et al. [39], that ruminant functions are correlated with microbiome structure and phylogenetically related taxa in ruminants. Heritable bacterial species attribute to heritable genetic factors and the study [45] indicated that phylogenetically related microbiome correlates with cow's capacity to harness energy from feed intake. La Reau et al. [46] resolved the phylogeny of Ruminococcus and

studied its strong association with the host. All Ruminococci require fermentable carbohydrates depending on the diet of their particular host Pitta et al. [47], studied rumen microbiome of the dairy cow for functional pathways characterized by lactation group and stages of lactation using a metagenomics approach. The study showcased that the phylogenetic distribution of ruminal bacterial populations at the phylum level is linked to different lactation periods. As an important observation, the authors noticed that when dairy cows transitioned from a non-lactation diet to a lactation diet, the abundance of Proteobacteria increased while Firmicutes tended to decrease. The study [36] further indicated that cattle diet can have a noticeable impact on the rumen microbiome phylogeny and abundance, which in turn influences host metabolism.

- **Phylogeny-based analyses dictating methane emissions by ruminants**

Methanogens are members of the domain Archaea, under kingdom Euryarchaeota. Interest in methanogens' phylogeny from ruminants has resulted in studying their role in methane emissions. Rumen microbiome has a strong association with methane emissions. The most common species of methanogens isolated from the rumen are strains of Methanobrevibacter, Methanomicrobium, Methanobacterium, and Methanosarcina [48]. Some of the interesting studies involving cattle and methane yields are discussed below. Rumen methanogens produce CH_4 by scavenging H2 and CO_2 produced by other diverse fermentative members of the microbiome present in cattle rumen [49]. Barbería et al. [50] estimated CH_4 emissions by ruminants using the phylogenetic linear model considering CH_4 as the response variable, Phylogeny, and Diet as additive random effects. The study interestingly found that most of the variance in differentiating functional phenotype of methane emissions rates is better explained by the random effects added to species phylogeny. The study [49] further highlighted those phylogenetic studies indicate a diverse and dynamic community of methanogens in cattle rumen. The study highlighted methane production by dairy cows is influenced by phylogenetic genetic composition and rumen microbial composition besides feed intake. The phylogenetic genetic effects contribute to inter-animal differences in CH_4 production [51]. The study [51] fitted the linear computational model to microbial abundance and additive phylogenetic genetic effects and these factors were found responsible for approximately 34% of the total phenotypic variation in CH_4 emissions. Henceforth, the study showcased that variation in methane emissions is attributed to additive genetics of the rumen microbiome in cows. The study by Archaea [52] showed increased heritability within the Thermoplasmatales highlighting the importance of collating phylogeny with the abundance of certain taxa across genetically related cows. In another interesting study by King et al. [53], Methanobrevibacter-related environmental clones were divided into categories based on their phylogenetic distribution and representation and were linked to methanogen populations of lactating Jersey and Holstein dairy cows under the same diet regimen. Potentially phylogeny was

found important means for estimating the methane synthesis potential of a methanogen community in different breeds. Shi et al. [55] measured methane yields of 22 sheep, by using deep metagenomic and metatranscriptomic sequencing techniques over the abundance of methanogenesis pathway microbial genes in high and low methane emitters. These genomic profiles were substantially increased in sheep with high methane yields. Henceforth, this study concluded that the differences in methane yields are due to the differential expression of genetically linked microbial genes on the phylogenetic tree. It would not have been possible to differentiate between the methane emissions just using abundance count data of microbial context. In the study by Shi et al.[55], authors identified rumen methanogens with methanogenesis pathway transcription profiles correlating with methane yields and providing new targets for CH_4 mitigation at different levels of microbiota phylogeny. Sasson et al. [39] identified 22 OTUs whose abundances were associated with rumen metabolic and physiological traits with measurable heritability. Interestingly, OTUs shared higher phylogenetic similarity between themselves suggesting that ruminant genetics and physiology are correlated with microbiome structure. In a very recent study by Min et al. [56], authors designed an experiment to quantify the effect of supplementation on diet and observed major phylogenic rumen microbiome changes in steers grazing winter wheat which further affected and reduced methane emissions and increased the animal performance. Maximum likelihood phylogenetic analyses of 16S rRNA sequences were conducted by Poulsen et al.[57] and found that Thermoplasmata is a novel group of methylotrophic methanogens in the bovine rumen as a source of carbon sources, contributing to methane emissions. The study by Lopes et al [43], performed a multivariate analysis of the phylogenetic profiling microbial data and found a negative correlation between Succinivibrionaceae (Proteobacteria phylum) and methane production. Lopes et al. described the bacterial composition and functions in the sheep rumen microbiome, highlighting some of the important carbohydrate-active enzymes (CAE) playing important role in degrading biomass during the animal's digestion process. Methane production by dairy cows is influenced by cow's individual genetic composition and rumen besides environmental factors was emphasized by Difford et al[51]. Each cow's additive genetic effects can also influence a variation in the abundance of a small percentage of rumen bacterial and archaeal taxa, and thereby contributing to variation in rumen microbiome composition and function. The study found out associations between methane emissions and rumen bacteria composition, which are known to produce methanogenesis substrates, suggesting bacteria driven methane production pathways.

- **Phylogeny-bases analyses using varied computational models**

Different phylogeny-driven workflows implemented in different studies (Table 2) are available to process the metagenomic information of the microbiome. Computational analysis of microbial functions requires the taxonomical assignment of the metagenome sampled from the environment. Although the approaches depending on only taxonomy would ignore the phylogenetic evolutionary distances which may otherwise prove useful in predictive modelling of microbiome. Variety of phylogeny-driven approaches [58] involve calculating sample to sample similarity based on the phylogenetic distances in a microbial sample; making use of evolutionary distances annotated on branches of the phylogenetic tree; or creating a feature space of environmental microbial communities using their biological assemblages. The current section gains insights from computational models using microbiome features and predictive learning using both the taxonomy and evolutionary distances as driven by microbial phylogeny (Table 2). A recent trend in the computational analysis is the use of deep learning (DL) solving problems end to end. DL models data as a nested hierarchy of concepts, with each concept defined in relation to other concepts [59]. DL usually dwells on Convolution Neural Networks (CNN)[60]. CNN works by learning filters that detect the local patterns. CNN has been used to predict functions from microbial abundance and phylogenetic profiles. An overview of the strategies for the phylogeny-based analysis of sequencing data of the ruminant microbiome is reported here (Table 2). Langille et al. [61] proposed a method of Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (i.e., PICRUSt), a predictive approach for determining the functional repertoire of 16S genes based on their taxonomic composition and the phylogenetic diversity of reference genomes derived from a phylogenetic tree. It uses an algorithm dwelling on ancestral-state reconstruction to identify closely related microbes with known full genome sequences to each OTUs/ASVs. The approach in [62], proposes a multi-class metagenomic classifier "MetaPhyl", regularizing the ML model of multinomial Logistic Regression [63] with a tree-based penalty function based on biological relatedness of microbial species on a phylogenetic tree. PhyloRelief proposed in [64], is driven by the Relief strategy [65] of selecting OTUs/ASVs features based on the phylogenetic weights annotated on tree branches. The approach progresses by associating phylogenetic weights with the clades (i.e., the branches connecting two OTUs/ASVs) and these weighted clades are ranked according to their contribution to the differentiation of the metagenomic sample.

Table 2. Phylogeny-Aware Modelling for Functional Metagenomic Analysis of Rumen based on the Computational models

Phylogeny-based Tools	Insights	Performance
PICRUSt [77],[78]	Using predicted metagenomes with the help of PICRUSt, metagenomic studies infer several functional pathways associated with cattle functions, weaning, diet induced that may contribute to effective ruminant metabolism, influencing energy balance.	PICRUSt highlighted significant differences in metabolic pathways for the inference of function of the rumen microbiome.
MetaPhyl [79]	The integrative approach incorporating phylogenetic tree structure into machine learning (ML) modelling achieved a high predictive performance for classifying cattle microbiomes into diet-types supplemented with oil, nitrate, a combination and control.	MetaPhyl indicated the comparable performance with other benchmarks at the lower phylogenetic levels of similarity cut-offs (65-90%) and outperformed the other methods at the higher resolution levels of phylogeny (> 90%)
Phylogenetic Isometric Log-Ratio (PhILR) [81]	The studies highlighted those methods for the identification of ruminant microbial genes that are significantly associated with a given phenotype should apply phylogeny-based transformation such as PhILR transform rather than analysis over the restrictive compositional nature of microbiome data.	It was observed that phylogenetically related OTUs were significantly more likely to covary than distantly correlated OTUs. PhILR transforms outperformed the classification of raw and log-transformed relative abundances.
PAAM-ML[68]	The approach regularizes abundances of OTUs/ASVs with branch lengths (distances) annotated on the phylogenetic tree as weights. Inclusion of branch lengths facilitate better modelling as fewer abundance taxa lineages might have undergone a significant evolutionary change (branch length is more) or vice versa. For this purpose, the designed data structure of PAAM in the process utilises both the tree topology as well as the evolutionary distances annotated on phylogenetic branches.	Results indicate significant classification performance, with high accuracy with ML over PAAM structure associated with rumen microbiome.
aMiSPU [79]	The tool combines individual microbial features into a group, and hence conduct group-based multivariate association analysis to investigate the statistical association between human microbiome compositions and a phenotype.	Combining taxon abundance information with phylogenetic distances provides an efficient computational approach to perform covariate analysis with the outcome of interest (phenotype) in a wide range of environmental applications. aMiSPU reported a statistically significant relationship between cattle microbiome and diet (p-value < 0.01)
Deep Learning (DL)[82]	The study employs a class activation map of the network and phylogenetic trees of cattle microbiome to perform classification based on different diets on which the cattle was fed.	DL achieved high classification performance (> 80 %)

Silverman et al. [66] proposed Phylogenetic Isometric Log-Ratio (PhILR), explained the inherently compositional nature of metagenomes yielding Isometric Log-Ratio (ILR) system to transform the compositional space of microbiome into a new coordinate system with an orthonormal basis, capturing the evolutionary relationships between metagenomes. The application of computational models using phylogenetic tree information node-by-node, intends to consider microbial taxa at different taxonomic levels. Wassan et al.[67] proposed a novel phylogeny and abundance aware machine learning modelling approach (PAAM-ML) for classifying microbial samples into their respective functional phenotypes integrating the abundance count of microbial species as well as relationships between them. The authors designed a feature space with phylogeny and abundance aware matrix structure (PAAM), inputted to machine learning (ML) models for the microbiome classification. The studies [15], [68], performed analysis by benchmarking various phylogeny-driven methods based on the integration of biological domain knowledge of relationships and non-phylogenetic methods based only on the raw abundances. The integrative approach incorporating phylogenetic tree structure into computational modelling provides better performance for classifying cattle microbiomes into functions. The results in [68], indicate that including microbial feature nodes at all levels of taxonomy, has the potential to increase the predictive performance of the sampled cattle microbiome. Wu et al. [69] devised a new multivariate regression test-driven statistical test method called the adaptive microbiome-based sum of the powered score (aMiSPU), for studying an overall association between the composition of a microbial community and a functional phenotype. This is useful for understanding -cattle microbiome composition and its functions. Phylogeny plays important role in differentiating functional phenotypes related to cattle [70]. Additionally, Co-abundance network analysis helps to study intra- and inter-domain interactions within the rumen microbiome and identify microbial groups related to function microbiome [71]. In a study by Tapio et al.[72], the authors investigated the advantages of including the phylogenetic depth of rumen microbial community analysis by evaluating microbial responses to functions. The study [72] introduced inter-domain co-occurrence network analysis to identify patterns of microbial interactions in rumen facilitating study of important dietary changes in cows, from forage to concentrate diets and the supplementation of diets with vegetable oil. Wallace et al. [54] identified phylogenetically linked core rumen microbiome forming co-occurrence networks linking structure with phenotype (methane emissions, rumen and blood metabolites, and milk production efficiency). In a study by Alvaro et al.[73], a network analysis of microbial genera (archaea, bacteria, fungi, and protists) and their genes was performed to study interactions within the ruminal microbiome affecting methane emissions. Co-abundance network analysis revealed the most abundant hydrogenotrophic Methanobacteriales with the key microbial genes involved in methanogenesis. Recent studies indicate that layers of phylogeny could be used to map them to CNN modelling. Nguyen et al. [59], proposed an approach named Met2Img, involving colour coding scheme for ranking OTUs in taxonomic orders from phylum to species. The presence and relationships of OTUs were marked by the

coloured pixels, and the absence was marked by the white pixel. This coding scheme was used to generate feature space for metagenomic classification. Reiman et al. [74], on the other hand, proposed a framework PopPhy-CNN based on the architecture of CNN to predict functions from microbial abundance profiles. The method used phylogenetic trees to indicate the spatial relationship of the microbes by embedding microbial abundance counts at each taxonomical level. The study in [75] , introduced Ph-CNN, a novel DL approach for the classification of microbiome data exploiting phylogenetic tree structure in the prediction phase. This study indicated promising results in model performance and biomarker detection. In the study by Zhu et al.[76], the deep forest model facilitated multiple processing layers to learn the representation of phylogenetic tree with multiple levels of abstraction, enhancing the performance of functional classification of metagenomic data which could be applied to rumen data. Furthermore, 10 studies (Table 3) targeting varied phenotypes and found most relevant in the Scoping Review, indicating optimization of cattle functions by making use of phylogenetic knowledge are highlighted in Table 3. In the studies indicated in Table 3, knowledge extracted from the quantitative and qualitative measures has been utilized to study cattle microbiome and their functions. Quantitative profiles are analyzed using OTUs/ASVs. Qualitative knowledge in terms of biological domain knowledge (obtained from phylogeny) has been utilized to model relationships between various microbiome features.

5. DISCUSSION

The aim of this review is to investigate the benefits of including phylogeny in analyses and increase the depth of the community analysis for describing and explaining ruminant responses to functional (phenotypic) changes. This review reports the latest assessment of functional analysis of microbiota using phylogeny and abundance of genes in the rumen ecosystem and summarizes the studies which allow for new insights into the structure and functions of these complex rumen residing microbial communities. Metagenomic data derived directly from an environmental sample, is useful for sampling a large number of genes and their phylogeny to understand the effect of microbial diversity. Phylogenetic placement of metagenomic data, highlights the importance of phylogeny identifying functions performed by an evolutionary line of species in cattle microbiome. Studies [73],[86],[92],[93],[94] of the cattle microbiota have previously linked exclusive methane production, feed efficiency and dietary effects to a bacterial and archaeal lineage. The linking of functionality to phylogeny has important implications and applications in correlating functions to host health. The first important step in studying rumen microbiome is to identify who are present and how they are related. However, it has to be realized that data about the community is important for studying the well-functioning rumen ecosystem. Progress has been made in analyzing the genetic potential within rumen ecosystems and studying their functions[40]. These developments have great promise for future research and are important in microbiome classification.

Table 3. Phylogeny-Aware Analysis of Cattle Functions

S.No	Phenotype	Ruminant	Microbiome from Its Phylogenetic perspective	Highlights	Year of publication, Reference
1	Diet	Angus bulls	Predominant bacterial phyla included Firmicutes higher in medium-grain backgrounding animals and Bacteroidetes was numerically higher in high grain diet-fed animals	The composition (abundance and phylogeny) of the rumen microbiome in all stages of beef production systems is important in host performance	2021, [83]
2	Methane Emissions	Cows. buffalos	Co-abundance network consisting of 549 microbial genes and 3349 connections, exhibits a clear modular structure with methane emissions trait-specific genes	The pattern of metagenomics abundance and phylogenetic relationships in the rumen microbial system has a strong association with methane emissions.	2017, [84]
3	Productivity and growth of ruminants	Tibetan lambs	Colonization of bacterial communities in lambs promotes functional development. The co-occurrence network showed a microbial variety of the rumen present in colonized formations and the role of the relationships and cattle productivity.	The study indicates that Bacteroidetes and Proteobacteria populations were significantly changed during the growing lambs.	2019, [85]
4	Methane emissions, rumen and blood metabolites, and milk production efficiency	Cows	Phylogenetically linked hierarchical structure associates host genetics and phenotype	Ruminococcus and Fibrobacteres are among the core heritable bacteria, consistent with their key role in cellulolytic, as is Succinovibrionaceae, which seems to be a key determinant between animal differences in methane emissions.	2019,[86]
5	Feed Efficiency	Cattle, sheep, moose, deer, and bison	The study performed phylogenetic characterization of the secondary metabolite biosynthetic gene clusters (BGCs) of rumen microbial genomes to build a network based on BGC similarity using BiGSCAPE[87]. Phylogenetic analyses of BGC indicated high inter-species diversity and thence the discovery of novel metabolites and probiotics to improve animal health and productivity.	The analysis identified 14,814 gene clusters from 8,160 rumen-specific genomes, indicating that rumen is a rich resource for secondary metabolites. The study focussed on expanding the phylogenetic diversity of known rumen bacteriocins and related peptides, identifying 4,326 putative bacteriocins, sactipeptide, lanthipeptide, and lasso peptide clusters for differentiating between high and low efficient steers.	2021, [87]
6	Diet	Holstein Cows	The rumen microbiome was dominated by Bacteroidetes (70%), Firmicutes (15–20%) and Proteobacteria (7%).	In this interesting study, the authors used metagenomic approaches to identify the rumen microbiome of the dairy cow for phylogeny and functional pathways by lactation group and stage of lactation. The study depicted that Bacteroidetes are primarily responsible for the first lactation of dairy cows. However, Firmicutes and Proteobacteria increase incrementally in second and third lactation dairy cows. Also, 70% of the CAZymes are oligosaccharide-breaking enzymes fermenting sugars in the diet. The rumen is a house of a complex and diverse group of microbes (related by phylogeny) living in a symbiotic relationship with the ruminant host.	2016, [47]

	Phenotype	Ruminant	Microbiome from Its Phylogenetic perspective	Highlights	Year of publication, Reference
7	Methane Emission	Cervidae (deer), Bovidae (bovid), and Moschidae (musk deer)	The study categorized the relationship between rumen methanogens and acetogenins and the host phylogeny with the help of phylosymbiosis analysis [89]. The rumen methanogen and acetogen communities do not show parallel host evolution, hence the strategies for mitigating methane production should be based on a species-specific rumen microbiome.	The study highlighted rumen methanogen and acetogen communities of 97 individual animals representing 14 ruminant species. The study found that Methanobrevibacter spp. and acetogenins associated with Eubacteriaceae are the most widespread methanogens and acetogens. The co-occurrence analysis showed that the variation of the predicted methane yields was characterized by the interactive patterns between methanogens, acetogens, and concentrations of rumen metabolites	2020,[90]
8	Methane emissions	Holstein cows	Methanomassilicoccales was found prevalent in the samples under study (mean relative abundance 35s indicating, Sporobacter potentially contributing to methane production. The study detected Sphaerochaeta be associated with estimated methane production.	The study highlighted that as phylogeny is differed between the bacteria and the archaea, it is important to consider differential relative abundance of taxa across genetically related cows. Archaea showed increased heritability within the Thermoplasmatales. Similarly Bacteria showed varirty too. This highlights the value of collating phylogeny with abundances into possible mechanisms detecting cattle functions across genetically related cows. The study used linear mixed model analysis to test for associations between bacterial and archaeal genome families with estimated methane emissions. Additive genetic effects drive the variation in the abundance of microbial taxa, and thereby contribute to variation in rumen microbiome composition and function.	2018, [92]
9	Feed Efficiency	Heifers and cows	Bacteroidetes and Firmicutes were the dominant phyla revealing significant differences with reference to feed efficiency groups.	The study highlighted that as phylogeny is differed between the bacteria and the archaea, it is important to consider differential relative abundance of taxa across genetically related cows. Archaea showed increased heritability within the *Thermoplasmatales.* Similarly Bacteria showed varirty too. This highlights the value of collating phylogeny with abundances into possible mechanisms detecting cattle functions across genetically related cows. The study used linear mixed model analysis to test for associations between bacterial and archaeal genome families with estimated methane emissions. dditive genetic effects drive the variation in the abundance of microbial taxa, and thereby contribute to variation in rumen microbiome composition and function.	2018, [91]
10	Feed efficiency	Beef Cattle	The study considered all levels of phylogeny in functional analysis with microbial metabolic functions.	The study found Ruminococcaceae (phylum Firmicutes) play important role in the effect of a barley-based high-grain diet with Succinivibrionaceae which belongs to the phylum Proteobacteria. Ruminants with low Residual feed intake(L-RFI) are considered to be feed efficient, whereas high-Residual Feed intake (H-RFI) individuals are considered to be inefficient. The active core microbiota consisted of six bacterial phyla, including Proteobacteria Firmicutes, Bacteroidetes, and Spirochaetes. Cyanobacteria, and Synergistetes. Since Lachnospiraceae are major butyrate producers, the activity of this family is associated with feed efficiency in beef cattle.	2017,[88]

Using scoping review methodology, we identified the evidence supporting the use of phylogenetic knowledge for cattle productivity with particular emphasis on the impact of diet and methane emissions as by-product of the cattle digestion process. An important finding of this review was the microbial genomic lineages from phylogeny associated with cattle functions; however, research supporting effect measures for analysis focusing on the combination of phylogeny with gene counts in any ruminant from which some analogies may be drawn is underway. The review supports that the presence of microbiome in cattle ruminants is linked to their dietary supplements, productivity, and methane emissions. Various studies in this review highlights important microbial species playing important role in cattle functions[47],[70],[86]. The analytical methods relying on low abundant taxa of cattle microbiome may not understand functions correctly if a biologically relevant change occurs during evolution in less abundant lineages. Henceforth, including a phylogenetic measure for the estimation of biological relationships between different microbial features in microbiome analysis is significant. Characterizing the relationship of microbiome composition and function is important in rumen ecosystem. Multiple microbial genes (OTUs/ASVs) from the same phylogenetic lineage could be present in a microbiome [87]. Also, phylogenetic diverse OTUs/ASVs may also perform the same action within a microbiome. Hence, relationships within and among lineages are important and are defined by phylogenetic trees from omics data sets.

Calculating phylogenetic diversity may play important role in estimating unique OTUs/ASVs in the defining microbial core. Phylogenetic analyses include study of microbial species diversity and closeness[95]. Most of rumen functional studies involving phylogeny assume microbial that closely related species respond similarly to phenotypes [18]. The concept of the core rumen microbiome could also include biological interactions comprised of interacting species. Networks analyses can generate hypotheses about interacting OTUs/ASVs while looking for local similarity analysis, created to understand the dynamics of rumen bacterial communities [96]. The appropriate method to analyze the core rumen microbiome depends on the ecological question in context. Including phylogeny of rumen microbiome ma y provide better understanding of cattle functions[15],[70],[83-87], [90-94]. Thus, applying phylogenetic definitions with abundance of microbial species of the cattle microbiome can better enhance ecological understanding.

The review highlights that domain knowledge of microbial phylogeny could be useful in discovering new microbial functions and could support downstream analysis in a biologically meaningful way. Prioritizing features in hierarchal feature space based on ancestral of OTUs/ASVs could further improve the predictive ability of phenotypic functions. Application of integration of phylogeny and abundances of microbial taxa has been successfully employed to engineer microbiome composition at different levels of taxonomical hierarchy for the prediction of metagenomic functions. The development of more effective approaches using DL is also likely as DL is emerging and the current approaches [74], [82], obscure the integration of phylogenetic

relatedness fully. The scope exists for developing new design principles based on phylogeny and DL, improving the performance results, and providing valuable insights into cattle functions. Regardless, additional research is needed to develop a whole computational framework considering both phylogeny and abundance count of cattle microbiome in functional analysis, demonstrating the efficacy of these factors to reduce methane emissions and improvising animal health with diet supplements and settings.

CONCLUSION

The cattle rumen contains a complex microbial ecosystem which play an important role in the digestion of plant materials by the ruminants. Microbial genes in cattle rumen are linked to animal health and performance. Understanding the phylogeny of the ruminant microbiome further aids in studying associations between microbial genes and functional pathways, influencing host phenotypes such as cattle diet, animal agriculture and processing biofuels. The current review highlights a study of comprehensive set of rumen microbial species along with a phylogenetic characterization of their taxonomic diversity and their functional significance. The review indicates that even though the bacterial taxa may vary considerably between cattle rumens, but they could be highly phylogenetically related. This further emphasize on the fact that the functional profile of cattle should select microbial taxa sharing similar genetic features. The review summarizes the association of host genetics with the phylogenetic and functional composition of the rumen microbiome. The review indicates that the Bacteroidales and Clostridiales order play important role in studying the functional composition of rumen microbiome.

Understanding the microbial context of cattle rumen will continue to get benefit from a range of computational methods using abundances as well as the phylogeny of the rumen microbiome. After reviewing the literature, we believe that maximizing the benefits of functional metagenome analysis depends on maintaining progress in microbial phylogeny and methods to count genes in cultural rumen microbiology. The approach incorporating phylogenetic tree structure into computational modelling provides better performance for classifying cattle microbiomes into functions. It is an interesting research area to judge how metagenome data could best be used to assist with the predictive modeling of rumen microbial metabolism. Our primary objective in this review was to provide a scope of rumen microbiome research over the last 12 years as it relates to cattle functions such as the effect of diet, methane emissions as a by-product, feed efficiency, and metabolism, fermentation. This review highlights how the ruminant functions may identify and exploit the phylogenetic relationships between the rumen microbiome and host traits of interest to cattle health and production. To our knowledge, a similar scoping review has not been published. The findings provide a baseline for including biological domain knowledge in the functional analysis of cattle microbiome, identifying a framework upon which data from more recent and future studies can be added. A global approach/framework that examines host/microbiome

relationships considering both phylogeny and genome counts is required to harness the full potential of the microbiome for sustainable ruminant production. Advancing sequencing technology and determining patterns and relationships in microbiome composition has redefined the study of rumen microbiome and created new opportunities to investigate complex microbial relationships to support functional analysis. Recent genome assemblies such as Hungate 1000 (www.hungate1000.org.nz/) and FibRumBa database (www.jcvi.org/rumenomics/) have paved the path to developing tools to process data and obtain the meaningful functional outcome. The host and rumen microbiome relationship could elucidate many driving factors affecting cattle health and productivity by considering biological domain knowledge.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1] M. Haynes, "Metagenomics," in *Brenner's Encyclopedia of Genetics: Second Edition*, 2013.

[2] R. J. Wallace *et al.*, "The rumen microbial metagenome associated with high methane production in cattle," *BMC Genomics*, 2015, doi: 10.1186/s12864-015-2032-0.

[3] J. J. Werner *et al.*, "Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys," *ISME J.*, vol. 6, no. 1, pp. 94–103, 2012, doi: 10.1038/ismej.2011.82.

[4] E. Pruesse, C. Quast, K. Knittel, … B. F.-N. acids, and undefined 2007, "SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB," *academic.oup.com*, Accessed: Aug. 17, 2022. [Online]. Available: https://academic.oup.com/nar/article-abstract/35/21/7188/2376260.

[5] H. Seedorf, S. Kittelmann, G. Henderson, P. J.- PeerJ, and undefined 2014, "RIM-DB: a taxonomic framework for community structure analysis of methanogenic archaea from the rumen and other intestinal environments," *peerj.com*, Accessed: Aug. 17, 2022. [Online]. Available: https://peerj.com/articles/494/.

[6] J. Gharechahi, S. Sarikhan, J.-L. Han, X.-Z. Ding, and G. H. Salekdeh, "Functional and phylogenetic analyses of camel rumen microbiota associated with different lignocellulosic substrates," *npj Biofilms Microbiomes 2022 81*, vol. 8, no. 1, pp. 1–14, Jun. 2022, doi: 10.1038/s41522-022-00309-9.

[7] E. Jami and I. Mizrahi, "Composition and Similarity of Bovine Rumen Microbiota across Individual Animals," *PLoS One*, vol. 7, no. 3, p. e33306, Mar. 2012, doi: 10.1371/JOURNAL.PONE.0033306.

[8] L. Bensoussan *et al.*, "Broad phylogeny and functionality of cellulosomal components in the bovine rumen microbiome," *Environ. Microbiol.*, vol. 19, no. 1, pp. 185–197, Jan. 2017, doi: 10.1111/1462-2920.13561.

[9] T. Zehavi, M. Probst, and I. Mizrahi, "Insights into culturomics of the rumen microbiome," *Front. Microbiol.*, vol. 9, no. AUG, p. 1999, Aug. 2018, doi: 10.3389/FMICB.2018.01999/BIBTEX.

[10] F. Li, A. L. A. Neves, B. Ghoshal, and L. L. Guan, "Symposium review: Mining metagenomic and metatranscriptomic data for clues about microbial metabolic functions in ruminants," *J. Dairy Sci.*, vol. 101, no. 6, pp. 5605–5618, Jun. 2018, doi: 10.3168/JDS.2017-13356.

[11] I. Mizrahi, "The Role of the Rumen Microbiota in Determining the Feed Efficiency of Dairy Cows," *Benef. Microorg. Multicell. Life Forms*, pp. 203–210, 2012, doi: 10.1007/978-3-642-21680-0_14.

[12] B. Koskella and M. A. Brockhurst, "Bacteria–phage coevolution as a driver of ecological and evolutionary processes in microbial communities," *FEMS Microbiol. Rev.*, vol. 38, no. 5, pp. 916–931, Sep. 2014, doi: 10.1111/1574-6976.12072.

[13] S. Namonyo, M. Wagacha, S. Maina, L. Wambua, and M. Agaba, "A metagenomic study of the rumen virome in domestic caprids," *Arch. Virol. 2018 16312*, vol. 163, no. 12, pp. 3415–3419, Sep. 2018, doi: 10.1007/S00705-018-4022-4.

[14] M. Kim, T. Park, and Z. Yu, "Metagenomic investigation of gastrointestinal microbiome in cattle," *Asian-Australasian J. Anim. Sci.*, vol. 30, no. 11, p. 1515, Nov. 2017, doi: 10.5713/AJAS.17.0544.

[15] P. Walsh *et al.*, "A metagenomics analysis of rumen microbiome," 2017, doi: 10.1109/BIBM.2017.8217980.

[16] D. Hadrich, "Microbiome research is becoming the key to better understanding health and nutrition," *Frontiers in Genetics*. 2018, doi: 10.3389/fgene.2018.00212.

[17] J. M. Brulc *et al.*, "Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 6, pp. 1948–1953, Feb. 2009, doi: 10.1073/PNAS.0806191105/SUPPL_FILE/0806191105SI.PDF.

[18] J. B. H. Martiny, S. E. Jones, J. T. Lennon, and A. C. Martiny, "Microbiomes in light of traits: A phylogenetic perspective," *Science (80-.).*, vol. 350, no. 6261, 2015, doi: 10.1126/science.aac9323.

[19] G. Henderson *et al.*, "Rumen microbial community composition varies with diet and host, but a core

microbiome is found across a wide geographical range," *Sci. Reports 2015 51*, vol. 5, no. 1, pp. 1–15, Oct. 2015, doi: 10.1038/srep14567.

[20] E. H. Crost, G. Le Gall, J. A. Laverde-Gomez, I. Mukhopadhya, H. J. Flint, and N. Juge, "Mechanistic insights into the cross-feeding of Ruminococcus gnavus and Ruminococcus bromii on host and dietary carbohydrates," *Front. Microbiol.*, vol. 9, no. NOV, p. 2558, Nov. 2018, doi: 10.3389/FMICB.2018.02558/BIBTEX.

[21] N. Malmuthuge and L. L. Guan, "Understanding host-microbial interactions in rumen: searching the best opportunity for microbiota manipulation," *J. Anim. Sci. Biotechnol. 2017 81*, vol. 8, no. 1, pp. 1–7, Jan. 2017, doi: 10.1186/S40104-016-0135-3.

[22] H. Zheng, H. Wang, R. Dewhurst, and R. Roehe, "Improving the Inference of Co-occurrence Networks in the Bovine Rumen Microbiome," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. PP, p. 1, 2018, doi: 10.1109/TCBB.2018.2879342.

[23] M. Kim, M. Morrison, and Z. Yu, "Status of the phylogenetic diversity census of ruminal microbiomes," *FEMS Microbiol. Ecol.*, vol. 76, no. 1, pp. 49–63, Apr. 2011, doi: 10.1111/J.1574-6941.2010.01029.X.

[24] R. D. Stewart *et al.*, "Assembly of 913 microbial genomes from metagenomic sequencing of the cow rumen," *Nat. Commun. 2018 91*, vol. 9, no. 1, pp. 1–11, Feb. 2018, doi: 10.1038/s41467-018-03317-6.

[25] R. Seshadri *et al.*, "Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection," *Nat. Biotechnol. 2018 364*, vol. 36, no. 4, pp. 359–367, Mar. 2018, doi: 10.1038/nbt.4110.

[26] R. D. Stewart, M. D. Auffret, A. Warr, A. W. Walker, R. Roehe, and M. Watson, "Compendium of 4,941 rumen metagenome-assembled genomes for rumen microbiome biology and enzyme discovery," *Nat. Biotechnol. 2019 378*, vol. 37, no. 8, pp. 953–961, Aug. 2019, doi: 10.1038/s41587-019-0202-3.

[27] D. H. Parks *et al.*, "Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life," *Nat. Microbiol. 2017 211*, vol. 2, no. 11, pp. 1533–1542, Sep. 2017, doi: 10.1038/s41564-017-0012-7.

[28] M. Morrison, S. C. Daugherty, W. C. Nelson, T. Davidsen, K. E. Nelson, and The North American Consortium for Fibrolytic Ruminal Bacteria, "The FibRumBa Database: A Resource for Biologists with Interests in Gastrointestinal Microbial Ecology, Plant Biomass Degradation, and Anaerobic Microbiology," *Microb. Ecol. 2009 592*, vol. 59, no. 2, pp. 212–213, Jul. 2009, doi: 10.1007/S00248-009-9562-4.

[29] Peters MD, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc. 2015 Sep;13(3):141-6. doi: 10.1097/XEB.0000000000000050. PMID: 26134548.

[30] M. J. Page and D. Moher, "Evaluations of the uptake and impact of the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) Statement and extensions: a scoping review," *Syst. Rev. 2018 61*, vol. 6, no. 1, pp. 1–14, Dec. 2017, doi: 10.1186/S13643-017-0663-8.

[31] C. N. Ncobela, A. T. Kanengoni, V. A. Hlatini, R. S. Thomas, and M. Chimonyo, "A review of the utility of potato by-products as a feed resource for smallholder pig production," *Anim. Feed Sci. Technol.*, vol. 227, pp. 107–117, May 2017, doi: 10.1016/J.ANIFEEDSCI.2017.02.008.

[32] C. M. Peter, D. B. Faulkner, N. R. Merchen, D. F. Parrett, T. G. Nash, and J. M. Dahlquist, "The effects of corn milling coproducts on growth performance and diet digestibility by beef cattle," *J. Anim. Sci.*, vol. 78, no. 1, pp. 1–6, Jan. 2000, doi: 10.2527/2000.7811.

[33] A. Sikora, M. Błaszczyk, M. Jurkowski, and U. Zielenkiewicz, "Lactic Acid Bacteria in Hydrogen-Producing Consortia: On Purpose or by Coincidence?," 2013, doi: 10.5772/50364.

[34] Y. Liu and W. B. Whitman, "Metabolic, Phylogenetic, and Ecological Diversity of the Methanogenic Archaea," *Ann. N. Y. Acad. Sci.*, vol. 1125, no. 1, pp. 171–189, Mar. 2008, doi: 10.1196/ANNALS.1419.019.

[35] S. Kruger Ben Shabat *et al.*, "Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants," *ISME J. 2016 1012*, vol. 10, no. 12, pp. 2958–2972, May 2016, doi: 10.1038/ismej.2016.62.

[36] P. R. Myer, T. P. L. Smith, J. E. Wells, L. A. Kuehn, and H. C. Freetly, "Rumen Microbiome from Steers Differing in Feed Efficiency," *PLoS One*, vol. 10, no. 6, p. e0129174, Jun. 2015, doi: 10.1371/JOURNAL.PONE.0129174.

[37] D. Mota-Rojas *et al.*, "Thermal biology in river buffalo in the humid tropics: neurophysiological and behavioral responses assessed by infrared thermography," *J. Anim. Behav. Biometeorol.*, vol. 9, no. 1, pp. 0–0, Oct. 2021, doi: 10.31893/JABB.21003/PDF/JABBNET-9-1-2103.PDF.

[38] E. P. Ryu and E. R. Davenport, "Host Genetic Determinants of the Microbiome Across Animals: From Caenorhabditis elegans to Cattle," *https://doi.org/10.1146/annurev-animal-020420-032054*, vol. 10, pp. 203–226, Feb. 2022, doi: 10.1146/ANNUREV-ANIMAL-020420-032054.

[39] G. Sasson *et al.*, "Heritable bovine rumen bacteria are phylogenetically related and correlated with the cow's capacity to harvest energy from its feed," *MBio*, vol. 8, no. 4, Jul. 2017, doi: 10.1128/MBIO.00703-17/SUPPL_FILE/MBO004173383ST2.TXT.

[40] J. C. Mccann, T. A. Wickersham, and J. J. Loor, "High-throughput Methods Redefine the Rumen Microbiome and Its Relationship with Nutrition and Metabolism," *Bioinform. Biol. Insights*, vol. 8, p.

BBI.S15389, Jan. 2014, doi: 10.4137/bbi.s15389.

[41]   K. C. Das and W. Qin, "Isolation and characterization of superior rumen bacteria of cattle (*Bos taurus*) and potential application in animal feedstuff," *Open J. Anim. Sci.*, vol. 2012, no. 04, pp. 224–228, Oct. 2012, doi: 10.4236/OJAS.2012.24031.

[42]   M. Zhou *et al.*, "Accessing Dietary Effects on the Rumen Microbiome: Different Sequencing Methods Tell Different Stories," *Vet. Sci. 2021, Vol. 8, Page 138*, vol. 8, no. 7, p. 138, Jul. 2021, doi: 10.3390/VETSCI8070138.

[43]   L. D. Lopes *et al.*, "Exploring the sheep rumen microbiome for carbohydrate-active enzymes," *Antonie van Leeuwenhoek 2015 1081*, vol. 108, no. 1, pp. 15–30, Apr. 2015, doi: 10.1007/S10482-015-0459-6.

[44]   E. McGovern, M. McGee, C. J. Byrne, D. A. Kenny, A. K. Kelly, and S. M. Waters, "Investigation into the effect of divergent feed efficiency phenotype on the bovine rumen microbiota across diet and breed," *Sci. Reports 2020 101*, vol. 10, no. 1, pp. 1–11, Sep. 2020, doi: 10.1038/s41598-020-71458-0.

[45]   F. Li *et al.*, "Host genetics influence the rumen microbiota and heritable rumen microbial features associate with feed efficiency in cattle," *Microbiome*, vol. 7, no. 1, pp. 1–17, Jun. 2019, doi: 10.1186/S40168-019-0699-1/FIGURES/7.

[46]   A. J. La Reau, J. P. Meier-Kolthoff, and G. Suen, "Sequence-based analysis of the genus Ruminococcus resolves its phylogeny and reveals strong host association," *Microb. Genomics*, vol. 2, no. 12, p. e000099, Dec. 2016, doi: 10.1099/MGEN.0.000099.

[47]   D. W. Pitta *et al.*, "Metagenomic assessment of the functional potential of the rumen microbiome in Holstein dairy cows," *Anaerobe*, vol. 38, pp. 50–60, Apr. 2016, doi: 10.1016/J.ANAEROBE.2015.12.003.

[48]   M. F. Whitford, R. M. Teather, and R. J. Forster, "Phylogenetic analysis of methanogens from the bovine rumen," *BMC Microbiol.*, vol. 1, no. 1, pp. 1–5, May 2001, doi: 10.1186/1471-2180-1-5/FIGURES/1.

[49]   A. Patra, T. Park, M. Kim, and Z. Yu, "Rumen methanogens and mitigation of methane emission by anti-methanogenic compounds and substances," *J. Anim. Sci. Biotechnol. 2017 81*, vol. 8, no. 1, pp. 1–18, Jan. 2017, doi: 10.1186/S40104-017-0145-9.

[50]   F. J. Pérez-Barbería, "Scaling methane emissions in ruminants and global estimates in wild populations," *Sci. Total Environ.*, vol. 579, pp. 1572–1580, Feb. 2017, doi: 10.1016/J.SCITOTENV.2016.11.175.

[51]   G. F. Difford *et al.*, "Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows," *PLOS Genet.*, vol. 14, no. 10, p. e1007580, Oct. 2018, doi: 10.1371/JOURNAL.PGEN.1007580.

[52]   S. Deusch, A. Camarinha-Silva, J. Conrad, U. Beifuss, M. Rodehutscord, and J. Seifert, "A structural and functional elucidation of the rumen microbiome influenced by various diets and microenvironments," *Front. Microbiol.*, vol. 8, no. AUG, p. 1605, Aug. 2017, doi: 10.3389/FMICB.2017.01605/BIBTEX.

[53]   E. E. King, R. P. Smith, B. St-Pierre, and A. D. G. Wright, "Differences in the rumen methanogen populations of lactating Jersey and holstein dairy cows under the same diet regimen," *Appl. Environ. Microbiol.*, vol. 77, no. 16, pp. 5682–5687, Aug. 2011, doi: 10.1128/AEM.05130-11/SUPPL_FILE/AEM-AEM05130-11-S01.DOC.

[54]   R. John Wallace *et al.*, "A heritable subset of the core rumen microbiome dictates dairy cow productivity and emissions," *Sci. Adv.*, vol. 5, no. 7, 2019, doi: 10.1126/SCIADV.AAV8391.

[55]   W. Shi *et al.*, "Methane yield phenotypes linked to differential gene expression in the sheep rumen microbiome," *Genome Res.*, vol. 24, no. 9, pp. 1517–1525, Sep. 2014, doi: 10.1101/GR.168245.113.

[56]   B. R. Min, W. E. Pinchak, M. E. Hume, and R. C. Anderson, "Effects of Condensed Tannins Supplementation on Animal Performance, Phylogenetic Microbial Changes, and In Vitro Methane Emissions in Steers Grazing Winter Wheat," *Anim. 2021, Vol. 11, Page 2391*, vol. 11, no. 8, p. 2391, Aug. 2021, doi: 10.3390/ANI11082391.

[57]   G. Henderson *et al.*, "Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range," *Sci. Rep.*, 2015, doi: 10.1038/srep14567.

[58]   M. Berlanga and M. Berlanga Herranz, "Molecular phylogeny of microorganisms," *Int. Microbiol. Off. J. Spanish Soc. Microbiol.*, 2010, doi: 10.2436/im.v13i4.53878.

[59]   T. H. Nguyen, Y. Chevaleyre, E. Prifti, N. Sokolovska, and J.-D. Zucker, "Deep Learning for Metagenomic Data: using 2D Embeddings and Convolutional Neural Networks," no. Nips, 2017, [Online]. Available: http://arxiv.org/abs/1712.00244.

[60]   M. Z. Alom *et al.*, "A State-of-the-Art Survey on Deep Learning Theory and Architectures," *Electron. 2019, Vol. 8, Page 292*, vol. 8, no. 3, p. 292, Mar. 2019, doi: 10.3390/ELECTRONICS8030292.

[61]   M. G. I. Langille *et al.*, "Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences," *Nat. Biotechnol.*, vol. 31, no. 9, pp. 814–821, 2013, doi: 10.1038/nbt.2676.

[62]   O. Tanaseichuk, J. Borneman, and T. Jiang, "Phylogeny-based classification of microbial communities," *Bioinformatics*, vol. 30, no. 4, pp. 449–456, 2014, doi: 10.1093/bioinformatics/btt700.

[63]   J. Friedman, T. Hastie, and R. Tibshirani, "Regularization Paths for Generalized Linear Models via Coordinate Descent," *J. Stat. Softw.*, 2010, doi: 10.18637/jss.v033.i01.

[64]   D. Albanese, C. De Filippo, D. Cavalieri, and C. Donati, "Explaining Diversity in Metagenomic

Datasets by Phylogenetic-Based Feature Weighting," *PLoS Comput. Biol.*, vol. 11, no. 3, pp. 1–18, 2015, doi: 10.1371/journal.pcbi.1004186.

[65]   R. J. Urbanowicz, M. Meeker, W. La Cava, R. S. Olson, and J. H. Moore, "Relief-based feature selection: Introduction and review," *Journal of Biomedical Informatics*. 2018, doi: 10.1016/j.jbi.2018.07.014.

[66]   J. Silverman, "Phylogenetic partitioning based ILR transform for metagenomics data." https://bioconductor.org/packages/release/bioc/html/philr.html.

[67]   J. T. Wassan, H. Wang, F. Browne, and H. Zheng, "PAAM-ML: A novel Phylogeny and Abundance aware Machine Learning Modelling Approach for Microbiome Classification," 2019, doi: 10.1109/BIBM.2018.8621382.

[68]   J. T. Wassan, H. Wang, F. Browne, and H. Zheng, "Phy-PMRFI: Phylogeny-Aware Prediction of Metagenomic Functions Using Random Forest Feature Importance," *IEEE Trans. Nanobioscience*, 2019, doi: 10.1109/TNB.2019.2912824.

[69]   C. Wu, J. Chen, J. Kim, and W. Pan, "An adaptive association test for microbiome data," *Genome Med.*, vol. 8, no. 1, pp. 1–12, 2016, doi: 10.1186/s13073-016-0302-3.

[70]   J. T. Wassan *et al.*, "An Integrative Framework for Functional Analysis of Cattle Rumen Microbiomes," 2019, doi: 10.1109/BIBM.2018.8621104.

[71]   H. Wang *et al.*, "Analysis of rumen microbial community in cattle through the integration of metagenomic and network-based approaches," *Proc. - 2016 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2016*, pp. 198–203, 2017, doi: 10.1109/BIBM.2016.7822518.

[72]   I. Tapio *et al.*, "Taxon abundance, diversity, co-occurrence and network analysis of the ruminal microbiota in response to dietary changes in dairy cows," *PLoS One*, vol. 12, no. 7, p. e0180260, Jul. 2017, doi: 10.1371/JOURNAL.PONE.0180260.

[73]   M. Martínez-Álvaro *et al.*, "Bovine host genome acts on rumen microbiome function linked to methane emissions," *Commun. Biol. 2022 51*, vol. 5, no. 1, pp. 1–16, Apr. 2022, doi: 10.1038/s42003-022-03293-0.

[74]   D. Reiman, A. A. Metwally, and Y. Dai, "PopPhy-CNN: A Phylogenetic Tree Embedded Architecture for Convolution Neural Networks for Metagenomic Data," *bioRxiv*, p. 257931, 2018, doi: 10.1101/257931.

[75]   D. Fioravanti *et al.*, "Phylogenetic convolutional neural networks in metagenomics," *BMC Bioinformatics*, vol. 19, no. 2, pp. 1–13, Mar. 2018, doi: 10.1186/S12859-018-2033-5/FIGURES/7.

[76]   Q. Zhu, Q. Zhu, M. Pan, X. Jiang, X. Hu, and T. He, "The Phylogenetic Tree based Deep Forest for Metagenomic Data Classification," *Proc. - 2018 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2018*, pp. 279–282, Jan. 2019, doi:

10.1109/BIBM.2018.8621463.

[77]   J. T. Wassan *et al.*, "Intelligent Computing Theories and Application," vol. 10362, pp. 421–427, 2017, doi: 10.1007/978-3-319-63312-1.

[78]   S. J. Meale *et al.*, "Weaning age influences the severity of gastrointestinal microbiome shifts in dairy calves," *Sci. Reports 2017 71*, vol. 7, no. 1, pp. 1–13, Mar. 2017, doi: 10.1038/s41598-017-00223-7.

[79]   J. T. Wassan *et al.*, "An Integrative Framework for Functional Analysis of Cattle Rumen Microbiomes," *Proc. - 2018 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2018*, pp. 1854–1860, Jan. 2019, doi: 10.1109/BIBM.2018.8621104.

[80]   C. Barelli *et al.*, "Habitat fragmentation is associated to gut microbiota diversity of an endangered primate: implications for conservation," *Sci. Reports 2015 51*, vol. 5, no. 1, pp. 1–12, Oct. 2015, doi: 10.1038/srep14862.

[81]   J. T. Wassan, H. Wang, F. Browne, and H. Zheng, "Microbial abundance analysis and phylogenetic adoption in functional metagenomics," *2017 IEEE Conf. Comput. Intell. Bioinforma. Comput. Biol. CIBCB 2017*, Oct. 2017, doi: 10.1109/CIBCB.2017.8058557.

[82]   T. Manning *et al.*, "Phylogeny-Aware Deep 1-Dimensional Convolutional Neural Network for the Classification of Metagenomes," *Proc. - 2018 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2018*, pp. 1826–1831, 2019, doi: 10.1109/BIBM.2018.8621543.

[83]   M. Zhou *et al.*, "Accessing Dietary Effects on the Rumen Microbiome: Different Sequencing Methods Tell Different Stories," *Vet. Sci. 2021, Vol. 8, Page 138*, vol. 8, no. 7, p. 138, Jul. 2021, doi: 10.3390/VETSCI8070138.

[84]   H. Wang *et al.*, "Integrated metagenomic analysis of the rumen microbiome of cattle reveals key biological mechanisms associated with methane traits," *Methods*, vol. 124, pp. 108–119, Jul. 2017, doi: 10.1016/J.YMETH.2017.05.029.

[85]   L. Wang *et al.*, "Dynamics and stabilization of the rumen microbiome in yearling Tibetan sheep," *Sci. Rep.*, vol. 9, no. 1, Dec. 2019, doi: 10.1038/S41598-019-56206-3.

[86]   R. John Wallace *et al.*, "A heritable subset of the core rumen microbiome dictates dairy cow productivity and emissions," *Sci. Adv.*, vol. 5, no. 7, 2019, doi: 10.1126/SCIADV.AAV8391.

[87]   C. L. Anderson and S. C. Fernando, "Insights into rumen microbial biosynthetic gene cluster diversity through genome-resolved metagenomics," *Commun. Biol.*, vol. 4, no. 1, Dec. 2021, doi: 10.1038/S42003-021-02331-7.

[88]   F. Li and L. L. Guan, "Metatranscriptomic Profiling Reveals Linkages between the Active Rumen Microbiome and Feed Efficiency in Beef Cattle," *Appl. Environ. Microbiol.*, vol. 83, no. 9, May 2017, doi: 10.1128/AEM.00061-17.

[89]   S. J. Lim and S. R. Bordenstein, "An introduction to

phylosymbiosis," *Proc. R. Soc. B*, vol. 287, no. 1922, Mar. 2020, doi: 10.1098/RSPB.2019.2900.

[90]   Z. Li *et al.*, "Comparative Microbiome Analysis Reveals the Ecological Relationships Between Rumen Methanogens, Acetogens, and Their Hosts," *Front. Microbiol.*, vol. 11, Jun. 2020, doi: 10.3389/FMICB.2020.01311.

[91]   P. R. Myer, T. P. L. Smith, J. E. Wells, L. A. Kuehn, and H. C. Freetly, "Rumen Microbiome from Steers Differing in Feed Efficiency," *PLoS One*, vol. 10, no. 6, p. e0129174, Jun. 2015, doi: 10.1371/JOURNAL.PONE.0129174.

[92]   G. F. Difford *et al.*, "Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows," *PLOS Genet.*, vol. 14, no. 10, p. e1007580, Oct. 2018, doi: 10.1371/JOURNAL.PGEN.1007580.

[93]   C. Matthews, F. Crispie, E. Lewis, M. Reid, P. W. O'Toole, and P. D. Cotter, "The rumen microbiome: a crucial consideration when optimising milk and meat production and nitrogen utilisation efficiency," *Gut Microbes*. 2019, doi: 10.1080/19490976.2018.1505176.

[94]   E. Hernandez-Sanabria, L. A. Goonewardene, Z. Wang, O. N. Durunna, S. S. Moore, and L. L. Guan, "Impact of feed efficiency and diet on adaptive variations in the bacterial community in the rumen fluid of cattle," *Appl. Environ. Microbiol.*, 2012, doi: 10.1128/AEM.05114-11.

[95]   D. R. Maddison, K. S. Schulz, and W. P. Maddison, "The tree of life web project," *Zootaxa*. 2007.

[96]   Z. Rasheed, "Data mining framework for metagenome analysis," p. 143, 2013.