SHORT COMMUNICATION

SIMB

Society for Industrial Microbiology
and Biotechnology

Analysis of the bovine rumen microbiome reveals a diversity of Sus-like polysaccharide utilization loci from the bacterial phylum *Bacteroidetes*

Carly P. Rosewarne · Phillip B. Pope · Jane L. Cheung ·
Mark Morrison

Abstract Several unique Sus-like polysaccharide utilization loci (PULs) were identified from bacteria resident in bovine rumen microbiomes through functional screening of a fosmid library. The loci were phylogenetically assigned to the genus *Prevotella* within the phylum *Bacteroidetes*. These findings were augmented by a bioinformatic re-evaluation of ruminal *Prevotella* genomes, revealing additional loci not previously reported in the literature. Analysis of Bacteroidales-affiliated genomes reconstructed from a bovine rumen metagenome in a previous study further expanded the diversity of Sus-like PULs resident in this microbiome. Our findings suggest that Sus-like systems represent an important mechanism for degradation of a range of plant-derived glycans in ruminants.

C. P. Rosewarne (✉)
CSIRO Animal, Food and Health Sciences, Riverside Life
Sciences Centre, PO Box 52, North Ryde, NSW 2113, Australia
e-mail: carly.rosewarne@csiro.au

P. B. Pope
Department of Chemistry, Biotechnology and Food Science,
Norwegian University of Life Sciences, Ås, Norway

J. L. Cheung · M. Morrison
CSIRO Animal, Food and Health Sciences, Queensland
Biosciences Precinct, St Lucia, QLD 4068, Australia

Present Address:
M. Morrison
The University of Queensland Diamantina Institute,
37 Kent Street, Woolloongabba, QLD 4102, Australia

Introduction

Mammalian gut microbial communities are dominated by members of the *Firmicutes* and *Bacteroidetes* [14]. The ability of bacteria from these phyla to metabolize otherwise indigestible complex polysaccharides is important for gut function and host health. The starch utilization system (Sus) of *Bacteroides thetaiotamicron* (Phylum *Bacteroidetes*) represents a well-studied gene cluster from a human gut symbiont [19]. In this system, SusD enables cell surface starch binding, and in association with the outer membrane proteins SusE and SusF, the affinity of starch binding is enhanced [28]. The oligosaccharides released by the extracellular α-amylase (SusG) are imported into the periplasmic space by SusC, which is a member of the TonB-dependent receptor family. Further maltooligosaccharide metabolism then occurs in the periplasmic space, mediated by SusA and SusB [7].

The review by Martens et al. [18] highlights that Sus-like PULs can be found in the genomes of many sequenced *Bacteroidetes*. Characteristic features of these cell-envelope systems include proteins with homology to SusC and SusD co-located with a diversity of carbohydrate active enzymes, providing evidence for their role in targeting a broad range of substrates. While previous studies have focused mainly on the importance of Sus-like PULs in glycan catabolism in the human gut, emerging data suggests that these loci are involved in co-ordinating plant biomass degradation in herbivore microbiomes. For example, *Prevotella bryantii* $B_1 4$ upregulates expression of a xylanolytic Sus-like PUL when grown on wheat arabinoxylan [8]. Functional screening of metagenomic libraries derived from the Svalbard reindeer rumen [23] and Tammar wallaby foregut [24] revealed a diversity of carbohydrate active enzymes affiliated with the *Bacteroidetes* arranged in Sus-like PULs with

activity against polysaccharide substrates in vitro. The results presented in this manuscript build on these observations by describing the hydrolytic potential encoded on Sus-like PULs harbored by members of the *Bacteroidetes* in the rumen of cattle. We first recovered a number of loci through functional screening of an enrichment culture derived metagenomic fosmid library for clones expressing hydrolytic activity against carboxymethylcellulose. Additional loci were recovered bioinformatically from sequenced rumen isolates belonging to the genus *Prevotella*, and from Bacteroidales-affiliated genomes reconstructed from a bovine rumen metagenomic dataset.

Materials and methods

Anaerobic culture of bovine rumen samples

Rumen grab samples were collected from six fistulated *Bos indicus* steers consuming Rhodes grass (*Chloris gayana*) in Rockhampton (QLD, Australia) using protocols approved by the Rendel Laboratory Animal Experimentation and Ethics Committee. In an anaerobic chamber, rumen samples were pooled before the digesta particles were separated from the rumen fluid by squeezing through two layers of cheesecloth. Particles were resuspended in an equal volume of sterile phosphate buffered saline and homogenized with an IKA UltraTurrax® TP18/10. Homogenate (0.01 vol. of 1:10 dilution) was added to 10 ml of liquid anaerobic medium 1 [3] containing 10 % (w/v) powdered Rhodes grass in triplicate, and incubated at 39 °C with overpressure of 100 kPa H_2 until an OD_{600} of 0.5 was reached (approximately 16 h). The resulting cultures are therefore reflective of rumen microbes that can divide rapidly under these conditions. Culture material (comprised of grass substrate and microbial cells) was pelleted at $10,000 \times g$ for 10 min at 4 °C. Genomic DNA was extracted from the enrichment cultures and aliquots of the homogenated inoculum using the NML method [27].

Microbiota community profiling and fosmid library construction and screening

Amplicons of the 16S ribosomal RNA gene were prepared for 454 GS FLX Titanium sequencing as previously described (amplified from 20 ng of input DNA using primers 8F15B and 515R14AM; [1]). The data was analyzed using QIIME 1.1.0 [5], based on sequence clustering at 97 % similarity, taxonomic assignment using RDP and rarefaction to provide 4,000 sequences per sample. The remaining metagenomic DNA from the triplicate enrichment cultures was pooled and used to prepare a large insert

fosmid library (pCC1FOS; Epicentre Corp.) using methods as previously described [27]. Clones capable of carboxymethylcellulose (CMC) hydrolysis were identified using a Congo red assay [23]. A total of ~2 × 10⁴ clones were screened. Specific selected clones were subsequently screened using the same assay for their ability to degrade xylan (from birchwood), xyloglucan (from tamarind), lichenan (from Icelandic moss) and glucomannan (from konjac; low viscosity). Substrates were purchased from Sigma-Aldrich or Megazyme International, and were prepared for use according to manufacturer's instructions; before inclusion into LB agar plates at a concentration of 0.05 % (w/v).

Functional annotation and analysis of sequences from Sus-like PULs

Fosmids from individual clones of interest were induced to high copy number according to the manufacturer's protocols. Fosmid DNA was extracted using the QIAprep spin miniprep kit (Qiagen), incorporating modifications for purification of large plasmids. The positional insertion of the cloned DNA within the vector was determined by Sanger sequencing with primers EPIFOS-F and EPIFOS-R (Epicentre Corp.). Equimolar amounts of each clone (quantified using the Invitrogen Quant-IT dsDNA BR kit and Qubit fluorometer) were pooled and a 3-kb paired-end library was prepared and subjected to pyrosequencing using 454 GS FLX chemistry (Macrogen Inc., Korea). The scaffolds from the resulting data were assembled using Newbler v2.6 and annotated using IMG/M-ER [16]. ORFs in reconstructed genomes from a published metagenomic survey of the bovine rumen [11] (genomes AC2a, AGa, AJ, AH, and AQ; available at http://portal.nersc.gov/project/jgimg/CowRumenRawData/submission/) were identified using MetaGeneMark [32]. ORFs in the genome of *Prevotella bryantii* B14 and *Prevotella ruminicola* 23 were identified in a previous study [26]. Carbohydrate active enzymes (CAZymes) and carbohydrate binding modules (CBMs) were identified in protein coding sequences from these various sources using dbCAN [31], a web resource that implements hidden Markov models (HMMs) for automated signature domain-based annotations representative of each individual family. Sus-like PULs were defined as operons encoding proteins with Pfam domains [25] belonging to the TonB-dependent receptor (PF00593) and SusD-like families (PF12771, PF14721, or PF14322), collocated with proteins encoding known CAZyme domains. For both the dbCAN and Pfam HMM assignments, a minimum e-value cut-off of 1×10^{-5} was used. All putative Sus-like PULs that were recovered from the in silico analysis have been described in this manuscript.

Nucleotide accession numbers

Annotated sequences from five fosmid scaffolds have been deposited in GenBank under accession numbers JX424616-28.

Results and discussion

Bacterial composition of fiber-adherent bovine rumen microbiome and cultures enriched on Rhodes grass

Amplicon pyrosequencing of 16S rRNA genes revealed differences in community structure resulting from in vitro cultivation of the fiber-adherent bovine rumen microbiome. The material used to inoculate the enrichments was comprised predominantly of sequences from the phyla *Bacteroidetes* (47.2 %) and *Firmicutes* (42.3 %). The *Bacteroidetes* fraction was dominated by sequences from the order Bacteroidales (71 % of *Bacteroidetes*; 33.5 % of total), while the *Firmicutes* fraction was dominated by sequences from the order Clostridiales (91 % of *Firmicutes*; 38.6 % of total). Much of this diversity could not be classified at the genus level. These results are comparable to other studies that have profiled the fiber-adherent fraction of the rumen microbiome from cattle consuming high-forage diets [4, 11, 22]. The cultures were dominated by representatives of recognized genera, including *Prevotella* (31.4 %), *Selenomonas* (19.1 %), *Psuedobutyrivibrio* (11.1 %), *Streptococcus* (8.4 %), and *Fibrobacter* (6.6 %). These organisms have been enriched in this experiment because they are more amenable to in vitro growth under the specific culture conditions used. Sequences affiliated to the genus *Prevotella* were distributed into 16 operational taxonomic units (OTUs). Ten of these OTUs contained over 95 % of the total *Prevotella* sequences in the enrichment dataset. Representative sequences from these 10 OTUs were compared to data from cultured isolates in Ribosomal Database Project Release 10.32 ([6]; Table S1). This indicated that the two predominant *Prevotella* OTUs (approximately 55 % of total *Prevotella* sequences) were closely related to *P. ruminicola* 23 and *P. bryantii* $B_1$4, two rumen species for which genome sequences are available [26]. The remaining OTUs were more similar to unsequenced isolates, suggesting that metagenomic DNA extracted from the cultures contains material from both sequenced and "novel" *Prevotella* strains.

Recovery of Sus-like PULs from fosmid library sequencing and the genomes of ruminal *Prevotella* isolates

Metagenomic DNA extracted from the enrichment cultures was pooled and used to construct a large insert fosmid library. A total of 142 clones expressing hydrolytic activity towards carboxymethylcellulose were recovered from a screen of 2×10^4 clones. It is probable that additional loci capable of CMC hydrolysis are present in the enrichment cultures but were not recovered due to biases associated with functional screening of fosmid clones in *E. coli*. Pooled DNA from each of the fosmid clones was sequenced and assembled into scaffolds. Mapping of insertion sites (based on Sanger sequencing with vector specific primers) indicated that several scaffolds were assembled from multiple fosmid clones (data not shown). The closest match to each scaffold larger than 10 kb ($n = 51$) was determined using Standard Nucleotide BLAST (BlastN) of the NCBI nr database [12]. Approximately 41 % of scaffolds were most closely affiliated to the genus *Prevotella* ($n = 21$), followed by *Fibrobacter* (29 %; $n = 15$) and *Streptococcus* (6 %; $n = 3$). The *Fibrobacter* derived scaffolds have a high degree of similarity to regions within the genome of *F. succinogenes* S85 [29] and were not analyzed further (data not shown).

A total of ten scaffolds encoding Sus-like PULs were recovered from the fosmid library. Each PUL encodes a putative SusC-like protein exhibiting homology to the TonB-dependent receptor family (PF07715) and a protein with homology to a SusD-like family (PF12771, PF14721, or PF14322). PhylopythiaS [21] was used to establish that each of these scaffolds was derived from a *Prevotella* genome. Insertion site mapping identified five scaffolds (designated Sc00026, 28, 33, 44 and 66) where the complete Sus-like PUL region could be accurately assigned to at least one fosmid clone (Fig. S1). For the remaining five scaffolds, both sites of insertion into the vector could not be mapped to an individual clone and as a result were excluded from further analysis. The genomic architecture of these loci is described in more detail in Fig. S2.

Concomitant with the diversity of predicted carbohydrate active enzymes (CAZymes [31]; Fig. 1), functional activity screening (Table 1) indicated that the five Sus-like PULs selected for further analysis have different substrate ranges. The substrates for this screen were chosen based on the CAZyme profiles of the fosmid clones. The Sc00026 PUL targets both carboxymethylcellulose and glucomannans. This scaffold encodes a 16S rRNA gene with 99 % sequence identity to *P. bryantii* $B_1$4. Comparison of the PUL-encoding contig from Sc00026 to the genome of *P. bryantii* $B_1$4 revealed a high level of similarity to a region of $B_1$4 previously shown to encode β-1,4-endoglucanase (GH5) and mannanase (GH26) activities ([10]; GenBank accession U96771 and Table S2). Association of the $B_1$4 glycoside hydrolases with a Sus-like PUL has not previously been reported, although the β-1,4-endoglucanase activity has been extensively studied. The $B_1$4 locus is unusual in that two β-1,4-endoglucanases with different
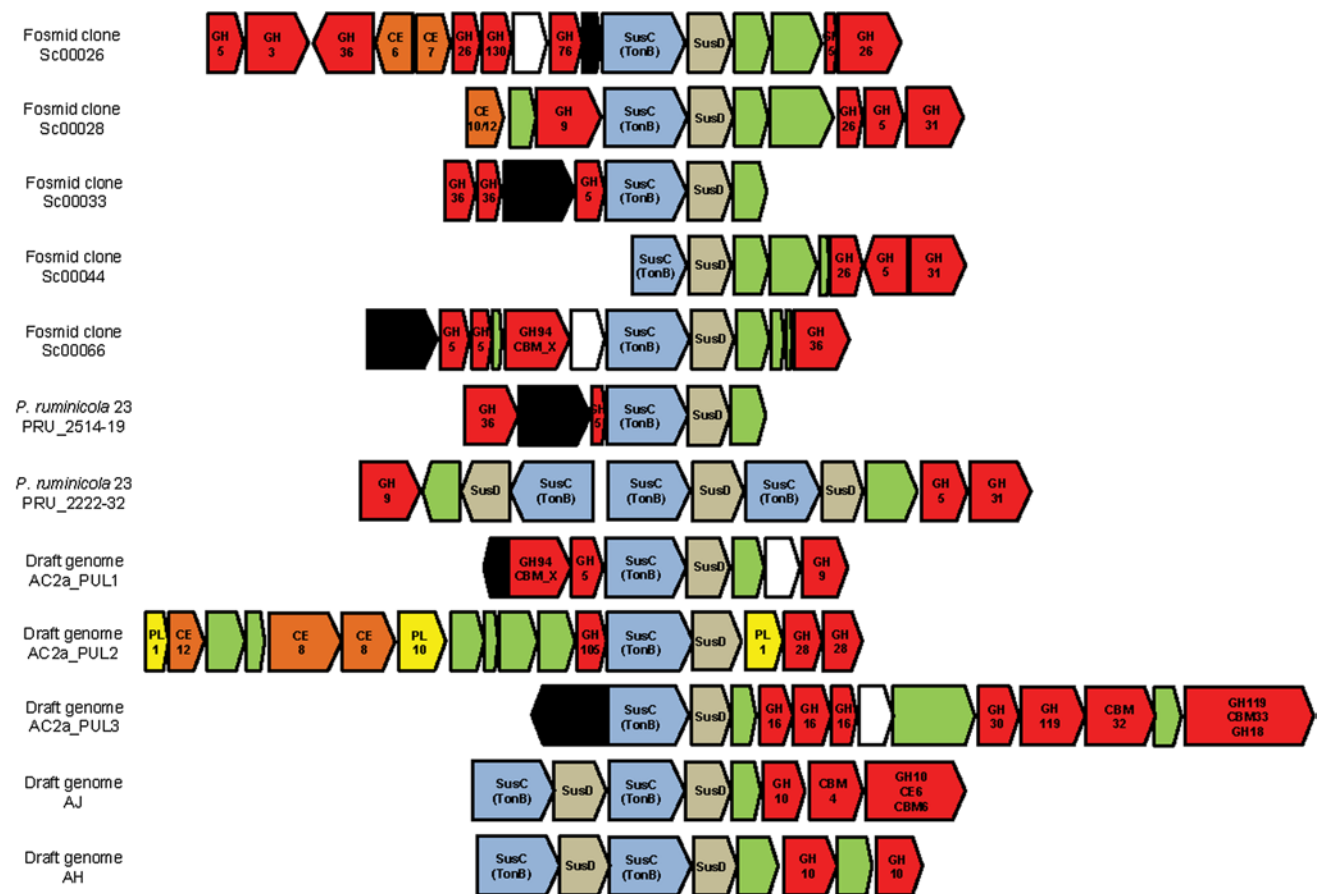
Fig. 1 Gene organization of selected PULs from bacteria resident in the bovine rumen. Loci were identified from functional screening of a fosmid library derived from anaerobic cultures enriched on Rhodes grass (this study); in the genome of *Prevotella ruminicola* 23 [26]; or from re-analysis of draft genomes reconstructed from a metagenomic survey [11]. *Green* genes represent putative outer membrane proteins of unknown function; *black* genes encode putative response regulators; *white* genes encode putative transmembrane proteins; TonB indicates members of the TonB-dependent receptor family of proteins that are predicted to transport solutes and macromolecules. *GH* glycoside hydrolase (*red*), *CBM* carbohydrate binding module, *CE* carbohydrate esterase (*orange*), *PL* polysaccharide lyase (*yellow*) (color figure online)

Table 1 Activity screening of fosmid clones on selected polysaccharide substrates

	Sc00026 (H36 A20)	Sc00028 (H39 N13)	Sc00033 (H6 M11)	Sc00044 (H30 F13)	Sc00066 (H13 J8)	Control (no fosmid)
Carboxymethylcellulose	+	+	+	+	+	−
Xylan	−	−	−	−	−	−
Xyloglucan	−	+	+	+	−	−
Lichenan	−	+	−	+	−	−
Glucomannan	+	−	−	−	−	−

Clones encompassing the entire Sus-like PUL (as shown in Fig. S1) were chosen for this assay as indicated in brackets below the scaffold ID. Substrates were incorporated into LB agar at a concentration of 0.05 % (w/v). Zones of clearing were visualized using a Congo red assay and recorded as a positive or negative result. The *E. coli* EPI300™-T1^R strain from the fosmid library kit was included as a negative control

molecular weights are transcribed from two adjacent open reading frames with a −1 frameshift [10, 20]. The Sc00026 PUL appears to have a similar structure, with the ORFs shown on the right of this locus in Fig. 1 potentially also encoding two endoglucanases that share a common N-terminal domain.

The Sus-like PUL encoded by Sc00066 only expresses activity against carboxymethylcellulose (Table 1). This

locus encodes two proteins with GH5 domains and one with a GH94 domain, indicative of cellobiosidase and cellobiose phosphorylase activities respectively. Accordingly, GH5a from Sc00066 is homologous (BlastX [12]; 75 % identity over 255 residues) to M40-2 (ACA61171), a protein cloned from buffalo rumen with degradative activity against *p*-nitrophenyl-D-cellobioside [9].

The PULs encoded on scaffolds Sc00028, 33 and 44 are capable of targeting carboxymethylcellulose and xyloglucans; while the Sc00028 and Sc00044 PULs are also able to degrade lichenan. According to the classification scheme proposed by Aspeborg et al. [2], the GH5 enzymes encoded by these PULs belong to subfamily 4. Members of this subfamily with activities against xyloglucan and lichenan have been reported previously [2]. There are two Sus-like PULs with GH5 family proteins in the genome of *P. ruminicola* 23 (Fig. 1). Comparison of this genome to the Sus-like PUL-encoding contigs of Sc00028, 33, 44 and 66 using BlastN [12] revealed 50–90 % coverage with 73–85 % sequence identity (Table S2). The PUL encoded by genes PRU_2514-19 is most similar to that found on Sc00033, while the locus at PRU_2222-32 includes three SusC and SusD-like proteins, a structure not found in any of the fosmid clones. A PUL with CMCase activity that encodes multiple Sus-like proteins was cloned from Svalbard reindeer rumen in a previous study [23]. Functional characterization of the two SusD-like proteins in this PUL indicated that different mechanisms are used to interact with various forms of cellulose, and that they are able to bind to distinct features in the plant cell wall microstructure [15]. An increasing body of evidence suggests that Sus-like PULs may be capable of cellulose degradation in herbivores, although additional targeted experiments are required in order to confirm this hypothesis.

Sus-like PULs recovered from a published bovine rumen metagenome

As part of their analysis into the structural and functional diversity of the fiber-adherent bovine rumen microbiome, Hess et al. [11] describe reconstruction of 15 bacterial genomes from a substantial metagenomic dataset. Five of these reconstructed genomes (AC2a, AH, AJ, AGa and AQ) are from deeply branching members of the order Bacteroidales. Searches of these genomes were undertaken to recover additional Sus-like PULs. Five of these loci have been included in Fig. 1. Three Sus-like PULs from the genome of AC2a are shown—the cellulase encoding PUL (AC2a_PUL1) has been described previously [11, 23]. The remaining two PULs from AC2a are predicted to target pectin (AC2a_PUL2) and

β-glucan (AC2a_PUL3). The draft genomes of AJ and AH contain PULs with two presumptive GH10 endoxylanases, suggesting a role in the degradation of hemicellulose similar to that found in *P. bryantii* B$_1$4 and other related bacteria [8].

A further 36 Sus-like PULs were uncovered in *P. bryantii* B$_1$4, *P. ruminicola* 23 and three of the five reconstructed rumen Bacteroidales genomes (AC2a, AH and AJ). Several additional SusC and SusD homologs were annotated in the genome sequences, but were not associated with recognized carbohydrate active enzymes in the dbCAN database (data not shown). The putative PULs (Table S3) are predicted to target a diverse array of glycans, expanding the potential substrate range of such systems in the bovine rumen. Several loci described in Table S3 (e.g., PBR_0326 to PBR_0345, PBR_0377 to PBR_0398, PRU_2666 to PRU_2716 and AJ_902 to AJ_918) are predicted to encode a large number of ORFs and include multiple Sus-like systems. Clustering of polysaccharide degradation genes is a strategy used by many bacteria to co-ordinate enzyme production and metabolic activities [13, 17]. Additional studies are required to confirm substrate specificity and to determine how the Sus-like PULs described in this study are regulated and expressed in response to the presence of different complex carbohydrates.

Conclusions

This paper describes more than 50 distinct Sus-like PULs from the bovine rumen microbiome, the most comprehensive collection described in herbivores to date. Each locus has been phylogenetically assigned to the predominant gut bacterial phylum *Bacteroidetes*. In *Prevotella ruminicola* 23, genes that were putatively associated with Sus-like PULs represent approximately 10 % of the total protein coding sequences in the complete genome. Association of a high proportion of the total genomic CAZyme repertoire with Sus-like genes has been shown previously for human gut *Bacteroides* isolates [17]. For many members of the *Bacteroidetes*, Sus-like PULs represent an important mechanism for metabolism of a diverse range of plant and host-derived polysaccharides. The diversity found within loci as described in this manuscript and others [23, 24, 30], in concert with previously observed differences in the substrate binding properties of SusD-like proteins [15] suggests that there is much to learn about the contribution of these cell-envelope-associated enzymatic complexes to plant biomass degradation in herbivores. This has implications for improving ruminant feed efficiency and discovery of novel systems for deconstruction of plant-based biofuel substrates.

References

1. Aguirre de Carcer D, Denman SE, McSweeney CS, Morrison M (2011) Strategy for modular tagged high-throughput amplicon sequencing. Appl Environ Microbiol 77:6310–6312

2. Aspeborg H, Coutinho PM, Wang Y, Brumer H 3rd, Henrissat B (2012) Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). BMC Evol Biol 12:186

3. Balch WE, Fox GE, Magrum LJ, Woese CR, Wolfe RS (1979) Methanogens: reevaluation of a unique biological group. Microbiol Rev 43:260–296

4. Brulc JM, Antonopoulos DA, Berg Miller ME, Wilson MK, Yannarell AC, Dinsdale EA et al (2009) Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. Proc Natl Acad Sci USA 106:1948–1953

5. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK et al (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336

6. Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ et al (2009) The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. Nucl Acids Res 37:D141–D145

7. D'Elia JN, Salyers AA (1996) Contribution of a neopullulanase, a pullulanase, and an alpha-glucosidase to growth of *Bacteroides thetaiotaomicron* on starch. J Bacteriol 178:7173–7179

8. Dodd D, Moon Y-H, Swaminathan K, Mackie RI, Cann IKO (2010) Transcriptomic analyses of xylan degradation by *Prevotella bryantii* and insights into energy acquisition by xylanolytic Bacteroidetes. J Biol Chem 285:30261–30273

9. Duan CJ, Xian L, Zhao GC, Feng Y, Pang H, Bai XL et al (2009) Isolation and partial characterisation of novel genes encoding acidic cellulases from metagenomes of buffalo rumens. J Appl Microbiol 107:245–256

10. Gardner RG, Wells JE, Fields MW, Wilson DB, Russell JB (1997) A *Prevotella ruminicola* B$_1$4 operon encoding extracellular polysaccharide hydrolases. Curr Microbiol 35:274–277

11. Hess M, Sczyrba A, Egan R, Kim T-W, Chokhawala H, Schroth G et al (2011) Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. Science 331:463–467

12. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL (2008) NCBI BLAST: a better web interface. Nucl Acids Res 36:W5–W9

13. Kelly W, Leahy SC, Altermann A, Yeoman CJ, Dunne JC, Kong Z et al (2010) The glycobiome of the rumen bacterium *Butyrivibrio proteoclasticus* B316^T highlights adaptation to a polysaccharide-rich environment. PLoS ONE 5:e11942

14. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS et al (2008) Evolution of mammals and their gut microbes. Science 320:1647–1651

15. Mackenzie AK, Pope PB, Pedersen HL, Gupta R, Morrison M, Willats WGT, Eijsink VGH (2012) Two SusD-like proteins encoded within a polysaccharide utilization locus of an uncultured ruminant *Bacteroidetes* bind strongly to cellulose. Appl Environ Microbiol 78:5935–5937

16. Markowitz VM, Chen IMA, Chu K, Szeto E, Palaniappan K, Grechkin Y et al (2012) IMG/M: the integrated metagenome data management and comparative analysis system. Nucl Acids Res 40:D123–D129

17. Martens EC, Lowe EC, Chiang H, Pudlo NA, Wu M, McNulty NP et al (2011) Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. PLoS Biol 9:e1001221

18. Martens EC, Koropatkin NM, Smith TJ, Gordon JI (2009) Complex glycan catabolism by the human gut microbiota: the *Bacteroidetes* Sus-like paradigm. J Biol Chem 284:24673–24677

19. Martens EC, Chiang HC, Gordon JI (2008) Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. Cell Host Microbe 4:447–457

20. Matsushita O, Russell JB, Wilson DB (1991) A bacteroides ruminicola 1,4-β-D-endoglucanase is encoded in two reading frames. J Bacteriol 173:6919–6926

21. Patil KR, Haider P, Pope PB, Turnbaugh PJ, Morrison M, Scheffer T, McHardy AC (2011) Taxonomic metagenome sequence assignment with structured output models. Nat Methods 8:191–192

22. Pitta D, Pinchak W, Dowd S, Osterstock J, Gontcharova V, Youn E et al (2010) Rumen bacterial diversity dynamics associated with changing from bermudagrass hay to grazed winter wheat diets. Microb Ecol 59:511–522

23. Pope PB, Mackenzie AK, Gregor I, Smith W, Sundset MA, McHardy AC et al (2012) Metagenomics of the Svalbard reindeer rumen microbiome reveals abundance of polysaccharide utilization loci. PLoS ONE 7:e38571

24. Pope PB, Denman SE, Jones M, Tringe SG, Barry K, Malfatti SA et al (2010) Adaptation to herbivory by the Tammar wallaby includes bacterial and glycoside hydrolase profiles different from other herbivores. Proc Natl Acad Sci USA 107:14793–14798

25. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C et al (2012) The Pfam protein families database. Nucl Acids Res 40:D290–D301

26. Purushe J, Fouts D, Morrison M, White B, Mackie RI, Coutinho P et al (2010) Comparative genome analysis of *Prevotella ruminicola* and *Prevotella bryantii*: insights into their environmental niche. Microb Ecol 60:721–729

27. Rosewarne CP, Pope PB, Denman SE, McSweeney CS, O'Cuiv P, Morrison M (2011) High-yield and phylogenetically robust methods of DNA recovery for analysis of microbial biofilms adherent to plant biomass in the herbivore gut. Microb Ecol 61:448–454

28. Shipman JA, Berleman JE, Salyers AA (2000) Characterization of four outer membrane proteins involved in binding starch to the cell surface of *Bacteroides thetaiotaomicron*. J Bacteriol 182:5365–5372

29. Suen G, Weimer PJ, Stevenson DM, Aylward FO, Boyum J, Deneke J et al (2011) The complete genome sequence of *Fibrobacter succinogenes* S85 reveals a cellulolytic and metabolic specialist. PLoS ONE 6:e18814

30. Wang L, Hatem A, Catalyurek UV, Morrison M, Zu Y (2013) Metagenomic insights into the carbohydrate active enzymes carried by the microorganisms adhering to solid digesta in the rumen of cows. PLoS ONE 8:e78507

31. Yin Y, Mao X, Yang JC, Chen X, Mao F, Xu Y (2012) dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucl Acids Res 40:W445–W451

32. Zhu W, Lomsadze A, Borodovsky M (2010) Ab initio gene identification in metagenomic sequences. Nucl Acids Res 38:e132