Differential bacterial capture and transport preferences facilitate competition for dietary fibers in the human gut

Maria Louise Leth[1], Morten Ejby[1], Christopher T. Workman[1], David Adrian Ewald[1], Signe Schultz Pedersen[1], Claus Sternberg[1], Martin Iain Bahl[2], Tine Rask Licht[2], Finn Lillelund Aachmann[3], Bjørge Westereng[4], Maher Abou Hachem[1]

1. Dept. of Biotechnology and Biomedicine, Technical University of Denmark, Lyngby, Denmark, 2. National Food Institute, Technical University of Denmark, Lyngby, Denmark, 3. NOBIPOL, Department of Biotechnology and Food Science, NTNU Norwegian University of Science and Technology, Sem Sælands vei 6-8, N-7491 Trondheim, Norway, 4. Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, P.O. Box 5003, N-1432 Ås.

Abstract

Metabolism of dietary glycans is an important factor in shaping the human gut microbiota. *Roseburia intestinalis* is a key degrader of the highly abundant dietary glycan xylan, but despite its relevance to human health, insight its xylan utilization strategies and competitiveness remains unexplored. *R. intestinalis* offer an attractive model to study features that promote competition for glycans by gut Firmicutes commensals. Here, we investigate xylan capture and transport mechanisms of *R. intestinalis* and show that this species deploys a large cell-attached modular xylanase that promotes multivalent and dynamic association to xylan by means of four xylan-binding modules including a representative of a novel family. This xylanase operates in concert with an ATP-binding cassette (ABC) transporter to mediate break-down and selective internalisation of substituted xylan-fragments. This apparatus supports competition between *R. intestinalis* with a model xylan-degrading *Bacteriodes* in mixed cultures. We show that the transport-protein conferring the capture of xylan-oligomers by *R. intestinalis* displays a strikingly different capture profile compared to the *Bacteroides* counterpart, which minimizes competition for preferred xylan-fragments. These findings highlight the differentiation of capture and transport preferences as a strategy to facilitate competition for abundant dietary fibers by human gut commensals.

Introduction

The human gut microbiota (HGM) is recognized as a determinant of human health and metabolic homeostatsis[1,2]. Specific signatures of the HGM are associated with local and systemic disorders including irritable-bowel disease, obesity, type 2 diabetes and colon cancer[3]. The composition of the HGM is greatly affected by dietary complex glycans, which are not digestible by the host[4,5]. Only a few species out of the hundreds present in the HGM are equipped to deconstruct distinct complex polysaccharides and ferment them into short chain fatty acids (SCFAs). The impact of SCFAs on host health and physiology remains an important aspect of the microbiota-host interaction. Particularly the SCFA butyrate, the preferred energy source for colonocytes, is known to have anti-inflammatory roles and reduce the risk of colon cancer and enteric colitis[6–9]. Butyrate producers belonging to the Firmicutes phylum are generally abundant in healthy individuals, but are markedly reduced in patients with inflammatory disorders[10,11]. Butyrate producers including *Roseburia* spp. are increased in metabolic syndrome patients after faecal transfer therapy, and correlate positively to improvement of insulin resistance[12]. Investigations of the metabolic preferences of butyrate producers and their interplay with major HGM commensals are instrumental to develop therapeutic interventions targeting butyrate-deficiency related disorders.

Roseburia is a common genus of *Clostridium* cluster XIVa within the Firmicutes that harbours prevalent butyrate producers[13,14] and has been shown to adhere to mucin, consistent with an intimate association with the host[15]. *Roseburia intestinalis* strains encode an impressive repertoire of carbohydrate active enzymes (CAZymes) compared to most other Firmicutes[16]. *R. intestinalis* represents one of the few HGM species that utilize the major hemicellulosic polysaccharide xylan[17,18]. Xylan is particularly abundant in cereal grains (arabinoxylan, AX), but is also found in fruits and vegetables (glucuronoxylan, GX)[19] (Fig. 1a). Xylan utilization by dominant gut commensals belonging to the *Bacteriodes* genus has been investigated in detail[20,21], but similar knowledge is lacking for Firmicutes counterparts.

Here, we show that *Roseburia intestinalis* L1-82 grows on acetyl, arabinosyl and 4-*O*-methyl-glucuronosyl decorated dietary-relevant xylans, and that it prefers for arabinoxylans from cereals. The growth is mediated by a multi-modular cell-attached xylanase and by an ABC transporter. The gene encoding this transporter, which displays selectivity for substituted xylo-oligomers, was the most upregulated gene in response to xylan, consistent with a paramount role during growth on this glycan. We have performed a detailed biophysical and enzymatic characterization of the enzymes and the transport protein that mediate extracellular breakdown, uptake and intracellular degradation of xylan into monosaccharides and acetate. These studies enabled modelling of xylan utilization by *R. intestinalis* and identification of two novel xylan-specific CAZyme families. We also showed that *R.*

intestinalis efficiently competes with a model xylan degrader belonging to the genus *Bacteroides*, when grown on soluble and insoluble xylans. The results emphasise the competitiveness of butyrate producing Firmicutes, and their status as primary xylan degraders in the human gut. Our findings highlight differential capture and transport preference as a key feature that potentiate competition for abundant dietary fibres such as xylan.

Results

Inducible cell-attached xylanase activity mediates growth of *R. intestinalis* on substituted xylans. Anaerobic growth of *R. intestinalis* L1-82 was measured as an increase in $OD_{600\ nm}$ and as a decrease in pH for growth on insoluble xylans (Fig. 1b-d). *R. intestinalis* L1-82 grows rapidly on soluble xylans with a preference for wheat arabinoxylan (WAX, μ_{max}=0.26 h^{-1}) compared to birch glucuronoxylan (BGX, μ_{max}=0.13 h^{-1}) (Fig. 1c). Interestingly, this bacterium also utilizes highly acetylated xylans and insoluble cereal arabinoxylans from wheat (InWAX) and oat spelt (OSX), but not corn bran glucuronoarabinoxylan (CBX). Xylo-oligosaccharides (xylobiose, X2) and xylan-derived monosaccharides (except glucuronic acid) were also efficiently utilized (Fig. 1b). Extracellular *endo*-1,4-β-xylanase (hereafter referred to as xylanase) activity was induced upon growth on BGX and WAX, followed by xylobiose (X2), despite poor growth on this substrate (Fig. 1e). The xylanase activity was cell-attached, but was released upon treatment of the cells with a high salt concentration (Fig. 1f), suggesting noncovalent attachment.

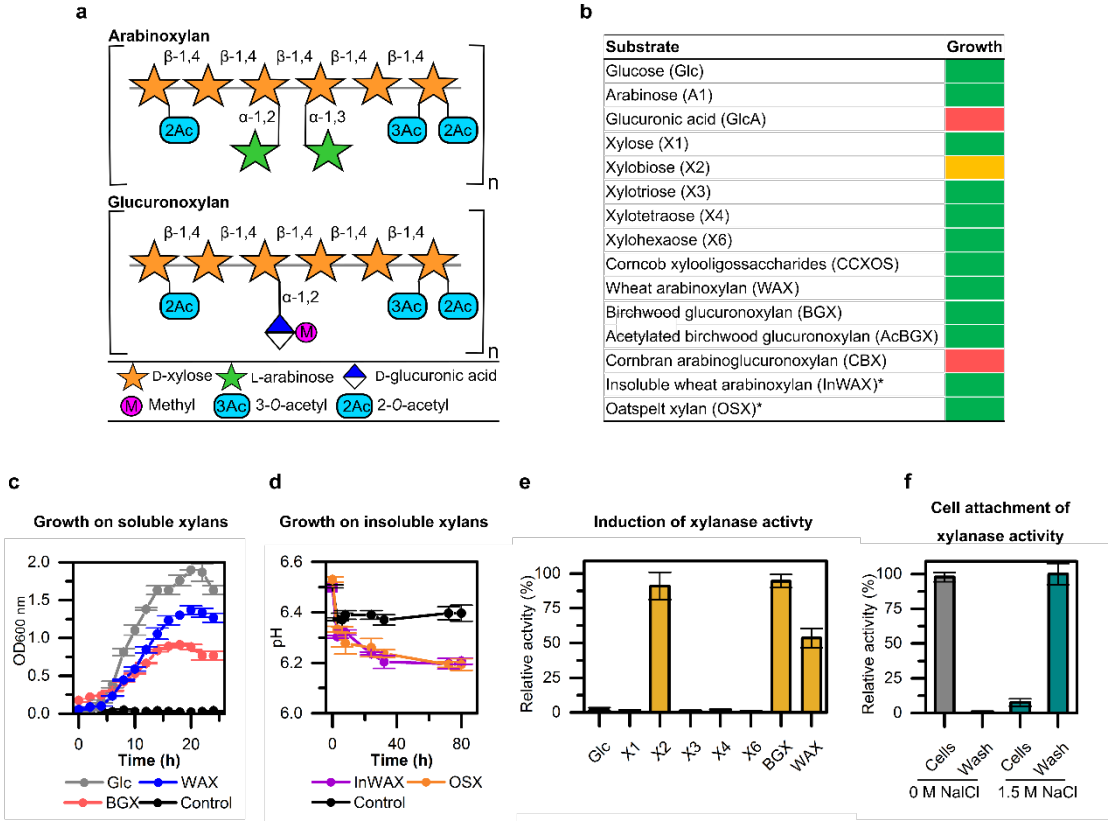


Figure 1 Growth of *R. intestinalis* and induction of extracellular activity. (**a**) Schematic representation of cereal arabinoxylan and glucuronoxylan present in dicots cell wall, e.g. in fruits and vegetables. (**b**) Growth level for 18 hours on xylans, oligomers thereof and monosaccharide components, with glucose as a control. Green: $OD_{600\ nm}$ increase >1.0 for soluble substrates and pH drop > 0.3 for insoluble xylans; yellow: 0.3 < $\Delta OD_{600\ nm}$ < 0.5; red: $\Delta OD_{600\ nm}$< 0.1. Asterisks indicate insoluble xylans (**c**) Growth curves on glucose, wheat arabinoxylan (WAX), birch glucuronoxylan (BGX) and a no carbon source control. (**d**) Growth on insoluble wheat arabinoxylan (InWAX) and oatspelt xylan (OSX). All growth measurements are means of triplicates with standard deviations. (**e**) Xylanase activity of *R. intestinalis* grown on glucose, xylo-oligosaccharides, BGX and WAX for 18 hours. (**f**) Cells grown on BGX were washed (PBS buffer ± 1.5 M NaCl) and

xylanase activity was measured in wash and cells fractions to verify localisation of the enzymes. Xylanase activity was measured using the DNS reducing sugar assay and data are triplicates with standard deviations.

Genes encoding an ABC transporter and a multi-modular xylanase are amongst the top upregulated in response to growth of *R. intestinalis* on xylan

To elucidate the genetic basis for growth on xylans, we performed an RNA-seq transcriptional analysis of *R. intestinalis* grown on WAX, BGX, xylose and glucose. Of the 4777 predicted genes, 1–3.5% were highly upregulated (Log2 fold-change > 5) on xylans compared to glucose (Supplementary Table 1), the majority being involved in carbohydrate and energy metabolism. Besides a separate locus encoding a multi-modular xylanase of glycoside hydrolase family 10 (GH10 according to the CAZy classification, http://www.cazy.org[22]), the top genes in the xylan transcriptomes cluster on a single locus (Fig. 2a,b). This locus contains eleven genes including four xylanolytic CAZymes of GH43, GH115, GH8, GH3. Only one (ROSINTL182_08192, LacI type, Pfam 00356) of three transcriptional regulator genes was highly upregulated. Strikingly, the top-upregulated gene in the xylan transcriptomes encodes a solute binding protein (SBP) of an ABC transporter. Furthermore, the genes encoding the permease components of this ABC transporter were amongst the top six upregulated by xylans. Signal peptides were only predicted for the xylanase and the transporter SBP, consistent with extracellular breakdown of xylan followed by capture and uptake of xylo-oligosaccharides by the ABC transporter. The expression and the localization of the transport SBP and the xylanase at the cell surface were corroborated using immunofluorescence microscopy (Fig. 2c). Two additional loci, unique to *R. intestinalis* L1-82, lacking in other *R. intestinalis* strains, were also upregulated albeit markedly less (Supplementary Fig. 2a-d). One of these loci encodes a second cell attached GH10 xylanase, which is also expressed at the cell surface (Supplementary Fig. 4c). The transcriptomic analysis also enabled us to assign the ABC-transporter mediating xylose import and to outline the genes involved in metabolism of xylose, arabinose and glucuronic acid (Supplementary Fig. 2e,f).

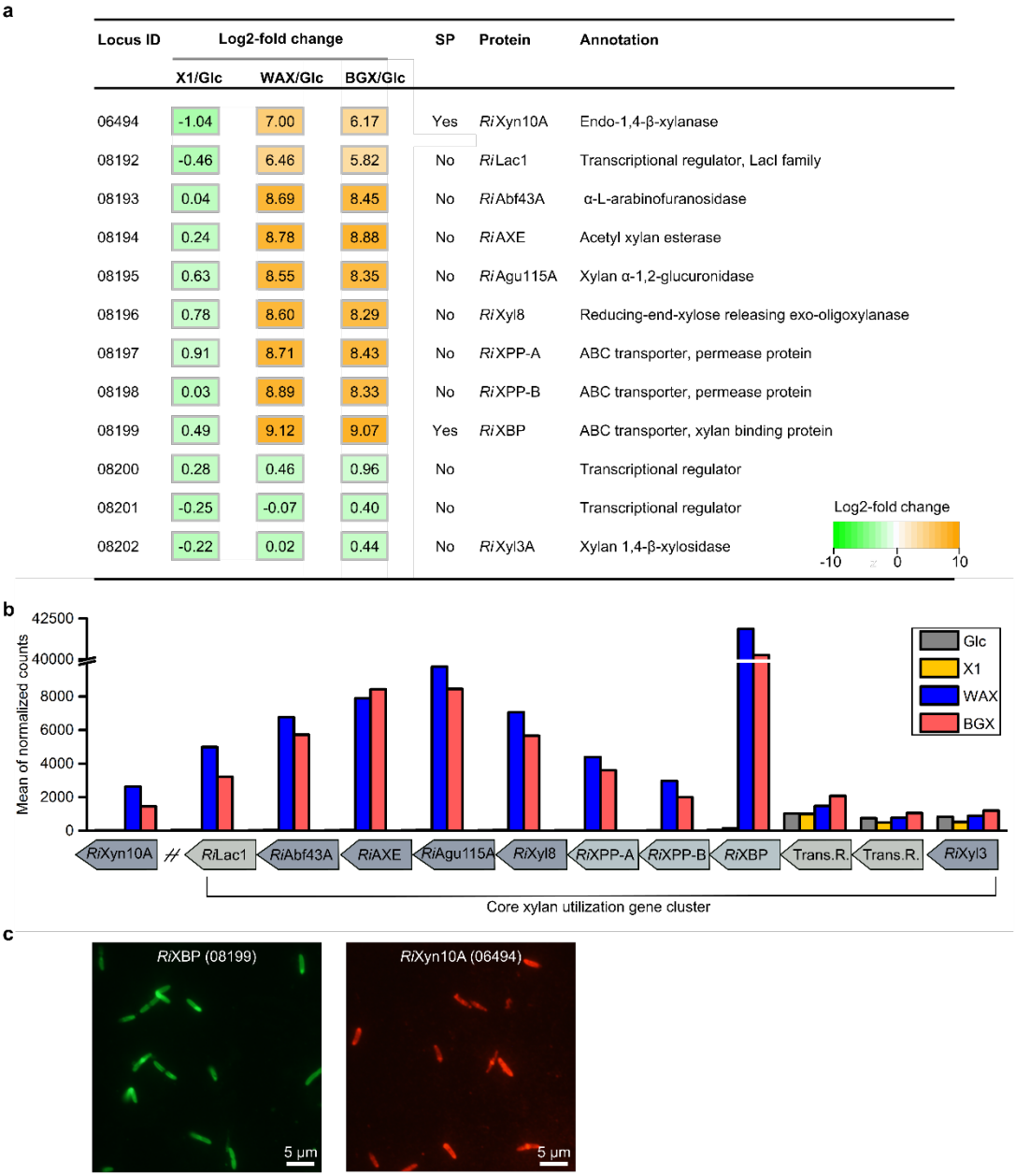


Figure 2 The core xylan utilization apparatus of *R. intestinalis*. (**a**) The RNA-Seq heatmap depicts Log2 fold changes of the top upregulated xylan utilization genes expressed by cells grown on xylose (X1), wheat arabinoxylan (WAX) and birch glucuronoxylan (BGX) relative to glucose (Glc). Formal locus tag numbers ROSINTL182_xxxxx are abbreviated with the last numbers after the hyphen and signal peptides (SP) were predicted using SignalP v.3.0. (**b**) Gene expression depicted as mean of the normalized Deseq2 gene counts for the core xylan utilization genes shown in (a). (**c**) Extracellular localization of *Ri*XBP and *Ri*Xyn10A, the solute binding protein of the xylo-oligosaccharide specific ABC transporter and the xylanase, respectively, were visualized by fluorescence microscopy of *R. intestinalis* cells using primary antibodies targeting these two proteins. No auto fluorescence was observed for cells without primary antibody (data not shown).

A new family of binding modules confers extended and dynamic xylan binding to the multi-modular xylanase in *R. intestinalis*

The highly upregulated *Ri*Xyn10A, which is conserved within the *R. intestinalis* species, is one of the largest known xylanases from human gut bacteria (Supplementary Fig. 3). The mature protein of *Ri*Xyn10A comprises an N-terminal unassigned domain (residues 28−165), a xylan binding module of CBM22, a catalytic module of GH10, a tandem repeat of CBM9 xylan binding modules, a bacterial Ig-like domain group 2 (BIG2, pfam02368)[23] and a Listeria-Bacteroides repeat domain (LBR, pfam09479)[24]. The two latter domains likely mediate cell attachment of the enzyme[23–25] in accordance with their positive charge, which is compatible with binding to the negatively charged cell surface (residues 1100-1356, pI>10).

To generate insight into the unique modularity of *Ri*Xyn10A, we characterized the enzyme and truncated versions thereof (Fig. 3a). *Ri*Xyn10A incubated with BGX, WAX and InWAX generated some linear, but mostly decorated oligosaccharides (Fig. 3b,c and Fig. 4). The enzyme was inactive on xylobiose (X2) and showed very low activity on xylotriose (X3) (Supplementary Fig. 4a). By contrast, xylotetraose (X4) and xylopentaose (X5) were hydrolyzed stoichiometrically, revealing the requirement for at least four substrate-binding sub-sites for efficient hydrolysis.

A BLASTP search of the N-terminal unassigned domain (CBMx) against UniProt gave no hits indicating the lack of homologues with assigned function. CBMx confers affinity to xylan as implied from the one-fold K_M increase when this domain was deleted (Fig. 3d). Affinity electrophoresis established CBMx to be a novel xylan-binding module and revealed a 30-fold stronger binding for WAX compared to BGX (Fig. 3e,f and Supplementary Fig. 4c,d). Surface plasmon resonance (SPR) analysis revealed the highest affinity towards xylohexaose (X6) consistent with the presence of a binding cleft large enough to accommodate at least six xylosyl units (Fig. 3e,g and Supplementary Fig. 6a-e). This analysis also indicated specificity to xylan as there was no measurable affinity to mannohexaose (Man6). The relatively low binding affinity to X6 (K_D≈0.5 mM) was corroborated using isothermal titration calorimetry (ITC) (Fig. 3e and Supplementary Fig. 5g,f). Deleting CBMx decreased the average K_D of *Ri*Xyn10A from 128 µM to 65.4 µM (*Ri*Xyn10AΔCBMx) (Supplementary Fig. 5h-k), asserting that at least one or more of other CBMs possess higher affinity compared to the N-terminal new module. Homologues (sequence identity 55−27%) of the new CBM are present mainly in other species belonging to *Clostridium* cluster XIVa (Supplementary 4e), which merits the assignment of these modules into a new CBM family.

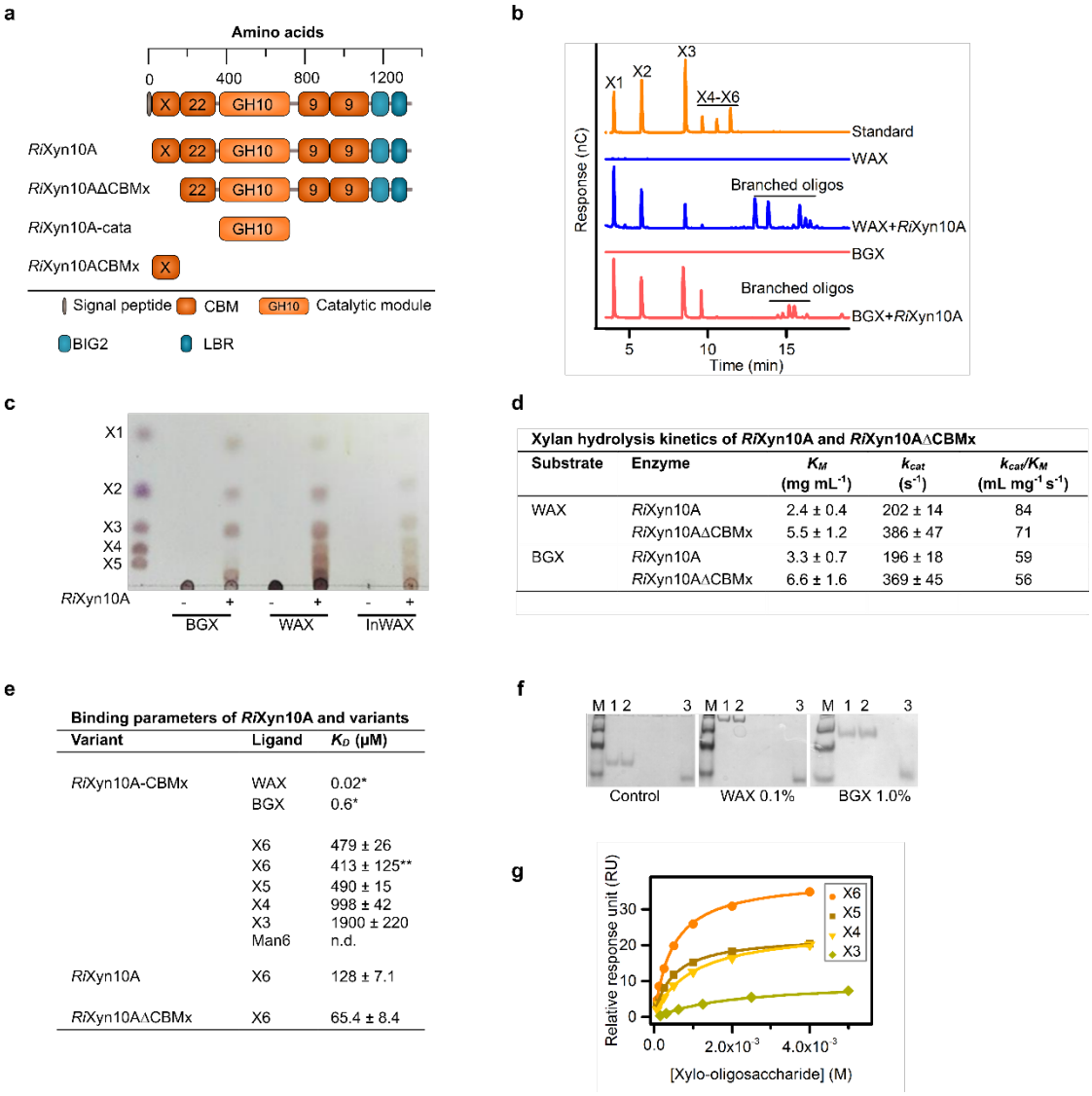


Figure 3 A novel low affinity xylan binding module mediates extended xylan binding to the xylanase *Ri*Xyn10A. (a) Domain organization of *Ri*Xyn10A and truncated variants. Carbohydrate binding module (CBM), novel CBM (CBMx), bacterial Ig-like domain group 2 (BIG2), Listeria-Bacteroides repeat domain (LBR). (**b,c**) Xylanase activity of *Ri*Xyn10A on WAX and BGX assayed by HPAEC-PAD and thin layer chromatography, respectively. (**d**) Hydrolysis kinetic parameters of *Ri*Xyn10A and *Ri*Xyn10AΔCBMx towards WAX and BGX. Data are means of triplicates with standard deviations. (**e**) Binding parameters of *Ri*Xyn10A and variants towards oligosaccharides. Dissociation constants (K_D) determined by surface plasmon resonance (SPR) are means of a duplicate with the standard deviations. * K_D (mg mL^{-1}) from affinity electrophoresis (AE), and ** K_D from isothermal titration calorimetry (ITC). (**f**) Binding of *Ri*Xyn10ΔCBMx to the negative control (no polysaccharide), WAX or BGX xylans analyzed using AE. Lane 1+2; *Ri*Xyn10A-CBMx (3.0 µg), Lane 3; ß-lactoglobulin negative control, M; marker (1.5 µg). (**g**) Binding isotherms of *Ri*Xyn10ΔCBMx binding to xylo-oligosaccharides. Solid lines are fits of a one binding site model to the SPR sensograms.

Preference of the binding protein of the ABC transporter that mediates uptake of xylan oligomers in *R. intestinalis*

We showed above that the action of xylanases produces mainly xylo-oligosaccharides decorated with arabinosyl and 4-*O*-methyl-glucuronosyl. No oligosaccharides were detectable (HPAEC-PAD analysis, data not shown) in spent supernatants from *R. intestinalis* growth on xylan, suggesting efficient uptake of oligomeric products. The transcriptional analysis (Fig. 2a) identified an ABC transporter likely to mediate the uptake of the xylo-

oligosaccharides hydrolysis products of *Ri*Xyn10A from WAX and BGX. The preference of SBPs associated with oligosaccharide-specific ABC transporters has been shown to correlate well to the uptake preference of bacteria[26,27]. We measured the affinity of *Ri*XBP, the SBP of the upregulated ABC transporter, on a range of xylo-oligosaccharide ligands (Table 1 and Supplementary Fig. 6). The preferred un-substituted ligand was X5 followed by X4, and the affinity decreased steeply for smaller or larger oligomers. Internal arabinosyl decorations (AX4) appeared to be preferred based on the 2.4-times higher affinity compared to the un-substituted X4. The tolerance and recognition of arabinosylated ligands is in agreement with the good growth on WAX. These results suggest that *Ri*XBP is selective in capturing internally branched xylo-oligosaccharides with a xylose backbone of 4–5 xylose residues.

Table 1: Binding energetics of the transport protein *Ri*XBP to xylo-oligosaccharides determined by ITC

Ligand	K_D (μM)	N_0	ΔH (kcal/mol)	TΔS (kcal/mol)	ΔG (kcal/mol)
X6	112.7 ± 7.5	1.19 ± 0.14	-9.01 ± 1.3	-3.6	-5.4
X5	10.3 ± 1.5	0.86 ± 0.01	-13.54 ± 0.3	-6.7	-6.8
X4	16.5 ± 2.6	0.68 ± 0.02	-12.8 ± 0.4	-6.3	-6.5
X3	225.7 ± 14.5	0.58 ± 0.23	-21.1 ± 9.5	-16.1	-5.0
X2	n.d.				
AX3	215.5 ± 95.2	0.26 ± 0.04	-44.3 ± 7.1	-39.4	-4.9
AX4	6.8 ± 1.2	0.58 ± 0.01	-12.3 ± 0.2	-7.0	-5.3

Data are means of a duplicate experiment with standard deviations. n.d. indicates no binding was observed. AX3 is an arabinotriose with a non-reducing end arabinosyl and AX4 is an arabino-xylotetraose with an arabinosyl decoration at the penultimate position from the non-reducing end (see Supplementary Fig. 6h,i).

R. intestinalis degrades internalized decorated xylo-oligosaccharides by the concerted action of three hydrolases and a novel family of acetyl esterases.

Xylo-oligosaccharides are degraded in the cytoplasm subsequent to their uptake. To gain insight into intracellular xylan-oligomer breakdown, we produced and characterized the α-glucuronidase *Ri*Agu115A (GH115), the α-L-arabinofuranosidase *Ri*Abf43A (GH43), two xylosidases *Ri*Xyl8 (GH8) and *Ri*Xyl3A (GH3) as well as *Ri*AXE (ROSITNL182_08194, GenBank accession EEU99941.1) from the core xylan utilization locus.

*Ri*Agu115A released 4-*O*-methyl-glucuronic acid (MeGlcA) from glucuronoxylans (BGX and BeGX) and from BGX pretreated with *Ri*Xyn10A (Fig. 4a and Supplementary Fig. 7a-c). The k_{cat}/K_m of *Ri*Agu115A was 16-fold higher on glucuronoxylan hydrolysate compared to intact glucuronoxylan (Supplementary Fig. 7c), indicating that *Ri*Agu115A preferentially accommodates glucuronoxylo-oligosaccharides, consistent with the intracellular localization of this enzyme. This enzyme also cleaves MeGlcA decorations at the penultimate xyloxyl to the reducing end, generated using a GH30

glucuronoxylanase (Supplementary Fig. 7b), but its activity was blocked by the presence of acetylations (Fig. 4d).

*Ri*Abf43A is an α-L-arabinofuranosidase that exclusively releases arabinose from WAX (Fig. 4a). Kinetic analysis towards WAX and arabino-xylotetraose (AX4) (Supplementary Fig. 7d) revealed recognition of internal arabinosyl substitutions, with a 13-fold increase in k_{cat} for oligosaccharides consistent with the intracellular localization.

Both *Ri*Xyl8 and *Ri*Xyl3A generated xylose from xylo-oligosaccharides, but lacked activity towards xylan (Supplementary Fig. 7g-k). *Ri*Xyl3A degraded xylo-oligosaccharides completely into monosaccharides, while *Ri*Xyl8 was inactive towards X2. Reduction of xylo-oligosaccharides with $NaBH_4$ abolished the activity of *Ri*Xyl8 assigning it as a reducing-end β-xylosidase[28] (Supplementary Fig. 7i), in contrast to *Ri*Xyl3A that recognizes non-reducing xylosyl moieties and maintains activity on reduced xylo-oligosaccharides. Thus, the concerted and overlapping activities of these enzymes (Supplementary Fig. 7) results in rapid depolymerization of arabinosyl and MeGlcA decorated xylo-oligosaccharides.

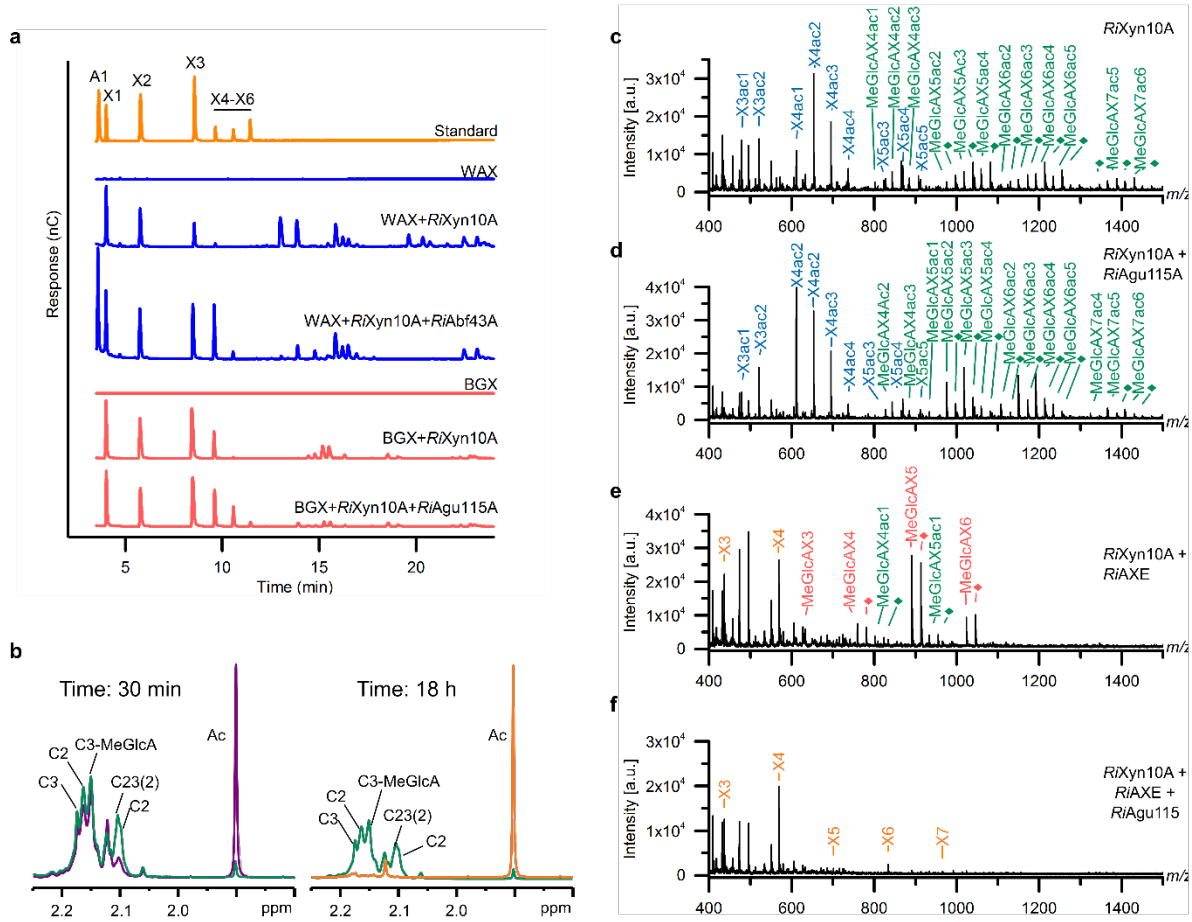


Figure 4 Intracellular xylo-oligosaccharide depolymerization (a) α-Glucuronidase and α-L-arabinofuranosidase activity on WAX and BGX for *Ri*Agu115A and *Ri*Abf43A, respectively, based on HPAEC-PAD analysis. (**b**) Time-resolved NMR for *Ri*AXE enzymatic deacetylation of acetylated birch glucuronoxylan (AcBGX) treated with *Ri*Xyn10A and *Ri*Agu115A. Deacetylation time course for the first 30 min and after 18 h (green 0 min, purple 30 min, orange 18 h). All verified signals with 2-*O*-acetylation decreased faster in the initial phase of the reaction. The proton spectra of the acetylated region shows nearly

complete deacetylation of the sample after 18 h. The signal at 2.13 ppm is likely attributed to another acetylated sugar residue. Acetyl groups are designated as: C2, 2-*O*-acetylated xylose; C3, 3-*O*-acetylated xylose, C23, 2,3-di-*O*-acetylated xylose; C3-MeGlcA; 4-*O*-methylglucuronic acid 2-*O*-substituted and 3-*O*-acetylated xylose; C23(2); signal for the 2-*O*-acetylated of C23. The assignment of the acetylated sugar signals were based on homo and heteronuclear NMR correlation experiments (Supplementary Fig. 8) (**c-f**) Hydrolysis products from AcBGX by (**c**) *Ri*Xyn10A, (**d**) *Ri*Xyn10A and *Ri*Agu115A, (**e**) *Ri*Xyn10A and *Ri*AXE, (**f**) *Ri*Xyn10A, *Ri*Agu115A and *Ri*AXE. Enzyme action was analyzed by MALDI-TOF MS; Xylo-oligosaccharides decorated with acetyl and methylglucuronic acid are in green, acetyl in blue, methylglucuronic acid in red, no sidechains in orange. Di-sodium adducts of a methylglucuronic acid decorated oligomers (diamonds) are colored as their corresponding single sodium adducts.

*Ri*AXE, which was un-assigned based on lack of hits in a BLASTP search of UniProt, was highly upregulated on xylans (Fig. 2a). This enzyme possesses the conserved residues in the SGNH lipases-esterases superfamily (Pfam cd00229), which also includes CAZy carbohydrate esterase families CE2, CE3, CE12 and CE16. We established that *Ri*AXE is an acetyl esterase, but low sequence identities to these families (<12%) merit assigning *Ri*AXE into a new carbohydrate esterase family. Indeed homologues of this enzyme are encoded by several *Clostridium* cluster XIVa strains from the human gut and by a range of Firmicutes (Supplementary Fig. 8i).

Assaying *Ri*AXE activity towards AcBGX oligosaccharides (generated with *Ri*Xyn10A) using NMR revealed efficient deacetylation of both C2 and C3, but with a preference for C2 decorations (Fig. 4b and Supplementary 8). Analysis of the deacetylation by MALDI-ToF MS left a single acetyl group on the AcBGX oligosaccharides (Fig. 4e). Inclusion of *Ri*Agu115A in this reaction resulted in complete deacetylation (Fig. 4f) suggesting that the presence of MeGlcA decorations protects acetylations in the proximity of the MeGlcA unit. Analysis of the deacetylation rates also unveiled the concerted action with *Ri*Agu115A and the preference to hydrolysates of *Ri*Xyn10A rather than intact xylan (Supplementary Fig. 8c,d). *Ri*AXE specifically recognizes acetylations on xylosyl units based on lack of activity on acetylated chitin and very low activity on acetylated mannan and cellulose monoacetate (Supplementary Fig. 8h). Taken together, the results showed that *Ri*AXE is an efficient xylan specific representative of a new acetyl esterase family.

In summary of the biochemical characterization presented above, we propose a model for the uptake and degradation of diet-derived acetylated arabinoxylan and glucuronoxylan by *R. intestinalis* L1-82 (Fig. 5a).

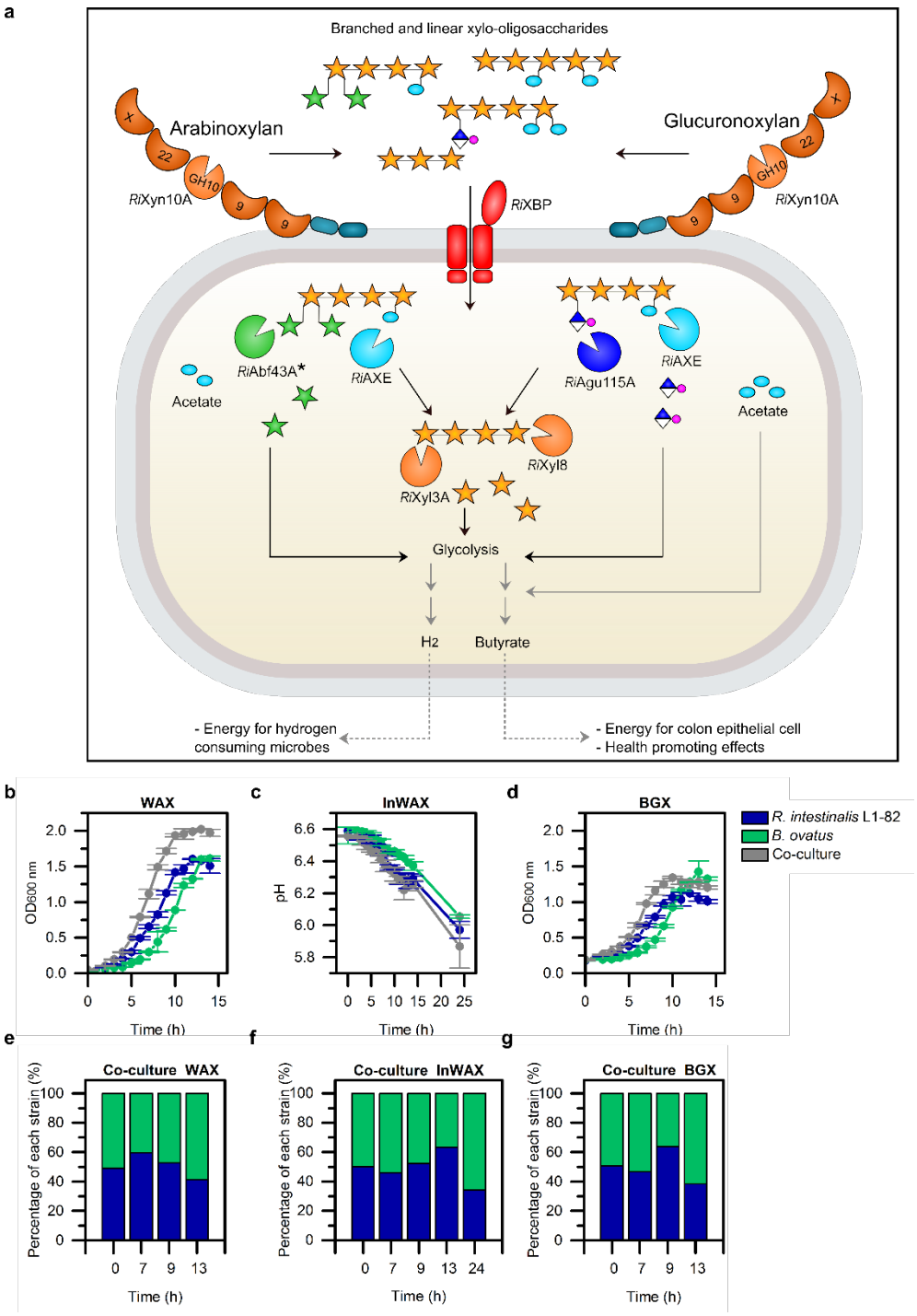


Figure 5 Model for xylan utilization by *R. intestinalis* and competition assay with *Bacteriodes ovatus* (a) *Ri*Xyn10A on the cell surface efficiently captures diet-derived acetylated arabinoxylan and acetylated glucuronoxylan by its CBMs and hydrolyzes it into linear and branched xylo-oligosaccharides, which are subsequently captured by *Ri*XBP for uptake into the cytoplasm. Internalized xylo-oligosaccharides are debranched and hydrolyzed into monosaccharides and acetate. Xylose and arabinose are converted to xylulose 5-phosphate before entering the pentose phosphate pathway, whereas methyl-glucuronic acid is converted to 2-oxo-3-deoxygalactonate 6-phosphate. These precursors enter glycolysis, which generates pyruvate, some of which is used to synthesize butyrate[29] that is externalized. The asterisk next to *Ri*Abf43A indicates that the enzyme is able to hydrolyze both α-1,2 and α-1,3 linked L-arabinose. Black solid arrows show steps established or confirmed in this study. Grey solid arrows indicate steps described in literature. Grey dashed arrows indicate that H_2 and butyrate are externalized by unknown mechanisms. To make the model more general for the *R. intestinalis* species, the second less upregulated extracellular xylanase *Ri*XynB, unique for the L1-82 strain, is not included in the model, although it is expressed at the cell surface. (**b-d**) Growth of monoculture and co-cultures of *R. intestinalis* and *B. ovatus* on WAX, InWAX

and BGX. Data are means of a triplicate with standard deviations. (**e-g**) Time course relative abundance during growth of co-cultures on xylans determined by qPCR. All data are means of a biological triplicate.

R. intestinalis competes with *Bacteriodes* for xylans

The growth potential of *R. intestinalis* was compared with *Bacteroides ovatus,* another efficient xylan degrader[20], by observing growth of individual cultures and in co-culture. Both strains displayed similar growth on xylan as carbon source (Fig. 5b-d and Supplementary Fig. 9a,b). In direct competition, both strains appeared to grow equally well, with no pronounced difference after 13-24 hours growth (Fig. 5e-g). The results indicate that *R. intestinalis* L1-82 is an efficient primary degrader of xylan that is able to compete with *B. ovatus.*

Discussion

The human gut is dominated by bacteria from two phyla: the Gram positive Firmicutes and the Gram-negative Bacteriodetes. Firmicutes are generally regarded as metabolic specialists, while Bacteroidetes (mainly from the *Bacteroides* genus) are regarded as generalists based on narrow versus broad glycan utilization capabilities, respectively[30]. The size and diversity of encoded CAZymes frequently reflects these metabolic labels. Although this generalisation applies to *R. intestinali*s, based on the relatively limited glycan growth profiles[5], this species possesses distinctively higher number of CAZymes than most known clostridial Firmicutes of the HGM[16]. *R. intestinali*s has been proposed as a key xylan degrader in the human gut along with specific species of *Bacteroides*[17,18]. Growth and enumeration of *R. intestinalis* on dietary xylans including wheat bran is reported both *in vitro* and *in vivo*[31,32]. Insight is lacking, however, on the preferences and the molecular machinery evolved by *R. intestinalis* to target distinct xylans as compared to species of *Bacteriodes*. In this study, we present a model that explains the molecular basis for the utilization of xylan by *R. intestinalis* L1-82 as a representative for prevalent butyrate producing clostridia (Fig. 5a). Our data establish that *R. intestinalis* is truly a primary degrader that is equipped with a highly efficient machinery for utilization of complex dietary xylans, including insoluble arabinoxylan from cereals. Identified key components of the *R. intestinalis* xylan utilization strategy include a multi-modular extracellular xylanase and an ABC transporter, which confer the capture, breakdown and internalization of substituted xylan oligomers. In the cytoplasm, internalized xylo-oligosaccharide are depolymerized without loss to competing species. We demonstrate for the first time the ability of *R. intestinalis* to grow on acetylated xylan, which reflects an adaptation to this abundant decoration in dietary xylans (Fig. 1b). Acetylated xylo-oligosaccharides could be metabolized after internalization due to an intracellular novel esterase family capable of removing C2, C3 and double acetylations (Fig. 4b and Supplementary Fig. 8).

The extracellular multi-modular xylanase *Ri*Xyn10A, the ABC transporter and enzymes conferring cytoplasmic breakdown of xylan oligomers were assigned as the core xylan utilization apparatus of *R. intestinalis* (Fig. 2a,b). This assignment was based on i) conservation of this apparatus within the *Roseburia* species, ii) highest transcriptional upregulation of the encoding genes on xylan (Fig. 2a), and iii) biochemical data from the present study. The two additional xylan-upregulated loci in *R. intestinalis* L1-82 (Supplementary Fig. 2) are lacking in *R. intestinalis* XB6B4 and *R. intestinalis* M50/1, both being able to grow on xylan[31]. The activity and expression of the xylanase *Ri*Xyn10B, encoded by one of these auxiliary loci (Supplementary Fig. 4g), supports the participation of more than one locus in xylan breakdown in *R. intestinalis* L1-82. The structural diversity of naturally occurring xylans may justify the acquisition and deployment of multiple loci to expand the targeted substrate range[20] and enforce the specialization on this abundant metabolic resource.

Our data support the role of the *R. intestinalis* core xylanase *Ri*Xyn10A in mediating the capture and breakdown of arabino- and glucuronoxylan (Fig. 1 and Fig. 3). This enzyme possesses four CBMs from two known and one novel xylan-binding families, representing the most complex modular organisation of HGM xylanases (Fig. 3a and Supplementary Fig. 3b). This organization is invariant within the *R. intestinalis* species, while other *Clostridium* XIVa taxa possess simpler enzymes lacking one or more of the *Ri*Xyn10A CBMs. The N-terminal CBMx of *Ri*Xyn10A displays approximately 7-fold lower affinity than the enzyme variant lacking this module (Fig. 3e). These data merit assigning this module into a novel low-affinity xylan-specific CBM family. Despite its relatively low affinity, CBMx is highly selective to arabinoxylan and clearly contributes to the overall affinity of the enzyme (Fig. 3e). Low-affinity CBMs maybe a trade-off between affinity, potentiated by multivalent cooperative

substrate binding, and reduction of turn-over due to the energetic penalty of breaking the bonds to the bound substrate during displacement from the active site (*i.e* maintenance of a relatively high k_{cat}/k_{off} ratio[33]). The extended binding mediated by the CBMs of *Ri*Xyn10A seems to confer an advantage in the capture and prolonged contact of the enzyme with xylan. Deletion of the binding modules caused a substantial decrease in the affinity of the catalytic module towards WAX and BGX (loss of curvature and deviation from Michaelis-Menten kinetics, data not shown). Multiplicity and variability of CBMs seem to be a signature of extracellular enzymes from butyrate producing Firmicutes[34,35]. By contrast, *Bacteriodes* members possess simpler outer-membrane anchored GH10 xylanases with an inserted tandem CBM4 repeat within the catalytic module[21]. Xylan capture by *Bacteriodes*, however, is additionally orchestrated by moderate affinity (K_D≈60 µM) xylan binding proteins that protrude away from the cell surface to facilitate binding[20].

R. intestinalis was able to compete with *B. ovatus* for soluble and insoluble xylans during the log-phase (Fig. 5e-g). *R. intestinalis* has been reported to be associated to insoluble xylans, including wheat bran, while species of *Bacteriodes* were more enriched in the solubilised xylan fractions[17,36]. The extended binding mediated by *Ri*Xyn10A may play an important role in the association to insoluble substrates. Indeed, the expression of this enzyme appeared similarly high in the mono- and mixed xylan cultures with *B. ovatus* (Supplementary Fig. 9d). These observations are different from the reported down-regulation of hydrolases by *Eubacterium rectale*, which is close taxonomic relative to *Roseburia*, during co-growth with *Bacteriodes thetaiotamicron* on a fiber rich diet in previously germ-free mice[37].

The gene encoding the binding protein (*Ri*XBP) of the ABC transporter that confers xylo-oligomer uptake in *R. Intestinalis* was the most upregualted in the xylan transcriptomes, reflecting the paramount role of oligosaccharide capture and transport in the densely populated ecological niche of the gut. The narrow preference of this protein for decorated backbone of 4−5 xyloxyl units aligned with the products of *Ri*Xyn10A (Fig. 3b and Fig. 4c). The affinity and size preference of *Ri*XBP were found to be very different from the corresponding protein from *Bifidobacterium*[27], which binds shorter XOS. Importantly, striking differences in binding affinities and prefernece are observed when *Ri*XBP is compared to the SusD-like xylan-binding counterpart from *Bacteriodes*. Indeed, both SusD-like proteins from *B. ovatus*, which mediate efficient capture and internalization of xylan-oligomers ≥ X6 by SusC TonB-dependent permeases, displayed no measurable binding to X4 and X5[20], the preferred ligands of *Ri*XBP. These differencial transport protein preferences are likely to be instrumental in establishing competitive uptake profiles to a select oligosacchrides of specific sizes and decorations for each taxon.

Our study highlights the molecular apparatus that *R. intestinalis*, as a model *Clostridium* group XIVa Firmicutes, has evolved to compete for abundant dietary glycans with other commensal gut bacteria. Strikingly complex enzymes with multiple ancillary modules mediate multivalent substrate capture and breakdown. Highly over-expressed ABC transporters mediate efficient capture and uptake of xylan oligomers that are conspicuously different from those of competing taxa. Based on these findings we propose that the differentiation of glycan capture and uptake preferences represent an adaptation strategy to minimize competition for major dietary fibers by different human gut taxa.

Taking these differential uptake preferences into account may guide the design of better therapeutic strategies aiming at restoring or boosting specific taxonomic groups in a safer and more controlled manner than currently practiced in faecal transfer therapy.

Methods

Chemicals

All chemicals were of analytical grade. Birchwood glucuronoxylan (BGX), beechwood glucuronoxylan BeGX), corncob xylo-oligosaccharides (CCXOS) and xylose were from Carl Roth (Karlsruhe, Germany). Cornbran xylan (CBX) was a kind gift from Dr. Madhav, Yadav, United States Department of Agriculture, Agricultural Research Service. Soluble wheat arabinoxylan (low viscosity 10 centiStokes (cSt)) (WAX), insoluble wheat arabinoxylan (high viscosity 48 cSt) (InWAX), xylobiose through to xylohexaose (X2−X6), arabinoxylotriose (AX3), arabinoxylotetraose (AX4) and mannohexaose (Man6) were from Megazyme (Wicklow, Ireland). D-Glucuronic acid was from Sigma Aldrich (St. Louis, MO, USA). L-arabinose was from VWR International Ltd (Lutterworth, Leicestershire, UK). Xylo-oligosaccharides Longlive 95P (XOS) were from Shandon Longlive Bio-technology (Shandong, China). Acetylated birchwood glucuronoxylan (AcBGX), acetylated aspen glucuronoxylan (AcAGX), acetylated spruce galactoglucomannan (AcSGGM) were prepared with steam explosion as described in[38]. Cellulose acetate was a kind gift from Alexander Deutschle, University of Hamburg, Germany. Acetylated chitin-oligosaccharides was prepared according to a well established procedure from chitinase degradation of a highly acetylated chitosan[39].

Growth experiments and RNA-seq transcriptional analysis

R. intestinalis DSM 14610 was grown in a Whitley DG250 Anaerobic Workstation (Don Whitley, UK) in YCFA medium[13] supplemented with autoclaved-sterilized 0.5% (w/v) carbohydrates. Cultures (5mL) were grown in triplicates and $OD_{600\ nm}$ and pH were measured to assess bacterial growth until the stationary phase was reached. pH measurements were only performed for insoluble substrates. Growth rates were calculated from the exponential growth phase.

For the RNA-seq analysis, total RNA was extracted at mid- to late log phase ($OD_{600\ nm}$ = 0.5−0.7) from biological triplicates cultures (10 mL) grown in YCFA supplemented with 0.5% (w/v) glucose, xylose, WAX or BGX. Cells were harvested (4000 g, 5 min, room temperature) and the pellets were frozen at -80°C until RNA extraction. The RNA was extracted using the RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol after enzymatic lysis followed by mechanical disruption of the cells. A DNase treatment was included to ensure removal of DNA. The purity and quantity of the extracted RNA were assessed by an Agilent 2100 Bioanalyzer (Agilent Technologies, UK). Removal of ribosomal RNA and library construction for RNAseq were performed using the ScriptSeq™ Complete Kit (Epicentre). High-throughput sequencing was performed in a single lane in paired end reads on an Illumina Hiseq 4000 platform at BGI (Copenhagen, Denmark). In total, 400 million paired-end reads were obtained and the read quality was assessed by FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The R1 reads were chosen for downstream analysis. Adaptor trimming and de-multiplexing was performed using custom python scripts (based on the Biopython SeqIO module[40]) and the FASTX-Toolkit v0.0.13.2 (http://hannonlab.cshl.edu/fastx_toolkit/). Reads were further trimmed with fastx_trimmer and subsequently, filtered with fastq_quality_filter with minimum quality score 30 (-q 30) where 95% of base-pairs meet the minimum quality score (-p 95). The resulting reads were kept if longer than 20 bps (-m 20). The *R. intestinalis* L1-82 reference genome and genome annotations are based on assembly *GCA_000156535.1_ASM15653v1*, obtained from NCBI (ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/bacteria/Roseburia_intestinalis/). Reads were mapped to the reference genome using Tophat2[41,42], and gene counts were determined with HTseq[43]. Differential gene expression was performed using DeSeq2 in R[44].

Xylanase activity measurements on whole cells

Cell-associated xylanase activity was determined by growing *R. intestinalis* cells in 800 µL YCFA containing 0.5% (w/v) xylo-oligosaccharides, WAX, BGX or glucose for 15 hours. Cells were harvested (4000 g, 5 min, room temperature), resuspended in phosphate-buffered saline (PBS) to a final $OD_{600\ nm}$ = 0.3 and the xylanase activity assayed using the DNS assay as described below. To determine the effect of high ionic strength on the localization of xylanase activity, *R. intestinalis* cells were grown in 6 mL YCFA containing 0.5% (w/v) BGX for 15 hours. Subsequently, the culture was divided into two 3 mL aliquots and harvested as described above. The cell pellets were resuspended in 300 µL PBS with or without 1.5 M NaCl. The suspensions were spun down and both pellets and supernatants were collected, washed with excess PBS and resuspended in 300 µL PBS. The xylanase activity of cells and wash fractions was assayed using the DNS assay.

Expression and purification of *R. intestinalis* proteins mediating xylan utilization

Open reading frames of the proteins without signal peptide, as predicted by SignalP v.3.0 (http://www.cbs.dtu.dk/services/SignalP-3.0), were amplified from *R. intestinalis* DSM 14610 genomic DNA using specific primers (Supplementary Fig. 1). The amplicons were cloned into the EcoRI and NcoI restriction sites of a pETM-11 vector (kind gift from Dr. Gunter Stier, EMBL, Center for Biochemistry, Heidelberg, Germany[45] or the XhoI and NcoI restriction site of a pET28a(+) vector (Novagen, Darmstadt, Germany) using In-Fusion cloning (Takara) to express proteins as fusions with either cleavable N-terminal His_6 tags or a C-terminal ones, respectively. Constructs were transformed into *Escherichia coli* DH5α (Novagen) and verified by sequencing. Recombinant plasmids were transformed into *E. coli* BL21(DE3) (Novagen) for recombinant protein production, which was carried out in LB medium supplemented with kanamycin (50 µg mL^{-1}). Shake flasks were inoculated by an overnight culture and grown at 37°C until $OD_{600\ nm}$≈0.5, then the cells were induced with 0.1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated for 16−40 hours at 16°C before harvest by centrifugation (10,000 g, 20 min, 4°C). Cell pellets were resuspended in buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 10 mM imidazole, 10% glycerol) and disrupted by a single passage in a high pressure homogenizer (Standsted Fluid Power, Essex, UK) at 1000 bar. The recombinant proteins were purified from the cell-free supernatants by affinity chromatography using 5 mL HisTrap HP columns (GE Healthcare, Uppsala, Sweden) using a standard protocol, which was followed by size exclusion chromatography using Hiload Superdex columns (GE Healthcare) installed on an ÄKTA-AVANT chromatograph (GE Healthcare) at 4°C. Protein purity was determined by SDS-PAGE and protein concentrations were measured spectrophotometrically and calculated from the theoretical molar extinction coefficients.

Enzymatic activity assays

Enzymatic assays were carried out in a 50 mM HEPES 0.005% (v/v) Triton X-100, pH 7.0 standard assay buffer unless otherwise stated.

The hydrolysis kinetics of full-length or truncated xylanases (10−200 nM) were assayed towards 1−9 mg mL^{-1} of BGX, WAX or InWAX (37°C, 900 µL, 12 min). Initial hydrolysis rates were determined by removing 200 µL aliquots every third minutes and quenching the reaction in 300 µL 3,5-dinitrosalicylic acid (DNS) reagent[46]. Next the samples incubated for 15 min at 90°C followed by $A_{540\ nm}$ measurement in 96 microtitre plates. Xylose was used as a standard (0−2.5 mM). Xylanase activity was assayed for *R. intestinalis* cells washed with PBS ± 1.5 M NaCl, and wash-fractions, with the following modification:

180 µL of 1% (w/v) BGX was incubated with 20 µL cell suspension or wash-fraction for 4 hours. The reaction was stopped by adding 300 µL DNS stop solution and the end point specific activity was measured as described above.

α-Glucuronidase hydrolysis kinetics were analyzed on 1−9 mg mL^{-1} BeGX or a hydrolysate thereof (prepared by incubation with 4 mM *Ri*Xyn10A xylanase for 15 hours at 37°C followed by heat inactivation). The initial rates of (*O*-methyl)-D-glucuronic acid release were measured using a coupled enzymatic assay (Megazyme). Reactions (770 µL) were incubated for 2 min at 37°C with 10−180 nM enzyme with intermittent removal of 175 µL aliquots every 15 s into 125 µL 1 M Tris pH 10 to quench the reaction. This was followed by mixing 270 µL of the stopped reaction with 45 µl of the NAD^+ and uronate dehydrogenase reagents. The conversion of NAD^+ to NADH was measured at A_{340} nm. Glucuronic acid was used as standard (0−500 µM).

β-Xylosidase kinetics of *Ri*Xyl8 and *Ri*Xyl3A were determined towards xylobiose (X2) through to xylohexaose (X6) (0.5-12 mM) in McIlvaine buffer pH 6.8 (10 mM citric acid and 20 mM sodium phosphate) as described in[47,48]. Reactions (350 µL) were incubated for 12 min at 37°C with 36−78 nM *Ri*Xyl3A or 2.4 nM *Ri*Xyn8. Aliquots of 50 µL were removed every 2 minutes and stopped in 250 µL *p*-bromoaniline (2% w/v) in glacial acetic acid with thiourea (4% w/v). The stopped reactions were incubated in darkness for 10 min at 70°C, followed by incubation at 37°C for 1 hour before measuring $A_{520\ nm}$. The concentration of released pentoses was determined using a xylose standard (0−5 mM).

The α-L-Arabinofuranosidase activity for *Ri*Abf43A was assayed in McIlvaine buffer pH 6.8 (10 mM citric acid and 20 mM sodium phosphate) using a coupled enzymatic L-arabinose/D-galactose assay (Megazyme) towards WAX (1−24 mg mL^{-1}). Reactions (75 µL) were incubated for 12 min at 37°C with 0.4−1.7 µM enzyme. Aliquots of 15 µL were removed every 2 min, and the enzyme was inactivated (10 min, 90°C) and thereafter 10 µL of this solution were mixed with 10 µL of the provided NAD^+, 20 µL of provided assay buffer and 2 µL galactose mutaotase/β-galactose dehydrogenase mix. The formation of NADH was measured as above. Arabinose was used as standard (0−5 mM).

The acetyl esterase specific activity of *Ri*AXE was determined in 250 µL reactions containing *para*-nitrophenyl-acetate (4 mM) and 0.14 µM enzyme. $A_{405\ nm}$ was measured every 60 s for 10 minutes at 37°C in a microtiter plate reader and *p*NP (0−1 mM) was used as standard. The specific activity was determined in units (U/mg), where a U is defined as the amount of enzyme that produces 1 µmol of *p*NP min^{-1}.

Kinetic parameters were calculated by fitting the Michaelis-Menten equation to the initial rate data using Graph Pad Prism 7. The catalytic efficiency k_{cat}/K_m, determined from the slope of the normalized initial rate (V_0/[E]) in the Michaelis-Menten plot, is reported when saturation was not attained. All experiments were performed in triplicates.

Action patterns of individual and mixtures of xylanolytic enzymes

Hydrolysis of xylan and xylo-oligosachharides was performed at 37°C for 15 hours in the standard assay buffer used above. Oligo-saccharide hydrolysates, to assay the sequential action of the debranching xylanolytic enzymes, were generated using *Ri*Xyn10A, which was separated by ultrafiltration (3 KDa cutoff) before the addition of debranching enzymes. The hydrolysis profiles were analyzed as detailed below. To verify the mode of reducing-end attack of *Ri*Xyl8, 30 mg XOS in standard assay buffer were reduced by $NaBH_4$ (1M in 100 µM NaOH). A total of 200 µL of the $NaBH_4$ was added dropwise to 800 µL of the xylo-oligossaccharides solution, which was kept on ice. As control 100 µM

NaOH was added to an 800 µL xylo-oligossaccharides solution. The mixture was incubated 1 hour at room temperature, then quenched by 400 µL 1 M acetic acid and diluted 10x in assay buffer.

Matrix-assisted laser desorption-ionization (MALDI)

Oligomeric products were analyzed with an Ultraflex MALDI ToF/ToF instrument (Bruker Daltonics, Bremen, Germany). The samples were applied with 2,5-dihydroxybenzoic acid (DHB) as matrix to a MTP 384 ground steel target plate (Bruker Daltonics). All spectra were obtained in positive reflection mode and processed using Bruker flexAnalysis 3.3.

Thin layer chromatography (TLC)

Aliquots of 1 µL of enzymatic reactions were spotted on a silica gel 60 F254 plate (Merck, Germany). The chromatography was performed in a butanol:acetic acid:water (2:1:1 v/v) mobile phase. The plates were dried at 50°C and carbohydrate hydrolysis products were visualized by spraying with a 5-methylresorcinol:ethanol:sulfuric acid (2:80:10 % v/v) developer and tarred briefly with an airfan at 350°C until bands appeared.

High performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD)

The release of xylo-oligossaccharides and monosaccharides was analyzed by HPAEC-PAD on an ICS-3000 (Dionex, CA, USA) using a 3x250mm CarboPac PA1 column, a 3x50 mm guard column and 10 µL injections. Xylo-oligosaccharide and standards were eluted with a two-step linear gradient; 0–20% sodium acetate gradient in 0.1 mM NaOH 0–30 min and 30–100% sodium acetate gradient in 0.1 mM NaOH 30–41 min with a flowrate at 0.25 mL min^{-1}. Monosaccharides and standards (0.1 mg mL^{-1}) of galactose, arabinose, glucose and xylose were eluted with with 1 mM KOH for 35 min at 0.25 mL min^{-1}.

NMR spectroscopy

To reduce the interference of the water signal, both oligomer substrate (AcBGX and AcSpruce mannan) were dissolved in 99.9% D_2O (Sigma-Aldrich) and lyophilized. Similar, 10 mL 50 mM phosphate buffer pH 7.0 was lyophilized and the powder was redissolved in 10 mL 99.9% D_2O. For the time-resolved NMR recordings: 4 mg oligomer substrate was dissolved in 500 µL 50 mM phosphate buffer pH 7.0 (99.9% D_2O) and transferred to a 5 mm NMR tube. The sample was preheated in the NMR spectrometer for ~10 min. Hereafter all recording parameters were set prior to the time-resolved NMR experiment. 2.5 µL of enzyme solution (to a final concentration of 64 nM) was added to the preheated substrate and mixed by inverting the sample three times. The sample was then immediately inserted into the preheated NMR spectrometer and the experiment was started (the time from adding the enzyme to the first spectra has been recorded was between 3-4 min. totally). The recorded spectrum is a pseudo-2D type experiment recording a 1D proton NMR spectrum every 5 min with in total 220 time points. The recorded 1D proton spectrum contains 32K data points and has a spectral width of 10 ppm, 24 scans with a 30° flip angle, and relaxation delay of 1 s (total recording time of 73s). For enzyme treatment, 2.5 µL of *Ri*Xyn10A and *Ri*Agu115A were added to the AcBGX sample to a final concentration of 167 nM and 13 nM, respectively, and the sample incubated at 37 °C for 24 hours prior to *Ri*AXE addition. All homo and heteronuclear NMR experiments were recorded on a BRUKER AVIIIHD 800 MHz (Bruker BioSpin AG, Fälladen, Switzerland) equipped with 5mm with cryogenic CP-TCI and all acquisitions were done at 37°C. For chemical shift assignment of AcBGX substrate, the following spectra were recorded: 1D proton, 2D double quantum filtered correlation spectroscopy (DQF-COSY), 2D total correlation spectroscopy (TOCSY) with 70 ms mixing time, 2D ^{13}C heteronuclear single quantum coherence (HSQC) with multiplicity editing, 2D ^{13}C Heteronuclear 2 Bond Correlation (H2BC), 2D ^{13}C HSQC-[^{1}H,^{1}H]TOCSY with 70 ms mixing time on protons and 2D

heteronuclear multiple bond correlation (HMBC) with BIRD filter to suppress first order correlations. The acetate signal to 1.903 ppm (pH 7.0 at 37 °C, in relation to 4,4-dimethyl-4-silapentane-1-sulfonic acid, DSS [49] was used as chemical shift reference for protons, while ^{13}C chemical shifts were referenced indirectly to acetate, based on the absolute frequency ratios[50]. The spectra were recorded, processed and analyzed using TopSpin 3.5 software (Bruker BioSpin).

Surface plasmon resonance (SPR)

Xylo-oligosaccharide binding to *Ri*Xyn10A, *Ri*Xyn10AΔCBMx and *Ri*Xyn10A-CBMx was analyzed using surface plasmon resonance (SPR) on a BIAcore T100 (GE Healthcare). Immobilization of the proteins on a CM5 chips was performed using a random amine coupling kit (GE Healthcare) according to the manufacture's protocol with 50-150 µg mL^{-1} protein in 10 mM sodium acetate pH 3.6-4.2, to a density of 1362, 10531 and 4041 response units (RU) for *Ri*Xyn10AΔCBMx, *Ri*Xyn10A and *Ri*Xyn10A-CBMx, respectively. The analysis comprised 90 s of association, 240 s of dissociation at 30 µL min^{-1}. Sensograms were recorded at 25°C in 20 mM phosphate/citrate buffer, pH 6.5, 150 mM NaCl, 0.005% (v/v) P20 (GE Healthcare). All solutions were filtered prior to analysis (0.22 µm). Experiments were performed in duplicates with seven concentrations in the range 156 µM−10 mM for X3, 75 µM−4 mM for X4, X6, Man6 and 62.5 µM−4 mM X5. Data analysis was carried out using the Biacore T100 evaluation software and dissociation constants (K_D) were determined by fitting a one-binding site model to the steady state sensograms. No binding was measured for Man6.

Isothermal titration calorimetry (ITC)

Titrations were performed using a Microcal ITC_{200} calorimeter (GE healthcare) at 25°C with *Ri*XBP (0.1mM) or *Ri*Xyn10AΔCBMx (0.25 mM) in the sample cell and xylo-oligosaccharides (2.2−5 mM) in 10 mM sodium phosphate pH 6.5 in the syringe. An initial injection of 0.5 µL, was followed by 19 x 2 µL injections separated by 120 s. The data were corrected for the heat of dilution, determined from buffer titration and a nonlinear single binding model was fitted to the normalized integrated binding isotherms using the MicroCal Origin software v7.0 to determine the thermodynamic binding parameters.

Affinity electrophoresis

Binding of CBMx to soluble WAX (0−0.1% w/v) or BGX (0−1.0% w/v) was assessed by affinity electrophoresis[51] in 10% native polyacrylamide gels (70 V, 3 hours, 4°C) using purified recombinant *Ri*Xyn10A-CBMx (3.0 µg) and ß-lactoglobulin (1.5 µg) as a negative control. The relative mobility (r) was calculated as the migration of *Ri*Xyn10A-CBMx relative to migration of the dye front. A linear regression of the 1/r versus xylan concentration allowed the determination of K_D as the intercept of this X-axis.

Western blot and immunofluorescence microscopy

Custom antibodies against the recombinant for the two xylanases *Ri*Xyn10A, *Ri*Xyn10B and the transport protein *Ri*XBP were raised in rats and rabbit, respectively (Eurogentec, Seraing, Belgium). The specificity of the antibodies was tested by western blots. The purified proteins of interest were transferred to an Amersham Protran Premium 0.2 µM nitrocellulose membrane (GE Healthcare). The membranes were blocked for 1 hour in 1% (w/v) BSA in TBST-buffer (Tris-buffered saline, 0.1% (v/v Tween 20) and incubated for 2 hours with the antisera (500x dilution in TBST-buffer). Subsequently, the membranes were washed three times in TBST-buffer and incubated for 2 hours with 6000x diluted secondary polyclonal goat anti-rabbit IgG-AP antibodies coupled to alkaline phosphatase (AP) (Dako,

Glostrup, Denmark) and rabbit anti-rat IgG-AP (Sigma). After three washes, the proteins were visualized by exposure to Sigma-Fast BCIP/NBT reagent (Sigma).
R. intestinalis cells were grown in 6 mL YCFA containing 0.5% (w/v) WAX to $OD_{600\ nm}$≈0.8, harvested (4000 g, 5 min, room temperature) and washed twice in PBS. The cells were resuspended in 3 mL 4% (w/v) paraformaldehyde in PBS and fixed by incubation on ice for 15 min. Thereafter the cells were washed twice in PBS and resuspended in 2 mL PBS. 50 µL of cell suspension were added to glass slides coated with poly-L-lysine, cells blocked for 1 hour in blocking buffer (1% (w/v) milk powder in PBS) and washed twice in PBS. For labelling, the cells were incubated with 50 µL anti-sera diluted 50x in blocking buffer for 2 hours, washed twice in PBS and incubated for 1 hour with 50 µL goat anti-rat IgG Alexa-Flour 555 or goat anti-rabbit IgG Alexa-Flour 488 (Thermo Scientific, Massachusetts, USA). Secondary antibodies were diluted 500x PBS. Finally, cells were washed two times in PBS, one drop of ProLong Gold antifade (Thermo Scientific, Massachusetts, USA) was applied and the cells secured with a cover slide. Fluorescence was visualized using Zeiss Axioplan 2 microscope equipped with a CoolSNAP cf color camera and a Zeiss Plan-Neofluar 100X/1.3NA, oil immersion objective.

Co-culture competition assay

Bacteriodes ovatus DSM 1896 and *R. intestinalis* DSM 14610 were grown anaerobically in 20 mL YCFA supplemented with 0.5% (w/v) glucose to late-log phase and an approximately equal number of cells (estimated by $OD_{600\ nm}$) were inoculated into CFA medium (YCFA lacking the yeast extract to reduce the risk of *B. ovatus* DSM 1896 growing on yeast extract[52]) containing 0.5% (w/v) WAX, BGX or InWAX. The co-cultures were grown in triplicates and samples (2 mL) were taken during growth. Genomic DNA was extracted from samples using DNAClean® Microbial DNA isolation kit (Qiagen). Relative bacterial abundance was estimated by qPCR. The extracted DNA was diluted to 0.5 ng μL^{-1} and amplified in technical triplicates using strain specific primers (Supplementary Fig. 0) The amplification mix contained 2 µl DNA, 5.5 µl LightCycler 480 SYBR Green I Master mix (Roche), 0.22 µL of each primer (10 pmol/µL) and 3 µL sterile water. Amplification conditions were 1 cycle of 95 °C for 5 min, 45 cycles of 95 °C for 10 s, 60 °C for 15 s and 72°C for 45 s using a LightCycler 480 II (Roche). Relative bacterial concentrations in each sample were estimated by comparing the gene copy numbers calculated using standard curves prepared with the respective reference DNA. Western blot was performed as described above but with cell cultures instead of purified proteins.

Acknowledgements

We wish to thank Bernard Henrissat, architecture et fonction des macromolécules biologiques, CNRS, Aix-Marseille University, for his advice and discussion on assigning the CBMx and the esterase into novel CAZy-families. We thank Dr. Madhav, Yadav, United States Department of Agriculture, Agricultural Research Service for the kind gift of cornbran xylan, and BioCHOS AS (Ås, Norway) for providing the chitooligo (CHOS) sample. Marzanna Due, Thanh Holm Madsen and Camilla Aaarup Christensen are thanked for technical help in cloning recombinant proteins and the performance of binding experiments. We also wish to thank Alexander Schultz, Helle Juel Martens and Micheal Hansen, PLEN, University of Copenhagen for the use of confocal laser scanning microscopy in the initial microscopy experiments. This project was funded by a Graduate School DTU Scholarship, Lyngby, Denmark. Additional fundings were from the Danish Research Council for Independent Research, Natural Sciences (DFF, FNU) by a Research Project 2 grant (Grant ID: 4002-00297B), a BIONÆR project (grant numbers 244259) and the Norwegian NMR Platform, NNP (FLA) from the Research Council of Norway and (226244).

References

1. Nicholson, J. K. *et al.* Host-gut microbiota metabolic interactions. *Science (80-.).* **108,** 1262–1268 (2012).

2. Sonnenburg, J. L. & Bäckhed, F. Diet–microbiota interactions as moderators of human metabolism. *Nature* **535,** 56–64 (2016).

3. Marchesi, J. R. *et al.* The gut microbiota and host health: a new clinical frontier. *Gut* 1–10 (2015).

4. David, L. A. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505,** 559–563 (2013).

5. Desai, M. S. *et al.* A dietary fiber-deprived gut microbiota degrades the colonic mucus barrier and enhances pathogen susceptibility. *Cell* **167,** 1339–1353.e21 (2016).

6. Xu, S. *et al.* Butyrate induces apoptosis by activating PDC and inhibiting complex I through SIRT3 inactivation. *Signal Transduct. Target. Ther.* **2,** e16035 (2017).

7. Donohoe, D. R. *et al.* The warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation. *Mol. Cell* **48,** 612–626 (2012).

8. Furusawa, Y. *et al.* Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* **506,** 254–254 (2014).

9. Morrison, D. J. & Preston, T. Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. *Gut Microbes* **7,** 189–200 (2016).

10. Takahashi, K. *et al.* Reduced abundance of butyrate-producing bacteria species in the fecal microbial community in Crohn's disease. *Digestion* **93,** 59–65 (2016).

11. Qin, J. *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490,** 55–60 (2012).

12. Vrieze, A. *et al.* Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology* **143,** 913–916.e7 (2012).

13. Duncan, S. H., Hold, G. L., Barcenilla, A., Stewart, C. S. & Flint, H. J. *Roseburia intestinalis* sp. nov., a novel saccharolytic, butyrate-producing bacterium from human faeces. *Int. J. Syst. Evol. Microbiol.* **52,** 1615–1620 (2002).

14. Louis, P. & Flint, H. J. Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. *FEMS Microbiol. Lett.* **294,** 1–8 (2009).

15. Van den Abbeele, P. *et al.* Butyrate-producing Clostridium cluster XIVa species specifically colonize mucins in an *in vitro* gut model. *ISME J.* **7,** 949–61 (2013).

16. El Kaoutari, A., Armougom, F., Gordon, J. I., Raoult, D. & Henrissat, B. The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat. Rev. Microbiol.* **11,** 497–504 (2013).

17. Mirande, C. *et al.* Dietary fibre degradation and fermentation by two xylanolytic bacteria *Bacteroides xylanisolvens* XB1AT and *Roseburia intestinalis* XB6B4 from the human intestine. *J. Appl. Microbiol.* **109,** 451–460 (2010).

18. Chassard, C., Goumy, V., Leclerc, M., Del'homme, C. & Bernalier-Donadille, A. Characterization of the xylan-degrading microbial community from human faeces. *FEMS Microbiol. Ecol.* **61,** 121–131 (2007).

19. Selvendran, R. R. Chemistry of plant cell walls and dietary fibre. *Scand. J. Gastroenterol.* **5521,** 33–41 (1987).

20. Rogowski, A. *et al.* Glycan complexity dictates microbial resource allocation in the large intestine. *Nat. Commun.* **6,** 7481 (2015).

21. Zhang, M. *et al.* Xylan utilization in human gut commensal bacteria is orchestrated by unique modular organization of polysaccharide-degrading enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **111,** E3708–E3717 | (2014).

22. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42,** (2014).

23. Kelly, G. *et al.* Structure of the cell-adhesion fragment of intimin from enteropathogenic *Escherichia coli*. *Nat Struct Mol Biol* **6,** 313–318 (1999).

24. Ebbes, M. *et al.* Fold and Function of the InlB B-repeat. *J. Biol. Chem.* **286,** 15496–15506 (2011).

25. Karlsson, E. N. *et al.* The modular xylanase Xyn10A from *Rhodothermus marinus* is cell-attached, and its C-terminal domain has several putative homologues among cell-attached proteins within the phylum Bacteroidetes. *FEMS Microbiol. Lett.* **241,** 233–242 (2004).

26. Ejby, M. *et al.* An atp binding cassette transporter mediates the uptake of α-(1,6)-linked dietary oligosaccharides in bifidobacterium and correlates with competitive growth on these substrates. *J. Biol. Chem.* **291,** 20220–20231 (2016).

27. Ejby, M. *et al.* Structural basis for arabinoxylo-oligosaccharide capture by the probiotic *Bifidobacterium animalis* subsp. lactis Bl-04. *Mol. Microbiol.* **90,** 1100–1112 (2013).

28. Honda, Y. & Kitaoka, M. A family 8 glycoside hydrolase from *Bacillus halodurans* C-125 (BH2105) is a reducing end xylose-releasing exo-oligoxylanase. *J. Biol. Chem.* **279,** 55097–55103 (2004).

29. Anand, S., Kaur, H. & Mande, S. S. Comparative *in silico* analysis of butyrate production pathways in gut commensals and pathogens. *Front. Microbiol.* **7,** 1–12 (2016).

30. Cockburn, D. W. & Koropatkin, N. M. Polysaccharide degradation by the intestinal microbiota and its influence on human health and disease. *J. Mol. Biol.* **428,** 3230–3252 (2016).

31. Sheridan, P. O. *et al.* Polysaccharide utilisation loci and nutritional specialisation in a dominant group of butyrate-producing human colonic Firmicutes. *Microb. Genomics* **2,** (2016).

32. Duncan, S. H. *et al.* Wheat bran promotes enrichment within the human colonic microbiota of butyrate-producing bacteria that release ferulic acid. *Environ. Microbiol.* **18,** 2214–2225 (2016).

33. Morrill, J. *et al.* The GH5 1,4-β-mannanase from Bifidobacterium animalis subsp. lactis Bl-04 possesses a low-affinity mannan-binding module and highlights the diversity of mannanolytic enzymes. *BMC Biochem.* **16,** 26 (2015).

34. Cockburn, D. W. *et al.* Molecular details of a starch utilization pathway in the human gut symbiont Eubacterium rectale. *Mol. Microbiol.* **95,** 209–230 (2015).

35. Ze, X. *et al.* Unique organization of extracellular amylases into amylosomes in the resistant starch-utilizing human colonic firmicutes bacterium *Ruminococcus bromii*. *MBio* **6,** 1–11 (2015).

36. De Paepe, K., Kerckhof, F.-M., Verspreet, J., Courtin, C. M. & Van de Wiele, T. Inter-individual differences determine the outcome of wheat bran colonization by the human gut microbiome. *Environ. Microbiol.* **0,** 1–17 (2017).

37. Mahowald, M. A. *et al.* Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. *Proc Natl Acad Sci U S A* **106,** 5859–5864 (2009).

38. Biely, P. *et al.* Mode of action of acetylxylan esterases on acetyl glucuronoxylan and acetylated oligosaccharides generated by a GH10 endoxylanase. *Biochim. Biophys. Acta - Gen. Subj.* **1830,** 5075–5086 (2013).

39. Sørbotten, A., Horn, S. J., Eijsink, V. G. H. & Vårum, K. M. Degradation of chitosans with chitinase B from Serratia marcescens. *FEBS J.* **272,** 538–549 (2005).

40. Cock, P. J. A. *et al.* Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25,** 1422–1423 (2009).

41. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14,** R36 (2013).

42. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9,** (2012).

43. Anders, S., Pyl, P. T. & Huber, W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–169 (2015).

44. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 550 (2014).

45. Dümmler, A., Lawrence, A.-M. & de Marco, A. Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. *Microb. Cell Fact.* **4,** 34 (2005).

46. Miller, G. L. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* **31,** 426–428 (1959).

47. Roe, J. H. & Rice, E. W. A photometric method for the determination of free pentoses in animal tissue. *J. Biol. Chem.* **173,** 507–512 (1948).

48. Deschatelets, L. & Yu, E. K. C. A simple pentose assay for biomass conversion studies. *Appl. Microbiol. Biotechnol.* **24,** 379–385 (1986).

49. Govind, V., Young, K. & Maudsley, A. A. Corrigendum: Proton NMR chemical shifts and coupling constants for brain metabolites. Govindaraju V, Young K, Maudsley AA, *NMR Biomed* . 2000; 13: 129-153. *NMR Biomed.* **28,** 923–924 (2015).

50. Zhang, H., Neal, S. & Wishart, D. S. RefDB: A database of uniformly referenced protein chemical shifts. *J. Biomol. NMR* **25,** 173–195 (2003).

51. Takeo, K. Affinity electrophoresis: Principles and applications. *Electrophoresis* **5,** 187–195 (1984).

52. Scott, K. P., Martin, J. C., Duncan, S. H. & Flint, H. J. Prebiotic stimulation of human colonic butyrate-producing bacteria and bifidobacteria, in vitro. *FEMS Microbiol. Ecol.* **87,** 30–40 (2014).

53. Larsen, N. *et al.* Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PLoS One* **5,** e9085 (2010).

54. Bergström, A. *et al.* Introducing GUt Low-Density Array (GULDA) - a validated approach for qPCR-based intestinal microbial community analysis. *FEMS Microbiol. Lett.* **337,** 38–47 (2012).

55. Walter, J. *et al.* Detection and identification of gastrointestinal *Lactobacillus* species by using denaturing gradient gel electrophoresis and species-specific PCR primers. *Appl. Environ. Microbiol.* **66,** 297–303 (2000).

56. Despres, J. *et al.* Xylan degradation by the human gut *Bacteroides xylanisolvens* XB1AT involves two distinct gene clusters that are linked at the transcriptional level. *BMC Genomics* **17,** 326 (2016).

57. St John, F. J., Hurlbert, J. C., Rice, J. D., Preston, J. F. & Pozharski, E. Ligand bound structures of a glycosyl hydrolase family 30 glucuronoxylan xylanohydrolase. *J. Mol. Biol.* **407,** 92–109 (2011).

58. Dodd, D. & Cann, I. K. Enzymatic deconstruction of xylan for biofuel production. *Glob Chang. Biol Bioenergy* **1,** 2–17 (2009).

59. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32,** 1792–1797 (2004).

60. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33,** msw054 (2016).

Supplementary (max 10 items (tables+figures)

Supplementary Figure 1

a

Primers

Gene	Accession number	Name	Orientation	Sequence (5' -> 3')
Cloning[a]				
ROSINTL182_ 06494 (AA27-1356)	EEV01588.1	*Ri*Xyn10A	Forward	**TTTCAGGGCGCCATG**GGGGTAAAAAAAGTTTTTACTGCAGAT
ROSINTL182_ 06494 (AA27-1356)	EEV01588.1	*Ri*Xyn10A	Reverse	**GACGGAGCTCGAATTTTA**CTACTTACTGATCTTTATCTTCTTTGCA
ROSINTL182_ 06494 (AA156-1356)	EEV01588.1	*Ri*Xyn10AΔCBMx	Forward	**TTTCAGGGCGCCATG**GCAGGAGCAGGCGATGCA
ROSINTL182_ 06494 (AA156-1356)	EEV01588.1	*Ri*Xyn10AΔCBMx	Reverse	**GACGGAGCTCGAATTTTA**CTACTTACTGATCTTTATCTTCTTTGCA
ROSINTL182_ 06494 (AA349-754)	EEV01588.1	*Ri*Xyn10A-cata	Forward	**TTTCAGGGCGCCATG**TCTATTGAGAAGGACATCCCGGA
ROSINTL182_ 06494 (AA349-754)	EEV01588.1	*Ri*Xyn10A-cata	Reverse	**GACGGAGCTCGAATTTTA**GGATGCATCTACATACGCCCA
ROSINTL182_ 06494 (AA27-165)	EEV01588.1	*Ri*Xyn10ACBMx	Forward	**TTTCAGGGCGCCATG**GGGGTAAAAAAAGTTTTTACTGCAGAT
ROSINTL182_ 06494 (AA27-165)	EEV01588.1	*Ri*Xyn10ACBMx	Reverse	**GACGGAGCTCGAATTTTA**ATCCCCCAATTTTGCA
ROSINTL182_ 08193	EEU99940.1	*Ri*Abf43A	Forward	**AGGAGATATACCATG**AGTATAGCAAAGAATCCGGTTC
ROSINTL182_ 08193	EEU99940.1	*Ri*Abf43A	Reverse	**GGTGGTGGTGCTCGA**AACCCGGTATTCCCTCATA
ROSINTL182_ 08194	EEU99941.1	*Ri*AXE	Forward	**AGGAGATATACCATG**AGTGGACCTGTGGCA
ROSINTL182_ 08194	EEU99941.1	*Ri*AXE	Reverse	**GGTGGTGGTGCTCGA**ATTCCACATAGCCAAAACCAA
ROSINTL182_ 08195	EEU99942.1	*Ri*Agu115A	Forward	**TTTCAGGGCGCCATG**GAAGCAATTTTGGTAAAGGATC
ROSINTL182_ 08195	EEU99942.1	*Ri*Agu115A	Reverse	**GACGGAGCTCGAATTTTA**TCATCTGTTCGTCCTCCTT
ROSINTL182_ 08196	EEU99943.1	*Ri*Xyl8	Forward	**AGGAGATATACCATG**AAAAGAGGAGCGTTTGAGA
ROSINTL182_ 08196	EEU99943.1	*Ri*Xyl8	Reverse	**GGTGGTGGTGCTCGA**AATAAATTCTATAATTGCCGCTCAG
ROSINTL182_ 08199	EEU99894.1	*Ri*XBP	Forward	**TTTCAGGGCGCCATG**GGAAACAAAGCAGCCG
ROSINTL182_ 08199	EEU99894.1	*Ri*XBP	Reverse	**GACGGAGCTCGAATTTTA**TTACTGATATTTTTTTGCTTCCTC
ROSINTL182_ 08202	EEU99897.1	*Ri*Xyl3A	Forward	**AGGAGATATACCATG**GAATTAAATCAGAATACAGAAAAACTG
ROSINTL182_ 08202	EEU99897.1	*Ri*Xyl3A	Reverse	**GGTGGTGGTGCTCGAA<u>TAA</u>**CATCAGACTTTCCACTGTTT
ROSINTL182_06338/ ROSINTL182_06339	EEV01752.1/ EEV01731.1	*Ri*Xyn10B	Forward	**TTTCAGGGCGCCATG**GCTGGGCAGGAAAATG
ROSINTL182_06338/ ROSINTL182_06339	EEV01752.1/ EEV01731.1	*Ri*Xyn10B	Reverse	**GACGGAGCTCGAATTTTA**CTATTTATCAGAATGAAATAAATTTTCAA

[a]Bold nucleotides indicate the sequences annealing to the vector.
[b]Underlined nucleotides indicate the changed codon and italics indicate the changed bases.

b

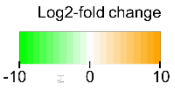
qPCR primers

Target bacteria	Orientation	Sequence (5' -> 3')	Reference
Roseburia spp.	Forward	TACTGCATTGGAAACTGTCG	[53]
Roseburia spp	Reverse	CGGCACCGAAGAGCAAT	[53]
Bacteroides spp.	Forward	CGATGGATAGGGGTTCTGAGAGGA	[54]
Bacteroides spp.	Reverse	GCTGGCACGGAGTTAGCCGA	[54]
Universal primer	Forward	ACTCCTACGGGAGGCAGCAGT	[55]
Universal primer	Reverse	GTATTACCGCGGCTGCTGGCAC	[55]

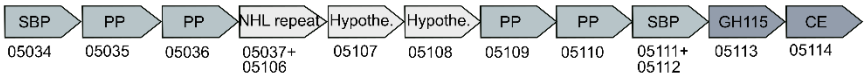
Supplementary Figure 1 Primers used in this study. (**a**) Cloning primers (**b**) qPCR primers

a

Locus ID	Log2-fold change			SP	Protein	Annotation
	X1/Glc	WAX/Glc	BGX/Glc			
05034	-0.84	6.48	5.54	Yes	SBP	ABC transporter, solute-binding protein
05035	-0.79	5.99	5.44	No	PP	ABC transporter, permease protein
05036	-0.80	6.04	5.26	No	PP	ABC transporter, permease protein
05037+05106	-0.33	6.44	5.95	Yes		NHL repeat protein
05107	-1.48	5.98	5.53	No		Hypothetical protein
05108	-0.42	5.91	5.48	No		Hypothetical protein
05109	-0.61	5.87	4.92	No	PP	ABC transporter, permease protein
05110	-0.63	5.11	3.92	No	PP	ABC transporter, permease protein
05111+05112	0.34	5.73	5.07	Yes	SBP	ABC transporter, solute-binding protein
05113	-0.77	4.95	4.78	No		Hypothetical protein
05114	-0.95	4.09	4.08	No	GH115	Xylan α-1,2-glucuronidase
05115	-0.73	4.35	4.17	No	CE	Putitativ esterase

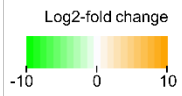


b

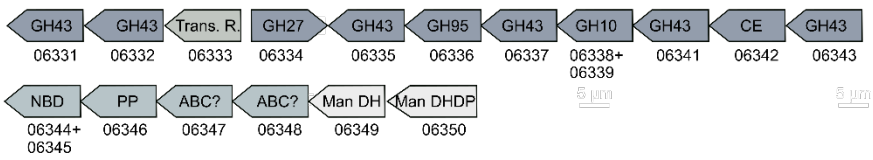


c

Locus ID	Log2-fold change			SP	Protein	Annotation
	X1/Glc	WAX/Glc	BGX/Glc			
06331	-0.63	5.26	4.76	No	GH43	β-xylosidase/α-L-arabinofuranosidase
06332	-1.60	5.46	4.74	No	GH43	β-xylosidase/α-L-arabinofuranosidase
06333	-0.70	0.44	2.24	No	AraC	Transcriptional regulator
06334	-1.22	1.03	2.44	No	GH27	α-galactosidase/β-L-arabinopyranosidase
06335	-0.69	3.58	3.44	No	GH43	β-xylosidase/α-L-arabinofuranosidase
06336	-1.04	3.61	3.21	No	GH95	α-L-galactosidase/α-L-fucosidase
06337	-0.57	3.45	4.08	No	GH43	β-xylosidase/α-L-arabinofuranosidase
06338+06339	-0.99	3.47	3.62	Yes	*Ri*Xyn10B	Endo-1,4-β-xylanase
06341	-1.12	4.09	3.80	No	GH43	β-xylosidase/α-L-arabinofuranosidase
06342	-1.32	4.28	4.07	No	CE1	Esterase
06343	-1.23	4.49	4.21	Yes	GH43	β-xylosidase/α-L-arabinofuranosidase
06344+06345	-1.57	4.49	4.68	No	ABC-NBD	ABC transporter, nucleotide binding domain
06346	-2.45	5.12	4.74	No	ABC-PP	ABC transporter, permerase protein
06347	-2.38	5.49	5.35	No		Hypothetical ABC transporter
06348	-1.62	4.52	4.04	No		Hypothetical ABC transporter
06349	-0.95	6.49	5.44	No	UxuA	Mannonate dehydratase
06350	-1.50	5.84	5.12	No	UxuB	Mannitol/D-arabinitol dehydrogenase domain protein



d

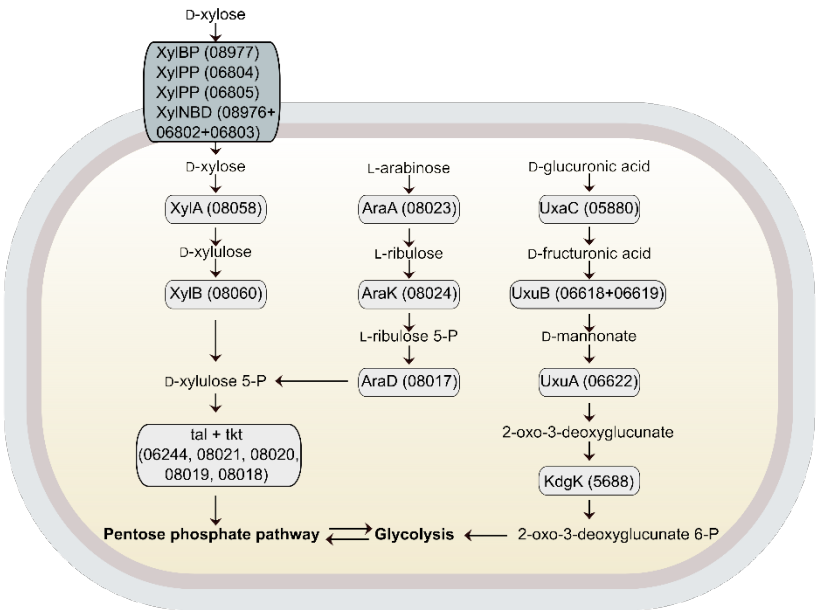


e

Locus ID	Log2-fold change			SP	Protein	Annotation
	X1/Glc	WAX/Glc	BGX/Glc			
08977	9.09	5.79	5.26	Yes	XylBP	ABC transporter, solute-binding protein
08976+06802+06803	8.98	6.15	5.65	No	XylNBD	ABC transporter, nucleotide binding domain
06804	8.73	5.56	4.43	No	XylPP	ABC transporter, permease protein
06805	8.42	5.50	4.75	No	XylPP	ABC transporter, permease protein
08058	5.51	4.34	4.16	No	XylA	Xylose isomerase
08060	5.05	4.28	3.77	No	XylB	Xylulokinase
08023	-0.66	5.89	3.11	No	AraA	L-arabinose isomerase
08024	-0.44	5.55	2.90	No	AraK	L-ribulokinase
08017	4.40	5.06	4.96	No	AraD	L-ribulose-5-phosphate 4-epimerase
05880	0.70	0.58	0.59	No	UxaC	Glucuronate isomerase
06618+06619	0.18	-0.24	5.77	No	UxuB	Mannonate oxidoreductase
06622	-1.06	1.11	4.06	No	UxuA	Mannonate dehydratase
05688	-0.18	-0.69	1.26	No	KdgK	2-keto-3-deoxy-D-gluconate kinase
06244	6.80	7.15	5.93	No	tkt	D-xylulose 5-phosphate/D-fructose 6-phosphate phosphoketolase
08021	5.29	5.72	4.85	No	tal	Fucose isomerase
08020	4.33	4.75	4.37	No	tkt	Transketolase
08019	4.06	4.44	3.76	No	tkt	Transketolase
08018	4.19	4.65	4.47	No	tkt	Transketolase

Log2-fold change
-10 0 10

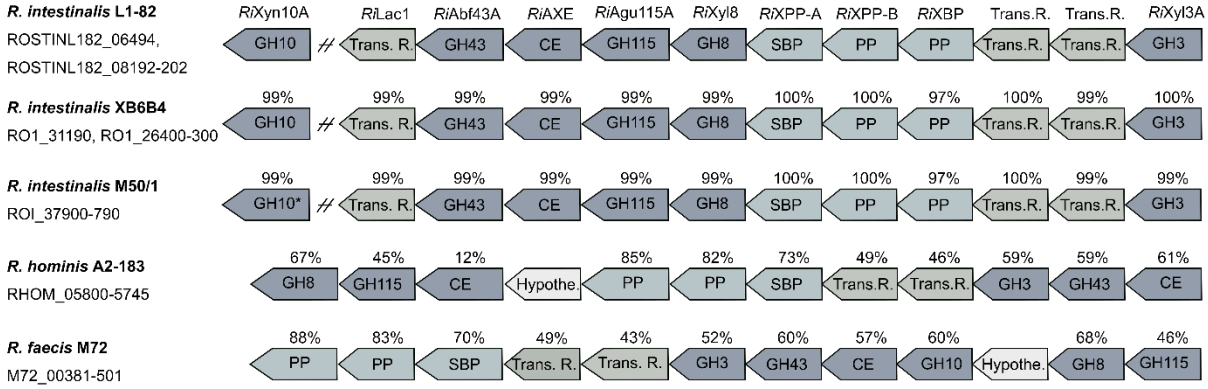
f



Supplementary Figure 2 *R. intestinalis* L1-82 unique xylan upregulated loci and xylose metabolism. (**a-d**) Identification of two additional xylan upregulated loci in *R. intestinalis* L1-82**,** which are not present in other strains within the species. (**e**) Xylose import and metabolism genes are also identified from the RNA-seq analysis, together with genes mediating the metabolism of arabinose and methylglucuronic acid**.** The RNA-Seq heatmap depicts Log2 fold changes of genes expressed by cells grown on xylose (X1), wheat arabinoxylan (WAX) and birch glucuronoxylan (BGX) relative to glucose (Glc). Formal locus tag numbers ROSINTL182_xxxxx are abbreviated with the last numbers after the hyphen and signal peptides (SP) were predicted using SignalP v.3.0. (**f**) Proposed model for the metabolism of monosaccharides xylose, arabinose and methylglucuronic acid in *R. intestinalis* L1-82. Representation of a gene with in more than one locus ID indicates that the gene spans these loci, which is an artefact awing to the lack of assembly of the contigs of the *R. intestinalis* L1-82 genome.

Supplementary Figure 3

a



b

Xylanases of GH10 from human gut Firmicutes and Bacteroidetes

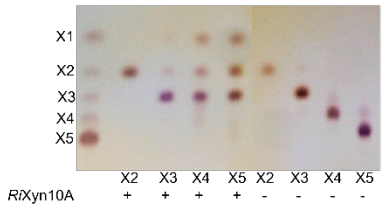
Phylum	Family	Strain	Accession number	Length (AA)	CBMs
Firmicutes	Lachnospiraceae	*Roseburia intestinalis* L1-82	ROSINTL182_06494	1356	X, 22, 9, 9
			ROSINTL182_6338-9	601	
		Roseburia intestinalis XB6B4	CBL13458.1	1356	X, 22, 9, 9
		Roseburia intestinalis M50/1	n.a.	1356	X, 22, 9, 9
		Roseburia faecis M72	CRL32809.1	1380	X, 22, 9, 9
		Eubacterium rectale T1-815	CRL34489.1	1028	X, 9, 9
		Butyrivibrio fibrisolvens 16/4	CBK74925.1	1153	9
			CBK75021.1	690	13, 2
		Hungatella hathewayi	CUO52114.1	421	
		Ruminococcus gnavus	WP_064787180.1	394	
	Ruminococcaceae	*Ruminococcus champanellensis* 18P13	CBL16579.1	633	22
			CBL17682.1	1268	22, 22, 6
		Ruminococcus callidus ATCC 27760	ERJ94429.1	1158	22, 22, 9
			ERJ87773.1	630	22
			ERJ97032.1	382	22
Bacteroidetes	Bacteroidaceae	*Bacteroides ovatus*	EDO13863.1	372	
			EDO10007.1[20]	376	
			EDO14247.1	573	
			EDO10010.1[20]	740	4, 4
			EDO14052.1	584	
			EDO10798.1	750	
		Bacteroides intestinalis DSM 17393	EDV05054.1	782	4, 4
			EDV05072.1[21]	746	4, 4
			EDV03684.1	738	
			EDV05059.1	910	
			EDV07678.1	725	
			EDV07007.1[21]	899	
		Bacteroides xylanisolvens XB1A	CBK67953.1[56]	754	4, 4
			CBH32823.1	378	

AA: amino acids, n.a.: GH10 is present, but not assigned in the genome,

Supplementary Figure 3 Conservation of *R. intestinalis* core xylan utilization genes within the *Roseburia* genus and the modular organization of human gut Firmicutes and Bacteriodetes xylanases. (**a**) Genes are denoted according to their protein products; glycoside hydrolase (GH), carbohydrate esterase (CE), transcriptional regulators (Trans.R.), ABC transporter solute binding protein (SBP), ABC transporter permease protein (PP) and hypothetical proteins (Hypothe.). Sequence identities to *R. intestinalis* L1-82 genes are shown above the genes; Locus IDs for the genes are denoted under the respective strains. , Asterisk indicate that the GH10 is not assigned in the genome. (**b**) Xylanases of GH10 from human gut Firmicutes and Bacteroidetes. AA: amino acids.

Supplementary Figure 4

a

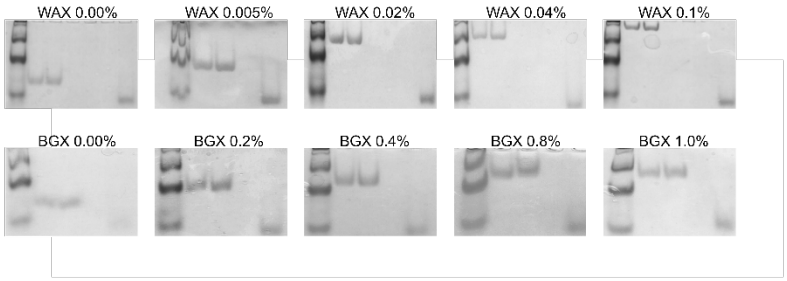


b

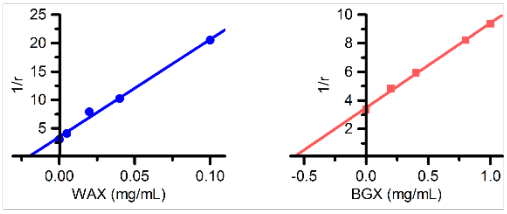
Specific activity for *Ri*Xyn10A and truncated variants (μmol min^{-1} mg^{-1})

	WAX	BGX	InWAX
*Ri*Xyn10A	87.2 ± 7.5	65.1 ± 3.2	6.5 ± 0.8
*Ri*Xyn10AΔCBMx	163.2 ± 0.8	118.6 ± 4.7	9.6 ± 0.3
*Ri*Xyn10A-cata	342.9 ± 13.2	165.3 ± 2.4	25.6 ± 0.8

c



d

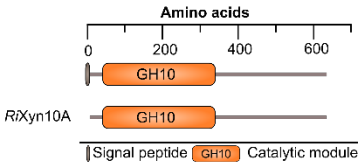


e

Homologs of CBMx

Strain	Accession number	Query cover	E-value	Identity
Roseburia intestinalis XB6B4	CBL13458.1	100%	4e-85	100%
Eubacterium rectale T1815	CRL34489.1	89%	5e-36	55%
Butyrivibrio sp. LC3010	WP_026509692.1	92%	1e-07	36%
Roseburia faecis M72	CRL32809.1	93%	9e-12	36%
Bacterium enrichment culture clone MC3F	AFU34339.1	86%	3e-07	30%
Lachnoclostridium phytofermentans ISDg	ABX41884.1	84%	5e-07	26%
Clostridium sp. KNHs205	WP_033165005.1	88%	1e-06	28%
Butyrivibrio sp. INIIa14	SCX91715.1	63%	2e-06	32%
Lachnospiraceae bacterium YSD2013	SCX14282.1	73%	1e-05	34%
Butyrivibrio sp. ob235	SEK63083.1	76%	2e-04	30%
Butyrivibrio sp. VCD2006	WP_026526370.1	72%	3e-04	27%

f

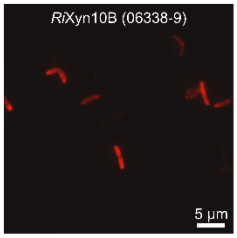


g

Kinetic parameters of *Ri*Xyn10B

	K_M (mg mL^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M (mL mg^{-1} s^{-1})
BGX	n.d.	n.d.	9.8
WAX	4.4 ± 0.8	413 ± 32	94
InWAX	n.d.	n.d.	2.3

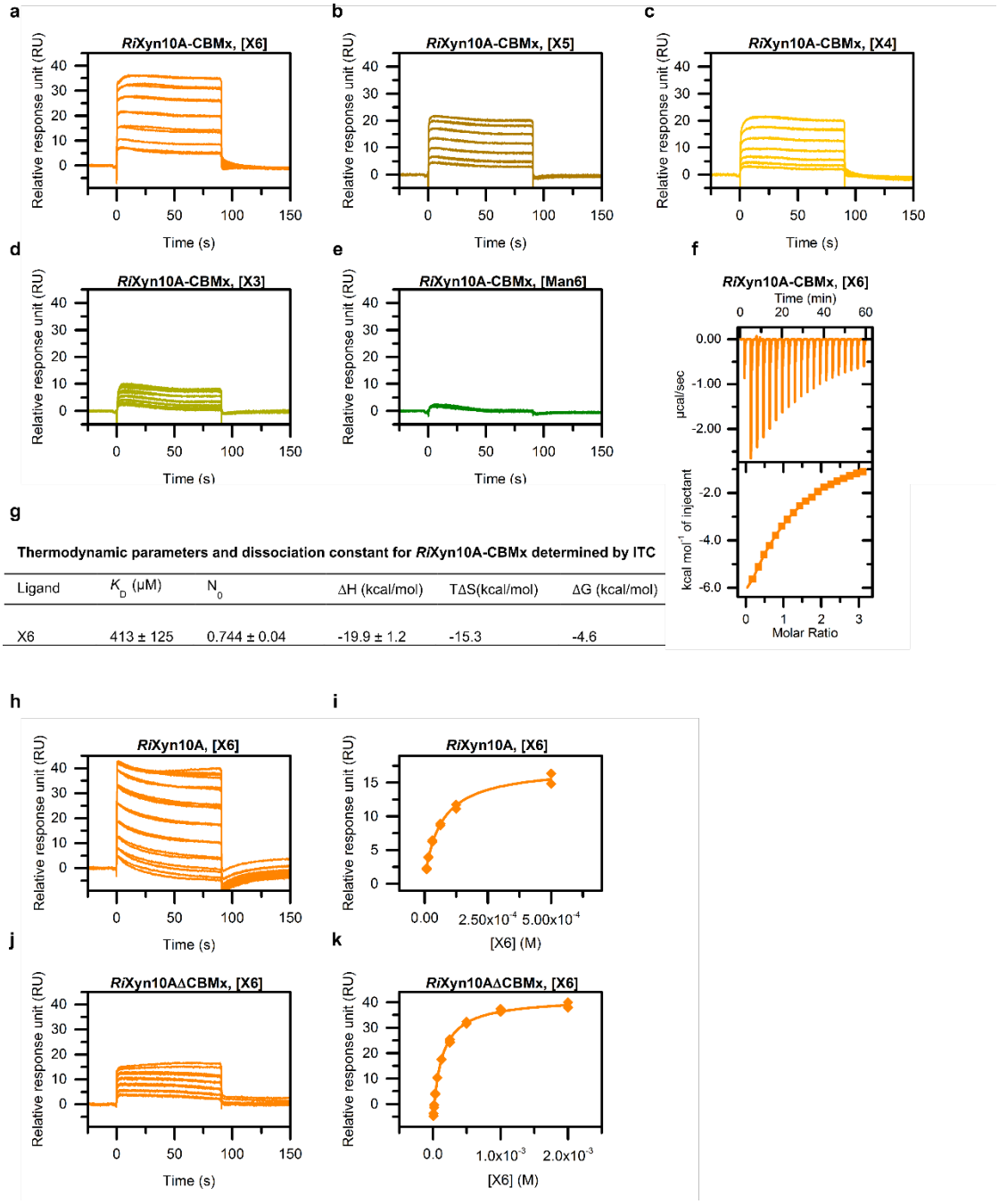
h



Supplementary Figure 4 Properties of extracellular xylanase from *R. intestinalis* (**a**) Action patterns of *Ri*Xyn10A on X2−X5 analyzed by TLC. (**b**) Specific activity of *Ri*Xyn10A and the truncated variants; *Ri*Xyn10AΔCBMx lacking the N-terminal module and *Ri*Xyn10A-cata, the catalytic module on WAX, InWAX and BGX. The data are means of triplicates with standard

deviations. (**c**) Binding of xylans to *Ri*Xyn10CBMx by affinity gel electrophoresis using native polyacrylamide gels with different concentrations of WAX (0.0-0.1% w/v) or BGX (0.0-1.0% w/v). No polysaccharides were added to the control. Lane 1+2; *Ri*Xyn10ACBMx (3.0 µg), Lane 3 ß-lactoglobulin (1.5 µg), M; marker. (**d**) Plot of 1/*r* against xylan concentration, where *r* is the relative migration distance of *Ri*Xyn10ΔCBMx in the presence of xylan in the gel. (**e**) Close homologs of CBMx identified by BLASTP against the non-redundant sequence database in NCBI. (**f**) Domain organization of the xylanase *Ri*Xyn10B encoded by a locus upregulated on xylan and unique for the of *R. intestinalis* L1-82 strain used in the study (**g**) Xylan hydrolysis kinetics of *Ri*Xyn10B on BGX, WAX and InWAX. (**h**) Extracellular localization of *Ri*Xyn10B visualized by fluorescence microscopy images of *R. intestinalis* cells labeled with anti-*Ri*Xyn10B primary antibodies. No auto fluorescence was observed for cell without primary antibody (data not shown).

Supplementary Figure 5



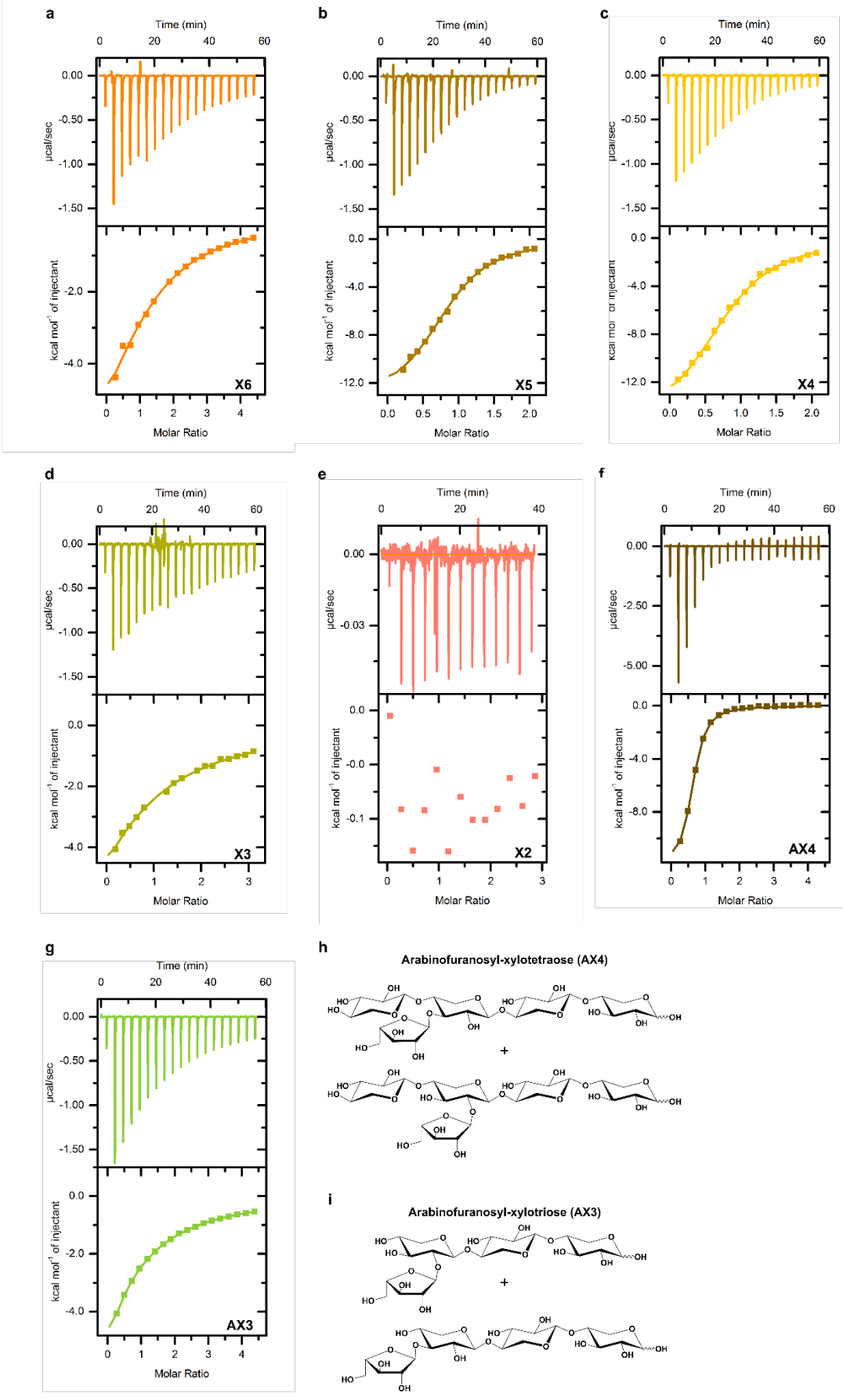
g

Thermodynamic parameters and dissociation constant for *Ri*Xyn10A-CBMx determined by ITC

Ligand	K_D (µM)	N_0	ΔH (kcal/mol)	TΔS(kcal/mol)	ΔG (kcal/mol)
X6	413 ± 125	0.744 ± 0.04	-19.9 ± 1.2	-15.3	-4.6

Supplementary Figure 5 Binding of CBMx and *Ri*Xyn10A to oligosaccharides. (**a-e**) Reference and blank corrected sensograms depict the binding of oligosaccharides to *Ri*Xyn10ACBMx. Analysis performed by SPR (**f,g**) ITC analysis of *Ri*Xyn10ACBMx binding to X6. (**k**) Reference and blank corrected sensograms and 1:1 fitted binding models depict the binding of X6 to *Ri*Xyn10A and *Ri*Xyn10AΔCBMx. Analysis performed by SPR.

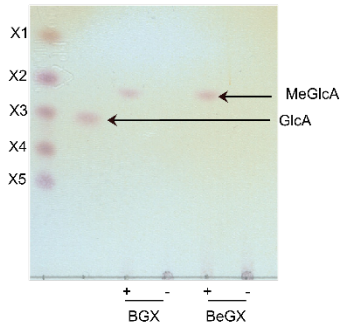
Supplementary Figure 6



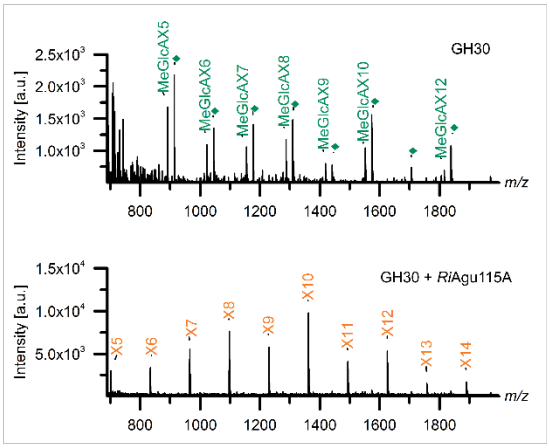
Supplementary Figure 6 The binding preference for *Ri*XBP associated to the ABC transporters that confers the uptake of xylan oligomers in *R. intestinalis*. (**a-g**) ITC analysis of *Ri*XBP binding to linear and branched xylo-oligosaccharides. (**h,i**) Structures of the branched xylo-oligosaccharides AX4 and AX3. Both compounds are mixtures with arabinofuranosyl decoration either at the C2 or C3 of xylosyl units.

Supplementary Figure 7

a



b



c

Kinetic parameters of *Ri*Agu115A	K_M (mg mL^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M (mL mg^{-1} s^{-1})
BeGX	n.d.	n.d.	2
BeGX + *Ri*Xyn10A	12 ± 3	395 ± 34	33

d

Kinetics *Ri*Abf43A	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})
AX4	0.8 ± 0.1	20 ± 1	25
	K_M (mg mL^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M (mL mg^{-1} s^{-1})
WAX	6.3 ± 0.4	12 ± 0	1.9

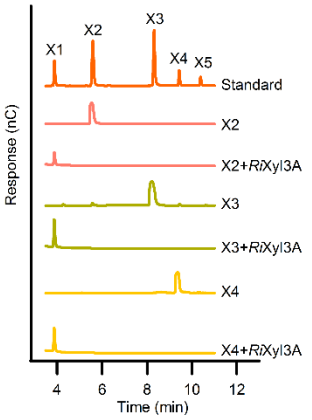
e

Kinetics *Ri*Xyl3A	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})
X2	2.7 ± 0.4	57 ± 3	21
X3	3.4 ± 0.3	60 ± 2	18
X4	2.4 ± 0.4	32 ± 2	13
X5	2.6 ± 0.5	36 ± 1	14
X6	2.1 ± 0.2	30 ± 1	15

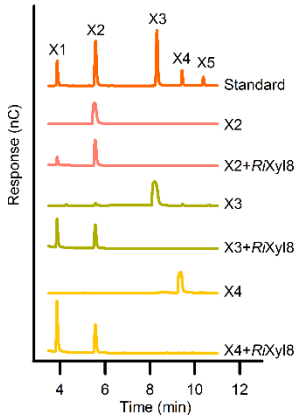
f

Kinetics *Ri*Xyl8	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})
X3	4.8 ± 0.97	1208 ± 124	251.7
X4	5.1 ± 1.46	892 ± 131	174.9

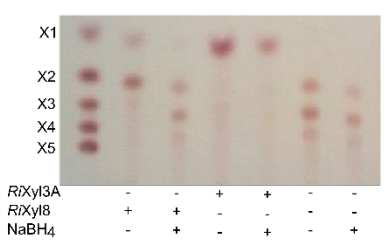
g



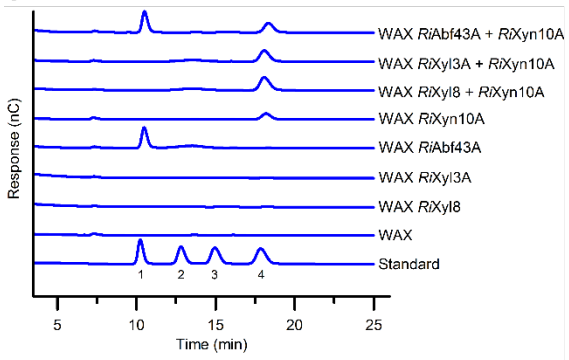
h



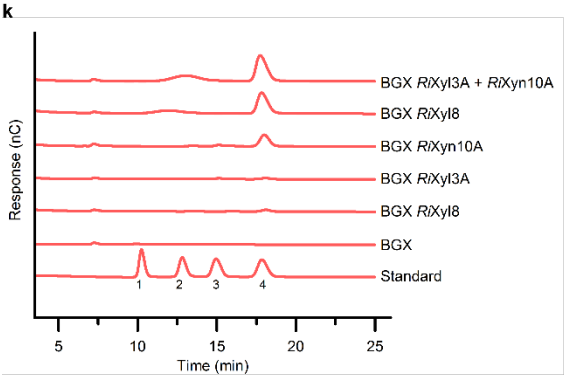
i



j



k



and MeGlcA. (**b**) Activity of *Ri*Agu115A on a BeBGX hydrolyzed using an in house GH30 xylanase monitored using MALDI-ToF MS analysis. Top panel only treatment with GH30 and bottom panel treatment with GH30 and *Ri*Agu115A. *Ri*Agu115A liberates MeGlcA from BGX preteated with a GH30, which generates xylo-oligosaccharides with a MeGlcA substitution at the penultimate xyloxyl to the reducing end[57], whereas a GH10 generates xylo-oligosaccharides with a MeGlcA substitution at the non-reducing end[58]. This data shows that the *Ri*Agu115A is able to act on both internal and terminal non-reducing end substitutions on glucuronoxylan-derived xylo-oligosaccharides. Di-sodium adducts of MeGlcA decorated oligomers (diamonds) are colored as their corresponding single sodium adducts. Kinetic parameters for the α-1,2-glucuronidase *Ri*Agu115A (**c**)**,** the α-L-arabinofuranosidase *Ri*Abf43A(**d**), the β-xylosidase *Ri*Xyl3A(**e**) and the β-xylosidase *Ri*Xyl8(**f**). Data are means of minimum a duplicates with standard deviations. (**g,h**) *Ri*Xyl3A and *Ri*Xyl8 hydrolysis of xylo-oligosaccharides analyzed with HPAEC-PAD. (**i**) Analysis of β-xylosidase activity for *Ri*Xyl3A and *Ri*Xyl8 towards xylo-oligosaccharides (XOS) by TLC. The "+" and "-" indicate the presence and absence of the different components, respectively. Reduction of the substrate with $NaBH_4$, which converts the reducing end unit to its alditol, provided evidence that *Ri*Xyl8 acts on the reducing end as the alditol is not accommodated in the active site. (**j,k**) Monosaccharide analysis of hydrolysis products from enzymatic treatment of WAX and BGX with *Ri*Xyn10A, *Ri*Abf43A, *Ri*Xyl3A and *Ri*Xyl8 by HPAEC-PAD. Standards were 1; arabinose, 2; galactose, 3; glucose, 4; xylose.

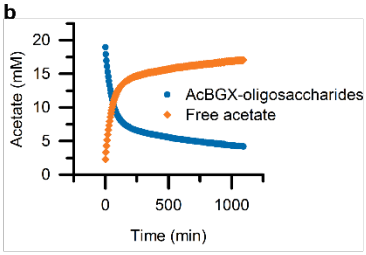
Supplementary Figure 8

a

Specific activity *Ri*AXE

	Specific activity (U/mg)
***p*NP-acetate**	4.7 ± 0.1

b



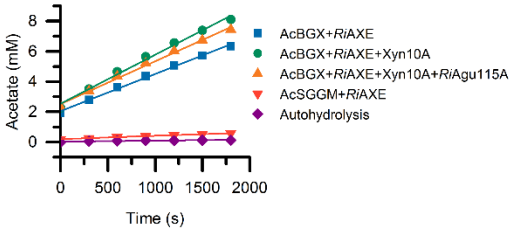
c

Rates for acetylation by *Ri*AXE

Substrate	Enzymes	*V* (μM s^{-1})	*V*/[E] (s^{-1})
AcBGX	*Ri*AXE	2.5	39.1
	*Ri*AXE+*Ri*Xyn10A	3.2	50
	*Ri*AXE+*Ri*Xyn10A+*Ri*Agu115A	2.8	43.8
AcSpruce mannan	*Ri*AXE	0.2	3.1
Autolyse		0.07	n.d.

V: rate, *V*/[E]: normalized rate by enzyme concentration

d

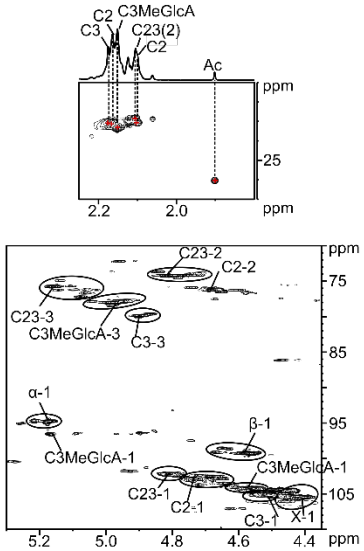


e

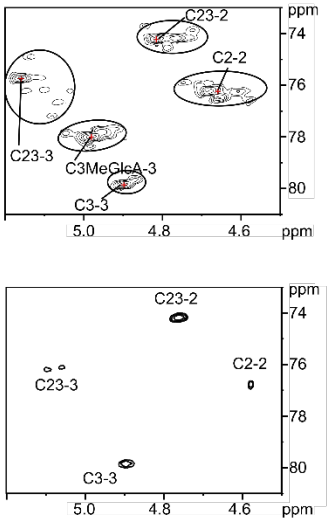
Assignment of chemical shift

Structural unit	Assignment						
	H-1; C-1	H-2; C-2	H-3; C-3	H-4; C-4	H-5; C-5	H-6; C-6	Ac-H; C
X	4.42; 105.4	3.19; 75.4	3.53; 76.4	3.78; 79,2	n.d	n.d	-
C2	4.68; 102.6	4.69; 76.1	3.79; 74.2	3.86; 78.9	n.d	n.d	2.10; 23.1 /2.16; 23.1
C3	4.47; 104.3	3.37; 75.4	4.89; 79.9	3.78; 79.1	n.d	n.d	2.17;23.2
C23	4.81; 102.2	4.81; 74.2	5.17; 74.1	4.05; 77.9	n.d	n.d	(2) 2.10; 22.9/ (2)2.12; 23.0
C3MeGlcA	4.57; 104.2	3.48; 73.6	4.98; 78.1	3.94;78.1	n.d	n.d	2.15; 23.3
MeGlcA	5.17; 96.6	3.56; 74.4	3.53; 73.3	n.d	n.d	n.d	-
α	5.18; 94.8	3.56;74.2	3.53;73.7	n.d	n.d	n.d	-
β	4.56; 99.3	3.25;76.7	3.52;77.9	3.72;79.7	n.d	n.d	-

f



g

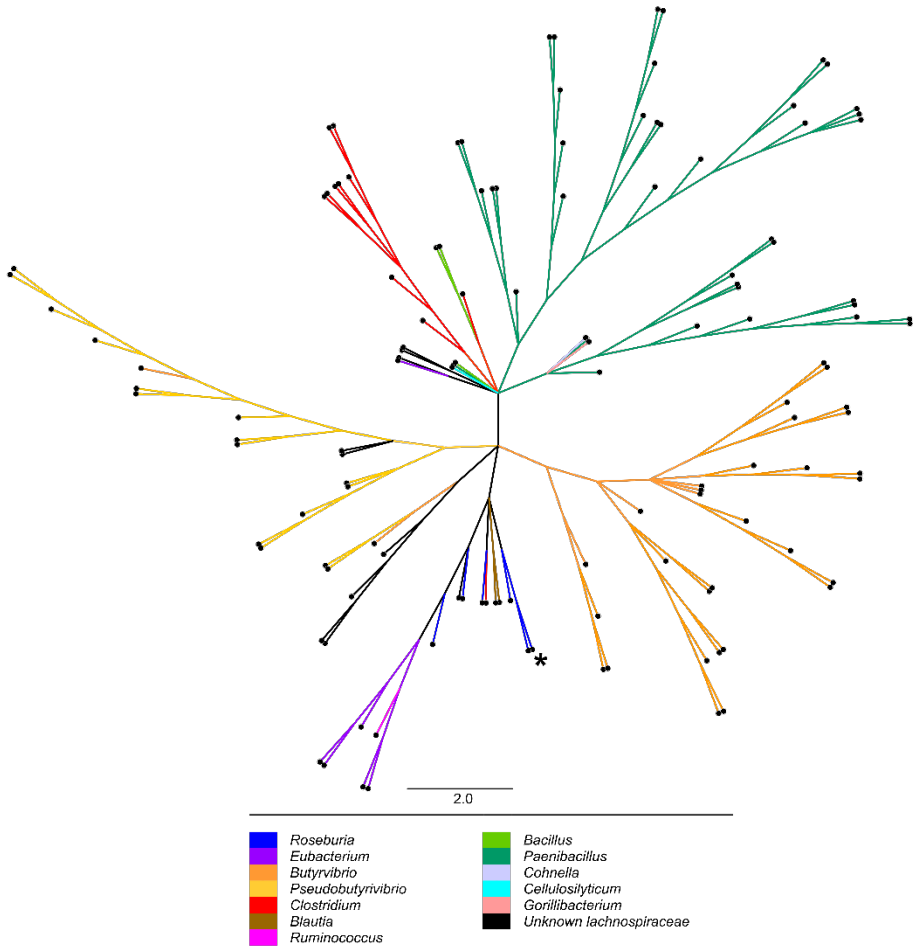


h

Esterase activity for *Ri*AXE

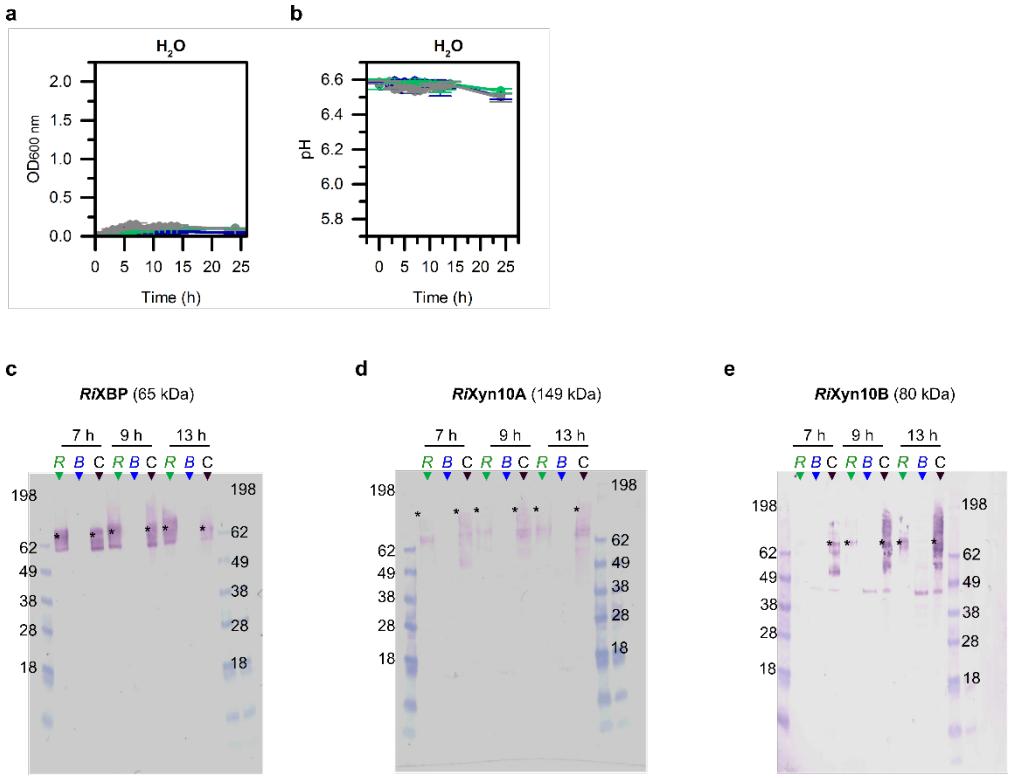
	AcBGX	AcAGX	AcSGGM	Cellulose mono acetate	AcChitin	InWAX
***Ri*AXE**	++	++	+	+	-	-
***Ri*AXE+ *Ri*Agu115A**	+++	n.d.	n.d.	n.d.	n.d.	n.d.

i



Supplementary Figure 8 Activity, specificity and taxonomic distribution of the novel xylan acetyl esterase *Ri*AXE. (**a**) Specific activity of *Ri*AXE on *p*NP-acetate. The data are means of a triplicate with standard deviations. (**b**) Time course deacetylation of AcBGX treated with *Ri*Xyn10A and *Ri*Agu115A by *Ri*AXE. (**c,d**) Estimated rates of deacetylation by *Ri*AXE on AcBGX and AcSpruce mannan with D_2O as a solvent, which may influence the absolute reaction rate (**e**) Assignment of chemical shifts. The most dominating signals of monosaccharide residues of *Ri*Xyn10A treated AcBGX assigned by by starting at the anomeric signal and then following the proton-proton connectivity using TOCSY, DQF-COSY/IP-COSY, ^{13}C H2BC and ^{13}C HSQC-[^{1}H,^{1}H]TOCSY. ^{13}C-HSQC is used for assigning the carbon chemical shifts. The ^{13}C HMBC spectrum provides information of the position of the acetyl group in the monosaccharide residues. The following designations are used: X; xylose, C2; 2-*O*-acetylated xylose, C3; 3-*O*-acetylated xylose, C23; 2,3-di-*O*-acetylated xylose; C3-MeGlcA; 4-*O*-methylglucuronic acid 2-*O*-substituted and 3-*O*-acetylated xylose, C23(2); signal for the 2-*O*-acetylated of C23; α, the α anomer of xylose, β, the β anomer of xylose. **-;** refers to the ring carbon number for the monosaccharides. (**f**) ^{13}C HSQC spectrum of *Ri*Xyn10A treated AcBGX. Top panel shows the acetyl region with the 1D proton projection and the bottom panel shows the spectral region for anomeric and *O*-acetylated xylose signals. While *Ri*Xyn10a treatment enhances signal-to-noise of resonances in the NMR spectra for the assignment, it also increases the total number of observable individual signals. (**g**) ^{13}C HSQC spectra for *O*-acetylated regions before (top panel) and after deactylation by *Ri*AXE (bottom panel). Nearly complete deacetylation of the AcBGX sample is accomplished during the time resolved NMR experiment. Chemical shift of the most dominating signal for the monosaccharide residue is mark by '**+**', peaks encircled by dotted lines indicate cluster of chemical shifts likely to belong to the same type of monosaccharide residue as dominating signal. (**h**) Esterase activity for *Ri*AXE towards different acetylated substrates analyzed using MALDI ToF MS. +++: Complete deacetylation, ++: almost complete acetylation (1 ≥ acetylations/oligosaccharide), +: minor deacetylation (1-2 acetylations/oligosaccharide), -: no deacetylation. (**i**) Phylogenetic tree of *Ri*AXE and homologs identified by a BLASTP search against the non-redundant database. Sequences with coverage >86% and identity >42% were selected. All sequences were identified as members of the Firmicutes phylum. The resulting 131 protein sequences were aligned using Muscle[59] and a phylogenetic tree constructed by the maximum likelihood algorithm in MEGA7[60]. Bootstraps were performed with 500 replicates. The phylogenetic tree was visualized using Figtree (http://tree.bio.ed.ac.uk/software/figtree). The asterisk indicate *Ri*AXE.

Supplementary Figure 9



Supplementary Figure 9 Co-culture experiment with *R. intestinalis* and *B. ovatus.* (**a**-**b**) Growth curves for monoculture and co-cultures after growth of *R. intestinalis* and *B. ovatus* with water as controls instead of carbon source. The western blots were carried out with (**c**) anti-*Ri*XBP, (**d**) anti-*Ri*Xyn10A, (**e**) anti-*Ri*Xyn10B. R: *R. intestinalis*, B: *B. ovatus,* C: co-culture of *R. intestinalis* and *B. ovatus.* Asterisk denotes the position of the band based on theoretical molecular mass. The molecular markers size is shown in kDa. Lower molecular mass signals than expected indicate proteolytic cleavage occurring particularly with the multi-modular *Ri*Xyn10A.