RVC OPEN ACCESS REPOSITORY – COPYRIGHT NOTICE

TITLE: Microbiota composition, gene pool and its expression in Gir cattle (Bos indicus) rumen under different forage diets using metagenomic and metatranscriptomic approaches

AUTHORS: Pandit, R J; Hinsu, A T; Patel, S H; Jakhesara,S J; Koringa, P G; Bruno, F; Psifidi, A; Shah, S V; Joshi, C G

JOURNAL: Systematic and Applied Microbiology

PUBLISHER: Elsevier

PUBLICATION DATE: 9 March 2018 (online)

DOI: https://doi.org/10.1016/j.syapm.2018.02.002

Microbiota composition, gene pool and its expression in Gir cattle (*Bos indicus*) rumen under different forage diet using metagenomic and metatranscriptomic approach

Ramesh J. Pandit[1], Ankit T. Hinsu[1], Shriram H. Patel[1], Subhash J. Jakhesara[1], Prakash G. Koringa[1], Fosso Bruno[2], Androniki Psifidi[3,4], S. V. Shah[5], Chaitanya G. Joshi*[1]

[1] Department of Animal Biotechnology, College of Veterinary Science and Animal Husbandry, Anand Agricultural University, Anand- 388 001, Gujarat, India.

[2] Institute of Biomembranes, Bioenergetics and Molecular Biotechnologies (IBIOM), National Research Council, Via Amendola 165/A, 70126, Bari, Italy.

[3] The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, UK.

[4] Department of Clinical Science and Services, Royal Veterinary College, North Mymms, Hertfordshire, UK.

[5] Livestock Research Station, College of Veterinary Science & Animal Husbandry, Anand Agricultural University, Anand-388 001, Gujarat, India.

Corresponding author at: Department of Animal Biotechnology, College of Veterinary Science and Animal Husbandry, Anand Agricultural University, Anand- 388 001, Gujarat, India, Tel.: +919227531075.

E-mail addresses: cgjoshi@rediffmail.com; cgjoshi@aau.in

Abstract

Zebu (*Bos indicus*) is a domestic cattle species originated in the Indian subcontinent and now widely domesticated in several continents. In this study, we are particularly interested in understanding the functionally active rumen microbiota of an important Zebu breed, the Gir, under different dietary regimes. We compared metagenomic and metatranscriptomic data at various taxonomic levels to elucidate the differential microbial population and its' functional dynamics in Gir cattle rumen during different roughage dietary regimes. Different proportions of roughage rather than the type of roughage (dry or green) modulated microbiome composition and the expression of their gene pool. Fibre degrading bacteria (i.e. *Clostridium*, *Ruminococcus*, *Eubacterium*, *Butyrivibrio*, *Bacillus* and *Roseburia*) were higher in the solid fraction of rumen ($P<0.01$) as compared to the liquid fraction, whereas bacteria that are considered to be utilizers of the degraded product (i.e. *Prevotella*, *Bacteroides*, *Parabacteroides*, *Paludibacter* and *Victivallis*) were dominant in the liquid fraction ($P<0.05$). Likewise, expression of fibre degrading enzymes and related Carbohydrate Binding Modules (CBMs) expressed in the solid fraction. When compared metagenomic and metatranscriptomic data, we found that some genera and species were transcriptionally more active although they were in low abundance, making an important contribution in fibre degradation and its further metabolism in rumen. This study also identified some of the transcriptionally active genera like *Caldicellulosiruptor* and *Paludibacter*, whose potential is less-explored in rumen. Overall, comparison of metagenomic shotgun and metatranscriptomic sequencing appeared to be a much richer source of information compared to conventional metagenomics analysis.

Key wards: Gir cattle, Fibre degrading enzymes, Metagenome, Metatranscriptome, Microbiome, Rumen.

Introduction

Zebu (*Bos indicus*) is one of the important domesticated cattle species worldwide especially in tropical countries. *Bos indicus* compared to *Bos taurus* is characterized by its heat tolerance [12] as well as resistance and resilience to parasites [40, 73]. India harbours rich genetic diversity of zebu cattle with currently 40 registered breeds at National Bureau of Animal Genetic Resources (NBAGR, India). Among these, Gir is one of the best milk producing indigenous cattle breed [9]. As a result of such profitable qualities, initially Gir cattle were used for improvement of breeds in Brazil [49] and then globally, mainly in African and Southeast Asian countries as well as the United States [61]. Despite such valued characteristics, the scientific studies targeting the microbial diversity and their function in Gir cattle rumen are missing from the literature.

Ruminants fulfil their nutritional requirements through grazing. However, the enzymes required for breakdown of plant constituents are absent in these animals and therefore they rely on the microbial symbionts residing in their rumen, an anaerobic fermentation sack, where breakdown of complex plant polymers occurs by enzymatic process of various microbes [23]. Rumen harbours a unique consortium of microbes which has been evolved into a complex and efficient system of lignocellulose degradation. This panel of enzymes are collectively known as Carbohydrate-Active enZYmes (CAZymes) and have been studied in cattle and buffalo rumen [20, 21, 68, 74]. However, it is crucial to understand the expression of such enzymes under different dietary regimes, especially in animals who thrive on high roughage diet. Among the microbial symbionts bacteria contribute much more in the ruminal fermentation as compared to fungi and protozoa. The rumen microbial diversity has been extensively studied using amplicon [19, 35, 41, 53, 57] and shotgun sequencing approaches [6, 52, 58, 72]. Considering the microbial composition, diet exerts significant impact on rumen microbial population [13, 14]. Moreover, the bacterial fermentation products especially Volatile Fatty Acids (VFAs) in the rumen have a direct correlation with milk production [16, 63].

Due to certain limitations of amplicon and shotgun sequencing, it is difficult to figure out which bacteria are actively engaged in fibre degradation and fermentation and what genes are transcribed. Metatranscriptome sequencing allows understanding of functional dynamics of the microbial communities. To this date, only a limited number of studies focusing on the understanding of actively transcribed genes in rumen microbial population have been carried out [5, 10, 55, 65]. Energy harvesting capacity of the ruminants depends on specific microbial symbionts [64] and it is also possible to modulate this population for increasing feed efficiency [38] and milk production [18, 75].

We have previously characterized rumen microbiome of Kankrej cattle, Jaffrabadi and Mehsani buffalo fed with different roughage concentrations in the diet using amplicon and shotgun sequencing and demonstrated that bacterial community in the liquid and solid fractions of rumen are diverse and different proportions of roughage in the diet exerted significant impact on rumen bacterial population [42, 45, 48, 53]. In the present study we further hypothesised that the importance of the role of bacterial species in the rumen is not only determine by their abundance but how functionally active they are. The main objectives of this study were to (a) understand the active bacterial population and their dynamics by comparing metagenome and metatranscriptome in rumen adopted to different roughage proportions in their diet (b) study the differentially expressed Carbohydrate Active enZYmes (CAZYmes) during different feed treatments; and (c) identify feed associated biomarkers of rumen microbiome.

Materials and Methods

Ethics Statement

All experimental procedures were approved by Institutional Animal Ethical Committee of College of Veterinary Science & Animal Husbandry, Anand Agricultural University, Anand vide letter no. AAU/GVC/CPCSEA-IAEC/108/2013 dated 05/10/2013, Anand, Gujarat, India.

Dietary treatments and sample collection

The experimental design and nutritive value of feed is similar to one described in our previous studies [42, 48, 53]. In brief, eight healthy, female, non-pregnant and non-lactating 3-4 years old, with an average weight of 250-300 Kg Gir cattle were reared at the Livestock Research Station, Anand Agricultural University, Anand, Gujarat, India (Latitude: 22.527413, Longitude: 72.97065). Before commencement of the experiment, animals were fed with diet as per National Royal Commission (NRC) [59] standards, India. The animals were divided into two groups with four animals each with one group nourished on green fodder as roughage and another group received dry fodder as roughage along with commercially available concentrate mixture. The composition of concentrate mixture used was the same as described previously by Parmar et al. [45] whereas, *Sorghum bicolor* was used as fodder. Briefly during the experiment, animals were passed through three dietary treatments which comprised of different proportions of roughage and concentrate mixture and access to fresh water. During the first treatment (G1), animals were fed with 50% roughage of respective fodder (Green or Dry) and 50% concentrate mixture; during the second treatment (G2), with 75% roughage and 25% concentrate mixture; and during the third treatment (G3), with 100% roughage as diet. This dietary schedule was followed twice in a day, in the morning and afternoon. Each dietary treatment continued for six weeks and then switched over to next. Before switching to next treatment, 500-700mL of rumen digesta was collected 3 hours post morning feeding using flexible stomach tube attached to vacuum pump. The rumen liquor was fractionated into solid (fibre) and liquid by filtering through two layered muslin cloth [30]. Each fraction was separately collected into sterile 2mL cryovials containing 1mL RNA protect Bacterial reagent (QIAGEN, USA), shipped to laboratory at -20°C and stored at -80°C till further processing. A total of 48 samples were collected and further processed for sequencing. A schematic representation of the experimental design is shown in supplementary fig. S1.

DNA and RNA isolation

Total DNA from 600µL liquid or ~0.5g solid samples was extracted using QIAamp DNA Stool Mini Kit (Qiagen, Germany) according to the manufacturer's instructions. Few modifications for solid fraction were implemented as described in our previous study [45]. DNA quantity and quality was examined using Qubit 2.0 Fluorometer (ThermoFisher Scientific, USA) and gel electrophoresis, respectively. For metatranscriptome sequencing, RNA was extracted using RNeasy Mini Kit (Qiagen, Germany) according to manufacturer's instructions with a few modifications in solid fraction samples. These modifications has been described in our previous study [15]. Quantity and quality of RNA was estimated using Qubit 2.0 fluorometer and RNA 6000 Nano kit on Agilent 2100 Bioanalyzer (Agilent Technologies, CA), respectively. Samples with RNA Integrity Number (RIN) > 6 were further processed for rRNA depletion using RiboMinus™ (Bacterial module) (ThermoFisher Scientific, USA) following the manufacturer's instructions.

Metagenomic shotgun and metatranscriptome library preparation and sequencing

A total of forty eight samples were subjected to shotgun (referred as MG here after) and metatranscriptome (referred as MT here after) barcoded libraries preparation using standard 400bp Ion Plus Fragment Library Kit and Ion Total RNA-Seq Kit v2 (ThermoFisher Scientific, USA) respectively, following the manufacturer's instructions. Approximately 500 ng of DNA and 250 ng of rRNA depleted mRNA was used for respective library preparation. Quantity and quality of libraries were determined using Qubit 2.0 Fluorometer (ThermoFisher Scientific, USA) and Agilent 2100 Bioanalyzer (DNA high sensitivity assay kit, Agilent Technology, CA), respectively. Total 24 sequencing runs (12 metatranscriptome and 12 shotgun) were carried out on Ion Torrent PGM (ThermoFisher Scientific, USA) using Ion 316 chip, wherein each run contained 4 barcoded libraries of biological replicate for each treatment.

Taxonomic and functional classification of metagenomic and metatranscriptomic data

To explore the taxonomic composition and functional potential of rumen metagenome and metatranscriptome, all the sequences were uploaded to the MG-RAST (Metagenomic Rapid Annotations using Subsystems Technology) server v3.4 [37]. Contaminating host-specific reads were filtered by mapping the reads against *Bos taurus* genome UMD v3.0 using bowtie [29] available within MG-RAST pipeline. Moreover, reads were also dynamically trimmed [4] with minimum Phred score of 15. Taxonomic classification was performed against RefSeq database with minimum E-value 1E-5 and minimum identity of 70%. Functional classification of bacterial hits were performed using KO database with above mentioned parameters. All analysis was carried out in MG-RAST (v4.0).

Expression of CAZy family genes

We used SortMeRNA [25] to remove rRNA reads from all the 48 metaRNA sequencing files using all the rRNA database available in this tool. The rRNA cleaned reads were pooled and *de novo* assembled using CLC Genomics workbench v7.0 (Qiagen, Germany) with the parameters: minimum 50% length overlap (among of the extending contig with minimum identity 80%) and minimum contig length 200bp. Protein sequences of carbohydrate active enzymes were downloaded from the dbCAN database [77], latest update on 24-July-2016 and clustered at 80% sequence similarity using Cd-hit [31]. BLASTx search (minimum E-value 1E-6) of all the *de novo* assembled contigs was performed against the clustered dbCAN database. For expression analysis, an *in-house* reference (.gbk) file was prepared using the contigs found to have hit during BLASTx search. Equal number of representative set of sequences were used for each comparison and reads were mapped using CLC Genomics workbench and subsequently expressed in terms of Reads Per Kilo base per Million mapped reads (RPKM). The RPKM values were log2 transformed and their heatmap was plotted.

Statistical analyses

Statistical Analysis of Metagenomic Profiles (STAMP) v2.1.3 software [44] was used for multi group comparison using ANOVA with Benjamini-Hochberg FDR corrections [2] and Games-Howell Post-hoc test ($P < 0.05$). For two group comparison, two sided Welch’s test [76] (confidence interval 95%) with Benjamini-Hochberg FDR ($P < 0.05$) was applied. Paleontological Statistics (PAST) v 2.17c software [11] was used to perform Principal

Coordinates Analysis (PCoA, Bray-Curtis dissimilarity), calculating diversity indices (1000 Bootstrap, 95% confidence) and performing diversity profiles analysis. PERMANOVA was used to elucidate the differences in microbial community among the two different fractions as well as among the different treatments of the fractions. Linear Discriminant Analysis (LDA) effect size (LEFSe) (http://huttenhower.sph.harvard.edu/lefse/) was used to identify differentially abundant taxa (biomarkers) in each sample group [62] of MG and MT with P-value < 0.05 and LDA score > 3.0. Furthermore, we used Venn diagram plotter available at http://bioinformatics.psb.ugent.be/webtools/Venn/ to identify differentially abundant genera (abundance >0.1%) among the MG and MT dataset during three sequential dietary regimes.

Results

Metagenome and metatranscriptome data

We analysed a total of 35.39 million reads (7.51 GB) of metatranscriptome and 35.69 million reads (7.97 GB) of metagenome data with an average of 737,367 and 743,556 sequences per sample for taxonomic and functional classification using MG-RAST (Supplementary table T1). Rarefaction curve for each sample of MG and MT data is depicted in Supplementary fig. 2 A & D. In the case of metatranscriptome, sequences generated are sufficient to define the rumen microbial community, while it is bit inadequate in the case of metagenome data. Distinct clusters for each treatment of MT and MG data (Supplementary fig. Supplementary fig. 2 B, C, E & F) indicated that rumen bacterial community shifts according to the change in the diet composition.

Taxonomic and functional classification of rumen microbial community using metagenome shotgun sequencing

PCoA based on taxonomic (species level) and function (KO) profile revealed distinct clustering of samples from solid and liquid fraction of the rumen content, which was expected (Supplementary fig. S3 A & B). Moreover, statistical analysis using PERMANOVA showed significant difference (P= 0.001) among the two groups, indicating diverse bacterial communities in the liquid and solid content. When we analysed only the liquid fractions of each dietary treatment, PCoA showed overlapping clusters (Supplementary fig. S3 C & D) with no significant difference between the three groups (P = 0.469; PERMANOVA). However, the structure of microbiota across the three different groups of the solid fraction was significantly different (Supplementary fig. S3 E & F, P = 0.001; PERMANOVA). This indicates that as proportion of roughage increased in the diet, microbial community was altered especially in the solid fraction of the rumen.

Phylum level analysis of Gir cattle rumen microbiome is shown in Supplementary fig. S4 and Supplementary table T2A. Among the different phyla, Bacteroidetes, Fibrobacteres, Lentisphaerae and Verrucomicrobia were most abundant in the liquid fraction whereas, Firmicutes and Actinobacteria were higher in the solid fractions. Moreover, the proportion of these phyla were significantly different (corrected P value < 0.05) across the liquid and solid fractions of the three dietary treatments. In the liquid fractions of each treatment, the proportion of Proteobacteria and Spirochaetes was increased while the abundance of Actinobacteria was decreased as roughage amount increased in the diet. On the other hand the percentage of Bacteroidetes and Firmicutes remained almost similar in the liquid fraction across the three treatments. In the three solid fractions, the percentage of Bacteroidetes was increased while this of Firmicutes and Spirochaetes was decreased as roughage increased in the diet. Classifying the Gir cattle rumen microbiome at genus level, showed that *Clostridium*, *Ruminococcus*, *Eubacterium*, *Butyrivibrio*, *Bacillus*, *Roseburia* and several other genera were in higher percentage in the solid fraction compared to the liquid fraction whereas, *Prevotella*, *Bacteroides*, *Parabacteroides*, *Paludibacter* and *Victivallis* were abundant in the liquid fraction. Furthermore, in the solid fractions, the percentage of *Prevotella* (16.74-27.5%), *Parabacteroides* (1.64-1.94%) *and Flavobacterium* (0.41-0.52%) increased whereas the abundance of *Clostridium* (11.83-9.01%), *Eubacterium* (5.07-4.3%), *Streptococcus* (0.62-0.52%) and *Faecalibacterium* (1.0-0.57) decreased with increased in roughage quantity. Across the liquid fractions, the proportions of *Clostridium* (3.88%), *Fibrobacter* (3.39%), *Parabacteroides* (2.45%), *Ruminococcus* (2.07%), *Eubacterium* (1.88%), *Paludibacter* (1.74%) and *Butyrivibrio* (1.54%) were the highest during the first dietary treatment (G1) (Fig. 1A and Supplementary table T2B). Considering the effect of green and dry roughage, *Prevotella* was found to be more prominent when animals were fed green roughage compared to dry roughage. This has the reverse effect on *Fibrobacter*. The percentage of *Bacteroides* was almost the same in both types of roughage while some genera depicted in Fig. 1B and Supplementary table T2C showed variation with different proportions of green and dry roughages although the difference was not statistically significant ($P>0.05$).

Looking into the different functional KO categories encoded by this bacterial community, within the metabolism category at level 2, amino acid and carbohydrate metabolism was most abundant in all treatments of liquid and the solid fractions. However, statistical analysis revealed that these categories differed significantly only in the liquid fractions of each treatment (Supplementary table T3A; P < 0.01). At level 3 of carbohydrate metabolism category, in the solid fraction, the proportion of reads related to fructose and mannose metabolism (PATH:ko00051) and pyruvate metabolism (PATH:ko00620) increased significantly (P<0.05) as roughage proportion increased (Fig. 1C, Supplementary table T3C). While looking into environmental information processing category al level 2, in the solid fractions, reads assigned to signal transduction category increased (P <0.05) as roughage proportion increased (Supplementary table T3D). In the case of cellular process at level3, none of the categories showed much variation across the three treatments (Supplementary table T3E). While looking at the effect of the type of roughage (green and dry), no much variation was observed in the overall metabolism and carbohydrate metabolism (P > 0.1) (Supplementary fig. S5).

Taxonomic and functional classification of rumen metatranscriptome

Similarly, in the case of metatranscriptome data, as expected, PCoA showed distinct clusters for each samples of liquid and solid fractions (Supplementary fig. S6 A&B; PERMNOVA, P = 0.001. However, when analysed for each dietary treatment, similar to metagenome data, we observed overlapping clusters in the liquid fraction with PERMANOVA, P = 0.002 (Supplementary fig. S6 C & D). For the solid fraction, PCoA revealed distinct clusters for each treatment for taxonomy while G1 and G3 overlaps in case of functions (Supplementary fig. S6 E & F) however, with significant difference between the three dietary groups (PERMANOVA, P = 0.002). Except *Bacteroidetes,* all the phyla differed significantly (corrected *P* value <0.05) between solid and liquid fraction of MT data (Fig. 2A, Supplementary table T4A). As roughage amount increased in diet, number of reads assigned to the genus *Butyrivibrio* (4.37, 6.93, & 7.23%), *Roseburia* (0.7, 1.82 & 1.91%), *Caldicellulosiruptor* (0.43, 0.98 & 1.84%) and Unclassified *Clostridiales* (0.27, 0.54 & 0.63%) increased in the solid fractions. In the liquid fractions many genera viz. *Clostridium*, *Parabacteroides*, Unclassified Bacteroidetes and *Lactobacillus* decreased as roughage proportion increased (Fig. 2B, Supplementary table T4B). On comparison of both the roughage treatment, bacteria belonging to the genera *Eubacterium*, *Ruminococcus*, *Roseburia* and *Caldicellulosiruptor* increased whereas, *Bacteroides*, *Parabacteroides*, *Paludibacter* and Unclassified *Bacteroidetes* decreased as the proportion of dry roughage increased in the diet. During green roughage, *Clostridium* (9.58%), *Eubacterium* (6.1%) and *Butyrivibrio* (5.03%), *Ruminococcus (*4.76%), *Paludibacter* (1.82%), *Bacillus* (0.94%) and *Roseburia* (0.98%) dominated in the G1 treatment (Fig. 2C, Supplementary table T4C).

In contrast to metagenome data, where amino acid metabolism was the most abundant category, in metatranscriptome data, carbohydrate metabolism was the most abundant category followed by amino acid metabolism (supplementary table T5A). Furthermore, at level3, pyruvate metabolism [PATH:ko00620], starch and sucrose metabolism [PATH:ko00500], butanoate metabolism [PATH:ko00650] and inositol phosphate metabolism [PATH:ko00562] dominated in solid fractions as compared to liquid. Moreover, butanoate metabolism increased as roughage amount increased in the diet (Fig. 2D). Within the environmental information processing category at level 2 the scenario for membrane transport and signal transduction in the solid fraction was reverse as it was observed in the metagenome dataset (supplementary table T5B). In the same category at level 3, ABC transporters [PATH:ko02010], and two-component system [PATH:ko02020] increased in the solid fraction as roughage increased (Supplementary table T5C). Within cellular process category at level 3, flagellar assembly [PATH:ko02040] related genes increased while genes related to bacterial chemotaxis [PATH:ko02030] decrease ($P < 0.05$) in the solid fractions as roughage increased (Supplementary table T5D). In case of metatranscriptome too, type of roughage (green and dry) does not revealed significant impact on metabolism and carbohydrate metabolism ($P > 0.5$) (Supplementary fig. S7).

Expression of CAZy family genes

Regarding the of CAZy expression profiles, in the solid fractions: a) in the G2 treatment, expression of various glycoside hydrolases, including cellulases [GH5 (CAL91974.1), GH9 (AEX9271.8) and GH45 (ACX75024.1)], hemicellulase [GH11 (AEQ15463.1)], xylanoglucanase [GH16 (AHW59387.1)] as well as various carbohydrate binding modules [CBM3 (CAL91977.1), CBM14 (AJO25038.1), CBM18 (BAB21577.1), CBM22 (AAT48119.1), CBM46 (ADU29456.1) & CBM63 (AKT42641.1)] was the highest compared to the other two treatments; b) in the G2 treatment, oligosaccharide degrading and debranching enzymes such as GH31 (AHF24694.1), GH77 (CBL24834.1) and CBM50 (CDI50176.1) was higher; c) in G1 treatment, expression of genes encoding for various carbohydrate binding modules, CBM13 (ADL52958.1), CBM77 (AIQ50255), CBM79 (ERJ89368.1) was higher (Supplementary fig. 8A, Supplementary table T6).

In the liquid fractions mainly oligosaccharide degrading and debranching enzymes expressed: a) in the G3 treatment, oligosaccharide degrading enzymes, GH13 (CRY95540.1, AEW20948.1, ADE81350.1), GH18 (ABY40380.1), GH39 (EEV37797.1); b) in the G2 treatment, during G2 treatment, GH10 (ALJ61518.1), GH13 (ADE82909.1), GH51

(ACM91037.1) and c) in G1 treatment, GH10 (AFU34339.1), GH13 (CAL92192.1), GH23 (AGZ39466.1), GH27 (ABE81161.1) and GH32 (CCE82332.1) all these are oligosaccharide degrading and debranching enzymes as well as CBMs, CBM37 (ADU23820.1), CBM13 (AIG56282.1), CBM22 (AAT48119.1) and CBM20 (CBZ54209.1) expressed at high extent. (Supplementary fig. 8B, Supplementary table T7).

Comparison of Gir cattle rumen metagenome and metatranscriptome

Since, the main objective of this study was to identify the active ruminal bacterial community, therefore we compared both the proportion of bacteria and their function in the MT and MG dataset. PCoA illustrated remarkable differences between MG and MT datasets in both taxonomic and functional profiles in the Gir cattle rumen (Supplementary fig. S9A & B). As was expected, the liquid and solid fractions of MT and MG datasets formed separate clusters (Supplementary fig. S9 C & D), suggesting the presence of huge variation between the detected bacterial community and their 'active' participation in the rumen. Abundance of Firmicutes, Proteobacteria and Actinobacteria was higher in metatranscriptome data. Conversely, Bacteroidetes and Fibrobacter were higher in metagenome data (Supplementary fig. S10).

Comparison of alpha diversity indices for the liquid and solid fractions of each treatment for metagenome and metatranscriptome dataset are presented in the supplementary table T8. In both datasets, the index dominance (D) was found to be increased as roughage proportion increased in the diet with the exception of the liquid fraction in the MG. This observation was also supported by the decrease in overall bacterial diversity based on Renyi index in the solid fraction (Supplementary fig. S11).

The abundance of *Proteobacteria, Spirochaetes, Fibrobacteres, Thermotogae Treponema* and *Rhodospirillum* was higher in the solid fractions of the MT compared to MG dataset across all the three treatments. In addition, the abundance of *Firmicutes* and *Actinobacteria* as well as the genera such as *Clostridium*, *Ruminococcus*, *Eubacterium*, *Butyrivibrio*, *Roseburia*, *Bacillus*, *Alkaliphilus*, *Atopobium*, *Lactobacillus* and *Caldicellulosiruptor* was higher in the MT dataset during G1 and G3 treatments. (Fig. 3A & 3B). Species level analysis (> 1% abundance) revealed that, *Clostridium proteoclasticum*, *Ruminococcus albus*, *Eubacterium rectale*, *Clostridium phytofermentans*, *Clostridium saccharolyticum*, *Clostridium thermocellum*, *Eubacterium eligens*, *Ruminococcus flavefaciens*, *Roseburia inulinivorans*, *Roseburia intestinalis* were highly active in the solid fraction of MT data (Supplementary table T9A).

Regarding the liquid fraction Proteobacteria and Spirochaetes were more abundant in the MT dataset, while Fibrobacteres, Lentisphaerae and Verrucomicrobia were higher in the MG dataset (Fig. 3C). At genus level, *Bacteroides*, *Parabacteroides*, *Porphyromonas*, *Paludibacter*, Unclassified Bacteroidetes, *Treponema* and Candidatus *Azobacteroides* were higher in the MT dataset across all the three dietary treatments. Conversely, *Prevotella*, *Fibrobacter*, *Ruminococcus* and *Victivallis* were higher in the MG dataset (Fig. 3D). Species level analysis (> 1% abundance) revealed that, several *Bacteroides* species, *Paludibacter propionicigenes*, *Porphyromonas gingivalis* and *Eubacterium rectale* were highly active in the liquid fraction of the MT dataset (Supplementary table T9B).

Comparison of the functional profiles of the MT and MG datasets revealed several important functional categories highly transcribing in the rumen bacteria community. In the solid fraction, carbohydrate metabolism category and within this pyruvate metabolism [PATH:ko00620], and inositol phosphate metabolism [PATH:ko00562] were more abundant in the MT compared to the MG dataset (Fig. 4A & 4B). In addition, butanoate metabolism [PATH:ko00650] was increased G3 treatment (Fig. 4B). In the liquid fractions, apart from carbohydrate metabolism, other functional categories that the bacteria were found to be actively involved were energy and lipid metabolism (Fig. 4C). Genes related to cell motility, bacterial chemotaxis [PATH:ko02030] and flagellar assembly [PATH:ko02040] progressively increased from the G1 to G3 treatment in both , the liquid and solid fraction of the MT dataset (Supplementary fig S12 A, B, C & D). In the environmental information processing category in the solid fraction signal transduction was abundant while membrane transport was less during three treatments while it was reverse in liquid fraction.

LEFSe analysis identified in the solid fraction 32 taxa (7 genera) and in the liquid fractions 13 taxa (4 genera) with significant difference among the six treatment groups (i.e. all the dietary treatments of both MG and MT datasets) (Fig. 5A , 5B,C & D), ($P < 0.05$ and LDA score > 4.0). When MG and MT datasets were compared, irrespectively of the dietary treatments, 14 taxa (4 genera) (Supplementary fig. S13A) and 10 taxa (3 genera) differed significantly ($P < 0.05$ and LDA score > 4.0) (Supplementary fig. S13B) in the solid and liquid fractions, respectively. Moreover, when only genus with abundance >0.1% were analysed, 38, 14 and 30 genera were exclusively present in the G1, G2 and G3 in the MT dataset, while 15, 1 and 3 genus were solely present in the MG dataset (Fig. 5E, also see Supplementary table T10).

Discussion

The host digestive enzymes and ruminal microbes are responsible for the digestion of plant fibres ingested by ruminants. Previous research has focused on various aspects of rumen microbial ecology and several studies have been conducted aiming to understand the microbiome composition and their function in the rumen in order to enhance feeding efficiency [26, 34, 39]. In the present study, we performed both metagenome shotgun and metatranscriptome sequencing to explore the active community in the rumen and understand in depth how ruminal microbes respond to different roughage /-concentrate proportions diets in the Gir cattle.

We found that different proportions of roughage in the diet affected both microbiome composition and their function in the rumen, especially in the solid fraction. These differences among the liquid and solid fraction identified in the MT dataset, and to a lesser extent in MG, suggested that there is a special role of specific bacteria in the fibre degradation and subsequent fermentation of released building blocks of different plant polymers. This results are in accordance with previous studies of rumen [15, 48, 53, 67, 78]. The shift in the microbial population observed during the different diet treatments suggested that different rumen bacteria show a preference for specific substrates and their abundance changes accordingly. These results are in accordance with previous studies in which animals were offered different roughage proportions [42, 69]. Moreover, green and dry roughages revealed almost similar profiles of rumen bacterial composition suggesting that the proportion of roughage and not the type (dry or green) affect the microbial population.

The primary role of the ruminal bacteria is to degrade the plant polysaccharides using potent hydrolytic enzymes. Our previous studies with different breeds of buffaloes [42, 47] and cattle [15] revealed that glycosyl hydrolases (GH) is the actively engaged CAZy family in the rumen microbes. In Gir cattle rumen, cellulolytic and xylan degrading fibrolytic enzymes (GH5, GH9, GH11, GH16 & GH45) were highly expressed in the solid fraction when the animals were provided 100% roughage in their diet. Interestingly, our findings are in accordance with previous studies of different cattle breeds [20, 66, 74], although they were fed with different diets which is rich in fibres. However, our findings diverse from those reported in buffalo [22] probably because of differential host-microbiota interaction since one is cattle and another is buffalo. However, further comparative study between cattle and buffalo is required in this direction. Moreover, since higher expression of cellulase, xylan and chitin binding modules further aids in the hydrolysis of different plant polymers would be expected that as fibre content increases in the diet, the expression of fibrolytic enzymes in the rumen bacteria would also increase. GH13 gene family, which is mostly consisting of a diverse group of amylases, was constantly expressed in the liquid fraction during the three treatments. As explained previously [43] in the rumen, alpha amylases break down starch into water soluble dextrins and oligosaccharides which cannot be further hydrolysed by this enzyme. Hence, over expressed GH13 family enzymes in the liquid fraction might be debranching enzymes including pullulanase and iso-amylase.

The comparison of taxonomic and functional profiles of MG and MT dataset revealed many remarkable differences which suggested that not only the bacterial abundance but their dynamism is important. For example, in the MG solid fraction most of the phyla were lower in proportion compared to MT; although, their abundance was lower, they were metabolically very active. In contrast, the higher proportion of *Bacteroidetes* (in G1 and G2 treatments), *Firmicutes* & *Actinobacteria* (in G1 treatment) and *Fibrobacteres* (in G2 Treatment) identified in the MG compared to MT dataset suggested that they were less active metabolically during the respective diet treatments.

Prevotella ruminicola was found to be the most abundant species in our study (both in the liquid and solid fraction). This result is in accordance with previous rumen metagenome studies [22, 45, 69]; however, we found that it was not very metabolically active. Instead, *Bacteroides*, *Parabacteroides*, Candidatus *Azobacteroides*, *Paludibacter* (*species propionicigenes*), *Porphyromonas*, *Flavobacterium*, *Capnocytophaga* (all belonging to phyla *Bacteroidetes*) were more active especially in the liquid fraction. Apart from *Bacteroide*s, *Paludibacter* have been reported in several cattle rumen studies [8, 24, 28, 32, 51], however, their special function in rumen remain unclear. *Paludibacter propionicigenes* have been also observed in sheep fore stomach [50]. This bacterium can utilize various sugars and produce propionate and acetate as major fermentation products [70] and their higher proportion in the rumen MT dataset suggest that this bacterium might play an important role in the fermentation of various. Similarly, in the solid fraction several fibrolytic bacteria such as *Clostridium*, *Ruminococcus*, *Eubacterium*, *Butyrivibrio*, *Roseburia*, *Treponema*, *Caldicellulosiruptor,* and *Rhodospirillum* were found to be major fibre degraders. In the study of Huws et al., 2016, *Clostridium*, *Ruminococcus*, *Eubacterium*, *Butyrivibrio*, *Roseburia* and other bacteria reported to remain attached to the feed and be in higher proportion during rumen incubation(post 2-4 h) [17]. The species of *Caldicellulosiruptor*, which found are thermophilic cellulolytic bacteria [56] capable of degrading plant bio mass using various glycoside hydrolases [3, 7, 36, 60, 71]. The higher proportion of reads assigned to this genus in our MT dataset implies that plays crucial role in plant fibre degradation in rumen along with the species of *Clostridium*, *Ruminococcus* and *Roseburia*. Moreover, *Treponema,* is a pectin lytic *Spirochetes* that has also been detected in previous rumen studies [1, 33, 46] and it has been reported to work synergistically with cellulolytic bacteria [27]. Furthermore, in the solid fraction, genes related to carbohydrate metabolism were abundant compared to the liquid fraction in the MT dataset. This result further indicates that the microbes associated with the solid fraction are actively engaged in degradation of complex plant carbohydrates into simpler sugars. The higher proportion of energy metabolism, translation and transcription related genes identified in the liquid fraction is also indicative of further metabolism of the released sugars in order to provide available energy to cattle.

Rumen's bacterial community is modified according to type of nutrient source available. The LEFSe analysis revealed the presence of regime-associated genera that might be abundant or active during particular diets. Moreover, higher expression of genes related to cell motility, bacterial chemotaxis, flagellar assembly and signal transduction was found in the solid fraction of G3 treatment. All these genes altogether, may be responsible for chemotaxis and supports the movement of bacteria towards specific feed substrates through the increase expression of specific signalling pathways. Our findings are also supported by previous study [54] showing that rumen bacterial communities repeatedly adapt to changes in dietary composition, nutrient concentration and environmental circumstances. Moreover, our functional analysis suggested that amino acid metabolism related genes were higher in the MG dataset. On the other hand, MT dataset revealed that bacteria in the solid fractions were actively involved in the carbohydrate metabolism, while those in the liquid fraction were engaged not only in the carbohydrate metabolism but also in energy and lipid metabolism. Higher abundance of genes related to pyruvate metabolism and butanoate metabolism in the solid fraction of MT data suggested a crucial role of fibrolytic bacteria (*Clostridium*, *Ruminococcus*, *Eubacterium*, *Butyrivibrio*, *Roseburia*, *Treponema*, *Caldicellulosiruptor*) in the degradation of plant fibre. In addition, the higher expression of inositol phosphate metabolism genes implies that ruminal bacteria also utilize organic phosphate. Higher expression in the liquid fraction of membrane transport as well as energy and lipid metabolism related genes was observed. This suggested that the liquid fraction bacterial community actively converts the degraded plant polysaccharide into useful energy for both themselves and host.

Conclusion

The present study revealed that metagenomics (MG) analysis alone is inadequate to understand in depth the cattle rumen bacterial community and their functions. We coupled with metatranscriptomics (MT) analysis in order to get a better understanding of bacterial community dynamics in the rumen. In conclusion, we observed that varying proportions of roughage in the diet modulate the bacterial population within the cattle rumen. Although *Prevotella* were found to be the most abundant genera in the rumen, they are less metabolically active compared to *Bacteroides*, *Parabacteroides*, Candidatus *Azobacteroides*, *Paludibacter* (*species propionicigenes*), *Porphyromonas*, *Flavobacterium* and *Capnocytophaga* within *Bacteroidetes* group. Several fibrolytic species belonging to bacterial genera *Clostridium*, *Ruminococcus*, *Eubacterium*, *Butyrivibrio*, *Roseburia, Caldicellulosiruptor, Rhodospirillum* and *Treponema* were found to be the major fibre degraders. Moreover, fibre degrading enzymes were expressed more in the solid fraction compared to the liquid fraction of digesta.

Acknowledgment: This research work was funded by Indian Council of Agricultural Research (ICAR) New Deli, India under the Niche area of excellence program on “Metagenomic analysis of ruminal microbes” ,grant number 10 2/2011-EPD; dated 21/10/2011.

References

[1] Bekele, A.Z., Koike, S., Kobayashi, Y., Phylogenetic diversity and dietary association of rumen *Treponema* revealed using group-specific 16S rRNA gene-based analysis. FEMS Microbiol Lett, 316 (2011) 51-60.

[2] Benjamini, Y., Hochberg, Y., Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the royal statistical society. Series B (Methodological), (1995) 289-300.

[3] Blumer-Schuette, S.E., Giannone, R.J., Zurawski, J.V., Ozdemir, I., Ma, Q., Yin, Y., Xu, Y., Kataeva, I., Poole, F.L., 2nd, Adams, M.W., Hamilton-Brehm, S.D., Elkins, J.G., Larimer, F.W., Land, M.L., Hauser, L.J., Cottingham, R.W., Hettich, R.L., Kelly, R.M., Caldicellulosiruptor core and pangenomes reveal determinants for noncellulosomal thermophilic deconstruction of plant biomass. J Bacteriol, 194 (2012) 4015-4028.

[4] Cox, M.P., Peterson, D.A., Biggs, P.J., SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinformatics, 11 (2010) 485.

[5] Dai, X., Tian, Y., Li, J., Luo, Y., Liu, D., Zheng, H., Wang, J., Dong, Z., Hu, S., Huang, L., Metatranscriptomic analyses of plant cell wall polysaccharide degradation by microorganisms in the cow rumen. Appl Environ Microbiol, 81 (2015) 1375-1386.

[6] Dai, X., Zhu, Y., Luo, Y., Song, L., Liu, D., Liu, L., Chen, F., Wang, M., Li, J., Zeng, X., Dong, Z., Hu, S., Li, L., Xu, J., Huang, L., Dong, X., Metagenomic insights into the fibrolytic microbiome in yak rumen. PloS one, 7 (2012) e40430.

[7] Dam, P., Kataeva, I., Yang, S.J., Zhou, F., Yin, Y., Chou, W., Poole, F.L., 2nd, Westpheling, J., Hettich, R., Giannone, R., Lewis, D.L., Kelly, R., Gilbert, H.J., Henrissat, B., Xu, Y., Adams, M.W., Insights into plant biomass conversion from the genome of the anaerobic thermophilic bacterium *Caldicellulosiruptor bescii* DSM 6725. Nucleic Acids Res, 39 (2011) 3240-3254.

[8] De Nardi, R., Marchesini, G., Li, S., Khafipour, E., Plaizier, K.J., Gianesella, M., Ricci, R., Andrighetto, I., Segato, S., Metagenomic analysis of rumen microbial population in dairy heifers fed a high grain diet supplemented with dicarboxylic acids or polyphenols. BMC Vet Res, 12 (2016) 29.

[9] Gaur, G., Kaushik, S., Garg, R., The Gir cattle breed of India-characteristics and present status. Animal Genetic Resources Information, 33 (2003) 21-29.

[10] Gullert, S., Fischer, M.A., Turaev, D., Noebauer, B., Ilmberger, N., Wemheuer, B., Alawi, M., Rattei, T., Daniel, R., Schmitz, R.A., Grundhoff, A., Streit, W.R., Deep metagenome and metatranscriptome analyses of microbial communities affiliated with an industrial biogas fermenter, a cow rumen, and elephant feces reveal major differences in carbohydrate hydrolysis strategies. Biotechnol Biofuels, 9 (2016) 121.

[11] Hammer, Ř., Harper, D., Ryan, P., PAST: Paleontological Statistics Software Package for Education and Data Analysis–Palaeontol. Electron. 4: 9. (2001).

[12] Hansen, P.J., Physiological and cellular adaptations of zebu cattle to thermal stress. Anim Reprod Sci, 82-83 (2004) 349-360.

[13] Henderson, G., Cox, F., Ganesh, S., Jonker, A., Young, W., Collaborators, G.R.C., Janssen, P.H., Erratum: Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. Sci Rep, 6 (2016) 19175.

[14] Hernandez-Sanabria, E., Goonewardene, L.A., Wang, Z., Durunna, O.N., Moore, S.S., Impact of feed efficiency and diet on adaptive variations in the bacterial community in the rumen fluid of cattle. Appl Environ Microbiol, 78 (2012) 1203-1214.

[15] Hinsu, A.T., Parmar, N.R., Nathani, N.M., Pandit, R.J., Patel, A.B., Patel, A.K., Joshi, C.G., Functional gene profiling through metaRNAseq approach reveals diet-dependent variation in rumen microbiota of buffalo (*Bubalus bubalis*). Anaerobe, 44 (2017) 106-116.

[16] Hurtaud, C., Rulquin, H., Verite, R., Effect of infused volatile fatty acids and caseinate on milk composition and coagulation in dairy cows. J Dairy Sci, 76 (1993) 3011-3020.

[17] Huws, S.A., Edwards, J.E., Creevey, C.J., Rees Stevens, P., Lin, W., Girdwood, S.E., Pachebat, J.A., Kingston-Smith, A.H., Temporal dynamics of the metabolically active rumen bacteria colonizing fresh perennial ryegrass. FEMS Microbiol Ecol, 92 (1) (2016) DOI: 10.1093/femsec/fiv137.

[18] Jami, E., White, B.A., Mizrahi, I., Potential role of the bovine rumen microbiome in modulating milk composition and feed efficiency. PloS one, 9 (2014) e85423.

[19] Jesus, R.B.d., Omori, W.P., Lemos, E.G.d.M., Souza, J.A.M.d., Bacterial diversity in bovine rumen by metagenomic 16S rDNA sequencing and scanning electron microscopy. Acta Scientiarum. Anim Sci, 37 (2015) 251-257.

[20] Jose, V.L., Appoothy, T., More, R.P., Arun, A.S., Metagenomic insights into the rumen microbial fibrolytic enzymes in Indian crossbred cattle fed finger millet straw. AMB Express, 7 (2017) 13.

[21] Jose, V.L., More, R.P., Appoothy, T., Arun, A.S., In depth analysis of rumen microbial and carbohydrate-active enzymes profile in Indian crossbred cattle. Syst Appl Microbiol, 40 (2017) 160-170.

[22] Kala, A., Kamra, D.N., Kumar, A., Agarwal, N., Chaudhary, L.C., Joshi, C.G., Impact of levels of total digestible nutrients on microbiome, enzyme profile and degradation of feeds in buffalo rumen. PloS one, 12 (2017) e0172051.

[23] Kamra, D.N., Rumen microbial ecosystem. Curr Sci, (2005) 124-135.

[24] Kim, M., Yu, Z., Quantitative comparisons of select cultured and uncultured microbial populations in the rumen of cattle fed different diets. J Anim Sci Biotechnol, 3(1) (2012) 28.

[25] Kopylova, E., Noe, L., Touzet, H., SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics, 28 (2012) 3211-3217.

[26] Krause, D.O., Denman, S.E., Mackie, R.I., Morrison, M., Rae, A.L., Attwood, G.T., McSweeney, C.S., Opportunities to improve fiber degradation in the rumen: microbiology, ecology, and genomics. FEMS Microbiol Rev, 27 (2003) 663-693.

[27] Kudo, H., Cheng, K.J., Costerton, J.W., Interactions between *Treponema bryantii* and cellulolytic bacteria in the in vitro degradation of straw cellulose. Can J Microbiol, 33 (1987) 244-248.

[28] Kumar, S., Indugu, N., Vecchiarelli, B., Pitta, D.W., Associative patterns among anaerobic fungi, methanogenic archaea, and bacterial communities in response to changes in diet and age in the rumen of dairy cows. Front Microbiol, 6 (2015) 781.

[29] Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol, 10 (2009) R25.

[30] Larue, R., Yu, Z., Parisi, V.A., Egan, A.R., Morrison, M., Novel microbial diversity adherent to plant biomass in the herbivore gastrointestinal tract, as revealed by ribosomal intergenic spacer analysis and rrs gene sequencing. Environ Microbiol, 7 (2005) 530-543.

[31] Li, W., Godzik, A., Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics, 22 (2006) 1658-1659.

[32] Li, Z., Wright, A.D., Liu, H., Fan, Z., Yang, F., Zhang, Z., Li, G., Response of the Rumen Microbiota of Sika deer (*Cervus nippon*) fed different concentrations of tannin rich plants. PloS one, 10 (2015) e0123481.

[33] Liu, J., Wang, J.K., Zhu, W., Pu, Y.Y., Guan, L.L., Liu, J.X., Monitoring the rumen pectinolytic bacteria *Treponema saccharophilum* using real-time PCR. FEMS Microbiol Ecol, 87 (2014) 576-585.

[34] Malmuthuge, N., Guan, L.L., Understanding host-microbial interactions in rumen: searching the best opportunity for microbiota manipulation. J Anim Sci Biotechnol, 8 (2017) 8.

[35] McCann, J.C., Wickersham, T.A., Loor, J.J., High-throughput methods redefine the rumen microbiome and its relationship with nutrition and metabolism. Bioinform Biol Insights, 8 (2014) 109-125.

[36] Meng, D.D., Ying, Y., Zhang, K.D., Lu, M., Li, F.L., Depiction of carbohydrate-active enzyme diversity in *Caldicellulosiruptor sp*. F32 at the genome level reveals insights into distinct polysaccharide degradation features. Mol Biosyst, 11 (2015) 3164-3173.

[37] Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J., Edwards, R.A., The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinfo, 9 (2008) 386.

[38] Mizrahi, I., The role of the rumen microbiota in determining the feed efficiency of dairy cows, in: Beneficial microorganisms in multicellular life forms, Springer, 2012, pp. 203-210.

[39] Morgavi, D.P., Kelly, W.J., Janssen, P.H., Attwood, G.T., Rumen microbial (meta)genomics and its application to ruminant production. Animal, 7 Suppl 1 (2013) 184-201.

[40] Morris, C.A., A review of genetic resistance to disease in *Bos taurus* cattle. Vet J, 174 (2007) 481-491.

[41] Myer, P.R., Smith, T.P., Wells, J.E., Kuehn, L.A., Freetly, H.C., Rumen microbiome from steers differing in feed efficiency. PloS one, 10 (2015) e0129174.

[42] Nathani, N.M., Patel, A.K., Mootapally, C.S., Reddy, B., Shah, S.V., Lunagaria, P.M., Kothari, R.K., Joshi, C.G., Effect of roughage on rumen microbiota composition in the efficient feed converter and sturdy Indian Jaffrabadi buffalo (*Bubalus bubalis*). BMC Genomics, 16 (2015) 1116.

[43] Ortega Cerrilla, M.E., Mendoza Martínez, G., Starch digestion and glucosemetabolism in the ruminant: a review. Interciencia, 28 (2003) 380-386.

[44] Parks, D.H., Tyson, G.W., Hugenholtz, P., Beiko, R.G., STAMP: statistical analysis of taxonomic and functional profiles. Bioinformatics, 30 (2014) 3123-3124.

[45] Parmar, N.R., Solanki, J.V., Patel, A.B., Shah, T.M., Patel, A.K., Parnerkar, S., Kumar, J.I., Joshi, C.G., Metagenome of Mehsani buffalo rumen microbiota: an assessment of variation

in feed-dependent phylogenetic and functional classification. J Mol Microbiol Biotechnol, 24 (2014) 249-261.

[46] Paster, B.J., Canale-Parola, E., *Treponema saccharophilum* sp. nov., a large pectinolytic spirochete from the bovine rumen. Appl Environ Microbiol, 50 (1985) 212-219.

[47] Patel, D.D., Patel, A.K., Parmar, N.R., Shah, T.M., Patel, J.B., Pandya, P.R., Joshi, C.G., Microbial and Carbohydrate Active Enzyme profile of buffalo rumen metagenome and their alteration in response to variation in the diet. Gene, 545 (2014) 88-94.

[48] Patel, V., Patel, A.K., Parmar, N.R., Patel, A.B., Reddy, B., Joshi, C.G., Characterization of the rumen microbiome of Indian Kankrej cattle (*Bos indicus*) adapted to different forage diet. Appl Microbiol Biotechnol, 98 (2014) 9749-9761.

[49] Pegorer, M.F., Vasconcelos, J.L., Trinca, L.A., Hansen, P.J., Barros, C.M., Influence of sire and sire breed (Gyr versus Holstein) on establishment of pregnancy and embryonic loss in lactating Holstein cows during summer heat stress. Theriogenology, 67 (2007) 692-697.

[50] Pei, C.X., Liu, Q., Dong, C.S., Li, H., Jiang, J.B., Gao, W.J., Diversity and abundance of the bacterial 16S rRNA gene sequences in forestomach of alpacas (*Lama pacos*) and sheep (*Ovis aries*). Anaerobe, 16 (2010) 426-432.

[51] Petri, R.M., Mapiye, C., Dugan, M.E., McAllister, T.A., Subcutaneous adipose fatty acid profiles and related rumen bacterial populations of steers fed red clover or grass hay diets containing flax or sunflower-seed. PloS one, 9 (2014) e104167.

[52] Pitta, D., Pinchak, W., Indugu, N., Vecchiarelli, B., Sinha, R., Fulford, J., Metagenomic analysis of the rumen microbiome of steers with wheat-induced frothy bloat. Front Microbiol, 7 (2016) 689.

[53] Pitta, D.W., Parmar, N., Patel, A.K., Indugu, N., Kumar, S., Prajapathi, K.B., Patel, A.B., Reddy, B., Joshi, C., Bacterial diversity dynamics associated with different diets and different primer pairs in the rumen of Kankrej cattle. PloS one, 9 (2014) e111710.

[54] Pitta, D.W., Pinchak, E., Dowd, S.E., Osterstock, J., Gontcharova, V., Youn, E., Dorton, K., Yoon, I., Min, B.R., Fulford, J.D., Wickersham, T.A., Malinowski, D.P., Rumen bacterial diversity dynamics associated with changing from bermudagrass hay to grazed winter wheat diets. Microb Ecol, 59 (2010) 511-522.

[55] Qi, M., Wang, P., O'Toole, N., Barboza, P.S., Ungerfeld, E., Leigh, M.B., Selinger, L.B., Butler, G., Tsang, A., McAllister, T.A., Forster, R.J., Snapshot of the eukaryotic gene expression in muskoxen rumen--a metatranscriptomic approach. PloS one, 6 (2011) e20521.

[56] Rainey, F.A., Donnison, A.M., Janssen, P.H., Saul, D., Rodrigo, A., Bergquist, P.L., Daniel, R.M., Stackebrandt, E., Morgan, H.W., Description of *Caldicellulosiruptor saccharolyticus* gen. nov., sp. nov: an obligately anaerobic, extremely thermophilic, cellulolytic bacterium. FEMS Microbiol Lett, 120 (1994) 263-266.

[57] Reilly, K., Carruthers, V.R., Attwood, G.T., Design and use of 16S ribosomal DNA-directed primers in competitive PCRs to enumerate proteolytic bacteria in the rumen. Microb Ecol, 43 (2002) 259-270.

[58] Ross, E.M., Moate, P.J., Bath, C.R., Davidson, S.E., Sawbridge, T.I., Guthridge, K.M., Cocks, B.G., Hayes, B.J., High throughput whole rumen metagenome profiling using untargeted massively parallel sequencing. BMC Genet, 13 (2012) 53.

[59] RoyalCommission, Royal Commission On Agriculture In India Vol IX, Government of India, 1927.

[60] Saleh, M.A., Han, W.J., Lu, M., Wang, B., Li, H., Kelly, R.M., Li, F.L., Two distinct alpha-L-Arabinofuranosidases in *Caldicellulosiruptor* species drive degradation of arabinose-based polysaccharides. Appl Environ Microbiol, 83(13) (2017) e00574-17.

[61] Sanders, J.O., History and development of Zebu cattle in the United States. J Anim Sci, 50 (1980) 1188-1200.

[62] Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., Huttenhower, C., Metagenomic biomarker discovery and explanation. Genome Biol, 12 (2011) R60.

[63] Seymour, W., Campbell, D., Johnson, Z., Relationships between rumen volatile fatty acid concentrations and milk production in dairy cows: a literature study. Anim Feed Sci Technol, 119 (2005) 155-169.

[64] Shabat, S.K., Sasson, G., Doron-Faigenboim, A., Durman, T., Yaacoby, S., Berg Miller, M.E., White, B.A., Shterzer, N., Mizrahi, I., Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants. ISME J, 10 (2016) 2958-2972.

[65] Shi, W., Moon, C.D., Leahy, S.C., Kang, D., Froula, J., Kittelmann, S., Fan, C., Deutsch, S., Gagic, D., Seedorf, H., Kelly, W.J., Atua, R., Sang, C., Soni, P., Li, D., Pinares-Patino, C.S., McEwan, J.C., Janssen, P.H., Chen, F., Visel, A., Wang, Z., Attwood, G.T., Rubin, E.M., Methane yield phenotypes linked to differential gene expression in the sheep rumen microbiome. Genome Res, 24 (2014) 1517-1525.

[66] Shinkai, T., Mitsumori, M., Sofyan, A., Kanamori, H., Sasaki, H., Katayose, Y., Takenaka, A., Comprehensive detection of bacterial carbohydrate-active enzyme coding genes expressed in cow rumen. Anim Sci J, 87 (2016) 1363-1370.

[67] Singh, K.M., Jisha, T.K., Reddy, B., Parmar, N., Patel, A., Patel, A.K., Joshi, C.G., Microbial profiles of liquid and solid fraction associated biomaterial in buffalo rumen fed green and dry roughage diets by tagged 16S rRNA gene pyrosequencing. Mol Biol Rep, 42 (2015) 95-103.

[68] Singh, K.M., Reddy, B., Patel, D., Patel, A.K., Parmar, N., Patel, A., Patel, J.B., Joshi, C.G., High potential source for biomass degradation enzyme discovery and environmental aspects revealed through metagenomics of Indian buffalo rumen. Biomed Res Int, 2014 (2014) 267189.

[69] Thoetkiattikul, H., Mhuantong, W., Laothanachareon, T., Tangphatsornruang, S., Pattarajinda, V., Eurwilaichitr, L., Champreda, V., Comparative analysis of microbial profiles in cow rumen fed with different dietary fiber by tagged 16S rRNA gene pyrosequencing. Curr Microbiol, 67 (2013) 130-137.

[70] Ueki, A., Akasaka, H., Suzuki, D., Ueki, K., Paludibacter propionicigenes gen. nov., sp. nov., a novel strictly anaerobic, Gram-negative, propionate-producing bacterium isolated from plant residue in irrigated rice-field soil in Japan. Int J Syst Evol Microbiol, 56 (2006) 39-44.

[71] VanFossen, A.L., Ozdemir, I., Zelin, S.L., Kelly, R.M., Glycoside hydrolase inventory drives plant polysaccharide deconstruction by the extremely thermophilic bacterium Caldicellulosiruptor saccharolyticus. Biotechnol Bioeng, 108 (2011) 1559-1569.

[72] Wallace, R.J., Rooke, J.A., McKain, N., Duthie, C.A., Hyslop, J.J., Ross, D.W., Waterhouse, A., Watson, M., Roehe, R., The rumen microbial metagenome associated with high methane production in cattle. BMC Genomics, 16 (2015) 839.

[73] Wambura, P.N., Gwakisa, P.S., Silayo, R.S., Rugaimukamu, E.A., Breed-associated resistance to tick infestation in *Bos indicus* and their crosses with Bos taurus. Vet Parasitol, 77 (1998) 63-70.

[74] Wang, L., Hatem, A., Catalyurek, U.V., Morrison, M., Yu, Z., Metagenomic insights into the carbohydrate-active enzymes carried by the microorganisms adhering to solid digesta in the rumen of cows. PloS one, 8 (2013) e78507.

[75] Weimer, P.J., Role of the riminal microbiome in the production and compositon of milk. (2014) 33-40.

[76] Welch, B.L., The significance of the difference between two means when the population variances are unequal. Biometrika, 29 (1938) 350-362.

[77] Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., Xu, Y., dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res, 40 (2012) W445-451.

[78] Zebeli, Q., Tafaj, M., Weber, I., Steingass, H., Drochner, W., Effects of dietary forage particle size and concentrate level on fermentation profile, in vitro degradation characteristics and concentration of liquid-or solid-associated bacterial mass in the rumen of dairy cows. Anim Feed Science Technol, 140 (2008) 307-325.

Figure legends:

Fig. 1: Error bar plot depicting microbial community composition in (A) liquid and solid fraction; (B) green and dry roughage and (C) functional classification of the carbohydrate metabolism category (L2) at level3 among the liquid and the solid fractions of cattle rumen digesta using metagenome sequencing. Taxonomic and functional (bacterial hits only) assignment was performed against RefSeq and KO database, respectively with minimum E-value 1E-5 and minimum identity of 70% using MG-RAST. MG stands for metagenome, G1 (50% roughage), G2 (75% roughage) and G3 (100% roughage) in the diet.

Fig. 2: Error bar plot showing microbial community composition (A & B) at phylum and genus level, respectively among the liquid and the solid fractions; (C) green and dry roughage and (D) functional classification carbohydrate metabolism (L2) at level3 among the liquid and solid fraction of cattle rumen digesta using metatranscriptome sequencing. Taxonomic and functional (bacterial hits only) assignment was performed against RefSeq and KO database, respectively with minimum E-value 1E-5 and minimum identity of 70% using MG-RAST. MT stands for metatranscriptome, G1 (50% roughage), G2 (75% roughage) and G3 (100% roughage) in the diet.

Fig. 3: Error bar plot depicting the differential microbial community composition at genus and phylum level in the solid and the liquid fractions of metagenome (MG) and metatranscriptome (MT) data during three sequential dietary treatments. A & B for the solid fractions and C & D for the liquid fractions of the cattle rumen digesta. Taxonomy was assigned using RefSeq with minimum E-value 1E-5 and minimum identity of 70% using MG-RAST. G1 (50% roughage), G2 (75% roughage) and G3 (100% roughage) in the diet.

Fig. 4: Differences in the functional categories in the solid and the liquid fractions of the metagenome (MG) and metatranscriptome (MT) data during three sequential dietary treatments. A) metabolism category (solid fractions) at level2, B) carbohydrate metabolism category at level3 (solid fractions) and C) metabolism category at level2 (liquid fractions Functional assignment (bacterial hits only) was performed against KO database with minimum E-value 1E-5 and minimum identity of 70% using MG-RAST. G1 (50% roughage), G2 (75% roughage) and G3 (100% roughage) in the diet.

Fig. 5: Differential biomarkers among the three dietary treatments of metagenome and metatranscriptome data. (A) & (B) represents cladogram and histogram, respectively for solid fractions of the two datasets (MG & MT) and (C) & (D) represents cladogram and histogram, respectively for liquid fractions of MG and MT dataset. Differentially abundant genera as biomarkers was determined using Kruskal-Wallis test ($P < 0.05$) with LDA score > 3.0. Cladogram represents the differentially abundant families and genera (only top 40% are plotted hare). The root of the cladogram denotes the domain bacteria. The taxonomic levels phylum and class are labelled, while family and genus are abbreviated. The size of each node represents their relative abundance. (E) Represents the differentially abundance genera (> 0.1%) amongst the six groups. G1 (50% roughage), G2 (75% roughage) and G3 (100% roughage) in the diet. MG and MT stands for metagenome and metatranscriptome, respectively.