

Differential bacterial capture and transport preferences facilitate co-growth on dietary xylan in the human gut

Maria Louise Leth¹, Morten Ejby¹, Christopher Workman¹, David Adrian Ewald¹, Signe Schultz Pedersen¹, Claus Sternberg¹, Martin Iain Bahl², Tine Rask Licht², Finn Lillelund Aachmann³, Bjørge Westereng⁴ and Maher Abou Hachem^{1*}

Metabolism of dietary glycans is pivotal in shaping the human gut microbiota. However, the mechanisms that promote competition for glycans among gut commensals remain unclear. *Roseburia intestinalis*, an abundant butyrate-producing Firmicute, is a key degrader of the major dietary fibre xylan. Despite the association of this taxon to a healthy microbiota, insight is lacking into its glycan utilization machinery. Here, we investigate the apparatus that confers *R. intestinalis* growth on different xylns. *R. intestinalis* displays a large cell-attached modular xylanase that promotes multivalent and dynamic association to xylan via four xylan-binding modules. This xylanase operates in concert with an ATP-binding cassette transporter to mediate breakdown and selective internalization of xylan fragments. The transport protein of *R. intestinalis* prefers oligomers of 4–5 xylosyl units, whereas the counterpart from a model xylan-degrading *Bacteroides* commensal targets larger ligands. Although *R. intestinalis* and the *Bacteroides* competitor co-grew in a mixed culture on xylan, *R. intestinalis* dominated on the preferred transport substrate xylo-tetraose. These findings highlight the differentiation of capture and transport preferences as a possible strategy to facilitate co-growth on abundant dietary fibres and may offer a unique route to manipulate the microbiota based on glycan transport preferences in therapeutic interventions to boost distinct taxa.

The human gut microbiota (HGM) is recognized as a determinant of human health and metabolic homeostasis^{1,2}. Specific signatures of the HGM are associated with local and systemic disorders, including irritable bowel disease, obesity, type 2 diabetes and colon cancer³. The composition of the HGM is greatly affected by non-digestible dietary glycans^{4,5}. Only a few species of the HGM are equipped to deconstruct and ferment distinct complex glycans into short-chain fatty acids (SCFAs)⁶. The effect of SCFAs on host health and physiology remains an important aspect of the microbiota–host interaction. The SCFA butyrate, the preferred energy source for colonocytes, possesses anti-inflammatory roles and reduces the risks of colon cancer and enteric colitis^{7–10}. Butyrate-producing Firmicutes are abundant in healthy individuals, but are markedly reduced in patients with inflammatory disorders^{11,12}. Butyrate producers including *Roseburia* spp. are increased in patients with metabolic syndrome after faecal transfer therapy and correlate positively to improvement of insulin resistance¹³. Investigations of the metabolic preferences of butyrate producers and their interplay with major HGM commensals are instrumental to develop interventions targeting butyrate deficiency-related disorders.

Roseburia is a common genus of *Clostridium* cluster XIVa within Firmicutes that harbour prevalent butyrate producers^{14,15}. This taxon adheres to mucin, reflecting intimate association with the host¹⁶. *Roseburia intestinalis* strains encode an impressive repertoire of carbohydrate-active enzymes (CAZymes) compared to most Firmicutes¹⁷. *R. intestinalis*, the taxonomically related *Eubacterium rectale*, and species from *Bacteroides* are the only

known HGM taxa that utilize the hemicellulose xylan^{18–20}. Xylan is particularly abundant in cereal grains (arabinoxylan) but is also found in fruits and vegetables (glucuronoxylan)²¹ (Fig. 1a). Xylan utilization by *Bacteroides*-dominant gut commensals has been investigated in detail^{22,23}, but similar knowledge is lacking for Firmicutes counterparts.

Here, we show that *R. intestinalis* L1-82 grows on dietary-relevant xylns, with a preference for cereal arabinoxylans. The growth is mediated by a multi-modular cell-attached xylanase and by an ATP-binding cassette (ABC) transporter. We have characterized the xylanolytic enzymes and the transport protein, which enabled modelling xylan utilization by *R. intestinalis* and the identification of two previously undescribed xylan-specific CAZyme families. *R. intestinalis* efficiently competes with a model xylan degrader from *Bacteroides* when grown on soluble and insoluble xylns. Strikingly, transport proteins that confer xylo-oligosaccharides capture in *R. intestinalis* and *Bacteroides* targeted ligands of different sizes, thus markedly reducing the competition for preferred ligands by either taxon. These results emphasize the competitiveness of butyrate-producing Firmicutes in targeting key dietary fibres such as xylan and highlight differential glycan capture and transport as an important feature in co-growth on abundant dietary fibres such as xylan.

Results

Inducible cell-attached xylanase activity mediates growth of *R. intestinalis* on substituted xylns. Anaerobic growth of *R. intestinalis* L1-82 was measured on soluble and insoluble xylns (Fig. 1b–d). *R. intestinalis* L1-82 grows rapidly on soluble xylns, especially wheat

¹Department of Biotechnology and Biomedicine, Technical University of Denmark, Lyngby, Denmark. ²National Food Institute, Technical University of Denmark, Lyngby, Denmark. ³NOBIPO, Department of Biotechnology and Food Science, NTNU Norwegian University of Science and Technology, Trondheim, Norway. ⁴Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås, Norway. *e-mail: maha@bio.dtu.dk

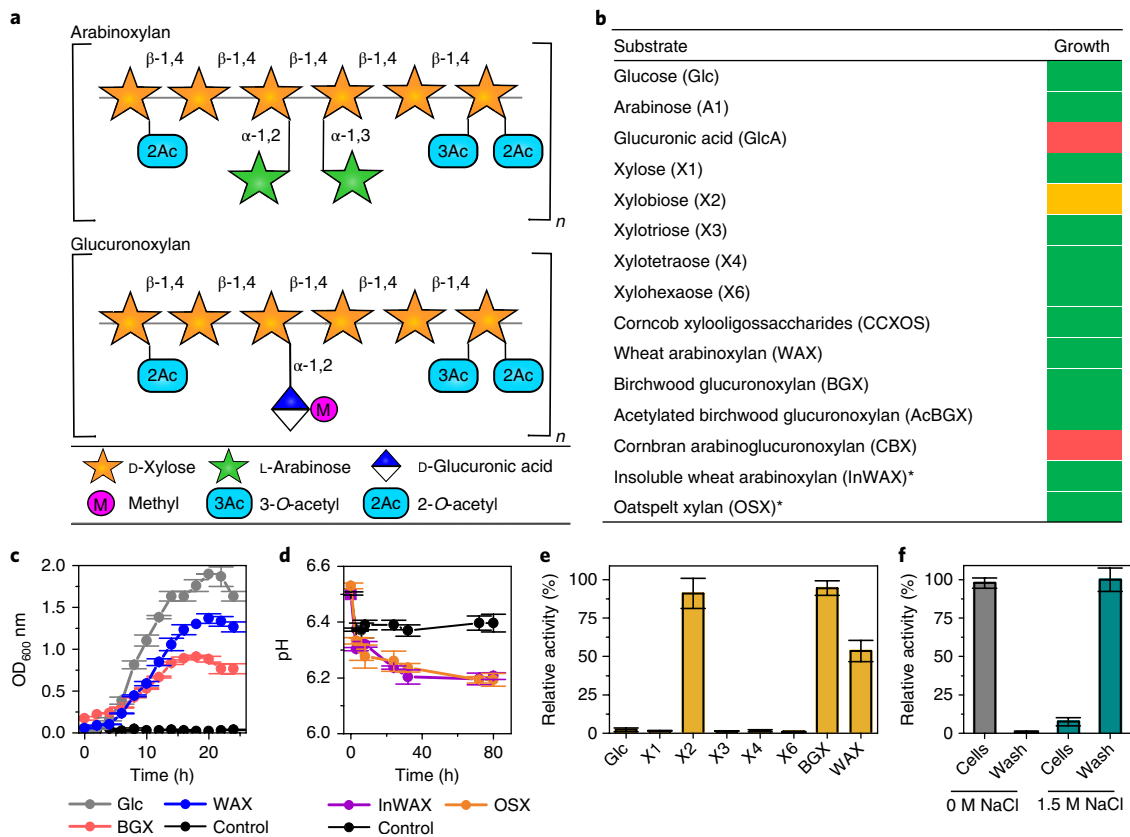


Fig. 1 | Growth of *R. intestinalis* and induction of extracellular activity. **a**, Schematic representation of cereal arabinoxylan and glucuronoxylan present in the dicot cell wall, for example, in fruits and vegetables. **b**, Growth level for 18 hours on xylans, oligosaccharide and monosaccharides thereof, with glucose as a control. Green: OD₆₀₀ increase of >1.0 for soluble substrates and a pH drop of >0.3 for insoluble xylans; yellow: 0.3 < ΔOD₆₀₀ < 0.5; red: ΔOD₆₀₀ < 0.1. Asterisks indicate insoluble xylans. Data are means of triplicates. **c**, Growth curves on glucose (Glc), WAX, BGX and a no-carbon-source control. **d**, Growth on InWAX and oatspelt xylan (OSX) and no-carbon-source control. All growth measurements (**c,d**) are means of triplicates with standard deviations. **e**, Xylanase activity of *R. intestinalis* cells 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 cell fractions to verify the localization of the enzymes. Xylanase activity (**e,f**) was measured using the DNS-reducing sugar assay, and data are triplicates with standard deviations.

arabinoxylan (WAX; $\mu_{\max} = 0.26 \text{ h}^{-1}$), compared to birch glucuronoxylan (BGX; $\mu_{\max} = 0.13 \text{ h}^{-1}$) (Fig. 1c). This bacterium also utilizes highly acetylated xylans and insoluble cereal arabinoxylans from wheat (InWAX) and oat spelt, but not corn bran glucuronoxylan (CBX). Xylo-oligosaccharides and xylan-derived monosaccharides (except glucuronic acid) were also utilized. Extracellular endo-1,4- β -xylanase (hereafter xylanase) activity was induced upon growth on BGX, WAX and the poor growth substrate xylobiose (X2) (Fig. 1e). Xylanase activity was cell attached, but was released in the presence of high-salt concentration (Fig. 1f), suggesting non-covalent attachment.

Genes encoding an ABC transporter and a multi-modular xylanase are top upregulated in response to growth on xylan. We performed an RNA sequencing (RNA-seq) transcriptional analysis of *R. intestinalis* grown on WAX, BGX, xylose and glucose. Of the 4,777 predicted genes, 1–3.5% were highly upregulated (\log_2 fold change > 5) on xylans compared to glucose (Supplementary Table 1) and the majority were involved in carbohydrate metabolism. Besides a separate locus encoding a multi-modular xylanase of glycoside hydrolase family 10 (GH10 according to CAZy classification²⁴), the top genes in the xylan transcriptomes cluster on a single locus (Fig. 2a,b) that contains 11 genes, including 4 xylanolytic CAZymes: GH43, GH115, GH8 and GH3. Only one (ROSINTL182_08192, LacI type) of three transcriptional regulators

was highly upregulated. Strikingly, the most upregulated gene in the xylan transcriptomes encodes a solute-binding protein (SBP) of an ABC transporter. The permease genes of this transporter were among the top six upregulated by xylans. Signal peptides were only predicted for the xylanase and the transporter SBP, which is consistent with the extracellular breakdown of xylan followed by capture and uptake of xylo-oligosaccharides by the ABC transporter. The expression and the localization of the SBP and the xylanase at the cell surface were corroborated using immunofluorescence microscopy (Fig. 2c). Two additional loci, unique to the *R. intestinalis* L1-82 strain, were also upregulated albeit markedly less (Supplementary Fig. 1a–d). One of these loci encodes a second active GH10 xylanase, which is expressed at the cell surface (Supplementary Fig. 1e). The transcriptomic analyses enabled the assignment of the xylose ABC transporter and the genes involved in intracellular metabolism of xylose, arabinose and glucuronic acid (Supplementary Fig. 2).

A previously undescribed family of binding modules confers extended and dynamic xylan binding to the multi-modular xylanase. The highly upregulated *RiXyn10A*, which is conserved within *R. intestinalis* (Supplementary Fig. 3), is one of the largest known xylanases from the HGM (Supplementary Table 2). *RiXyn10A* comprises an amino-terminal unassigned domain (residues 28–165), a CBM22 xylan-binding module, a GH10 catalytic module, two tandem CBM9 xylan-binding modules, a bacterial immunoglobulin-like

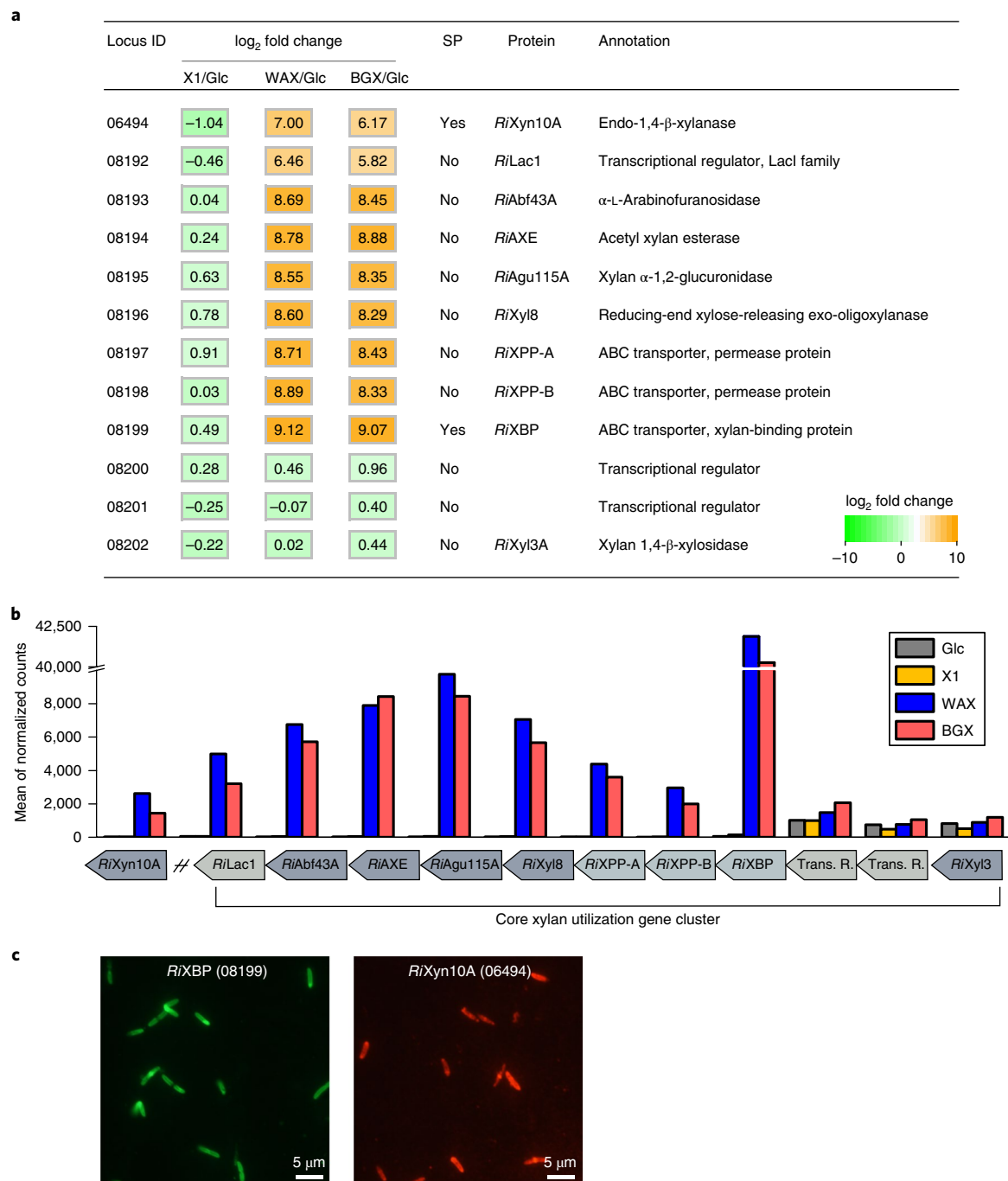


Fig. 2 | The core xylan utilization apparatus of *R. intestinalis*. **a**, The RNA-seq heatmap depicts log₂ fold changes of the top upregulated xylan utilization genes expressed by cells grown on xylose (X1), WAX and BGX relative to glucose. Experiments are performed in biological triplicates, and locus tag numbers ROSINTL182_xxxxx are abbreviated with the last numbers after the underscore. **b**, Gene expression depicted as the mean of the normalized Deseq2 gene counts for the core xylan utilization genes shown in **a**. Trans. R., transcriptional regulator; SP, signal peptide. **c**, Representative micrographs showing the extracellular localization of *RiXBP* and *RiXyn10A*, the SBP of the xylo-oligosaccharide-specific ABC transporter and the xylanase, respectively, were visualized by fluorescence microscopy of *R. intestinalis* cells in biological triplicates using primary antibodies against these two proteins.

domain group 2 (pfam02368)²⁵ and a *Listeria-Bacteroides* repeat domain (pfam09479)²⁶. The two latter positively charged domains (residues 1,100–1,356; pI > 10) (Fig. 3a) probably mediate attachment of the enzyme to the cell^{25–27}.

RiXyn10A incubated with BGX, WAX and InWAX generated linear and decorated oligosaccharides (Fig. 3b,c), but was inactive on highly and heterogeneously substituted arabinoglucuronoxylan from corn bran, consistent with the growth profile (Fig. 1b). The enzyme was inactive on X2 and showed poor activity on xylotri-ose

(X3) (Supplementary Fig. 4a). Xylotetraose (X4) and xylopentaose (X5) were efficiently hydrolysed, revealing the requirement for the occupancy of at least four substrate-binding subsites for high turnover.

A BLASTp search of the N-terminal domain (previously undescribed carbohydrate-binding module (CBMx)) against UniProt gave no hits indicating the lack of homologues with assigned function. CBMx confers affinity to xylan based on a two-times higher Michaelis constant (K_m) when CBMx was deleted from *RiXyn10A*

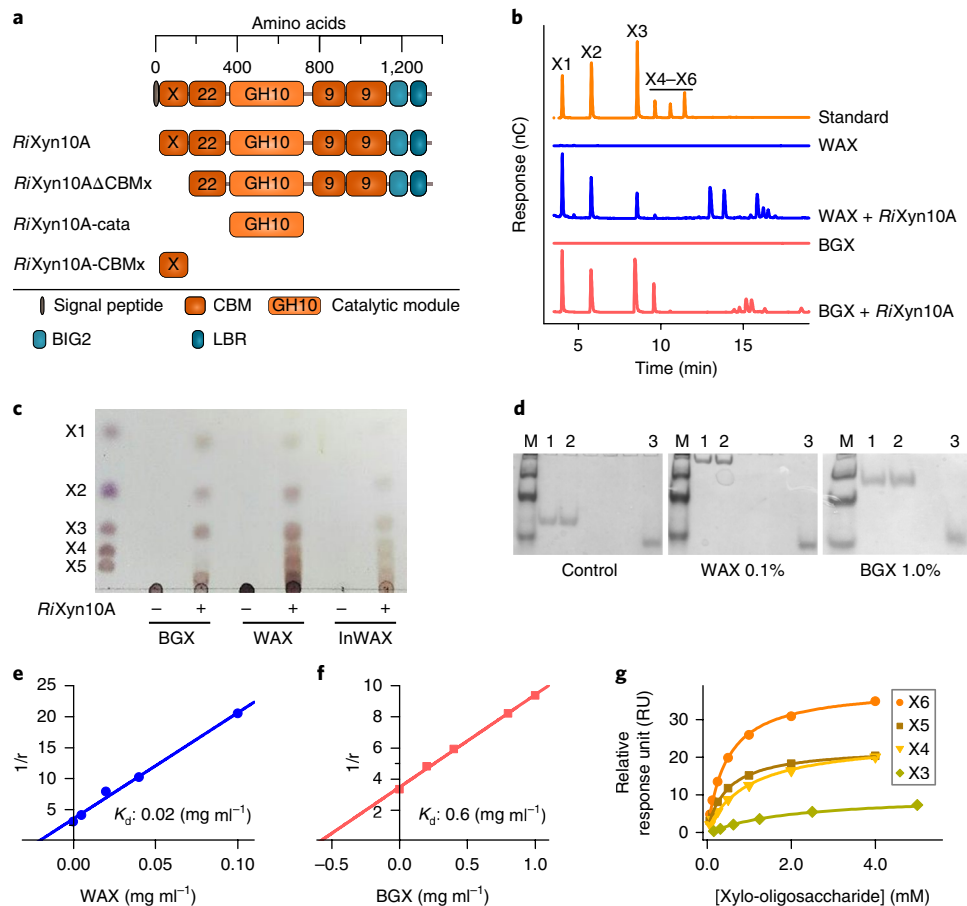


Fig. 3 | A low-affinity xylan-binding module mediates extended xylan binding to the xylanase *RiXyn10A*. **a**, Domain organization of *RiXyn10A* and truncated variants, all produced recombinantly without the native signal peptide. BIG2, bacterial immunoglobulin-like domain group 2; CBM, carbohydrate-binding module; CBMx, previously undescribed CBM; LBR, *Listeria-Bacteroides* repeat domain. **b**, Xylanase activity of *RiXyn10A* on WAX and BGX assayed by HPAEC-PAD. The peaks eluting after X6 are probably decorated xylo-oligosaccharides. **c**, *RiXyn10A* activity analysed using thin layer chromatography. The - and + indicate absence or presence of *RiXyn10A*, respectively. **d**, Binding of *RiXyn10A*-CBMx to the negative control (no polysaccharide), WAX or BGX xylans analysed using affinity electrophoresis. The uncropped gels are in Supplementary Fig. 4h. Lanes 1+2, *RiXyn10A*-CBMx (3.0 μ g); lane 3, β -lactoglobulin negative control (1.5 μ g); M, marker. **e, f**, Binding data from **d** is depicted as a plot of $1/r$ against xylan concentration, where r is the relative migration distance of *RiXyn10A*-CBMx in the presence of xylans in the gel. The K_d from this analysis is shown in each plot. **g**, Binding isotherms of *RiXyn10*-CBMx to xylo-oligosaccharides. Solid lines are fits of a one-binding site model to the SPR sensograms. Experiments were performed in duplicates, except in **g**, which were performed in triplicates.

(Table 1 and Supplementary Fig. 4b–g). Affinity electrophoresis established CBMx as a xylan-binding module and revealed a 30-fold higher dissociation constant (K_d) on WAX compared with BGX

Table 1 | Xylan hydrolysis kinetics of *RiXyn10A* and truncated variants

Substrate	Enzyme	K_m (mg ml ⁻¹)	k_{cat} (s ⁻¹)	k_{cat}/K_m (ml mg ⁻¹ s ⁻¹)
WAX	<i>RiXyn10A</i>	2.4 ± 0.4	202 ± 14	84
	<i>RiXyn10A</i> ΔCBMx	5.5 ± 1.2	386 ± 47	71
	<i>RiXyn10A</i> -cata	ND	ND	34
BGX	<i>RiXyn10A</i>	3.3 ± 0.7	196 ± 18	59
	<i>RiXyn10A</i> ΔCBMx	6.6 ± 1.6	369 ± 45	56
	<i>RiXyn10A</i> -cata	ND	ND	16

Kinetics of the *RiXyn10A*-cata are not modelled by the Michaelis–Menten expression, and catalytic efficiencies are estimated from the linear regression of the initial rate data. Data are means of triplicates with standard deviations. ND indicates low substrate affinity precluding determination of the kinetic parameters.

(Fig. 3d–f and Supplementary Fig. 4h). Surface plasmon resonance (SPR) analysis revealed the highest affinity towards xylohexaose (X6) (Table 2, Fig. 3g and Supplementary Fig. 5a–e). The relatively low binding affinity to X6 ($K_d \approx 0.5$ mM) was corroborated using isothermal titration calorimetry (ITC) (Table 2, Supplementary Fig. 5f and Supplementary Table 3). Deleting CBMx decreased the average K_d of *RiXyn10A* from 128 μ M to 65.4 μ M (*RiXyn10A*ΔCBMx) (Table 2 and Supplementary Fig. 5g–j), asserting that one or more of the other CBMs possess a higher affinity compared to the N-terminal module. Homologues (sequence identity: 55–27%) of CBMx are present mainly in other *Clostridium* XIVa cluster Firmicutes (Supplementary Table 4).

Preference of the binding protein of the ABC transporter that mediates uptake of xylo-oligosaccharides. Complex xylo-oligosaccharides decorated with arabinosyl and 4-O-methylglucuronosyl are produced by *RiXyn10*, based on the decrease levels in some of these products and the increase levels in arabinose, 4-O-methylglucuronic acid (MeGlcA) and un-substituted xylo-oligosaccharides after treatment with debranching enzymes (see the next section). Transcriptional analysis (Fig. 2a) identified a

Table 2 | Binding parameters of the xylan-binding module CBMx and RiXyn10A and variants

Variant	Ligand	K_d (μM)
RiXyn10A-CBMx	WAX	0.02 ^a
	BGX	0.6 ^a
	X6	479 ± 26
	X6	413 ± 125 ^b
	X5	490 ± 15
	X4	998 ± 42
	X3	1,900 ± 220
	Man6	ND
RiXyn10A	X6	128 ± 7.1
RiXyn10AΔCBMx	X6	65.4 ± 8.4

K_d determined by SPR are the means of a duplicate with the standard deviations. ^a K_d (mg ml^{-1}) was determined by affinity electrophoresis. ^b K_d was determined by ITC. The SPR experiments are performed in triplicates and the data are the means with standard deviations. The affinity electrophoresis and the ITC data are from single experiments. The error estimate in the ITC experiment is from the fit of a one-binding site model to the binding isotherm.

plausible ABC transporter of xylo-oligosaccharides. The preference of SBPs associated with oligosaccharide-specific ABC transporters correlates with the bacterial uptake preference^{28,29}. We measured the affinity of RiXBP, the SBP of the top upregulated ABC transporter on xylo-oligomers (Table 3 and Supplementary Fig. 6a–i). The preferred un-substituted ligand was X5 followed by X4, and the affinity decreased steeply for smaller or larger oligosaccharides. Internal arabinosyl decorations (AX4) were preferred based on the 2.4-times higher affinity compared to X4. The tolerance and recognition of arabinosylated ligands are supported by the good growth on WAX. In summary, RiXBP is selective in capturing branched xylo-oligosaccharides with preference for a backbone of 4–5 xylose residues.

The concentrations of oligosaccharides in spent supernatants from *R. intestinalis* cultures during growth on xylan were too low for reliable detection using high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD), which suggests an efficient uptake of oligomeric products (Supplementary Fig. 6j,k).

***R. intestinalis* degrades internalized decorated xylo-oligosaccharides by the concerted action of three hydrolases and a previously undescribed family of acetyl esterases.** Xylo-oligosaccharides are degraded in the cytoplasm after uptake. We characterized the α -glucuronidase RiAgu115A (GH115), the α -L-arabinofuranosidase RiAbf43A (GH43), two xylosidases RiXyl8 (GH8) and RiXyl3A (GH3) as well as RiAXE (ROSITNL182_08194; GenBank accession [EEU99941.1](https://www.ncbi.nlm.nih.gov/nuclot/EEU99941.1)) from the core xylan utilization locus.

RiAgu115A released MeGlcA from glucuronoxylans (BGX and beechwood glucuronoxylan) and from BGX pretreated with RiXyn10A (Fig. 4a,c–f and Supplementary Fig. 7a). The k_{cat}/K_m of RiAgu115A was 16-fold higher on glucuronoxylan hydrolysate compared to intact glucuronoxylan (Supplementary Table 5), consistent with the intracellular enzyme localization. This enzyme also cleaves MeGlcA decorations at the xylosyl penultimate to the reducing end (Supplementary Fig. 7b,c), but its activity was blocked by acetylations (Fig. 4d). The released MeGlcA is metabolized intracellularly in agreement with the transcriptomic data (Supplementary Fig. 2a). The lack of growth on glucuronic acid (Fig. 1b) is probably due to the lack of an uptake system for this monosaccharide.

RiAbf43A is an α -L-arabinofuranosidase that exclusively releases arabinose from WAX (Fig. 4a and Supplementary Fig. 7d). Kinetic analysis against WAX and AX4 (Supplementary Table 6) revealed

Table 3 | The preference of the xylo-oligosaccharide transport protein from *R. intestinalis*

Ligand	K_d (μM)	N_0	ΔH (kcal per mol)	$T\Delta S$ (kcal per mol)	ΔG (kcal per mol)
X6	112.7	1.19	−9.01	−3.6	−5.4
X5	10.3	0.86	−13.54	−6.7	−6.8
X4	16.5	0.68	−12.8	−6.3	−6.5
X3	225.7	0.58	−21.1	−16.1	−5.0
X2	ND				
AX3	215.5	0.26	−44.3	−39.4	−4.9
AX4	6.8	0.58	−12.3	−7.0	−5.3

Binding energetics of the transport protein RiXBP to xylo-oligosaccharides determined by ITC. Data are the means of duplicate experiments. N_0 , ΔH , $T\Delta S$ and ΔG denote the binding stoichiometry, enthalpy, entropy and Gibbs free energy, respectively. ND indicates affinity too low to be determined. The cartoons to the right side depict the analysed substrates using the same code as in Fig. 1a, and the chemical structures of arabinosyl-decorated oligomers AX3 and AX4 are shown in Supplementary Fig. 6h,i.

recognition of internal arabinosyl substitutions. Degradation of WAX by xylanases results in the production of significant amounts of double-arabinosyl-substituted oligomers. The lack of accumulation of these to detectable levels (Supplementary Fig. 6j) is suggestive of their uptake and debranching by RiAbf43A or other upregulated arabinofuranosidases.

Both RiXyl3A and RiXyl8 generated xylose from xylo-oligosaccharides, but lacked activity towards xylan (Supplementary Tables 7,8 and Supplementary Fig. 7d–g). RiXyl3A degraded xylo-oligosaccharides to xylose, whereas RiXyl8 was inactive towards X2. Reduction of xylo-oligosaccharides abolished the activity of RiXyl8, assigning it as a reducing-end β -xylosidase³⁰ (Supplementary Fig. 7h), in contrast to RiXyl3A that recognizes non-reducing xylosyl units. The concerted and overlapping activities of these enzymes (Supplementary Fig. 7) result in rapid depolymerization of arabinosyl- and MeGlcA-decorated xylo-oligosaccharides. Induction of extracellular xylanase activity by X2, but not by X3–X6, is intriguing as X2 is transiently formed during cytoplasmic degradation of X3–X6. A possible explanation is that the induction is mediated by X2 binding to an extracellular regulatory protein, similar to the xylose sensory system in clostridial Firmicutes³¹.

RiAXE, which lacks functionally described homologues, was highly upregulated on xylans (Fig. 2a). This enzyme shares conserved residues with the SGNH lipase esterase superfamily (Pfam [cd00229](https://www.ncbi.nlm.nih.gov/CLUSTAL/Accession/cd00229)), but possesses low sequence identities to esterases from CAZy. Homologues of this enzyme are encoded by *Clostridium* cluster XIVa strains from the human gut and a range of Firmicutes (Supplementary Fig. 8a). Assaying RiAXE activity towards acetylated BGX (AcBGX) oligosaccharides (generated with RiXyn10A) using NMR revealed efficient deacetylation of both 2-O-acetylated xylose (C2) and 3-O-acetylated xylose (C3), but with a preference for C2 decorations (Fig. 4b, Supplementary Tables 9,10 and Supplementary Fig. 8). RiAXE left a single acetyl group on the AcBGX oligosaccharides (Fig. 4e). Inclusion of RiAgu115A in this reaction resulted in complete deacetylation (Fig. 4f), suggesting that close MeGlcA decorations protect closeby acetylations. Analysis of the deacetylation rates unveiled the concerted action with RiAgu115A and preference to hydrolysates of RiXyn10A rather than intact xylan (Supplementary Fig. 8b,c). RiAXE specifically recognizes acetylations on xylosyl units based on the lack of activity on acetylated chitin and very low activity on acetylated mannan and cellulose monoacetate (Supplementary Table 11). Accordingly, RiAXE is an efficient xylan-specific representative of a previously undescribed acetyl esterase family.

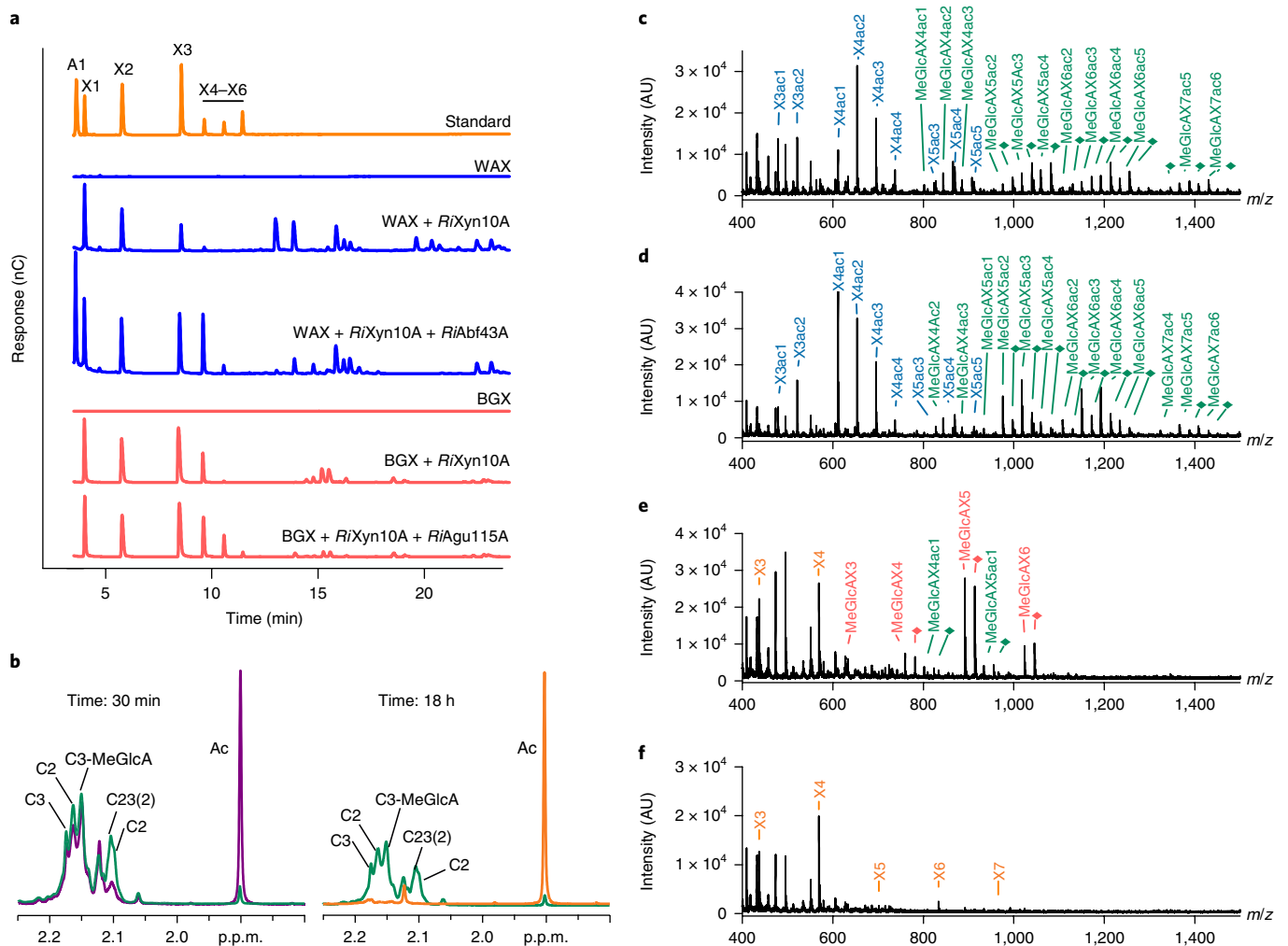


Fig. 4 | Intracellular xylo-oligosaccharide depolymerization. a, α -L-Arabinofuranosidase and α -glucuronidase activity on WAX and BGX for *RiAbf43A* and *RiAgu115A*, respectively, based on HPAEC-PAD analysis. **b**, Time-resolved NMR for *RiAXE* deacetylation of AcBGX treated with *RiXyn10A* and *RiAgu115A*. Deacetylation time course for the first 30 min and after 18 h (green: 0 min, purple: 30 min and 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 show nearly complete deacetylation after 18 h. The signal at 2.13 p.p.m. is probably attributed to another acetylated sugar residue. Acetyl groups are designated as: C2; C3; C23, 2,3-di-O-acetylated xylose; C3-MeGlcA, 4-O-methyl glucuronic acid 2-O-substituted and 3-O-acetylated xylose; C23(2), signal for the 2-O-acetylated C23. The assignment of the acetylated sugar signals was based on homonuclear and heteronuclear NMR correlation experiments (Supplementary Fig. 8). **c-f**, Hydrolysis products from AcBGX by *RiXyn10A* (**c**), *RiXyn10A* and *RiAgu115A* (**d**), *RiXyn10A* and *RiAXE* (**e**), and *RiXyn10A*, *RiAgu115A* and *RiAXE* (**f**). Enzyme action was analysed by MALDI-TOF mass spectrometry. Xylo-oligosaccharides decorated with acetyl and MeGlcA are in green, acetyl in blue, MeGlcA in red and no sidechains in orange. Di-sodium adducts of a MeGlcA-decorated oligosaccharides (diamonds) are coloured as their corresponding single sodium adducts. Experiments were performed in duplicates, except for the NMR, where they were performed once. AU, arbitrary units.

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

***R. intestinalis* competes with *Bacteroides* for xylans.** The growth potential of *R. intestinalis* was compared to the efficient xylan-degrader *Bacteroides ovatus*²², by observing the growth of individual cultures and in co-culture. Both strains displayed similar growth on xylan as a carbon source (Fig. 5b–d and Supplementary Fig. 9a,b). In competition, both strains appeared to grow well on xylans, whereas *R. intestinalis* dominated the co-culture on X4 after 7 hours of growth (Fig. 5e–h). The results indicate that *R. intestinalis* is an efficient primary degrader of xylan that is able to compete with *B. ovatus* and outcompete this bacterium on preferred smaller xylo-oligosaccharides.

Discussion

The human gut is dominated by bacteria from the Firmicutes and Bacteroidetes phyla. Firmicutes are regarded as metabolic specialists, whereas Bacteroidetes (mainly *Bacteroides*) are considered generalists based on narrow versus broad glycan utilization capabilities, respectively⁶. The size and diversity of encoded CAZymes reflect these metabolic labels, which applies to *R. intestinalis*, based on the limited glycan growth profiles⁵. However, this species possesses distinctively larger CAZymes than most known gut clostridial Firmicutes¹⁷. *R. intestinalis* has been proposed to be a key xylan degrader in the human gut along with specific species of *Bacteroides*^{18,19}. Enumeration of *R. intestinalis* on xylans including wheat bran is reported in vitro and in vivo^{20,32}. However, insight is lacking on the preferences and the molecular machinery evolved by *R. intestinalis* to target xylan compared to *Bacteroides* members. In this study, we present a model underpinning the molecular basis

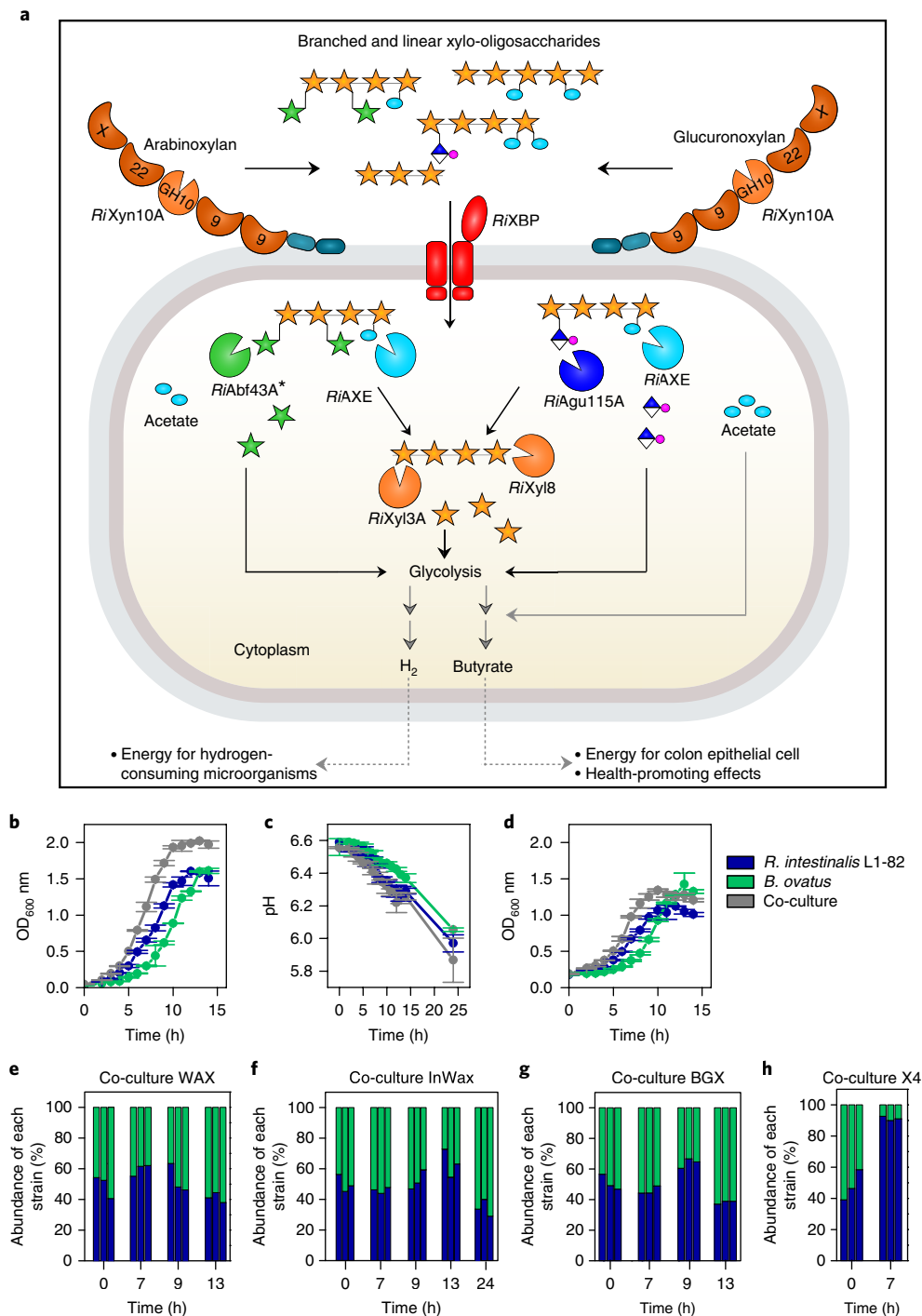


Fig. 5 | Model for xylan utilization by *R. intestinalis* and competition assay with *B. ovatus*. **a**, *RiXyn10A* on the cell surface efficiently captures diet-derived acetylated arabinoxylan and acetylated glucuronoxylan by its CBMs and hydrolyses it into linear and decorated xylo-oligosaccharides, which are subsequently captured by *RiXBP* for uptake into the cytoplasm. Internalized xylo-oligosaccharides are debranched and hydrolysed into monosaccharides and acetate. Xylose and arabinose are converted to xylulose 5-phosphate before entering the pentose phosphate pathway, whereas MeGlcA is converted to 2-oxo-3-deoxygalactonate 6-phosphate. These precursors enter glycolysis, which generates pyruvate, some of which is converted to butyrate⁵⁴. The asterisk next to *RiAbf43A* indicates activity on both α -1,2-linked and α -1,3-linked L-arabinose. The black solid arrows show steps established or confirmed in this study. The grey solid arrows indicate steps for the production of H₂ and butyrate by *R. intestinalis*¹⁵, which has been demonstrated for this species during xylose fermentation¹⁹. The grey dashed arrows indicate that H₂ and butyrate are externalized by unknown mechanisms. To make the model more general for the *R. intestinalis* species, the second less upregulated extracellular xylanase *RiXynB*, which is unique for the L1-82 strain, is not included in the model, although it is expressed at the cell surface. The figure keys in Figs. 1a and 3a also apply here. **b-d**, Growth of the monoculture and co-cultures of *R. intestinalis* and *B. ovatus* on WAX (**b**), InWAX (**c**) and BGX (**d**). Data are the means of a biological triplicate with standard deviations. **e-h**, Time course of the relative strain abundance during growth of co-cultures on xylans and X4 determined by quantitative PCR. All data are the means of a biological triplicate.

for xylan utilization by *R. intestinalis* L1-82 as a representative for prevalent butyrate-producing clostridia (Fig. 5a). *R. intestinalis* is a primary degrader equipped with a highly efficient machinery for the utilization of complex dietary xylans, including insoluble arabinoxylan from cereals. 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 decorated xylo-oligosaccharides. In the cytoplasm, internalized oligomers are degraded without loss to competing species. *R. intestinalis* grows on acetylated xylan, which reflects an adaptation to this abundant decoration in dietary xylans (Fig. 1b). Internalized acetylated xylo-oligosaccharides are metabolized by an intracellular previously unknown esterase family that is capable of removing C2, C3 and double acetylations (Fig. 4b and Supplementary Fig. 8).

The extracellular xylanase *RiXyn10A*, the ABC transporter and enzymes conferring cytoplasmic breakdown of xylan oligosaccharides were assigned as the core xylan utilization apparatus of *R. intestinalis* (Fig. 2a,b). This was based on (1) conservation of this apparatus within this species (Supplementary Fig. 3), (2) the highest transcriptional upregulation (Fig. 2a), and (3) biochemical data from the present study. The two additional xylan-upregulated loci in *R. intestinalis* L1-82 (Supplementary Fig. 1) are lacking in the xylan-utilizing *R. intestinalis* XB6B4 and *R. intestinalis* M50/1 (ref. 20). The expression and activity of the xylanase *RiXyn10B*, encoded by an auxiliary locus (Supplementary Table 12 and Supplementary Fig. 1c,e), supports the participation of more than one locus in xylan breakdown in *R. intestinalis* L1-82. Multiplicity of xylan utilization loci is implicated in targeting a larger structural diversity of naturally occurring xylans by *Bacteroides*²², which may also apply for *R. intestinalis*.

The *R. intestinalis* core xylanase *RiXyn10A* mediates the capture and breakdown of arabinoxylan and glucuronoxylan (Figs. 1,3). This enzyme possesses four CBMs from two known and one previously undescribed xylan-binding families, representing the most complex organization of HGM xylanases (Fig. 3a and Supplementary Table 2). This organization is conserved within currently sequenced *R. intestinalis* species, whereas other *Clostridium* XIVa taxa possess simpler enzymes that lack one or more of the *RiXyn10A* CBMs. The N-terminal CBMx of *RiXyn10A* displays approximately 7-fold lower affinity for X6 than the average affinity measured for the enzyme variant lacking this module (Table 2). Nonetheless, CBMx is selective for arabinoxylan and contributes to the overall affinity of the enzyme (Table 1). Low-affinity CBMs may potentiate multivalent cooperative substrate binding, with less reduction of k_{cat} (that is, promote relatively high k_{cat}/k_{off} ³³), compared to higher-affinity counterparts that evoke higher energetic penalty during substrate displacement. The extended binding mediated by the CBMs seems to confer an advantage in the capture and prolonged contact of the enzyme with xylan. Deletion of the CBMs (*RiXyn10A*-cata) caused a substantial decrease in the apparent affinity towards WAX and BGX and deviation from Michaelis–Menten kinetics (Table 1 and Supplementary Fig. 4b–g). These findings are consistent with the importance of CBMs in catalysis under substrate limitations, whereas similar turnover rates were obtained by the catalytic module and the full-length *RiXyn10A* at high (9 mg ml⁻¹) substrate concentrations. Multiplicity and variability of CBMs seem to be a signature of extracellular enzymes from butyrate-producing Firmicutes^{34,35}. By contrast, *Bacteroides* members possess simpler xylanases with an inserted tandem CBM4 repeat within the catalytic module²³. However, xylan capture by *Bacteroides* is additionally orchestrated by moderate-affinity ($K_d \approx 60 \mu\text{M}$) binding proteins²².

R. intestinalis was able to compete with *B. ovatus* for xylans during the log-phase (Fig. 5e–g). Notably, *R. intestinalis* seemed to out-compete *B. ovatus* after propagation of the co-culture (in the late log phase) in fresh medium for two additional passages, underscoring

the competitiveness of the xylan utilization machinery of this Firmicute (Supplementary Fig. 9c). *R. intestinalis* association to insoluble xylans, including wheat bran was reported, whereas *Bacteroides* spp. were enriched in solubilized xylan fractions^{18,36}. The extended binding by *RiXyn10A* may have an important role in the association to insoluble substrates. Indeed, the expression of this enzyme seemed similarly high in the mono- and mixed xylan cultures with *B. ovatus* (Supplementary Fig. 9e), in contrast to the reported downregulation of hydrolases by the taxonomically related *E. rectale* during co-growth with *Bacteroides thetaiotamicron* on a fibre-rich diet in mice³⁷.

The gene encoding the binding protein (*RiXBP*) of the ABC transporter that confers xylo-oligosaccharide uptake in *R. intestinalis* was the most upregulated in the xylan transcriptomes, attesting the crucial role of oligosaccharide capture and transport in the densely populated gut ecological niche. The affinity and size preference of *RiXBP* were found to be very different from the corresponding protein from *Bifidobacterium*²⁹, which prefers shorter xylo-oligosaccharides with a different side-chain decoration pattern. Importantly, striking differences in binding affinities and preference are observed when *RiXBP* is compared to the SusD-like xylan-binding counterpart from *Bacteroides*. Indeed, both SusD-like proteins from *B. ovatus*, which mediate the capture and internalization of xylo-oligosaccharides $\geq X6$ by SusC TonB-dependent permeases, displayed no measurable binding to X4 and X5 (ref. 22), the preferred ligands of *RiXBP* (Table 3). These differential transport protein preferences are likely to be instrumental in establishing competitive uptake profiles to select oligosaccharides of specific sizes and decorations for each taxon. This is supported by the dominance of *R. intestinalis* when the co-culture with *B. ovatus* was grown on X4 (Fig. 5h).

Our study highlights the molecular apparatus that *R. intestinalis*, as a model *Clostridium* group XIVa Firmicute, has evolved to compete for abundant dietary glycans with other dominant commensal bacteria. Strikingly, complex enzymes with multiple ancillary modules mediate multivalent substrate capture and breakdown. Highly expressed ABC transporters mediate efficient capture and uptake of xylan oligosaccharides with a different preference than the corresponding transport systems of currently known competing taxa. These findings suggest that the differentiation of glycan capture and uptake preferences may contribute to co-growth on dietary fibres by establishing competitive windows for distinct fibre breakdown oligomers by different taxa. Our study sets the stage for further investigations to evaluate the generality of this hypothesis.

Methods

Chemicals. All chemicals were of analytical grade. BGX, beechwood glucuronoxylan, corn cob xylo-oligosaccharides and xylose were from Carl Roth (Karlsruhe, Germany). CBX was a kind gift from M. Yadav (US 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. D-Glucuronic acid was from Sigma-Aldrich. L-Arabinose was from VWR International. Xylo-oligosaccharides Longlive 95 P (XOS) were from Shandong Longlive Bio-technology. AcBGX, acetylated aspen glucuronoxylan and acetylated spruce galactoglucomannan were prepared with steam explosion as previously described³⁸. Cellulose acetate was a kind gift from A. Deuschle (University of Hamburg, Germany). Acetylated chitin-oligosaccharides were prepared as previously described³⁹.

Growth experiments and RNA-seq transcriptional analysis. *R. intestinalis* DSM 14610 was grown in a Whitley DG250 Anaerobic Workstation (Don Whitley Scientific) in YCFA medium¹⁴ supplemented with autoclaved-sterilized 0.5% (w/v) carbohydrates. Cultures (5 ml) were grown in triplicates, and optical density at 600 nm (OD₆₀₀) and pH (for insoluble substrates) were measured to assess bacterial growth until the stationary phase was reached. Growth rates were calculated from the exponential growth phase.

For the RNA-seq analysis, total RNA was extracted at mid- to late-log phase (OD₆₀₀ = 0.5–0.7) from biological triplicate cultures (10 ml) grown in YCFA

supplemented with 0.5% (w/v) glucose, xylose, WAX or BGX. Cells were harvested (4,000g for 5 min at 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). Removal of ribosomal RNA and library construction for RNA-seq 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 Beijing Genomics Institute (BGI). 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 were performed using custom python scripts (based on the Biopython SeqIO module⁴⁰) and the FASTX-Toolkit v0.0.13.2 (http://hannonlab.cshl.edu/fastx_toolkit/). Reads were further trimmed with fastq-trimmer and subsequently filtered with fastq_quality_filter with minimum quality score of 30 (-q 30), where 95% of base pairs (bp) meet the minimum quality score (-p 95). The resulting reads were kept if they were longer than 20 bp (-m 20). The *R. intestinalis* L1-82 reference genome and genome annotations are based on assembly GCA_000156535.1_ASM15653v1, obtained from the NCBI (ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/bacteria/Roseburia_intestinalis/). Reads were mapped to the reference genome using Tophat2 (refs ^{41,42}), and gene counts were determined with HTseq⁴³. Differential gene expression was performed using DESeq2 in R⁴⁴.

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 h. Cells were harvested (4,000g for 5 min at room temperature), resuspended in PBS to $\text{OD}_{600} = 0.3$ and xylanase activity was assayed using the 3,5-dinitrosalicylic acid (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 h. Subsequently, the culture was divided into two 3 ml aliquots and harvested as described above. 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 (wash fractions) were collected. Cell pellets were 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 Table 13). Amplicons were cloned into the EcoRI and NcoI restriction sites of a pETM-11 (a kind gift from G. Stier, EMBL, Center for Biochemistry, Heidelberg, Germany)⁴⁵ or the XhoI and NcoI restriction site of a pET28a(+) (Novagen) using In-Fusion cloning (Takara) to express proteins as fusions with either cleavable N-terminal His₆ tags or C-terminal His₆ tags, respectively. Standard protocols were used for recombinant protein expression and purification using His-affinity and size-exclusion chromatography.

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. Hydrolysis kinetics of full-length or truncated xylanases (10–200 μM) were assayed towards 1–9 mg ml^{-1} of BGX, WAX or InWAX (900 μl at 37°C for 12 min). Initial hydrolysis rates were determined by removing 200 μl aliquots every 3 min and quenching the reaction in 300 μl DNS reagent⁴⁶. Next, samples were incubated for 15 min at 90°C followed by absorbance at 540 nm (A_{540}) measurement in 96-well microtitre plates. Xylose was used as a standard (0–2.5 mM). Xylanase activity was assayed for *R. intestinalis* cells washed with PBS \pm 1.5 M NaCl, and wash fractions, as above with the following modification: 180 μl of 1% (w/v) BGX was incubated with 20 μl cell suspension or wash fraction for 4 h.

Hydrolysis kinetics of α -glucuronidase were analysed on 1–9 mg ml^{-1} beechwood glucuronoxylan or a hydrolysate thereof (prepared by incubation with 4 mM RiXyn10A xylanase for 15 h at 37°C followed by heat inactivation). The initial rates of MeGlcA 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. Conversion of NAD⁺ to NADH was measured at A_{340} . Glucuronic acid was used as standard (0–500 μM).

Hydrolysis kinetics of RiXyl8 and RiXyl3A were determined towards X2 through to X6 (0.5–12 mM) in McIlvaine buffer pH 6.8 (10 mM citric acid and 20 mM sodium phosphate) as described previously^{17,48}. Reactions (350 μl) were incubated for 12 min at 37°C with 36–78 nM RiXyl3A or 2.4 nM RiXyn8. Aliquots of 50 μl were removed every 2 min 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 h before measuring A_{520} . The concentration of released pentoses was determined using a xylose standard (0–5 mM)⁴⁹.

α -L-Arabinofuranosidase activity for RiAbf43A 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 at 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 mutaoxase/ 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 RiAXE was determined in 250 μl reactions containing para-nitrophenyl-acetate (4 mM) and 0.14 μM enzyme. A_{405} was measured every 60 s for 10 min at 37°C in a microtitre plate reader and pNP (0–1 mM) was used as standard. The specific activity was determined in units (U) per mg, where a U is defined as the amount of enzyme that produces 1 μmol of pNP min^{-1} .

Kinetic parameters were calculated by fitting the Michaelis-Menten equation to the initial rate data using GraphPad Prism 7. The catalytic efficiency $k_{\text{cat}}/K_{\text{m}}$, 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-oligosaccharides was performed at 37°C for 15 h in the standard assay buffer used above. Oligosaccharide hydrolysates, used to assay the sequential action of the debranching xylanolytic enzymes, were generated using RiXyn10A, which was separated by ultrafiltration (3 kDa cut-off) before the addition of debranching enzymes. The hydrolysis profiles were analysed as detailed below. To verify the mode of reducing-end attack of RiXyl8, 30 mg XOS in standard assay buffer were reduced by NaBH₄ (1 M in 100 μM NaOH). A total of 200 μl of the NaBH₄ was added dropwise to 800 μl of the xylo-oligosaccharides solution, which was kept on ice. As a control, 100 μM NaOH was added to an 800 μl xylo-oligosaccharides solution. The mixture was incubated for 1 h at room temperature, then quenched by 400 μl 1 M acetic acid and diluted 10 \times in assay buffer.

Matrix-assisted laser desorption/ionization. Oligosaccharides were analysed with an Ultraflex matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF)/TOF instrument (Bruker Daltonics). 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 and HPAEC-PAD. Aliquots of 1 μl of enzymatic reactions were spotted on a silica gel 60 F254 plate (Merck). 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 at 350°C until bands appeared. Release of xylo-oligosaccharides and monosaccharides was analysed by HPAEC-PAD on an ICS-3000 (Dionex) using a 3 \times 250 mm CarboPac PA1 column, a 3 \times 50 mm guard column and 10 μl injections. Xylo-oligosaccharide and standards were eluted with mobile phase of constant 0.1 mM NaOH (flow rate: 0.35 ml min^{-1}) and a two-step linear gradient of sodium acetate: 0–25 min of 0–75 mM and 25–30 min of 75–400 mM. Monosaccharides and standards (0.1 mg ml^{-1}) of galactose, arabinose, glucose and xylose were eluted with 1 mM KOH for 35 min at 0.25 ml min^{-1} .

NMR spectroscopy. For the time-resolved NMR recordings: 4 mg AcBGX or acetylated spruce galactoglucomannan were dissolved in 500 μl 50 mM phosphate buffer pH 7.0 (99.9% D₂O). 2.5 μl of RiAXE to a final concentration of 64 nM was added. The recorded spectrum is a pseudo-two dimensional (2D)-type experiment recording a 1D proton NMR spectrum every 5 min with in total 220 time points. The 1D proton spectrum was recorded with 24 scans using a 30 $^{\circ}$ flip angle and a relaxation delay of 1 s (total recording time of 73 s). For enzyme treatment, 2.5 μl of RiXyn10A and RiAgu115A were added to the AcBGX sample to 167 nM and 13 nM, respectively, and the sample was incubated at 37°C for 24 h prior to RiAXE addition. All homonuclear and heteronuclear NMR experiments were recorded on a BRUKER AVIIIHD 800 MHz (Bruker BioSpin) equipped with 5 mm with cryogenic CP-TCI, and all acquisitions were done at 37°C . For the chemical shift assignment of AcBGX, the following spectra were recorded: 1D proton, 2D double-quantum-filtered correlation spectroscopy (DQF-COSY), 2D total correlation spectroscopy (TOCSY), 2D ¹³C-heteronuclear single-quantum coherence (HSQC), 2D ¹³C-heteronuclear 2 bond correlation (H2BC), 2D ¹³C-HSQC-[¹H,¹H]TOCSY and 2D heteronuclear multiple bond correlation (HMBC). The acetate signal to 1.903 p.p.m. (pH 7.0 at 37°C , in relation to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS)⁵⁰) was used as chemical shift reference for protons, whereas ¹³C chemical shifts were referenced indirectly to acetate, based on the absolute

frequency ratios⁵¹. The spectra were recorded, processed and analysed using TopSpin 3.5 software (Bruker BioSpin).

SPR. Xylo-oligosaccharide binding to *RiXyn10A*, *RiXyn10A*ΔCBMx and *RiXyn10A*-CBMx was analysed using 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 manufacturer's protocol with 50–150 μg ml⁻¹ protein in 10 mM sodium acetate pH 3.6–4.2, to a density of 1,362, 10,531 and 4,041 response units (RU) for *RiXyn10A*ΔCBMx, *RiXyn10A* and *RiXyn10A*-CBMx, respectively. The analysis comprised 90 s of association and 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 and 0.005% (v/v) P20 (GE Healthcare). 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 and Man6 and 62.5 μM–4 mM for X5. Data analysis was carried out using the Biacore T100 evaluation software, and K_d were determined by fitting a one-binding site model to the steady-state sensograms.

ITC. Titrations were performed using a Microcal ITC200 calorimeter (GE Healthcare) at 25 °C with *RiXBP* (0.1 mM) or *RiXyn10A*ΔCBMx (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 × 2 μl injections separated by 120 s. The data were corrected for the heat of dilution, determined from buffer titration and a non-linear single-binding model was fitted to the normalized integrated binding isotherms using the MicroCal Origin software v7.0 to determine the binding thermodynamics.

Affinity electrophoresis. Binding of CBMx to WAX (0–0.1% w/v) or BGX (0–1.0% w/v) was assessed by affinity electrophoresis⁵² in 10% native polyacrylamide gels (70 V for 3 h at 4 °C) using purified recombinant *RiXyn10A*-CBMx (3.0 μg) and β-lactoglobulin (1.5 μg) as a negative control. The relative mobility (r) was calculated as the migration of *RiXyn10A*-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 *RiXyn10A*, *RiXyn10B* and the transport protein *RiXBP* were raised in rats and rabbit, respectively (Eurogentec[®]). The specificity of the antibodies was tested by western blots using a standard protocol. The membranes were blocked for 1 h in 1% (w/v) BSA in TBST buffer (Tris-buffered saline, 0.1% (v/v) Tween 20) and incubated for 2 h with the anti-sera (500× dilution in TBST-buffer). Subsequently, the membranes were washed three times in TBST buffer and incubated for 2 h with 6,000× diluted secondary polyclonal goat anti-rabbit IgG-AP antibodies coupled to alkaline phosphatase (AP) (Dako) 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 OD₆₀₀ ≈ 0.8, harvested (4,000g for 5 min at 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 were blocked for 1 h 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 50× in blocking buffer for 2 h, washed twice in PBS and incubated for 1 h with 50 μl goat anti-rat IgG Alexa-Fluor 555 or goat anti-rabbit IgG Alexa-Fluor 488 (Thermo Scientific). Secondary antibodies were diluted 500× PBS. Finally, cells were washed twice in PBS, one drop of ProLong Gold antifade (Thermo Scientific) was applied and the cells were 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 × 100/1.3 NA, oil immersion objective.

Co-culture competition assay. *B. 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 OD₆₀₀) were inoculated into CFA medium (YCFA lacking the yeast extract to minimize *B. ovatus* growth on yeast extract⁵³) containing 0.5% (w/v) WAX, BGX, InWAX or X4. The co-cultures were grown in triplicates and samples (2 ml) were taken during growth. In the propagation experiment, the co-culture was passaged into fresh media after 9 h of growth (start OD₆₀₀ = 0.01), then grown for 12 h and passaged again into fresh media and grown for 12 h. 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 Table 14). 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⁻¹) 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 were estimated by comparing 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.

Life Sciences Reporting Summary. Further information on experimental design is available in the Life Sciences Reporting Summary.

Code availability. Adaptor trimming and demultiplexing was performed using custom python scripts (based on the Biopython SeqIO module⁴⁰) and the FASTX-Toolkit0.0.13.2 (http://hannonlab.cshl.edu/fastx_toolkit/). The fastx_barcode_splitter script (from FASTX-toolkit) was modified to demultiplex large fastq files provided from BGI. The modified version of this script is available upon request.

Data availability. The proteins characterized in this study are available from the NCBI with the following accession numbers: [EEV01588.1](https://.ncbi.nlm.nih.gov/nucl/EEV01588.1) (ROSINTL182_06494), [EEU99940.1](https://.ncbi.nlm.nih.gov/nucl/EEU99940.1) (ROSINTL182_08193), [EEU99941.1](https://.ncbi.nlm.nih.gov/nucl/EEU99941.1) (ROSINTL182_08194), [EEU99942.1](https://.ncbi.nlm.nih.gov/nucl/EEU99942.1) (ROSINTL182_08195), [EEU99943.1](https://.ncbi.nlm.nih.gov/nucl/EEU99943.1) (ROSINTL182_08196), [EEU99894.1](https://.ncbi.nlm.nih.gov/nucl/EEU99894.1) (ROSINTL182_08199) and [EEU99897.1](https://.ncbi.nlm.nih.gov/nucl/EEU99897.1) (ROSINTL182_08202). The authors declare that the data supporting the findings of this study are available within the paper and the supplementary information or from the corresponding author on request. The fastx_barcode_splitter script (from the FASTX-toolkit) was modified to demultiplex large fastq files provided from the BGI. The modified version of this script is available upon request.

Received: 25 September 2017; Accepted: 19 February 2018;

Published online: 02 April 2018

References

- Nicholson, J. K. et al. Host-gut microbiota metabolic interactions. *Science* **336**, 1262–1267 (2012).
- Sonnenburg, J. L. & Bäckhed, F. Diet-microbiota interactions as moderators of human metabolism. *Nature* **535**, 56–64 (2016).
- Marchesi, J. R. et al. The gut microbiota and host health: a new clinical frontier. *Gut* **65**, 330–339 (2016).
- David, L. A. et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559–563 (2013).
- Desai, M. S. et al. A dietary fiber-deprived gut microbiota degrades the colonic mucus barrier and enhances pathogen susceptibility. *Cell* **167**, 1339–1353 (2016).
- 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).
- Xu, S. et al. Butyrate induces apoptosis by activating PDC and inhibiting complex I through SIRT3 inactivation. *Signal Transduct. Target. Ther.* **2**, e16035 (2017).
- 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).
- Furusawa, Y. et al. Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* **504**, 446–450 (2013).
- 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).
- 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).
- Qin, J. et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55–60 (2012).
- Vrieze, A. et al. Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology* **143**, 913–916 (2012).
- 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).
- 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).
- Van den Abbeele, P. et al. Butyrate-producing *Clostridium* cluster XIVa species specifically colonize mucins in an in vitro gut model. *ISME J.* **7**, 949–961 (2013).
- El Kaoutari, A., Armougom, F., Gordon, J. I., Raoult, D. & Henricsson, B. The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat. Rev. Microbiol.* **11**, 497–504 (2013).
- 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).

19. 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).
20. Sheridan, P. O. et al. Polysaccharide utilization loci and nutritional specialization in a dominant group of butyrate-producing human colonic Firmicutes. *Microb. Genomics* **2**, e000043 (2016).
21. Selvendran, R. R. Chemistry of plant cell walls and dietary fibre. *Scand. J. Gastroenterol.* **5521**, 33–41 (1987).
22. Rogowski, A. et al. Glycan complexity dictates microbial resource allocation in the large intestine. *Nat. Commun.* **6**, 7481 (2015).
23. 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. USA* **111**, E3708–E3717 (2014).
24. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490–D495 (2014).
25. Kelly, G. et al. Structure of the cell-adhesion fragment of intimin from enteropathogenic *Escherichia coli*. *Nat. Struct. Mol. Biol.* **6**, 313–318 (1999).
26. Ebbes, M. et al. Fold and function of the InlB B-repeat. *J. Biol. Chem.* **286**, 15496–15506 (2011).
27. 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).
28. 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).
29. Ejby, M. et al. Structural basis for arabinoxylo-oligosaccharide capture by the probiotic *Bifidobacterium animalis* subsp. lactis BI-04. *Mol. Microbiol.* **90**, 1100–1112 (2013).
30. 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).
31. Sun, Z. et al. A novel three-component system-based regulatory model for D-xylose sensing and transport in *Clostridium beijerinckii*. *Mol. Microbiol.* **95**, 576–589 (2015).
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 BI-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 amyloosomes in the resistant starch-utilizing human colonic firmicutes bacterium *Ruminococcus bromii*. *mBio* **6**, e01058-15 (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.* **19**, 3251–3267 (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. USA* **106**, 5859–5864 (2009).
38. Biely, P. et al. Mode of action of acetylxyylan esterases on acetyl glucuronoxylan and acetylated oligosaccharides generated by a GH10 endoxylanase. *Biochim. Biophys. Acta* **1830**, 5075–5086 (2013).
39. Sorbotten, 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**, 357–359 (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. Dilokpimol, A. et al. Enzymatic synthesis of β -xylosyl-oligosaccharides by transxylosylation using two β -xylosidases of glycoside hydrolase family 3 from *Aspergillus nidulans* FGSC A4. *Carbohydr. Res.* **346**, 421–429 (2011).
50. Govind, V., Young, K. & Maudsley, A. A. Proton NMR chemical shifts and coupling constants for brain metabolites. *NMR Biomed.* **13**, 129–153 (2000).
51. Zhang, H., Neal, S. & Wishart, D. S. RefDB: a database of uniformly referenced protein chemical shifts. *J. Biomol. NMR* **25**, 173–195 (2003).
52. Takeo, K. Affinity electrophoresis: principles and applications. *Electrophoresis* **5**, 187–195 (1984).
53. 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).
54. Anand, S., Kaur, H. & Mande, S. S. Comparative in silico analysis of butyrate production pathways in gut commensals and pathogens. *Front. Microbiol.* **7**, 1945 (2016).

Acknowledgements

We thank B. Henrissat, architecture et fonction des macromolécules biologiques, CNRS, Aix-Marseille University and the curator of CAZy, for his advice and discussions on the assignment of the novel CBMx and the esterase. We also thank M. Yadav, US Department of Agriculture, Agricultural Research Service, for the kind gift of cornbran xylan, and BioCHOS AS for providing the chitoooligo (CHOS) sample. M. Due, T. Holm Madsen and C. Aarup Christensen are thanked for their technical help in cloning recombinant proteins and the performance of binding experiments. We also wish to thank A. Schultz, H. Juel Martens and M. Hansen, PLEN, University of Copenhagen, for the use of the 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 (DFP, FNU) by a Research Project 2 grant (grant ID: 4002-00297B), a BIONÆR project (grant number: 244259) and the Norwegian NMR Platform, NNP (F.L.A.) from the Research Council of Norway (226244). Carlsberg Foundation is acknowledged for an ITC instrument grant (2011-01-0598).

Author contributions

Growth analysis was performed by M.L.L. Transcriptomic analysis was done by M.L.L., C.W. and D.A.E. Enzyme characterization was performed by M.L.L., M.E., S.S.P., F.L.A. and B.W. qPCR was done by M.L.L. and M.I.B. Microscopy was performed by M.L.L. and C.S. Experiments were designed by M.L.L. and M.A.H. The manuscript written by M.L.L. and M.A.H. with contributions from T.R.L., B.W. and F.L.A. Figures were prepared by M.L.L.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41564-018-0132-8>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to M.A.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

▶ Experimental design

1. Sample size

Describe how sample size was determined.

No sample-size calculations have been performed in this study. Biological triplicates were performed in all experiments involving bacteria (growth, transcriptomics, qPCR) according to standard scientific methods.

2. Data exclusions

Describe any data exclusions.

No data were excluded from the data analysis

3. Replication

Describe whether the experimental findings were reliably reproduced.

All attempts at replication were successful.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

No experimental groups were allocated in this study.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Blinding was not relevant in this study, which does not include animals or human participants.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- The test results (e.g. P values) given as exact values whenever possible and with confidence intervals noted
- A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

► Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

All software used are commercially available and details of which software version is stated in the method section.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

► Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

All acetylated polysaccharides are available from the authors on request. Otherwise no unique materials were used.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

Polyclonal goat anti-rabbit IgG-AP (Dako) LOT: 20024392, Polyclonal anti-rat IgGAP (Sigma aldrich) LOT 035M4827V. Alexa Flour 488 goat anti rabbit IgG(thermo) LOT 1829924. Alexa Flour 555 goat anti rat (Thermo) LOT 1783431. Three custom made antibodies were made against recombinant proteins in rat and rabbit by Eurogentec. See details in method section.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eukaryotic cell lines were used.

b. Describe the method of cell line authentication used.

Not applicable

c. Report whether the cell lines were tested for mycoplasma contamination.

Not applicable

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

Not relevant

► Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

No animals were used.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

The study did not involve human research participants.