



Evaluation of Rumen Microbes For Milk Fat Yield in Cattle Based on 16S rRNA Amplicon Metagenomic Sequencing

M. Saifudeen¹, K. Anilkumar¹, T.V. Aravindakshan¹, Jamuna Valsalan², K. Ally³, V.L. Gleeja⁴

10.18805/ajdfr.DR-2176

ABSTRACT

Background: Ruminant microbes support growth, lactation and maintenance of animals by providing energy and protein yielding nutrients. Metagenomics is the study of microbial communities directly in their natural environments such as soil, water, gut samples etc. The 16S rRNA gene sequencing is one of the basic genetic markers for the characterisation of prokaryote microorganism community. Current study exploited complex gut microbial communities present in cattle rumen based on variation in milk fat yield.

Methods: The study was conducted in cattle reared at University Livestock Farm, Kerala Veterinary and Animal Sciences University. Twelve healthy cattle in the initial stages of lactation (less than three months) were selected and divided into two groups of six each based on high and low milk fat composition. Rumen samples were collected and DNA was isolated. Amplicons of 16S rRNA has been generated from metagenomic DNA targeting V3 and V4 hyper variable regions. Qiime2 was used for analysis for 16S rRNA amplicon sequencing.

Result: Results revealed that a total of 38,28,241 sequences were obtained after demultiplexing. Majority of sequence length ranged around 295 to 301 nucleotides for forward reads and 216 to 300 nucleotides for reverse reads. After denoising, a total of 6,62,498 sequences were obtained. Clustering (de-replication) produced 711 unique sequence features. In kingdom level of taxonomy, for bacterial kingdom, there was no significant difference in abundance of organisms for high milk fat yield and low milk fat yield cows. For archaea, significant difference observed between abundance of high milk fat and low milk fat groups at five per cent level in the current study.

Key words: Amplicon sequencing, Cattle, Metagenomics, Milk fat, Rumen microbes.

INTRODUCTION

The rumen hosts a complex microbial community comprised mainly of anaerobic bacteria, methanogens, protozoa and fungi (Dewhurst *et al.*, 2001). Ruminant microbes support growth, lactation and maintenance of ruminants by providing energy and protein yielding nutrients (Flint *et al.*, 2008). Jami *et al.* (2014) found out that abundance of specific rumen bacterial groups had higher correlation with physiological parameters such as milk yield and composition. Metagenomics is the study of microbial communities directly in their natural environments such as soil, water and gut samples (Wooley *et al.*, 2010; Di Bella *et al.*, 2013). Metagenomics enabled us to understand the varieties of microorganisms, their functions, collaboration and advancement in a specific biological system (Al-Ajlan and El-Allali, 2018). High throughput sequencing (HTS) technologies used in metagenomics can yield millions of short DNA reads in a single run for the environmental samples (Zhang *et al.*, 2017b).

16S rRNA gene sequencing is one of the basic genetic markers for the characterisation of prokaryote microorganism community (Engelbrektson *et al.*, 2010). Amplicon sequencing on the Illumina MiSeq is a very useful approach for screening large numbers of complex microbial community samples to identify samples that show major perturbations in the composition of microbial genera (Youssef *et al.*, 2009). While conducting 16S rRNA gene sequencing study, one or several hypervariable re-gions

¹Department of Animal Genetics and Breeding, College of Veterinary and Animal Sciences, Mannuthy, Thrissur-680651, Kerala, India.

²Centre for Advanced Studies in Animal Genetics and Breeding, College of Veterinary and Animal Sciences, Mannuthy, Thrissur-680 651, Kerala, India.

³Department of Animal Nutrition, College of Veterinary and Animal Sciences, Mannuthy, Thrissur-680 651, Kerala, India.

⁴Department of Statistics, College of Veterinary and Animal Sciences, Mannuthy, Thrissur-680 651, Kerala, India.

Corresponding Author: Saifeer M. Saifudeen, Department of Animal Genetics and Breeding, College of Veterinary and Animal Sciences, Mannuthy, Thrissur-680 651, Kerala, India.

Email: saifeermsaifudeen@gmail.com

How to cite this article: Saifudeen, S.M., K. Anilkumar, Aravindakshan, T.V., Valsalan, J., Ally, K. and Gleeja V.L. (2024). Evaluation of Rumen Microbes For Milk Fat Yield In Cattle Based on 16S rRNA Amplicon Metagenomic Sequencing Asian Journal of Dairy and Food Research. DOI: 10.18805/ajdfr.DR-2176.

Submitted: 03-11-2023 **Accepted:** 01-06-2024 **Online:** 03-07-2024

were amplified using broad-range primers that each bind to a conserved region and are sequenced (Caporaso *et al.*, 2011).

Rumen microbial organisms of cattle has been found to be associated with enzymatic digestion and fermentation of plant polysaccharides. Ruminant microbes support growth, lactation and maintenance of animals by providing energy

and protein yielding nutrients. Microbial community variation in the rumen of cattle is of great interest due to its connection with economical traits such as milk fat. Current study was focused on comparative analysis of rumen metagenome of high milk fat and low milk fat yielding cows for understanding major rumen microbial communities at kingdom level of taxonomy and to study the effect of rumen microfloral abundance on milk fat yield.

MATERIALS AND METHODS

The study was conducted in cattle under field progeny testing of crossbred cattle (FPT) scheme reared at University Livestock Farm, Mannuthy, Kerala Veterinary and Animal Sciences University on January to March, 2022. Twelve healthy cattle in the initial stages of lactation (less than three months) were selected and divided into two groups of six each based on high and low milk fat composition. 16S rRNA based metagenome sequencing study was conducted to study the taxonomic abundance, diversity and other characters of microbial organism in rumen with respect to milk fat. Cattle with milk fat percentage more than 4.25 and less than 3.5 were considered as high milk fat and low milk fat groups respectively.

All animals were maintained on standard ration (roughage: concentrate ratio of 60:40) on dry matter basis. Rumen liquor was collected from the rumen of each crossbred cattle, three hours after morning feeding. A flexible stomach tube connected to a vacuum pump was introduced for sample collection. The DNA was isolated from rumen liquor of crossbred cattle using cetyl trimethyl ammonium bromide (CTAB) based buffer for cell lysis followed by purification with phenol: chloroform: isoamyl alcohol (Minas *et al.*, 2011). Paired-end sequencing amplified library was analysed and loaded onto the Illumina MiSeq platform for cluster generation followed by paired-end sequencing. Amplicons of 16S rRNA has been generated from metagenomic DNA targeting V3 and V4 hyper variable regions.

Fastq quality check was performed using Fastqc software to analyse quality parameters for the sequence reads. Based on quality report of fastq files, the sequence reads were trimmed, to retain high quality sequences for further analysis. The adapter trimming was performed using the software Trim-galore. Reads which passed the quality control processes were then subjected to detailed bioinformatics analysis. The demultiplexing was done by Illumina MiSeq with the help of corresponding barcodes assigned. Present study used Qiime2 for further analysis for 16S rRNA amplicon sequencing. Quantitative Insights into Microbial Ecology (QIIME) was used to process data from a high-throughput 16S rRNA sequencing study, beginning with multiplexed sequence reads from sequencing instrument and finishing with taxonomic and phylogenetic profiles and comparisons of the samples in the study (Kuczynski *et al.*, 2012). Demultiplexed sequence reads were denoised into amplicon sequence variants (ASVs) and

clustered into operational taxonomic units (OTUs) for reducing sequence errors and for dereplicating sequences. Representative sequences obtained from denoising and clustering method was used as input for taxonomic analysis. Kingdom level taxonomic abundance data were obtained in terms of number of sequence reads for all the six high milk fat yielding cows and six low fat yielding cows.

Test of normality Shapiro-Wilk test was done to check whether the the taxonomic abundance for kingdom level was approximately normally distributed or not in case of both bacteria and archaea. In order to convert non-normal data to normal distribution, log transformation was done. Normally distributed data sets (both before and after transformation) were compared using Independent Sample t test. For those group of data where normality assumptions were not satisfied even after transformation of the data, non-parametric test Mann-Whitney U was applied to find the statistical differences between two groups. Statistical analysis was carried out using SPSS version 24.

RESULTS AND DISCUSSION

Raw sequences of twelve rumen samples, six samples from cattle having comparatively high fat in milk (HFS1, HFS2, HFS3, HFS4, HFS5 and HFS6) and six samples from cattle having comparatively low fat in milk (LFS1, LFS2, LFS3, LFS4, LFS5 and LFS6) were obtained. Number of sequences obtained from the samples under study after demultiplexing process was given in descending order in

Table 1: Summary of sequence counts obtained after demultiplexing

Samples under study	Sequence counts
HFS1	761185
HFS5	592495
LFS5	570664
LFS1	283397
LFS2	280790
HFS6	264017
HFS2	263809
HFS4	222590
LFS6	182921
LFS3	181725
HFS3	160410
LFS4	64238
Total sequences	3828241

Table 2: Nucleotide based demultiplexed sequence length summary

Percentile	Forward reads	Reverse reads
2	48	47
9	68	68
25	295	216
50	299	230
75	300	246
91	301	275
98	301	300

Table 3: Summary of sequences obtained after denoising procedure.

Samples	Input sequences	Denoised	Merging paired end reads	Sequences after denoising	Percentage of denoised
HFS1	761185	407944	232557	161159	21.17
HFS2	263809	93884	49249	35871	13.60
HFS3	160410	42903	20097	15537	9.69
HFS4	222590	73634	37833	28701	12.89
HFS5	592495	151434	71012	61433	10.37
HFS6	264017	108552	53729	46737	17.70
LFS1	283397	73376	26031	22087	7.79
LFS2	280790	131449	75760	55675	19.83
LFS3	181725	84660	31251	30401	16.73
LFS4	64238	26570	18724	15202	23.67
LFS5	570664	348707	194580	155087	27.18
LFS6	182921	83938	42776	34608	18.92

Table 1. HFS1 had highest number of raw reads with more than seven lakh sixty thousand sequences whereas one of the sample from low fat group (LFS4) had got comparatively fewer number of sequences which was around sixty four thousand. A total of 38, 28, 241 sequences were obtained after demultiplexing.

Table 2 described the overall length summary of demultiplexed sequences. Sequence length summary were found out using a random sampling of 10000 out of 3828241 sequences without replacement. Majority of sequence length (25 percentile and above) ranged around 295 to 301 nucleotides for forward reads and 216 to 300 nucleotides for reverse reads.

Table 3 showed the results obtained after denoising of all the twelve DNA samples. Number of input sequences were given in second column. Second last column denoted the frequency of features. Last column of the table represents the per cent of sequences which we could use for further analysis after denoising and reducing the sequence errors. Among high fat samples, we could use 21.17 and 9.69 per cent of the sequences from HFS1 and HFS3 respectively. Among low fat samples, maximum number of useful sequences obtained from LFS5 (27.18 per cent) and minimum number of useful sequences obtained from LFS1 (7.79 per cent). HFS1 had around one lakh sixty thousand sequences for further taxonomy analysis which was the highest number of sequences obtained among all the samples. Similarly the sample with lowest sequence frequency was LFS4 which had only 15202 sequences. After

Table 4: Shapiro-Wilk test of normality for kingdom.

Kingdom	Type	W value
Bacteria	HF	0.76*
	LF	0.73*
Bacteria (log transformed)	HF	0.97NS
	LF	0.93NS
Archaea	HF	0.91NS
	LF	0.82NS

NS-Non significant (values were sampled from normal population)

*-significant at five per cent level (values were sampled from skewed population).

denoising, a total of 6,62,498 sequences were obtained which could be used for further analysis. Clustering (de-replication) produced 100 per cent operational taxonomic units (OTUs). After clustering, 711 sequence features were obtained. Out of total 6,62,498 sequences, there were 711 unique features. For further taxonomical analysis, 711 unique features were used.

Kingdom level taxonomic abundance of both bacteria and archaea were separately identified. Results revealed that abundance data of Archaea is normally distributed whereas abundance data of bacteria is not normally distributed (positively skewed) (Table 4). In order to convert non-normal data to normally distributed data, log transformation was done for bacterial kingdom abundance data. After log transformation, data was found to be normally distributed. Statistical analysis was carried out with the transformed data. However, results were presented with the

Table 5: Independent sample t test for kingdom.

Kingdom	Type	Mean $\pm$ SE	t value
Bacteria	HF	52970.17 $\pm$ 20990.80	0.136NS
	LF	50446.67 $\pm$ 21158.71	
Archaea	HF	5268.17 $\pm$ 1037.92	2.914*
	LF	1730.00 $\pm$ 629.97	

*Significant at five percent level; NS-non significant.

mean and standard error values, which were transformed back in to original units.

In kingdom level of taxonomy, for bacterial kingdom, there was no significant difference in abundance of organisms for high milk fat and low milk fat groups. For archaea, significant difference observed between abundance of high milk fat and low milk fat groups at five per cent level in the current study (Table 5). Mean archaeal abundance obtained for high milk fat yielding and low milk fat yielding cows were 5268.17 ± 1037.92 and 1730.00 ± 629.97 respectively. Archaeal abundance in low milk fat group is significantly less when compared to high milk fat group. We know that volatile fatty acid acetate is the precursor for milk fat and higher acetate production will leads to the elimination of hydrogen from rumen as methane. Significantly higher abundance of methanogenic archaea in rumen of high milk fat yielding cows is evident in current study also.

Coming to the relative frequency of bacterial and archaeal organisms in HF cows under study, 87.32 per cent and 94.70 per cent bacterial abundance observed in high fat and low fat group respectively. Parmar *et al.* (2017) reported that bacterial domain was the most abundant rumen microbial population with an average of 88 per cent abundance in Gir and Kankrej, 90 per cent in Holstein cattle and 84 per cent in Jersey cattle. Similarly archaeal abundance was found 12.68 per cent and 5.30 per cent respectively for high fat and low fat group. In agreement with the bacterial abundance observed in low fat group of the current study, Brulc *et al.* (2009) found that more than 95 per cent of microbial abundance was contributed by bacterial domain in Angus Simmental cross steers.

ACKNOWLEDGEMENT

We are thankful to the Dean, college of veterinary and animal sciences, Thrissur and other officials of Kerala Veterinary and Animal Sciences University for providing necessary facilities for conduction of our research work.

Conflict of Interest

There is no conflict of interest noticed between authors

REFERENCES

- Al-Ajlan, A. and El-Allali, A. (2018). Feature selection for gene prediction in metagenomic fragments. *Bio Data Min.* 11: 1-12.
- Brulc, J.M., Antonopoulos, D.A., Miller, M.E.B., Wilson, M.K., Yannarell, A.C., Dinsdale, E.A., Edwards, R.E., Frank, E.D., Emerson, J.B., Wacklin, P. and Coutinho, P.M. (2009). Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proc. Nati. Acad. Sci.* 106: 1948-1953.
- Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N. and Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *PNAS.* 108: 4516-4522.
- Dewhurst, R.J., Wadhwa, D., Borgida, L.P. and Fisher, W.J. (2001). Rumen acid production from dairy feeds. 1. Effects on feed intake and milk production of dairy cows offered grass or corn silages. *J. Dairy Sci.* 84: 2721-2729.
- Di Bella, J.M., Bao, Y., Gloor, G.B., Burton, J.P. and Reid, G. (2013). High throughput sequencing methods and analysis for microbiome research. *J. Microbiol. Methods.* 95: 401-414.
- Engelbrektson, A., Kunin, V., Wrighton, K.C., Zvenigorodsky, N., Chen, F., Ochman, H. and Hugenholtz, P. (2010). Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *ISME J.* 4: 642-647.
- Flint, H.J., Bayer, E.A., Rincon, M.T., Lamed, R. and White, B.A. (2008). Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nat. Rev. Microbiol.* 6: 121-131.
- Jami, E., White, B.A. and Mizrahi, I. (2014). Potential role of the bovine rumen microbiome in modulating milk composition and feed efficiency. *PLoS one.* [on line]. 9(1). Available: <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0085423>. [22 May 2014].
- Kuczynski, J., Stombaugh, J., Walters, W.A., Gonzalez, A., Caporaso, J.G. and Knight, R. (2012). Using QIIME to analyze 16S rRNA gene sequences from microbial communities. *Curr. Protoc. Microbiol.* 27: 1-5.
- Minas, K., McEwan, N.R., Newbold, C.J. and Scott, K.P. (2011). Optimization of a high-throughput CTAB-based protocol for the extraction of qPCR-grade DNA from rumen fluid, plant and bacterial pure cultures. *FEMS Microbiol. Lett.* 325: 162-169.
- Parmar, N.R., Pandit, P.D., Purohit, H.J., Kumar, J.N. and Joshi, C.G. (2017). Influence of diet composition on cattle rumen methanogenesis: a comparative metagenomic analysis in Indian and exotic cattle. *Indian J. Microbiol.* 57: 226-234.
- Wooley, J.C., Godzik, A. and Friedberg, I. (2010). A primer on metagenomics. *PLoS Comput. Biol.* 6: 115-124.
- Youssef, N., Sheik, C.S., Krumholz, L.R., Najar, F.Z., Roe, B.A. and Elshahed, M.S. (2009). Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. *Appl. Environ. Microbiol.* 75: 5227-5236.
- Zhang, S. W., Jin, X.Y. and Zhang, T. (2017).b. Gene prediction in metagenomic fragments with deep learning. *Biomed. Res. Int.* 10: 1155-1164.