Gut Microbes 2:2, 86-90; March/April 2011; © 2011 Landes Bioscience

Mechanistic insight into polysaccharide use within the intestinal microbiota

David N. Bolam[1] and Justin L. Sonnenburg[2]

[1]Institute for Cell and Molecular Biosciences; Newcastle University; Medical School; Newcastle-Upon-Tyne, UK; [2]Department of Microbiology and Immunology; Stanford University School of Medicine; Stanford, CA USA

It is becoming increasingly clear that diet is one of the major factors that drives the function and composition of the intestinal microbiota. The diet of humans is highly diverse when considering different populations or even a single individual over a relatively short period of time. However, we are just beginning to understand the mechanisms that connect dietary change to intestinal microbiota dynamics. The community of microbes within our distal digestive tract influences numerous aspects of our biology, and aberrant shifts in its composition appear to be associated with several diseases. It is, therefore, necessary to understand how our behaviour and environmental factors, such as changes in diet, impact our intestinal residents. Here we look to recent work to highlight some of the major questions on the horizon for understanding the key role that the Bacteroidetes play in the commerce of dietary polysaccharides within the intestine.

Key words: Bacteroidetes, intestinal microbiota, dietary polysaccharides, fructan, polysaccharide utilization-locus, germ free mice, hybrid two-component system

Submitted: 12/13/10

Revised: 02/16/11

Accepted: 02/21/11

DOI: 10.4161/gmic.2.2.15232

Correspondence to:
David N. Bolam; Email: d.n.bolam@ncl.ac.ukl
and Justin L. Sonnenburg; Email: jsonnenburg@stanford.edu

Addendum to: Sonnenburg ED, Zheng H, Joglekar P, Higginbottom SK, Firbank SJ, Bolam DN, et al. Specificity of polysaccharide use in intestinal Bacteroides species determines diet-induced microbiota alterations. Cell 2010; 141:1241–52; PMID: 20603004; DOI: 10.1016/j.cell.2010.05.005.

We have recently elucidated a pathway present in gut-resident Bacteroides species that enables the use of fructans, a class of dietary plant and microbial polysaccharides.[1] Using *Bacteroides thetaiotaomicron* as a model, we performed a genetic and biochemical dissection of a polysaccharide utilization locus (PUL) that encodes the fructan utilization pathway. PULs encode machinery that performs a remarkable feat: converting extracellular polysaccharides into glycolytic substrates in a step-wise process of cell-surface binding, degradation and import. These loci are built around orthologs of *susD* and *susC*, genes within *B. thetaiotaomicron*'s starch utilization system (Sus) that encode outer-membrane polysaccharide binding and import proteins, respectively (**Fig. 1**).[2,3] Our work builds on a rapidly expanding field that is dedicated to understanding the role that PULs play in Bacteroidetes adaptation to the dynamic nutrient environment of the intestine.[4-8]

PULs have been widely expanded within the Bacteroidetes.[4] The duplication and diversification of PUL function is one obvious mode of evolution that enables intestinal bacterial strains to exploit new niches and substrates within the gut. PULs also appear to be shared via lateral transfer, which results in the acquisition of polysaccharide use in the recipient strain.[9] A mode of PUL diversification became apparent through our study: altering gene content and specificity of existing (i.e., non-duplicated) PULs to access new substrates. We examined various forms of the fructan-utilization PUL that are conserved across six Bacteroides species (those with complete genome sequences at the time of the study). However, despite similarity in PUL gene content and order with other Bacteroides species, *B. thetaiotaomicron* has a distinctive fructan-utilization functionality. It grows very well on levan (a β2-6-linked fructan), but very poorly on inulin (a β2-1 fructan), which is in stark contrast to all but one of the other Bacteroides species that efficiently use inulin but not levan (the exception being *B. vulgatus*, which can use neither inulin or levan). We discovered that the

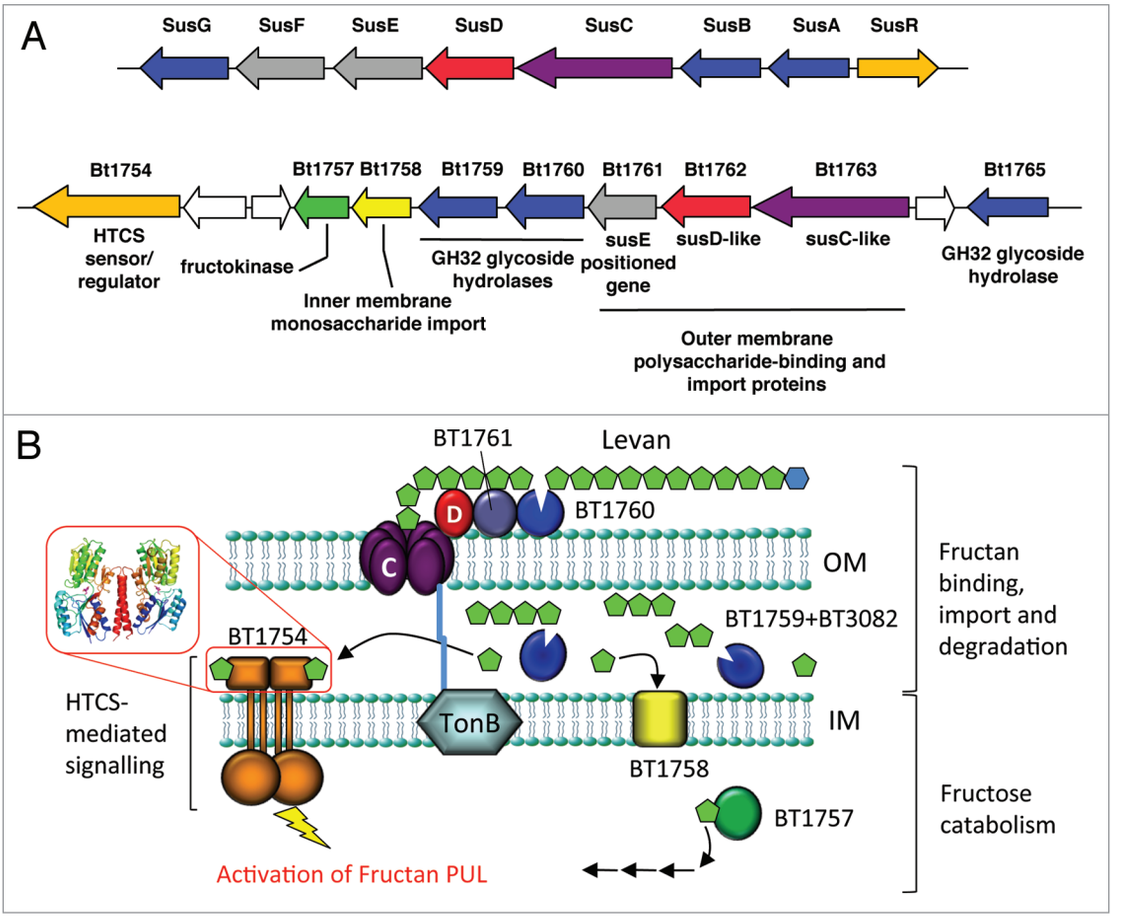

Figure 1. Organization of the starch and levan PULs in *Bacteroides thetaiotaomicron* and model of levan utilization. **A.** Genomic organization of *B. thetaiotaomicron*'s prototypic Sus locus (top) and levan utilization locus (bottom). Genes of similar function are coded by color; intervening unrelated genes are white. **B.** Based on our studies and previous genetic and biochemical studies focused on other PULs of *B. thetaiotaomicron*,[1-5,7,8] we have developed a working model of β2-6 fructan utilization in *B. thetaiotaomicron*. In this model, levan is bound by the SusD homolog (D), BT1762 and the susE-positioned gene product, BT1761, on the surface of the bacterium. The polysaccharide is then cleaved into oligosaccharides by the extracellular endo-acting levanase, BT1760. Oligosaccharides are actively imported by the SusC homolog (C), BT1763 (a TonB dependent porin). In the periplasm, exo-acting glycoside hydrolases, BT1759 and BT3082, liberate monosaccharide fructose (green pentagons) from the oligosaccharides. Free fructose binds to the periplasmic sensor domain of the homodimeric HTCS, BT1754 (inset red box shows structure of BT1754 periplasmic domain homodimer bound to fructose), activating the regulatory protein and resulting in upregulation of the PUL. The inner membrane monosaccharide importer, BT1758, transports periplasmic fructose into the cell where it is shunted into the glycolytic pathway upon phosphorylation by the PUL-encoded fructokinase, BT1757. Some transport across the inner membrane of the disaccharide products of fructan degradation (sucrose and levanbiose) also occurs and these are broken down to their constituent monosaccharides by the putative intracellular fructosidase, BT1765 (not shown). Prior to activation, low levels of all components of the PUL are expressed constitutively such that *B. thetaiotaomicron* is always in a 'prepared' state.

key to *B. thetaiotaomicron*'s ability to use levan is the alteration of genes within its PUL. A cell surface endo-acting β2-6-specific glycoside hydrolase-encoding gene (*BT1760*), which is absent in the other species, has been inserted into *B. thetaiotaomicron*'s fructan PUL (**Fig. 1**). In addition to acquiring this important enzyme, *B. thetaiotaomicron* appears to have modified the specificity of its polysaccharide binding *susD*-like gene product within the locus, *BT1762*. Unlike the β2-1 specificity that is apparent in the SusD-like proteins of the inulin-using species tested to date (Bolam D, unpublished), BT1762 has a strict β2-6 fructan specificity, despite its apparent orthology to the β2-1-binding SusD-like proteins of other species.

The cell surface components of *B. thetaiotaomicron*'s fructan utilization system serve two important roles for fructan use. The first is to bind and degrade β2-6 polymers and translocate the resulting oligosaccharides into the periplasm where non-specific exo-acting fructosidases stand ready to generate fructose monomers. The second role is to serve as a fructan-screening-system that ensures the free fructose generated within the periplasm is derived from β2-6 linkages. This screening function is important to compensate for the lack of specificity toward linkage exhibited by the locus-associated signalling sensor (BT1754), which activates the locus when the appropriate substrate is present. *BT1754* encodes a hybrid two-component system (HTCS) that traverses the inner membrane. Our structural and biophysical studies demonstrate that the periplasmic sensor domain of BT1754 binds to monomeric fructose,

which, therefore, acts as a proxy signal for the presence of levan within the intestinal environment (**Fig. 1**). The lack of specificity displayed by the BT1754 sensor for levan-derived oligosaccharides may seem unexpected due to the strong linkage preference of the adjacent locus that it controls (β2-6>>β2-1). However, the selectivity of the cell surface binding and degradation machinery for β2-6 linkages ensures the periplasmic-located sensor will encounter fructose that is primarily derived from levan. Precisely why the system has evolved such a defined specificity is unclear as it would seem advantageous for *B. thetaiotaomicron* to be able to target both types of fructan. Whatever the reason, specificity for one type of fructan appears to be the norm as none of the six Bacteroides species we have assayed so far grow equally well on inulin and levan—one linkage is always strongly preferred. These data suggest that promiscuity in the case of fructans is disadvantageous. The localization of specificity elements to the cell surface enables a transcriptional response that, despite the sensor's inability to discriminate linkage, is linkage-specific due to the sequestration of the sensor within the periplasm. One question that arises is whether PUL-regulating sensors that are blind to linkage permit more evolutionary plasticity of PUL gene content and specificity. As genome data accumulates, it will be interesting to determine if a less-specific signal corresponds to increased diversity, and perhaps more rapid evolution, of the accompanying PUL across the genus and phylum.

The work summarized above raises many issues that remain to be addressed regarding the mechanisms of polysaccharide utilization by Sus-like systems, not only related to fructan use, but also to the vast array of plant- and host-derived polysaccharides utilized by gut Bacteroidetes species through their PUL-encoded systems.[4,7]

Enzymes

When trying to construct an accurate model of PUL-mediated glycan harvest, determining the cellular location of the PUL encoded glycosidic-linkage breaking enzymes becomes as important as understanding the enzymes' specificity (**Fig. 1**). Analysis of the predicted cellular locations of the glycoside hydrolases (GH) from the four efficient inulin-using Bacteroides fructan PULs studied suggest that all four Bacteroides species have at least one enzyme that is likely to be an outer membrane lipoprotein, and thus may be cell surface located. Studies with the only two Sus-like systems characterized to date (the *B. thetaiotaomicron* starch and levan systems) have shown that the surface located enzyme from each of these cleaves its polymeric substrate in an endo-like fashion, producing oligosaccharides that are transported into the cell via the SusC porin.[1,4,8] In *B. caccae* and *B. ovatus* the putative surface located enzymes are GH91s, a family known to possess endo-inulinase activity.[10] However, in *B. fragilis* and *B. uniformis* the putative surface enzyme is a GH32 that is closely related to *B. thetaiotaomicron* enzymes that we have shown to be non-linkage specific exo-fructosidases (see ref. 1, Fig. S6; BF3177 and BACUNI_01159), which leads to the question: are these enzymes endo-acting inulinases or are they actually producing fructose extracellularly? The latter possibility seems unlikely as the production of a monosaccharide would negate the requirement for any periplasmic glycoside hydrolases and a SusC ortholog for transport of oligosaccharides.

B. vulgatus is the only sequenced Bacteroides to date that lacks *susC*- or *susD*-like genes within its fructan PUL and cannot use inulin or levan. However, *B. vulgatus* is able to grow well on short-chain β2-1 fructo-oligosaccharides (FOS; Sonnenburg J, unpublished). Intriguingly, *B. vulgatus* contains only a single fructan-degrading enzyme (a GH32; BVU_1663), which is predicted to be an outer membrane located lipoprotein and may therefore be surface located. Comparison of the sequence of BVU_1663 with the *B. thetaiotaomicron* GH32s reveals it is most closely related to BT1765, an exo-fructosidase enzyme that displays a strong preference for short chain fructo-oligosaccharides over polymeric fructans. If BVU_1633 displays a similar substrate size preference to BT1765, this may explain the ability of *B. vulgatus* to utilize FOS, but not inulin. The enzyme's predicted location would suggest that the production of fructose occurs extracellularly. Thus, the likely mechanism of FOS utilization by this bacterium provides a ready explanation for its lack of genes encoding SusC and SusD homologs, as it does not require such an elaborate system to import fructose across the outer membrane. It remains to be determined if *B. vulgatus* has specialized in the use of FOS present in the diet, cross-feeds on short fructans generated extracellularly by other long-chain fructan-users, or has lost the ability to use long-chain fructans for another reason.

The data described for *B. vulgatus* supports the view that SusC- and SusD-like proteins are required for utilization of polymeric fructans. However, it is currently unclear whether an extracellular endo-acting enzyme is also an absolute requirement. With respect to the efficient polymeric inulin users, an alternative possibility is that some currently uncharacterized Bacteroides species lack an extracellular fructan-utilizing enzyme altogether. In this case the inulin chains would be threaded through the SusC porin without being cleaved outside the cell and all degradation would occur in the periplasm. Although this is clearly not the case in *B. thetaiotaomicron*'s levan utilization system,[1] it may be possible with inulin. This polysaccharide has a relatively short average degree of polymerization of only ~25 fructose units (for Chicory inulin) and may not require external processing by the cell for transport through the outer membrane via the SusC/D system.[11] We are currently working to determine whether SusC/D systems are always paired with extracellular endo-polysaccharidases, or whether a sub-class of PULs work in the absence of extracellular oligosaccharide generation.

The apparent redundant specificity and periplasmic location of two of the *B. thetaiotaomicron* GH32 enzymes, BT3082 and BT1759, provides another unanswered question. The redundancy contrasts with the prototypic Sus that relies upon a single periplasmic glucosidase, SusB, which is capable of cleaving four different α-glucosidic linkages.[12] One possibility is that one of the enzymes is actually optimized to deal with more branched

form of fructans such as graminins, which are mixed β2-1/2-6 linked fructans that are found mainly in grass species such as wheat and barley.[13] While it is not known if *B. thetaiotaomicron* can actually utilize these highly branched substrates, it would seem an obvious rationale for the duplicity of apparently similar enzyme specificities. Indeed, the presence of four predicted fructan-processing enzymes encoded within the fructan PULs of several of the other inulin-using Bacteroides species strengthens the idea that fructan use in these strains is more complex than simply β2-1 or β2-6. Interestingly, *B. fragilis* contains only two GH32s, suggesting it specifically targets inulin and that branched oligosaccharides are precluded from entry into the periplasm by the specificity of the *B. fragilis* SusD. Activity screening with branched forms of fructan, combined with biochemical and gene knockout studies, should enable us to dissect the role that some of these apparently redundant enzymes play.

Polysaccharide Binding Proteins

Homologs of the SusC and SusD outer membrane proteins are the defining feature of a Sus-system, and genes encoding these proteins are almost always found as a pair in the Bacteroidetes genomes sequenced to date, indicating the importance of both proteins in polysaccharide utilization in these organisms. In addition, in the prototypic *B. thetaiotaomicron* starch PUL, downstream of *susD* is a gene that encodes SusE, a surface located lipoprotein, which also displays a starch binding function (**Fig. 1**). Despite the conservation of *sus*D-like genes in Bacteroidetes PULs, the importance of the SusD homolog in polysaccharide utilization may vary. SusD from the prototypic Sus has been shown to be essential for the utilization of starch and malto-oligosaccharides larger than five sugars.[5] In contrast, our data suggest that a SusD-homolog is not necessarily essential to each system. The knockout of the *sus*D (*BT1762*) in the levan PUL retains some, albeit significantly retarded, growth on levan (doubling time ~7x slower than wild-type). While the impact of this growth defect would be severe in a competitive gut environment, the retention of some growth in vitro in the absence of the SusD-homolog may be informative of how the fructan PUL differs mechanistically from the starch utilization system. Indeed, the mechanism by which the SusD homolog functions is not known. SusD is known to form a complex with SusC in the starch system.[3,14] One possibility is that the protein simply acquires and holds the polysaccharide chain in the correct position so it may 'thread' through the SusC porin and/or be degraded by cell surface endo-acting polysaccharidases. Perhaps in some systems, other surface polysaccharide binding proteins, like the *susE*-positioned gene product (e.g., BT1761 from the levan PUL, see below), play a role that is similar and partially redundant to that of the SusD-like protein, although such compensation is clearly not universal as illustrated by the essential role of SusD in starch utilization.

In addition to the possible variable function of the SusD homolog in different Sus systems, the role and conservation of the *susE*-positioned gene product remains unclear. The *susE*-positioned gene products are lipoproteins that in several cases have been shown to share the trait of polysaccharide binding, but display no sequence similarity to the prototypical protein from the *B. thetaiotaomicron* starch PUL.[1-4] While the levan PUL of *B. thetaiotaomicron* has a *susE*-positioned gene, *BT1761*, inspection of the fructan PULs of the inulin-using species reveals no such candidate gene downstream of the *susD* homolog. Closer analysis indicates that the products of genes located either upstream of the *susC* in *B. fragilis* (*BF4324*) or downstream of a GH32 gene in *B. uniformis* (*BACUNI_01157*) share ~65% identity, are putative lipoproteins and have no homology to known proteins. Thus, it is possible that despite their non-conserved location, the products of these genes play a similar role to the *susE*-positioned genes in other PULs. No such *susE*-positioned gene candidates exist in the *B. ovatus* or *B. caccae* inulin PULs, two species that share almost identical PUL structure. However, analysis of the sequence of the GH91 enzymes from both organisms indicates that all four are predicted to be outer membrane lipoproteins and therefore may all be exposed on the cell surface. One of the GH91 enzymes in each organism is ~300 amino acids longer than the adjacent GH91, leading to the possibility that the polysaccharide binding activity of a *susE*-positioned gene product has been incorporated into the cell surface enzyme as an ancillary domain, similar to the discrete carbohydrate binding module that has recently been shown to be a component of SusG, the surface located glycoside hydrolase from the *B. thetaiotaomicron* starch utilization system.[8] Structural and biochemical studies are underway to investigate this possibility.

HTCS Signal Recognition

Identification of fructose as the molecular trigger for activation of the Bacteroides fructan PUL raises the question of whether other HTCS-controlled PULs are also activated by simple monosaccharides or by more complex oligosaccharide cues. Several lines of evidence suggest the latter is true. Firstly, fructose is somewhat unusual in that it is only found in fructans or sucrose and not as a component of other complex glycans. When a Bacteroides cell in the large intestine detects fructose, the likely source is polymeric fructan (free fructose and sucrose are taken up by the host in the proximal small intestine). Unlike fructose, most other monosaccharides are components of many different complex glycans with variable linkages and additional sugars. Therefore, an oligosaccharide that contains this complex structural information is required for specific recognition of the parent glycan to enable activation of the appropriate PUL. Secondly, the ~300 aa periplasmic binding protein fold adopted by the fructose-binding sensor domain of BT1754 appears to be a unique 'invention.' Homologs of BT1754 appear to be strictly associated with fructan-use PULs in the other Bacteriodes. Most PUL-associated HTCS in sequenced Bacteroidetes contain a much larger conserved periplasmic domain of ~800 aa, which is predicted to form a double β-propeller structure.[15,16] A possible evolutionary rationale for the conservation of this domain in these HTCS is that it has the plasticity to adapt to binding a wide range of different complex glycans, thus providing Bacteroidetes species with

the broad glycan-sensing ability required for survival in the dynamic nutrient environment of the large intestine. The view that these ubiquitous β-propeller HTCS sensors recognize more complex ligands than BT1754 is supported by studies with *Prevotella bryantii*, a ruminal Bacteroidetes, which has been shown to require xylo-oligosaccharides larger than six sugars for activation of its HTCS controlled xylan PUL.[17] Further studies are underway in our lab to define the precise identity of the molecular cues recognized by Bacteroidetes PUL-associated HTCS.

Translating PUL Encoded Mechanisms into Microbiota Function in vivo

Pursuit of a detailed mechanistic understanding of PUL function will be required to fully understand and possibly predict how Bacteroidetes respond to host dietary change. In our recent study, we used defined communities of two Bacteroides species cohabitating within gnotobiotic mice to test bacterial response to dietary fructan.[1] We demonstrated that a mechanistic understanding of just a very small fraction of the ~10,000 total genes within the two-species microbiome could lead to accurate predictions of how a defined change in diet (addition of dietary inulin, in this case) impacts community composition.

The relative competitive advantage required for persistence of a species in the microbiota and the selective pressures required for persistence or alterations of PULs within a strain is still unclear. As we survey more Bacteroidetes isolates, it is evident that some are better adapted to utilize inulin than others. In the case of *B. thetaiotaomicron*, it is likely that the costs associated with poor inulin-use are off-set by the advantages gained in optimizing its fructan PUL to use β2-6-linked levan. However, it is not clear if other isolates that exhibit sub-optimal inulin use are best adapted to a substrate that deviates structurally from the inulin preparation that we use in our laboratory experiments, or are just poorly adapted to inulin. Addressing this question will be aided by increased availability of additional pure polysaccharides with well-defined structures. Poor use of a polysaccharide by a strain could result from the absence of that substrate in the hosts' diets—surely some polysaccharides have played a minimal selective role in the evolutionary history of certain strains. It is also possible that being poor at using a specific polysaccharide, even if it is abundant in the intestinal environment, is adaptive to the extent that such a strain may out-compete other strains that are completely unable to use that substrate—a "survival of the fitter" hypothesis. Whatever the cause, the retention of the ability to utilize inulin in some Bacteroides species at a level that is sub-optimal indicates that within their specific niche the trait confers a survival advantage. While it is impossible to re-run the evolutionary history of each strain, these questions can be addressed using experimental models in which community composition and diet can be tightly controlled. The merging of genomics, biochemistry, and genetics to generate hypotheses that can be tested on simplified model communities, or an intact human microbiota living within gnotobiotic mice fed defined diets, will provide a powerful pipeline to unravel these important questions.

Acknowledgments

We thank Erica Sonnenburg, Payal Joglekar and Harry Gilbert for helpful discussions and Sara Fisher for editing the manuscript. This work was funded in part by grants from the NIH-NIDDK (R01DK085025) and BBSRC (BBF0141631).

Note

Species names refer to type strains. It is expected that the attributes that have been described for the type strains will vary between other strains that share a high degree of 16S rRNA sequence identity to these type strains (i.e., qualify as the same "species").

References

1. Sonnenburg ED, Zheng H, Joglekar P, Higginbottom SK, Firbank SJ, Bolam DN, et al. Specificity of polysaccharide use in intestinal Bacteroides species determines diet-induced microbiota alterations. Cell 2010; 141:1241-52.
2. Reeves AR, Wang GR, Salyers AA. Characterization of four outer membrane proteins that play a role in utilization of starch by *Bacteroides thetaiotaomicron*. J Bacteriol 1997; 179:643-9.
3. Shipman JA, Berleman JE, Salyers AA. Characterization of four outer membrane proteins involved in binding starch to the cell surface of *Bacteroides thetaiotaomicron*. J Bacteriol 2000; 182:5365-72.
4. Martens EC, Koropatkin NM, Smith TJ, Gordon JI. Complex glycan catabolism by the human gut microbiota: The Bacteroidetes Sus-like paradigm. J Biol Chem 2009; 284:24673-7.
5. Koropatkin NM, Martens EC, Gordon JI, Smith TJ. Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. Structure 2008; 16:1105-15.
6. Flint HJ, Bayer EA, Rincon MT, Lamed R, White BA. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. Nat Rev Microbiol 2008; 6:121-31.
7. Martens EC, Chiang HC, Gordon JI. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. Cell Host Microbe 2008; 4:447-57.
8. Koropatkin NM, Smith TJ. SusG: A unique cell-membrane-associated alpha-amylase from a prominent human gut symbiont targets complex starch molecules. Structure 2010; 18:200-15.
9. Hehemann JH, Correc G, Barbeyron T, Helbert W, Czjzek M, Michel G. Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. Nature 2010; 464:908-12.
10. Jung WS, Hong CK, Lee S, Kim CS, Kim SJ, Kim SI, et al. Structural and functional insights into intramolecular fructosyl transfer by inulin fructotransferase. J Biol Chem 2007; 282:8414-23.
11. Roberfroid MB. Introducing inulin-type fructans. Br J Nutr 2005; 93:13-25.
12. Kitamura M, Okuyama M, Tanzawa F, Mori H, Kitago Y, Watanabe N, et al. Structural and functional analysis of a glycoside hydrolase family 97 enzyme from *Bacteroides thetaiotaomicron*. J Biol Chem 2008; 283:36328-37.
13. Van Den Ende W, Clerens S, Vergauwen R, Van Riet L, Van Laere A, Yoshida M, et al. Fructan 1-exohydrolases. beta-(2,1)-trimmers during graminan biosynthesis in stems of wheat? Purification, characterization, mass mapping and cloning of two fructan 1-exohydrolase isoforms. Plant Physiol 2003; 131:621-31.
14. Cho KH, Salyers AA. Biochemical analysis of interactions between outer membrane proteins that contribute to starch utilization by *Bacteroides thetaiotaomicron*. J Bacteriol 2001; 183:7224-30.
15. Menke M, Berger B, Cowen L. Markov random fields reveal an N-terminal double beta-propeller motif as part of a bacterial hybrid two-component sensor system. Proc Natl Acad Sci USA 2010; 107:4069-74.
16. Mascher T, Helmann JD, Unden G. Stimulus perception in bacterial signal-transducing histidine kinases. Microbiol Mol Biol Rev 2006; 70:910-38.
17. Miyazaki K, Hirase T, Kojima Y, Flint HJ. Medium- to large-sized xylo-oligosaccharides are responsible for xylanase induction in *Prevotella bryantii* B14. Microbiology (Reading, England) 2005; 151:4121-5.