AEM
Journals.ASM.org

A Polysaccharide Utilization Locus from an Uncultured *Bacteroidetes* Phylotype Suggests Ecological Adaptation and Substrate Versatility

A. K. Mackenzie,[a] A. E. Naas,[a] S. K. Kracun,[b] J. Schückel,[b] J. U. Fangel,[b] J. W. Agger,[a] W. G. T. Willats,[b] V. G. H. Eijsink,[a] P. B. Pope[a]

Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås, Norway[a]; Department of Plant Biology and Biotechnology, University of Copenhagen, Copenhagen, Denmark[b]

Recent metagenomic analyses have identified uncultured bacteria that are abundant in the rumen of herbivores and that possess putative biomass-converting enzyme systems. Here we investigate the saccharolytic capabilities of a polysaccharide utilization locus (PUL) that has been reconstructed from an uncultured *Bacteroidetes* phylotype (SRM-1) that dominates the rumen microbiome of Arctic reindeer. Characterization of the three PUL-encoded outer membrane glycoside hydrolases was performed using chromogenic substrates for initial screening, followed by detailed analyses of products generated from selected substrates, using high-pressure anion-exchange chromatography with electrochemical detection. Two glycoside hydrolase family 5 (GH5) endoglucanases (GH5_g and GH5_h) demonstrated activity against β-glucans, xylans, and xyloglucan, whereas GH5_h and the third enzyme, GH26_i, were active on several mannan substrates. Synergy experiments examining different combinations of the three enzymes demonstrated limited activity enhancement on individual substrates. Binding analysis of a SusE-positioned lipoprotein revealed an affinity toward β-glucans and, to a lesser extent, mannan, but unlike the two SusD-like lipoproteins previously characterized from the same PUL, binding to cellulose was not observed. Overall, these activities and binding specificities correlated well with the glycan content of the reindeer rumen, which was determined using comprehensive microarray polymer profiling and showed an abundance of various hemicellulose glycans. The substrate versatility of this single PUL putatively expands our perceptions regarding PUL machineries, which so far have demonstrated gene organization that suggests one cognate PUL for each substrate type. The presence of a PUL that possesses saccharolytic activity against a mixture of abundantly available polysaccharides supports the dominance of SRM-1 in the Svalbard reindeer rumen microbiome.

Microbes belonging to the *Bacteroidetes* phylum are numerically dominant in the gut of mammals, where they contribute to plant cell wall degradation. Our current understanding regarding their enzymatic mechanisms, regulated by what are otherwise known as polysaccharide utilization loci (PULs), has largely been developed from in-depth analysis of cultivated *Bacteroidetes* strains within the human distal gut (1–3) and a single *Prevotella* rumen isolate (4). PULs have been defined as gene clusters that encode cell envelope-associated enzyme systems that enable the bacterium to respond to, bind, and degrade specific glycans and import the released oligosaccharides (2). The original PUL was described as a starch utilization system (Sus) and was built around two core proteins that are anchored on the outer membrane (SusC/D). The SusC proteins are a group of outer membrane-spanning proteins that can import solutes and macromolecules into the periplasm (5, 6), whereas the SusD protein coordinates polysaccharide binding at the cell surface (7). The remaining genes often encoded by PULs include (and are not limited to) other outer membrane-binding proteins (SusE and SusF), inner membrane-bound sugar transporters, regulatory proteins, and putative glycoside hydrolases (GHs) that dictate which target carbohydrates the PUL can hydrolyze. To date there exists an already impressive catalogue of PUL target substrates, including starch, alginate, various hemicelluloses, and host mucin glycans (1, 4, 8–10).

Metagenomic studies of lignocellulose-degrading rumen microbiomes have frequently detected a numerical abundance of uncultured *Bacteroidetes* species (11, 12) as well as an uncharacterized inventory of putative PUL-encoded carbohydrate-active enzymes within several reconstructed genomes (12–15). For example, using metagenomic binning methods (PhyloPythiaS) (16), we recently reconstructed a draft genome representing a dominant and novel *Bacteroidetes* phylotype (SRM-1) that constituted roughly 11% of the total microbial community resident in the rumen of Svalbard reindeer, a wild herbivore that exclusively eats arctic vegetation (12). Further annotation revealed a 30-kb PUL that encodes at least 13 genes (genes a to m), including two consecutive pairs of SusC/D homologues and seven carbohydrate-active enzymes with putative activities against beta-(1,4)-glucans (GH5, GH94), xylans (GH5, GH5-CE7), and mannans (GH26, GH130, GH2) (12). While previous binding studies of the two SusD-like genes in this cluster demonstrated binding to crystalline cellulose (17), the predicted activity of the associated outer membrane GHs (two GH5 enzymes and GH26) would suggest that the PUL is targeted to various hemicellulose substrates, such as mannans and soluble β-glucans. The abundance of this uncultured phylotype and the perceived functional inconsistencies of the PUL-associated lipoproteins and enzymes made this particular PUL a target conducive to detailed characterization. Therefore, we

Received 19 September 2014 **Accepted** 14 October 2014

Accepted manuscript posted online 17 October 2014

Citation Mackenzie AK, Naas AE, Kracun SK, Schückel J, Fangel JU, Agger JW, Willats WGT, Eijsink VGH, Pope PB. 2015. A polysaccharide utilization locus from an uncultured *Bacteroidetes* phylotype suggests ecological adaptation and substrate versatility. Appl Environ Microbiol 81:187–195. doi:10.1128/AEM.02858-14.

Editor: H. L. Drake

Address correspondence to Phillip B. Pope, phil.pope@nmbu.no.

doi:10.1128/AEM.02858-14

TABLE 1 Primers used to clone SRM-1 proteins

Primer name	Sequence	GenBank accession no.
GH5_g_NT_LIC	TTAAGAAGGAGATATACTATGCAGAAGCCCGAACCGGCCGCCGAGCCGGCT	KM887864
GH5_g_CT_LIC	AATGGTGGTGATGATGGTGCGCTTTCCAGTTGTCCCAATAGGCCTGCAGGAT	KM887864
GH5_h_NT_LIC	TTAAGAAGGAGATATACTATGACGCCCGTCGCCCCGAATGACACGCCGGAC	KM887865
GH5_h_CT_LIC	AATGGTGGTGATGATGGTGCGCATATTTCGGTGAATGGTCATAAACGGATTG	KM887865
GH26_i_NT_LIC	TTAAGAAGGAGATATACTATGGCCCGGCAGCCCTCCACCCCCGAGGATCCC	KM887866
GH26_i_CT_LIC	AATGGTGGTGATGATGGTGCGCCAGTACGCTGTTGGGGTTGTTGTCCGTCCA	KM887866
SusE_f_NT_LIC	TTAAGAAGGAGATATACTATGATCCCGCAGGCAGATAACCAGCTGGCGGAC	KM887863
SusE_f_CT_LIC	AATGGTGGTGATGATGGTGCGCCTCGATGTAGTAGAGTTTCAGCGGGGTGTA	KM887863

have biochemically characterized the three outer membrane enzymes (GH5 and GH26) as well as the SusE-positioned lipoprotein to complement the earlier work on the two SusD-like lipoproteins (17). To aid in interpreting the implications of the functional characterization of the SRM-1 PUL, we also analyzed the ecosystem, i.e., the reindeer rumen, from which the SRM-1 phylotype was retrieved, for the presence of target plant polysaccharide substrates.

MATERIALS AND METHODS

Comprehensive microarray polymer profiling. Since SRM-1 was recovered from a wild animal ingesting arctic vegetation, cell wall glycan epitopes were analyzed using a previously described high-throughput semiquantitative microarray-based method (18, 19). Rumen contents were analyzed from the same Svalbard reindeer samples which were used to generate the original metagenome and SRM-1 draft genome (12). These samples were collected in late winter (31 January 2010) from two wild adult females grazing on their natural winter pastures (20) in Bjørndalen, near Longyearbyen, Svalbard, Norway. The reindeer were euthanized, and approximately 50 g of whole rumen contents (solid and liquid phases) was immediately transferred to sterile containers and frozen at −80°C (12). The pH of the Svalbard reindeer rumen has been previously determined to be approximately 6.75 in winter (21). As before, the two samples were mixed (vortexed) and pooled (equal weight), and cell wall glycans were sequentially extracted from approximately 50 mg of whole reindeer rumen contents (4 replicates) using 50 mM diaminocyclohexane-tetraacetic acid (CDTA; pH 7.5) and 4 M NaOH with 1% (vol/vol) $NaBH_4$, which are known to solubilize pectins and noncellulosic polysaccharides, respectively. For each extraction, 300 μl of solvent was added to each tube, followed by incubation at room temperature with shaking for 1 h. After centrifugation at 2,500 × *g* for 10 min, the supernatants were retained and left undiluted or diluted 5- and 25-fold in phosphate-buffered saline (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.7 mM KH_2PO_4, pH 7.5), and the undiluted and diluted samples were printed in sextuplet onto nitrocellulose membranes. Each replicate was therefore represented by an 18-spot subarray (three concentrations and six printing replicates). Arrays were probed with monoclonal antibodies (MAbs) or carbohydrate-binding modules (CBMs) (18), scanned, and uploaded into ImaGene (version 6.0) microarray analysis software. The maximal mean spot signal was set to 100%, and all other values within that data set were adjusted accordingly. A mean spot signal minimum was set at 5%.

Heterologous expression and purification of enzymes. Genes encoding signal peptide-free versions of GH5_g, GH5_h, GH26_i, and SusE-

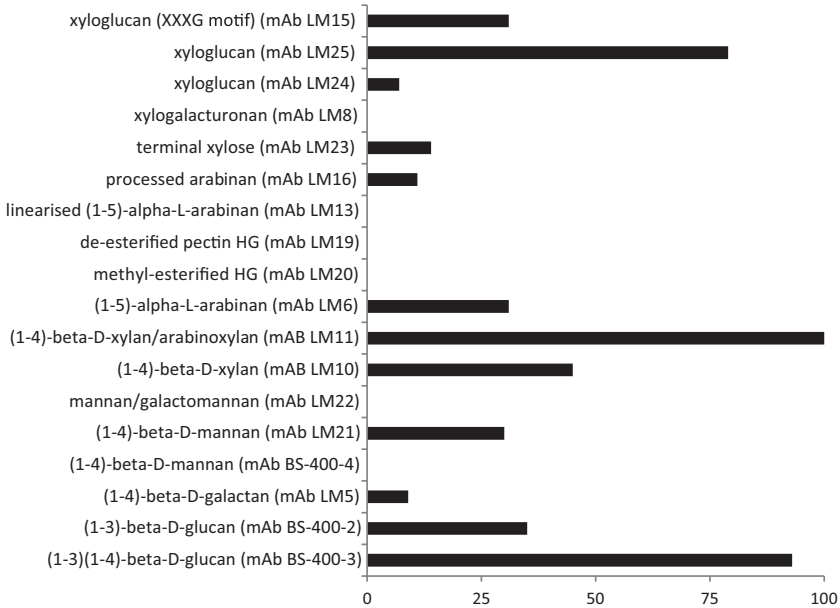


FIG 1 Occurrence of cell wall hemicelluloses in the Svalbard reindeer rumen. The bar graph shows the relative abundance of hemicelluloses recognized by various probes in whole reindeer rumen contents; the bar length is proportional to the mean spot signal. Numerical values were generated from the mean spot signals derived from four extraction replicates and six printing replicates of each of three sample concentrations. Glycans were sequentially extracted using NaOH and probed with a range of monoclonal antibodies (MAbs), as indicated. The highest mean spot signal in the data set was assigned a value of 100%, and the values for all other signals were adjusted accordingly. A signal minimum of 5% was imposed. HG, homogalacturonan.

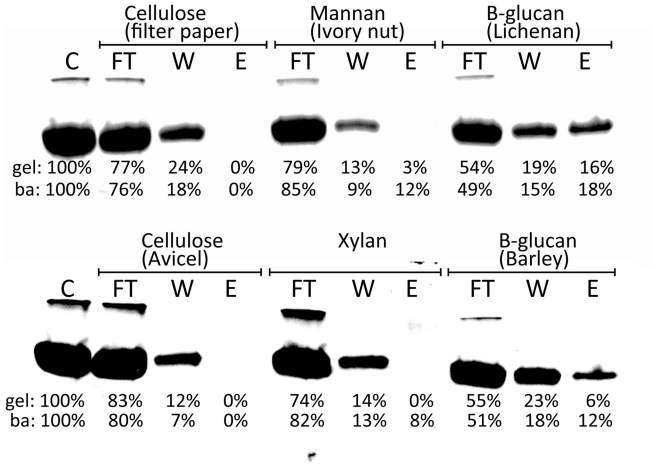


FIG 2 Binding of SRM-1 SusE-positioned lipoproteins to plant polysaccharides. SDS-PAGE analysis of fractions from pulldown assays using cellulose (filter paper and Avicel crystalline cellulose), mannan (ivory nut), xylan, and β-glucan (lichenan and barley). Lanes: C, control protein loaded without substrate; FT, unbound protein (ca. 68 kDa) from supernatant fractions collected after 1 h incubation and centrifugation; W, wash fraction containing protein washed off the substrate; E, eluted protein fractions where protein was released from the polysaccharides by incubation with urea. The original concentration of SusE-positioned lipoproteins used for binding assessments was loaded on the gel as a marker reference (1 mg/ml). Protein recovery after each stage is indicated as a percentage calculated against the control using the intensity of the gel band (gel) or the Bradford assay (ba).

positioned proteins were cloned from a fosmid clone derived from the environmental DNA of the Svalbard reindeer rumen microbiome (12). Genes without N-terminal signal sequences were cloned into the pNIC-CH expression vector (22) by ligation-independent cloning (LIC) using the primers listed in Table 1 (22). Transformants were verified by sequencing. *Escherichia coli* BL21 harboring the respective plasmids was precultured for 8 h in Luria broth at 37°C and inoculated to 0.125% in an overnight culture at 18°C. Expression was induced by the addition of IPTG (isopropyl-β-D-thiogalactopyranoside) to a final concentration of 0.75 mM at an optical density at 600 nm of 0.5 to 1.0, followed by incubation for 24 h at 18°C. Cells were harvested by centrifugation (4,500 × *g*, 10 min) and frozen at −20°C. Frozen pellets were resuspended in lysis buffer (100 mM Tris-HCl, pH 8.5, 500 mM NaCl, 5 mM imidazole, 0.1 mg/ml lysozyme), and the cells were disrupted by pulsed sonication. Cell debris was removed by centrifugation (10,000 × *g*, 10 min), and soluble proteins were purified using ion-exchange chromatography on a 5-ml HiTrap DEAE FF column (GE Healthcare, Little Chalfont, United Kingdom). The recombinant His-tagged proteins were eluted using a linear NaCl gradient (from 0 to 1 M NaCl in 50 mM Tris-HCl, pH 8.5), and pooled fractions containing the protein of interest were then loaded onto a 5-ml HisTrap HP Ni Sepharose column (GE Healthcare, Little Chalfont, United Kingdom) preequilibrated with 100 mM Tris-HCl, pH 8.0, 500 mM NaCl, 5 mM imidazole. The proteins were eluted using a linear imidazole gradient (100 mM Tris-HCl, pH 8.0, 500 mM NaCl, 500 mM imidazole), and the eluted fractions were concentrated using Sartorius Vivaspin concentrators (polyethersulfone membrane) with a 10-kDa cutoff. The final purification step was performed using gel filtration (HiLoad Superdex 75; GE Healthcare) in a buffer containing 50 mM Tris-HCl with 500 mM NaCl. The proteins were concentrated, and the buffer was changed to 10 mM Tris-HCl, pH 7.5. Protein purity was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and estimated to be near 95% for all proteins; protein concentrations were determined using the Bradford assay (Bio-Rad) with bovine serum albumin as a standard.

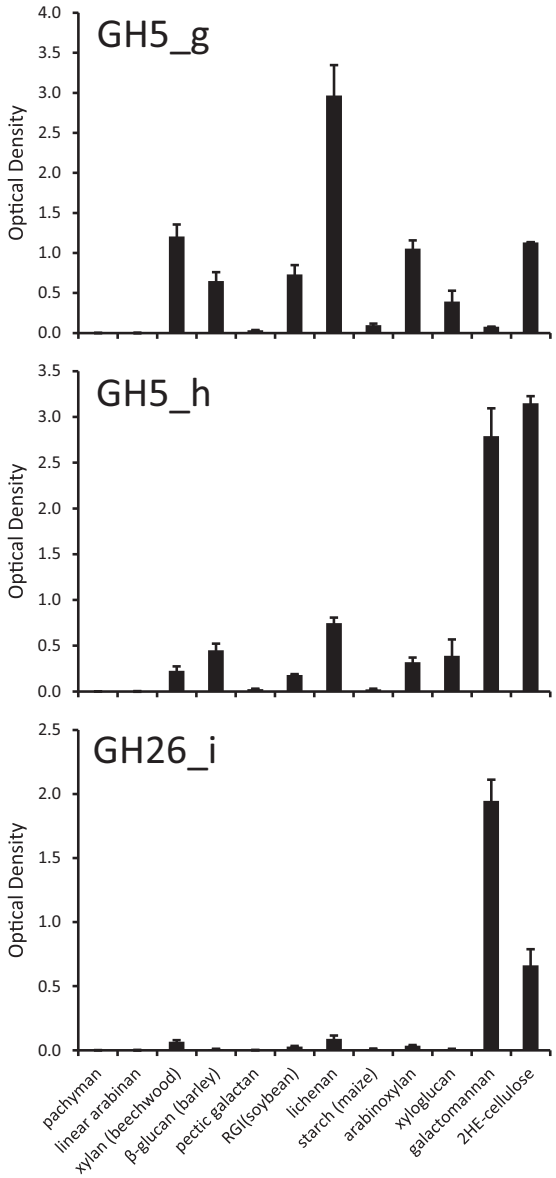
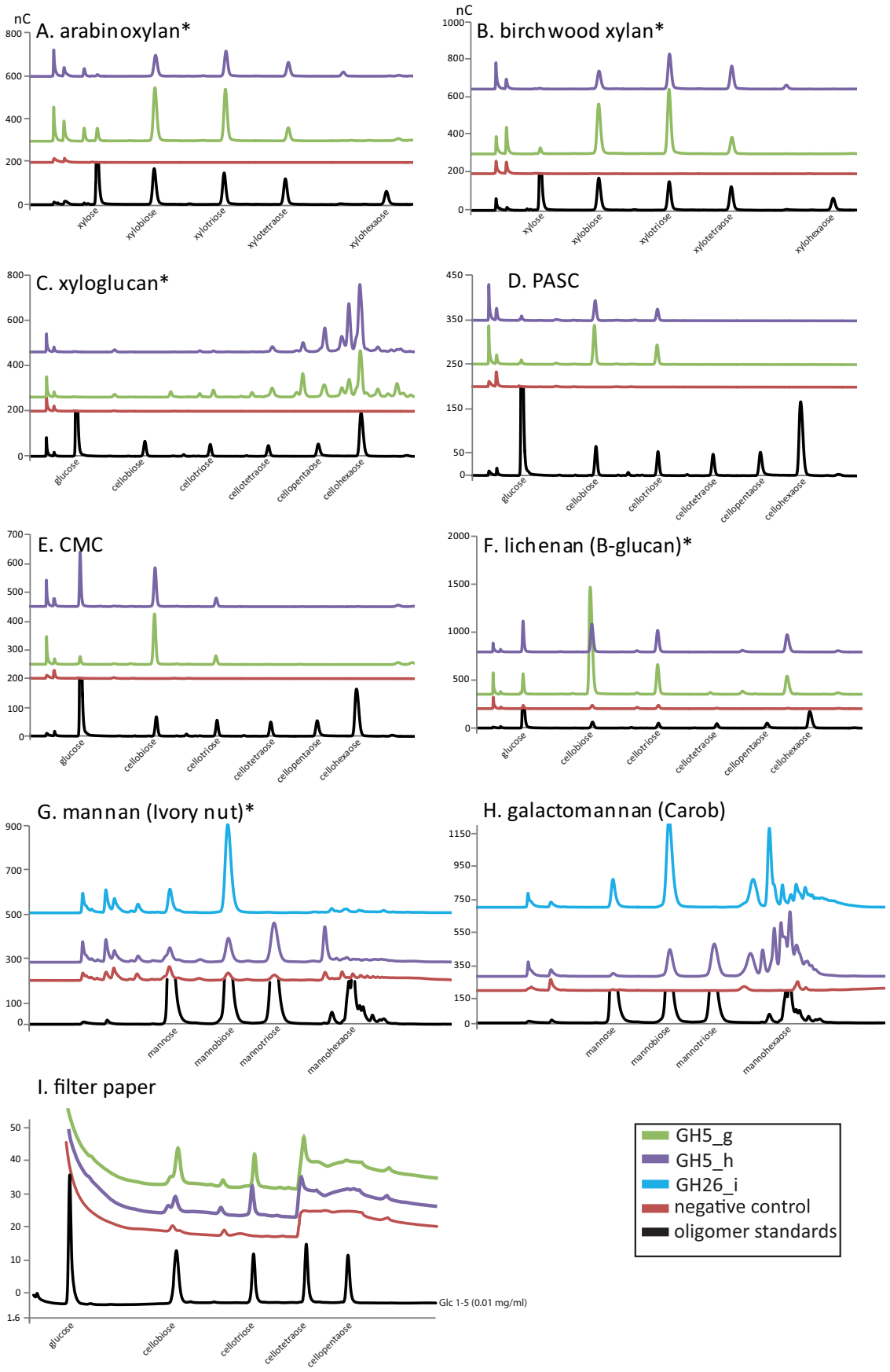


FIG 3 Substrate specificity screening of SRM-1 outer membrane enzymes. The substrate specificities of the two GH5 enzymes and the GH26 enzyme were determined by chromogenic hydrogel substrate screening. Values are reported as the optical density of the filtrate measured against that of the negative controls at 404 nm, 517 nm, 505 nm, and 630 nm for yellow, red, blue, and green chromogenic substrates, respectively. Error bars represent standard deviations between three replicates. RGI, rhamnogalacturonan I; 2HE, 2-hydroxyethyl.

Chromogenic substrate specificity screening. Specific chromogenic substrates were made essentially as described previously (23) and washed with water until no residual color remained. One hundred fifty microliters of buffer (50 mM potassium phosphate pH 7.5), 100 μl of each chromogenic substrate, and 10 μg of enzyme were added to the wells of 96-well filter plates (catalog number MSHVN4550; Millipore, Bedford, MA, USA). The plates were sealed with adhesive PCR plate seals (catalog number AB-0558; Thermo Scientific, MA, USA) and incubated at 40°C with shaking for 1 h. The plates were spun down (2,500 × *g*, 10 min), and the absorption of the filtrate was measured against that of the negative controls at 404 nm, 517 nm, 505 nm, and 630 nm for yellow, red, blue, and green chromogenic substrates, respectively.


nC
A. arabinoxylan*
xylose
xylobiose
xylotriose
xylotetraose
xylohexaose
nC
B. birchwood xylan*
xylose
xylobiose
xylotriose
xylotetraose
xylohexaose
C. xyloglucan*
glucose
cellobiose
cellotriose
cellotetraose
cellopentaose
cellohexaose
D. PASC
glucose
cellobiose
cellotriose
cellotetraose
cellopentaose
cellohexaose
E. CMC
glucose
cellobiose
cellotriose
cellotetraose
cellopentaose
cellohexaose
F. lichenan (B-glucan)*
glucose
cellobiose
cellotriose
cellotetraose
cellopentaose
cellohexaose
G. mannan (Ivory nut)*
mannose
mannobiose
mannotriose
mannohexaose
H. galactomannan (Carob)
mannose
mannobiose
mannotriose
mannohexaose
I. filter paper
glucose
cellobiose
cellotriose
cellotetraose
cellopentaose
Glc 1-5 (0.01 mg/ml)
GH5_g
GH5_h
GH26_i
negative control
oligomer standards

Enzymatic assays. Screening of the pH activity profiles showed that the optimum pH was approximately pH 6.0 for the GH5 and GH26 enzymes, and this pH was used for all subsequent characterization. The activities of the enzymes carboxymethyl cellulose (CMC; Sigma-Aldrich, MO, USA), phosphoric acid-swollen cellulose (PASC) (24), β-glucans (Megazyme, Wicklow, Ireland), birch wood xylan (Carl Roth GmbH, Germany), arabinoxylan (Megazyme), tamarind xyloglucan (Megazyme), and mannans (Megazyme) toward filter paper (Whatman no. 1, milled) were determined. The SRM-1 enzyme activity reaction mixture contained 0.1 M citric acid–0.2 M Na_2HPO_4 buffer (pH 6.0), 0.25 μM enzyme, and 1% (wt/vol) substrate and was incubated with shaking for 1 h at 40°C in a total volume of 200 μl. The reaction was stopped by the addition of 50 μl 0.1 M NaOH. For assays for determination of activity toward filter paper, the conditions were 1% substrate (wt/vol) and 100 nM enzyme with an overnight incubation time. The reactions were stopped by boiling (5 min), before soluble cellodextrins were quantified. Soluble oligosaccharide products were quantified against a standard curve by high-pressure anion-exchange chromatography with electrochemical detection using a Dionex ICS-3000 system with a CarboPac PA1 column with 0.1 M NaOH at a flow rate of 0.25 ml/min. Oligosaccharides were eluted by a multistep linear gradient going from 0.1 M NaOH to 0.1 M NaOH–0.1 M sodium acetate (NaOAc) over 35 min, 0.1 M NaOH–0.3 M NaOAc over 25 min, and 0.1 M NaOH–1 M NaOAc over 5 min, before reconditioning with 0.1 M NaOH for 9 min. Time course synergy experiments were done with 1% (wt/vol) lichenan β-glucan (Megazyme) and 0.5% (wt/vol) Konjac glucomannan (Megazyme). Substrates were incubated at 40°C, with horizontal shaking, with a total of 56 nM enzyme (for the three-enzyme cocktail, 18.67 nM each; for the two-enzyme cocktail, 28 nM each; for one enzyme, 56 nM) in a total volume of 1,056 μl. Aliquots of 150 μl were removed at 10-min intervals, and the reactions were stopped by the addition of 50 μl 0.1 M NaOH. An equal amount of 3,5-dinitrosalicylic acid (DNS) reagent was added (200 μl), and the amount of reducing sugars relative to that of a cellobiose standard curve (diluted with 50 μl 0.1 M NaOH) was determined using the DNS assay (25). The degree of synergism was defined as the ratio of the cellobiose equivalents produced by the combined action of glycoside hydrolases on a given substrate to the total cellobiose equivalents produced by their independent actions (calculated at 10 min).

Pulldown assays. Binding assays were performed under conditions similar to those used previously for SusD-like lipoproteins from the SRM-1 PUL (17). Briefly, crystalline cellulose (Avicel and filter paper [Whatman no. 1, milled to 0.5 mm]), and the insoluble fractions (room temperature) of β-glucan (lichenan and barley; Megazyme), ivory nut mannan, and xylan were washed twice (500 μl), resuspended to 2% (wt/vol) in a total volume of 200 μl buffer (20 mM MES [morpholineethanesulfonic acid] buffer, pH 6.0) along with 1 mg/ml protein, and incubated at 40°C with horizontal shaking. The substrate and bound protein were pelleted by centrifugation (10,000 × *g*), and the supernatant containing unbound protein (referred to as the "flowthrough") was carefully removed. The pellet was washed with 200 μl buffer and incubated with shaking for 15 min before the supernatant was again removed by centrifugation (referred to as the "wash"). The pellet was washed with an additional 800 μl for 15 min before the supernatant was removed. To elute the proteins, the pellets were resuspended in 100 μl 8 M urea, vortexed briefly, and incubated for 10 min at room temperature. The flowthrough, initial wash, and elution fractions were analyzed by SDS-PAGE.

Nucleotide sequence accession numbers. The GenBank accession numbers for GH5_g, GH5_h, GH26_i, and the SusE-positioned protein are KM887864, KM887865, KM887866, and KM887863, respectively.

RESULTS AND DISCUSSION

Cell wall glycan content of the Svalbard reindeer rumen. Svalbard reindeer inhabit the high arctic archipelago of Svalbard under austere nutritional conditions, particularly in winter, when access to grass is restricted and their diet is supplemented by dicots (flowering plants) and bryophytes (nonvascular mosses). Previous measurements of Svalbard reindeer rumen contents have identified cellulose (10 to 23%) and hemicellulose (16 to 29%) to be the major glycan constituents in the ingested diet (20). Hemicellulose consists of different heteropolymers that vary depending on the plant species, including xylan, glucuronoxylan, arabinoxylan, glucomannan, β-glucan, and xyloglucan. Therefore, as a prelude to our enzymological assessment, we examined the relative abundances of the different hemicellulosic substrates in more detail by using comprehensive microarray polymer profiling (Fig. 1). Xylans and β-glucans, both of which are common in the cell walls of grasses (26), were abundant in the reindeer rumen, suggesting that grass species are still part of the host's daily feed intake in winter (20). Xyloglucans and mannans were also abundant (Fig. 1), which was expected, given that these polysaccharides are abundant in both mosses and dicots (26, 27).

Biochemical characterization of outer membrane glycoside hydrolases and binding proteins associated with the SRM-1 hemicellulose-degrading PUL. Biochemical characterization of the two GH5 (GH5_g, GH5_h) enzymes, the GH26 enzyme, and the SusE-positioned lipoprotein was pursued to provide insight into the saccharolytic capabilities of the SRM-1 PUL. Analysis of the various fractions from the binding experiments by SDS-PAGE (Fig. 2) showed that the SusE-positioned protein from the SRM-1 PUL bound to β-glucans and, to a much lesser extent, mannan. Unlike the SusD-like lipoproteins, no binding to crystalline cellulose was observed (17). Both of the GH5 enzymes were annotated as subfamily 4 members (28), which are typically extracellular bacterial enzymes that have previously been reported to be endoglucanases, xyloglucanases, xylanases, and licheninases. Differences between the two enzymes were noted in the N terminus, where GH5_g has an additional bacterial Ig-like domain (PF13205) and GH5_h has an additional BACON domain (PF13004). All purified GH enzymes were initially screened using chromogenic substrates (Fig. 3). Despite their identical CAZy carbohydrate-active enzyme annotation, the two GH5 endoglucanases exhibited differences in substrate specificity, with GH5_g demonstrating activity against xylans, β-glucans (lichenan and barley), xyloglucan, and soluble cellulose derivatives (2-hydroxyethyl cellulose). While detectable GH5_h activity against β-glucans, xylan, and cellulosic substrates was demonstrated, mannan activity was additionally detected. These results were confirmed by analysis of the products generated from

FIG 4 SRM-1 GH5 and GH26 product analysis. High-pH anion-exchange chromatography-pulsed amperometric detection chromatograms of products generated from xylans (A and B), xyloglucan (C), amorphous cellulose (D and E), lichenan (F), mannans (G and H), and crystalline filter paper (I) digested by GH5 and GH26 enzymes. Enzyme assays were performed for 60 min (A to H) or overnight (I), and the enzymes were inactivated by NaOH (A to H) or boiling (I). Oligomers (Megazyme) were mixed to known concentrations and used as standards, whereas assay mixtures without added enzymes were used as a negative control. GH26 showed no activity toward the substrates used in the assays whose results are presented in panels A to F, and those results have been excluded from the graphs for clarity. GH5_g showed no activity toward mannans, and results for GH5_g have been excluded from panels G and H for clarity. *, substrates which were determined to be abundant in the reindeer rumen contents (Fig. 1).

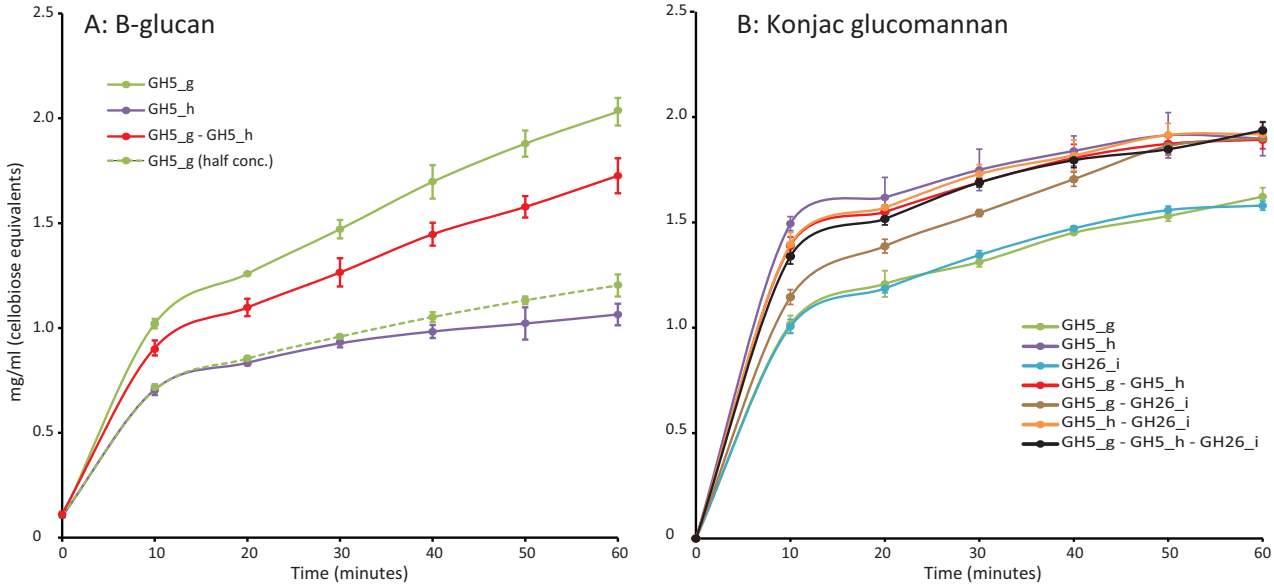


FIG 5 Characterization of enzyme cocktails containing various SRM-1 PUL glycoside hydrolases. Time course analysis of enzymatic reaction mixtures (equal concentrations) containing GH5_g, GH5_h, GH26_i, and different combinations of the three enzymes. Their activities against β-glucan (A) and glucomannan (B) substrates were determined by DNS analysis of the products (cellobiose equivalents). Error bars represent standard deviations for three replicates. Time course synergy experiments were done with either 1% (wt/vol) lichenan β-glucan or 0.5% (wt/vol) Konjac glucomannan, plus a total of 56 nM enzyme (GH5_g, GH5_h, GH26_i, or combinations thereof).

natural polymeric substrates (Fig. 4). Hydrolysis of β-glucan, CMC, and PASC by both GH5 enzymes produced degree of polymerization 1 (DP1) to DP3 oligosaccharides, whereas their activity against xyloglucan produced longer oligosaccharides (>DP4) and no disaccharides. Very weak activity was observed by both GH5 enzymes on crystalline cellulose (filter paper), producing trace amounts of DP2 to DP4 oligosaccharides (Fig. 4). Both GH5 enzymes generated dimer, trimer, and tetramer xylooligosaccharides on xylan, whereas only GH5_h and GH26 were active on mannan, producing mannooligosaccharides with different length distributions (Fig. 4). GH26 did not generate oligosaccharides from any natural substrate other than mannan, despite chromogenic analysis suggesting some activity on 2HE-cellulose (Fig. 3). Surprisingly, enzyme cocktails containing various combinations of GH5_g, GH5_h, and GH26_i seemingly exhibited limited synergistic effects on the degradation of β-glucan and glucomannan substrates (Fig. 5). No enzyme cocktail exceeded the performance of the best individual enzyme (GH5_g on β-glucan, GH5_h on glucomannan); however, the combined action of GH5_g and GH26_i on glucomannan was higher than the independent actions of GH5_g (degree of synergism, 1.13 at 10 min) and GH26_i (degree of synergism, 1.14 at 10 min).

Proposed model of the SRM-1 PUL. A model of the SRM-1 PUL was reconstructed using a combination of our new activity information, information from previous studies regarding SusD-like lipoproteins (17), and homology with the model starch utilization system of *Bacteroides thetaiotaomicron* (Fig. 6A). Collectively, the results suggest that the SRM-1 PUL functions by binding to plant cell walls via the actions of the outer membrane SusD ($n = 2$) and SusE-positioned lipoproteins, which demonstrate affinity toward cellulose, β-glucans, and, to a lesser extent, mannan (Fig. 6A). This in turn would allow the GH5 and GH26 enzymes to act upon hemicellulosic substrates in close proximity, producing oligosaccharides that are imported via the actions of SusC-like transporters. The cellobiose produced via the GH5 enzymes is predicted to be converted to glucose via the actions of a periplasmic GH94 cellobiose phosphorylase. Similarly, the mannooligosaccharides produced from the GH26 enzyme are predicted to be degraded into mannose via the actions of a putative GH130 mannooligosaccharide phosphorylase, whereas galactose monomers are predicted to be released via GH2 β-galactosidase activity (Fig. 6A). Although xylooligosaccharides were released by GH5 activity on xylan, no enzymes involved in downstream xyloside hydrolysis were identified in the SRM-1 PUL. However, beta-xylosidases have been identified in the SRM-1 genome (12). In addition, the incomplete gene_a encodes both a GH5 endoglucanase and a CE7 acetyl xylan esterase domain, suggesting deacetylation of xylans and xylooligosaccharides (Fig. 6A). By analogy to archetypal starch PULs (2), the final transportation of monomer sugars is predicted to be facilitated by an inner membrane transporter (Fig. 6A) prior to further utilization.

Concluding remarks. Despite the importance of the rumen microbiome, a comprehensive understanding of the microbial ecology and enzymology of plant biomass deconstruction in these ecosystems is still being developed. Enzymes, cellulosome components, and PULs have been described for several renowned cultured representatives, such as *Prevotella ruminicola* (29), *Fibrobacter succinogenes* (30), *Butyrivibrio fibrisolvens* (31), *Ruminococcus flavefaciens* (32), and *Prevotella bryantii* (4). More recently, metagenomic studies have revealed enzymes from uncultured and deeply branched lineages, which constitute the majority of microorganisms within rumen microbiomes (15). While PULs originating from uncultured rumen bacteria have previously been detected (12, 14, 33, 34), here we have taken additional measures to partially describe the enzymatic capabilities of one such PUL encoded within a numerically abundant uncultured phylotype. The

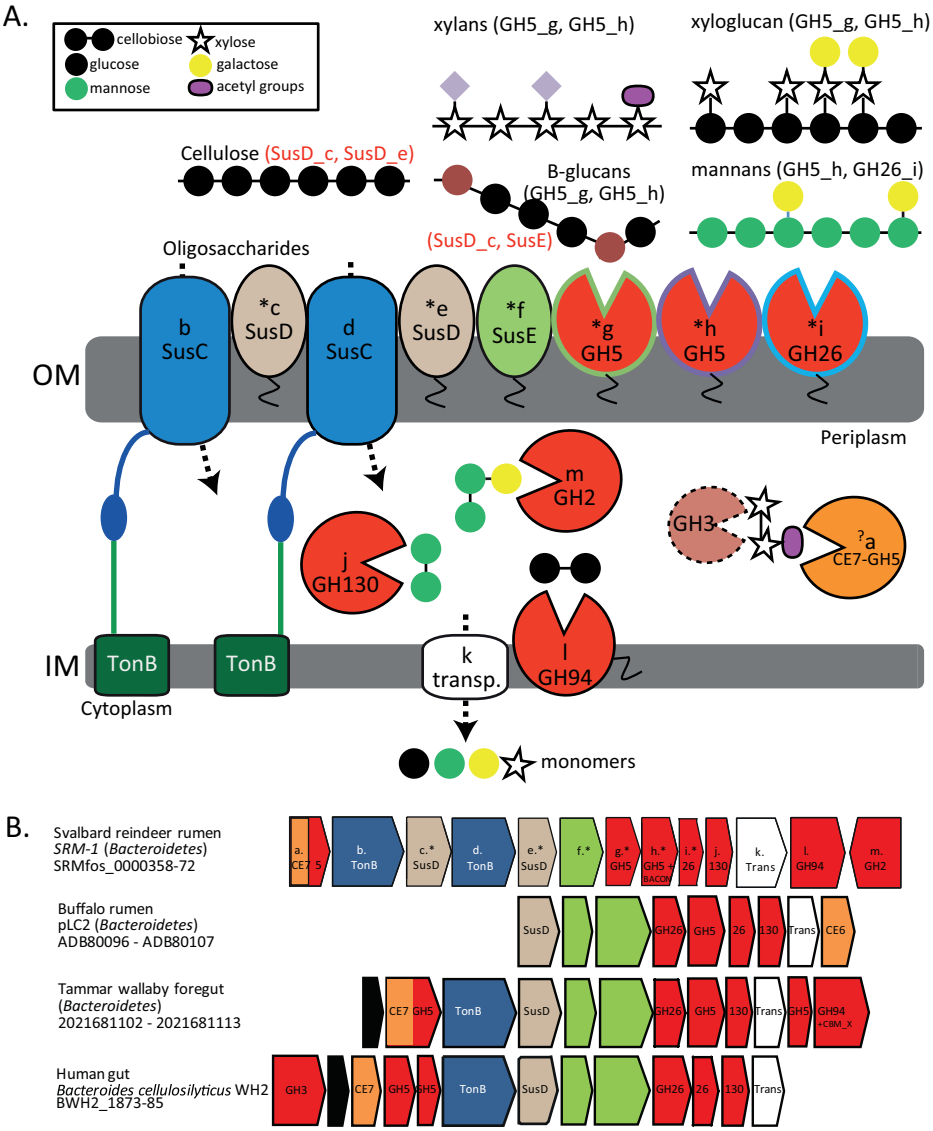


FIG 6 Gene organization of the SRM-1 PUL and hypothetical model of the encoded enzyme system indicating the predicted localization and activity of its components. (A) Predicted proteins marked with an asterisk were subjected to biochemical characterization, and their activities are indicated in parentheses (red type, binding activity; black type, enzyme activity). Binding data for SusD-like lipoproteins (SusD_c and SusD_e, previously termed SusD.1 and SusD.2, respectively) are from Mackenzie et al. (17); the other data are from the present study. Taken together, the biochemical data for the outer membrane lipoproteins (SusD-like and SusE-positioned lipoproteins) and enzymes (GH26 and 2 GH5 enzymes) suggest that substrates are bound to the cell surface and degraded into oligosaccharides prior to being transported via the actions of SusC-like transporters (5, 6). These oligosaccharides are proposed to be further hydrolyzed to monomers by periplasmic GHs (GH94, GH130, and GH2) and finally transported to the cytoplasm via a glycoside sugar transporter. Because gene_a was incomplete (it was missing N terminus), its cell localization could not be predicted. Beta-xylosidases (GH3; indicated by a dotted line) were not encoded in the SRM-1 PUL but have previously been identified in the SRM-1 genome (12). OM, outer membrane; IM, inner membrane. (B) Genomic comparisons of the SRM-1 PUL with PULs from publicly available metagenomes that encode GH5 and GH26 representatives and a PUL from the human gut bacterium *Bacteroides cellulosilyticus* that is upregulated by mannan substrates (37). The figure was partly adapted from a paper by Naas et al. (34).

gene organization of the SRM-1 PUL demonstrated interesting synteny with PULs uncovered from other gastrointestinal microbiomes from ruminants, marsupial foregut fermenters, and the human gut, representing both uncultured phylotypes and well-characterized cultured species (Fig. 6B). All of these listed PULs encoded one or more GH5, GH26, and GH130 representatives, suggesting that such loci are somewhat conserved and widely distributed in ruminants and nonruminants alike. Biochemically, the SRM-1 PUL possesses several intriguing features. Most notable were the inconsistencies between the detected hemicellulase activities of the GHs and the cellulose-binding affinities of the SusD-like lipoproteins. Interestingly, these observations bear similarities to published findings for GH5 and GH26 mannanases in *Cellvibrio japonicas*, which also contain cellulose-binding CBMs (35). In addition, Zhang et al. observed that recalcitrant mannan is fully accessible to GH26 mannanases that are appended to a cellulose-binding CBM (36).

Despite the structural variance in the list of polysaccharides against which previously described PULs have demonstrated activity, characterization studies consistently indicate that PULs have narrow specificity regarding their target glycan. In other

words, there seems to be one cognate PUL for each substrate. This has been elegantly illustrated via transcriptomic studies of *B. thetaiotaomicron*, *Bacteroides ovatus*, and *Bacteroides cellulosilyticus*, which have shown a plethora of different PULs that are upregulated in response to their specific hemicellulose or pectin target (1, 37). Furthermore, characterized GH5 and GH9 enzymes within a *B. ovatus* xyloglucan PUL were documented to lack activity against any other hemicellulose or cellulose substrate, even though they were members of the more functionally versatile CAZy families (10). In contrast, the GH5 enzymes originating from the SRM-1 PUL have versatile activity targeting multiple substrates, including β-glucans, mannans, xylan, xyloglucan, and amorphous cellulose, although it is unlikely that the SRM-1 PUL is dedicated to cellulosic substrates in nature. The limited observations of synergy suggest that this one PUL targets different major components of the accessible plant biomass, rather than working collectively to degrade individual substrates (Fig. 5). While the ability of the SRM-1 PUL to degrade the most readily available polysaccharides in its environment is a seemingly innocuous observation, it does illustrate the ecological relevance of the SRM-1 phylotype and provides a rationale for its dominance and adaptation to the host-associated microbial ecosystem.

ACKNOWLEDGMENTS

We are especially grateful for the support from The Research Council of Norway's FRIPRO program (214042) and the European Commission Marie Curie International Incoming Fellowship (awarded to P.B.P.; PIIF-GA-2010-274303). A.K.M. is supported by a grant from the Norwegian Research Council (190965/S60). J.S.'s contribution to this paper has received funding from the European Union Seventh Framework Programme (FP7 2007–2013) under the WallTraC project (grant agreement no. 263916).

REFERENCES

1. **Martens EC, Lowe EC, Chiang H, Pudlo NA, Wu M, McNulty NP, Abbott DW, Henrissat B, Gilbert HJ, Bolam DN, Gordon JI.** 2011. Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. PLoS Biol **9:**e1001221. http://dx.doi.org/10.1371/journal.pbio.1001221.
2. **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. http://dx.doi.org/10.1074/jbc.R109.022848.
3. **Koropatkin NM, Cameron EA, Martens EC.** 2012. How glycan metabolism shapes the human gut microbiota. Nat Rev Microbiol **10:**323–335. http://dx.doi.org/10.1038/nrmicro2746.
4. **Dodd D, Moon YH, Swaminathan K, Mackie RI, Cann IK.** 2010. Transcriptomic analyses of xylan degradation by Prevotella bryantii and insights into energy acquisition by xylanolytic Bacteroidetes. J Biol Chem **285:**30261–30273. http://dx.doi.org/10.1074/jbc.M110.141788.
5. **Schauer K, Rodionov DA, de Reuse H.** 2008. New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'? Trends Biochem Sci **33:**330–338. http://dx.doi.org/10.1016/j.tibs.2008.04.012.
6. **Ferguson AD, Deisenhofer J.** 2002. TonB-dependent receptors—structural perspectives. Biochim Biophys Acta **1565:**318–332. http://dx.doi.org/10.1016/S0005-2736(02)00578-3.
7. **Koropatkin NM, Martens EC, Gordon JI, Smith TJ.** 2008. Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. Structure **16:**1105–1115. http://dx.doi.org/10.1016/j.str.2008.03.017.
8. **Hehemann JH, Kelly AG, Pudlo NA, Martens EC, Boraston AB.** 2012. Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. Proc Natl Acad Sci U S A **109:**19786–19791. http://dx.doi.org/10.1073/pnas.1211002109.
9. **Sonnenburg ED, Zheng H, Joglekar P, Higginbottom SK, Firbank SJ, Bolam DN, Sonnenburg JL.** 2010. Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. Cell **141:**1241–1252. http://dx.doi.org/10.1016/j.cell.2010.05.005.
10. **Larsbrink J, Rogers TE, Hemsworth GR, McKee LS, Tauzin AS, Spadiut O, Klinter S, Pudlo NA, Urs K, Koropatkin NM, Creagh AL, Haynes CA, Kelly AG, Cederholm SN, Davies GJ, Martens EC, Brumer H.** 2014. A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes. Nature **506:**498–502. http://dx.doi.org/10.1038/nature12907.
11. **Edwards JE, McEwan NR, Travis AJ, Wallace RJ.** 2004. 16S rDNA library-based analysis of ruminal bacterial diversity. Antonie Van Leeuwenhoek **86:**263–281. http://dx.doi.org/10.1023/B:ANTO.0000047942.69033.24.
12. **Pope PB, Mackenzie AK, Gregor I, Smith W, Sundset MA, McHardy AC, Morrison M, Eijsink VGH.** 2012. Metagenomics of the Svalbard reindeer rumen microbiome reveals abundance of polysaccharide utilization loci. PLoS One **7:**e38571. http://dx.doi.org/10.1371/journal.pone.0038571.
13. **Pope PB, Denman SE, Jones M, Tringe SG, Barry K, Malfatti SA, McHardy AC, Cheng J-F, Hugenholtz P, McSweeney CS, Morrison M.** 2010. Adaptation to herbivory by the Tammar wallaby includes bacterial and glycoside hydrolase profiles different to other herbivores. Proc Natl Acad Sci U S A **107:**14793–14798. http://dx.doi.org/10.1073/pnas.1005297107.
14. **Rosewarne CP, Pope PB, Cheung JL, Morrison M.** 2014. Analysis of the bovine rumen microbiome reveals a diversity of Sus-like polysaccharide utilization loci from the bacterial phylum Bacteroidetes. J Ind Microbiol Biotechnol **41:**601–606. http://dx.doi.org/10.1007/s10295-013-1395-y.
15. **Hess M, Sczyrba A, Egan R, Kim TW, Chokhawala H, Schroth G, Luo S, Clark DS, Chen F, Zhang T, Mackie RI, Pennacchio LA, Tringe SG, Visel A, Woyke T, Wang Z, Rubin EM.** 2011. Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. Science **331:**463–467. http://dx.doi.org/10.1126/science.1200387.
16. **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. http://dx.doi.org/10.1038/nmeth0311-191.
17. **Mackenzie AK, Pope PB, Pedersen HL, Gupta R, Morrison M, Willats WG, Eijsink VGH.** 2012. Two SusD-like proteins encoded within a polysaccharide utilization locus of an uncultured ruminant Bacteroidetes phylotype bind strongly to cellulose. Appl Environ Microbiol **78:**5935–5937. http://dx.doi.org/10.1128/AEM.01164-12.
18. **Moller I, Sørensen I, Bernal AJ, Blaukopf C, Lee K, Øbro J, Pettolino F, Roberts A, Mikkelsen JD, Knox JP, Bacic A, Willats WG.** 2007. High-throughput mapping of cell-wall polymers within and between plants using novel microarrays. Plant J **50:**1118–1128. http://dx.doi.org/10.1111/j.1365-313X.2007.03114.x.
19. **Moller IE, Pettolino FA, Hart C, Lampugnani ER, Willats WG, Bacic A.** 2012. Glycan profiling of plant cell wall polymers using microarrays. J Vis Exp **70:**e4238. http://dx.doi.org/10.3791/4238.
20. **Sørmo W, Haga ØE Gaare E, Langvatn R, Mathiesen SD.** 1999. Forage chemistry and fermentation chambers in Svalbard reindeer (Rangifer tarandus platyrhynchus). J Zool Lond **247:**247–256. http://dx.doi.org/10.1111/j.1469-7998.1999.tb00988.x.
21. **Orpin CG, Mathiesen SD, Greenwood Y, Blix AS.** 1985. Seasonal changes in the ruminal microflora of the high-arctic Svalbard reindeer (*Rangifer tarandus platyrhynchus*). Appl Environ Microbiol **50:**144–151.
22. **Aslanidis C, de Jong PJ.** 1990. Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res **18:**6069–6074. http://dx.doi.org/10.1093/nar/18.20.6069.
23. **Ten LN, Im W, Kim MK, Kang MS, Lee ST.** 2004. Development of a plate technique for screening of polysaccharide-degrading microorganisms by using a mixture of insoluble chromogenic substrates. J Microbiol Methods **56:**375–382. http://dx.doi.org/10.1016/j.mimet.2003.11.008.
24. **Wood TM.** 1988. Preparation of crystalline, amorphous, and dyed cellulase substrates. Methods Enzymol **160:**19–25. http://dx.doi.org/10.1016/0076-6879(88)60103-0.
25. **Miller GL.** 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal Chem **31:**426–428. http://dx.doi.org/10.1021/ac60147a030.
26. **Vogel J.** 2008. Unique aspects of the grass cell wall. Curr Opin Plant Biol **11:**301–307. http://dx.doi.org/10.1016/j.pbi.2008.03.002.
27. **Popper ZA, Fry SC.** 2003. Primary cell wall composition of bryophytes

and charophytes. Ann Bot **91:**1–12. http://dx.doi.org/10.1093/aob/mcg013.
28. **Aspeborg H, Coutinho PM, Wang Y, Brumer H, Henrissat B.** 2012. Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). BMC Evol Biol **12:**186. http://dx.doi.org/10.1186/1471-2148-12-186.
29. **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. http://dx.doi.org/10.1007/s002849900253.
30. **Qi M, Jun HS, Forsberg CW.** 2008. Cel9D, an atypical 1,4-beta-D-glucan glucohydrolase from *Fibrobacter succinogenes*: characteristics, catalytic residues, and synergistic interactions with other cellulases. J Bacteriol **190:**1976–1984. http://dx.doi.org/10.1128/JB.01667-07.
31. **Hazlewood GP, Davidson K, Laurie JI, Romaniec MP, Gilbert HJ.** 1990. Cloning and sequencing of the celA gene encoding endoglucanase A of *Butyrivibrio fibrisolvens* strain A46. J Gen Microbiol **136:**2089–2097. http://dx.doi.org/10.1099/00221287-136-10-2089.
32. **Kirby J, Martin JC, Daniel AS, Flint HJ.** 1997. Dockerin-like sequences in cellulases and xylanases from the rumen cellulolytic bacterium *Ruminococcus flavefaciens*. FEMS Microbiol Lett **149:**213–219. http://dx.doi.org/10.1016/S0378-1097(97)00078-5.
33. **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:**e78507. http://dx.doi.org/10.1371/journal.pone.0078507.
34. **Naas AE, Mackenzie AK, Mravec J, Schückel J, Willats WG, Eijsink VG, Pope PB.** 2014. Do rumen *Bacteroidetes* utilize an alternative mechanism for cellulose degradation? mBio **5**(4):e01401-14. http://dx.doi.org/10.1128/mBio.01401-14.
35. **Hogg D, Pell G, Dupree P, Goubet F, Martín-Orúe SM, Armand S, Gilbert HJ.** 2003. The modular architecture of *Cellvibrio japonicus* mannanases in glycoside hydrolase families 5 and 26 points to differences in their role in mannan degradation. Biochem J **371:**1027–1043. http://dx.doi.org/10.1042/BJ20021860.
36. **Zhang X, Rogowski A, Zhao L, Hahn MG, Avci U, Knox JP, Gilbert HJ.** 2014. Understanding how the complex molecular architecture of mannan-degrading hydrolases contributes to plant cell wall degradation. J Biol Chem **289:**2002–2012. http://dx.doi.org/10.1074/jbc.M113.527770.
37. **McNulty NP, Wu M, Erickson AR, Pan C, Erickson BK, Martens EC, Pudlo NA, Muegge BD, Henrissat B, Hettich RL, Gordon JI.** 2013. Effects of diet on resource utilization by a model human gut microbiota containing *Bacteroides cellulosilyticus* WH2, a symbiont with an extensive glycobiome. PLoS Biol **11:**e1001637. http://dx.doi.org/10.1371/journal.pbio.1001637.