Archives of Microbiology (2019) 201:1385–1397
https://doi.org/10.1007/s00203-019-01706-z

ORIGINAL PAPER

Understanding the alteration in rumen microbiome and CAZymes profile with diet and host through comparative metagenomic approach

Varsha Bohra[1] · Nishant A. Dafale[1] · Hemant J. Purohit[1]

Received: 27 May 2019 / Revised: 2 July 2019 / Accepted: 11 July 2019 / Published online: 23 July 2019

Abstract

Rumen microbial community harbors a distinct genetic reservoir of potent carbohydrate-active enzymes (CAZyme) that functions efficiently for the deconstruction of plant biomass. Based on this premise, metagenomics approach was applied to characterize the rumen microbial community and identify carbohydrate-active genes of *Bos taurus* (cow) and *Bubalus bubalis* (buffalo) fed on green or dry roughage. Metadata was generated from the samples: green roughage-fed cow (NDC_GR), buffalo (NDB_GR) and dry roughage-fed cow (NDC_DR), buffalo (NDB_DR). Phylogenetic analysis revealed the dominance of Bacteroidetes, Firmicutes, Proteobacteria, Actinobacteria and Fibrobacter in all the four samples, covering 90–96% of the total bacterial population. On finer resolution, higher abundance of bacterial genera *Fibrobacter*, *Bacteroides*, *Clostridium*, *Prevotella* and *Ruminococcus* involved in plant biomass hydrolysis was observed in NDB_DR. Functional annotation using dbCAN annotation algorithm identified 28.13%, 8.08% 10.93% and 12.53% of the total contigs as putatively carbohydrate-active against NDC_GR, NDB_GR, NDC_DR and NDB_DR, respectively. Additional profiling of CAZymes revealed an over representation and diversity of putative glycoside hydrolases (GHs) in the animals fed on dry roughage with substantial enrichments of genes encoding GHs from families GH2, GH3, GH13 and GH43. GHs of families GH45, GH12, GH113, GH128, GH54 and GH27 were observed exclusively in NDB_DR metagenome. A higher abundance of cellulases, hemicellulases, debranching and oligosaccharide hydrolyzing enzymes was revealed in NDB_DR metagenome. Accordingly, it can be concluded that buffalo rumen microbiome are more efficient in plant biomass hydrolysis. The present study provides a deep understanding of the shifts in microbial community and plant polysaccharide deconstructing capabilities of rumen microbiome in response to changes in the feed type and host animal. Activity-specific microbial consortia procured from these animals can be used further for efficient plant biomass hydrolysis. The study also establishes the utility of rumen microbiome as a unique resource for mining diverse lignocellulolytic enzymes.

Keywords Glycoside hydrolase · MG-RAST · dbCAN · CAZyme · Microbial community · Biomass hydrolysis

Introduction

The rumen chamber in ruminant digestive tract has evolved into an efficient fermentation unit, with a remarkable ability for plant biomass utilization. The rumen is inhabited by specialized microflora that harbors complex system for the microbial attachment and hydrolysis of plant polysaccharides into their oligomers and monomers, which are else indigestible by the host (Pope et al. 2010; Yeoman and White 2014). Amid the various microbial domains (bacteria, archaea, fungi, protozoa and bacteriophage) inhabiting the rumen, bacteria are the most dominant representing about 95% of the total microbial population and contribute far more to ruminal fermentation than other inhabitants. These organisms produce a range of

Communicated by Erko Stackebrandt.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00203-019-01706-z) contains supplementary material, which is available to authorized users.

✉ Nishant A. Dafale
na_dafale@neeri.res.in

[1] CSIR - National Environmental Engineering Research Institute, Nagpur, India

enzymes collectively known as CAZymes, including cellulases and hemicellulases attributed for recalcitrant plant polysaccharides deconstruction. Accordingly, rumen can be used as an excellent niche for exploring CAZymes to be used in economic and efficient conversion of plant biomass to biofuels and other value-added products (Brulc et al. 2009; Aggarwal et al. 2017). It has been stated that, the symbiotic microbial communities residing inside the rumen are characteristic of both, the concerned animal species and its diet (Smith et al. 2013; Reddy et al. 2014). When cow and buffalo are provided with same feed type, buffalo utilizes feed more proficiently due to a higher abundance of cellulolytic bacteria present in buffalo rumen as compared to cattle rumen (Wanapat et al. 2000; Chanthakhoun et al. 2012). Dietary interventions and nature of the forage (green or dry roughage) provided to the host have also been proposed to influence rumen microbiome for improved feed digestibility (Thomas et al. 2011; Stiverson et al. 2011). In tropical countries like India, ruminants such as *Bos taurus* (cow) and *Bubalus bubalis* (buffalo) represent a substantial proportion of domesticated animal species and are primarily fed on nutrient-poor, high fiber content seasonal crop residues, commonly termed as roughage. Roughages are further classified into two types; green and dry roughages depending on their quality. *Green roughage* (GR) generally has high moisture content, relatively soluble and is readily hydrolyzed by bacteria. *Dry roughage* (DR) has lesser moisture content, less soluble and rigid to hydrolyze but is mostly fed in the scarcity or unavailability of green roughage, i.e., during summer or drought condition (Hoffmann and Kovacs 2011; Prajapati et al. 2016). Therefore, it becomes essential to reveal differences prevailing between buffalo and cow rumen microbiome under different feeding scenario which is the basis to exploit rumen bacteria for plant biomass deconstruction. However, due to the absence of suitable techniques to analyze such composite ecosystem in the past, the impact of host and diet on composition and function of rumen microbiome is currently under the scanner. Application of small subunit (SSU) rRNA gene sequence analysis have demonstrated the inability of culture-dependent methods to represent the entire bacterial community as large number of the microbial component remains uncultured and uncharacterized, mostly a major fraction of the fibrolytic bacteria (Larue et al. 2005). Understanding carbohydrate utilization in ruminants is expanding with the advent of "Metagenomics". The technique has allowed an in-depth understanding of rumen microbiome and uncovered diverse functional genes/enzymes in the system. Profound understanding of the ruminal microbial community can help in providing opportunities for the conversion of plant biomass into bio based products. In the view of the above, Illumina platform was used for sequencing rumen metagenomes of cow and buffalo, with the objective of exploring phylogenetic profile and CAZymes diversity in ruminants under the similar and varied (green and dry roughage) feeding scenarios. The data obtained in our study were analyzed comparatively for the presence of unique CAZyme families in the rumen ecosystem fed with either type of the diet. The results enhance our information on manipulating microbial communities of livestock to improve plant biomass deconstruction. We believe, study with the purpose to ascertain the effects of feed type and animal species on the rumen microbiota and CAZymes diversity between buffalo and cow, with modern microbial molecular techniques has not been reported.

Methodology

Experimental design and sample collection

The experimental animals, a cow and a buffalo, were selected at Gourakshan Kendra, Nagpur, India. Sampling was carried out at two different time intervals: post-monsoon and pre-monsoon. For post-monsoon sampling, both the animals were nourished with readily available green roughage, whereas for pre-monsoon sampling the animals were nourished with readily available dry roughage for nearly 4 months before sampling point. About 50 ml of the rumen digesta samples were acquired using a sterile stomach tube in 250 ml of autoclaved screw capped bottle using a vacuum pump as illustrated by Singh et al. (2014). Rumen samples were placed in dry ice, brought to the laboratory and immediately treated for DNA extraction.

Simultaneously, culturable approach was employed to screen potent cellulolytic bacteria from the collected rumen samples. Based on plate zymography, RAPD profiling, ability to produce diverse enzymes using chromogenic substrates and genome sequencing *P. polymyxa* ND25, *P. polymyxa* ND24, *B. subtilis* ND23, *M. arborescence* ND21 and *E. cloacae* ND22 were selected for designing the microbial consortia NDMC-1 (Bohra et al. 2019a).

Metagenomic DNA extraction

The metagenomic DNA extraction of rumen samples was accomplished by means of QIAmp DNA stool mini kit (Prajapati et al. 2016). DNA was quality checked on 0.8% agarose gel and quantified on Qubit® 2.0 Fluorometer (Life Technologies). A260/280 ratio was determined on Nanodrop ND-1000 spectrophotometer (Thermo Scientific, India).

Metagenome library construction, sequencing and assembly

NEB Next Ultra DNA Library Preparation Kit was utilized for constructing the paired-end sequencing library from 200 ng DNA. The library was quality and quantity checked in Bioanalyzer 2100 equipped with high sensitivity DNA chip (Agilent Technologies) and quantified again on Qubit

fluorometer. Libraries were further laden onto Illumina platform for cluster generation and sequencing. CLC Genomics Workbench 6.0 was employed for De novo assembly of high quality PE reads by means of default parameters. The resulting metadata were deposited in the Metagenomics Rapid Annotation Using Subsystem Technology (MG-RAST) server at http://metagenomics.anl.gov/.

Bioinformatics analysis of metadata

Bioinformatics analysis of metadata from all the four samples was performed using MG-RAST server. Taxonomic and hierarchical classification of reads that passed the quality filters were executed using the RefSeq database (Pandit et al. 2018), whereas functional annotation was performed against the Subsystem database (Overbeek et al. 2013) with the maximum e value cutoff: e^{-5} and the minimum identity cutoff: 60%. Data normalization was performed before subsequent analysis. The significance of differences (q value) among microbial phyla, genera and various categories of functional gene was statistically tested using G test (w/Yates') + Fisher's test built-in Statistical Analysis of Metagenomic Profiles (STAMP) software package. An effect size of < 1.00 was fixed for "difference between the proportion' of effect size filter 1 (Jose et al. 2017a).

Mining of carbohydrate-active enzyme using dbCAN CAZyme annotation algorithm

A gene-calling program, FragGeneScan (Ye 2015) was used for predicting putative protein-coding regions in the metadata. This program is put up on a hidden Markov model and combines sequencing error models for improved estimation of the protein-coding region in smaller reads. The identified coding regions were further examined for the occurrence of CAZymes by dbCAN carbohydrate-active enzymes (CAZy) annotation algorithm against the Carbohydrate-Active Enzymes database (Do et al. 2018; Huang et al. 2017). The output so obtained were analyzed manually for determining the diversity and abundance of the various CAZymes classes: glycoside hydrolases (GHs), carbohydrate esterases (CEs), polysaccharide lyases (PLs), glycosyl transferases (GTs), auxiliary activities (AAs) and carbohydrate-binding modules (CBMs) in the rumen metagenomes. The coding sequences predicted as GHs were further characterized by similarity search against NCBI's protein data bank (pdb) and non-redundant (nr) database using the BLASTP algorithm (Bohra et al. 2018, 2019b). Venn diagram was constructed for the comparison of CAZyme, GHs and CBMs within NDC_GR, NDC_DR, NDB_GR and NDB_DR to show the number of proteins shared or unique within a particular relationship.

NCBI accession numbers of rumen microbiome samples

The sequencing data of the present study is submitted to MG-RAST with accession IDs: NDC_GR (mgm4826198.3), NDC_DR (mgm4826199.3), NDB_GR (mgm4826197.3) and NDB_DR (mgm4826196.3).

Results

Culturable approach for designing of microbial consortia

All the four rumen samples were explored for the potent cellulolytic bacterial population. The distinct isolates obtained were further screened using RAPD, plate zymography and were broadly classified based on enzyme activity index (EAI). The group exhibiting EAI > 3 were studied for their efficiency to produce diverse cellulolytic and hemicellulolytic enzymes on chromogenic substrate. Of these five isolates were selected for genome sequencing and designing microbial consortia NDMC-1 based on their diverse enzymes activities (Bohra et al. 2019a).

Yield and purity assessment of extracted metagenomic DNA

The extracted metagenomic DNA was quality checked by 0.8% agarose gel electrophoresis and quantified using ND1000 nanodrop spectrophotometry. The quantification of metagenomic DNA is commonly performed to determine the average DNA concentration and its purity in a solution. Generally, absorbance at 260 nm (A260) is used to determine the concentration of DNA and A260/280 ratio is calculated to assess purity. A260/280 ratio below 1.7 indicates the protein contamination. A260/280 ratio of all four extracted metagenomic DNA ranged between 1.78–1.89 and DNA concentration ranged between 204–338 ng/µL. DNA extraction results showed that the DNA isolation procedure that we opted gave the best output in terms of both enough and intact metagenomic DNA.

Metagenome sequence data statistics

Illumina next generation sequencing led to the thorough elucidation of rumen microbial diversity and their functional capacity to hydrolyze lignocellulosic biomass in cow and buffalo rumen under different (roughage) feeding scenario. Metagenome sequencing of DNA from all the four samples generated about 255.71 Mb metadata with 2,80,220 raw reads and 2,72,700 reads post quality filtering having an average read length ranging from 673 to 1343 base pairs. The

metagenomic data analysis statistics of individual samples calculated by in-house Perl scripts are detailed in Table 1. As evident from the MG-RAST annotation, the metagenomic libraries were found to have very rare number of SSU rRNA hits, comprising 0.009–0.13% of the total reads per rumen microbiome sample while 47–66% of the reads were assigned to various functional subsystems. Similar data have been used earlier to inspect rumen microbial ecology with reference to aromatic compounds metabolism, resistome, mobilome and stress, dormancy and spore producing genes in bacteria (Singh et al. 2014; Prajapati et al. 2016, Reddy et al. 2014). In this study, the metagenomic sequences were explicitly used to assess lignocellulosic biomass hydrolyzing capabilities of two different ruminant organisms, i.e., cow and buffalo when fed with seasonal roughage diet.

Taxonomic diversity of rumen microbiome samples

At the domain level, with the abundance of more than 93%, bacteria were found to be the most abundant domain in all the samples followed by archaea ranging between ~ 0.6 and 5.3%. The proportion of eukaryotes ranged between 0.4 and 1%, while the rest comprised virus and unassigned sequences (Fig. S1).Taxonomic assignment of bacterial taxa using RefSeq database showed the dominance of phylum Bacteroidetes in the rumen microbiome of both, green and dry roughage-fed animals. Bacteroidetes were followed by Firmicutes, Proteobacteria, Actinobacteria and Fibrobacter and constituted about 90–96% of the total phyla across each of the four rumen metagenomes. Rest 4–10% majorly comprised phyla Chlorobi, Cyanobacteria, Spirochaetes, Verrucomicrobia, Elusimicrobia, Fusobacteria, Chloroflexi, Lentisphaerae, Planctomycetes, Acidobacteria, Tenericutes, Thermotogae, Synergistetes and others (Fig. 1a).

Unfolding microbial diversity in response to dietary changes and ruminant animal

Effect of roughage type on the rumen microbial community of cow

As the feeding type was shifted from GR to DR in cow, significant decrease in the abundance of Bacteroidetes, Fibrobacteres coupled with an increase in Firmicutes, Proteobacteria, and Actinobacteria was observed (Fig. 1b). At genus level significant changes in 11 different genera were observed. *Prevotella* and *Bacteroides* were found to be the most abundant genera in both the samples from cow rumen. However, their abundance declined when the diet is switched from green to dry roughage type. *Bacillus*, *Parabacteroides*, *Clostridium*, *Acinetobacter* were found more abundant genera in NDC_DR when compared to NDC_GR (Fig. 1c).

Effect of roughage type on the rumen microbial community of buffalo

With the shift in feeding type from GR to DR in buffalo significant changes were observed in the abundance of 10 different phyla. An increase in the abundance of Firmicutes, Bacteroidetes, Fibrobacteres and Actinobacteria coupled with the

Table 1 Sequence details and data assembly statistics of rumen metagenome from NDC_GR, NDB_GR, NDC_DR and NDB_DR by using in-house Perl scripts

Parameters	NDC_GR	NDC_DR	NDB_GR	NDB_DR
MG-RAST Accession No.	mgm4826198.3	mgm4826199.3	mgm4826197.3	mgm4826196.3
Total number of bases uploaded	15,345,590 bp	97,696,253 bp	40,859,065 bp	101,808,170 bp
Total number of sequences uploaded	22,804	72,739	50,170	134,507
Mean sequence length of bp uploaded	673 ± 263 bp	1343 bps ± 3969 bp	814 ± 575 bp	757 ± 1784 bp
Mean GC count uploaded	48 ± 9%	47 ± 9%	43 ± 10%	47 ± 10%
Artificial duplicate reads	0	0	0	83
Number of sequences failed QC	0	0	0	7520
Total number of bases post-QC	15,345,590 bp	97,696,253 bp	40,859,065 bp	84,077,919 bp
Total number of sequences post-QC	22,804	72,739	50,170	126,987
Mean sequence length post-QC	673 ± 263 bp	1343 bps ± 3969 bp	814 ± 575 bp	662 ± 407 bp
Mean GC count post-QC	48 ± 9%	47 ± 9%	43 ± 10%	47 ± 10%
Predicted protein features	27,160	110,841	64,566	151,826
Predicted rRNA features	2	266	12	202
Identified protein features	14,087 (51.87%)	73,000 (65.86%)	33,244 (51.49%)	71,649 (47.19%)
Identified rRNA features	2	96	12	61
Identified functional categories	7,202	55,246	24,350	8130
α Diversity	127	464	263	432

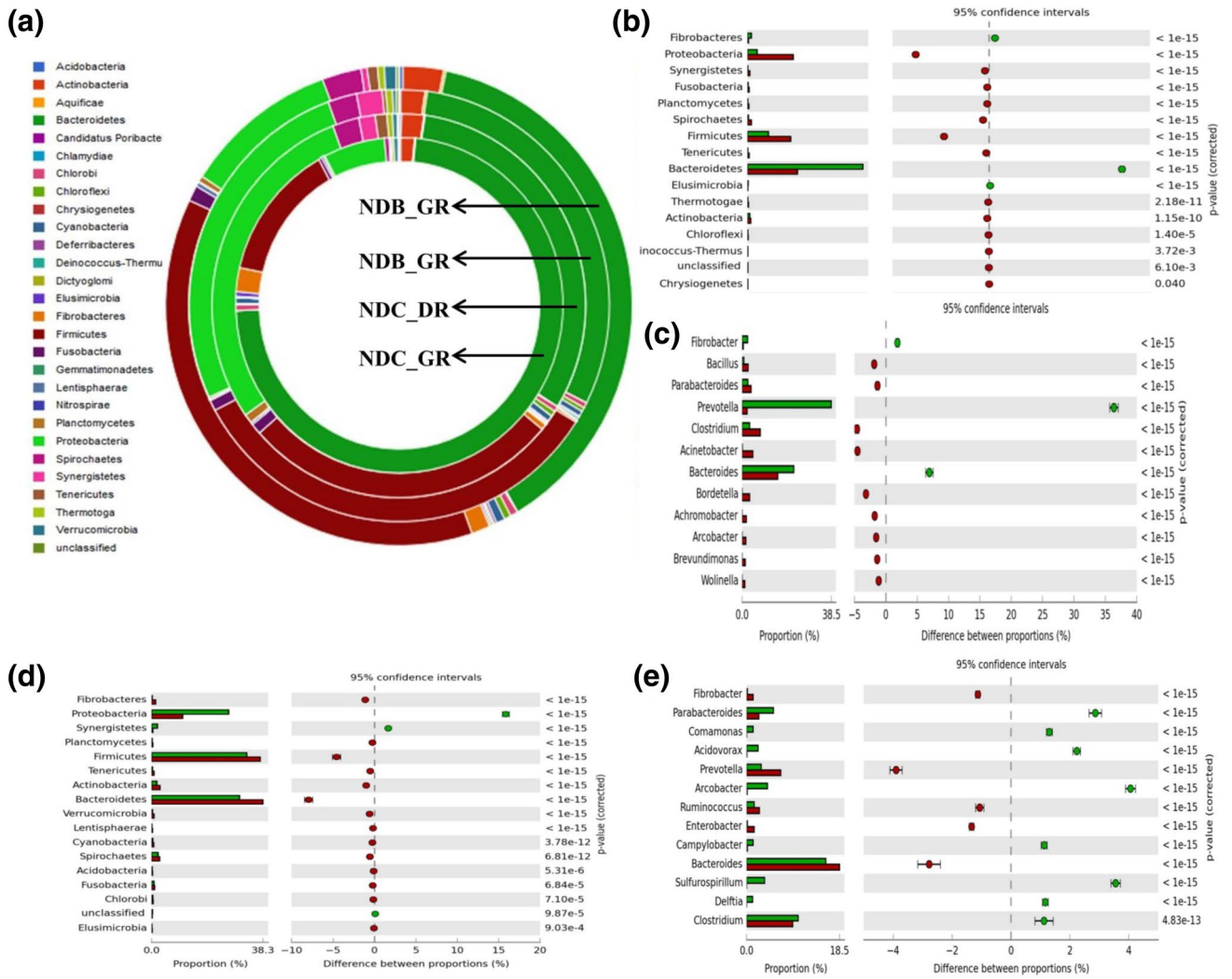

Fig. 1 **a** Taxonomic distribution of bacterial phyla based on predicted proteins and rRNA in metagenome sample from NDC_GR, NDB_GR, NDC_DR and NDB_DR. Statistical analysis using STAMP based on phylum level taxonomic assignments in **b** NDC_GR and NDC_DR, **d** NDB_GR and NDB_DR. Statistical analysis using STAMP based on genus-level taxonomic assignments in **c** NDC_GR and NDC_DR, **e** NDB_GR and NDB_DR. Brown color indicates dry roughage and green color indicates green roughage-fed animals

decrease in Proteobacteria was observed (Fig. 1d). A further genus-level classification of the phyla revealed *Prevotella*, *Bacteroides*, *Clostridium*, *and Parabacteroides* to be the most abundant genera in both the samples from buffalo rumen. The abundance of *Clostridium*, *Parabacteroides* declined while *Prevotella*, *Bacteroides* increased as the diet is switched from green to dry roughage type (Fig. 1e). Buffalo fed on GR shared members of the genera *Micrococcus* with the buffalo fed on DR. The members of *Tetragenococcus* were only identified in the sample NDB_GR.

Interspecies differences in the ruminal microbial community under the same (green roughage) feeding scenario

Phylogenetic comparison of animals fed on green roughage type revealed the significant difference in the abundance of genera Bacteroidetes, Firmicutes, Proteobacteria and Fibrobacter in their metagenomes. Bacteroidetes and Fibrobacter were found to be abundant in cow rumen metagenome whereas Firmicutes and Proteobacteria were abundant in

buffalo rumen fed on green roughage. At genus level *Prevotella*, *Bacteroides*, *Fibrobacter* were abundant genera in cow rumen while *Clostridium*, *Parabacteroides* abundant in buffalo rumen. The maximum difference was found in the abundance of genera *Prevotella* (Figs. S2a, S2b). In addition to this, members of *Candidatus Zinderia* and *Intrasporangium* were identified only in the samples of animal fed with GR.

Interspecies differences in the ruminal microbial community under the same (dry roughage) feeding scenario

Animals fed on dry roughage have a higher profusion of Bacteroidetes, Firmicutes, and Proteobacteria covering more than 60% of the bacterial community. Relatively higher abundance of Bacteroidetes and Firmicutes is apparent in buffalo rumen whereas Proteobacteria was abundant in cow rumen metagenome. At genus level *Clostridium*, *Prevotella* and *Bacteroides* dominated both the rumen metagenome with higher abundance in buffalo rumen (Figs. S2c, S2d). Members of *Thalassobium* were identified only in the samples of animal fed with DR. Members of *Wigglesworthia* were only identified in NDC_DR. Similarly, members of the genera *Leptospirillum* and *Phaeobacter* were found in the NDB_DR.

Functional profiling of the rumen metagenome datasets

Functional annotation of the rumen metagenomes was performed using subsystem database. Out of the 24 subsystems identified, the reads assigned to carbohydrates metabolism were most numerous followed by reads associated with clustering-based subsystems. The abundance of genes responsible for carbohydrates metabolism were found to be 14.52% in NDC_GR, 12.14% in NDB_GR, 11.99% in NDC_DR and 13.66% in NDB_DR as depicted in Fig. S3. A significant difference in the abundance of genes encoding various subsystems with the change in diet from green to dry roughage (Fig. 2) and change in ruminant animal (Fig. S4) was statistically tested using STAMP.

Mining rumen metagenomes to identify CAZymes

The annotation for CAZymes present in metadata was performed against the CAZy database through dbCAN annotation algorithm. dbCAN analysis of assembled sequences identified 28.13 and 8.08% of the total contigs from NDC_GR and NDB_GR, respectively, having resemblance with at least one of the CAZyme. This was in contrast with samples NDC_DR and NDB_DR, exhibiting 10.93 and 12.53% abundance for CAZyme genes associated with the lignocellulosic biomass hydrolysis. The CAZyme encoding contigs were further categorized into diverse categories of CAZymes viz. GHs, CEs, GTs, PLs, CBMs and SLH. GHs catalytic modules were the most predominant category found in our dataset which included the genes involved in plant polysaccharide hydrolysis. GHs was found to be increased with the shift in the diet from green (NDC_GR = 39.51%, NDB_GR = 41.04%) to dry (NDC_DR = 41.80%, NDB_DR = 42.55%) roughage. The enzymes classified to functional group GTs were found to be the second most abundant group in all four samples representing 19–24% of the total CAZymes. CBM modules were the third dominant CAZyme class comprising 20.81, 16.14, 15.38 and 17.60%, respectively, in NDC_GR, NDB_GR, NDC_DR and NDB_DR sample. Approximately 11–14% of the contigs found to be associated with CE families. Of these, CE1 family encoding feruloyl esterases, crucial for lignin deconstruction, was the most prominent CE family with 23.44, 22.35, 25.43 and 25.85% abundance against NDC_GR, NDB_GR, NDC_DR and NDB_DR sample, respectively. The presence of nearly 2–5% PLs and 0.7–2.5% AAs were also detected in the metagenome of studied rumen sample. Dry roughage-fed

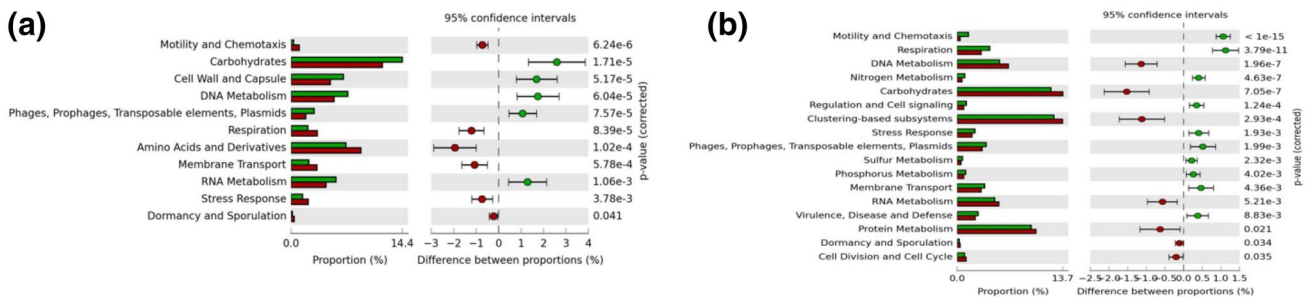

Fig. 2 Statistical analysis using STAMP based on functional assignments at subsystem level in **a** NDC_GR and NDC_DR, **b** NDB_GR and NDB_DR. Brown color indicates dry roughage and green color indicates green roughage-fed animals

rumen samples NDC_DR and NDB_DR displayed a greater abundance of genes associated with GHs and CEs than the green roughage NDC_GR, NDB_GR. The percentage abundance of various CAZyme classes is depicted in Fig. 3.

Comparative jVenn plotting of CAZymes in rumen metagenomes

Venn diagram was constructed for the comparison of CAZyme (Fig. S5), GHs (Fig. S6) and CBMs (Fig. S7) annotated within NDC_GR, NDC_DR, NDB_GR and NDB_DR to show the number of proteins shared or unique within a particular relationship. Overall, 178, 295, 208 and 294 unique CAZyme were annotated in NDC_GR, NDC_DR, NDB_GR and NDB_DR metagenome. Of these 10, 7, 32 and 34 CAZyme were exclusively annotated in respective metagenomes. GHs accounted for about 39–43% of the total CAZyme identified in the studied metagenomes. 89, 155, 114 and 154 unique GHs were annotated in NDC_GR, NDC_DR, NDB_GR and NDB_DR metagenomes. Of these 3, 12, 2 and 14 CAZyme were exclusively present in respective metagenomes. Similar to GHs, the most diverse CBMs (49) with 8 exclusive CBMs were annotated in NDB_DR metagenome. Comparatively high CAZyme (294), GHs (154) and CBMs (49) along with the presence of elite CAZyme 34, GHs (14) and CBMs (8) indicates the proficiency of NDB_DR toward efficient biomass hydrolysis. The significantly enormous number of enzymes categorized into GH family in the rumen metagenome of NDB_DR (154 families) and NDC_DR (155 families) shows a more complex process of plant biomass breakdown in them as compared to others.

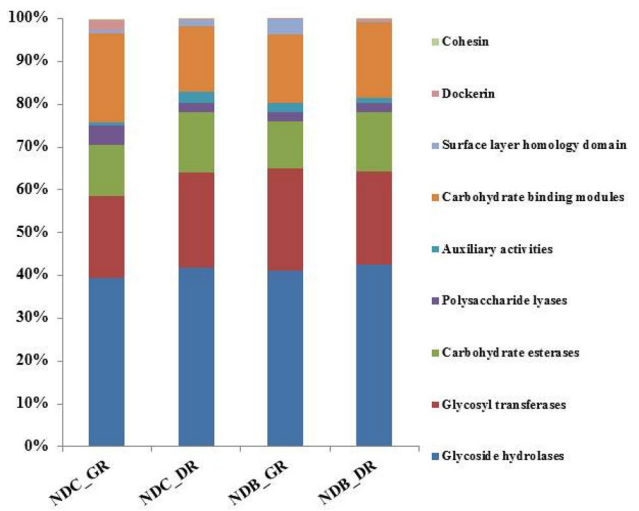

Fig. 3 Comparison of predicted CAZymes classes: GHs, GTs, CEs, CBMs, PLs and AAs annotated in the rumen metagenome samples from NDC_GR, NDB_GR, and NDC_DR and NDB_DR

Comparative metagenomics of GHs and CBMs responsible for plant biomass hydrolysis in rumen metagenomes

Of the total GHs identified, 83.99, 73.84, 74.39 and 77.78% of them were identified to be directly involved in plant polysaccharide hydrolysis against NDC_GR, NDB_GR, NDC_DR and NDB_DR, respectively. Four families of GHs (GH5, GH9, GH44, GH45) identified to be majorly cellulolytic comprised 7.42, 4.01, 3.20 and 4.40% of the total GHs in NDC_GR, NDB_GR, NDC_DR and NDB_DR, respectively, with the dominance of GH5 and GH9. Enzymes dedicated for hydrolysis of oligosaccharides (GH1, GH2, GH3, GH4, GH13, GH18, GH20, GH27, GH29, GH30, GH31, GH32, GH35, GH36, GH38, GH39, GH42, GH43, GH52, GH59, GH57, GH92 GH95, GH97, GH116, GH120, GH130) were the most abundant representing 47.56, 46.87, 46.90, 47.66% of the total GHs, respectively. The analysis also revealed the presence of 9.8–14.6% GHs from GH1, GH3, GH5, and GH9 families, which have an essential role in the breakage of β-1,4-linkage in cellulose polymer and release of glucose residue. Enzymes chiefly accountable for the deconstruction of xylan main chains including, endo-1,4-β-xylanase and endo-1,3-β-xylanase are presented by GH8, GH10, GH11, GH12, GH26, GH28, GH53, GH98, GH113 and GH141 families. Debranching enzymes from family GH23, GH33, GH36, GH51, GH54, GH67, GH77, GH78, GH84, GH103, GH106, GH110 and GH127 responsible for the depolymerization of hemicellulose were also annotated. In total both of these accounted for about 29, 23, 24 and 25% of the total GHs, respectively. Out of the annotated GH families, sequences affiliated to GH3, GH13 and GH43 were most abundant and in total contributed 25.29, 18.94, 19.36 and 24.14% against NDC_GR, NDB_GR, NDC_DR and NDB_DR, respectively. GH3, GH13 and GH43 family encodes for glucan 1,3-β-glucosidase, β-glucosidase, xylan 1,4-β-xylosidase, α-L-arabinofuranosidase, glucan 1,4-β-glucosidase, β-xylosidase, oligo-α-glucosidase, α-L-arabinofuranosidase, arabinanase, β-*N*-acetyl hexosaminidase and xylanase (Fig. 4). CAZymes often exhibit linked modular structure known as CBMs which helps in adhesion to the carbohydrates. In total, fourteen CBMs were annotated in rumen microbiome that associates with both cellulase and hemicellulose. Additionally, 8 and 13 CBMs exclusively associating with cellulase and hemicellulose gene, respectively, were also found in the rumen microbiome. CBM4, CBM6, CBM32, CBM37 and CBM44 were found to be the most abundant covering 5–9% of the GHs (Fig. 5).

A detail of major enzymatic activity carried out by individual GH family, helpful in biomass deconstruction is provided in Table S1. Endo-β-1, 4-glucanases accountable for the breakage of β-1,4-linked glucose residues in long chain cellulose comprised only 0.2–0.4% contigs of

Enzyme	GH family	NDC_GR	NDC_DR	NDB_GR	NDB_DR
Cellulase	GH5	5.57	2.15	3.05	3.46
	GH9	1.62	1.00	0.96	0.89
	GH44	0.23	0.05	0.00	0.00
	GH45	0.00	0.00	0.00	0.06
Endohemicellulase	GH8	1.39	0.26	0.00	2.40
	GH10	2.78	1.10	0.96	1.56
	GH11	0.00	0.05	0.00	0.06
	GH12	0.00	0.00	0.00	0.06
	GH26	3.25	0.52	0.48	1.45
	GH28	4.18	2.26	1.61	3.68
	GH53	3.48	0.37	0.32	1.23
	GH98	0.23	0.00	0.00	0.00
	GH113	0.00	0.00	0.16	0.28
	GH141	0.93	0.37	0.00	0.06
Xyloglucanases	GH16	0.93	1.52	2.57	0.06
	GH17	0.23	0.10	0.32	0.00
	GH55	0.00	0.10	0.16	0.06
	GH64	0.00	0.00	0.16	0.00
	GH74	0.46	0.79	0.64	0.72
	GH81	0.00	0.05	0.00	0.06
	GH94	0.00	0.47	0.32	0.22
	GH128	0.00	0.00	0.00	0.06
Debranching enzymes	GH23	2.32	4.09	5.62	3.40
	GH33	1.16	0.31	1.77	1.39
	GH36	0.46	0.84	0.48	1.06
	GH51	0.23	0.94	0.48	0.33
	GH54	0.00	0.05	0.00	0.11
	GH67	0.23	0.26	0.00	0.33
	GH77	2.32	1.47	2.41	1.34
	GH78	2.78	4.67	2.41	3.40
	GH84	0.00	0.10	0.32	0.11
	GH110	0.00	0.26	0.16	0.17
	GH103	0.00	0.47	0.48	0.00
	GH106	0.93	1.31	0.48	0.61
	GH127	0.70	1.52	0.64	1.23
Oligosaccharide degrading enzymes	GH1	0.23	0.73	2.73	0.72
	GH2	2.78	4.04	1.61	2.34
	GH3	7.19	5.93	5.14	7.25
	GH4	0.70	1.00	2.09	0.72
	GH13	7.66	7.87	5.94	9.48
	GH18	0.70	1.05	0.48	1.34
	GH20	1.62	3.10	2.73	0.22
	GH27	0.00	0.21	0.00	0.61
	GH29	2.32	2.83	2.57	3.07
	GH30	0.70	0.79	1.44	0.78
	GH31	2.78	2.26	2.09	2.68
	GH32	1.39	0.94	1.61	0.89
	GH35	0.93	0.31	0.48	0.61
	GH36	0.46	0.84	0.48	1.06
	GH38	0.23	0.26	0.64	0.78
	GH39	0.00	0.26	0.16	0.22
	GH42	0.00	0.63	0.32	0.33
	GH43	10.44	5.56	7.87	7.41
	GH52	0.00	0.05	0.16	0.11
	GH59	0.00	0.05	0.00	0.06
	GH57	0.46	1.52	1.44	1.39
	GH92	1.62	2.94	3.05	2.06
	GH95	1.62	0.89	0.80	0.61
	GH97	3.02	1.47	1.61	0.95
	GH116	0.00	0.16	0.00	0.11
	GH120	0.00	0.16	0.00	0.33
	GH130	0.70	1.05	1.44	1.51

Fig. 4 Distribution of predicted GH family responsible for cellulose and hemicellulose hydrolysis in rumen metagenome samples from NDC_GR, NDB_GR, NDC_DR and NDB_DR. GH families are assigned to different categories based on CAZy database (http://www.cazy.org). Heat Map lists the percentages of different GH families

the total GHs in the metagenome. In contrast, genes relating to hemicellulose hydrolysis dominated in the rumen samples in all treatment groups. A large proportion of hemicellulolytic enzymes; 36.89, 37.24, 27.81 and 38.18% were identified in sample NDC_GR, NDB_GR, NDC_DR and NDB_DR, respectively. Hemicellulose debranching enzymes like α-L-arabinofuranosidases, α-L-rhamnosidases, and α-glucuronidases that play an important role in depolymerization of hemicellulose were also identified. Endo-1, 4-β-xylanase and endo-1, 3-β-xylanase involved in deconstructing basic xylan chains were estimated to be 2.05 and 6.23% of the total GHs, however debranching enzymes, α-L-arabinofuranosidases, α-L-rhamnosidases, and α-glucuronidases and other oligosaccharide degrading enzyme exhibiting important role in hydrolyzing hemicellulose were profusely annotated in the metagenome samples (Table 2).

Ruminant animal and diet-based alteration in CAZymes distribution

Intraspecies alteration of CaZyme composition in cow rumen with a change in roughage type

The abundance of GHs in cow rumen increased with the shift in the diet from green (39.51%) to dry (41.80%) roughage. As compared to NDC_GR, GH11, GH55, GH81, GH94, GH54, GH84, GH110, GH103, GH27, GH39, GH42, GH52, GH59, GH116 and GH120 involved in lignocellulosic biomass hydrolysis were exclusively annotated in NDC_DR metagenome samples. In addition to various other CAZYmes, oligo-1, 6-glucosidase, arabinan endo-1,5-α-L arabinosidase, mannan endo-1,4-β-mannosidase, α-glucuronidase were annotated only in NDC_DR metagenome sample.

Intraspecies alteration of CaZyme composition in cow rumen with a change in roughage type

With the shift in the diet from green to dry roughage, an abundance of GHs increased from 41.04 to 42.55%. However, the percentage of GHs directly involved in biomass hydrolysis was more in NDB_GR, NDB_ DR comprised members from diverse GHs family. GH45, GH8, GH11, GH12, GH141, GH81, GH28, GH54, GH67, GH27, and GH59 were reported only in NDB_ DR. whereas absent in NDB_GR samples. α-Glucuronidase was annotated only in NDB_ DR with various other CAZymes annotated in NDB_GR.

Interspecies differences in CaZyme profile under the same (green roughage) feeding scenario

Significant difference exists in *CaZyme profile w*hen both the animals were allowed to feed on green roughage. Green roughage-fed buffalo has higher proportion of GHs (41.04%) as compared to green roughage-fed cow (39.51%). The GH profile was more diverse in NDB _GR metagenome as GH39, GH42, GH52, GH55, GH64, GH84, GH94, GH103, GH110, GH113 families were exclusively annotated in them; whereas, GH8, GH44, GH67, GH98, GH141 were only annotated in NDC _GR metagenome. Also oligo-1, 6-glucosidase, mannan endo-1,4-β-mannosidase, arabinan

Carbohydrate substrate	CBM family	NDC_GR	NDC_DR	NDB_GR	NDB_DR
Cellulose/ Hemicellulose binding module	CBM4	1.10	0.61	0.46	0.57
	CBM6	2.11	0.29	0.26	0.40
	CBM9	0.00	0.50	0.13	0.28
	CbM11	0.00	0.09	0.00	0.12
	CBM16	0.46	0.39	0.40	0.43
	CBM28	0.00	0.02	0.00	0.00
	CBM37	1.37	0.81	0.92	1.14
	CBM44	1.01	0.59	0.20	0.64
	CBM59	0.00	0.00	0.13	0.00
	CBM65	0.00	0.00	0.00	0.02
	CBM72	0.27	0.00	0.00	0.02
	CBM76	0.00	0.02	0.07	0.09
	CBM78	0.00	0.02	0.00	0.05
	CBM81	0.00	0.02	0.00	0.00
Cellulose binding module	CBM1	0.09	0.00	0.00	0.00
	CBM2	0.00	0.07	0.00	0.17
	CBM3	0.00	0.00	0.00	0.02
	CBM8	0.09	0.00	0.00	0.00
	CBM30	0.00	0.24	0.07	0.09
	CBM46	0.00	0.04	0.13	0.05
	CBM63	0.00	0.00	0.00	0.05
	CBM64	0.00	0.00	0.07	0.00
Hemicellulose binding module	CBM13	0.82	0.20	0.59	0.14
	CBM15	0.00	0.00	0.00	0.02
	CBM22	0.18	0.18	0.07	0.21
	CBM23	0.00	0.02	0.00	0.09
	CBM29	0.00	0.00	0.00	0.02
	CBM35	1.19	0.15	0.26	0.40
	CBM36	0.09	0.00	0.00	0.02
	CBM42	0.00	0.00	0.00	0.05
	CBM54	0.09	0.13	0.13	0.07
	CBM62	0.18	0.02	0.13	0.09
	CBM57	0.27	0.07	0.00	0.02
	CBM61	1.28	0.13	0.26	0.36
	CBM32	2.20	1.89	1.25	1.78

Fig. 5 Comparison of CBM linked to various cellulases and hemicellulose in rumen metagenome samples from NDC_GR, NDB_GR, NDC_DR and NDB_DR. CBM families are assigned to different categories based on CAZy data base (http://www.cazy.org). Heat Map lists the percentages of different CBMs

endo-1,5-α-L-arabinosidase were annotated in NDB _GR metagenome.

Interspecies differences in CaZyme profile under the same (dry roughage) feeding scenario

Dry roughage displays similar effect on the *CaZyme profile* of cow and buffalo as observed with green roughage. NDB_DR metagenome display higher abundance of GHs (42.55%) as compared to NDC_DR (41.80%). GH12, GH45, GH113, GH128 were annotated only in NDB_DR metagenome. The diverse CaZyme profile of NDB_DR indicates its efficiency toward hydrolyzing complex biomass.

Discussion

Ruminants utilize carbohydrate-rich agricultural waste residues as substantial energy resources. The host itself is unable to synthesize any enzymes required for deconstructing the plant biomass and primarily depend on rumen microbiota to liberate the energy from plant polysaccharides in the form of carbohydrates and sugars. The fiber deconstructing enzymes produced by rumen microflora are of supreme importance in the research related to biofuel generation. Rumen microbial community is crucial for the hydrolysis of plant biomass and varies significantly with changes in the host's age and diet (Kittelmann and Janssen 2011; Li et al. 2012). Substantial differences in the rumen microbiome have also been

Table 2 Percentage abundance of predicted cellulase and hemicellulose enzymes annotated in the rumen metagenome of NDC_GR, NDB_GR, NDC_DR and NDB_DR

Glycoside hydrolases		NDC_GR	NDC_DR	NDB_GR	NDB_DR
3.2.1.4	Cellulase	0.23	0.42	0.13	0.22
3.2.1.8	Endo-1,4-β-xylanase	6.26	2.05	3.05	2.73
3.2.1.10	Oligo-1,6-glucosidase	0.00	0.47	0.80	0.89
3.2.1.21	β-Glucosidase	2.78	2.89	4.01	5.24
3.2.1.22	α-Galactosidase	1.86	2.20	3.53	3.57
3.2.1.23	β-Galactosidase	18.33	8.50	8.83	13.88
3.2.1.25	β-Mannosidase	0.23	1.36	1.61	1.73
3.2.1.37	Xylan 1,4-β-xylosidase	0.93	1.84	3.69	2.62
3.2.1.39	Glucan endo-1,3-β-D glucosidase	0.23	0.10	0.00	0.00
3.2.1.40	α-L-Rhamnosidase	2.78	3.15	2.89	3.57
3.2.1.51	α-L-Fucosidase	3.25	2.26	2.57	3.29
3.2.1.52	Hexosaminidase	0.70	3.31	6.26	2.79
3.2.1.55	α-L-Arabinofuranosidase	2.09	1.89	2.73	2.68
3.2.1.73	Licheninase	0.23	0.10	0.32	0.11
3.2.1.78	Mannan endo-1,4-β-mannosidase	0.00	0.05	0.32	0.00
3.2.1.99	Arabinan endo-1,5-α-L-arabinosidase	0.00	0.31	0.64	0.17
3.2.1.139	α-Glucuronidase	0.00	0.21	0.00	0.17
	Glycoside hydrolases incidence	39.91	31.11	41.38	43.65

reported across discrete animal species (Jami and Mizrahi 2012). In this study, we compared both the rumen microflora and functional genes that putatively encode CAZymes in the rumen metadata from cow and buffalo fed on green and dry roughage diet. The metagenomic approach was used for unfolding the composite rumen biome at a level, not been accomplished previously using traditional culture-dependent studies. Previous studies on rumen ecology have reported the presence of core bacterial community chiefly from the phyla Firmicutes and Bacteroidetes, in spite of the variation in the breed, diet and other environmental aspects (Kittelmann and Janssen 2011; Petri et al. 2013; Thoetkiattikul et al. 2013; Henderson et al. 2015). These findings are consistent with the results obtained in our study, where phyla Firmicutes and Bacteroidetes dominated all four rumen samples irrespective of diet and species, further affirming the existence of core microbiome in the rumen. However, the diversity and abundance of bacterial community in buffalo rumen were found to be higher as compared to cow's rumen similar to the study conducted by Chanthakhoun et al. (2012), who reported a high abundance of cellulolytic bacteria in buffalo rumen than cattle rumen. The difference in community composition between the two animal species is indicated by a significantly higher abundance of Firmicutes and lower abundance of Bacteroidetes in Buffalo rumen as compared to cow rumen. Higher Firmicutes abundance in host rumen is correlated with enhanced hydrolysis of plant biomass and efficient feed utilization (Myer et al. 2015). Accordingly, in this study, the higher richness of Firmicutes in buffalo rumen indicates its higher efficiency in hydrolyzing plant biomass over cow rumen microbiome. Parallel to the findings of Jose et al. (2017a), the abundance of bacterial genus *Ruminococcus* was relatively higher while the abundance of *Prevotella* was lesser in buffalo as compared to cows. *Prevotella* is reported for protein and polysaccharide digestion (Flint et al. 2012), whereas *Ruminococcus* is proficient in the deconstruction of cellulose and hemicellulose (Biddle et al. 2013). The high richness of *Ruminococcus* in buffalo rumen reflects its efficiency for biomass deconstruction. Cellulose and hemicellulose in plant biomass are considered as one of the prevalent renewable resources of fermentable sugars that could be utilized further in numerous industrial processes, like ethanol generation. However, the existence of extremely resilient lignin remains the major bottleneck in accessibility of these fermentable sugars (Alvira et al. 2010). Ruminants with their highly composite and proficient microbial populations can facilitate the hydrolysis of lignocellulose biomass (Hess et al. 2011). The natural ability of rumen microflora to produce a range of potential CAZymes that hydrolyze the harsh plant biomass has been effectively utilized for treating diverse agricultural waste. Ruminant's capability to deconstruct the harsh feed into fermentable sugars is entirely accredited to polysaccharide hydrolyzing enzymes produced by rumen microflora. The complex lignocellulosic degradation machinery of rumen makes it an excellent source for mining CAZymes (Yue et al. 2013; Patel et al. 2014). The CAZy database (http://www.cazy.org) is an assemblage of a diverse group of enzymes exclusively responsible for carbohydrate metabolism. The database provides information concerning all CAZyme families; GHs, GTs, PLs, CEs, CBMs and AAs; thus allowing one to inspect all recognized families and enzymes accountable for cellulolytic

and hemicellulolytic activity. Of various CAZyme family, GHs are the most abundant and diverse group of enzymes accountable for the breakage of glycosidic bonds in plant polysaccharide, comprising 50% of the enzyme classified in CAZyme database (Cantarel et al. 2009). GHs accounted for about 39–43% of the total CAZyme identified. Previous studies have reported the presence of 72 and 78 distinct families of GHs, respectively, from buffalo and cattle rumen that is responsible for the hydrolysis of various plant cell wall constituents (Patel et al. 2014; Jose et al. 2017a). The significantly enormous enzymes categorized into GH family in the rumen metagenome of NDB_DR (154 families) and NDC_GR (155 families) show a more complex process of lignocellulose breakdown compared to others. In line with the other studies (Wang et al. 2013; Pope et al. 2012), a higher proportion of genes from GH5 and GH9 families encoding endoglucanase were present than other GH families. GH5 family proteins are reported to exhibit endo-xylanase mannanase and exoglucanase activity in addition to endoglucanase activity (Dai et al. 2012). The abundance of genes encoding putative GH5 family in NDC_GR indicates that rumen microbiome of cow may comprise a unique approach for plant polycarbohydrate deconstruction. Similar to previously reported studies on ruminants metagenome (Pope et al. 2012; Dai et al. 2012), contigs associated with GH6 and GH7 families cannot be predicted in the metagenome explored in this study. However, Jose et al. (2017b) had identified contig belonging to GH6 family in Indian crossbred cattle rumen microbiome. Enzymes belonging to GH48 family are reported to be one of the components in most of the cellulosome system and therefore the most abundant enzyme subunit in cellulosome-producing bacteria (Artzi et al. 2018; Shinoda et al. 2018). Its absence in studied metagenome points the likelihoods of the non-cellulosomal mode of plant biomass deconstruction in ruminants. Other than GHs, activities of other groups of CAZymes including GTs, CBMs, CEs, PLs, and AA is required for complete depolymerization of plant polysaccharides. GTs, previously stated to be responsible for catalyzing the breakage of glycosidic linkage (Mota et al. 2018), were found to be the second most profuse CAZy family (19–24% of total CAZyme). Of various GT family recognized, GT2 and GT4 families (comprising chitin synthase, cellulose synthase, α-glucosyltransferase, etc.) were the most numerous comprising > 50% of the total GTs. CBMs which do not exhibit the enzymatic activity of their own but helps in binding CAZymes (GHs, CEs, and AAs) to polysaccharide, thus potentiating their activities (Du et al. 2018; Bernardes et al. 2019), accounted for 15–21% of CAZymes in our study. Nearly 11.2–14% of the CAZymes were found to have similarity with class CEs that facilitate the action of various GHs through ester hydrolysis of plant polysaccharides. PLs responsible for cleaving glycosidic bonds (Bertucci et al. 2019) and AAs targeting the insoluble polymers (Sandhu et al. 2018) were the lesser annotated CAZyme classes in studied metagenome.

Conclusion

Ruminants harbor a vast and diverse microbial community that functions in the utilization of lignocellulosic feedstuffs. In the present study, metagenomics approach was employed to portray the rumen microbial diversity and identify the major plant biomass hydrolyzing enzymes of two ruminant animals *Bos taurus* (cow) and *Bubalus bubalis* (buffalo) bred on either green or dry roughage. Results indicate that the rumen bacterial community develops to hydrolyze particular forage is quite specific and vary with variation in feed type and animal species. Phylogenetic profiling of rumen microbiome samples revealed the abundance of bacterial genera *Bacteroides, Fibrobacter, Clostridium, Prevotella*, and *Ruminococcus* in all the samples, however, significant differences were observed between the animals fed on different forage type. Further analysis revealed the existence of enormously diverse CAZyme belonging to classes: GHs, CEs, PLs, GTs, CBMs and AAs, with the dominance of GHs. Of the total GHs identified, 83.99, 73.84, 74.39 and 77.78% against NDC_GR, NDB_GR, NDC_DR and NDB_DR, respectively, were found to be responsible for direct plant polysaccharide deconstruction. The genes related to cellulose and hemicellulose deconstruction were found in the highest abundance in all the sequenced metagenomes. Due to the presence of higher bacterial population present study concludes that the rumen of buffalo as the most efficient plant biomass fermenter.

The study establishes rumen as an important reservoir for gene mining studies with the purpose of prospecting industrially important novel CAZymes simultaneously providing an extended collection of ruminal CAZymes that can be inspected further for obtaining potential enzymes for industrial applicability and development of efficient fermentation systems for bioconversion of plant biomass to high-value products.

Acknowledgements Miss Varsha Bohra thanks the Department of Science and Technology (DST) of India for awarding Junior Research Fellowship (JRF). The funding from DBT project G-1-2282 is gratefully acknowledged for carrying out the work. The authors are thankful to Nagpur Veterinary Hospital and Gourakshan Kendra, Nagpur, for providing rumen samples from the animals. The manuscript has been checked for plagiarism by Knowledge Resource Centre, CSIR-NEERI, Nagpur, India, and assigned KRC No.: CSIR-NEERI/KRC/2019/MARCH/EBGD/1.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

Aggarwal NK, Goyal V, Saini A, Yadav A, Gupta R (2017) Enzymatic saccharification of pretreated rice straw by cellulases from *Aspergillus niger* BK01. 3 Biotech 7(3):158. https://doi.org/10.1007/s13205-017-0755-0

Alvira P, Tomás-Pejó E, Ballesteros MJ, Negro MJ (2010) Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: a review. Bioresour Technol 101:4851–4861. https://doi.org/10.1016/j.biortech.2009.11.093

Artzi L, Dadosh T, Milrot E, Moraïs S, Levin-Zaidman S, Morag E, Bayer EA (2018) Colocalization and disposition of cellulosomes in *Clostridium clariflavum* as revealed by correlative super resolution imaging. MBio 9(1):e00012–18. https://doi.org/10.1128/mBio.00012-18

Bernardes A, Pellegrini VOA, Curtolo F, Camilo CM, Mello BL, Johns MA, Scott JL, Guimaraes FEC, Polikarpov I (2019) Carbohydrate binding modules enhance cellulose enzymatic hydrolysis by increasing access of cellulases to the substrate. Carbohydr Polym 211:57–68. https://doi.org/10.1016/j.carbpol.2019.01.108

Bertucci M, Calusinska M, Goux X, Rouland-Lefèvre C, Untereiner B, Ferrer P, Gerin PA, Delfosse P (2019) Carbohydrate hydrolytic potential and redundancy of anaerobic digestion microbiome exposed to acidosis uncovered by metagenomics. Appl Environ Microbiol 15:9. https://doi.org/10.1128/AEM.00895-19

Biddle A, Stewart L, Blanchard J, Leschine S (2013) Untangling the genetic basis of fibrolytic specialization by Lachnospiraceae and Ruminococcaceae in diverse gut communities. Diversity. 5:627–640. https://doi.org/10.3390/d5030627

Bohra V, Dafale NA, Purohit HJ (2018) *Paenibacillus polymyxa* ND25: candidate genome for lignocellulosic biomass utilization. 3 Biotech 8:248. https://doi.org/10.1007/s13205-018-1274-3

Bohra V, Dafale NA, Hathi Z, Purohit HJ (2019a) Genomic annotation and validation of bacterial consortium NDMC-1 for enhanced degradation of sugarcane bagasse. Ann Microbiol. https://doi.org/10.1007/s13213-019-01462-x

Bohra V, Dafale NA, Tikariha H (2019b) Genomically defined *Paenibacillus polymyxa* ND24 for efficient cellulase production utilizing sugarcane bagasse as a substrate. Appl Biochem Biotechnol 187:266–281. https://doi.org/10.1007/s12010-018-2820-5

Brulc JM, Antonopoulos DA, Miller ME, Wilson MK, Yannarell AC, Dinsdale EA, Edwards RE, Frank ED, Emerson JB, Wacklin P, Coutinho PM, Henrissat B, Nelson KE, White BA (2009) Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. Proc Natl Acad Sci USA 106:1948–1953. https://doi.org/10.1073/pnas.0806191105

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henris-sat B (2009) The carbohydrate-active enzymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res 37:D233–D238. https://doi.org/10.1093/nar/gkn663

Chanthakhoun V, Wanapat M, Kongmun P, Cherdthong A (2012) Comparison of ruminal fermentation characteristics and microbial population in swamp buffalo and cattle. Livest Sci 143:172–176. https://doi.org/10.1016/j.livsci.2011.09.009

Dai X, Zhu Y, Luo Y, Song L, Liu D, Liu L, Chen F, Wang M, Li J, Zeng X, Dong Z (2012) Metagenomic insights into the fibrolytic microbiome in yak rumen. PLoS One 7(7):e40430. https://doi.org/10.1371/journal.pone.0040430

Do TH, Le NG, Dao TK, Nguyen TMP, Le TL, Luu HL, Nguyen KHV, Van LN, Le LA, Phung TN, Van Straalen NM (2018) Metagenomic insights into lignocellulose-degrading genes through Illumina-based de novo sequencing of the microbiome in Vietnamese native goats' rumen. J Gen Appl Microbiol 64(3):108–116. https://doi.org/10.2323/jgam.2017.08.004

Du L, Ma L, Ma Q, Guo G, Han X, Xiao D (2018) Hydrolytic boosting of lignocellulosic biomass by a fungal lytic polysaccharide monooxygenase, AnLPMO15 g from *Aspergillus niger*. Ind Crops Prod 126:309–315. https://doi.org/10.1016/j.indcrop.2018.10.029

Flint HJ, Scott KP, Louis P, Duncan SH (2012) The role of the gut microbiota in nutrition and health. Nat Rev Gastroenterol Hepatol 9:577–589. https://doi.org/10.1038/nrgastro.2012.156

Henderson G, Cox F, Ganesh S, Jonker A, Young W, Abecia L, Angarita E, Aravena P, Arenas GN, Ariza C, Collaborators GRC (2015) Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. Sci Rep 5:14567. https://doi.org/10.1038/srep14567

Hess M, Sczyrba A, Egan R, Kim TW, Chokhawala H, Schroth G, Luo S, Clark DS, Chen F, Zhang T, Mackie RI (2011) Metagenomic discovery of biomass degrading genes and genomes from cow rumen. Science 331:463–467. https://doi.org/10.1126/science.1200387

Hoffmann R, Kovacs B (2011) Compare to different green roughage nutritional value and productivity on Southern Transdanubian region. Bull UASVM Agric 68:168–173. https://doi.org/10.15835/buasvmcn-agr:6436

Huang L, Zhang H, Wu P, Entwistle S, Li X, Yohe T, Yi H, Yang Z, Yin Y (2017) dbCAN-seq: a database of carbohydrate-active enzyme (CAZyme) sequence and annotation. Nucl Acids Res 46(D1):D516–D521. https://doi.org/10.1093/nar/gkx894

Jami E, Mizrahi I (2012) Composition and similarity of bovine rumen microbiota across individual animals. PLoS One 7(3):e33306. https://doi.org/10.1371/journal.pone.0033306

Jose VL, Appoothy T, More RP, Arun AS (2017a) Metagenomic insights into the rumen microbial fibrolytic enzymes in Indian crossbred cattle fed finger millet straw. AMB Express 7(1):13. https://doi.org/10.1186/s13568-016-0310-0

Jose VL, More RP, Appoothy T, Arun AS (2017b) In depth analysis of rumen microbial and carbohydrate-active enzymes profile in Indian crossbred cattle. Syst Appl Microbiol 40(3):160–170. https://doi.org/10.1016/j.syapm.2017.02.003

Kittelmann S, Janssen PH (2011) Characterization of rumen ciliate community composition in domestic sheep, deer, and cattle, feeding on varying diets, by means of PCR-DGGE and clone libraries. FEMS Microbiol Ecol 75:468–481. https://doi.org/10.1111/j.1574-6941.2010.01022.x

Larue R, Yu Z, Parisi VA, Egan AR, Morrison M (2005) 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 7:530–543. https://doi.org/10.1111/j.1462-2920.2005.00721.x

Li RW, Connor EE, Li C, Baldwin VIRL, Sparks ME (2012) Characterization of the rumen microbiota of pre-ruminant calves using metagenomic tools. Environ Microbiol 14:129–139. https://doi.org/10.1111/j.1462-2920.2011.02543.x

Mota TR, de Oliveira DM, Marchiosi R, Ferrarese-Filho O, dos Santos WD (2018) Plant cell wall composition and enzymatic deconstruction. AIMS Bioeng 5(1):63–77. https://doi.org/10.3934/bioeng.2018.1.63

Myer PR, Smith TP, Wells JE, Kuehn LA, Freetly HC (2015) Rumen microbiome from steers differing in feed efficiency. PLoS One 10(6):e0129174. https://doi.org/10.1371/journal.pone.0129174

Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V (2013) The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res 42(D1):D206–D214. https://doi.org/10.1093/nar/gkt1226

Pandit RJ, Hinsu AT, Patel SH, Jakhesara SJ, Koringa PG, Bruno F, Psifidi A, Shah SV, Joshi CG (2018) Microbiota composition, gene pool and its expression in Gir cattle (*Bos indicus*) rumen under different forage diets using metagenomic and metatranscriptomic approaches. Syst Appl Microbiol 41(4):374–385. https://doi.org/10.1016/j.syapm.2018.02.002

Patel DD, Patel AK, Parmar NR, Shah TM, Patel JB, Pandya PR, Joshi CG (2014) Microbial and carbohydrate active enzyme profile of buffalo rumen metagenome and their alteration in response to variation in the diet. Gene 545:88–94. https://doi.org/10.1016/j.gene.2014.05.003

Petri RM, Schwaiger T, Penner GB, Beauchemin KA, Forster RJ, McKinnon JJ, McAllister TA (2013) Characterization of the core rumen microbiome in cattle during transition from forage to concentrate as well as during and after an acidotic challenge. PLoS One 8(12):e83424. https://doi.org/10.1371/journal.pone.0083424

Pope PB, Denman SE, Jones M, Tringe SG, Barry K, Malfatti SA, McHardy AC, Cheng JF, Hugenholtz P, McSweeney CS, Morrison M (2010) Adaptation to herbivory by the Tammar wallaby includes bacterial and glycoside hydrolase profiles different from other herbivores. Proc Natl Acad Sci 107:14793–14798. https://doi.org/10.1073/pnas.1005297107

Pope PB, Mackenzie AK, Gregor I, Smith W, Sundset MA, McHardy AC, Morrison M, Eijsink VG (2012) Metagenomics of the Svalbard reindeer rumen microbiome reveals abundance of polysaccharide utilization loci. PLoS One 7(6):e38571. https://doi.org/10.1371/journal.pone.0038571

Prajapati VS, Purohit HJ, Raje DV, Parmar N, Patel AB, Jones OA, Joshi CG (2016) The effect of a high-roughage diet on the metabolism of aromatic compounds by rumen microbes: a metagenomic study using Mehsani buffalo (*Bubalus bubalis*). Appl Microbiol Biotechnol 100(3):1319–1331. https://doi.org/10.1007/s00253-015-7239-0

Reddy B, Singh KM, Patel AK, Antony A, Panchasara HJ, Joshi CG (2014) Insights into resistome and stress responses genes in *Bubalus bubalis* rumen through metagenomic analysis. Mol Biol Rep 41(10):6405–6417. https://doi.org/10.1007/s11033-014-3521-y

Sandhu SK, Mathur A, Gupta R, Puri SK, Adsul M (2018) Cellulosic biomass-hydrolyzing enzymes. In: Singhania R, Agarwal R, Kumar R, Sukumaran R (eds) Waste to wealth. Springer, Singapore, pp 441–456. https://doi.org/10.1007/978-981-10-7431-8_19

Shinoda S, Kurosaki M, Kokuzawa T, Hirano K, Takano H, Ueda K, Haruki M, Hirano N (2018) Comparative biochemical analysis of cellulosomes isolated from *Clostridium clariflavum* DSM 19732 and *Clostridium thermocellum* ATCC 27405 grown on plant biomass. Appl Biochem Biotechnol. https://doi.org/10.1007/s12010-018-2864-6

Singh KM, Reddy B, Patel AK, Panchasara H, Parmar N, Patel AB, Shah TM, Bhatt VD, Joshi CG (2014) Metagenomic analysis of buffalo rumen microbiome: effect of roughage diet on Dormancy and Sporulation genes. Meta Gene 2:252–268. https://doi.org/10.1016/j.mgene.2014.01.005

Smith DL, Smith T, Rude BJ, Ward SH (2013) Comparison of the effects of heat stress on milk and component yields and somatic cell score in Holstein and Jersey cows. J Dairy Sci 96:3028–3033. https://doi.org/10.3168/jds.2012-5737

Stiverson J, Morrison M, Yu Z (2011) Populations of select cultured and uncultured bacteria in the rumen of sheep and the effect of diets and ruminal fractions. Int J Microbiol. https://doi.org/10.1155/2011/750613

Thoetkiattikul H, Mhuantong W, Laothanachareon T, Tangphatsornruang S, Pattarajinda V, Eurwilaichitr L, Champreda V (2013) Comparative analysis of microbial profiles in cow rumen fed with different dietary fiber by tagged 16S rRNA gene pyrosequencing. Curr Microbiol 67:130–137. https://doi.org/10.1007/s00284-013-0336-3

Thomas F, Hehemann JH, Rebuffet E, Czjzek M, Michel G (2011) Environmental and gut bacteroidetes: the food connection. Front Microbiol 2:93. https://doi.org/10.3389/fmicb.2011.00093

Wanapat M, Puramongkon T, Siphuak W (2000) Feeding of cassava hay for lactating dairy cows. Asian-Aust J Anim Sci 13:478–482. https://doi.org/10.5713/ajas.2000.478

Wang L, Hatem A, Catalyurek UV, Morrison M, Yu Z (2013) Metagenomic insights into the carbohydrate-active enzymes carried by the microorganisms adhering to solid digesta in the rumen of cows. PLoS One 8(11):e78507. https://doi.org/10.1371/journal.pone.0078507

Ye Y (2015) FragGeneScan: predicting genes in short and error-prone reads. In: Nelson KE (ed) Encycl Metagenom Genes Genomes Metagenom Basics, Methods, Databases Tools, vol 1. Springer, US, pp 181–183. https://doi.org/10.1007/978-1-4899-7478-5

Yeoman CJ, White BA (2014) Gastrointestinal tract microbiota and probiotics in production animals. Annu Rev Anim Biosci 2:469–486. https://doi.org/10.1146/annurev-animal-022513-114149

Yue ZB, Li WW, Yu HQ (2013) Application of rumen microorganisms for anaerobic bioconversion of lignocellulosic biomass. Bioresour Technol 128:738–744. https://doi.org/10.1016/j.biortech.2012.11.073