



Metagenome Analysis of Protein Domain Collocation within Cellulase Genes of Goat Rumen Microbes

SooYeon Lim, Jaehyun Seo¹, Hyunbong Choi¹, Duhak Yoon², Jungrye Nam,
Heebal Kim, Seoae Cho^{3,*}, and Jongsoo Chang^{1,*}

Department of Agricultural Biotechnology and Research Institute for Agriculture and Life Sciences,
Seoul National University, Seoul 151-742, Korea

ABSTRACT: In this study, protein domains with cellulase activity in goat rumen microbes were investigated using metagenomic and bioinformatic analyses. After the complete genome of goat rumen microbes was obtained using a shotgun sequencing method, 217,892,109 pair reads were filtered, including only those with 70% identity, 100-bp matches, and thresholds below E^{-10} using METAIDBA. These filtered contigs were assembled and annotated using blastN against the NCBI nucleotide database. As a result, a microbial community structure with 1431 species was analyzed, among which *Prevotella ruminicola* 23 bacteria and *Butyrivibrio proteoclasticus* B316 were the dominant groups. In parallel, 201 sequences related with cellulase activities (EC.3.2.1.4) were obtained through blast searches using the enzyme.dat file provided by the NCBI database. After translating the nucleotide sequence into a protein sequence using Interproscan, 28 protein domains with cellulase activity were identified using the HMMER package with threshold E values below 10^{-5} . Cellulase activity protein domain profiling showed that the major protein domains such as lipase GDSL, cellulase, and Glyco hydro 10 were present in bacterial species with strong cellulase activities. Furthermore, correlation plots clearly displayed the strong positive correlation between some protein domain groups, which was indicative of microbial adaption in the goat rumen based on feeding habits. This is the first metagenomic analysis of cellulase activity protein domains using bioinformatics from the goat rumen. (**Key Words:** Goat Rumen, Shot-gun Sequencing, Metagenome, Protein Domain)

INTRODUCTION

Goats have an extremely varied diet including the tips of woody shrubs, trees, and lignocellulosic agricultural by-products. Symbiotic microbes in the rumen of these herbivores play key roles in providing the hosts with various nutrients. Enzymes secreted by rumen microbes are essential for the conversion of cellulose and hemi-cellulose

into simple sugars, which are metabolized to volatile fatty acids by rumen microbes. Produced volatile fatty acids serve as energy sources for ruminants. Many studies have investigated the symbiotic microorganisms in the rumen because of their link to economically or environmentally important traits such as feed conversion efficiency, methane production (Hegarty, 1999; Guan et al., 2008; Hess et al., 2011). There have been various studies about the correlation between rumen microbiota and their role in nutrients digestion for sheep and cattle. Especially, information for the microbial digestion consortia in goat rumen was expected to provide its species distinct characteristics compared to those of other ruminant animals (McAllister et al., 1994).

A key challenge in this study was identifying rumen microbial profiles, which are associated and potentially predictive of these traits. Thus, methods for profiling the rumen microbial population should be relatively inexpensive and efficient to allow a large number of individuals to be profiled (Ross et al., 2012). Untargeted

* Corresponding Authors: Seoae Cho. Tel: +82-2-876-8820, Fax: +82-2-876-8827, E-mail: seoae@cnkgenomics.com / Jongsoo Chang. Tel: +82-2-3668-4636, Fax: +82-2-3668-4187, E-mail: jschang@knou.ac.kr

¹ Department of Agricultural Science, Korea National Open University, 169 Dongsung-dong, Jongno-gu, Seoul, 110-791, Korea.

² Department of Animal Science, Kyungpook National University, Sangju, 741-711, Korea.

³ C&K genomics Inc. 514 Main Bldg., Seoul National University Research Park, San 4-2 Boncheon-dong, Gwanak-gu, Seoul, 151-742, Korea.

Submitted Apr. 16, 2013; Accepted Apr. 29, 2013; Revised May 11, 2013

rumen bacterial communities contain numerous novel gene sequences based on deep sequencing of pooled samples of true biological variation. The rumen metagenome profile included the counts of reads that aligned to each contig, which could be analyzed using metagenomic tools and correlation plots. The composition of the microbial population differs between goat species and based on their diet. Analysis of microorganisms in the rumen fluid of different herbivores revealed bacteria (10^{10} to 10^{11} cells/ml, representing more than 50 genera), ciliate protozoa (10^4 to 10^6 /ml, from 25 genera), anaerobic fungi (10^3 to 10^5 zoospores/ml, representing six genera), and bacteriophages (10^8 to 10^9 ml). These numbers represented only a small fraction of the microbial species in rumens of animals on fiber-based diets since less than 10 to 20% of microbial populations are cultivable on synthetic media (Zhou et al., 2011). However, metagenomic research has generated genetic information on the entire microbial community, which is important because 99% of microbes cannot be isolated or cultured. The metagenomic method provides a global microbial gene pool without the need to culture of the microorganisms. In this study, we analyzed the complete genome of goat rumen microbes obtained using a shotgun sequencing method. This differed from previous studies on microbes based on 16 rRNA. Also, our results were filtered under strict conditions and provided high-quality results on the rumen microbe community and cellulase activity protein domains.

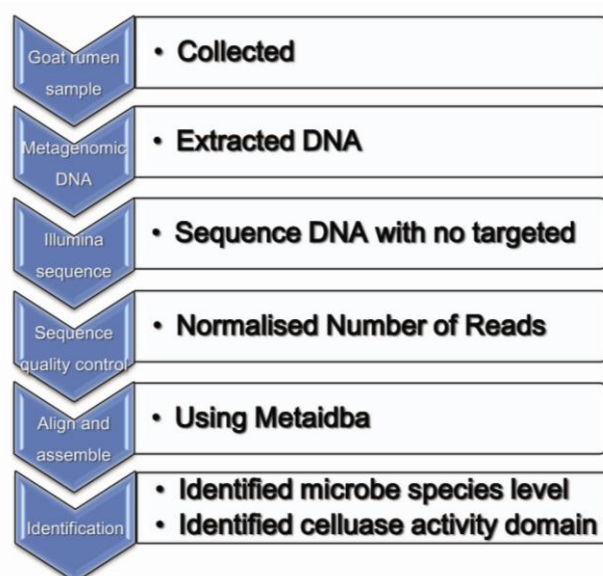
MATERIALS AND METHODS

Sampling and extraction of genomic DNA

Rumen fluid was collected from a 1-yr-old Korean native goat and Saanen hybrid raised on Timothy (*Phleumpratense*) hay at a private goat farm in the Cheonan City area and slaughtered at a local slaughter house. Rumen fluid was filtered through four layers of cheesecloth. Genomic DNA was isolated from rumen fluid using the Wizard Genomic DNA Purification Kit (Promega, US) according manufacturer's protocol. Gel electrophoresis was performed with 1% agarose gel at 50 V for 2 h to check both quality and quantity of isolated genomic DNA.

DNA shotgun paired-end library preparation

Random DNA fragmentation was performed using the Covaris S2 System, and the DNA library was prepared using TruSeqDNA Sample Prep.Kit (Illumina, US). Briefly, DNA fragments were repaired to blunt-ended DNA by fill-in and exonuclease after A-tailing was conducted to prevent the formation of adapters, dimers, and concatemers. Adaptors were ligated to genomic DNA inserts at a molar ratio of 10:1. The DNA samples were then amplified via polymerase chain reaction (PCR) using two universal



Supplementary Figure 1. Each pair read, scaffold and contig from the shotgun sequencing of goat rumen microbes.

primers. One primer contained an attachment site for the flow cell and the other contained sequencing sites for the index read. After gel electrophoresis of the PCR product, 600 to 700-bp fragments (including the insert and adapter) were selected and purified for genomic sequencing.

Genomic sequencing

Genomic DNA sequences were generated using the Illumina HiSeq2000 platform. Briefly, only library fragments with proper adapters at both ends were amplified using P5 and P7 primers on the flow cell. Clonal clusters were generated using TreSeq PE Cluster kitV3-cBot-HS (ILPE-401-3001; Illumina). Using the HiSeq2000 platform with TruSeq SBS Kit v3-HS (200 cycles; ILFC-401-3001; Illumina) 435,784, 218 reads were obtained.

Metagenomic bioinformatics application

Each pair read, scaffold, and contig of the shotgun sequencing of goat rumen microbes was summarized in Supplementary Figure 1 and Supplementary Table 1. Whole genomic DNA of collected goat rumen microbes were extracted for Illumina sequencing without DNA targeting. This shotgun sequencing generated 217,892,109 pair reads, which were filtered based on 70% identity, more than a 100-bp match, and a threshold below E^{-10} based on METAIDBA (Peng et al., 2011). These filtered 1,373,011 scaffolds were assembled and annotated to 114,031 contigs

Supplementary Table 1. Assembly and annotation statistics

Assembly	Annotation
Total pair reads	217,892,109
Scaffolds	1,373,011
Contigs	114,031

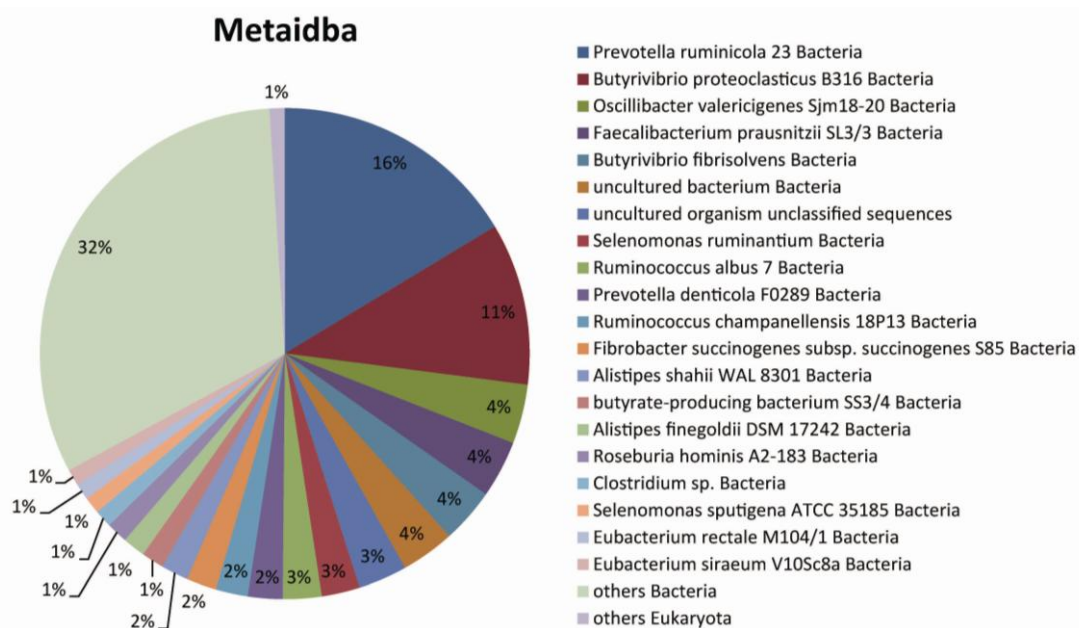


Figure 1. Analysis of the goat rumen microbial community structure at the species level.

using blastN against the NCBI nucleotide database. The domains of these 201 protein sequences were assigned to the cellulase (EC.3.2.1.4) database and translated using the HMMER package with threshold E values below 10^{-5} . Finally, these annotated genomic sequences were assigned for both identification of microbial species and cellulase-like protein domains

RESULTS AND DISCUSSION

Microbial community structure in goat rumen

The isolated genes in rumen fluid were classified into a total of 1,704 organisms, among which each 181 and 1431 ID corresponded to plant and bacteria, respectively. Using the METAIDBA metagenomic bioinformatic program, 114,031 sequences were classified into 1431 species; their population structure at the species level is graphically depicted in Figure 1. *Prevotella ruminicola* 23 bacteria and *Butyrivibrio proteoclasticus* B316 bacteria were the dominant populations, accounting for 16% and 11%, respectively.

The majority of goat rumen bacteria identified in this study have been previously reported in the rumens of cow or lamb, such as *Prevotella ruminicola* 23, *Butyrivibrio proteoclasticus* B316, and *Butyrivibrio fibrisolvens* (Bryant and Small, 1956; Van Gylswyk and Van Der Toorn, 1986; McKain et al., 1992; Moon et al., 2008). Also, some microorganisms such as butyrate-producing bacterium SS3/4 have been identified in the human colon. Previous studies have revealed the detailed rumen metabolism of *Fibrobacter succinogenes* subsp. *Fibrobacter succinogenes*

S85 and *Selenomonas ruminantium* (Heinrichova et al., 1989; Chow and Russell, 1992).

Protein domains with cellulase activity

Cellulase protein ID was obtained from the enzyme.dat file provided by the NCBI database. As a result, 201 sequences related with cellulase activity were obtained through blast searches using the NCBI BLAST program. In total, 28 protein domains with cellulase activity are summarized in Table 1. For other ruminant animals, Toyoda and coworkers analyzed the cellulose-binding proteins from sheep rumens, which consisted of endo-glucanases, proteins from fiber degrading bacterium and exo-glucanases, respectively (Toyoda et al., 2009). For cattle, constructed metagenomic library and identified 22 clones with distinct hydrolytic activities such as 12 esterases, nine endo- β -1,4-glucanases and one cyclodextrin (Ferrer et al., 2005). Considering the close correlation between rumen microbial ecology and its enzymatic functions according to the other ruminal livestock (Krause et al., 2013), list of cellulase-like protein domain list of this study can provide a clue to the characterization of Korean native goat rumen.

Profile of protein domains with cellulase activity

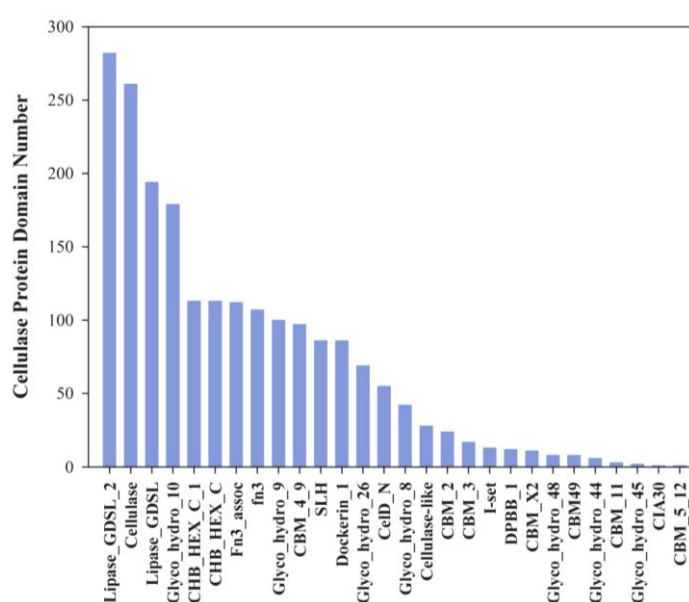
After 28 protein domains with cellulase activity were identified, the richness of each domain was analyzed (Supplementary Figure 2). Some of protein domains were overlapped to same part of sequences and also counted. The dominant bacteria had a larger number of protein domains, which suggested that strong cellulase activities were related to bacterial survival in the goat rumen. Protein domains

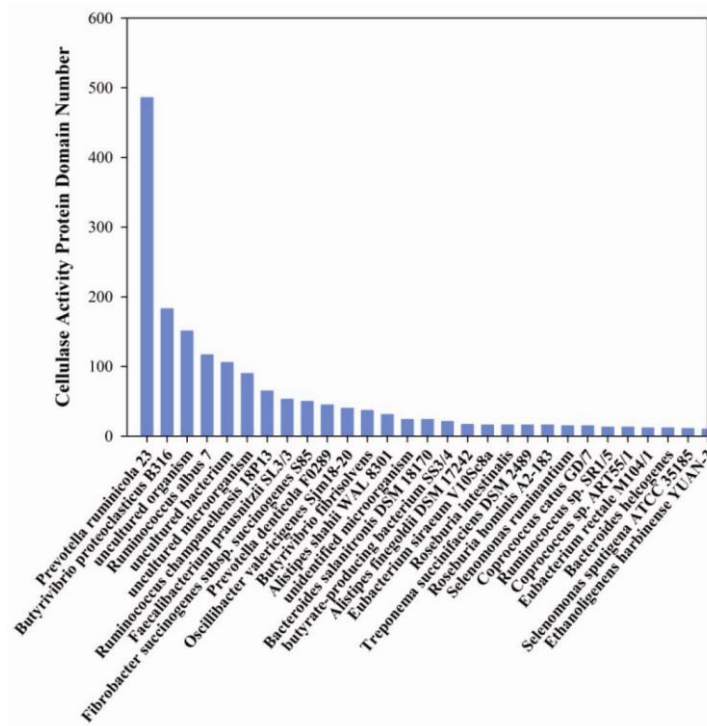
Table 1. List of protein domains with cellulase activity in goats

Domain	Accession	Superfamily	References
CBM_11	Pfam03425	cl15292	
CBM_2	Pfam 00553	cl02709	(Xu et al., 1995)
CBM_3	Pfam00942	cl03026	(Poole et al., 1992; Tormo et al., 1996)
CBM_4_9	Pfam02018	cl03406	(Johnson et al., 1996)
CBM49	Pfam09478	cl02709	(Mosbah et al., 2000)
CBM_5_12	Pfam02839	cl00046	
CBM_X2	Pfam03442	cl04075	(Mosbah et al., 2000; Kosugi et al., 2004)
eID_N	Pfam02927	cl09101	(Dominguez et al., 1996)
Cellulase	Pfam00150	cl15381	
Cellulase-like	Pfam12876	cl15381	
CHB_HEX_C	Pfam03174	cl09101	(Tews et al., 1996)
CHB_HEX_C_1	Pfam13290	cl09101	
CIA30	Pfam08547	cl15292	(Walker et al., 1992; Janssen et al., 2002)
Dockerin_1	Pfam00404	cl02860	(Shoham et al., 1999; Lytle et al., 2000)
DPBB_1	Pfam 03330	cl04011	(Takase et al., 1987; Castillo et al., 1999; Mizuguchi et al., 1999)
fn3	Pfam00041	cl00065	(Kornbliht et al., 1985; Bazan et al., 1990; Little et al., 1994)
Fn3_assoc	Pfam13287	cl09101	
Glyco_hydro_10	Pfam00331	cl01495	
Glyco_hydro_26	Pfam02156	cl09200	
Glyco_hydro_44	Pfam12891	cl15148	(Kitago et al., 2007)
Glyco_hydro_45	Pfam02015	cl03405	
Glyco_hydro_48	Pfam02011	no ref	
Glyco_hydro_8	Pfam01270	cl01351	(Alzari et al., 1996)
Glyco_hydro_9	Pfam00759	cl02959	
I-set	Pfam07679	no ref	
Lipase_GDSL	Pfam 00657	cl01053	(Upton et al., 1995)
Lipase_GDSL_2	pfam13472	cl01053	(Molgaard et al., 2000)
SLH	Pfam00395	cl02857	(Mesnage et al., 2000)

with high richness such as lipase GDSL, cellulase, and Glyco hydro 10 were also identified in the goat rumen microbes. Both lipase GDSL and lipase GDSL_2 have been reported to have molecular function of cellulose binding

(Galagan et al., 2005; Wortman et al., 2009). This is speculated as one of the main reason for its high detection. Next, the number of protein domains in each microbe was investigated (Supplementary Figure 3). Prevalent bacteria

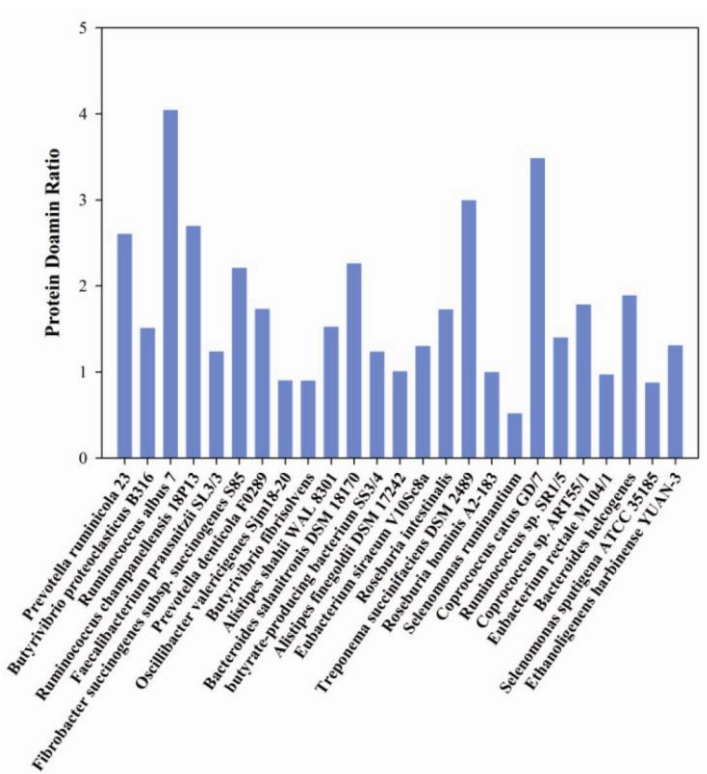
**Supplementary Figure 2.** Richness of each protein domain with cellulase activity in the goat rumen microbial population.



Supplementary Figure 3. Protein domains with cellulase activity were present in over 1% of the dominant microbe species.

such as *Prevotella ruminicola* 23 bacteria and *Butyrivibrio proteoclasticus* B316 contained a large number of cellulase protein domains, implying that these bacteria play a role in the degradation of cellulose in the goat rumen. Finally, the

protein domain ratio in each bacterial species, of which definition was the portion of cellulase-like protein domain to the assembled and annotated contigs, was analyzed to evaluate the richness of protein domain with cellulase



Supplementary Figure 4. Ratio of protein domain in each bacteria species.

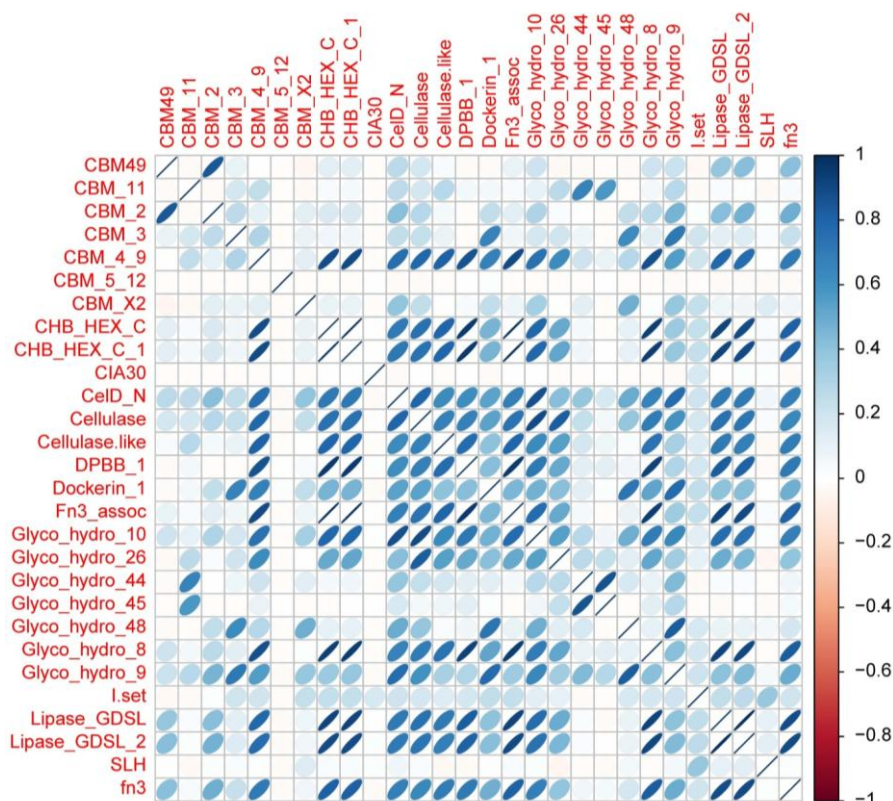


Figure 2. Correlation plot between each cellulase activity protein domain in goat rumen.

activity in the dominant bacterial species (Supplementary Figure 4). The dominant bacteria showed a ratio greater than 1, suggesting that they have high cellulase activity. A correlation plot among 28 protein domains (Figure 2) confirmed the strong positive correlation between some protein domain groups. For example, CHB_HEX_c and CHB_HEX_c-1, CHB_HEX_c and fn3 asso, and CHB_HEX_c -1 and fn3 asso had a positive correlation greater than 0.99.

Another group of lipase GDSLs, lipase GDSL_2, also showed a positive correlation greater than 0.99. To determine whether the goat rumen microbe profile was predictive of the rumen fluid metagenome profile, we correlated every rumen metagenome profile with every cellulase activity protein domain. We then determined whether the correlations were higher for samples from the same animal than for between animal samples. The results suggested that rumen fluid samples had strong correlations with each protein domain. Microbial community structure and specific protein domains with cellulase activity in the goat rumen have been identified using metagenomic analysis with both shotgun sequencing and bioinformatics. This study demonstrated that specific dominant bacterial species and protein domains have strong positive correlations, suggesting adaption to the unique feeding habits of goats.

CONCLUSIONS

In this study, microbial community structure and specific protein domains with cellulase activity in the goat rumen were identified using metagenomic analysis with both shotgun sequencing and bioinformatics. As a result, the presence of both specific dominant bacterial species such as *Prevotella ruminicola* 23, *Butyrivibrio proteoclasticus* B316, and *Butyrivibrio fibrisolvens* were identified among 1,431 bacteria in rumen fluid. At the same time, 28 protein domains with cellulase-like activity such as lipase GDSL, cellulase, and Glyco hydro 10 were identified with strong positive correlations, suggesting adaption to the unique feeding habits of goats.

ACKNOWLEDGMENT

This work was supported by a grant PJ0081912013 from Bio-Green 21 Program, Rural Development Administration, Republic of Korea.

REFERENCES

- Alzari, P. M., and R. Dominguez. 1996. The crystal structure of endoglucanase CelA, a family 8 glycosyl hydrolase from *Clostridium thermocellum*. Structure 4:265-275.

- Bazan, J. F. 1990. Structural design and molecular evolution of a cytokine receptor superfamily. *Proc. Natl. Acad. Sci. USA* 87: 6934-6938.
- Brick, D. J., M. J. Brumlik, J. T. Buckley, J.-X. Cao, P. C. Davies, S. Misra, T. J. Tranbarger, and C. Upton. 1995. A new family of lipolytic plant enzymes with members in rice, arabidopsis and maize. *FEBS Lett.* 377:475-480.
- Bryant, M. P., and N. Small. 1956. The anaerobic monotrithous butyric acid-producing curved rod-shaped bacteria of the rumen. *J. Bacteriol.* 72:16-21.
- Castillo, R. M., K. Mizuguchi, V. Dhanaraj, A. Albert, T. L. Blundell, and A. G. Murzin. 1999. A six-stranded double-psi β barrel is shared by several protein superfamilies. *Structure* 7: 227-236.
- Chow, J. M., and J. B. Russell. 1992. Effect of pH and monensin on glucose transport by *Fibrobacter succinogenes*, a cellulolytic ruminal bacterium. *Appl. Environ. Microbiol.* 58: 1115-1120.
- Dominguez, R., H. Souchon, M.-B. Lascombe, and P. M. Alzari. 1996. The crystal structure of a family 5 endoglucanase mutant in complexed and uncomplexed forms reveals an induced fit activation mechanism. *J. Mol. Biol.* 257:1042-1051.
- Ferrer, M., Olga V. Golyshina, Tatyana N. Chernikova, Amit N. Khachane, Dolores Reyes-Duarte, Vitor A. P. Martins Dos Santos, Carsten Strompl, Kieran Elborough, Graeme Jarvis, Alexander Neef, Michail M. Yakimov, Kenneth N. Timmis, and Peter N. Golyshin. 2005. Novel hydrolase diversity retrieved from a metagenome library of bovine rumen microflora. *Environ. Microbiol.* 7:1996-2010.
- Galagan, J. E., S. E. Calvo, C. Cuomo, L.-J. Ma, J. R. Wortman, S. Batzoglou, S.-I. Lee, M. Bastuerkmen, C. C. Spevak, J. Clutterbuck, V. Kapitonov, J. Jurka, C. Scazzocchio, M. L. Farman, J. Butler, S. Purcell, S. Harris, G. H. Braus, and B. W. Birren. 2005. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature*. 438:1105-1115.
- Guan, L. L., J. D. Nkrumah, J. A. Basarab, and S. S. Moore. 2008. Linkage of microbial ecology to phenotype: correlation of rumen microbial ecology to cattle's feed efficiency. *FEMS Microbiol. Lett.* 288:85-91.
- Hegarty, R. S. 1999. Reducing rumen methane emissions through elimination of rumen protozoa. *Aust. J. Agric. Res.* 50:1321-1328.
- Heinrichova, K., M. Wojciechowicz, and A. Ziolecki. 1989. The pectinolytic enzyme of *Selenomonas ruminantium*. *J. Appl. Microbiol.* 66:169-174.
- Hess, M., A. Sczyrba, R. Egan, T. W. Kim, H. Chokhawala, G. Schroth, S. Luo, D. S. Clark, F. Chen, T. Zhang, R. I. Mackie, L. A. Pennacchio, S. G. Tringe, A. Visel, T. Woyke, Z. Wang, and E. M. Rubin. 2011. Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science (New York, NY)* 331:463-467.
- Janssen, R., J. Smeitink, R. Smeets, and L. van den Heuvel. 2002. CIA30 complex I assembly factor: a candidate for human complex I deficiency? *Hum. Genet.* 110:264-270.
- Johnson, P. E., M. D. Joshi, P. Tomme, D. G. Kilburn, and L. P. McIntosh. 1996. Structure of the N-terminal cellulose-binding domain of *Cellulomonas fimi* CenC determined by nuclear magnetic resonance spectroscopy. *Biochemistry* 35:14381-14394.
- Kitago, Y., S. Karita, N. Watanabe, M. Kamiya, T. Aizawa, K. Sakka, and I. Tanaka. 2007. Crystal structure of Cel44A, a glycoside hydrolase family 44 endoglucanase from *Clostridium thermocellum*. *J. Biol. Chem.* 282:35703-35711.
- Kornblihtt, A. R., K. Umezawa, K. Vibe-Pedersen, and F. Baralle. 1985. Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene. *EMBO J.* 4:1755-1759.
- Kosugi, A., Y. Amano, K. Murashima, and R. H. Doi. 2004. Hydrophilic domains of scaffolding protein CbpA promote glycosyl hydrolase activity and localization of cellulosomes to the cell surface of *Clostridium cellulovorans*. *J. Bacteriol.* 186: 6351-6359.
- Krause, D. O., T. G. Nagaraja, A. D. G. Wright, and T. R. Callaway. 2013. Board-invited review: Rumen microbiology: Leading the way in microbial ecology. *J. Anim. Sci.* 91:331-341.
- Little, E., P. Bork, and R. F. Doolittle. 1994. Tracing the spread of fibronectin type III domains in bacterial glycohydrolases. *J. Mol. Evol.* 39:631-643.
- Lytle, B. L., B. F. Volkman, W. M. Westler, and J. Wu. 2000. Secondary structure and calcium-induced folding of the *Clostridium thermocellum* dockerin domain determined by NMR spectroscopy. *Arch. Biochem. Biophys.* 379:237-244.
- McAllister, T. A., H. D. Bae, G. A. Jones, and K.-J. Cheng. 1994. Microbial attachment and feed digestion in the rumen. *J. Anim. Sci.* 72:3004-3018.
- Mølgaard, A., S. Kauppinen, and S. Larsen. 2000. Rhamnogalacturonan acetyltransferase elucidates the structure and function of a new family of hydrolases. *Structure* 8:373-383.
- McKain, N., R. J. Wallace, and N. D. Watt. 1992. Selective isolation of bacteria with dipeptidyl aminopeptidase type I activity from the sheep rumen. *FEMS Microbiol. Lett.* 95:169-173.
- Mesnage, S., T. Fontaine, T. Mignot, M. Delepierre, M. Mock, and A. Fouet. 2000. Bacterial SLH domain proteins are non-covalently anchored to the cell surface via a conserved mechanism involving wall polysaccharide pyruvylation. *EMBO J.* 19:4473-4484.
- Mizuguchi, K., V. Dhanaraj, T. L. Blundell, and A. G. Murzin. 1999. N-ethylmaleimide-sensitive fusion protein (NSF) and CDC48 confirmed as members of the double-psi beta-barrel aspartate decarboxylase/formate dehydrogenase family. *Structure (London, England: 1993)* 7:R215-R216.
- Moon, C. D., D. M. Pacheco, W. J. Kelly, S. C. Leahy, D. Li, J. Kopečný, and G. T. Attwood. 2008. Reclassification of *Clostridium proteoclasticum* as *Butyrivibrio proteoclasticus* comb. nov., a butyrate-producing ruminal bacterium. *Intl J. Syst. Evol. Microbiol.* 58:2041-2045.
- Mosbah, A., A. Belaïch, O. Bornet, J.-P. Belaïch, B. Henrissat, and H. Darbon. 2000. Solution structure of the module X2_1 of unknown function of the cellulosomal scaffolding protein CipC of *Clostridium cellulolyticum*. *J. Mol. Biol.* 304:201-217.
- Peng, Y., Henry C. M. Leung, S. M. Yiu, and Francis Y. L. Chin. 2011. Meta-IDBA: a *de Novo* assembler for metagenomic data. *Bioinformatics* 27:i94-i101.
- Poole, D. M., E. Morag, R. Lamed, E. A. Bayer, G. P. Hazlewood, and H. J. Gilbert. 1992. Identification of the cellulose-binding

- domain of the cellulosome subunit S1 from *Clostridium thermocellum* YS. FEMS Microbiol. Lett. 99:181-186.
- Ross, E. M., P. J. Moate, C. R. Bath, S. E. Davidson, T. I. Sawbridge, K. M. Guthridge, B. G. Cocks, and B. J. Hayes. 2012. High throughput whole rumen metagenome profiling using untargeted massively parallel sequencing. BMC Genet. 13:53.
- Shoham, Y., R. Lamed, and E. A. Bayer. 1999. The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides. Trends Microbiol. 7:275-281.
- Takase, I., F. Ishino, M. Wachi, H. Kamata, M. Doi, S. Asoh, H. Matsuzawa, T. Ohta, and M. Matsuhashi. 1987. Genes encoding two lipoproteins in the leuS-dacA region of the *Escherichia coli* chromosome. J. Bacteriol. 169:5692-5699.
- Tews, I., A. Perrakis, A. Oppenheim, Z. Dauter, K. S. Wilson, and C. E. Vorgias. 1996. Bacterial chitinase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. Nat. Struct. Biol. 3:638-648.
- Tormo, J., R. Lamed, A. J. Chirino, E. Morag, E. A. Bayer, Y. Shoham, and T. A. Steitz. 1996. Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose. EMBO J. 15:5739-5751.
- Toyoda, A., W. Iio, M. Mitsumori, and H. Minato. 2009. Isolation and identification of cellulose-binding proteins from sheep rumen contents. Appl. Environ. Microbiol. 75:1667-1673.
- Van Gylswyk, N., and J. Van Der Toorn. 1986. Enumeration of *Bacteroides succinogenes* in the rumen of sheep fed maize-straw diets. FEMS Microbiol. Lett. 38:205-209.
- Walker, J. E., J. M. Arizmendi, A. Dupuis, I. M. Fearnley, M. Finel, S. M. Medd, S. J. Pilkington, M. J. Runswick, and J. M. Skehel. 1992. Sequences of 20 subunits of NADH: ubiquinone oxidoreductase from bovine heart mitochondria: Application of a novel strategy for sequencing proteins using the polymerase chain reaction. J. Mol. Biol. 226:1051-1072.
- Wortman, J. R., J. M. Gilsenan, V. Joardar, J. Deegan, J. Clutterbuck, M. R. Andersen, D. Archer, M. Bencina, G. Braus, P. Coutinho, H. von Dohren, J. Doonan, A. J. Driessen, P. Durek, E. Espeso, E. Fekete, M. Flipphi, C. G. Estrada and G. Turner. 2009. The 2008 update of the *Aspergillus nidulans* genome annotation: a community effort. Fungal Genet. Biol. 46:S2-13.
- Xu, G.-Y., E. Ong, N. R. Gilkes, D. G. Kilburn, D. Muhandiram, M. Harris-Brandts, J. P. Carver, L. E. Kay, and T. S. Harvey. 1995. Solution structure of a cellulose-binding domain from *Cellulomonas fimi* by nuclear magnetic resonance spectroscopy. Biochemistry 34:6993-7009.
- Zhou, J., B. Copeland, C. Zhang, Z. Liu, S. Bhatti, R. Sauve, and S. Zhou. 2011. Identification of prokaryotic organisms in goat rumen based on metagenomic DNA sequences. J. Res. Biol. 6: 451-455.