





Synergy between Cell Surface Glycosidases and Glycan-Binding Proteins Dictates the Utilization of Specific Beta(1,3)-Glucans by Human Gut *Bacteroides*

Guillaume Déjean,^a Kazune Tamura,^{a,b} Adriana Cabrera,^a Namrata Jain,^{a,c} Nicholas A. Pudlo,^d Gabriel Pereira,^d Alexander Holm Viborg,^a  Filip Van Petegem,^b Eric C. Martens,^d  Harry Brumer^{a,b,c,e}

^aMichael Smith Laboratories, University of British Columbia, Vancouver, British Columbia, Canada

^bDepartment of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia, Canada

^cDepartment of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada

^dDepartment of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, Michigan, USA

^eDepartment of Botany, University of British Columbia, Vancouver, British Columbia, Canada

Guillaume Déjean and Kazune Tamura contributed equally. The order of the authors has been determined according to the details in the Acknowledgments. Regarding the two first authors, briefly, G.D. spearheaded all microbiology, biochemistry, and enzymology; K.T. led all structural biology and assisted with enzymology.

ABSTRACT The human gut microbiota (HGM) has far-reaching impacts on human health and nutrition, which are fueled primarily by the metabolism of otherwise indigestible complex carbohydrates commonly known as dietary fiber. However, the molecular basis of the ability of individual taxa of the HGM to address specific dietary glycan structures remains largely unclear. In particular, the utilization of $\beta(1,3)$ -glucans, which are widespread in the human diet as yeast, seaweed, and plant cell walls, had not previously been resolved. Through a systems-based approach, here we show that the symbiont *Bacteroides uniformis* deploys a single, exemplar polysaccharide utilization locus (PUL) to access yeast $\beta(1,3)$ -glucan, brown seaweed $\beta(1,3)$ -glucan (laminarin), and cereal mixed-linkage $\beta(1,3)/\beta(1,4)$ -glucan. Combined biochemical, enzymatic, and structural analysis of PUL-encoded glycoside hydrolases (GHs) and surface glycan-binding proteins (SGBPs) illuminates a concerted molecular system by which *B. uniformis* recognizes and saccharifies these distinct β -glucans. Strikingly, the functional characterization of homologous $\beta(1,3)$ -glucan utilization loci (1,3GUL) in other *Bacteroides* further demonstrated that the ability of individual taxa to utilize $\beta(1,3)$ -glucan variants and/or $\beta(1,3)/\beta(1,4)$ -glucans arises combinatorially from the individual specificities of SGBPs and GHs at the cell surface, which feed corresponding signals to periplasmic hybrid two-component sensors (HTCSs) via TonB-dependent transporters (TBDTs). These data reveal the importance of cooperativity in the adaptive evolution of GH and SGBP cohorts to address individual polysaccharide structures. We anticipate that this fine-grained knowledge of PUL function will inform metabolic network analysis and proactive manipulation of the HGM. Indeed, a survey of 2,441 public human metagenomes revealed the international, yet individual-specific, distribution of each 1,3GUL.

IMPORTANCE *Bacteroidetes* are a dominant phylum of the human gut microbiota (HGM) that target otherwise indigestible dietary fiber with an arsenal of polysaccharide utilization loci (PULs), each of which is dedicated to the utilization of a specific complex carbohydrate. Here, we provide novel insight into this paradigm through functional characterization of homologous PULs from three autochthonous *Bacteroides* species, which target the family of dietary $\beta(1,3)$ -glucans. Through detailed biochemical and protein structural analysis, we observed an unexpected diversity in the substrate specificity of PUL glycosidases and glycan-binding proteins with regard to $\beta(1,3)$ -glucan linkage and branching patterns. In combination, these in-

Citation Déjean G, Tamura K, Cabrera A, Jain N, Pudlo NA, Pereira G, Viborg AH, Van Petegem F, Martens EC, Brumer H. 2020. Synergy between cell surface glycosidases and glycan-binding proteins dictates the utilization of specific beta(1,3)-glucans by human gut *Bacteroides*. mBio 11:e00095-20. <https://doi.org/10.1128/mBio.00095-20>.

Editor Laurie E. Comstock, Brigham and Women's Hospital/Harvard Medical School

Copyright © 2020 Déjean et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Harry Brumer, brumer@mssl.ubc.ca.

Received 15 January 2020

Accepted 9 March 2020

Published 7 April 2020

dividual enzyme and protein specificities support taxon-specific growth on individual $\beta(1,3)$ -glucans. This detailed metabolic insight, together with a comprehensive survey of individual 1,3GULs across human populations, further expands the fundamental roadmap of the HGM, with potential application to the future development of microbial intervention therapies.

KEYWORDS *Bacteroides*, *Bacteroidetes*, dietary fiber, glycan-binding protein, glycoside hydrolase, polysaccharide utilization locus, polysaccharides

The human gut microbiota (HGM) is a complex community that underpins our nutrition and overall well-being (1, 2) yet is also associated with some diseases (3–6), depending on its particular composition and physiology. A key challenge in manipulating HGM dynamics toward healthful outcomes is a limited understanding of the ecological forces that shape this community within individuals (7, 8). The catabolism of complex dietary carbohydrates is a key driver of HGM structure and metabolic function (9, 10). Thus, resolving a detailed roadmap of the glycan utilization mechanisms deployed by individual members of the HGM is central to the development of dietary and microbial interventions to promote human health.

An explosion of (meta)genome sequence data continues to reveal substantial taxon-level variation in the metabolic potential of the HGM, yet a lack of functional data restricts our ability to fully explain or predict these differences and eventually to use this knowledge to engineer changes to the HGM. *Bacteroides* species, in particular, are predominant autochthonous members of the HGM that metabolize a wide variety of complex glycans into short-chain fatty acids (SCFAs) (11), which is enabled by a plethora of PULs in their genomes (12). As exemplified by the $\beta(1,3)$ -glucan utilization loci (1,3GULs) elucidated here, PULs encode concerted molecular systems of surface glycan binding proteins (SGBPs), carbohydrate-active enzymes (CAZymes), TonB-dependent transporters (TBDTs), and sensor/regulators to recognize, capture, import, and saccharify individual substrates (13) (Fig. 1). Recently, a number of seminal, integrated PUL studies combining genetics, biochemistry, and structural biology have highlighted how strain-level genomic variation dictates nutrient specificities (14–27).

Amorphous β -glucans are ubiquitous polysaccharides in the human diet, which can be delineated broadly by backbone linkage (Fig. 2A): mixed-linkage $\beta(1,3)/\beta(1,4)$ -glucans (MLGs) are commonly found in cereal crops such as oats and barley, while the $\beta(1,3)$ -glucan callose is found as a component of plant cell walls (28). Edible fungi, including yeasts, contain $\beta(1,6)$ -glucans (22) and $\beta(1,3)$ -glucans (28). $\beta(1,3)$ -Glucans also occur in the cell walls of seaweeds (29). Many $\beta(1,3)$ -glucans also contain $\beta(1,6)$ -linked branches, the length and frequency of which vary according to source (e.g., yeast β -glucan and laminarin) (28). It is known generally that a range of taxa in the HGM, including from the phyla *Firmicutes*, *Actinobacteria*, and *Bacteroidetes*, metabolize β -glucans to produce short-chain fatty acids (19, 22, 30–32). In addition to this nutritional benefit (6), β -glucans have been associated with health-promoting effects against cancer, diabetes/metabolic syndrome, and inflammation (33–36). However, the molecular mechanisms underpinning these effects have not been fully elucidated.

Previous studies on MLG and $\beta(1,6)$ -glucan utilization by the symbiotic *Bacteroides* revealed the molecular details by which the different, dedicated PUL-encoded machineries target these two distinct classes of β -glucans (19, 22, 26). Here, we functionally dissected an exemplar 1,3GUL from *Bacteroides uniformis* ATCC 8492 to provide molecular insight into $\beta(1,3)$ -glucan utilization, thereby resolving a key outstanding deficit in our understanding of β -glucan metabolism by the HGM. Notably, this included solving the first tertiary structure and resolving the catalytic mechanism of a member of the new glycoside hydrolase family 158 (GH158). Building upon these foundational results, we subsequently demonstrated that the individual abilities of three *Bacteroides* species to metabolize distinct $\beta(1,3)$ -glucans and/or MLG is dictated by the cumulative specificities and contributions of their respective SGBPs, cell surface GHs, and other sensor/transport functions encoded by partially homologous 1,3GULs. Finally, we

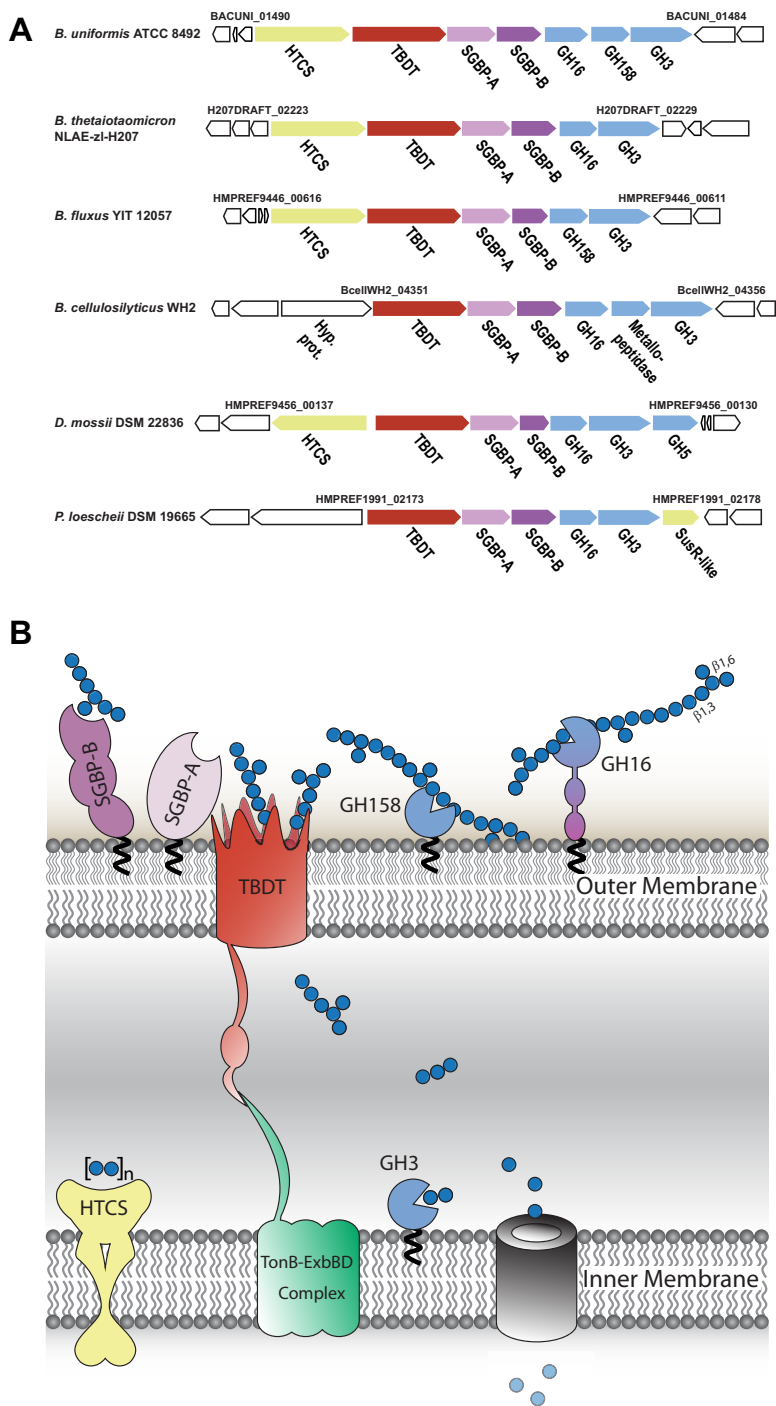


FIG 1 $\beta(1,3)$ -Glucan utilization systems in the order *Bacteroidales*. (A) Strain-specific, syntenic $\beta(1,3)$ -glucan utilization loci (1,3GUL) from *Bacteroides uniformis*, *B. thetaiotaomicron*, *B. fluxus*, *B. cellulosilyticus*, *Dysgonomonas mossii*, and *Prevotella loescheii*. Genome locus tags constituting 1,3GUL boundaries are indicated. Predicted or confirmed (this work) functional annotations are denoted below each gene. HTCS, hybrid two-component system sensor/regulator; SusR-like sensor-regulator; TBDT, SusC-like TonB-dependent transporter; SGBP-A, SusD-like cell-surface glycan-binding protein; SGBP-B, sequence-divergent cell-surface glycan-binding protein; GH n , member of glycoside hydrolase family n . (B) Model of *B. uniformis* laminarin utilization based on the present study and by analogy with the archetypal starch utilization system (Sus) (10, 13). Gene products are colored analogously to panel A, and predicted N-terminal lipidation following signal peptidase II cleavage is indicated with a black squiggle.

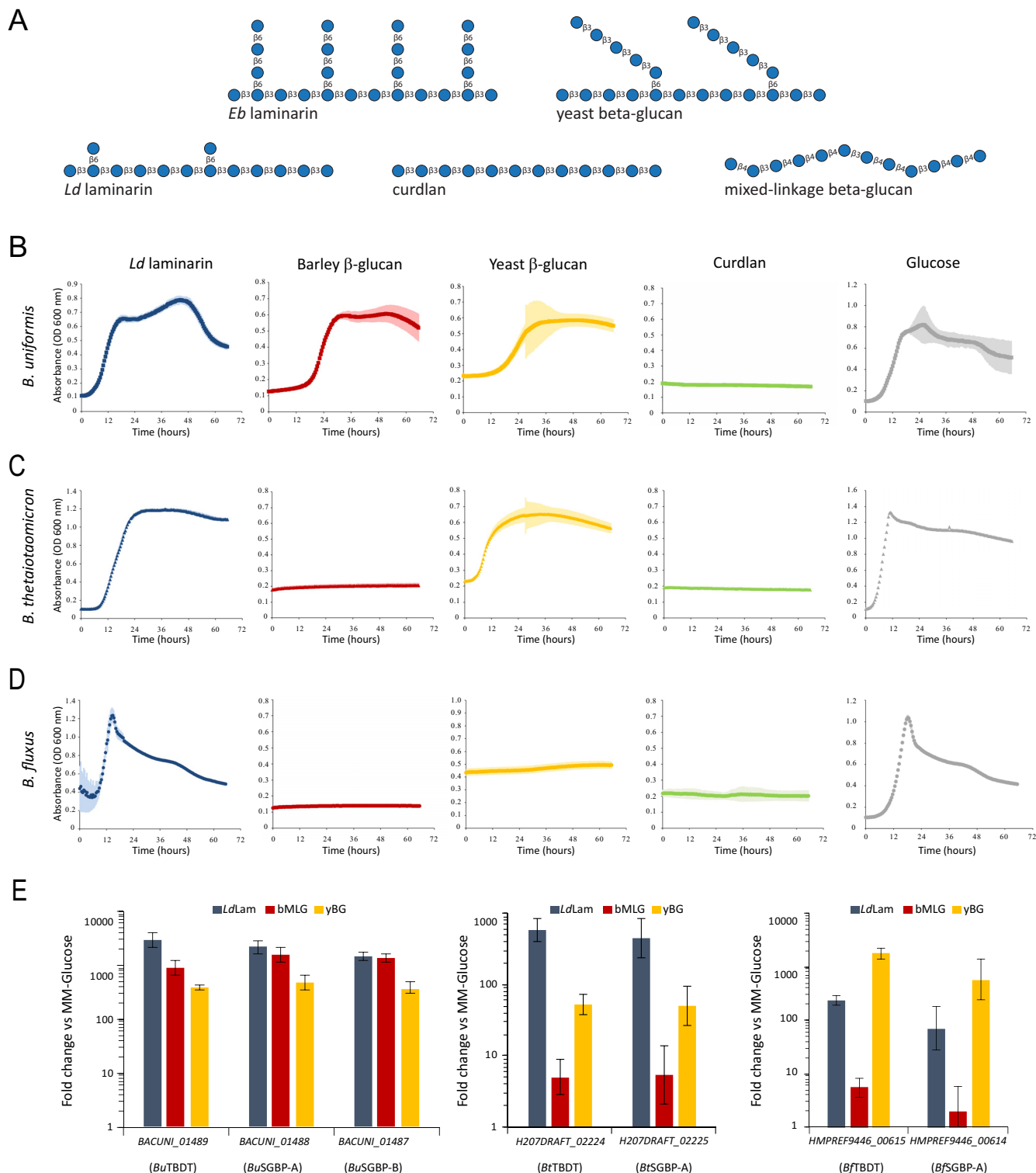


FIG 2 $\beta(1,3)$ -Glucan utilization by *Bacteroides*. (A) Structures of β -glucans used in this study. Bacterial curdlan is a representative unbranched $\beta(1,3)$ -glucan, analogous to plant callose. Algal laminarins are $\beta(1,3)$ -glucans with $\beta(1,6)$ -linked branches. *Laminaria digitata* laminarin has infrequent, single $\beta(1,6)$ -glucosyl branches [$\beta(1,3)$ to $\beta(1,6)$ molar ratio, 7:1], while *Eisenia bicyclis* laminarin has a high frequency of branches with degrees of polymerization of up to three $\beta(1,6)$ -glucosyl residues [$\beta(1,3)$ to $\beta(1,6)$ molar ratio, 3:2]. Yeast $\beta(1,3)$ -glucan contains longer $\beta(1,3)$ -glucan branches extended from $\beta(1,6)$ -linked branch points. Cereal mixed-linkage $\beta(1,3)/\beta(1,4)$ -glucans (MLGs) are linear chains of $\beta(1,4)$ -linked cellotriosyl and cellotetraosyl units linked by $\beta(1,3)$ bonds. (B to D) Growth curves of *B. uniformis* ATCC 8492 (B), *B. thetaiotaomicron* NLAE-zl-H207 (C), and *B. fluxus* YIT 12057 (D) in minimal medium containing the indicated carbon source at 0.5%, wt/vol. Points represent averages from $n = 6$ technical replicates (microplate wells) for polysaccharides and $n = 3$ for glucose; error bars represent the standard errors of the means. (E) RNA abundance for core PUL genes of *Bacteroides* spp. quantified by qRT-PCR. Bacteria were grown to mid-log phase in minimal medium containing glucose as the sole carbon source and subsequently exposed to different β -glucans ($n = 3$, expression measurements from individual cultures, relative to glucose control; error bars represent the standard errors of the means).

found, through metagenomic analysis, that the prevalence of 1,3GULs in human gut microbiomes is species-dependent but broadly distributed worldwide.

RESULTS

***B. uniformis* possesses a distinct PUL upregulated during growth on $\beta(1,3)$ -glucans.** *B. uniformis* ATCC 8492 exhibited robust growth in minimal medium containing either glucose or the branched $\beta(1,3)$ -glucans from *Laminaria digitata* [*L. digitata* laminarin, or *LdLam*, which contains single $\beta(1,6)$ -linked glucosyl branches] and *Saccharomyces cerevisiae* [yeast β -glucan, *yBG*, which contains longer $\beta(1,6)$ -linked glucosyl branches extended by $\beta(1,3)$ -linked glucose units]. No growth was observed on *Alcaligenes faecalis* curdlan [an unbranched $\beta(1,3)$ -glucan analogous to plant callose], which is poorly soluble and forms a gel in water (Fig. 2B and Table S1 at <https://doi.org/10.14288/1.0388792>). Interestingly, *B. uniformis* also grew well on barley MLG [*bMLG*, which has an ~2.5:1 ratio of $\beta(1,4)$ to $\beta(1,3)$ backbone linkages] (Fig. 2B and Table S1 at <https://doi.org/10.14288/1.0388792>) despite lacking a canonical mixed-linkage glucan utilization locus (MLGUL) homologous to the *Bacteroides ovatus* MLGUL (19).

We identified in *B. uniformis* a putative $\beta(1,3)$ -glucan utilization locus that encodes a TonB-dependent transporter (TBDT; a *SusC* homolog, BACUNI_01489) and a cell surface glycan-binding protein (SGBP-A; a *SusD* homolog, BACUNI_01488) as canonical PUL signatures (37), as well as an additional nonhomologous SGBP (*BuSGBP-B*; BACUNI_01487), three glycoside hydrolases (GH16 subfamily 3, GH158, and GH3; BACUNI_01486 to _01484), and a hybrid two-component system (HTCS) transcriptional regulator (BACUNI_01490) (Fig. 1). In particular, the GH complement was suggestive of a role in $\beta(1,3)$ -glucan hydrolysis: GH16 subfamily 3 (GH16_3 [38]) contains known *endo*-laminarinases (among other *endo*- β -glucanases) (19) and GH3 contains *exo*- β -glucosidases (among others) (39). Notably, during the course of this study, GH158 emerged as a new family whose founding member was shown to hydrolyze a chemically derivatized $\beta(1,3)$ -glucan in a high-throughput screen (40). Concordant with this proposed PUL specificity, when we probed the expression of core genes encoding the TBDT and both SGBPs (BACUNI_01487-01489), we found that they were strongly upregulated in the presence of laminarin, *bMLG*, and *yBG* as sole carbon sources versus a glucose control (Fig. 2E). The syntenic TBDT and SGBP-A (*SusC/SusD* homologs) in a partially homologous *Bacteroides cellulosilyticus* WH2 PUL (Fig. 1) were shown previously to be similarly upregulated in the presence of laminarin and *bMLG* (41). Notably, *B. ovatus* ATCC 8483, which possesses an MLGUL (19) but not a homologous 1,3GUL (Fig. 1), does not grow on laminarin (37).

Biochemical basis of $\beta(1,3)$ -glucan recognition and degradation by *B. uniformis*. By analogy with other PUL-encoded systems, we propose a working model of concerted $\beta(1,3)$ -glucan saccharification and uptake by the proteins of the 1,3GUL (Fig. 1). This model involves polysaccharide capture at the cell surface by at least one SGBP, backbone hydrolysis by at least one *endo*-glucanase, transport of oligosaccharide fragments through the outer membrane by the TBDT, and ultimate saccharification by an *exo*- β -glucosidase in the periplasm. Indeed, signal peptide analysis with SignalP, LipoP, and PSORTb predicted that *BuGH16* is localized at the cell surface via N-terminal Cys lipidation (PSORTb score, 9.7), whereas *BuGH3*, which also harbors a type II signal peptide, is predicted to be periplasmic (PSORTb score, 9.4). Interestingly, PSORTb was unable to predict the localization of *BuSGBP-A*, *BuSGBP-B*, and *BuGH158* (score, <2.5) despite the presence of a type II signal peptide, a +2 serine residue in each, and expected extracellular localization of at least the SGBPs by analogy with the archetypal starch utilization system (9, 10).

(i) *BuGH16* is a broad-specificity *endo*- $\beta(1,3)/\beta(1,4)$ -glucanase. To investigate the catalytic role of *BuGH16* in surface polysaccharide breakdown, the hydrolytic activity of the recombinant protein was screened against a library of polysaccharides (Table S2 at <https://doi.org/10.14288/1.0388792>). *BuGH16* displayed activity toward the $\beta(1,3)$ -glucans *LdLam*, *Eisenia bicyclis* laminarin [*EbLam*, which contains more frequent $\beta(1,6)$ -linked glucosyl branches of up to three $\beta(1,6)$ -linked glucosyl residues (42);

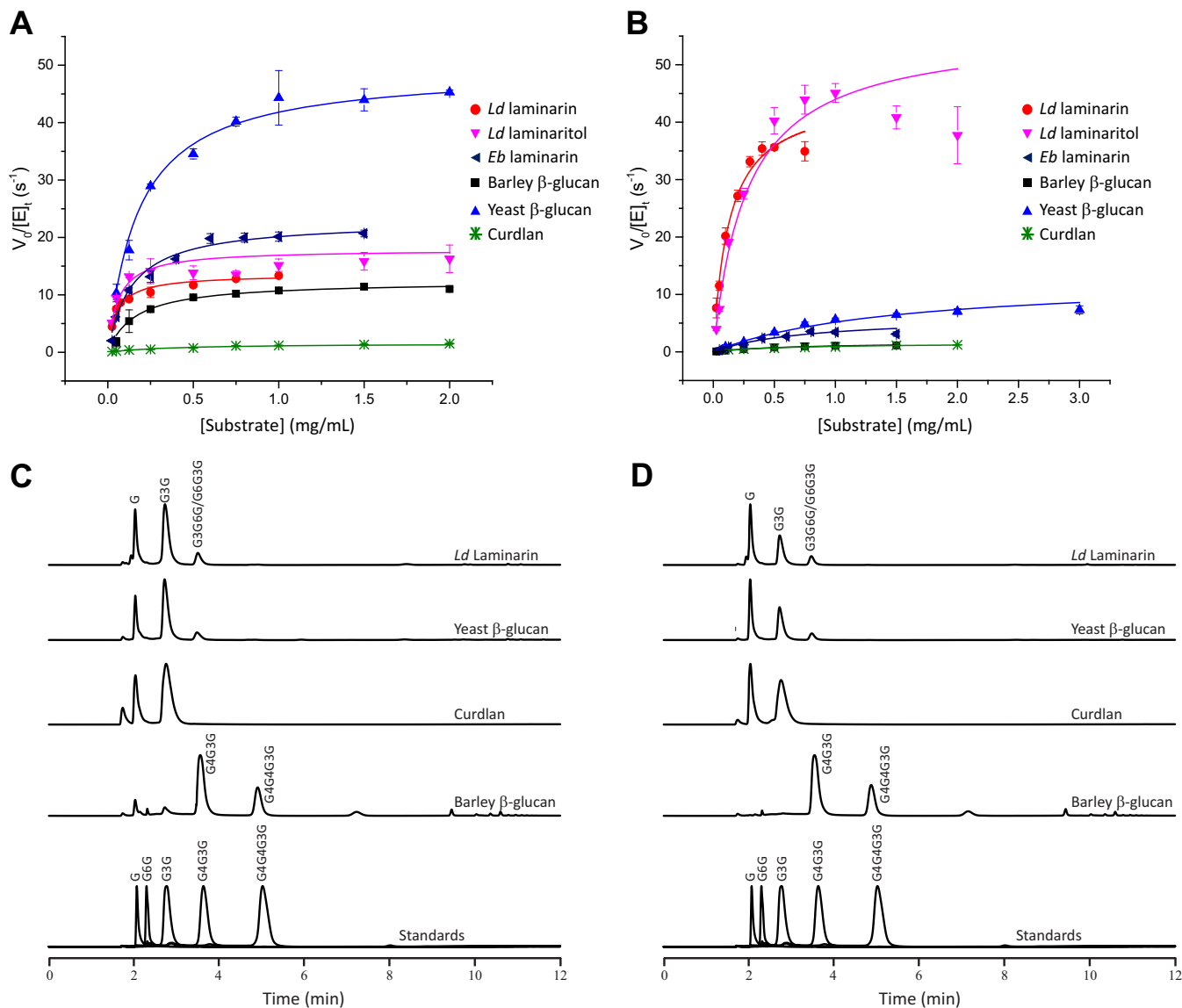


FIG 3 Hydrolysis of β-glucans by *BuGH16* and *BuGH158*. (A and B) Initial-rate kinetics analysis of *BuGH16* (A) and *BuGH158* (B). $V_o/[E]_t$ = initial rate per total enzyme concentration. Curves represent fits of the Michaelis-Menten equation to the average data points ($n = 3$); error bars represent standard deviations from the means. (C and D) HPAEC-PAD analysis of limit-digest products of *BuGH16* (C) and *BuGH158* (D) at 37°C.

Fig. 2A], yBG, and curdlan. *BuGH16* was also active on bMLG. Using laminarin as the substrate, the pH optimum was 6.0 (consistent with extracellular function in the human large intestine [43]), and the maximum temperature of activity was ca. 50°C (Fig. S1 at the URL mentioned above).

Subsequent Michaelis-Menten kinetics at the optimum pH and 37°C confirmed that *BuGH16* is a broadly specific β(1,3)-glucanase with similar catalytic efficiencies toward laminarins and yBG (Fig. 3A and Table S2 at the URL mentioned above). The k_{cat}/K_m value of *BuGH16* on bMLG was 4-fold lower than those for these all-β(1,3)-backbone glucans (Fig. 3A and Table S2 at the URL mentioned above). The enzyme was very poorly active on the unbranched β(1,3)-glucan, curdlan (k_{cat}/K_m value ca. 1% of that on the laminarins) (Fig. 3A and Table S2 at the URL mentioned above). *LdLam* and yBG ultimately were hydrolyzed to glucose, laminaribiose, and a trisaccharide (G3G6G or G6G3G), as determined by high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) (Fig. 3C) and confirmed by matrix-assisted laser desorption ionization mass spectrometry of the per-*O*-acetylated crude

mixture (data not shown). The limit-digest products were glucose and laminaribiose from curdlan, and G4G3G and G4G4G3G from bMLG (Fig. 3C). Analyses of the hydrolysis products over time suggest that *BuGH16* operates by an *endo*-dissociative mode of action (Fig. S2 at the URL mentioned above).

Despite extensive attempts using the full-length protein and the GH16 module only (Fig. S3 at the URL mentioned above), we were not able to obtain a crystal structure of *BuGH16* to explain the observed substrate promiscuity. However, protein phylogeny unambiguously places *BuGH16* in the “ β -bulge laminarinase/MLGase” clade (Fig. S4 at <https://doi.org/10.14288/1.0388792>), corresponding to GH16 subfamily 3 (38), which is known to contain laminarinases with secondary MLGase activity, as well as bona fide MLGases (19). These enzymes require a $\beta(1,3)$ linkage between the -1 and -2 subsites (44), but beyond the -2 subsite, requirements are less stringent, leading to widespread promiscuity in polysaccharide and bond cleavage specificity [$\beta(1,3)$ versus $\beta(1,4)$] (19). Notably, structural homology modeling and superposition with an exemplar laminarinase from the marine bacterium *Zobellia galactanivorans* DsijT (45) rationalizes the ability of *BuGH16* to accommodate highly branched $\beta(1,3)$ -glucans, i.e., laminarins and yBG (Fig. S5 at the URL mentioned above).

In addition to the catalytic module, *BuGH16* comprises a pair of PFAM 13004 domains (Fig. S3 at the URL mentioned above). Despite initial bioinformatics predictions, carbohydrate binding has not been demonstrated for any PFAM 13004 domain to date (15). Likewise, affinity gel electrophoresis (AGE) analysis of a recombinant protein consisting of the two PFAM 13004 domains of *BuGH16* revealed that they do not bind cognate polysaccharides of the 1,3GUL (Fig. S6 at the URL mentioned above). These domains likely serve a spacer function, as in *B. ovatus* GH5 (*BoGH5*) (15), and are analogous to all β -sheet domains in SGBPs (16, 26, 46).

(ii) *BuGH158* is a strictly specific, retaining *endo*- $\beta(1,3)$ glucanase with a triose phosphate isomerase (TIM) barrel fold. *BuGH158* is a member of a newly established GH family, the distantly related founding member of which, Vvad_PD1638 (sequence identity 29%) (Fig. S7 at <https://doi.org/10.14288/1.0388792>), was shown to be active on the artificial proxy substrate, carboxymethyl-curdlan, in a high-throughput screen (40). Hence, we performed detailed kinetic and product analysis to more precisely delineate the specificity of *BuGH158* in the context of the 1,3GUL (Fig. 3B and Table S2 at the URL mentioned above). In contrast to *BuGH16*, *BuGH158* is highly specific for *LdLam*, with a k_{cat}/K_m value ca. 2 orders of magnitude higher than those for *EbLam*, yBG, MLG, and curdlan (Fig. 3B; also see and Table S2 and Fig. S1 and S2 at the URL mentioned above). The corresponding hydrolysis products were identical to those of *BuGH16* (Fig. 3D).

To provide the first three-dimensional insights into substrate specificity and catalysis in GH158, we determined the tertiary structure of the enzyme to 1.8 Å by X-ray crystallography (Table S3 at the URL mentioned above). *BuGH158* consists of an N-terminal (α/β)₈ TIM barrel domain and a C-terminal, eight-stranded immunoglobulin (Ig)-like domain in contact with helices $\alpha 7$ and $\alpha 8$ of the TIM barrel. Additional loops contribute to the extensive contact between domains, with one from the Ig-like domain extending above the TIM barrel to shape the active-site cleft (Fig. 4A). A Dali structure similarity search returned members of clan GH-A as the top 20 structural homologs (Table S4 at the URL mentioned above), confirming the bioinformatics prediction that GH158 constitutes an evolutionarily distinct family within this clan (40). Correspondingly, superposition with the top two GH2 and GH5 results show that the TIM barrel core is well conserved and that family-level differences arise from the presence or absence of accessory domains (Fig. 4B and C).

All clan GH-A enzymes are predicted to use a canonical Koshland double-displacement mechanism employing a covalent glycosyl enzyme intermediate and resulting in overall retention of the anomeric stereochemistry at the site of polysaccharide backbone hydrolysis (39). This was confirmed for *BuGH158* by nuclear magnetic resonance (NMR) analysis of the hydrolysis of 2-chloro-4-nitrophenyl laminaribioside (G3G-CNP), which constitutes the first stereochemical determination for the family (Fig.

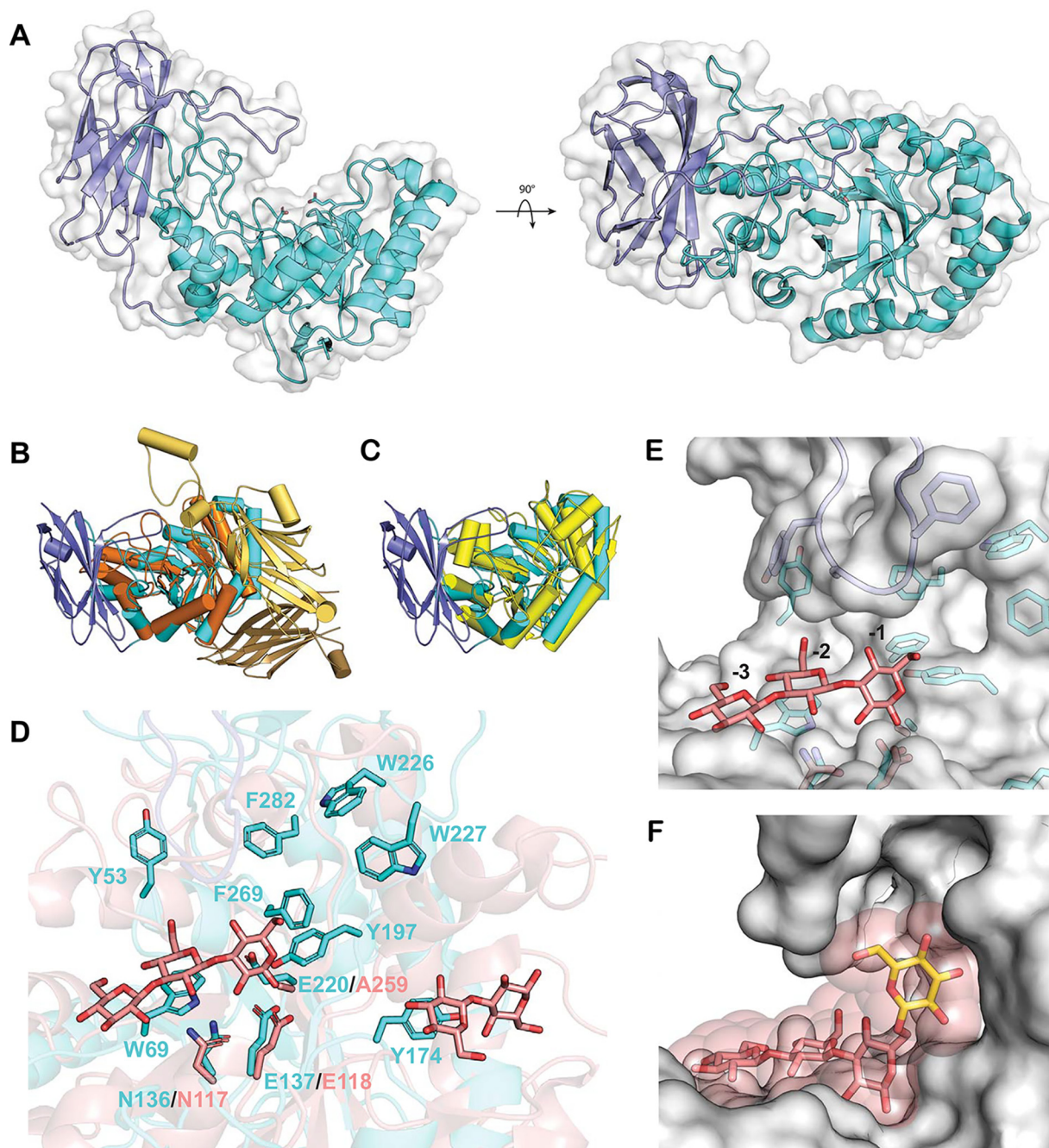


FIG 4 *BuGH158* tertiary structure. (A) Overall structure of *BuGH158*, with the TIM barrel domain colored cyan, the Ig-like domain colored slate, and a semitransparent surface in white. The catalytic residues E137 and E220 are shown as sticks. (B) Superposition of *BuGH158* (colored as described for panel A) with the *Bifidobacterium dentium* GH2 β -glucuronidase, which has two additional domains N terminal to the TIM barrel (PDB entry 5Z1B, orange). (C) Superposition of *BuGH158* (colored as described for panel A) and the *Chrysonilia sitophila* GH5 β (1,4)-mannanase (PDB entry 4AWE, yellow), which has no additional domains. (D) Superposition of *BuGH158* (colored as described for panel A) with a laminaritriose complex of a GH17 *endo*- β (1,3)-glucanase from *Solanum tuberosum* (StGH17; PDB entry 4GZJ, rose). *BuGH158* aromatic residues in the active-site cleft are shown in cyan sticks, and homologous catalytic residues in StGH17 are shown in rose sticks. (E) Surface representation of *BuGH158* superposed with the StGH17 laminaritriose ligand (rose sticks) occupying the negative subsites as labeled, revealing a pocket near 6-OH of the glucose in the -1 subsite. (F) A glucose molecule, shown as yellow sticks, attached to the subsite -1 glucose via a β (1,6)-linkage. The transparent sphere about each atom of the modeled glucose and laminaritriose represent a van der Waals radius of 1.5 Å.

S8 at the URL mentioned above). Clan GH-A members present the conserved catalytic acid/base and nucleophile residues on loops immediately following strands β 4 and β 7, corresponding to E137 and E220, respectively, in *BuGH158* (Fig. 4D). Additionally, a conserved asparagine (N136) precedes the general acid/base (E137) and is anticipated

to engage in hydrogen-bonding interactions with the substrate (Fig. 4D) (39). Indeed, site-directed mutation of N136, E137, or E220 to alanine completely abolished catalytic activity (data not shown).

GH158 is the fourth clan GH-A family, in addition to GH17, -128, and -148, now known to contain an *endo*- β (1,3)-glucanase activity (39, 40). The active-site cleft surrounding the catalytic sidechains is rich in surface-exposed aromatic residues oriented to engage in stacking interactions with the substrate (Fig. 4D). Structural alignment with a *Solanum tuberosum* GH17 *endo*- β (1,3)-glucanase:laminaritrose (G3G3G) complex (47) (root mean square deviations, 3.1 Å; sequence identity, 20.6%) infers the directionality of polysaccharide binding in *BuGH158* (Fig. 4D). Crucially, this superposition also reveals that 6-OH of the glucosyl residue in the -1 subsite is oriented toward a pocket lined with aromatic residues that may accommodate a single β (1,6)-linked glucose branch in *BuGH158* (Fig. 4E). Indeed, a glucosyl residue could easily be manually docked into this position without clashes (Fig. 4F). This structural feature provides a plausible explanation for the strict *LdLam* specificity and poor activity on yBG, the latter of which contains longer β (1,6)-linked glucose branches (48) (Fig. 2A).

(iii) Periplasmic saccharification of imported oligosaccharides is mediated by *BuGH3*, a specific *exo*- β (1,3)-glucosidase. GH3 is known to contain members with a diversity of *exo*- β -glycosidases (39), warranting detailed kinetic characterization of *BuGH3* in the context of the 1,3GUL. Initial substrate screening on chromogenic pNP glycosides revealed that *BuGH3* is an *exo*- β -glucosidase (no other activity detected on a panel of pNP substrates; see Table S5 at <https://doi.org/10.14288/1.0388792>). Subsequent Michaelis-Menten kinetics on diverse β -gluco-oligosaccharides further established that *BuGH3* is a specific β (1,3)-glucosidase poised to efficiently handle the hydrolysis products of the predicted cell surface GH16 and GH158 enzymes (Table S5 and Fig. S9 at the URL mentioned above). Indeed, incubation with the limit-digest products of either *BuGH16* or *BuGH158* confirmed that *BuGH3* is capable of completely degrading all oligosaccharide products of laminarin and bMLG to glucose (Fig. S2 and S10 at the URL mentioned above).

Despite possessing an apparently broad ability to hydrolyze β (1,3)-, β (1,4)-, and β (1,6)-glucosides, the particular preference of *BuGH3* for β (1,3)-linkages is highlighted by a two-orders-of-magnitude higher k_{cat}/K_m value for laminaribiose (G3G) over cellobiose (G4G) and gentiobiose (G6G) (Table S5 at the URL mentioned above). Concordant with this finding, catalytic efficiency toward mixed-linkage trisaccharides with a β (1,3)-glucose at the nonreducing end (G3G4G) was 2 orders of magnitude higher than that with a β (1,4)-glucose at the nonreducing end (G4G3G) (Table S5 at the URL mentioned above). Laminari-oligosaccharides of increasing degree-of-polymerization were hydrolyzed with comparable k_{cat} values, but K_m values decreased between the di- and trisaccharide, after which K_m values leveled off, suggesting that the *BuGH3* has two positive subsites of kinetic significance.

(iv) *BuSGBP-B* mediates β -glucan specificity. The 1,3GUL encodes two potential SGBPs: the SusD-homolog *BuSGBP-A* and the sequence-divergent, “SusE-positioned” (13) *BuSGBP-B*. Notably, qualitative screening of a library of soluble polysaccharides by affinity gel electrophoresis (AGE), as well as isothermal titration calorimetry (ITC), indicated that *BuSGBP-A* does not bind any likely substrates, including *LdLam* and bMLG (Fig. 5; see also Tables S6 and S7 and Fig. S11 and S12 at <https://doi.org/10.14288/1.0388792>). SGBP-A homologs that do not bind polysaccharide currently are rare but not entirely unknown (25); indeed, the primary role of this *Bacteroides* PUL component appears to be its indispensable structural association with the cognate TBDT (16, 26, 46, 49).

In contrast, AGE demonstrated that *BuSGBP-B* was able to bind *LdLam*, *EbLam*, yBG, and bMLG (Fig. 5). The smaller shift observed for *LdLam* than for *EbLam* (Fig. 5) recapitulates observations for an SGBP-A from a marine *Bacteroidetes*, *Gramella* sp. (50). Quantitative ITC indicated an order-of-magnitude-higher affinity constant (K_a) for *LdLam* over bMLG (Table S6 at the URL mentioned above), revealing the high specificity of *BuSGBP-B* for β (1,3)-glucan. The smallest laminari-oligosaccharide bound by

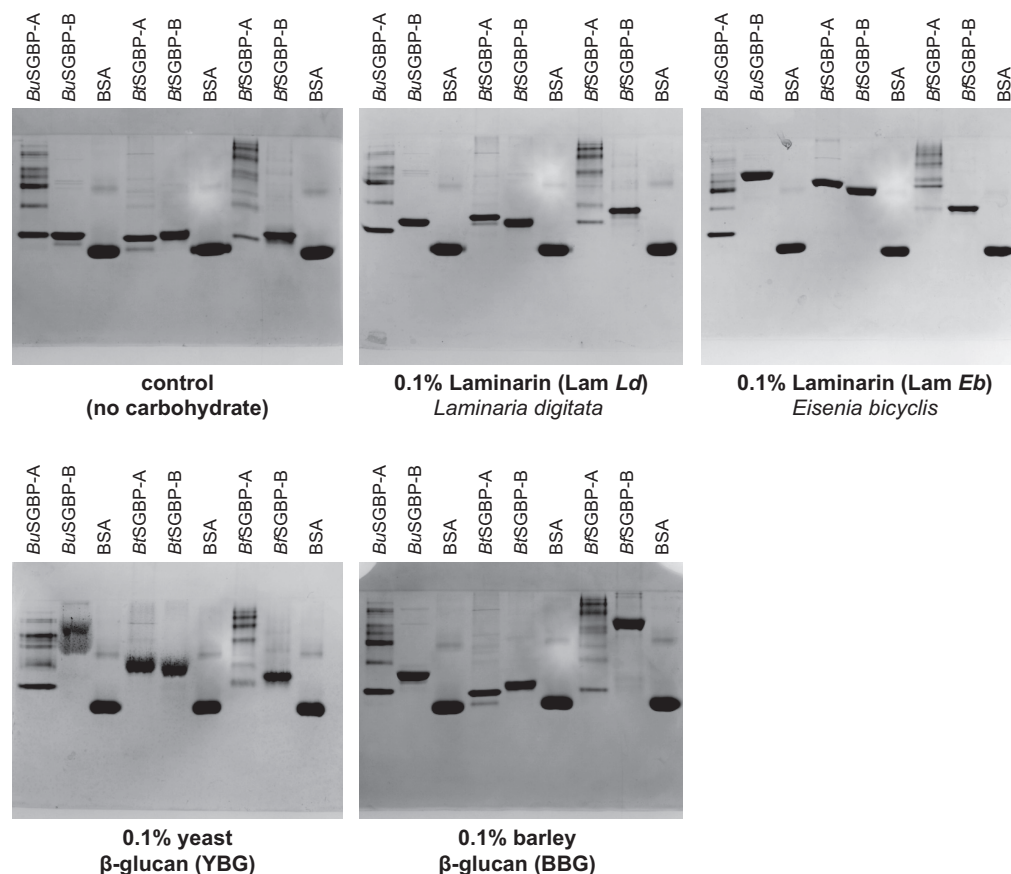


FIG 5 Binding of *Bacteroides* SGBPs to $\beta(1,3)$ -glucans. Native polyacrylamide (10%) gel electrophoresis containing 0.1% polysaccharide, with bovine serum albumin (BSA) as a control protein.

BuSGBP-B was the trisaccharide (G3G3G), and affinity increased with degree of polymerization (Table S7 and Fig. S12 at the URL mentioned above). Neither SGBP recognized insoluble crystalline cellulose [$\beta(1,4)$ -glucan] in pull-down assays (data not shown), unlike SGBPs that bind MLG (26), which further underscores the $\beta(1,3)$ -glucan specificity of *BuSGBP-B*.

Divergent GH and SGBP specificities collectively dictate the range of β -glucans utilized by *Bacteroides* species. Using the *B. uniformis* 1,3GUL as the archetype, we identified several homologous 1,3GULs from the *Bacteroidaceae*, *Porphyromonadaceae*, and *Prevotellaceae* families (order *Bacteroidales*), of which five representatives are shown in Fig. 1. All comprise a syntenic TBDT, SGBP-A, SGBP-B, and GH3 as the core set of conserved genes. However, the predicted *endo*-glucanase (GH16 and GH158) content is notably variable, and the SGBPs have particularly low sequence similarity, including among related *Bacteroides* species (Fig. 1 and Table S8 at <https://doi.org/10.14288/1.0388792>). With a focus on this genus, we tested the growth of *B. thetaiotaomicron* and *B. fluxus vis-à-vis B. uniformis* to determine how heterogeneity in 1,3GUL gene content might affect the utilization of individual β -glucans. Strikingly, whereas all three species were able to grow on *LdLam* as the sole carbon source, only *B. uniformis* and *B. thetaiotaomicron* grew on *yBG*, while only *B. uniformis* grew on *bMLG* (Fig. 2B to D and Table S1 at the URL mentioned above).

To elucidate the molecular basis of this species-specific β -glucan utilization, we characterized the transcriptional response of the *B. thetaiotaomicron* and *B. fluxus* 1,3GUL (as a measure of HTCS specificity) and the biochemistry of the predicted surface *endo*-glucanases and SGBPs *vis-à-vis* the *B. uniformis* system. In *B. thetaiotaomicron* and *B. fluxus*, the genes encoding the TBDT and SGBP-A were strongly upregulated by

LdLam and *yBG*, whereas they were only very weakly activated by *bMLG* (<10-fold) (Fig. 2E). These results are generally concordant with the observed growth phenotypes among all three *Bacteroides* species, although we unexpectedly observed strong up-regulation of the *B. fluxus* 1,3GUL (*Bf1,3GUL*) with *yBG*, on which *B. fluxus* is not able to grow. The *Bf1,3GUL* still was highly activated after dialyzing the substrate to remove any potential small oligosaccharide signals (Fig. S13 at the URL mentioned above), suggesting that this system is capable of sufficiently cleaving *yBG* to release an oligosaccharide that activates this PUL but subsequently unable to grow appreciably on the bulk polysaccharide. These distinct regulatory profiles likely involve a combination of surface GH and transporter specificity, combined with the sensing abilities of the respective HTCS regulators. The latter signaling specificity is notable in light of high protein sequence similarity among species (80 to 85% amino acid identity and 88 to 91% similarity versus the *BuHTCS*; Table S8 at the URL mentioned above), and the results below show that at least *B. thetaiotaomicron* GH16 (*BtGH16*) can cleave *bMLG* but does generate an activating cue.

Whereas *B. uniformis* encodes both a GH16_3 member and a GH158 member in its 1,3GUL, *B. thetaiotaomicron* possesses only a GH16_3 member while *B. fluxus* possesses only a GH158 member. Enzymology revealed that *BtGH16* is a predominant laminarinase with lower, yet comparable, activity on *yBG* and *bMLG*, similar to *BuGH16* (Fig. S14 and Table S2 at the URL mentioned above; the homologous *Prevotella loescheii* GH16 [*PlGH16*] also had broad activity). In contrast, *BfGH158* was highly specific for *LdLam*, with low activity on *yBG* and very poor activity on *bMLG*, similar to *BuGH158* (Table S2 at the URL mentioned above).

We also assessed the β -glucan specificities of the SGBPs from the *B. thetaiotaomicron* and *B. fluxus* 1,3GULs for comparison with those of *B. uniformis*. *BtSGBP-A*, *BtSGBP-B*, *BfSGBP-A*, and *BfSGBP-B* each bound the all- $\beta(1,3)$ -linked *LdLam*, *EbLam*, and *yBG*, as shown by AGE and ITC, although *BfSGBP-A* and *-B* interacted only weakly with *yBG* (Fig. 5 and Table S6 and Fig. S11 at the URL mentioned above). Strikingly, *BfSGBP-B* migration was strongly hindered by the $\beta(1,3)/\beta(1,4)$ -linked *bMLG* in AGE, whereas that of *BfSGBP-A*, *BtSGBP-A*, and *BtSGBP-B* was unaffected. Indeed, ITC indicated that *BfSGBP-B* bound *bMLG* with an affinity similar to that of *BuSGBP-B* (Table S6 at the URL mentioned above).

These biochemical data show that whereas orthologous GH specificity is well predicted by the family to which it belongs, considerable diversity exists in the specificity of syntenic SGBPs encoded by 1,3GULs of *B. uniformis*, *B. thetaiotaomicron*, and *B. fluxus*. When combined with microbiological data, a pattern emerges whereby GH content combined with SGBP specificity predicts the range of β -glucan congeners a *Bacteroides* species can utilize. In parallel, the signaling specificity or promiscuity through the single HTCS sensor associated with each system must also be able to respond to the diversity of cues generated from cleavage of related substrates. Of the three species examined, only *B. uniformis* possesses the full complement of synergistic HTCS, GH, and SGBP specificities to enable growth on laminarin, *yBG*, and *bMLG* (summarized in Fig. 6).

Metagenomic survey reveals global 1,3GUL distribution in the human gut. To determine the prevalence of 1,3GULs in the human gut, we surveyed publicly available gut metagenomic data from 2,441 adults across five different continents (North American, South America, Africa, Europe, and Asia) (Fig. 7). Despite similar genetic synteny, these 1,3GULs have different nucleotide sequences (pairwise identities for *B. fluxus*/*B. thetaiotaomicron*, 44%; *B. fluxus*/*B. uniformis*, 60%; *B. thetaiotaomicron*/*B. uniformis*, 44%), allowing us to use the nucleotide sequence to distinguish the presence or absence of a specific 1,3GUL in each metagenomic sample. The *B. uniformis* type 1,3GUL is the most abundant across all samples (48% of the total population), perhaps reflecting the high prevalence of this species in humans from industrialized populations. This is followed by the *B. thetaiotaomicron* type 1,3GUL (26% of samples) and *B. fluxus* (0.53% of samples). The low prevalence of *B. fluxus* may be explained by its low

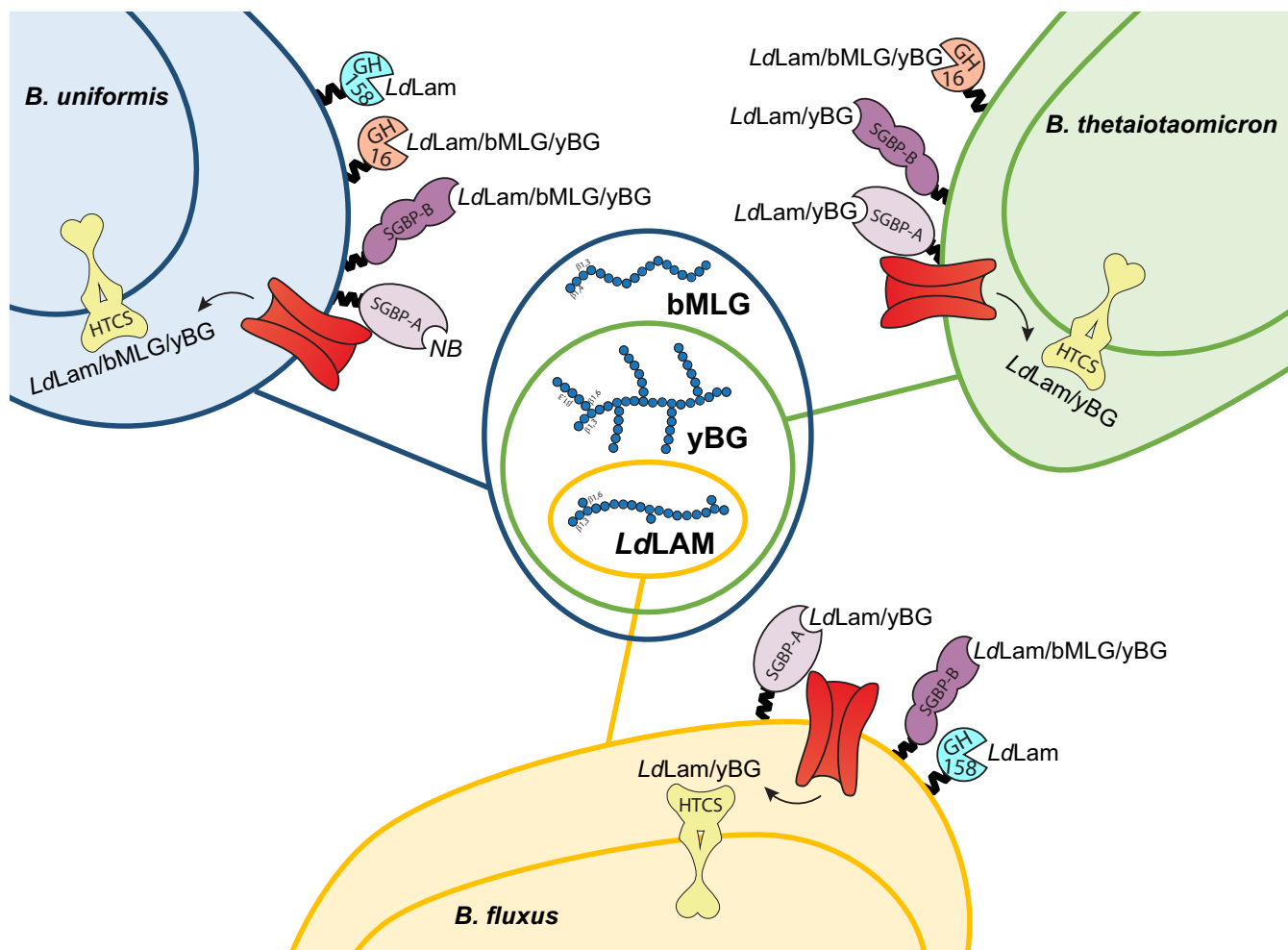


FIG 6 Diagram of β -glucan utilization by *Bacteroides* of the HGM. For each species, the binding or catalytic specificity of the proteins encoded by the endogenous 1,3GUL is indicated on the cell surface (Lam, laminarin; bMLG, barley mixed-linkage β -glucan; yBG, yeast β -glucan; NB, no binding). The capacity of each species to utilize individual β -glucans for growth, which arises as a combination of these individual specificities, as well as periplasmic transport/sensing, is indicated in the Venn diagram.

abundance and prevalence in European, North American, and Asian populations, consistent with previous observations for the xyloglucan utilization locus (15). Although 1,3GULs are widely distributed throughout human populations (59%), we do not see correlation with any particular geographic region or population. This ubiquity may be explained by the prevalence of $\beta(1,3)$ -glucans in different diets, since edible fungi and yeast fermentation products are common worldwide. Strikingly, we were unable to detect any of the three identified 1,3GULs in the indigenous Hadza and Yanomami tribes, which may be due to a high prevalence of *Prevotella* and not *Bacteroides* in these populations (51, 52). The present study, in combination with our previous metagenomic survey of MLG utilization (19), reveals a broadly represented potential for specific β -glucan metabolism to establish niches for individual human gut bacteria.

DISCUSSION

The human large intestine is a highly competitive ecosystem, in which access to a wide range of carbohydrates confers selective advantage. Here, we outline a model in which a single PUL enables *B. uniformis* to utilize a range of β -glucan congeners (Fig. 1). In this model, *LdLam*, *yBG*, or *MLG* (Fig. 2) is bound by *BuSGBP-B* on the surface of the bacterium, whereas the *SusD* homolog *BuSGBP-A* is passive in this step. Depending on the backbone linkage composition and extent of branching, individual β -glucans are cleaved into oligosaccharides by one or both cell surface-anchored *endo*-glucanases.

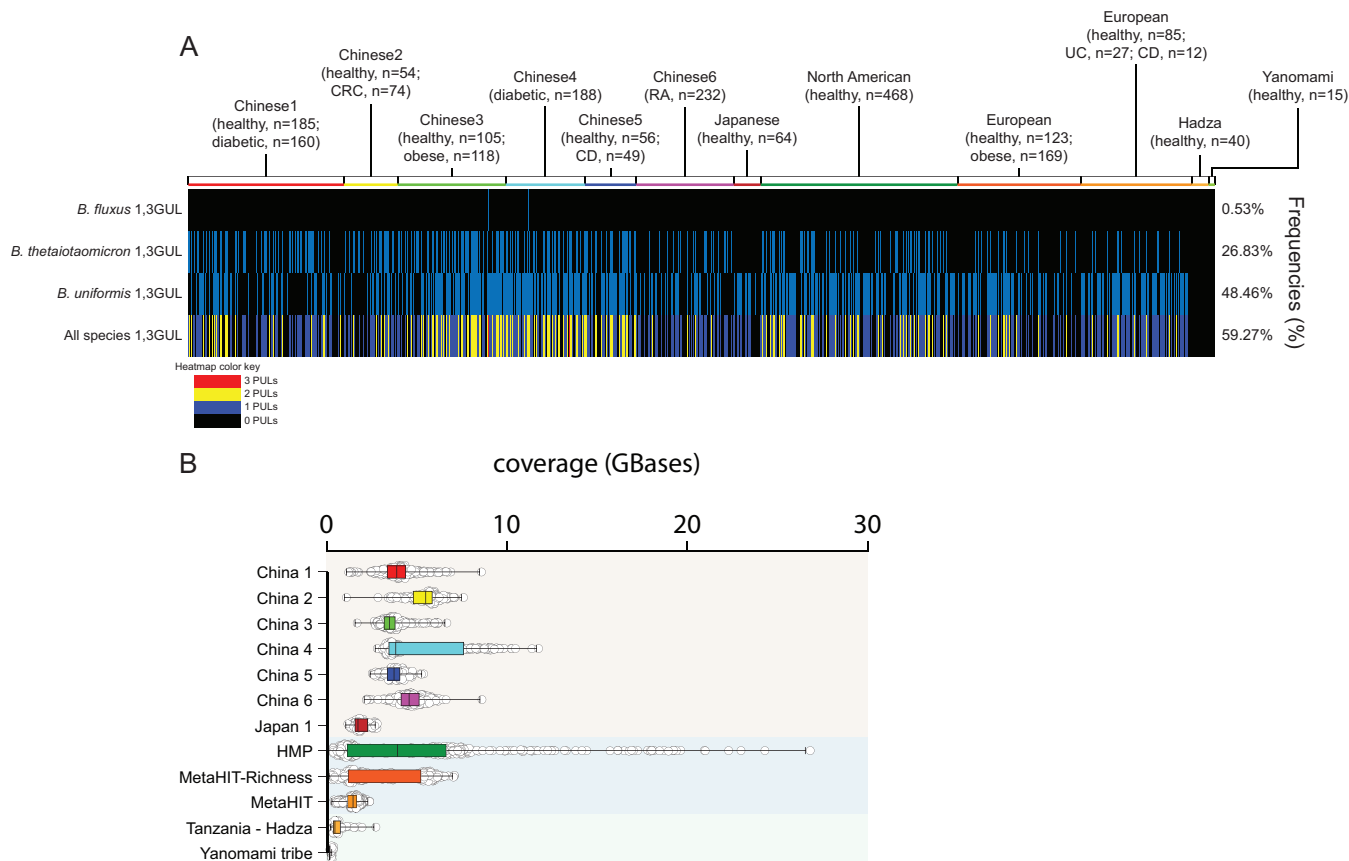


FIG 7 Prevalence of 1,3-GULs in 2,441 human metagenomes. (A) Each line represents the presence (blue) or absence (black) of a specific 1,3GUL species related to a single human gut metagenomic sample. The bottom row represents the total number of 1,3GUL that each individual possesses, colored according to the legend. The frequency of 1,3GUL incidence across all 2,441 individuals is shown on the right. (B) Coverage variation of each metagenomic sample is indicated by individual samples (open circles) and a box plot with the mean.

Whereas *BuGH158* is a highly specific laminarinase, *BuGH16* is a generalist *endo*- β (1,3)-glucanase that can accommodate backbone β (1,4)-glucosyl linkages (as in MLG) and long/frequent β (1,6)-glucosyl branches (as in *EbLam* and *yBG*). Oligosaccharides produced by these *endo*-glucanases are actively imported by the associated TBDT into the periplasm, where the *exo*- β -glucosidase *BuGH3* completes saccharification. Notably, the poor activity of *BuGH3* toward β (1,6) and β (1,4) linkages suggests that other periplasmic glucosidases encoded outside 1,3GUL assist with oligosaccharide debranching and complete MLG oligosaccharide hydrolysis. The genome of *B. uniformis* ATCC 8492 encodes three members of GH30, which is known to contain *exo*- β (1,6)-glucosidases active on laminarin (53), as well as twenty additional members of GH3, which could comprise a potent *exo*- β (1,4)-glucosidase.

The *Bacteroides* 1,3GULs characterized here share partial homology with various laminarin-targeting PULs from marine *Bacteroidetes*, including *Gramella* (54), *Formosa* (55), and other closely related genera within the family *Flavobacteriaceae* (56), raising questions of evolutionary origin of 1,3GULs in the HGM. Indeed, PULs that target porphyran (a galactan-based seaweed polysaccharide) are thought to have been acquired by the HGM from the marine *Bacteroidetes* *Zobellia galactanivorans* via a horizontal gene transfer event (57). In addition to syntenic 1,3GULs, recently characterized *Formosa* species also harbor enlarged laminarin-targeting PULs containing other enzymes. In this system, laminarin must first be debranched by a GH30 *exo*- β (1,6)-glucosidase before being hydrolyzed by a GH17 *endo*- β (1,3)-glucanase, which has very low activity on β (1,6)-glucose-branched laminarin (55). In general, GH17 endoglucanases have a very narrow active-site cleft that does not effectively accom-

modate such branches (47, 55). This is unlike *BuGH16* and *BuGH158* enzymes, which allow the 1,3GUL system to bypass the initial debranching requirement. Thus, although GH16 and GH17 laminarin-active enzymes may display similar overall $\beta(1,3)$ linkage specificities (58), differences in active-site structure may delineate distinct roles in the stepwise total hydrolysis of laminarins.

Comparative genomic and biochemical analysis revealed variations in GH content that, together with complementary HTCS, GH, and SGBP specificity, dictate selective β -glucan utilization. Among the *Bacteroides* species tested, *B. uniformis* is unique in its ability to utilize the trifecta of β -glucans containing contiguous or individual $\beta(1,3)$ linkages, viz., *LdLam*, *yBG*, and *bMLG*. In contrast, *B. thetaiotaomicron* can only access *LdLam* and *yBG*, while *B. fluxus* is further restricted to utilizing *LdLam* only (summarized in Fig. 6). None of the species tested were able to utilize curdlan, perhaps due to its poor solubility in water. Orthologous GHs from syntenic 1,3GUL share similar specificities, yet the GH complement alone is insufficient to confer growth on a particular β -glucan. For example, *B. thetaiotaomicron* produces a GH16 member that is an approximately equally competent laminarinase, *yBGase*, and *MLGase* (see Table S2 at <https://doi.org/10.14288/1.0388792>). However, an inability to capture *bMLG* at the cell surface, due to a lack of polysaccharide binding by the cognate SGBPs (Fig. 5), and an inability to transport or recognize breakdown products in the periplasm (Fig. 2) precludes growth on this abundant dietary glycan. It is interesting to ponder to what extent losses or gains of function are correlated and arise (a)synchronously through the evolution of these distinct protein families more widely among the *Bacteroidetes*.

In this regard, *B. fluxus* is the least versatile of the three *Bacteroides* species explored here. The inability of *B. fluxus* to utilize *bMLG* clearly is the consequence of three factors (Fig. 6): (i) the lack of a polyspecific GH16 ortholog encoded by its 1,3GUL (cf. *B. uniformis* and *B. thetaiotaomicron*), (ii) the extremely poor *MLGase* activity of *BfGH158* (Table S2 at the URL mentioned above), and (iii) an inability to transport/sense this cereal polysaccharide in the periplasm (Fig. 2). However, we note that *B. fluxus* possesses an SGBP-B able to bind both laminarin and *bMLG* (Fig. 5 and Table S6 at the URL mentioned above), which is correlated with *MLG* utilization in the absence of an *MLG*-binding SGBP-A in *B. uniformis* (Fig. 6). Further, the inability of *B. fluxus* to grow on *yBG* is consistent with weak binding by *BfSGBP-A* and *BfSGBP-B*, poor catalytic efficiency of *BfGH158* toward *yBG*, and the lack of a compensatory GH16 ortholog. This is despite an HTCS that is equally highly responsive to *yBG* and *LdLam* (Fig. 2), as well as both SGBP types binding both of these 1,3- β -glucans. Hence, we might anticipate that a gain-of-function mutation to introduce *yBG* hydrolase activity, for example, through acquisition of a versatile GH16 member or broadening the substrate scope of the extant GH158, would result in growth on *yBG*. Testing this hypothesis will depend on the future generation of a transformation system for *B. fluxus* or discovery of a corresponding wild-type strain.

The present study underscores that dietary specificity of related gut commensal strains is gleaned most precisely by systems-based approaches involving genomic, biochemical, and structural biological dissection of PULs (14–27), which otherwise could not have been predicted based on sequencing data alone. Whereas the GH complement of PULs has been shown to drive specificity in the levan/fructan system (14), to our knowledge the present study describes the first case in which the interplay of GHs and SGBPs, underpinned by HTCS specificity, collectively dictate glycan utilization among species. Indeed, in most PULs characterized to date, the specificities of the vanguard GH and SGBPs are concordant (15, 17, 19, 23, 25, 26, 46). On the other hand, the *B. fluxus* case reveals how a limited endoglycanase in the context of polyspecific SGBPs and an HTCS can restrict nutrient range. Thus, the evolution of both the gene complement and the specificity of individual components within a PUL allows bacteria to access to distinct nutrient niches in the competitive human gut environment. In the context of human nutrition and health, the detailed characterization presented here provides a validated set of molecular markers (Fig. 1) to identify $\beta(1,3)$ -glucan utilization potential among members of the HGM (and other microbiota). This insight may prove

especially transformative for our future ability to select strains and dietary formulations in tailoring microbial intervention therapies (59).

MATERIALS AND METHODS

Substrates. Tamarind seed xyloglucan, barley beta-glucan (60), konjac glucomannan, carob galactomannan, *Alcaligenes faecalis* curdlan, and yeast beta-glucan (48) were purchased from Megazyme International (Bray, Ireland). Carboxymethyl cellulose was purchased from Acros Organics (Morris Plains, NJ, USA). Hydroxyethyl cellulose was purchased from Amresco (Solon, OH, USA). *Xanthomonas campestris* xanthan gum was purchased from Spectrum (New Brunswick, NJ, USA). Laminarins from *Laminaria digitata* (61) and *Eisenia bicyclis* (42) were purchased from Sigma-Aldrich (St. Louis, MO, USA) and Carbosynth (Compton, UK), respectively. Laminarin from *Laminaria digitata* was reduced to laminaritol, as described previously (62), to reduce background responses in the bicinchoninic acid (BCA) reducing-sugar enzyme kinetics assay.

Laminaribiose (G3G), laminaritriose (G3G3G), laminaritetraose (G3G3G3G), laminaripentaose (G3G3G3G3G), mixed-linkage glucotriose A (G3G4G), mixed-linkage glucotriose B (G4G3G), and cellotriase (G4G4G) were purchased from Megazyme. Gentiobiose (G6G) was purchased from Carbosynth (Compton, UK). Cellobiose (G4G) was purchased from Acros Organics. G3G-CNP was synthesized by glycosylation of α -laminaribiosyl bromide (63) with 2-chloro-4-nitrophenol under phase-transfer conditions (63, 64), the details of which will be published elsewhere.

Bacterial growth experiments. *Bacteroides uniformis* ATCC 8492 (NCBI accession no. NZ_AAYH00000000.2, JGI accession no. 641380447), *Bacteroides thetaiotaomicron* NLAE-zl-H207 (NCBI accession no. NC_004663.1, JGI accession no. 2515154063), and *Bacteroides fluxus* YIT12057 (NCBI accession no. NZ_AFBN00000000.1, JGI accession no. 651324011) were grown in tryptone yeast extract glucose (TYG) medium at 37°C under anaerobic conditions (Coy anaerobic chamber; 85% N₂, 10% H₂, 5% CO₂). These cultures were used to inoculate minimal medium containing glucose as the sole carbon source (MM-Glc), followed by incubation at 37°C for 20 h. One-milliliter samples then were centrifuged and the bacterial pellets were gently resuspended in MM containing no carbohydrate (MM-NC). These suspensions were diluted 1:50 in MM-NC before being used to inoculate MM containing 0.5% carbohydrate. Growth experiments were performed in replicates of 6 (laminarin, yeast β -glucan, barley β -glucan, and curdlan) or 3 (glucose and H₂O) in 96-well plates. Growth was monitored by measuring the A₆₀₀, and data were processed using Gen5 software (BioTek). Growth was quantified in each assay by first identifying a minimum time point (A_{min}) at which the A₆₀₀ had increased by 10% over a baseline reading taken during the first 500 min of incubation. We then identified the time point at which the A₆₀₀ reached its maximum (A_{max}) immediately after exponential growth. The growth rate for each well was defined by $(A_{max} - A_{min}) / (T_{max} - T_{min})$, where T_{max} and T_{min} are the corresponding time values for each absorbance. Cultures where the density did not increase by at least 0.1 (A₆₀₀) were considered to have no growth.

Quantitative reverse transcription-PCR (qRT-PCR). *B. uniformis* was capable of growth on all carbon sources of interest in this study and, therefore, was cultured directly in 3 ml of MM containing 0.5% (wt/vol) carbohydrate, as described above. *B. thetaiotaomicron* and *B. fluxus* were precultured on MM-Glc, pelleted, washed, resuspended twice in MM-NC, and inoculated to an A₆₀₀ of ~0.3 in 4 ml of MM containing 0.5% (wt/vol) carbohydrate. Bacterial cultures were harvested in triplicate (at mid-log phase [A₆₀₀ ~0.6] for *B. uniformis* and after 5 h of incubation for *B. thetaiotaomicron* and *B. fluxus*), placed in RNA protect (Qiagen) for immediate stabilization of RNA, and then stored at -20°C. RNA was extracted and purified with the RNeasy minikit (Qiagen), and RNA purity was assessed by spectrophotometry. One μ g of RNA was used for reverse transcription and synthesis of the cDNA (SuperScript VILO master mix; Invitrogen). Quantitative PCRs (20 μ l final volume) using specific primers were performed with a SensiFast SYBR Lo-ROX kit (Bioline) on a 7500 Fast real-time PCR system (Applied Biosystems) (see Table S9 at <https://doi.org/10.14288/1.0388792>). Data were normalized to 16S rRNA transcript levels, and changes in expression levels were calculated as fold change compared with levels for cultures of MM containing glucose.

Bioinformatics, gene cloning, and site-directed mutagenesis. Potential 1,3GULs were identified by homology searches of sequences available in the Joint Genome Institute Integrated Microbial Genomes and Metagenomes (JGI-IMG/M) Database (65) and other sources (41). Signal peptides and subcellular localization were predicted by protein sequence analysis (66–68). The gene fragments corresponding to BACUNI_01484 to BACUNI_01488 (encoding BuGH3, BuGH158, BuGH16, BuSGBP-B, and BuSGBP-A), H207DRAFT_02225 to H207DRAFT_02227 (encoding BtSGBP-A, BtSGBP-B, and BtGH16), HMPREF9446_00612 to HMPREF9446_00614 (encoding BfGH158, BfSGBP-B, and BfSGBP-A), and HMPREF1991_02176 (encoding PIGH16) were PCR amplified from genomic DNA using the Q5 high-fidelity polymerase (NEB) with primers designed to exclude signal peptides and lipidation cysteines (66, 69) (Table S10 at the URL mentioned above). The PCR introduced Sall/XhoI (NEB) restriction sites to the flanks of BACUNI_01484 and _01486 target genes, and the amplified DNA products were ligated into the expression vector pET28a such that the encoded recombinant proteins contain an N-terminal His₆ tag. The rest of the PCR products contained appropriate pMCSG complementary sequences for subsequent ligation-independent cloning into pMCSG53 or pMCSG-GST plasmids, providing an N-terminal His₆ tag or an N-terminal His₆-glutathione S-transferase (GST) tag (70). Successful cloning was confirmed by colony PCR (GoTaq polymerase from Promega) and sequencing (Genewiz). The site-directed mutants (71) BuGH158 N136A, E137A, and E220A were generated using pMCSG53::BuGH158 as a template DNA (Table S11 at the URL mentioned above).

Recombinant protein production and purification. Recombinant proteins were produced in *Escherichia coli* BL21(DE3) cells cultured in TB broth containing ampicillin (50 μ g·ml⁻¹) or kanamycin

(50 $\mu\text{g}\cdot\text{ml}^{-1}$) at 37°C (200 rpm). Cells were grown to mid-exponential phase (optical density at 600 nm [OD_{600}], ~ 0.4 to 0.6). Overexpression was induced by adding isopropyl β -D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM, and the cultures were further grown at 16°C (200 rpm) for 18 h. The cells were harvested by centrifugation and sonicated, and His₆-tagged recombinant proteins were purified via immobilized nickel affinity chromatography (His-Trap; GE Healthcare) utilizing a gradient elution of up to 100% elution buffer containing 20 mM sodium phosphate, pH 7.4, 500 mM NaCl, and 500 mM imidazole in a BioLogic fast protein liquid chromatography system (Bio-Rad). The purity of the recombinant proteins was determined by SDS-PAGE (Fig. S15 and S16 at <https://doi.org/10.14288/1.0388792>), and their concentrations were determined from calculated molar extinction coefficients at 280 nm using an Epoch microplate spectrophotometer (BioTek).

Selenomethionine-labeled protein was produced by inhibition of methionine biosynthesis in *E. coli* BL21(DE3) (72). Briefly, cells were grown in 1 liter of M9 minimal medium supplemented with 100 $\mu\text{g}\cdot\text{ml}^{-1}$ ampicillin at 37°C with shaking until the OD_{600} reached 0.6. At this point, 100 mg each of L-lysine, L-threonine, and L-phenylalanine and 50 mg each of L-leucine, L-isoleucine, L-valine, and L-selenomethionine were added to the medium and shaken for a further 15 min before inducing expression with 0.5 mM IPTG. The culture was transferred to 16°C and incubated for an additional 24 h. Nickel affinity chromatography was conducted as described above using HEPES buffer instead of sodium phosphate. Native and selenomethionine-labeled proteins for crystallography were further purified by size exclusion chromatography through a Superdex 75 resin (GE Healthcare Life Sciences) in an XK 16/100 column (GE Healthcare Life Sciences) run in 10 mM HEPES, pH 7.0, at 0.8 $\text{ml}\cdot\text{min}^{-1}$.

Affinity gel electrophoresis. Affinity polyacrylamide gel electrophoresis was performed for 180 min at 80 V and room temperature on nonreducing 10% (wt/vol) polyacrylamide gels containing a polysaccharide concentration of 0.1% (wt/vol) (or water for a control), essentially as previously described (15, 26, 46). Five micrograms of the tested SGBP proteins, along with bovine serum albumin (BSA) used as noninteracting negative-control protein, were loaded on the gels.

ITC. Isothermal titration calorimetry (ITC) was performed using the MicroCal VP-ITC titration calorimeter equilibrated to 25°C, essentially as previously described (15, 26, 46). The proteins (20 to 100 μM) were placed in the sample cell, and the syringe was loaded with 2.5 mg/ml polysaccharide or 0.5 to 2 mM oligosaccharide. Following an initial injection of 2 μl , 25 subsequent injections of 10 μl were performed with stirring at 280 rpm, and the resulting heat of the reaction was recorded. Integrated heats were fit to a single-site model using Microcal Origin v7.0 to derive n , K_{d} , and ΔH values.

Carbohydrate analytical method. High-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) was performed on CarboPac PA200 guard and analytical columns connected in series on a Dionex ICS-5000 high-performance liquid chromatography system, operated by Chromeleon software, version 7 (Thermo Scientific), essentially as previously described (73). Solvent A was ultrapure water, solvent B was 1 M sodium hydroxide, and solvent C was 1 M sodium acetate (anhydrous bio ultra-grade; Sigma-Aldrich). The injection volume was 10 μl , and the gradient was the following: 0 to 5 min, 10% B and 3.5% C; 5 to 12 min, 10% B, linear gradient from 3.5 to 30% C; 12.0 to 12.1 min, 50% B, 50% C; 12.1 to 13.0 min, exponential gradient (curve setting, 9) of B and C back to initial conditions; 13 to 17 min, initial conditions.

Enzymatic assays. Polysaccharide hydrolysis was quantified using a BCA reducing-sugar assay (74). Assays were conducted in a final volume of 100 μl at the optimum pH and 37°C for 10 min. Reactions were terminated by the addition of an equal volume (100 μl) of BCA reagent. Color was developed by heating to 80°C for 20 min before reading A_{563} . Glucose (25 to 150 μM) was used to generate a standard curve for quantitation. The pH and temperature optima of each enzyme were initially determined using the same enzyme reaction assay to quantify reducing ends over 10 min of incubation with 1.0 mg/ml laminarin in different buffers at 50 mM: sodium citrate (pH 3.0 to 6.5), sodium phosphate (pH 6.5 to 8.0), and glycine (pH 9.0 to 10.5). To determine Michaelis-Menten parameters, different concentrations of polysaccharide solutions were used over the range of 0.025 to 3 $\text{mg}\cdot\text{ml}^{-1}$ with the appropriate concentration of enzyme for 10 min, and the numbers of reducing ends released were quantified as described above.

The release of glucose from oligosaccharides was quantified using the D-glucose HK assay kit from Megazyme (Bray, Ireland), modified for use as a continuous assay exactly as previously described (19).

To measure enzyme activity on chromogenic glycosides, the release of *para*-nitrophenyl was monitored by following the A_{405} in a 1-cm-path-length quartz cuvette with a Cary 60 UV-visible spectrophotometer (Agilent Technologies). Reaction mixtures in 250 μl at the optimum pH and 37°C were assayed with nine different substrate concentrations, and the rate was calculated using an extinction coefficient determined according to the buffer used. Endpoint assays were used for pH and temperature optima (same range as that described above) of BuGH3. Reactions were terminated after 10 min by the addition of 100 μl of 1 M Na_2CO_3 to raise the pH, and the A_{405} was measured in 96-well plates on an Epoch microplate spectrophotometer (BioTek). An extinction coefficient of 18,100 $\text{M}^{-1}\cdot\text{cm}^{-1}$ was used for these assays. Continuous assays were used for initial-rate saturation kinetics and reactions, initiated by adding 25 μl of enzyme solution to 225 μl of the remaining assay mixture in the optimum pH buffer at 37°C.

Crystallization and structure determination. Initial sitting-drop crystal screens were set up in 96-well plates using a Phoenix robot (Art Robbin Instruments) and were stored at room temperature. A hit was obtained in pHClear (Qiagen), condition E12 (0.1 M bicine, pH 9.0, 1.6 M ammonium sulfate), with purified BuGH158 at 22.7 $\text{mg}\cdot\text{ml}^{-1}$ and optimized in larger hanging drops in a grid screen by varying pH and ammonium sulfate concentrations in 24-well plates. Crystals were cryoprotected in crystallization solution supplemented with 2 M lithium acetate before flash freezing with liquid nitrogen, and diffraction data were collected at the Advanced Photon Source (APS) beamline 23ID-D. Selenomethionine-labeled

crystals were obtained in the same grid screen as the native crystals but required macroseeding with fine needles to obtain crystals of sufficient thickness. Crystals were cryoprotected in crystallization solution supplemented with 4 M lithium chloride before flash freezing with liquid nitrogen. Diffraction data were collected at the Stanford Synchrotron Radiation Lightsource (SSRL) beamline 9-2 at the selenium absorption peak, inflection, and high-energy remote wavelengths as determined by a fluorescence scan. Data sets were indexed and integrated with XDS (75), space groups were determined with Pointless (76), and data reduction was performed with Aimless (77). Phasing by multiple anomalous dispersion, density modification, and initial model building was performed in AutoSol (78) in the Phenix suite (79). This selenomethionine-labeled structure was used as the search model for molecular replacement with the native crystal data using Phaser (80). After initial refinement in Phenix.refine (81), iterative rounds of manual model building and refinement were conducted with Coot (82) and Refmac5 (83), respectively, in the CCP4 suite (84). The quality of the model was monitored throughout using Molprobity (85). A structure similarity search was conducted using the Dali server (86).

¹H-NMR determination of catalytic mechanism. BuGH158 in 50 mM sodium phosphate, pH 7.0, and laminaribiose- β -CNP were independently lyophilized and resuspended in 99.9% D₂O. After recording an initial ¹H-NMR spectrum of the substrate (Bruker Avance 400-MHz spectrometer), the enzyme was added to obtain final concentrations of 20 μ M BuGH158 and 10 mM laminaribiose- β -CNP. Spectra were recorded at appropriate time intervals thereafter to observe the first-formed product anomer and subsequent mutarotation.

Survey of human metagenomic data sets. Available cohorts of human gut metagenomic sequence data (National Center for Biotechnology Information projects PRJNA422434 [87], PRJEB10878 [88], PRJEB12123 [89], PRJEB12124 [90], PRJEB15371 [91], PRJEB6997 [92], PRJDB3601 [93], PRJNA48479 [94], PRJEB4336 [95], PRJEB2054 [96], PRJNA392180 [51], and PRJNA527208 [97]) were searched for the presence of 1,3GUL nucleotide sequences from *B. fluxus* (12.5 kb), *B. thetaiotaomicron* (12.6 kb), and *B. uniformis* (14.5 kb) using the following workflow. Each 1,3GUL nucleotide sequence was used separately as a template, and then magic-blast v1.5.0 (98) was used to recruit raw Illumina reads from the available metagenomic data sets with an identity cutoff of 97%. The alignment files next were used to generate a coverage map using bedtools v2.29.0 (99) to calculate the percent coverage of each sample against each reference. We considered a metagenomic data sample to be positive for a particular 1,3GUL if it had at least 70% of the corresponding 1,3GUL nucleotide sequence covered (since the three 1,3GUL sequences were very similar in size, no normalization was made for PUL template variation).

Data availability. The crystal structure data sets generated (coordinate files and structure factors) have been deposited in the Protein Data Bank (PDB) under the accession code 6PAL. We declare that the data supporting the findings of this study are available within the article and the supplemental material, the latter of which is available at <https://doi.org/10.14288/1.0388792>.

ACKNOWLEDGMENTS

Work at the University of British Columbia was generously supported by operating grants from the Canadian Institutes for Health Research (MOP-137134 and MOP-142472) and infrastructure support from the Canadian Foundation for Innovation (project number 30663) and the British Columbia Knowledge Development Fund. K.T. was partially supported by a 4-year doctoral fellowship from the University of British Columbia. E.C.M. acknowledges funding from the U.S. National Institutes of Health (R01 DK118024).

We thank the General Medical Sciences and Cancer Institutes Structural Biology Facility at the Advanced Photon Source (GM/CA@APS) for access to beamline 23ID-D. GM/CA@APS has been funded in whole or in part with Federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357.

We thank the Stanford Synchrotron Radiation Lightsource (SSRL) at the SLAC National Accelerator Laboratory (Menlo Park, CA, USA) for access to beamline 9-2, the use of which is supported by the U.S. DOE, Office of Science, Office of Basic Energy Sciences, under contract no. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM103393). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH.

We thank Stephanie Forget (Brumer group, UBC) for assistance with preliminary NMR analysis of GH158 reaction stereochemistry.

G.D. performed bacterial growth analyses, qPCR transcript analyses, comparative bioinformatics, GH and SGBP cloning, recombinant protein production, and biochem-

ical/biophysical/kinetic analyses. K.T. performed recombinant protein production and purification, all crystallography and tertiary structural analyses, and phylogenetic analysis of GH16, PFAM 13004 domain production and analysis and assisted with the characterization of *B. uniformis* enzymes (pH and temperature rate analyses, BuGH3 Glc-pNP kinetics). A.C. assisted G.D. with GH cloning, production, and characterization. N.J. synthesized laminaribiose-CNP and performed NMR analysis together with K.T. N.A.P. assisted G.D. with bacterial growth analyses in Michigan. G.P. performed human metagenome analysis. A.H.V. performed phylogenetic analysis of GH158. F.V.P. supervised crystallography and data analysis. H.B. and E.C.M. conceived the study and supervised research. G.D. drafted the manuscript, which was revised by K.T. and H.B.

We have no competing interests to declare.

REFERENCES

- Backhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JL. 2005. Host-bacterial mutualism in the human intestine. *Science* 307:1915–1920. <https://doi.org/10.1126/science.1104816>.
- Clemente JC, Ursell LK, Parfrey LW, Knight R. 2012. The impact of the gut microbiota on human health: an integrative view. *Cell* 148:1258–1270. <https://doi.org/10.1016/j.cell.2012.01.035>.
- Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR. 2007. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci U S A* 104:13780–13785. <https://doi.org/10.1073/pnas.0706625104>.
- Hooper LV, Littman DR, Macpherson AJ. 2012. Interactions between the microbiota and the immune system. *Science* 336:1268–1273. <https://doi.org/10.1126/science.1223490>.
- Schwabe RF, Jobin C. 2013. The microbiome and cancer. *Nat Rev Cancer* 13:800–812. <https://doi.org/10.1038/nrc3610>.
- Sonnenburg JL, Backhed F. 2016. Diet-microbiota interactions as moderators of human metabolism. *Nature* 535:56–64. <https://doi.org/10.1038/nature18846>.
- Costello EK, Stagaman K, Dethlefsen L, Bohannan BJ, Relman DA. 2012. The application of ecological theory toward an understanding of the human microbiome. *Science* 336:1255–1262. <https://doi.org/10.1126/science.1224203>.
- Heintz-Buschart A, Wilmes P. 2018. Human gut microbiome: function matters. *Trends Microbiol* 26:563–574. <https://doi.org/10.1016/j.tim.2017.11.002>.
- David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, Ling AV, Devlin AS, Varma Y, Fischbach MA, Biddinger SB, Dutton RJ, Turnbaugh PJ. 2014. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505:559–563. <https://doi.org/10.1038/nature12820>.
- Koropatkin NM, Cameron EA, Martens EC. 2012. How glycan metabolism shapes the human gut microbiota. *Nat Rev Microbiol* 10:323–335. <https://doi.org/10.1038/nrmicro2746>.
- Porter NT, Martens EC. 2017. The critical roles of polysaccharides in gut microbial ecology and physiology. *Annu Rev Microbiol* 71:349–369. <https://doi.org/10.1146/annurev-micro-102215-095316>.
- Lapebie P, Lombard V, Drula E, Terrapon N, Henrissat B. 2019. Bacteroidetes use thousands of enzyme combinations to break down glycans. *Nat Commun* 10:7. <https://doi.org/10.1038/s41467-019-10068-5>.
- Grondin JM, Tamura K, Dejean G, Abbott DW, Brumer H. 2017. Polysaccharide utilization loci: fuelling microbial communities. *J Bacteriol* 199:e00860-16. <https://doi.org/10.1128/JB.00860-16>.
- Sonnenburg ED, Zheng H, Joglekar P, Higginbottom SK, Firbank SJ, Bolam DN, Sonnenburg JL. 2010. Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell* 141:1241–1252. <https://doi.org/10.1016/j.cell.2010.05.005>.
- Larsbrink J, Rogers TE, Hemsworth GR, McKee LS, Tauzin AS, Spadiut O, Klintner S, Pudlo NA, Urs K, Koropatkin NM, Creagh AL, Haynes CA, Kelly AG, Cederholm SN, Davies GJ, Martens EC, Brumer H. 2014. A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes. *Nature* 506:498–502. <https://doi.org/10.1038/nature12907>.
- Cameron EA, Kwiatkowski KJ, Lee B-H, Hamaker BR, Koropatkin NM, Martens EC. 2014. Multifunctional nutrient-binding proteins adapt human symbiotic bacteria for glycan competition in the gut by separately promoting enhanced sensing and catalysis. *mBio* 5:e01441-14. <https://doi.org/10.1128/mBio.01441-14>.
- Rogowski A, Briggs JA, Mortimer JC, Tryfona T, Terrapon N, Lowe EC, Basle A, Morland C, Day AM, Zheng H, Rogers TE, Thompson P, Hawkins AR, Yadav MP, Henrissat B, Martens EC, Dupree P, Gilbert HJ, Bolam DN. 2015. Glycan complexity dictates microbial resource allocation in the large intestine. *Nat Commun* 6:7481. <https://doi.org/10.1038/ncomms8481>.
- Cuskin F, Lowe EC, Temple MJ, Zhu Y, Cameron E, Pudlo NA, Porter NT, Urs K, Thompson AJ, Cartmell A, Rogowski A, Hamilton BS, Chen R, Tolbert TJ, Piens K, Bracke D, Verweken W, Hakki Z, Speciale G, Muñoz-Munoz JL, Day A, Peña MJ, McLean R, Suits MD, Boraston AB, Atherly T, Ziemer CJ, Williams SJ, Davies GJ, Abbott DW, Martens EC, Gilbert HJ. 2015. Human gut Bacteroidetes can utilize yeast mannan through a selfish mechanism. *Nature* 517:165–169. <https://doi.org/10.1038/nature13995>.
- Tamura K, Hemsworth GR, Dejean G, Rogers TE, Pudlo NA, Urs K, Jain N, Davies GJ, Martens EC, Brumer H. 2017. Molecular mechanism by which prominent human gut Bacteroidetes utilize mixed-linkage beta-glucans, major health-promoting cereal polysaccharides. *Cell Rep* 21:417–430. <https://doi.org/10.1016/j.celrep.2017.09.049>.
- Cartmell A, Lowe EC, Basle A, Firbank SJ, Ndeh DA, Murray H, Terrapon N, Lombard V, Henrissat B, Turnbull JE, Czejek M, Gilbert HJ, Bolam DN. 2017. How members of the human gut microbiota overcome the sulfation problem posed by glycosaminoglycans. *Proc Natl Acad Sci U S A* 114:7037–7042. <https://doi.org/10.1073/pnas.1704367114>.
- Bagenholm V, Reddy SK, Bouraoui H, Morrill J, Kulcinskaja E, Bahr CM, Aurelius O, Rogers T, Xiao Y, Logan DT, Martens EC, Koropatkin NM, Stalbrand H. 2017. Galactomannan catabolism conferred by a polysaccharide utilization locus of *Bacteroides ovatus*: enzyme synergy and crystal structure of a beta-mannanase. *J Biol Chem* 292:229–243. <https://doi.org/10.1074/jbc.M116.746438>.
- Temple MJ, Cuskin F, Basle A, Hickey N, Speciale G, Williams SJ, Gilbert HJ, Lowe EC. 2017. A Bacteroidetes locus dedicated to fungal 1,6-beta-glucan degradation: unique substrate conformation drives specificity of the key endo-1,6-beta-glucanase. *J Biol Chem* 292:10639–10650. <https://doi.org/10.1074/jbc.M117.787606>.
- Ndeh D, Rogowski A, Cartmell A, Luis AS, Baslé A, Gray J, Venditto I, Briggs J, Zhang X, Labourel A, Terrapon N, Buffetto F, Nepogodiev S, Xiao Y, Field RA, Zhu Y, O'Neil MA, Urbanowicz BR, York WS, Davies GJ, Abbott DW, Ralet M-C, Martens EC, Henrissat B, Gilbert HJ. 2017. Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature* 544:65–70. <https://doi.org/10.1038/nature21725>.
- Cartmell A, Muñoz-Munoz J, Briggs JA, Ndeh DA, Lowe EC, Basle A, Terrapon N, Stott K, Heunis T, Gray J, Yu L, Dupree P, Fernandes PZ, Shah S, Williams SJ, Labourel A, Trost M, Henrissat B, Gilbert H. 2018. A surface endogalactanase in *Bacteroides thetaiotaomicron* confers keystone status for arabinogalactan degradation. *Nat Microbiol* 3:1314–1326. <https://doi.org/10.1038/s41564-018-0258-8>.
- Luis AS, Briggs J, Zhang X, Farnell B, Ndeh D, Labourel A, Basle A, Cartmell A, Terrapon N, Stott K, Lowe EC, McLean R, Shearer K, Schuckel J, Venditto I, Ralet MC, Henrissat B, Martens EC, Mosimann SC, Abbott DW, Gilbert HJ. 2018. Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic Bacteroides. *Nat Microbiol* 3:210–219. <https://doi.org/10.1038/s41564-017-0079-1>.
- Tamura K, Foley MH, Gardill BR, Dejean G, Schnizlein M, Bahr CME,

- Creagh A, van Petegem F, Koropatkin NM, Brumer H. 2019. Surface glycan-binding proteins are essential for cereal beta-glucan utilization by the human gut symbiont *Bacteroides ovatus*. *Cell Mol Life Sci* 76: 4319–4340. <https://doi.org/10.1007/s00018-019-03115-3>.
27. Brillite J, Urbanowicz PA, Luis AS, Basle A, Paterson N, Rebello O, Hendel J, Ndeh DA, Lowe EC, Martens EC, Spencer DIR, Bolam DN, Crouch LI. 2019. Complex N-glycan breakdown by gut *Bacteroides* involves an extensive enzymatic apparatus encoded by multiple co-regulated genetic loci. *Nat Microbiol* doi:10.1038/s41564-019-0466-x. <https://doi.org/10.1038/s41564-019-0466-x>.
 28. Stone BA, Clarke AE. 1992. Chemistry and biology of (1-3)-beta-glucans. LaTrobe University Press, Carlton, Australia.
 29. O'Sullivan L, Murphy B, McLoughlin P, Duggan P, Lawlor PG, Hughes H, Gardiner GE. 2010. Prebiotics from marine macroalgae for human and animal health applications. *Mar Drugs* 8:2038–2064. <https://doi.org/10.3390/md8072038>.
 30. Salyers AA, Palmer JK, Wilkins TD. 1977. Laminarinase (beta-glucanase) activity in *Bacteroides* from the human colon. *Appl Environ Microbiol* 33:1118–1124. <https://doi.org/10.1128/AEM.33.5.1118-1124.1977>.
 31. Sheridan PO, Martin JC, Lawley TD, Browne HP, Harris HM, Bernalier-Donadille A, Duncan SH, O'Toole PW, Scott KP, Flint HJ. 2016. Polysaccharide utilization loci and nutritional specialization in a dominant group of butyrate-producing human colonic Firmicutes. *Microb Genom* 2:e000043. <https://doi.org/10.1099/mgen.0.000043>.
 32. Zhao J, Cheung PC. 2011. Fermentation of beta-glucans derived from different sources by bifidobacteria: evaluation of their bifidogenic effect. *J Agric Food Chem* 59:5986–5992. <https://doi.org/10.1021/jf200621y>.
 33. Cloetens L, Ulmuis M, Johansson-Persson A, Akesson B, Onning G. 2012. Role of dietary beta-glucans in the prevention of the metabolic syndrome. *Nutr Rev* 70:444–458. <https://doi.org/10.1111/j.1753-4887.2012.00494.x>.
 34. Driscoll M, Hansen R, Ding C, Cramer DE, Yan J. 2009. Therapeutic potential of various beta-glucan sources in conjunction with anti-tumor monoclonal antibody in cancer therapy. *Cancer Biol Ther* 8:218–225. <https://doi.org/10.4161/cbt.8.3.7337>.
 35. El Khoury D, Cuda C, Luhovyy BL, Anderson GH. 2012. Beta glucan: health benefits in obesity and metabolic syndrome. *J Nutr Metab* 2012: 851362. <https://doi.org/10.1155/2012/851362>.
 36. Stier H, Ebbeskotte V, Gruenwald J. 2014. Immune-modulatory effects of dietary yeast beta-1,3/1,6-D-glucan. *Nutr J* 13:38. <https://doi.org/10.1186/1475-2891-13-38>.
 37. Martens EC, Lowe EC, Chiang H, Pudlo NA, Wu M, McNulty NP, Abbott DW, Henrissat B, Gilbert HJ, Bolam DN, Gordon JL. 2011. Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. *PLoS Biol* 9:e1001221-16. <https://doi.org/10.1371/journal.pbio.1001221>.
 38. Viborg AH, Terrapon N, Lombard V, Michel G, Czjzek M, Henrissat B, Brumer H. 2019. A subfamily roadmap of the evolutionarily diverse glycoside hydrolase family 16 (GH16). *J Biol Chem* 294:15973–15986. <https://doi.org/10.1074/jbc.RA119.010619>.
 39. CAZypedia Consortium. 2018. Ten years of *CAZypedia*: a living encyclopedia of carbohydrate-active enzymes. *Glycobiology* 28:3–8. <https://doi.org/10.1093/glycob/cwx089>.
 40. Helbert W, Poulet L, Drouillard S, Mathieu S, Loiodice M, Couturier M, Lombard V, Terrapon N, Turchetto J, Vincentelli R, Henrissat B. 2019. Discovery of novel carbohydrate-active enzymes through the rational exploration of the protein sequences space. *Proc Natl Acad Sci U S A* 116:6063–6068. <https://doi.org/10.1073/pnas.1815791116>.
 41. McNulty NP, Wu M, Erickson AR, Pan CL, Erickson BK, Martens EC, Pudlo NA, Muegge BD, Henrissat B, Hettich RL, Gordon JL. 2013. Effects of diet on resource utilization by a model human gut microbiota containing *Bacteroides cellulosilyticus* WH2, a symbiont with an extensive glycomiome. *PLoS Biol* 11:e1001637. <https://doi.org/10.1371/journal.pbio.1001637>.
 42. Menshova RV, Ermakova SP, Anastyuk SD, Isakov VV, Dubrovskaya YV, Kusaykin MI, Um BH, Zvyagintseva TN. 2014. Structure, enzymatic transformation and anticancer activity of branched high molecular weight laminaran from brown alga *Eisenia bicyclis*. *Carbohydr Polym* 99: 101–109. <https://doi.org/10.1016/j.carbpol.2013.08.037>.
 43. Evans DF, Pye G, Bramley R, Clark AG, Dyson TJ, Hardcastle JD. 1988. Measurement of gastrointestinal pH profiles in normal ambulant human subjects. *Gut* 29:1035–1041. <https://doi.org/10.1136/gut.29.8.1035>.
 44. Planas A. 2000. Bacterial 1,3-1,4-beta-glucanases: structure, function and protein engineering. *Biochim Biophys Acta* 1543:361–382. [https://doi.org/10.1016/s0167-4838\(00\)00231-4](https://doi.org/10.1016/s0167-4838(00)00231-4).
 45. Labourel A, Jam M, Legentil L, Sylla B, Hehemann JH, Ferrieres V, Czjzek M, Michel G. 2015. Structural and biochemical characterization of the laminarinase ZgLamC(GH16) from *Zobellia galactanivorans* suggests preferred recognition of branched laminarin. *Acta Crystallogr D Biol Crystallogr* 71:173–184. <https://doi.org/10.1107/S139900471402450X>.
 46. Tauzin AS, Kwiatkowski KJ, Orlovsky NI, Smith CJ, Creagh AL, Haynes CA, Wawrzak Z, Brumer H, Koropatkin NM. 2016. Molecular dissection of xyloglucan recognition in a prominent human gut symbiont. *mBio* 7:e02134-15. <https://doi.org/10.1128/mBio.02134-15>.
 47. Wojtkowiak A, Witek K, Hennig J, Jaskolski M. 2013. Structures of an active-site mutant of a plant 1,3-beta-glucanase in complex with oligosaccharide products of hydrolysis. *Acta Crystallogr D Biol Crystallogr* 69:52–62. <https://doi.org/10.1107/S0907444912042175>.
 48. Manners DJ, Masson AJ, Patterson JC. 1973. Structure of a beta-(1-3)-D-glucan from yeast-cell walls. *Biochem J* 135:19–30. <https://doi.org/10.1042/bj1350019>.
 49. Glenwright AJ, Pothula KR, Bhamidimarri SP, Chorev DS, Basle A, Firbank SJ, Zheng HJ, Robinson CV, Winterhalter M, Kleinekathofer U, Bolam DN, van den Berg B. 2017. Structural basis for nutrient acquisition by dominant members of the human gut microbiota. *Nature* 541:407–411. <https://doi.org/10.1038/nature20828>.
 50. Mystkowska AA, Robb C, Vidal-Melgosa S, Vanni C, Fernandez-Guerra A, Höhne M, Hehemann J-H. 2018. Molecular recognition of the beta-glucans laminarin and pustulan by a SusD-like glycan-binding protein of a marine *Bacteroidetes*. *FEBS J* 285:4465–4481. <https://doi.org/10.1111/febs.14674>.
 51. Smits SA, Leach J, Sonnenburg ED, Gonzalez CG, Lichtman JS, Reid G, Knight R, Manjurano A, Changalucha J, Elias JE, Dominguez-Bello MG, Sonnenburg JL. 2017. Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science* 357:802–806. <https://doi.org/10.1126/science.aan4834>.
 52. De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, Massart S, Collini S, Pieraccini G, Lionetti P. 2010. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci U S A* 107:14691–14696. <https://doi.org/10.1073/pnas.1005963107>.
 53. Becker S, Scheffel A, Polz MF, Hehemann JH. 2017. Accurate quantification of laminarin in marine organic matter with enzymes from marine microbes. *Appl Environ Microbiol* 83:14. <https://doi.org/10.1128/AEM.03389-16>.
 54. Kabisch A, Otto A, König S, Becher D, Albrecht D, Schuler M, Teeling H, Amann RL, Schweder T. 2014. Functional characterization of polysaccharide utilization loci in the marine *Bacteroidetes* “*Gramella forsetii*” KT0803. *ISME J* 8:1492–1502. <https://doi.org/10.1038/ismej.2014.4>.
 55. Unfried F, Becker S, Robb CS, Hehemann JH, Markert S, Heiden SE, Hinzke T, Becher D, Reintjes G, Krüger K, Avci B, Kappelmann L, Hahnke RL, Fischer T, Harder J, Teeling H, Fuchs B, Barbeyron T, Amann RL, Schweder T. 2018. Adaptive mechanisms that provide competitive advantages to marine bacteroidetes during microalgal blooms. *ISME J* 12:2894–2906. <https://doi.org/10.1038/s41396-018-0243-5>.
 56. Panshin I, Huang S, Meier-Kolthoff JP, Tindall BJ, Rohde M, Verborg S, Lapidus A, Han J, Trong S, Haynes M, Reddy TBK, Huntemann M, Pati A, Ivanova NN, Mavromatis K, Markowitz V, Woyke T, Goker M, Klenk HP, Kyrpides NC, Hahnke RL. 2016. Comparing polysaccharide decomposition between the type strains *Gramella echinicola* KMM 6050(T) (DSM 19838(T)) and *Gramella portivictoriae* UST040801-001(T) (DSM 23547(T)), and emended description of *Gramella echinicola* Nedashkovskaya et al. 2005 emend. Shahina et al. 2014 and *Gramella portivictoriae* Lau et al. 2005. *Stand in Genomic Sci* 11:16. <https://doi.org/10.1186/s40793-016-0163-9>.
 57. Hehemann JH, Correc G, Barbeyron T, Helbert W, Czjzek M, Michel G. 2010. Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* 464:908–912. <https://doi.org/10.1038/nature08937>.
 58. Badur AH, Ammar EM, Yalamanchili G, Hehemann JH, Rao CV. 2020. Characterization of the GH16 and GH17 laminarinases from *Vibrio breoganii* T10. *Appl Microbiol Biotechnol* 104:161–171. <https://doi.org/10.1007/s00253-019-10243-0>.
 59. Zmora N, Suez J, Elinav E. 2019. You are what you eat: diet, health and the gut microbiota. *Nat Rev Gastroenterol Hepatol* 16:35–56. <https://doi.org/10.1038/s41575-018-0061-2>.
 60. Lazaridou A, Biliaderis CG. 2007. Molecular aspects of cereal beta-glucan functionality: physical properties, technological applications and physi-

- ological effects. *J Cereal Sci* 46:101–118. <https://doi.org/10.1016/j.jcs.2007.05.003>.
61. Read SM, Currie G, Bacic A. 1996. Analysis of the structural heterogeneity of laminarin by electrospray-ionisation-mass spectrometry. *Carbohydr Res* 281:187–201. [https://doi.org/10.1016/0008-6215\(95\)00350-9](https://doi.org/10.1016/0008-6215(95)00350-9).
 62. Abdek-Akher M, Hamilton JK, Smith F. 1951. The reduction of sugars with sodium borohydride. *J Am Chem Soc* 73:4691–4692. <https://doi.org/10.1021/ja01154a061>.
 63. Viladot JL, Moreau V, Planas A, Driguez H. 1997. Transglycosylation activity of *Bacillus* 1,3-1,4-beta-D-glucan 4-glucohydrolases. Enzymic synthesis of alternate 1,3-,1,4-beta-D-glucooligosaccharides. *J Chem Soc Perkin Trans 1*:2383–2388. <https://doi.org/10.1002/chin.199750215>.
 64. Ibatullin FM, Baumann MJ, Greffe L, Brumer H. 2008. Kinetic analyses of retaining endo-(xylo)glucanases from plant and microbial sources using new chromogenic xylogluco-oligosaccharide aryl glycosides. *Biochemistry* 47:7762–7769. <https://doi.org/10.1021/bi8009168>.
 65. Chen IA, Chu K, Palaniappan K, Pillay M, Ratner A, Huang J, Huntemann M, Varghese N, White JR, Seshadri R, Smirnova T, Kirton E, Jungbluth SP, Woyke T, Eloe-Fadrosh EA, Ivanova NN, Kyrpidis NC. 2019. IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res* 47:D666–D677. <https://doi.org/10.1093/nar/gky901>.
 66. Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, Krogh A. 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci* 12:1652–1662. <https://doi.org/10.1110/ps.0303703>.
 67. Nielsen H. 2017. Predicting secretory proteins with SignalP. *Methods Mol Biol* 1611:59–73. https://doi.org/10.1007/978-1-4939-7015-5_6.
 68. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ, Brinkman F. 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26:1608–1615. <https://doi.org/10.1093/bioinformatics/btq249>.
 69. Paetzel M, Karla A, Strynadka NCJ, Dalbey RE. 2002. Signal peptidases. *Chem Rev* 102:4549–4579. <https://doi.org/10.1021/cr010166y>.
 70. Eschenfeldt WH, Makowska-Grzyska M, Stols L, Donnelly MI, Jedrzejczak R, Joachimiak A. 2013. New LIC vectors for production of proteins from genes containing rare codons. *J Struct Funct Genomics* 14:135–144. <https://doi.org/10.1007/s10969-013-9163-9>.
 71. Hutchison CA, Phillips S, Edgell MH, Gillam S, Jahnke P, Smith M. 1978. Mutagenesis at a specific position in a DNA-sequence. *J Biol Chem* 253:6551–6560.
 72. Doublé S. 2007. Production of selenomethionyl proteins in prokaryotic and eukaryotic expression systems. *Methods Mol Biol* 363:91–108. https://doi.org/10.1007/978-1-59745-209-0_5.
 73. McGregor N, Arnal G, Brumer H. 2017. Quantitative kinetic characterization of glycoside hydrolases using high-performance anion-exchange chromatography (HPAEC). *Methods Mol Biol* 1588:15–25. https://doi.org/10.1007/978-1-4939-6899-2_2.
 74. Arnal G, Attia MA, Asohan J, Brumer H. 2017. A low-volume, parallel copper-bicinchoninic acid (BCA) assay for glycoside hydrolases. *Methods Mol Biol* 1588:3–14. https://doi.org/10.1007/978-1-4939-6899-2_1.
 75. Kabsch W. 2010. XDS. *Acta Crystallogr D Biol Crystallogr* 66:125–132. <https://doi.org/10.1107/S0907444909047337>.
 76. Evans P. 2006. Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* 62:72–82. <https://doi.org/10.1107/S0907444905036693>.
 77. Evans PR, Murshudov GN. 2013. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* 69:1204–1214. <https://doi.org/10.1107/S0907444913000061>.
 78. Terwilliger TC, Adams PD, Read RJ, McCoy AJ, Moriarty NW, Grosse-Kunstleve RW, Afonine PV, Zwart PH, Hung LW. 2009. Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* 65:582–601. <https://doi.org/10.1107/S0907444909012098>.
 79. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221. <https://doi.org/10.1107/S0907444909052925>.
 80. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Crystallogr* 40: 658–674. <https://doi.org/10.1107/S0021889807021206>.
 81. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. 2012. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68:352–367. <https://doi.org/10.1107/S0907444912001308>.
 82. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486–501. <https://doi.org/10.1107/S0907444910007493>.
 83. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA. 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67:355–367. <https://doi.org/10.1107/S0907444911001314>.
 84. Potterton L, Agirre J, Ballard C, Cowtan K, Dodson E, Evans PR, Jenkins HT, Keegan R, Krissinel E, Stevens K, Lebedev A, McNicholas SJ, Nicholls RA, Noble M, Pannu NS, Roth C, Sheldrick G, Skubak P, Turkenburg J, Uski V, von Delft F, Waterman D, Wilson K, Winn M, Wojdyr M. 2018. CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr D Struct Biol* 74:68–84. <https://doi.org/10.1107/S2059798317016035>.
 85. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, Snoeyink J, Richardson JS, Richardson DC. 2007. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35:W375–W383. <https://doi.org/10.1093/nar/gkm216>.
 86. Holm L, Laakso LM. 2016. Dali server update. *Nucleic Acids Res* 44: W351–W355. <https://doi.org/10.1093/nar/gkw357>.
 87. Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, Liang S, Zhang W, Guan Y, Shen D, Peng Y, Zhang D, Jie Z, Wu W, Qin Y, Xue W, Li J, Han L, Lu D, Wu P, Dai Y, Sun X, Li Z, Tang A, Zhong S, Li X, Chen W, Xu R, Wang M, Feng Q, Gong M, Yu J, Zhang Y, Zhang M, Hansen T, Sanchez G, Raes J, Falony G, Okuda S, Almeida M, LeChatelier E, Renault P, Pons N, Batto J-M, Zhang Z, Chen H, Yang R, Zheng W, Li S, Yang H, Wang J, Ehrlich SD, Nielsen R, Pedersen O, Kristiansen K, Wang J. 2012. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490: 55–60. <https://doi.org/10.1038/nature11450>.
 88. Yu J, Feng Q, Wong SH, Zhang D, Liang QY, Qin Y, Tang L, Zhao H, Stenvang J, Li Y, Wang X, Xu X, Chen N, Wu WKK, Al-Aama J, Nielsen HJ, Kiekerich P, Jensen BAH, Yau TO, Lan Z, Jia H, Li J, Xiao L, Lam TYT, Ng SC, Cheng AS-L, Wong VW-S, Chan FKL, Xu X, Yang H, Madsen L, Datz C, Tilg H, Wang J, Brünner N, Kristiansen K, Arumugam M, Sung JJ-Y, Wang J. 2017. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut* 66:70–78. <https://doi.org/10.1136/gutjnl-2015-309800>.
 89. Liu RX, Hong J, Xu XQ, Feng Q, Zhang DY, Gu YY, Shi J, Zhao SQ, Liu W, Wang XK, Xia HH, Liu ZP, Cui B, Liang PW, Xi LQ, Jin JB, Ying XY, Wang XL, Zhao XJ, Li WY, Jia HJ, Lan Z, Li FY, Wang R, Sun YK, Yang ML, Shen YX, Jie ZY, Li JH, Chen XM, Zhong HZ, Xie HL, Zhang YF, Gu WQ, Deng XX, Shen BY, Xu X, Yang HM, Xu GW, Bi YF, Lai SH, Wang J, Qi L, Madsen L, Wang JQ, Ning G, Kristiansen K, Wang WQ. 2017. Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. *Nat Med* 23:859–868. <https://doi.org/10.1038/nm.4358>.
 90. Gu YY, Wang XK, Li JH, Zhang YF, Zhong HZ, Liu RX, Zhang DY, Feng Q, Xie XY, Hong J, Ren HH, Liu W, Ma J, Su Q, Zhang HM, Yang JL, Wang XL, Zhao XJ, Gu WQ, Bi YF, Peng YD, Xu XQ, Xia HH, Li F, Xu X, Yang HM, Xu GW, Madsen L, Kristiansen K, Ning G, Wang WQ. 2017. Analyses of gut microbiota and plasma bile acids enable stratification of patients for anti-diabetic treatment. *Nat Commun* 8:12. <https://doi.org/10.1038/s41467-017-01682-2>.
 91. He Q, Gao Y, Jie ZY, Yu XL, Laursen JM, Xiao L, Li Y, Li LL, Zhang FM, Feng Q, Li XP, Yu JH, Liu C, Lan P, Yan T, Liu X, Xu X, Yang HM, Wang J, Madsen L, Brix S, Wang JP, Kristiansen K, Jia HJ. 2017. Two distinct metacommunities characterize the gut microbiota in Crohn's disease patients. *GigaScience* 6:1–11. <https://doi.org/10.1093/gigascience/gix050>.
 92. Zhang X, Zhang D, Jia H, Feng Q, Wang D, Liang D, Wu X, Li J, Tang L, Li Y, Lan Z, Chen B, Li Y, Zhong H, Xie H, Jie Z, Chen W, Tang S, Xu X, Wang X, Cai X, Liu S, Xia Y, Li J, Qiao X, Al-Aama JY, Chen H, Wang L, Wu Q-J, Zhang F, Zheng W, Li Y, Zhang M, Luo G, Xue W, Xiao L, Li J, Chen W, Xu X, Yin Y, Yang H, Wang J, Kristiansen K, Liu L, Li T, Huang Q, Li Y, Wang J. 2015. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. *Nat Med* 21: 895–905. <https://doi.org/10.1038/nm.3914>.
 93. Nishijima S, Suda W, Oshima K, Kim SW, Hirose Y, Morita H, Hattori M. 2016. The gut microbiome of healthy Japanese and its microbial and functional uniqueness. *DNA Res* 23:125–133. <https://doi.org/10.1093/dnares/dsw002>.
 94. Lloyd-Price J, Mahurkar A, Ahnavard GR, Rabtree JC, Rvis JO, Hall ABR, Rady AB, Reasy HHC, McCracken C, Giglio MG, McDonald D, Franzosa EA,

- Knight R, White O, Huttenhower C. 2017. Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* 550:61–66. <https://doi.org/10.1038/nature23889>.
95. Le Chatelier E, MetaHIT Consortium, Nielsen T, Qin JJ, Prifti E, Hildebrand F, Falony G, Almeida M, Arumugam M, Batto JM, Kennedy S, Leonard P, Li JH, Burgdorf K, Grarup N, Jorgensen T, Brandslund I, Nielsen HB, Juncker AS, Bertalan M, Levenez F, Pons N, Rasmussen S, Sunagawa S, Tap J, Tims S, Zoetendal EG, Brunak S, Clement K, Dore J, Kleerebezem M, Kristiansen K, Renault P, Sicheritz-Ponten T, de Vos WM, Zucker JD, Raes J, Hansen T, Bork P, Wang J, Ehrlich SD, Pedersen O, Meta H. 2013. Richness of human gut microbiome correlates with metabolic markers. *Nature* 500:541–546. <https://doi.org/10.1038/nature12506>.
96. Qin J, MetaHIT Consortium, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto J-M, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Li S, Jian M, Zhou Y, Li Y, Zhang X, Li S, Qin N, Yang H, Wang J, Brunak S, Doré J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, Bork P, Ehrlich SD, Wang J. 2010. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464:59–65. <https://doi.org/10.1038/nature08821>.
97. Contevelle LC, Oliveira-Ferreira J, Vicente A. 2019. Gut microbiome biomarkers and functional diversity within an Amazonian semi-nomadic hunter gatherer group. *Front Microbiol* 10:1743. <https://doi.org/10.3389/fmicb.2019.01743>.
98. Boratyn GM, Thierry-Mieg J, Thierry-Mieg D, Busby B, Madden TL. 2019. Magic-BLAST, an accurate RNA-seq aligner for long and short reads. *BMC Bioinformatics* 20:19. <https://doi.org/10.1186/s12859-019-2996-x>.
99. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842. <https://doi.org/10.1093/bioinformatics/btq033>.