Version of Record: https://www.sciencedirect.com/science/article/pii/S1931312821005771
Manuscript_b71921580c7464e09d08c64740d76976

Polysaccharide utilization loci in *Bacteroides* determine population fitness and community-level interactions

Jun Feng[1,2], Yili Qian[2], Zhichao Zhou[3], Sarah Ertmer[4], Eugenio I. Vivas[3,5], Freeman Lan[2], Joshua J. Hamilton[2], Federico E. Rey[3], Karthik Anantharaman[3] and Ophelia S. Venturelli[1,2,3,4,6,*]

[1]The Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI 53706, USA

[2]Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706, USA

[3]Department of Bacteriology, University of Wisconsin-Madison, Madison, WI 53706, USA

[4]Department of Chemical & Biological Engineering, University of Wisconsin-Madison, Madison, WI 53706, USA

[5]Gnotobiotic Animal Core Facility, University of Wisconsin-Madison, Madison, WI 53706, USA

[6]Lead contact

*Correspondence: venturelli@wisc.edu

SUMMARY

Polysaccharide utilization loci (PULs) are co-regulated bacterial genes that sense nutrients and enable glycan digestion. Human gut microbiome members, notably *Bacteroides,* contain numerous PULs that enable glycan utilization and shape ecological dynamics. To investigate the role of PULs on fitness and inter-species interactions, we develop a CRISPR-based genome-editing tool to study 23 PULs in *Bacteroides uniformis* (BU). BU PULs show distinct glycan-degrading functions and transcriptional coordination that enables the population to adapt upon loss of other PULs. Exploiting a BU mutant barcoding strategy, we demonstrate that *in vitro* fitness and BU colonization in the murine gut are enhanced by deletion of specific PULs and modulated by glycan availability. PULs mediate glycan-dependent interactions with butyrate producers that depend on the degradation mechanism and glycan utilization ability of the butyrate producer. Thus, PULs determine community dynamics and butyrate production and provide a selective advantage or disadvantage depending on the nutritional landscape.

INTRODUCTION

A core function of the human gut microbiome is to transform dietary and host derived polysaccharides (i.e. glycans) into metabolites that regulate our energy balance, provide colonization resistance to intestinal pathogens and maintain gut homeostasis (Desai et al., 2016; Fan and Pedersen, 2021; Koh et al., 2016; Lloyd-Price et al., 2016). *Bacteroides* harbor a broad repertoire of carbohydrate active enzymes (CAZymes) that degrade chemically diverse glycans, providing a unique metabolic function that benefits the host (Cantarel et al., 2012; El Kaoutari et al., 2013; Porter and Martens, 2017). Consequently, *Bacteroides* thrive as one of the most abundant and stable groups of organisms in the human gut (El Kaoutari et al., 2013).

In *Bacteroides*, utilization of complex glycans is frequently mediated by polysaccharide utilization loci (PULs), which consist of sets of co-regulated genes for sensing nutrient availability, glycan capture, uptake and digestion (Wexler and Goodman, 2017). PULs can have activity for a single glycan or a set of chemically similar glycans (Déjean et al., 2020; Larsbrink et al., 2014; Ndeh et al., 2020; Sonnenburg et al., 2010). In some cases, multiple PULs can contribute to the utilization of chemically complex glycans (Cartmell et al., 2018; Rogowski et al., 2015). The abundance of genome sequencing data has enabled the identification of many PULs in gut bacteria and facilitated their biochemical characterization (Briliute et al., 2019; Larsbrink et al., 2014; Luis et al., 2018; Martens et al., 2008; Martens et al., 2011; Ndeh et al., 2017). However, there remain many PULs with unknown functions (Lapébie et al., 2019).

Glycan utilization is a major driver of inter-species interactions in the gut microbiome. Competition for a given glycan can generate negative inter-species interactions (Tuncil et al., 2017). By contrast, utilization of polysaccharide breakdown products (PBPs) generated via extracellular glycan digestion or released metabolic byproducts (derived from intracellular metabolism) can enhance the growth of specific community members (Rakoff-Nahoum et al., 2014; Rakoff-Nahoum et al., 2016). However, glycans can also be degraded via selfish mechanisms where PBPs are not released into the extracellular environment (Cuskin et al., 2015; Rakoff-Nahoum et al., 2016). Gut bacteria that produce beneficial molecules such as butyrate have a narrower range of glycan degrading capabilities (El Kaoutari et al., 2013; Sheridan et al., 2016). Therefore, the unselfish release of PBPs and metabolic byproducts from *Bacteroides* could establish new metabolic niches that could be exploited by butyrate producers. However, we do not fully understand how glycan utilization via PULs in *Bacteroides* modulates inter-species interactions and beneficial gut microbiome functions such as butyrate production.

BU is one of the most abundant and prevalent species in the gut microbiome (Almeida et al., 2021; Qin et al., 2010). However, we lack an understanding of the contributions of individual

PULs to the fitness of BU in response to diverse glycans. Equipped with a CRISPR genome editing system for *Bacteroides*, we investigate the contribution of 23 (55 total) PULs in *B. uniformis* DSM 6597 to fitness by combining barcoded PUL deletion mutants into a single culture (pooled mutant barcoding strategy) (Terrapon et al., 2017). We discover glycan utilization functions and the transcriptional coordination of PULs in response to complex glycans. Notably, while the presence of a given PUL can provide a fitness advantage in certain environments, we show that PULs can have a negative impact on fitness and colonization ability of the mammalian gut. Finally, we demonstrate that PULs in BU can shape ecological dynamics and the production of the beneficial metabolite butyrate.

RESULTS

Development of genetic tools for constructing barcoded PUL deletion strains in BU

To investigate the contributions of PULs to BU fitness, we developed a genome editing tool for *Bacteroides*. Current gene manipulation methods for *Bacteroides* frequently involve two-step selection and counterselection (Baughn and Malamy, 2002; Garcia-Bayona and Comstock, 2019; Koropatkin et al., 2008), thus often limiting their generalizability to diverse *Bacteroides* isolates. Developing CRISPR-Cas genome editing tools in *Bacteroides* could expand our ability to understand and engineer health-relevant functions. To this end, we used the type V CRISPR-Cas nuclease from *Francisella novicida* U112 (FnCpf1) due to several unique features for genome editing, including its small size (i.e. lower fitness burden), endoribonuclease domain and functionality without requiring RNase activity (Zetsche et al., 2015). To construct the CRISPR-FnCpf1 system and develop gene expression manipulation tools in BU, we constructed a library of ribosome binding sites (RBSs), constitutive and inducible promoters and shuttle plasmids (Lim et al., 2017; Mimee et al., 2015; Whitaker et al., 2017) (**Figure S1**). Equipped with the shuttle plasmids (pFW1000 and pFW2000), constitutive promoter P_{BT1311}, and IPTG inducible promoter, we designed CRISPR-FnCpf1 systems for gene manipulation in *Bacteroides*. The FnCpf1 harboring plasmids varied in the type of replication origin and presence/absence of IPTG-inducible *E. coli* recombinase *recT* for enhancing gene manipulation efficiency (**Figure S2A**).

The CRISPR-FnCpf1 *Bacteroides* genome editing system was successfully used to delete 23 PULs in BU that were selected based on predicted glycan degrading activities and cluster length (**Supplementary Data 1**) (Almagro Armenteros et al., 2019; Kelley et al., 2015; Terrapon et al., 2017). The gene deletion efficiency is defined as the number of correct mutant colonies divided by the total number of tested colonies and ranged between 3-100% for deletion of these PULs. We found that plasmid pFW2500 harboring *recT* had a higher gene deletion efficiency in

BU (**Figure S2G**) and the construction of ~50% PUL mutants required RecT. We created two double deletions of *PUL22* and *PUL23* due to their proximity or *PUL11* and *PUL43* based on bioinformatic prediction of their functions (Larsbrink et al., 2014). To quantify the abundance of each strain when combined into a single culture, we simultaneously deleted the tryptophanase *tryP* (*BACUNI_RS07455*) and introduced a unique 4-base pair DNA barcode downstream (**Figure 1A**). We deleted *tryP* to reduce the production of potential toxic compounds linked to chronic kidney disease for future therapeutic applications (Devlin et al., 2016). Finally, we demonstrated CRISPR-FnCpf1 mediated gene deletions in *Bacteroides thetaiotamicron* (BT) and *Bacteroides fragilis* and a gene insertion in BU (**Figures S2D-S2F**). In sum, CRISPR-FnCpf1 coupled to RecT expression can be used for efficient genome engineering in *Bacteroides*.

Effects of PUL deletions on the growth of BU

BU can utilize diverse dietary and host-derived glycans in the human gut microbiome (Benitez-Paez et al., 2017; Cartmell et al., 2018; Desai et al., 2016; Li et al., 2017; Luis et al., 2018; Pluvinage et al., 2018; Sonnenburg et al., 2010). We characterized the glycan utilizing ability of BU by growing the wild-type strain (BU WT) in *Bacteroides* minimal media (Bacic and Smith, 2008) individually supplemented with 22 glycans, some of which were chosen based on previous studies (Benitez-Paez et al., 2017; Cartmell et al., 2018; Desai et al., 2016; Li et al., 2017; Luis et al., 2018; Pluvinage et al., 2018; Sonnenburg et al., 2010). Our results showed that 50% of the tested glycans could support growth (**Figure S3A**), indicating that BU can utilize a wide range of chemically diverse glycans.

To quantify the effect of a given PUL deletion on the growth of BU in the presence of individual glycans, we fit a logistic growth model to time-series measurements of absorbance at 600 nm (OD_{600}) (Venturelli et al., 2018). This model captures microbial growth as a function of its exponential growth rate and carrying capacity (i.e. steady-state abundance) (**Figures 1B and S3B**). To account for any potential effects of the genome integrated barcode and *tryP* deletion on fitness, barcoded BU WT (Δ*tryP-24*) was used as the reference strain. The growth of BU WT and the Δ*tryP-24* were similar in glucose media, demonstrating that the genome modifications did not impact the growth of BU (**Figure S3C**).

We examined the fold change of the steady-state abundance or growth rate for each PUL mutant compared to Δ*tryP-24* in media with individual glycans (**Figure 1B; Supplementary Data 2**). In the presence of glucose, the inferred parameters of the PUL mutants were similar to Δ*tryP-24*, demonstrating that these PUL deletions did not impact the growth of BU in media with simple carbon sources. However, certain PUL deletions reduced the steady-state abundance of BU in

the presence of glycogen, type II mucin, arabinogalactan, or galactomannan. By contrast, certain PUL deletions enhanced steady-state abundance in media with xyloglucan, inulin, laminarin or pectin compared to Δ*tryP-24*. The steady-state abundance of Δ*PUL18* was the most significantly reduced across the majority of glycans, suggesting that Δ*PUL18* has promiscuity for multiple glycans.

We applied a threshold in the percent change of the inferred growth parameters compared to Δ*tryP-24* (<40% steady-state abundance and/or <20% growth rate) to identify a subset of mutants that have a substantial impact on BU fitness. For instance, the bipartite network in **Figure 1C** indicated that *PUL18* had multiple glycan degradation activities. In addition, *PUL12*, *PUL17, PUL11/PUL43,* contributed to glucomannan and galactomannan, inulin or xyloglucan utilization, respectively. These mutants were further characterized using higher resolution time-series OD_{600} measurements (**Figure 1D**). Whereas some mutants were unable to grow on a given glycan, the growth of other mutants were diminished but not abolished (Δ*PUL12*-galactomannan, Δ*PUL18*-type II mucin and Δ*PUL21*-laminarin), suggesting that other genes in BU could contribute to utilization of these glycans.

The sequence and glucomannan/galactomannan utilization functions of *PUL12* are similar to a previously reported PUL (*BACOVA_02088-97*) in *Bacteroides ovatus* (BO) (Bågenholm et al., 2017). While both PULs contain an outer membrane glycoside hydrolase family 26 (GH26), other glycoside hydrolases differed between *PUL12* and *BACOVA_02088-97* including an α-galactosidase (GH36 or *BACOVA_02091*). This enzyme is critical for the galactomannan utilization and was present in *BACOVA_02088-97* and absent in *PUL12* (Bågenholm et al., 2017). *PUL12* had greater specificity for glucomannan, whereas *BACOVA_02088-97* showed preference for galactomannan, suggesting that the differences in these enzymes contributed to the variation in glycan preferences.

PUL18 has similarity to a previous reported BT starch and pullulan utilization loci, as both PULs contain GH13 enzymes (α-amylase and pullulanase; **Supplementary Data 1**) (Koropatkin et al., 2008; Martens et al., 2009a). However, *PUL18* also contains an endo-β-1,4-galactanase (GH53) and β-galactosidase (GH2), which could enable utilization of pectin and pectic galactan. To test the possibility that *PUL18* consists of two separate PULs, we measured the expression of specific *PUL18* genes in media with glycogen, pectic galactan or type II mucin using real-time quantitative reverse transcription PCR (qRT-PCR). All measured genes in *PUL18* were upregulated in media with glycogen, pectic galactan or type II mucin compared to media with glucose, indicating a similar pattern of regulation across the measured genes (**Figures S4A, S4B and S4E**). These data suggest that *PUL18* consists of a single PUL.

While utilization of pullulan, glycogen, pectin and pectic galactan are supported by the enzyme content of *PUL18*, this PUL does not contain enzymes predicted to utilize type II mucin (**Supplementary Data 1**). To explore this further, we performed RNA-seq on BU WT in the presence of type II mucin or glucose. Our results showed that *PUL5*, *15*, *18* and *28* were significantly up-regulated in response to type II mucin compared to glucose, with *PUL18* displaying the highest fold changes than the other up-regulated PULs (**Figures S4D and S4E; Supplementary Data 3,** Sequencing data is available from Zenodo: doi:10.5281/zenodo.5207518). These data combined with the substantial growth defect of the Δ*PUL18* in media with type II mucin suggests that *PUL18* contributes either directly or indirectly to type II mucin utilization (**Figure 1D**). However, we cannot rule out the possibility that the up-regulated PULs in the presence of type II mucin contribute to utilization of potential contaminants in the type II mucin extract.

PUL17 has the unique capability to utilize both inulin (β2-1 fructan) and levan (β2-6 fructan), which has not been previously observed in *Bacteroides* (**Figures 1D and S3D**). We predicted the functional activities and sub-cellular localization of the PUL enzymes **(Supplementary Data 1**) (Almagro Armenteros et al., 2019; Bhasin et al., 2005; Kelley et al., 2015; Terrapon et al., 2017; Yu et al., 2014; Yu et al., 2010). A comparison of BU PULs characterized in this study with previous experimentally characterized PUL functions in *Bacteroides* is described in **Supplementary Data 1** (Bågenholm et al., 2017; Déjean et al., 2020; Déjean et al., 2019; Hemsworth et al., 2016; Koropatkin et al., 2008; Larsbrink et al., 2014; Luis et al., 2018; Martens et al., 2008; Ndeh et al., 2017; Rakoff-Nahoum et al., 2016; Reddy et al., 2016; Shipman et al., 2000; Shipman et al., 1999; Sonnenburg et al., 2010; Spence et al., 2006; Tancula et al., 1992; Tauzin et al., 2016).

To determine the abundance and prevalence of BU PULs in the human gut microbiome, we performed bioinformatic analysis of metagenomic sequencing datasets (Almeida et al., 2019; Pasolli et al., 2019) (**Methods**, **Figure S5A**). Whereas *PUL12*, *17*, *21*, *11* and *43* were found in over 53% of BU MAGs, *PUL18* was observed in ~36%, indicating that *PUL18* is less conserved in BU than the other PULs (genome completeness >90%, **Figure S5B**). Although these PULs were found in *Bacteroides* genomes beyond BU, they were infrequently observed in gut species excluding *Bacteroides* (**Figure S5C**). However, our method does not identify potential orthologous PULs. In sum, these PULs were ubiquitous in BU but rarely found in other gut species, highlighting the specificity of these PULs to BU.

Xyloglucan utilization in BU is due to the coordination of two PULs

A xyloglucan utilization pathway (XyGULs) previously reported in BO (Hemsworth et al., 2016; Larsbrink et al., 2014) has similarity to *PUL11* in BU. In addition, a second xyloglucan utilization pathway was predicted in BU (Larsbrink et al., 2014) (*PUL43* and *PUL44* in this work) (**Figure 2A**). To investigate the functions of these PULs, we constructed the double deletion strains Δ*PUL11_43* and Δ*PUL11_44*. We found that Δ*PUL11*, Δ*PUL43,* Δ*PUL44* and Δ*PUL11_44* grew in the presence of xyloglucan, whereas Δ*PUL11_43* failed to grow. This implies that *PUL11* and *43* have redundant roles in xyloglucan utilization (**Figure 2B**). While Δ*PUL43* had a similar growth response to Δ*tryP-24*, Δ*PUL11* and Δ*PUL11_44* exhibited a longer lag phase and higher maximum abundance than Δ*tryP*-24 (**Figure 2B**), suggesting that *PUL11* enables a faster response to xyloglucan but reduces the maximum biomass.

To test whether *PUL11* and *43* were transcriptionally co-regulated, we performed RNA-seq of Δ*tryP-24*, Δ*PUL11*, Δ*PUL43*, Δ*PUL44* or Δ*PUL11_44* in the presence of xyloglucan or the control glucose (**Supplementary Data 3,** Sequencing data is available from Zenodo: doi:10.5281/zenodo.5204678). The expression of *PUL44* genes was not upregulated in xyloglucan compared to glucose media, consistent with the negligible effect of the *PUL44* deletion on growth in media with xyloglucan (**Figures 2B, 2C and S4C**). Therefore, our results indicate that *PUL44* does not contribute to xyloglucan utilization and thus *PUL43* and *44* have independent functions. In Δ*tryP-24*, *PUL11* and *43* were upregulated in media with xyloglucan compared to the glucose condition (**Figures 2C and S4C**), with *PUL11* exhibiting higher transcript abundance on average than *PUL43*. The xyloglucan-dependent induction of *PUL11* and *43* genes corroborated their key roles in xyloglucan utilization in BU.

To explore potential transcriptional coordination, we studied how the expression of *PUL11* or *43* were impacted by the deletions of *PUL43* or *11* in media with xyloglucan, respectively. Notably, the expression of *PUL43* was significantly up-regulated in Δ*PUL11* and Δ*PUL11_44* compared to Δ*tryP-24* in both the RNA-seq and qRT-PCR data, whereas the expression of *PUL11* did not depend on the presence of *PUL43* (**Figures 2D and 2E**). In sum, these data suggest that the transcriptional coordination between *PUL11* and *43* may enable BU to adapt to the loss of *PUL11* in xyloglucan (**Figure 2B**). Notably, many PULs distributed across the BU genome were significantly up- or down-regulated in the PUL deletion strains compared to Δ*tryP-24* in media with xyloglucan (**Figure 2F**). These data indicate the deletion of a given PUL can have global effects on PUL gene expression patterns in BU.

We examined the co-occurrence of *PUL11* and *43* across BU MAGs in the human gut microbiome (Almeida et al., 2019; Pasolli et al., 2019). We found that both *PUL11* and *43* were present in ~50% of BU MAGs (genome completeness >90%) and rarely found in MAGs of other

gut organisms (**Figure S5D**). The frequent co-occurrence of *PUL11* and *43* in BU suggests that redundant PULs for xyloglucan utilization may provide a fitness advantage for BU.

PULs can negatively or positively impact fitness of BU in competitive environments

The presence of a given PUL can vary across isolates of the same gut species (Fehlner-Peach et al., 2019) (**Figure S5B**) and PULs have been shown to evolve within individuals, potentially due to the variable selection pressures of host diet (Zhao et al., 2019). Therefore, we hypothesized that PULs may provide a benefit or cost to microbial fitness depending on the nutritional landscape. To test this hypothesis, we combined the 22 PUL mutants and Δ*tryP-24* into a mutant pool and characterized their abundances in the presence of single glycans using serial passaging (**Figure 3A**). The absolute abundance of each mutant was determined by multiplying the relative abundance measured by barcode sequencing by OD_{600} (Sequencing data is available from Mendeley data: doi: 10.17632/3888sct8c5.1).

The OD_{600} of the mutant pool varied in the presence of individual glycans, with pectic galactan and pectin exhibiting the highest and lowest total biomass across conditions, respectively (**Figure 3B**). The growth impairment of certain PUL mutants (e.g. Δ*PUL12*-glucomannan or galactomannan; Δ*PUL17*-inulin, Δ*PUL18*-glycogen or pectic galactan or pectin; Δ*PUL21*-laminarin) were consistent in both monoculture and the mutant pool (**Figures 1D and 3C**). However, growth of specific PUL deletion strains in monoculture and the mutant pool deviated in the presence of certain glycans (e.g. Δ*PUL18*-type II mucin or pullulan) (**Figures 1D and 3C**), suggesting that interactions between mutants could impact growth. Our results showed Δ*tryP-24* was outcompeted by PUL mutants in media with glycogen, pectic galactan or pectin (**Figure S6**). These results indicate that specific PULs can provide a fitness cost that depends on the nutrient environment.

We analyzed the Shannon diversity of the mutant pool in the presence of individual glycans to provide insight into the strength of growth selection. The Shannon diversity in pectic galactan was substantially lower compared to other glycan conditions, indicating that pectic galactan provided a strong growth selection for certain BU mutants (**Figure 3D**). By contrast, the Shannon diversity was high in media with glucose, inulin or xyloglucan, demonstrating that these carbon sources provided a weaker growth selection for the BU mutants. The high Shannon diversity in these conditions was consistent with the robust monoculture growth of the majority of PUL mutants in media with glucose, inulin or xyloglucan (**Figure 1B**).

Deletion of key PULs enhances in vivo colonization

To understand how PULs impact the fitness of BU in the mammalian gut environment in response to nutrient availability, we colonized gnotobiotic mice with the BU mutant pool. Each group of mice was fed a different diet that varied in the type and abundance of microbiota accessible carbohydrates (MAC) (**Supplementary Data 4**). Characterization of temporal changes in mutant abundance *in vivo* could provide insights into interactions between PULs in BU and the host as well as how the diet modulates these interactions. We introduced the mutant pool into germ-free mice fed a diet containing high MAC (high fiber diet) or MAC free (fiber free diet or FFD) (**Figure 4A**). To understand the effects of diet-derived glycans on PUL mutant colonization, separate groups of mice fed the FFD received a given glycan (inulin, glucomannan or pectic galactan) in the drinking water (**Figure 4A**). In addition, we included a high fat and low MAC diet (high fat diet) based on the hypothesis that BU would shift its metabolic niche towards utilization of host-derived glycans due to the limited availability of diet-derived glycans (Koropatkin et al., 2012; Mahowald et al., 2009).

To understand how variation in the diet impacted the overall fitness of the mutant pool in the murine gut, we characterized the absolute abundance of BU in ceca at the end of the experiment. The absolute abundance of BU was significantly lower in the FFD group and highest in the high fiber group (**Figure 4B**), indicating that higher MAC enhances BU colonization ability. In addition, Δ*PUL37* was highly abundant in all conditions except the high fat diet, whereas Δ*PUL18* was present at variable levels in all diets except FFD-pectic galactan (**Figure 4B,** Sequencing data is available from Mendeley data: doi: 10.17632/3888sct8c5.1). The temporal changes in the composition of the mutant pool varied across diets, indicating that diet is a critical variable shaping the colonization ability of the PUL mutants.

We analyzed the maximum rate of change of the Shannon diversity of the BU mutant pool to quantify the strength of selection across diets. The Shannon diversity of the mutant pool decreased most rapidly in the high fat diet, whereas a gradual decrease in diversity was observed in the FFD and FFD supplemented with inulin (FFD-inulin) (**Figure 4C**). The substantial decrease in diversity of the mutant pool across all diets indicates a strong selection for a subset of PUL deletion strains that could stably colonize the murine gut. At the final time point, the mutant pools in the FFD and FFD-inulin groups displayed higher Shannon diversities than the other diets. Notably, the trends in Shannon diversity were consistent both *in vitro* and *in vivo*, with the diversity being lowest and highest in media with pectic galactan or inulin/glucose, respectively (**Figures 3D and 4C**). This implies that nutrient availability was one of the major driving factors of Shannon diversity *in vivo*. Principal component analysis over all time points revealed that Δ*PUL18* and

Δ*PUL37* contributed most significantly to the temporal changes in the mutant pool composition and the abundance of Δ*PUL18* distinguished the high fat diet from all other diets (**Figure 4D**).

The abundance of most mutants decreased over time in FFD, FFD-inulin and FFD-glucomannan groups, potentially due to competition with the high fitness mutants Δ*PUL18* and Δ*PUL37* (**Figures 4C and 4E**). However, Δ*PUL12* was more abundant in FFD than FFD-glucomannan for a period and Δ*PUL17* exhibited moderately higher abundance in FFD than FFD-inulin, consistent with the *in vitro* characterization of these strains (**Figures 1D and 4E**). Δ*PUL18* was unable to colonize mice in FFD-pectic galactan but persisted as a high fraction of the mutant pool in the mice fed all other diets, consistent with the critical role of this PUL in pectic galactan utilization identified *in vitro* (**Figure 1D**). Therefore, the glycan-utilization functions of PULs identified *in vitro* can be used to predict the colonization ability of BU in the mammalian gut in response to specific diets. To determine the expression of PULs in the mammalian gut, we monocolonized germ-free mice fed a high fiber diet with BU WT. We measured the fold change in the expression of representative genes in a subset of PULs (12, 17, 18, 21 and 37) compared to a glucose media control using qRT-PCR. Our results showed that most of these genes were up-regulated, suggesting that these PULs are actively expressed in response to host colonization (**Figures S4F**).

The temporal changes in mutant pool composition were significantly different in mice fed the FFD and high fat diets although both diets were deficient in diet-derived glycans. Therefore, the colonization ability of the mutant pool in high fat diet could be attributed to differences in dietary fat content or host responses to the high fat diet. Notably, Δ*tryP-24* colonized at a low abundance only in mice fed the FFD and FFD-inulin diets. The diminished colonization ability of Δ*tryP-24* in competition with other mutants mirrored the low fitness of Δ*tryP-24 in vitro* across many conditions (**Figure S6**). In sum, our data indicate that PULs play a major role in the colonization ability of BU due to an interplay of diet-derived nutrient availability and potential microbe-microbe and microbe-host interactions.

PULs are major drivers of inter-species interactions with butyrate producers

Butyrate is linked to numerous health benefits including maintaining homeostasis of the gut environment (Duncan et al., 2007; Fischbach and Sonnenburg, 2011; Makki et al., 2018). We hypothesized that glycans utilized by BU could influence inter-species interactions with butyrate producers. To this end, we characterized the growth of BU WT, Δ*PUL12*, Δ*PUL17*, Δ*PUL18*, Δ*PUL21* and Δ*PUL11_43* and four highly prevalent butyrate producers in the human gut microbiome, including *Anaerostipes caccae* (AC), *Coprococcus comes* (CC), *Eubacterium rectale*

(ER) and *Roseburia intestinalis* (RI), in monoculture and in pairwise communities in media with single carbon sources (**Figure 5A**). To decipher inter-species interactions, we fit a generalized Lotka-Volterra model (gLV) to time-series measurements of species absolute abundance based on 16S rDNA sequencing and OD_{600} (**Figures 5A and S7; Supplementary Data 2**). The gLV model is a set of coupled differential equations that describes species growth dynamics due to each organism's growth rate and interactions with each community member (Venturelli et al., 2018). We used a Markov-Chain Monte Carlo method to infer the parameters of the gLV model (Haario et al., 2001).

Visualizing the inferred inter-species interaction coefficients as a network highlighted that BU WT can substantially enhance the growth of butyrate producers in media with inulin, laminarin, pectic galactan, pectin or pullulan (**Figure 5B**). Most of these inferred interactions vanished in communities composed of a PUL mutant and butyrate producer, demonstrating the critical role of PULs in mediating inter-species interactions. For example, BU and each butyrate producer co-existed in media with inulin, but the growth of all butyrate producers was abolished in co-culture with Δ*PUL17* (**Figures 5A, 5B and S7**). In cases where the butyrate producer could utilize the glycan, the inferred inter-species interactions with BU WT or the PUL mutant exhibited major differences in directionality and sign. For instance, ER could utilize pullulan and its growth was enhanced by BU WT and inhibited by Δ*PUL18* (**Figure 5B**). Further, the growth of BU WT was inhibited by ER whereas the growth of Δ*PUL18* was enhanced by ER. Together, these data demonstrate the key role of PULs in shaping ecological networks.

Previous work has demonstrated that the abundance of butyrate producers can be used to predict butyrate production (Clark et al., 2021). Therefore, we constructed a linear regression model to predict the end-point butyrate concentration based on the abundance of the butyrate producers (**Methods**). Using the trained regression models, the predicted butyrate concentration was highly correlated to the measured butyrate concentration (Pearson r=0.9, p=7.1e-18) (**Figure 5C**). This finding implies that butyrate production is primarily determined by the abundance of the butyrate producer, while the butyrate produced per cell is largely independent of the type of carbon source, as well as the presence/absence of BU. Therefore, these results demonstrate PUL-dependent glycan utilization by BU impacts butyrate production by modulating the growth of the butyrate producers. In sum, BU PULs have a critical role in determining inter-species interactions influencing butyrate production.

PULs influence BU-butyrate producer interactions via three major mechanisms

We sought to understand the molecular basis of the inferred inter-species interactions between BU and the butyrate producers. We hypothesized that the inferred positive interactions supporting the butyrate producer growth in co-culture with BU WT could be due to the release of metabolic byproducts (derived from intracellular metabolism) or energy rich PBPs degraded extracellularly.

To determine if the released compounds (including metabolic byproducts and PBPs) from BU could impact butyrate producer growth, we characterized the growth of butyrate producers in BU conditioned glycan media. To eliminate the effect of external pH modification, the pH of the conditioned glycan media was adjusted to match fresh media (**Figure S8A**). We computed the ratio of total growth (area under the curve or AUC) of each butyrate producer in conditioned glycan media (AUC_{CM}) to the corresponding fresh media (AUC_{FM}). This metric showed that the conditioned glycan media enhanced the growth of butyrate producers across many conditions. For instance, a growth benefit for all butyrate producers was observed in conditioned inulin or laminarin media, whereas a negligible growth benefit was observed for other conditions such as CC in conditioned xyloglucan media (**Figure S8B and Supplementary Data 5**). Therefore, our results demonstrate that the effect of released compounds from BU on the growth of each butyrate producer was determined by the specific glycan, consistent with the glycan-dependent variation in the inferred inter-species interaction networks of the communities (**Figure 5B**).

We next characterized the effects of PBPs released by BU due to outer-membrane glycan degrading enzymes (**Figures S8C and S8D; Supplementary Data 1**) (Cartmell et al., 2018; Rakoff-Nahoum et al., 2014; Sonnenburg et al., 2010). To test this possibility, we treated glycan media with BU cell membrane fractions (**Methods; Figure S8C**) and determined the fold change in the AUC of the growth of a given butyrate producer in the cell membrane treated (AUC_{MT}) to untreated media (AUC_{FM}) (**Figure S8D; Supplementary Data 5**). Our results suggest that BU releases PBPs that can be utilized by a butyrate producer if $AUC_{MT}\ AUC_{FM}^{-1}$ was much larger than one and the associated PULs were predicted to have only outer-membrane glycan degrading enzymes (e.g. *PUL17*-inulin) (**Figure S8D; Supplementary Data 1 and 5**).

We used the combined results of these experiments to predict whether the PUL mechanism of glycan degradation was selfish or unselfish (i.e. released PBPs). For instance, for inulin and laminarin, all butyrate producers displayed substantial growth benefits in co-culture with BU, BU conditioned glycan media and cell membrane treated glycan media compared to their respective controls (**Figures S7 and S8; Supplementary Data 5**). Therefore, these data suggest that the PULs responsible for utilization of inulin (*PUL17*) and laminarin (*PUL21*) in BU degrade their respective glycans via an unselfish mechanism. For xyloglucan, we observed a growth enhancement of specific butyrate producers (e.g. AC, ER or RI) in BU conditioned xyloglucan

media and in cell membrane treated xyloglucan media, but not in co-culture with BU (**Figure S7; Supplementary Data 5**). These data suggest that BU can release xyloglucan breakdown products and AC, ER and RI have the capability to utilize these products, but the community interactions are more complex.

Based on our data, we propose three classes of mechanisms where PULs in BU can influence inter-species interactions. In Mechanism A (**Figure 6A**), the butyrate producer can utilize PBPs and metabolic byproducts such as acetate (Duncan et al., 2004) released by BU but not the glycan. In this case, butyrate producer growth and butyrate production are enhanced in co-culture with BU WT and not supported in co-culture with the PUL mutant deficient in utilization of the given glycan. For example, the interactions between AC and BU or Δ*PUL17* in media with inulin were consistent with Mechanism A. Whereas AC did not display growth or butyrate production in inulin media or in co-culture with Δ*PUL17*, significant growth and butyrate production were observed in the presence of BU WT (**Figures 6B and 6C**). In addition, the inulin degradation product fructose was detected in BU conditioned inulin media and growth of AC was observed in media with fructose as the primary carbon source, suggesting that the growth of AC was enhanced by inulin breakdown products released by BU (**Figures 6D and 6E**).

In Mechanism B, BU and the butyrate producer can utilize a given glycan and thus the butyrate producer exhibits growth in co-culture with BU WT or the PUL mutant (**Figure 6F**). However, the presence or absence of a PUL required for utilization of a certain glycan can shape the inter-species interaction network. For example, RI displayed growth and butyrate production in pullulan media and in co-culture with BU WT or Δ*PUL18* (**Figures 6G and 6H**). The growth of BU WT was inhibited by RI whereas RI enhanced the growth of Δ*PUL18*, which was captured by the inferred inter-species interaction network (**Figures 5B, S7B and S7C**).

In Mechanism C, the butyrate producer lacks the capability to utilize the glycan and PBPs (**Figure 6I**). However, metabolic byproducts released by BU could lead to a moderate or even negligible growth benefit (e.g. BU and CC in xyloglucan media) (**Figures 6J and 6K**). Growth and butyrate production of CC was not observed in co-culture with BU WT or Δ*PUL11_43*, indicating that any potential released compounds from BU did not benefit CC. In addition, the BU conditioned xyloglucan media or BU cell membrane treated xyloglucan media did not enhance the growth of CC (**Figures 6L and S8; Supplementary Data 5**). In sum, the PUL specific mechanism of glycan degradation and utilizing capabilities of the butyrate producer are critical variables shaping community dynamics and butyrate production.

DISCUSSION

We developed a CRISPR-FnCpf1 genome editing tool to comprehensively study the role of 23 PULs on the fitness of BU, providing a deeper understanding of how diverse PULs contribute to glycan utilization, microbial interactions, and colonization of the mammalian gut. We show that PULs can provide a fitness benefit or cost depending on the nutrient landscape. In addition, we showed that the deletion of given PUL impacts global PUL gene expression patterns, highlighting unknown mechanisms of transcriptional coordination. Finally, PULs can shape ecological networks via competition for glycans or release of PBPs and metabolic byproducts, thus shaping the dynamics and butyrate production capability of gut communities.

The observed transcriptional coordination between *PUL11* and *43* for xyloglucan utilization could be due to cross-regulation of *PUL43* by a sensor-regulator in *PUL11*, mirroring the transcriptional linkage observed between xylan-degrading PULs in *Bacteroides xylanisolvens* (Despres et al., 2016). The shifts in the gene expression patterns of PUL genes across the BU genome in response to a given PUL deletion could arise from cross-regulation by a sensor-regulator in the deleted PULs. Alternatively, elimination of certain enzymes in these PULs could result in differences in the PBP profile, which in turn could alter the regulation of disparate PULs across the BU genome by direct sensing of these altered PBP profiles. Finally, global regulators that respond to carbon limitation could couple the expression of disparate PULs (Schwalm et al., 2016; Townsend et al., 2020).

Overall, our data show that the effects of PULs on BU colonization were complex: the presence of a given PUL could improve fitness in specific nutrient conditions, whereas it can reduce fitness in other conditions. Therefore, while a given PUL can enhance fitness by enabling utilization of key nutrients, expression of PUL genes may impose substantial energetic costs when they are not needed for growth (Venturelli et al., 2015). Supporting this hypothesis, BT has been shown to constitutively express certain PULs at a low-level (Martens et al., 2009b; Sonnenburg et al., 2010; Sonnenburg et al., 2005). This regulatory strategy may allow the cells to rapidly adapt to temporally changing nutrient conditions at the cost of unnecessary gene expression.

Bioinformatic analysis of *PUL37* predicted enzymes that could target degradation of glycosphingolipids on host cells (**Supplementary Data 1**) (Ryckman et al., 2020). Immunoglobulin A (IgA) regulates the growth and functional activities of gut microbiota. For instance, IgA targeted pectin and fructan utilization PULs in BT, leading to a decrease in PUL gene expression (Joglekar et al., 2019). To regulate the potential host-derived glycan degrading activities of *PUL18* or *37*, IgA could target proteins in these PULs. However, we cannot exclude other possibilities including the role of *PUL18* or *PUL37* as receptors for prophage (Reyes et al.,

2013) or interactions with capsular polysaccharide synthesis, which can provide protection from the host immune system and phages (Martens et al., 2009b; Sonnenburg et al., 2005). Future work will elucidate the mechanisms that enable high colonization ability of Δ*PUL18* and *ΔPUL37*. These mechanisms could be exploited as potential engineering targets for manipulating the colonization ability of BU in the mammalian gut.

BU enabled the growth of butyrate producers in media with inulin, laminarin, pectic galactan, pectin and pullulan, suggesting that the corresponding PULs (*PUL17*, *21* and *18*) degrade glycans via unselfish mechanisms. Release of PBPs by BT has been shown enhance growth and butyrate production of *E. ramulus* and AC (Chia et al., 2020; Rodriguez-Castano et al., 2019). In addition, the host butyrate transporter *Mct-1* was up-regulated in the presence of both BT and ER compared to mice monocolonized with ER (Mahowald et al., 2009). We found that interactions between BU and butyrate producers can be more nuanced when both species are able to utilize a given glycan. For example, both species could compete for the available glycan and the butyrate producer could enhance the growth of BU. In sum, our results suggest that inter-species interactions between *Bacteroides* and butyrate producers mediated by glycan competition, release of PBPs and metabolic byproducts are critical determinants of ecological dynamics. Since we did not directly measure PBPs and metabolic byproducts, we cannot rule out the possibility that other interaction mechanisms could impact the growth of the butyrate producers (e.g. release of signaling molecules by BU that activates metabolic networks in the butyrate producers or other high energy substrates beyond PBPs).

PULs harbored by gut bacteria provide the host with the unique capability to transform chemically diverse glycans into molecules that serve as nutrients for the host and shape the ecological dynamics of the gut microbiome. A deeper understanding of the contributions of PULs to microbial fitness and community-level functions could guide the design of microbiome interventions to optimize nutrient extraction from food or restore homeostasis of gut microbiota following a disturbance. For example, PULs could be harnessed as a control knob to modulate colonization or key inter-species interactions to promote beneficial functions performed by gut microbiota.

ACKNOWLEDGEMENTS

We would like to thank Yu-Yu Cheng, Bryce M. Connors, Ryan L. Clark, Erin Ostrem Loss, Susan E. Hromada for helpful advice; Michael Fischbach, Anthony Rush, Justin Sonnenburg and Michelle St Onge for providing plasmids; Mcsean Mcgee, David Stevenson and Joshua Coon for assistance in butyrate measurements; Aaron Hoskins and Xingyang Fu for providing access to

key equipment. This research was sponsored by the National Institutes of Health under Grant Number R35GM124774, National Institute of Biomedical Imaging and Bioengineering under grant number R01EB030340, Multi University Research Initiative (MURI) under grant number W911NF-19-1-0269 and Department of Energy under grant number DE-FC02-07ER64494.

AUTHOR CONTRIBUTIONS

Conceptualization, O.S.V. and J.F.; Methodology, J.F., O.S.V., Y.Q., C.Z., K.A., and F.E.R.; Investigation, J.F. and E.I.V.; Writing – Original Draft, J.F., O.S.V., Y.Q., S.E., C.Z., and K.A.; Writing – Review & Editing, J.F., O.S.V., Y.Q., C.Z., K.A., and F.E.R.; Formal Analysis, J.F., O.S.V., Y.Q., S.E., and C.Z.; Resources, F.L. and J.H.H.; Funding acquisition, O.S.V.

DECLARATION OF INTERESTS

The authors do not have a conflict of interest.

FIGURE LEGENDS

Figure 1. Effects of polysaccharide utilization loci (PULs) on growth of *B. uniformis* in response to individual glycans. (A) Schematic of the two-stage procedure to construct the barcoded PUL deletion mutants using the CRISPR-FnCpf1 genome editing tool. First, the barcode-tagged mutants were constructed and then each PUL was deleted using FnCpf1. Gray square brackets indicates that the plasmid used to delete a subset of PULs harbored *recT*. The growth of each mutant was characterized in media with individual glycans. Plasmid features include p15A, *E. coli* plasmid replication origin; pB8-51, *Bacteroides* plasmid replication origin; *lacI*, transcription factor; P_{lacO23}, IPTG inducible promoter. **(B)** Biclustering heatmap of the fold changes in steady-state abundance (k_i in the logistic growth model) for each PUL mutant to Δ*tryP-24* in media with single carbon sources. Asterisks represent low confidence (coefficient of variation > 0.2) in the inferred parameter value (**Supplementary Data 2**). In these conditions, the steady-state abundance was set to the maximum OD_{600}. **(C)** Bipartite network of PULs and glycans based on thresholds in the fold changes of inferred steady-state abundance and/or growth rate of each PUL deletion mutant to Δ*tryP-24*. **(D)** Time-series measurements of OD_{600} of PUL mutants in media with specific glycans highlighted in the bipartite network in **(C)**. Lines denote the mean and the shaded regions represent 95% confidence interval of 3 biological replicates.

Figure 2. Co-regulation of polysaccharide utilization loci (PULs) for xyloglucan utilization in *B. uniformis*. (A) Schematic of gene organization of *PUL11*, *PUL43* and *PUL44*. GenBank locus tags are shown for each gene. SusC: SusC-like TonB-dependent transporter (Red); SusD: SusD-like cell-surface glycan-binding protein (Pink); SBP, sugar binding protein (Purple); HTCS, hybrid two-component system (Blue). All genes encoding glycoside hydrolyses are shaded in Green. Genes with unknown functions are shaded in gray. **(B)** Time-series measurements of OD_{600} of PUL mutants and Δ*tryP-24* in media with glucose (top) or xyloglucan (bottom). Lines denote the mean and shaded regions represent 95% confidence interval of 3 biological replicates. **(C)** Bar plots of reads per kilobase per million mapped reads (RPKM) for each gene in *PUL11*, *PUL43* and *PUL44* in Δ*tryP-24* in media with glucose or xyloglucan. Data points denote 2 biological replicates. The colored bars represent mean RPKM value of the genes shown in the same order as panel **(A)**. **(D)** Bar plot of the log2 fold changes of RPKM of Δ*PUL11* to Δ*tryP-24* (left), Δ*PUL43* to Δ*tryP-24* (middle), or Δ*PUL11_44* to Δ*tryP-24* (right) in xyloglucan media (n=2, *p<0.05, unpaired t-test). **(E)** Bar plot of the fold change of *BACUNI_RS15340* in *PUL43* to Δ*tryP-24* in xyloglucan media using qRT-PCR (n = 6, ***p<0.001; Δ*PUL11* vs Δ*tryP-24* p=2.9e-5; Δ*PUL11_44* vs Δ*tryP-24* p=2.3e-5, unpaired t-test). Bars indicate the mean of replicates +/- 1 s.d. and data points show individual measurements. **(F)** Heatmap of log2 fold changes of RPKM for PUL genes in Δ*PUL11*, Δ*PUL43,* Δ*PUL44* or Δ*PUL11_44* compared to Δ*tryP-24* in xyloglucan media. The x-axis represents individual genes contained in the indicated PULs. The indicated PULs (x-axis) contained at least one gene with an absolute value of the log2 fold change greater than 2.

Figure 3. Impact of polysaccharide utilization loci (PULs) on *B. uniformis* pooled mutant fitness in media with different carbon sources. (A) Schematic representing experimental design of the pooled barcoded PUL mutants and Δ*tryP-24* in media with different carbon sources. The mutant pool was passaged every 24 hr using a 1:20 dilution. The absolute abundance of each mutant was determined by barcode sequencing and OD_{600} measurements. **(B)** Stacked bar plots represent the mean relative abundance of each mutant in media with a given carbon source (top). Bar plots denote the OD_{600} of each condition (bottom). Data points represent 3 biological replicates. **(C)** Violin plots represent the distribution of the absolute abundance of the PUL mutants in media with a given carbon source. Colored data points represent the absolute abundance of the indicated mutant within this distribution (n = 3 biological replicates). **(D)** Heatmap of Shannon diversity of mutant pool across different passages. The values represent the mean of 3 biological replicates.

Figure 4. Impact of polysaccharide utilization loci (PULs) on the colonization ability of *B. uniformis* in germ-free mice fed different diets. (A) Schematic of experimental design to evaluate the effect of PULs on BU colonization ability in germ-free mice fed different diets. Top: mice were fed a high fiber diet, fiber free diet (FFD) or high fat diet a week prior to oral gavage and then maintained on the same diet for two weeks following oral gavage (n = 5 for high fiber group and n = 4 for other groups). Bottom: mice were fed the FFD a week prior to oral gavage and then provided with drinking water supplemented with inulin, pectic galactan or glucomannan (n = 4). On day 0, mice were orally gavaged with the mutant pool and Δ*tryP-24*. Time-series measurements of fecal samples were performed. The cecal samples were collected at the end of the experiment. **(B)** Stacked bar plot of the absolute abundance of mutants in cecal samples (CFU g^{-1}) in each group of mice. Data points denote 2 independent CFU measurements for each mouse. The two most abundance mutants are highlighted. Asterisks denote a statistically significant difference in the CFU of each group compared to FFD group based on unpaired t-test (n = 8; *p<0.05, ***p<0.001; High fiber diet vs FFD, p=4.9e-9; High fat diet vs FFD, p=0.01271; FFD-Inulin vs FFD, p=2.8e-6; FFD-Glucomannan vs FFD, p=1.0e-4; FFD-Pectic galactan vs FFD, p=0.01393). **(C)** (i) Stacked bar plots of relative abundance of mutants in fecal samples as a function of time (top). Line plots of Shannon diversity of the mutant pool as a function of time (bottom). Asterisks indicate statistically significant difference of Shannon diversity of each group on day 14 compared to FFD group based on unpaired t-test (n=4-5; **p<0.01, ***p<0.001; High fiber diet vs FFD, p=3.9e-6; High fat diet vs FFD, p=1.4e-4; FFD-Glucomannan vs FFD, p=0.00425; FFD-Pectic galactan vs FFD, p=7.9e-5). (ii) Categorial scatter plot of the maximum slope of the Shannon diversity as a function of time. The colored data points represent individual mice and the black data point represents the mean. **(D)** Principal component analysis (PCA) as a function of time. Colors represent different diets based on the legend in **(A)**. The size of the data points is proportional to the time of measurement. The PCA loadings are denoted by the black lines. Symbols represent different mice. **(E)** Line plots of the relative abundance of PUL mutants or Δ*tryP-24* in mice fed different diets as a function of time. The colors represent different diets based on the legend in **(A)**. Data points denote individual mice and lines represent the mean. The asterisks denote a statistically significant difference in relative abundance between the FFD and each indicated diet based on unpaired t-test (n=4-5, p<0.05).

Figure 5. *B. uniformis* polysaccharide utilization loci (PULs) mediate glycan-dependent inter-species interactions influencing butyrate production. (A) Schematic representing the

experimental design investigating PUL mediated inter-species interactions between Inter-species interactions were deduced using a generalized Lotka-Volterra (gLV) model informed by time-series data of species absolute abundance. Representative data (middle) shows time-series measurements of absolute abundance based on 16S rDNA sequencing and OD_{600} measurements in inulin media (**Fig. S7**). Species include BU WT (blue), BU PUL mutant (purple) and butyrate producers (BPs): *A. caccae* (AC, orange), *C. comes* (CC, green), *E. rectale* (ER, yellow) and *R. intestinalis* (RI, red). The color of each growth curve indicates the species. The first row represents monoculture conditions and the second and third rows denote the growth of each species in the BU WT, butyrate producer (BP) or BU PUL mutant, BP coculture in inulin media. **(B)** Inferred networks of inter-species interactions between BU WT or a given PUL mutant and each BP in media with different carbon sources. Node size represents the maximum mean OD_{600} measured in monoculture in the indicated media. For species whose maximum mean OD_{600} was less than 0.5, their node sizes were set to OD_{600}=0.5 for visibility in the network. The width of an edge connecting node j to node i represents the magnitude of the median of the inferred marginal distribution of their inter-species interaction coefficient (a_{ij}). An edge is colored red (blue) if the interaction is positive (negative). If the 25% and the 75% quantiles of the a_{ij} marginal distribution have different signs, we represent the edge with a dashed line, indicating lack of certainty. Inter-species interactions where the magnitude of the median of the marginal a_{ij} distribution is less than 0.01 are not included in the network. **(C)** Scatter plot of predicted and measured butyrate concentrations in monoculture and coculture experiments. Predicted butyrate concentrations are computed according to the linear regression model (**Methods)**. Marker horizontal position represents predicted butyrate concentration based on mean end-point BP abundance, and horizontal error bars represent 1 s.d. of predicted butyrate concentration given the uncertainty in BP abundance measurements.

Figure 6. Mechanisms of PUL mediated interactions between *B. uniformis* (BU) and butyrate producers. (A) Schematic of Mechanism A: (i) butyrate producer (BP) such as *A. caccae* (AC) is unable to utilize the given glycan but can utilize PBPs released by BU. (ii) Deletion of a given PUL such as *PUL17* eliminates release of PBPs and thus AC is unable to grow in co-culture with Δ*PUL17* in inulin media. Bar plot of absolute abundance **(B)** and butyrate production **(C)** of BU WT, Δ*PUL17* and *A. caccae* (AC) monoculture and coculture in inulin media after 48 hr of growth. **(D)** Fructose concentrations in inulin or BU conditioned inulin media. **(E)** Growth of AC in glucose and fructose media. Lines denote the mean and the shaded regions represent 95% confidence interval. **(F)** Schematic of Mechanism B: (i) BP such as *R. intestinalis* (RI) can utilize

pullulan and thus compete with BU. (ii) RI can grow in co-culture with Δ*PUL18* in pullulan media**.** Bar plot of absolute abundance **(G)** and butyrate production **(H)** of BU WT, Δ*PUL18* and RI monoculture and coculture in pullulan media after 48 hr of growth. **(I)** Schematic of Mechanism C: (i) BP such as *C. comes* (CC) lacks the capability to utilize both the given glycan and PBPs. (ii) CC is unable to grow in coculture with Δ*PUL11_43* in xyloglucan media. Bar plot of absolute abundance **(J)** and butyrate production **(K)** of BU WT, Δ*PUL11_43* and CC monoculture and co-culture in xyloglucan media after 48 hr of growth. **(L)** Bar plot of the fold change of the area under curve (AUC) of CC in conditioned xyloglucan media and cell membrane treated xyloglucan media to AUC of CC in xyloglucan media. Bars indicate the mean of replicates ± 1 s.d. and data points show individual measurements (n>=3 biological replicates). PBPs: Polysaccharide breakdown products.

STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCE TABLE
- RESOURCE AVAILABILITY
    - Lead contact
    - Materials availability
    - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
    - Gnotobiotic mouse experiments
    - Microbial strains and growth conditions
- METHOD DETAILS
    - Plasmid construction
    - Conjugation
    - Characterization of shuttle plasmid stability
    - Markerless gene deletion and insertion in *Bacteroides*
    - Construction of barcoded BU strains
    - Luciferase assay to quantify gene expression in *Bacteroides*
    - Glycan utilization characterization of BU strains
    - Growth characterization of BU WT and Δ*tdk* with 5-fluorodeoxyuridine
    - *B. fragilis* Δ*xyl* xylose utilization assay
    - Characterization of *B. thetaiotamicron* Δ*levan* strain
    - Pooled barcoded BU mutant experiments
    - CFU counting for gnotobiotic mouse experiments
    - DNA extraction from fecal and cecal samples
    - Gene expression measurements based on qRT-PCR of representative BU PULs in germ-free mice
    - Butyrate producer community experiments and sample collection
    - Bacterial genome DNA extraction and next-generation sequencing
    - Characterization of butyrate producer growth in BU conditioned media

    - Characterization of butyrate producer growth in cell membrane treated glycan media
    - Measurements of fructose
    - Growth characterization of AC with fructose as the primary carbon source
- QUANTIFICATION AND STATISTICAL ANALYSIS
    - Bioinformatic analysis of species and barcoded strain abundances
    - Whole-genome transcriptional profiling of BU
    - Quantitative reverse transcription PCR (qRT-PCR) of *in vitro* cultures
    - Bioinformatic analysis of PULs in human gut microbiome metagenome datasets
    - Logistic growth and generalized Lotka-Volterra (gLV) models
    - Parameter inference for logistic growth and gLV models
    - Regression model for butyrate concentration

SUPPLEMENTARY DATA LEGENDS

Supplementary Data 1 Predicted PUL enzyme functions and the comparison of *B. uniformis* PULs characterized in this study with previous reported PULs in *Bacteroides*, related to Figure 1.

Supplementary Data 2 Inferred parameters of the logistic growth model and the generalized Lotka-Volterra model, related to Figure 1 and 5.

Supplementary Data 3 Reads per kilobase per million mapped reads (RPKM) of each transcript in *B. uniformis* in xyloglucan, type II mucin or glucose media based on RNA-seq, related to Figure 1 and 2

Supplementary Data 4 Detailed information about experiments including strains, plasmids, primers, sequences of genetic parts, smedia, and mouse diet compositions, related to Figures 1-6 in this study

Supplementary Data 5 Table of fold changes in the total growth of each butyrate producer modified by *B. uniformis* compared to the total butyrate producer growth in the absence of *B. uniformis* in media with individual glycans, related to Figures 5 and 6

REFERENCES

Almagro Armenteros, J.J., Tsirigos, K.D., Sonderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol *37*, 420-423.

Almeida, A., Mitchell, A.L., Boland, M., Forster, S.C., Gloor, G.B., Tarkowska, A., Lawley, T.D., and Finn, R.D. (2019). A new genomic blueprint of the human gut microbiota. Nature *568*, 499-504.

Almeida, A., Nayfach, S., Boland, M., Strozzi, F., Beracochea, M., Shi, Z.J., Pollard, K.S., Sakharova, E., Parks, D.H., and Hugenholtz, P. (2021). A unified catalog of 204,938 reference genomes from the human gut microbiome. Nat Biotechnol *39*, 105-114.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. http://wwwbioinformaticsbabrahamacuk/projects/fastqc.
Bacic, M.K., and Smith, C.J. (2008). Laboratory maintenance and cultivation of *Bacteroides* species. Curr Protoc Microbiol *9*, 13C-11.
Bågenholm, V., Reddy, S.K., Bouraoui, H., Morrill, J., Kulcinskaja, E., Bahr, C.M., Aurelius, O., Rogers, T., Xiao, Y., and Logan, D.T. (2017). Galactomannan catabolism conferred by a polysaccharide utilization locus of *Bacteroides ovatus*. J Biol Chem *292*, 229-243.
Baughn, A.D., and Malamy, M.H. (2002). A mitochondrial-like aconitase in the bacterium *Bacteroides fragilis*: implications for the evolution of the mitochondrial Krebs cycle. Proc Natl Acad Sci USA *99*, 4662-4667.
Benitez-Paez, A., Gomez Del Pulgar, E.M., and Sanz, Y. (2017). The glycolytic versatility of *Bacteroides uniformis* CECT 7771 and its genome response to oligo and polysaccharides. Front Cell Infect Microbiol *7*, 383.
Bhasin, M., Garg, A., and Raghava, G.P. (2005). PSLpred: prediction of subcellular localization of bacterial proteins. Bioinformatics *21*, 2522-2524.
Briliute, J., Urbanowicz, P.A., Luis, A.S., Basle, A., Paterson, N., Rebello, O., Hendel, J., Ndeh, D.A., Lowe, E.C., Martens, E.C.*, et al.* (2019). Complex N-glycan breakdown by gut *Bacteroides* involves an extensive enzymatic apparatus encoded by multiple co-regulated genetic loci. Nat Microbiol *4*, 1571-1581.
Buchfink, B., Xie, C., and Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. Nat Methods *12*, 59-60.
Bushnell, B. (2014). BBMap: A Fast, Accurate, Splice-Aware Aligner. sourceforgenet/projects/bbmap/.
Cantarel, B.L., Lombard, V., and Henrissat, B. (2012). Complex carbohydrate utilization by the healthy human microbiome. PloS one *7*, e28742.
Cartmell, A., Muñoz-Muñoz, J., Briggs, J.A., Ndeh, D.A., Lowe, E.C., Baslé, A., Terrapon, N., Stott, K., Heunis, T., and Gray, J. (2018). A surface endogalactanase in *Bacteroides thetaiotaomicron* confers keystone status for arabinogalactan degradation. Nat Microbiol *3*, 1314-1326.
Chaumeil, P.A., Mussig, A.J., Hugenholtz, P., and Parks, D.H. (2019). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics *36*, 1925-1927.
Chia, L.W., Mank, M., Blijenberg, B., Aalvink, S., Bongers, R.S., Stahl, B., Knol, J., and Belzer, C. (2020). *Bacteroides thetaiotaomicron* fosters the growth of butyrate-producing *Anaerostipes caccae* in the presence of lactose and total human milk carbohydrates. Microorganisms *8*, 1513.
Clark, R.L., Connors, B.M., Stevenson, D.M., Hromada, S.E., Hamilton, J.J., Amador-Noguez, D., and Venturelli, O.S. (2021). Design of synthetic human gut microbiome assembly and function. Nat Commun *12*, 1-16.
Cuskin, F., Lowe, E.C., Temple, M.J., Zhu, Y., Cameron, E.A., Pudlo, N.A., Porter, N.T., Urs, K., Thompson, A.J., and Cartmell, A. (2015). Human gut *Bacteroidetes* can utilize yeast mannan through a selfish mechanism. Nature *517*, 165-169.
Datsenko, and A., K. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci USA *97*, 6640-6645.
Déjean, G., Tamura, K., Cabrera, A., Jain, N., Pudlo, N.A., Pereira, G., Viborg, A.H., Van Petegem, F., Martens, E.C., and 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.
Déjean, G., Tauzin, A.S., Bennett, S.W., Creagh, A.L., and Brumer, H. (2019). Adaptation of syntenic xyloglucan utilization loci of human gut *Bacteroidetes* to polysaccharide side chain diversity. Appl Environ Microbiol *85*, e01491-01419.

Desai, M.S., Seekatz, A.M., Koropatkin, N.M., Kamada, N., Hickey, C.A., Wolter, M., Pudlo, N.A., Kitamoto, S., Terrapon, N., Muller, A.*, et al.* (2016). A dietary fiber-deprived gut microbiota degrades the colonic mucus barrier and enhances pathogen susceptibility. Cell *167*, 1339-1353.
Despres, J., Forano, E., Lepercq, P., Comtet-Marre, S., Jubelin, G., Chambon, C., Yeoman, C.J., Berg Miller, M.E., Fields, C.J., Martens, E.*, et al.* (2016). Xylan degradation by the human gut *Bacteroides xylanisolvens* XB1A(T) involves two distinct gene clusters that are linked at the transcriptional level. BMC Genomics *17*, 326.
Devlin, A.S., Marcobal, A., Dodd, D., Nayfach, S., Plummer, N., Meyer, T., Pollard, K.S., Sonnenburg, J.L., and Fischbach, M.A. (2016). Modulation of a Circulating Uremic Solute via Rational Genetic Manipulation of the Gut Microbiota. Cell Host Microbe *20*, 709-715.
Duncan, S.H., Belenguer, A., Holtrop, G., Johnstone, A.M., Flint, H.J., and Lobley, G.E. (2007). Reduced dietary intake of carbohydrates by obese subjects results in decreased concentrations of butyrate and butyrate-producing bacteria in feces. Appl Environ Microbiol *73*, 1073-1078.
Duncan, S.H., Holtrop, G., Lobley, G.E., Calder, A.G., Stewart, C.S., and Flint, H.J. (2004). Contribution of acetate to butyrate formation by human faecal bacteria. Br J Nutr *91*, 915-923.
El Kaoutari, A., Armougom, F., Gordon, J.I., Raoult, D., and Henrissat, B. (2013). The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. Nat Rev Microbiol *11*, 497-504.
Fan, Y., and Pedersen, O. (2021). Gut microbiota in human metabolic health and disease. Nat Rev Microbiol *19*, 55-71.
Fehlner-Peach, H., Magnabosco, C., Raghavan, V., Scher, J.U., Tett, A., Cox, L.M., Gottsegen, C., Watters, A., Wiltshire-Gordon, J.D., Segata, N.*, et al.* (2019). Distinct polysaccharide utilization profiles of human intestinal *Prevotella copri* Isolates. Cell Host Microbe *26*, 680-690.
Fischbach, M.A., and Sonnenburg, J.L. (2011). Eating for two: how metabolism establishes interspecies interactions in the gut. Cell Host Microbe *10*, 336-347.
Garcia-Bayona, L., and Comstock, L.E. (2019). Streamlined genetic manipulation of diverse *Bacteroides* and *Parabacteroides* isolates from the human gut microbiota. mBio *10*, e01762.
Goodman, A.L., Kallstrom, G., Faith, J.J., Reyes, A., Moore, A., Dantas, G., and Gordon, J.I. (2011). Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. Proc Natl Acad Sci USA *108*, 6252-6257.
Haario, H., Saksman, E., and Tamminen, J. (2001). An adaptive Metropolis algorithm. Bernoulli *7*, 223-242.
Hemsworth, G.R., Thompson, A.J., Stepper, J., Sobala, L.F., Coyle, T., Larsbrink, J., Spadiut, O., Goddard-Borger, E.D., Stubbs, K.A., Brumer, H.*, et al.* (2016). Structural dissection of a complex *Bacteroides ovatus* gene locus conferring xyloglucan metabolism in the human gut. Open Biol *6*, 160142.
Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform *11*, 119.
Jo Vandesompele, K.D.P., Filip Pattyn, Bruce Poppe, Nadine Van Roy, Anne De Paepe, Frank Speleman (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol *3*, 1-12.
Joglekar, P., Ding, H., Canales-Herrerias, P., Pasricha, P.J., Sonnenburg, J.L., and Peterson, D.A. (2019). Intestinal IgA regulates expression of a fructan polysaccharide utilization locus in colonizing gut commensal *Bacteroides thetaiotaomicron*. mBio *10*, e02324.
Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc *10*, 845-858.
Koh, A., De Vadder, F., Kovatcheva-Datchary, P., and Backhed, F. (2016). From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. Cell *165*, 1332-1345.
Koropatkin, N.M., Cameron, E.A., and Martens, E.C. (2012). How glycan metabolism shapes the human gut microbiota. Nat Rev Microbiol *10*, 323-335.

Koropatkin, N.M., Martens, E.C., Gordon, J.I., and Smith, T.J. (2008). Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. Structure *16*, 1105-1115.
Lapébie, P., Lombard, V., Drula, E., Terrapon, N., and Henrissat, B. (2019). *Bacteroidetes* use thousands of enzyme combinations to break down glycans. Nat Commun *10*, 2043.
Larsbrink, J., Rogers, T.E., Hemsworth, G.R., McKee, L.S., Tauzin, A.S., Spadiut, O., Klinter, S., Pudlo, N.A., Urs, K., Koropatkin, N.M.*, et al.* (2014). A discrete genetic locus confers xyloglucan metabolism in select human gut *Bacteroidetes*. Nature *506*, 498-502.
Li, M., Shang, Q., Li, G., Wang, X., and Yu, G. (2017). Degradation of marine algae-derived carbohydrates by *Bacteroidetes* isolated from human gut microbiota. Mar Drugs *15*, 92.
Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.
Lim, B., Zimmermann, M., Barry, N.A., and Goodman, A.L. (2017). Engineered regulatory systems modulate gene expression of human commensals in the gut. Cell *169*, 547-558.
Lloyd-Price, J., Abu-Ali, G., and Huttenhower, C. (2016). The healthy human microbiome. Genome Med *8*, 51.
Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.
Luis, A.S., Briggs, J., Zhang, X., Farnell, B., Ndeh, D., Labourel, A., Baslé, A., Cartmell, A., Terrapon, N., and Stott, K. (2018). Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic *Bacteroides*. Nat Microbiol *3*, 210-219.
Mahowald, M.A., Rey, F.E., Seedorf, H., Turnbaugh, P.J., Fulton, R.S., Wollam, A., Shah, N., Wang, C., Magrini, V., and Wilson, R.K. (2009). Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. Proc Natl Acad Sci USA *106*, 5859-5864.
Makki, K., Deehan, E.C., Walter, J., and Backhed, F. (2018). The impact of dietary fiber on gut microbiota in host health and disease. Cell Host Microbe *23*, 705-715.
Martens, E.C., Chiang, H.C., and Gordon, J.I. (2008). Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. Cell Host Microbe *4*, 447-457.
Martens, E.C., Koropatkin, N.M., Smith, T.J., and Gordon, J.I. (2009a). Complex glycan catabolism by the human gut microbiota: the *Bacteroidetes* Sus-like paradigm. J Biol Chem *284*, 24673-24677.
Martens, E.C., Lowe, E.C., Chiang, H., Pudlo, N.A., Wu, M., McNulty, N.P., Abbott, D.W., Henrissat, B., Gilbert, H.J., Bolam, D.N.*, et al.* (2011). Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. PLoS Biol *9*, e1001221.
Martens, E.C., Roth, R., Heuser, J.E., and Gordon, J.I. (2009b). Coordinate regulation of glycan degradation and polysaccharide capsule biosynthesis by a prominent human gut symbiont. J Biol Chem *284*, 18445-18457.
Mimee, M., Tucker, A.C., Voigt, C.A., and Lu, T.K. (2015). Programming a human commensal bacterium, *Bacteroides thetaiotaomicron*, to sense and respond to stimuli in the murine gut microbiota. Cell Syst *1*, 62-71.
Ndeh, D., Baslé, A., Strahl, H., Yates, E.A., McClurgg, U.L., Henrissat, B., Terrapon, N., and Cartmell, A. (2020). Metabolism of multiple glycosaminoglycans by *Bacteroides thetaiotaomicron* is orchestrated by a versatile core genetic locus. Nat Commun *11*, 1-12.
Ndeh, D., Rogowski, A., Cartmell, A., Luis, A.S., Baslé, A., Gray, J., Venditto, I., Briggs, J., Zhang, X., and Labourel, A. (2017). Complex pectin metabolism by gut bacteria reveals novel catalytic functions. Nature *544*, 65-70.
Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res *25*, 1043-1055.

Pasolli, E., Asnicar, F., Manara, S., Zolfo, M., Karcher, N., Armanini, F., Beghini, F., Manghi, P., Tett, A., Ghensi, P*., et al.* (2019). Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. Cell *176*, 649-662.
Pluvinage, B., Grondin, J.M., Amundsen, C., Klassen, L., Moote, P.E., Xiao, Y., Thomas, D., Pudlo, N.A., Anele, A., and Martens, E.C. (2018). Molecular basis of an agarose metabolic pathway acquired by a human intestinal symbiont. Nat Commun *9*, 1043.
Porter, N.T., and Martens, E.C. (2017). The critical roles of polysaccharides in gut microbial ecology and physiology. Annu Rev Microbiol *71*, 349-369.
Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., and Yamada, T. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. Nature *464*, 59-65.
Rakoff-Nahoum, S., Coyne, M.J., and Comstock, L.E. (2014). An ecological network of polysaccharide utilization among human intestinal symbionts. Curr Biol *24*, 40-49.
Rakoff-Nahoum, S., Foster, K.R., and Comstock, L.E. (2016). The evolution of cooperation within the gut microbiota. Nature *533*, 255-259.
Reddy, S.K., Bågenholm, V., Pudlo, N.A., Bouraoui, H., Koropatkin, N.M., Martens, E.C., and Stålbrand, H. (2016). A β-mannan utilization locus in *Bacteroides ovatus* involves a GH36 α-galactosidase active on galactomannans. FEBS Lett *590*, 2106-2118.
Rey, F.E., Faith, J.J., Bain, J., Muehlbauer, M.J., Stevens, R.D., Newgard, C.B., and Gordon, J.I. (2010). Dissecting the *in vivo* metabolic potential of two human gut acetogens. J Biol Chem *285*, 22082-22090.
Reyes, A., Wu, M., McNulty, N.P., Rohwer, F.L., and Gordon, J.I. (2013). Gnotobiotic mouse model of phage-bacterial host dynamics in the human gut. Proc Natl Acad Sci USA *110*, 20236-20241.
Rodriguez-Castano, G.P., Dorris, M.R., Liu, X., Bolling, B.W., Acosta-Gonzalez, A., and Rey, F.E. (2019). *Bacteroides thetaiotaomicron* starch utilization promotes quercetin degradation and butyrate production by *Eubacterium ramulus*. Front Microbiol *10*, 1145.
Rogowski, A., Briggs, J.A., Mortimer, J.C., Tryfona, T., Terrapon, N., Lowe, E.C., Basle, A., Morland, C., Day, A.M., Zheng, H*., et al.* (2015). Glycan complexity dictates microbial resource allocation in the large intestine. Nat Commun *6*, 7481.
Ryckman, A.E., Brockhausen, I., and Walia, J.S. (2020). Metabolism of glycosphingolipids and their role in the pathophysiology of lysosomal storage disorders. Int J Mol Sci *21*, 6881.
Schwalm, N.D., Townsend, G.E., and Groisman, E.A. (2016). Multiple signals govern utilization of a polysaccharide in the gut bacterium *Bacteroides thetaiotaomicron*. mBio *7*, e01342.
Sheridan, P.O., Martin, J.C., Lawley, T.D., Browne, H.P., Harris, H.M., Bernalier-Donadille, A., Duncan, S.H., O'Toole, P.W., Scott, K.P., and Flint, H.J. (2016). Polysaccharide utilization loci and nutritional specialization in a dominant group of butyrate-producing human colonic Firmicutes. Microb Genom *2*, e000043.
Shipman, J.A., Berleman, J.E., and Salyers, A.A. (2000). Characterization of four outer membrane proteins involved in binding starch to the cell surface of *Bacteroides thetaiotaomicron*. J Bacteriol *182*, 5365-5372.
Shipman, J.A., Cho, K.H., Siegel, H.A., and Salyers, A.A. (1999). Physiological characterization of SusG, an outer membrane protein essential for starch utilization by *Bacteroides thetaiotaomicron*. J Bacteriol *181*, 7206-7211.
Sonnenburg, E.D., Zheng, H., Joglekar, P., Higginbottom, S.K., Firbank, S.J., Bolam, D.N., and Sonnenburg, J.L. (2010). Specificity of polysaccharide use in intestinal *Bacteroides* species determines diet-induced microbiota alterations. Cell *141*, 1241-1252.
Sonnenburg, J.L., Xu, J., Leip, D.D., Chen, C.-H., Westover, B.P., Weatherford, J., Buhler, J.D., and Gordon, J.I. (2005). Glycan foraging *in vivo* by an intestine-adapted bacterial symbiont. Science *307*, 1955-1959.

Spence, C., Wells, W.G., and Smith, C.J. (2006). Characterization of the primary starch utilization operon in the obligate anaerobe *Bacteroides fragilis*: regulation by carbon source and oxygen. J Bacteriol *188*, 4663-4672.
Tagawa, J., Inoue, T., Naito, M., Sato, K., Kuwahara, T., Nakayama, M., Nakayama, K., Yamashiro, T., and Ohara, N. (2014). Development of a novel plasmid vector pTIO-1 adapted for electrotransformation of *Porphyromonas gingivalis*. J Microbiol Methods *105*, 174-179.
Tancula, E., Feldhaus, M., Bedzyk, L., and Salyers, A. (1992). Location and characterization of genes involved in binding of starch to the surface of *Bacteroides thetaiotaomicron*. J Bacteriol *174*, 5609-5616.
Tauzin, A.S., Kwiatkowski, K.J., Orlovsky, N.I., Smith, C.J., Creagh, A.L., Haynes, C.A., Wawrzak, Z., Brumer, H., Koropatkin, N.M., Sonnenburg, J.*, et al.* (2016). Molecular dissection of xyloglucan recognition in a prominent human gut symbiont. mBio *7*, e02134-02115.
Terrapon, N., Lombard, V., Drula, E., Lapébie, P., Al-Masaudi, S., Gilbert, H.J., and Henrissat, B. (2017). PULDB: the expanded database of polysaccharide utilization Loci. Nucleic Acids Res *46*, D677-D683.
Townsend, G.E., Han, W., Schwalm, N.D., Hong, X., Bencivenga-Barry, N.A., Goodman, A.L., and Groisman, E.A. (2020). A master regulator of *Bacteroides thetaiotaomicron* gut colonization controls carbohydrate utilization and an alternative protein synthesis factor. mBio *11*, e03221.
Tuncil, Y.E., Xiao, Y., Porter, N.T., Reuhs, B.L., Martens, E.C., and Hamaker, B.R. (2017). Reciprocal prioritization to dietary glycans by gut bacteria in a competitive environment promotes stable coexistence. mBio *8*, e01068.
Venturelli, O.S., Carr, A.V., Fisher, G., Hsu, R.H., Lau, R., Bowen, B.P., Hromada, S., Northen, T., and Arkin, A.P. (2018). Deciphering microbial interactions in synthetic human gut microbiome communities. Mol Syst Biol *14*, e8157.
Venturelli, O.S., Zuleta, I., Murray, R.M., and El-Samad, H. (2015). Population diversification in a yeast metabolic program promotes anticipation of environmental shifts. PLoS Biol *13*, e1002042.
Wang, J., Shoemaker, N.B., Wang, G.-R., and Salyers, A.A. (2000). Characterization of a *Bacteroides* mobilizable transposon, NBU2, which carries a functional lincomycin resistance gene. J Bacteriol *182*, 3559-3571.
Wexler, A.G., and Goodman, A.L. (2017). An insider's perspective: *Bacteroides* as a window into the microbiome. Nat Microbiol *2*, 17026.
Whitaker, W.R., Shepherd, E.S., and Sonnenburg, J.L. (2017). Tunable expression tools enable single-cell strain distinction in the gut microbiome. Cell *169*, 538-546.
Yu, C.S., Cheng, C.W., Su, W.C., Chang, K.C., Huang, S.W., Hwang, J.K., and Lu, C.H. (2014). CELLO2GO: a web server for protein subCELlular LOcalization prediction with functional gene ontology annotation. PloS one *9*, e99368.
Yu, N.Y., Wagner, J.R., Laird, M.R., Melli, G., Rey, S., Lo, R., Dao, P., Sahinalp, S.C., Ester, M., Foster, L.J.*, et al.* (2010). PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics *26*, 1608-1615.
Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., Van Der Oost, J., and Regev, A. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell *163*, 759-771.
Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z., Busk, P.K., Xu, Y., and Yin, Y. (2018). dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res *46*, W95-W101.
Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics *30*, 614-620.

Zhao, S., Lieberman, T.D., Poyet, M., Kauffman, K.M., Gibbons, S.M., Groussin, M., Xavier, R.J., and Alm, E.J. (2019). Adaptive evolution within gut microbiomes of healthy people. Cell Host Microbe *25*, 656-667.
Zhu, A., Ibrahim, J.G., and Love, M.I. (2019). Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. Bioinformatics *35*, 2084-2092.

STAR METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial strains		
All bacteria strains are listed in **Supplementary Data 4**	N/A	N/A
Oligonucleotides		
All the oligonucleotides are listed in **Supplementary Data 4**	N/A	
Recombinant DNA		
All Recombinant DNA (plasmids) are listed in **Supplementary Data 4**	N/A	N/A
Experimental Models: Organisms/Strains		
Mus musculus C57BL/6	Germ-free mice from the gnotobiotic mouse facility at the University of Wisconsin-Madison	N/A
Chemicals		
Calcium chloride	Sigma-Aldrich	Cat# C5670-100G
Copper(II) sulfate	Sigma-Aldrich	Cat# 209198-100G
Pyridoxal hydrochloride	Sigma-Aldrich	Cat# P6155-5G
Thymidine	Sigma-Aldrich	Cat# T1895-1G
Xanthine sodium salt	Sigma-Aldrich	Cat# X3627-1G
Folic Acid	Alfa Aesar	Cat# J62937-06
Orotic Acid potassium salt	Sigma-Aldrich	Cat# O2875-100G
Ca-D-pantothenate	TCI	Cat# P0012

Cobalamin (B12)	TCI	Cat# C0449
Pyridoxine HCl	DOT Scientific	Cat# DSP50240-50
Riboflavin	DOT Scientific	Cat# DSR64040-100
Tetrahydrofolic Acid	Cayman Chemical	Cat# 18263
Thiamine HCl	Sigma-Aldrich	Cat# T1270-25G
p-Aminobenzoic Acid	DOT Scientific Inc	Cat# DSA20050-25
Pyridoxamine	Chem Impex	Cat# 01461
Ammonium Chloride	Fisher Scientific	Cat# A661-500
Magnesium Chloride	Fisher Scientific	Cat# BP214-500
Potassium phosphate monobasic	Alfa Aesar	Cat# 11594-1KG
Nicotinamide	Sigma-Aldrich	Cat# N0636-100G
Haemin	Sigma-Aldrich	Cat# H9039-1G
Biotin	DOT Scientific	Cat# DSB40040-1
L-alanine	DOT Scientific	Cat# DSA20060-100
L-arginine	DOT Scientific	Cat# DSA50010-1000
L-asparagine	DOT Scientific	Cat# DSA50030-100
L-Aspartic acid	DOT Scientific	Cat# DSA50060-100
L-Cysteine	DOT Scientific	Cat# DSC81020-500
L-Glutamic acid	Sigma-Aldrich	Cat# G1251-500G
L-glutamine	Acros Organics	Cat# 119951000
L-Glycine	Acros Organics	Cat# 120070010
L-histidine	VWR	Cat# 1B1164-100G
L-isoleucine	DOT Scientific	Cat# DSI54020-25
L-leucine	DOT Scientific	Cat# DSL22000-500
L-lysine	DOT Scientific	Cat# DSL37040-500
L-methionine	DOT Scientific	Cat# DSM22060-100
L-phenylalanine	DOT Scientific	Cat# DSP2060-100
L-proline	DOT Scientific	Cat# DSP50200-100
L-serine	DOT Scientific	Cat# DSS22020-100

L-threonine	DOT Scientific	Cat# DST21060-100
L-tryptophan	DOT Scientific	Cat# DST60080-25
L-valine	DOT Scientific	Cat# DSV42020-100
L-tyrosine	Sigma-Aldrich	Cat# T3754-50G
MOPS	Fisher Scientific	Cat# BP308-500
Sodium Bicarbonate	Sigma-Aldrich	Cat# S5761-500G
Glucose	VWR	Cat# BDH9230-500G
Xylose	MP Biomedicals	Cat# 103300
Fructose	Sigma-Aldrich	Cat# F0127-100G
Sodium chloride	DOT Scientific	Cat# DSS23020-1000
EDTA	Sigma-Aldrich	Cat# EDS-100G
10x AUGC Solution	Fisher Scientific	Cat# NC9659158
ATCC Mineral Supplement	ATCC	Cat# MD-TMS
Tris·HCl pH 8	Research Products International	Cat# T60050-1000.0
Sodium n-dodecyl sulfate	Alfa Aesar	Cat# J64241
Phenol/chloroform/isoamyl alcohol	Invitrogen	Cat# 15593049
Isopropyl β-D-1-thiogalactopyranoside	Gold Biotechnology	Cat# I2481C50
2,6-Diaminopimelic acid	Sigma-Aldrich	Cat# D1377-5G
Sodium acetate	Sigma-Aldrich	Cat# S5636-250G
5-Fluorodeoxyuridine	Sigma-Aldrich	Cat# F0503-100MG
PBS	MP Biomedicals	Cat# PBS10X02
Isopropanol	LabChem	Cat# LC157504
Ethanol	Koptec	Cat# 2716
Sulfuric acid	Sigma-Aldrich	Cat# 258105-2.5L
Molecular grade water	VWR	Cat# VWRL0201-0500
Potassium hydroxide	Sigma-Aldrich	Cat# 06005-1KG
Sodium propionate	Sigma-Aldrich	Cat# P1880-100G

Sodium succinate dibasic hexahydrate	Sigma-Aldrich	Cat# S2378-100G
Pefabloc SC	Sigma-Aldrich	Cat# 11429868001
Nitrogen	Airgas	Cat# NI200
Mixed gas	Airgas	Cat# X03NI75C30057W3
Anaerobe Basal Broth	Oxoid	Cat# CM0957
Brain Heart Infusion Broth	Sigma-Aldrich	Cat# 53286-500G
Luria Bertoni Broth	Sigma-Aldrich	Cat# L3022-1KG
Carbenicillin	IBI Scientific	Cat# IB02025
Erythromycin	Sigma-Aldrich	Cat# E5389-5G
Gentamicin	Sigma-Aldrich	Cat# G1264-5G
AE buffer	Qiagen	Cat# 19077
Phusion High-Fidelity DNA Polymerase	Fisher Scientific	Cat# F530L
BioMix™ Red Polymerase Mix	Bioline	Cat# BIO-25006
Kapa HiFi HotStart ReadyMix	Roche	Cat# 07958960001
SYBR Green PCR Master Mix	Fisher Scientific	Cat# 4309155
Not I	NEB	Cat# R3189S
Nsi I	NEB	Cat# R0127S
Kpn I	NEB	Cat# R3142S
BamH I	NEB	Cat# R3136S
Sal I	NEB	Cat# R3138S
rLysozyme™ Solution	EMD Millipore	Cat# 71110-6000ku
DNase I	Thermo Fisher Scientific	Cat# EN0521
DNase I	Sigma-Aldrich	Cat# 11284932001
Acid-Phenol:Chloroform, pH 4.5 (with IAA, 125:24:1)	Thermo Fisher Scientific	Cat# Ambion 9720
Baseline-ZERO™ DNase	Epicentre	Cat# DB0715K
Glass beads, acid-washed	Sigma-Aldrich	Cat# G1277
Arabinogalactan	Sigma-Aldrich	Cat# 10830-25G
Arabinoxylan	Megazyme	Cat# P-WAXYL

Amylopectin (maize)	Sigma-Aldrich	Cat# 10120-250G
Agarose	Sigma-Aldrich	Cat# A9539-10G
Chondroitin sulfate	Sigma-Aldrich	Cat# C9819-5G
Dextran (Leuconostoc spp)	Sigma-Aldrich	Cat# 31392-10G
Galactan (potato)	Sigma-Aldrich	Cat# 50009-1G
Galactomannan (carob)	Sigma-Aldrich	Cat# 91783-1G
Glucomannan	Megazyme	Cat# P-GLCML
Glycogen	Sigma-Aldrich	Cat# G8751-5G
Gum Arabic (acacia tree)	Sigma-Aldrich	Cat# G9752-500G
Inulin (chicory)	Sigma-Aldrich	Cat# I2255-10G
Inulin (chicory)	Acros Organics	Cat# AC457100250
Laminarin (Laminaria digitata)	Sigma-Aldrich	Cat# L9634-1G
Lichenin (Icelandic Moss)	Megazyme	Cat# P-LICHN
Levan	Megazyme	Cat# P-LEVAN
Pectic galactan (Potato)	Megazyme	Cat# P-PGAPT
Pectin (Citrus)	Gojira Fine Chemicals	Cat# PE1006
Pullulan (Aureobasidium pullulans)	Sigma-Aldrich	Cat# P4516-5G
Starch	Sigma-Aldrich	Cat# S9765-100G
Type II mucin (porcine stomach)	Sigma-Aldrich	Cat# M2378-100G
Type III mucin (porcine stomach)	Sigma-Aldrich	Cat# M1778-100G
Xylan (Beechwood)	Megazyme	Cat# P-XYLNBE-10G
Xyloglucan (Tamarind)	Megazyme	Cat# P-XYGLN
Critical commercial assays		
Nano-Glo Luciferase Assay System kit	Promega	Cat# N1110
RNeasy Mini Kit	Qiagen	Cat# 74124
Zymo DNA Clean & Concentrator™-5 kit	Zymo Research	Cat# D4013
Qiagen DNeasy Blood and Tissue Kit	Qiagen	Cat# 69506

RNase-Free DNase Set	Qiagen	Cat# 79254
RNAprotect Bacteria Reagent	Qiagen	Cat# 76506
Agilent High Sensitivity RNA ScreenTape	Agilent	Cat# 5067-5579
BugBuster	EMD Millipore	Cat# 70921-4
Quant-iT™ dsDNA Assay Kit	Invitrogen	Cat# Q33120
MiSeq Reagent Kit v3	Illumina	Cat# MS-102-3003
MiSeq Reagent Nano Kit v2	Illumina	Cat# MS-103-1003
iScript Select cDNA Synthesis Kit	Bio-Rad Laboratories	Cat# 1708897
High Sensitivity Fructose Assay Kit	Sigma-Aldrich	Cat# MAK180-1KT
Software and Algorithms		
FastQC tool v0.11.8	Babraham Bioinformatics	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
BBTools suite v38.42	Sourceforge	https://sourceforge.net/projects/bbmap/
FeatureCounts package v1.6.4	Liao et al., 2013	PMID: 24227677
DESeq2 Bioconductor library v4.0.3	Love et al., 2014	PMID: 25516281
R v4.0.4	R Core Team	www.r-project.org
apeglm shrinkage estimator	Zhu et al., 2019	PMID: 30395178
PEAR (Paired-End reAd mergeR) v0.9.10	Zhang et al., 2014	PMID: 24142950
Prodigal v2.6.3	Hyatt et al., 2010	PMID: 20211023
DIAMOND BLASTP v0.9.28.129	Buchfink et al., 2015	PMID: 25402007
CheckM v1.0.11	Parks et al., 2015	PMID: 25977477
GTDB-Tk v0.1.3	Chaumeil et al., 2019	PMID: 31730192
MCMC algorithm	Haario et al., 2001	https://projecteuclid.org/journals/bernoulli/volume-7/issue-2/An-adaptive-Metropolis-algorithm/bj/1080222083.full

SignalP 5.0	Armenteros et al., 2019	PMID: 30778233
CELLO2GO	Yu et al., 2014	PMID: 24911789
PSORTb 3.0	Yu et al., 2010	PMID: 20472543
PSLpred	Bhasin et al., 2005	PMID: 15699023
Phyre2	Kelley et al., 2015	PMID: 25950237
Python 3.7	Python	https://www.python.org/downloads/release/python-370/
MATLAB R2016a	MathWorks	https://www.mathworks.com/help/images/release-notes-R2016a.html
Deposited Data		
Pooled BU mutant or 16S rRNA gene sequencing data of BU and individual butyrate producers	This paper; Mendeley Data	doi: 10.17632/3888sct8c5.1
RNA-seq data of BU in media with xyloglucan or glucose	This paper; Zenodo	doi:10.5281/zenodo.5204678
RNA-seq data of BU in media with type II mucin or glucose	This paper; Zenodo	doi:10.5281/zenodo.5207518
Other		
Chow diet	Purina	Cat# LabDiet 5021
High fat diet	Envigo	Cat# TD.88137
Fiber free diet	Envigo	Cat# TD.190849

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ophelia S. Venturelli (venturelli@wisc.edu).

Materials availability

Plasmids and bacterial strains generated in this study are available upon request from the Lead Contact.

Data and code availability

- The RNA-seq data of BU in media with xyloglucan or glucose are available from Zenodo (doi:10.5281/zenodo.5204678). The RNA-seq data of BU in media with type II mucin or glucose are available from Zenodo (doi:10.5281/zenodo.5207518).
- Code for analyses of the pooled BU mutant experiments, raw sequencing data for the pooled BU mutant experiments and the co-cultures of BU and individual butyrate producers are deposited at Mendeley Data (doi: 10.17632/3888sct8c5.1). Code for the analysis of RNA-seq data and computational modeling are available from Zenodo (doi:10.5281/zenodo.5722360)
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Gnotobiotic mouse experiments

All germ-free mouse experiments were performed following protocols approved by the University of Wisconsin-Madison Animal Care and Use Committee. Three diets were used in this experiment: high fiber diet (Chow diet, Purina, LabDiet 5021), high fat diet (Envigo, TD.88137) and fiber free diet (Envigo, TD.190849) (**Supplementary Data 4**). Note that the high fiber diet contains higher fiber compared to the High fat diet or Fiber free diet and was thus referred to as high fiber diet. The barcoded BU control strain Δ*tryP-24* and PUL deletion mutants were grown at 37 ºC anaerobically in Anaerobe Basal Broth (ABB, Oxoid) for 12-16 hr. Cultures were diluted by 20-fold into ABB media and then grown at 37 ºC anaerobically until the culture reached exponential phase (OD_{600} ~1.0). All strains were mixed in equal proportions based on OD_{600} and transferred to Hungate tubes (Chemglass) on ice prior to oral gavage. We used 8-week old C57BL/6 gnotobiotic male mice (wild-type) fed the specific diets a week prior to oral gavage (**Figure 4A**). At this time, 0.2 mL of mutant pool was introduced into the mice by oral gavage inside a Biological Safety Cabinet (BSC) and the mice were housed in biocontainment cages (Allentown Inc.) for the duration of the experiment. Mice were maintained on the same experimental diets with autoclaved water or water supplemented with glycans for two weeks after colonization. The concentration of inulin and pectic galactan in the drinking water was 5 g L^{-1}. 2 g L^{-1} of glucomannan solution were prepared and autoclaved. The autoclaved solution was centrifuged at 3,200 x g for 5 min, and the supernatant of the glucomannan solution was used for the mouse experiments. Groups of mice (4-5 mice) fed a given diet were co-housed in a single

cage. Fecal samples were collected every 2-3 days after oral gavage. At the end of the experiment, mice were euthanized, and the cecal samples were collected for NGS sequencing and CFU plating.

To determine the fold changes in expression of representative genes in key PULs in BU in the mammalian gut compared to glucose media, 6-week old C57BL/6 gnotobiotic male mice (wild-type) were fed the high fiber diet a week prior to oral gavage (**Figure 4A**). The BU WT was grown at anaerobically at 37 ºC in DM35 media overnight. We introduced 0.2 mL of BU WT overnight culture into the mice by oral gavage inside a BSC and the mice were co-housed in a biocontainment cage. Mice were maintained on high fiber diet for a week and then euthanized and the ceca were collected for RNA extraction.

Microbial strains and growth conditions

Detailed information of the strains used in this study is provided in **Supplementary Data 4**. All anaerobes were cultured in an anaerobic chamber with an atmosphere of 83% N_2, 2% H_2 and 15% CO_2 at 37 ºC. For most experiments, *Bacteroides* strains, AC, CC were grown in ABB media, ER was grown in ABB media with the addition of 3.3 mM sodium acetate (Sigma) and RI was grown in Brain Heart Infusion Broth (BHI, Sigma). We used *E. coli* pir2 (Invitrogen) for cloning and maintenance of plasmids with R6K origin (pNBU2-ermGb derivatives and pFW1000 derivatives, **Supplementary Data 4**). *E. coli* DH5α (Thermo Fisher Scientific) was used for cloning and maintenance of plasmids with p15A, pSC101ts and ColE1 origins (pFW2000 derivatives, pFW3000 and pFW4000). We used *E. coli* BW29427 (*E. coli* Genetic Stock Center, CGSC) for *E. coli-Bacteroides* conjugations. All *E. coli* strains were grown aerobically in Luria Bertoni (LB, Sigma) media. To support the growth of *E. coli* BW29427, we supplemented LB media with 25 µM of 2,6-Diaminopimelic acid (DAP, Sigma). We used the following antibiotics when required including 100 µg mL^{-1} carbenicillin (Carb, IBI Scientific), 25 µg mL^{-1} erythromycin (Erm, Sigma) and 200 µg mL^{-1} gentamicin (Gm, Sigma). We used 1 mM of Isopropyl β-D-1-thiogalactopyranoside (IPTG, Gold Biotechnology) for the induction of FnCpf1.

METHOD DETAILS

Plasmid construction

All the plasmids and primers used in this work are described in **Supplementary Data 4** and sequences of genetic elements are listed in **Supplementary Data 4**. The P_BfP1E6 promoter fused to 18 previously reported RBSs were cloned into pNBU2-ermGb at the Not I restriction site

with NanoLuc gene to characterize the strength of each RBS in BU (Koropatkin et al., 2008; Wang et al., 2000; Whitaker et al., 2017). We constructed plasmids with 20 BU native promoters as well as three previously reported strong promoters in BT (P_{BT1311}, P_{cfiA} and P_BfP1E6) by cloning each promoter with NanoLuc gene into pNBU2-ermGb using the Not I restriction site. In addition, RBS8 was fused to the promoters $P_{BU18065}$, $P_{BU18270}$, $P_{BU15675}$ and P_{BfP1E6}. The P_{BT1311}-lacI and P_{lacO23} were amplified from pMM773 (Mimee et al., 2015), and inserted into pNBU2-ermGb with NanoLuc gene using the Pst I and Not I sites, respective, generating pFW001. The P_{BT1311} promoter controlling *rhaR* regulator from BT was amplified and inserted into pMM656 (Mimee et al., 2015) at Not I site, generating pFW023. The NanoLuc gene was inserted into pNBU2_erm-TetR-P1T_DP-GH023 (Lim et al., 2017) at BamH I site, generating pFW064.

To construct the shuttle plasmid pFW1000, the *Bacteroides* replication origin PB8-51(Tagawa et al., 2014) (synthesized by Twist Biosciences) was cloned using the Nsi I and Kpn I restriction sites in pNBU2-ermGb by replacing the NBU2 integrase gene and *attN2* site. To generate the shuttle plasmids pFW2000, pFW3000 and pFW4000, the R6K origin on pFW1000 was replaced by the *E. coli* origins p15A, pSC101ts or ColE1, respectively.

The three plasmids pFW1004, pFW2100 and pFW2500 were used for genetic manipulation in BU. The P_{BT1311}-*lacI*-P_{lacO23} sequence amplified from pFW025 was cloned into pFW1000 and pFW2000 using the BamH I site, generating pFW1001 and pFW2001, respectively. Subsequently, the *Fncpf1* gene amplified from pT7FnCpf1 was cloned into pFW1001 and pFW2001 using the BamH I site, generating pFW1004 and pFW2100, respectively. Next, the IPTG inducible promoter P_{lacO23}(Mimee et al., 2015) and *recT* cloned from *E. coli* DH5α were cloned into pFW2100 using the Sal I site, generating plasmid pFW2500.

To construct plasmids for deletion of genes or PULs, the promoters P_{BT1311} or P_{BfP1E6} controlling crRNA, *rrnB* T2 terminator as well as two homologous arms (~500-bp for single gene deletion and ~1000-bp for PUL deletion) were cloned into pFW1004, pFW2100 or pFW2500 using the BamH I site, yielding a set of plasmids for targeted gene or PUL deletions. For plasmids used for *rhaR* modification, the promoter P_{BT1311} controlling the crRNA, *rrnB* T2 terminator as well as two ~800-bp homologous arms containing the gene fragment for modification were cloned into pFW2100, yielding the final plasmids pFW2028. The gene fragment replacement (for the generation of *rhaR'*) of inactive *rhaR* (BU) gene was designed based on the RhaR sequence from WP_005834782.1. For *gfp* gene insertion, the P_BfP1E6 promoter controlling *gfp* amplified from pWW3452 (Whitaker et al., 2017) was inserted into pFW2033 (*PUL13* deletion plasmid) at Afl II site, yielding the final plasmid pFW2059.

Conjugation

All plasmids (pNBU2-ermGb derivatives and shuttle plasmids derivatives) were introduced into *Bacteroides* via conjugation with *E. coli* BW29427, which harbors the conjugative machinery integrated into the chromosome(Datsenko and A., 2000). To perform the conjugation, *E. coli* BW29427 was grown until early stationary phase. Next, cell pellets were collected by centrifugation at 3,200 x g for 5 min and washed once with fresh LB media. The washed cell pellets were combined with *Bacteroides* cultures (OD_{600} approximately equal to 0.5-0.6) at a ratio of 1:10 (donor:recipient, v/v). The mixed culture was pelleted and resuspended in 0.2 mL BHI media and then spotted onto BHIAD (BHI + 10%ABB + DAP) plates prior to anaerobic incubation at 37 ºC for 24 hr. Next, we collected the cells from the plate and resuspended into 1 mL BHI. The culture was plated on BHIAEG (BHI + 10%ABB + Erm + Gm) plates using a range of dilutions and anaerobically incubated at 37 ºC for 2 days. The pNBU2-ermGb plasmid derivatives harbor the IntN2 tyrosine integrase, which mediates the recombination between the *attN* site on the plasmid and one of the two *attBT* sites at the 3' end of the $tRNA^{Ser}$ gene (*BACUNI_RS18270* and *BACUNI_RS18350*) on the BU chromosome(Wang et al., 2000). Thus, all pNBU2-ermGb plasmid derivatives were integrated onto the chromosome following conjugation (Mimee et al., 2015). As all the shuttle plasmids derivatives contain the *Bacteroides* replication origin pB8-51 and lack the *intN2* gene, these plasmids may continue to be replicated and potentially maintained over time.

Characterization of shuttle plasmid stability

The BU strains harboring shuttle plasmids pFW1000, pFW2000, pFW3000 and pFW4000 were first grown at 37 ºC in ABB media with erythromycin for 12-16 hr. Next, these strains were diluted 20-fold into fresh ABB media lacking antibiotics every 12 hr for 15 passages. The same volume of the diluted cell cultures was plated on BHIA (BHI + 10% ABB) and BHIAE (BHI + 10% ABB + Erm) plates to determine the number of colonies containing the plasmid compared to the total number of colonies per unit volume. We evaluated plasmid stability based on the ratio of CFU on antibiotic selective plates to the total number of CFU on plates without antibiotic selection. Colony forming units (CFU) were determined at the following passages: 1, 3, 5, 7, 9, 11,13 and 15.

Markerless gene deletion and insertion in *Bacteroides*

The plasmids pFW1004, pFW2100 and pFW2500 were used as the original vector for genome editing (**Figure S2A; Supplementary Data 4**). Plasmids for gene deletion and insertion were first transformed into *E. coli* BW29427 and then introduced into the *Bacteroides* strains using conjugation. The transconjugants were then grown anaerobically in ABB media supplemented

with Erm at 37 ºC for 12-16 hr. The cell cultures were diluted 10-fold and plated on BHIAEI (BHI + 10% ABB + Erm + IPTG) plates. The plates were anaerobically incubated at 37 ºC for 2 days until the colonies were observed. We performed colony PCR using BioMix Red PCR mix (Bioline) to screen for genome modifications. Colonies with the correct genome modifications based on the colony PCR results were anaerobically cultured in ABB media and passaged three times every 12-16 h with a dilution factor of 20-fold. After this period, the cell cultures were diluted 10^{-3} to 10^{-4} and plated on BHIA agar plates and incubated at 37 ºC for 1-2 days until the colonies were observed. Next, single colonies were picked and streaked on BHIA and BHIAE agar plates for replica plating. Colonies that could only grow on BHIA plates were selected as the final mutants that had lost the plasmids. We evaluated the efficiency of genome modification by dividing the number of correct mutant colonies by the total number of tested colonies.

Construction of barcoded BU strains

To distinguish the PUL deletion mutants in the mutant pool, randomly generated DNA barcodes (4 bp) were introduced onto the chromosome of BU prior to gene deletion. We constructed Δ*PUL12*, Δ*PUL34* and Δ*PUL49* by first deleting the given PUL and then introducing the barcode into the PUL deletion background. In other cases, the barcoded strains were generated by introducing the 4 bp barcode onto the chromosome while simultaneously deleting the tryptophanase gene (*tryP*) (**Figure 1A**). To this end, a library of pFW2026 plasmids were constructed and introduced into the BU WT via conjugation. The barcoded plasmids were sequenced using Sanger Sequencing (Functional Biosciences) prior to gene deletion. The final set of *tryP* deletion mutants were used as the barcode-tagged strains in this study.

Luciferase assay to quantify gene expression in *Bacteroides*

All NanoLuc luciferase assays to quantify gene expression were performed with cell lysate according to the procedure described in the Nano-Glo Luciferase Assay System kit (Promega). For these experiments, *Bacteroides* strains were first anaerobically cultured at 37 ºC in ABB media for 12-16 hr. Next, cultures were inoculated into BHI media and incubated at 37 ºC anaerobically until the OD_{600} reached ~0.6. To characterize the strengths of the RBS and promoter sequences, 4 mL of the cultures was centrifuged at 13,800 x g for 3 min. Cell pellets were lysed by resuspending into 400 µL of 1X BugBuster (EMD Millipore) in 1X PBS (MP Biomedicals). To characterize the inducible promoters, 20 mL of cultures were harvested via centrifugation at 3,200 x g for 10 min. Cell pellets were resuspended into 400 µL of 1X BugBuster in 1X PBS containing 0.5 µL of rLysozyme™ Solution (EMD Millipore). Next, 10 µL of the cell lysate was mixed with an

equal volume of NanoLuc Reaction buffer containing the substrate. The relative light units (RLU) were measured using a plate reader (Spark 10M, Tecan). The luminescence value was normalized to the OD_{600} value of the culture prior to lysis.

Characterization of BU growth in media with individual glycans

The BU WT strain was incubated in ABB media at 37 ºC anaerobically without shaking for 12-16 hr and then inoculated into ABB media again and incubated at 37 ºC anaerobically until the culture reached exponential phase (OD_{600} ~1.0). Next, the cell pellets were collected by centrifugation at 3,200 x g for 10 min, and then washed with BMM-C media (*Bacteroides* minimal media (Bacic and Smith, 2008) (BMM) without glucose, **Supplementary Data 4**). The washed cell pellets were resuspended into BMM-C media to a final OD_{600} of approximately 1. These cultures were inoculated into a 96-well plate (Greiner Bio-One) containing BMM-glycan (BMM-C media supplemented with a given glycan, **Supplementary Data 4**) to an initial OD_{600} of 0.05. These plates were incubated at 37 ºC anaerobically. Cell growth determined by OD_{600} was monitored using a Tecan F200 plate reader every 12 hr or 30 min depending on the experimental design. All the glycans used in this study are listed in **Supplementary Data 4**.

Growth characterization of BU WT and BU Δ*tdk* with 5-fluorodeoxyuridine

The BU WT and Δ*tdk* strains were grown in ABB media at 37ºC anaerobically for 12-16 hours. Next, cultures were diluted by 20-fold into ABB media and then incubated at 37 ºC anaerobically until the culture reached exponential phase (OD_{600} ~1.0). Next, the cultures were diluted into ABB and ABB containing 200 µg/mL 5-fluorodeoxyuridine (FudR, Sigma) to an initial OD_{600} of 0.025. The growth of the strains was monitored based on OD_{600} on a Tecan F200 plate reader every 30 min for 24 hr.

***B. fragilis* Δ*xyl* xylose utilization assay**

The *B. fragilis* DSM 2151 wild-type and *B. fragilis* Δ*xyl* strains were grown at 37ºC anaerobically in ABB media for 12-16 hr. The cultures were then diluted by 20-fold into ABB media and then incubated at 37 ºC anaerobically until the culture reached exponential phase (OD_{600} ~1.0). The cell pellets were collected by centrifugation at 3,200 x g for 10 min and washed with BMM-C media. The cell pellets were resuspended into BMM-C at an OD_{600} of approximately 1 and then inoculated into BMM-xylose (5 g L^{-1} xylose, MP Biomedicals) to an OD_{600} of 0.025 in a 96-well plate (Greiner Bio-One). The growth of the strains was measured based on OD_{600} using a Tecan F200 plate reader and measurements were taken every 30 min for 54 hr.

Characterization of *B. thetaiotamicron* Δ*levan* strain

The *B. thetaiotaomicron* ATCC 29148 (VPI-5482) wild-type and *B. thetaiotamicron* Δ*levan* (deletion of *BT1754-BT1765*) strains were grown at 37 ºC anaerobically in ABB media for 12-16 hr. Cultures were diluted by 20-fold into ABB media and then incubated at 37 ºC anaerobically until the culture reached exponential phase (OD_{600} ~1.0). Cell pellets were collected by centrifugation at 3,200 x g for 10 min, and then washed with BMM-C media. The cell pellets were resuspended into BMM-C to an OD_{600} of approximately 1 and then inoculated into BMM-levan (5 g L^{-1} levan, Sigma) using an initial OD_{600} of 0.05 in a 96-well plate (Greiner Bio-One). Cell growth was determined based on OD_{600} using a Tecan F200 plate reader every 30 min for 60 hr.

Pooled barcoded BU mutant experiments

The barcoded BU control strain Δ*tryP-24* and PUL deletion mutants were grown at 37 ºC anaerobically in ABB media for 12-16 hr. Cultures were then diluted by 20-fold into ABB media and grown at 37 ºC anaerobically until the culture reached exponential phase (OD_{600} ~1.0). Next, all the strains were mixed in equal proportions based on OD_{600} and then washed with DM29 media (lacking a carbon source, **Supplementary Data 4**). Next, the cells were resuspended in DM29 and diluted to a final OD_{600} of approximately 1. The mixture of strains was inoculated into 2 mL 96-deep-well plates (Nest Scientific) containing 1 mL DM29-glucose or DM29 supplemented with different glycans (**Supplementary Data 4**) to an initial OD_{600} of 0.05 and incubated at 37 ºC anaerobically for varying lengths of time and passaging was performed by diluting the cultures by 20-fold up to two times. After 48 hr of cultivation, OD_{600} was measured by Tecan F200 plate reader and cell pellets were collected for NGS sequencing (**Figure 3A**).

Colony forming unit counting for gnotobiotic mouse experiments

The cecal contents weighing 0.2-0.3 mg were collected into sterilized Eppendorf tubes and then resuspended into 1 mL anaerobic ABB media. The suspended contents were homogenized by two 3.2 mm stainless steel beads (BioSpec Products) with intermittent vortex (BR-2000, Bio-Rad) for 5 min. The homogenized contents were then transferred into 9 mL of ABB media. This mixture was diluted 10^3 to 10^5 times for CFU plating. We plated 100 µL of the diluted solutions on BHIA plates and incubated in an anaerobic chamber at 37 °C for 36-48 hr until colonies were visible on the plates. We computed the CFU and divided this number by the measured weight of each cecal content.

DNA extraction from fecal and cecal samples

The DNA extraction for fecal and cecal samples was performed as described previously with some modifications (Goodman et al., 2011). Fecal samples (~50 mg) were transferred into solvent-resistant screw-cap tubes (Sarstedt Inc) with 500 µL of 0.1 mm zirconia/silica beads (BioSpec Products) and one 3.2 mm stainless steel bead (BioSpec Products). The samples were resuspended in 500 µL of Buffer A (200 mM NaCl (DOT Scientific), 20 mM EDTA (Sigma) and 200 mM Tris·HCl pH 8.0 (Research Products International)), 210 µL 20% SDS (Alfa Aesar) and 500 µL phenol/chloroform/isoamyl alcohol (Invitrogen). Cells were lysed by mechanical disruption with a bead-beater (BioSpec Products) for 3 min twice to prevent overheating. Next, cells were centrifuged for 5 min at 8,000 x g at 4°C, and the supernatant was transferred to a Eppendof tube. We added 60 µL 3M sodium acetate (Sigma) and 600 µL isopropanol (LabChem) to the supernatant and incubated on ice for 1 hr. Next, samples were centrifuged for 20 min at 18,000 x g at 4°C. The harvested DNA pellets were washed once with 500 µL of 100% ethanol (Koptec). The remaining trace ethanol was removed by air drying the samples. Finally, the DNA pellets were then resuspended into 200 µL of AE buffer (Qiagen). The crude DNA extracts were purified by Zymo DNA Clean & Concentrator™-5 kit (Zymo Research) for NGS sequencing.

Gene expression measurements of representative PULs in BU in germ-free mice

The RNA extraction for cecal samples was performed as described previously with some modifications (Rey et al., 2010). Specifically, cecal samples (50~100 mg) were transferred into solvent-resistant screw-cap tubes (Sarstedt Inc) with 500 µL of acid washed beads (212-300 µm, Sigma). The samples were resuspended in 500 µL of Acid-Phenol:Chloroform (with IAA, 125:24:1, pH 4.5, Thermo Fisher Scientific), 500 µL of Buffer B (200mM NaCl (DOT Scientific), 20 mM EDTA (Sigma), pH 8.0), 210 µL 20% SDS (Alfa Aesar). Cells were lysed by mechanical disruption with a bead-beater (BioSpec Products) for 2 min twice to prevent overheating. Next, cells were centrifuged for 10 min at 8,000 x g at 4°C, and the supernatant was transferred to a Eppendof tube. We added 60 µL 3M sodium acetate (pH 5.5, Sigma) and 600 µL isopropanol (LabChem) to the supernatant and incubated at –80°C for 6 min. Next, samples were centrifuged for 15 min at 18,200 x g at 4°C. The harvested RNA pellets were washed once with 900 µL of 100% ethanol (Koptec). Finally, the RNA pellets were then resuspended into 100 µL of RNase-free water. The crude RNA extracts were purified by the RNeasy Mini kit (Qiagen) with a DNase I (Qiagen) treatment step to eliminate DNA in the sample. To remove any remaining DNA, the RNA products were treated with Baseline-ZERO DNase I (Epicentre) and then purified again using the RNeasy Mini kit.

We performed cDNA synthesis with 0.5-1 µg of total purified RNA using the iScript Select cDNA Synthesis Kit (Bio-Rad Laboratories). We performed Real Time quantitative PCR (RT-qPCR) on the Bio-Rad CFX connect Real-Time PCR instrument with SYBR™ Green PCR Master Mix (Thermo Fisher Scientific). We computed the fold changes of the target genes by normalizing to the two reference genes 16S rRNA gene and *gyrA* (*BACUNI_RS15400*, encoding DNA gyrase A) using the geometric mean (Jo Vandesompele, 2002). We computed the fold change using the equation 2^x where $x = (Ct_{control} - \sqrt{Ct_{16Sc}\, Ct_{gyrAc}}) - (Ct_{mice} - \sqrt{Ct_{16Sm}\, Ct_{gyrAm}})$ where $Ct_{control}, Ct_{mice}, Ct_{16Sc}, Ct_{gyrAc}, Ct_{16Sm}, Ct_{gyrAm}$ denote the Ct value of the target gene in glucose media, the target gene in mouse ceca, 16S rRNA gene in glucose media, *gyrA* in glucose media, 16S rRNA in mouse ceca and *gyrA* in the mouse ceca, respectively. We used three technical replicates per gene for PCR.

Butyrate producer community experiments and sample collection

All BU strains and butyrate producers (*A. caccae* DSM 14662, *C. comes* ATCC 27758, *E. rectale* ATCC 33656 and *R. intestinalis* DSMZ 14610) were inoculated into ABB media with the exception of *R. intestinalis* which was inoculated into BHI media and grown at 37 ºC anaerobically. After 16-24 hr, the cultures were diluted by 20-fold into the same media until the culture reached exponential phase (OD_{600}~1.0). The cultures were centrifuged at 3,200 x g for 10 min, and then washed with DM29 media. The washed cells were then resuspended into the same DM29 media and the OD_{600} was diluted to 1. The BU and each butyrate producer cultures were mixed in equal proportions based on OD_{600} and inoculated into DM29-glucose (5 g L^{-1}) or DM29 media supplemented with glycogen, glucomannan, inulin, laminarin, pectic galactan, pectin, pullulan or xyloglucan (**Supplementary Data 4**). The pairwise communities and monocultures were introduced into 2 mL 96-deep-well plates (Nest Scientific) to an initial OD_{600} of 0.05. We have 4 plates for each experiment and each plate was taken out for sampling every 12 hr for a total of 48 hr. At each time point, OD_{600} was measured with Tecan F200 (with 5-10 dilution based on the density of the cell culture) for the monitor of cell growth and cell pellets were collected for NGS sequencing.

We performed butyrate measurements at a single time point. Specifically, 2 µL of H_2SO_4 (Sigma) was added to the supernatant of each sample to precipitate any components that was incompatible with the mobile phase. The samples were then centrifuged at 3,200 x g for 10 min and then filtered through a 0.2 µm filter (Pall Corporation) using a vacuum manifold (KNF Neuberger) before transferring to HPLC vials (Thermo Scientific). Butyrate concentrations were measured with an Agilent 1260 infinity HPLC system equipped with a quaternary pump, chilled

(4°C) autosampler, vacuum degasser, refractive index detector, Aminex HPX-87H column and Cation-H guard column (300x7.8mm, BioRad). We used 0.02 N H_2SO_4 as the mobile phase with a flow rate of 0.5 mL min^{-1} at a column temperature of 50°C. The injection volume of all samples was 50 µL and the run time was 30 min. Data analysis was performed using the Chem Station Rev.C01.08 software (Agilent Technologies).

Bacterial genome DNA extraction and next-generation sequencing

All the genomic DNA extraction and next-generation sequencing sample preparation was performed as described previously (Clark et al., 2021). Briefly, bacterial genome DNA extraction was carried out using a modified version of the Qiagen DNeasy Blood and Tissue Kit protocol in 96-well plates(Clark et al., 2021). Genomic DNA concentrations were measured using the Quant-iT™ dsDNA Assay Kit (Invitrogen) and then normalized to 1 ng $µL^{-1}$ or 2 ng $µL^{-1}$ for genomic DNA extracted from fecal and cecal samples by diluting in molecular grade water (VWR International) using a Tecan Evo Liquid Handling Robot. We performed PCR using custom dual-indexed primers(Clark et al., 2021; Venturelli et al., 2018) targeting the V3-V4 region of the 16S rRNA gene using the diluted genomic DNA samples as template. These libraries were purified using the DNA Clean & Concentrator™-5 kit (Zymo) and eluted in water. Sequencing was performed on an Illumina MiSeq using MiSeq Reagent Kit v3 (600-cycle) to generate 2x300 paired end reads. For the sequencing of barcoded BU strains, the 200 bp amplicon libraries containing 4-bp barcodes were generated using the procedure described above and PCR amplified with custom dual-indexed primers listed in **Supplementary Data 4**. The obtained libraries were sequenced on an Illumina MiSeq using MiSeq Reagent Nano Kit v2 (500-cycles) or MiSeq Reagent Kit v3 (600-cycle) to generate 2x250 paired end reads or 2x300 paired end reads.

Characterization of butyrate producer growth in BU conditioned media

BU was grown anaerobically at 37 ºC in ABB media for 12-16 hr. Cells were harvest by centrifugation at 3,200 x g for 10 min, washed with DM29 and then inoculated into 12 mL DM29 media supplemented with different glycans and incubated at 37ºC anaerobically for 16 hr. Next, cell pellets were collected with centrifugation at 3,200 x g for 10 min and washed twice with DM29. The washed cell pellets were resuspended into same volume of the same media and incubated at 37°C for 3 hr to allow cell growth and glycan utilization. The supernatants were collected by centrifugation at 3,200 x g at 4°C for 30 min and the pH was adjusted to the same value as fresh media 5N KOH (Alfa Aesar) with the Mettler Toledo InLab Micro pH electrode. The pH adjusted conditioned media was filtered twice using a 0.2 µm filter (Whatman) to remove BU cells. Butyrate

producers were grown in ABB (AC, CC and ER) or BHI (RI) at 37 ºC anaerobically for 16-24 hr, and passaged in the same media with a dilution of 20-fold until they reached exponential phase (OD_{600}~1.0). Cells were harvested by centrifugation at 3,200 x g for 10 min and washed with DM29 and resuspended in DM29 to an OD_{600} of approximately 1. The cultures were inoculated into both fresh media and conditioned media with an initial OD_{600} of 0.05 and anaerobically grown at 37 ºC in a 96-well plate (Greiner Bio-One) without shaking. Cell growth was monitored by plate reader (Tecan F200).

Characterization of butyrate producer growth in cell membrane treated glycan media

The preparation of the cell membrane fraction including inner and outer membrane anchoring enzymes was modified based on the procedures described from Millipore Sigma (https://www.sigmaaldrich.com/technical-documents/protocols/biology/purifying-challenging-proteins/cell-disruption-and-membrane-preparation.html). BU was grown anaerobically at 37 ºC in ABB media for 12-16 hr. Cells were harvest by centrifuge at 3,200 x g for 10 min and washed with DM29 and inoculated into 20 mL of DM29-glycan media (**Supplementary Data 4**) and grown anaerobically at 37ºC for 16 hr. Cell pellets were collected and washed with DM29 at 3,200 x g, 4°C for 1 hr. The washed cell pellets were resuspended into 6 mL of DM29 with the addition of 60 µL of pefabloc (100 mM, Sigma) and 6 µL of DNase I (20 mg mL^{-1}, Sigma). Cells were lysed via sonication (5 s on and 5 s off for 2.5 min, performed twice, Sonicator 3000, Misonix). The cell lysis was centrifuged at 3,200 x g at 4 °C for 10 min. The supernatants were collected and filtered twice with 0.45 µm filter (Whatman) to remove the remaining intact cells. The collected supernatants were centrifuged at 300,000 x g for 2 hr with Optima MAX-XP ultracentrifuge with SW 55 Ti Swinging-Bucket Rotor (Beckman Coulter). The pellets consisting of the cell membrane fractions were resuspended into 10 mL of the same glycan containing media and incubated in the anaerobic chamber at 37°C for 16 hr. The cell membrane treated glycan media as well as the respective fresh glycan containing media were filtered with 0.2 µm filter (Whatman). Butyrate producers were grown in ABB (AC, CC and ER) or BHI (RI) at 37ºC anaerobically for 16-24 hr, and passaged in the same media with a dilution of 20-fold until they reached exponential phase (OD_{600}~1.0). Cells were harvested by centrifugation at 3,200 x g for 10 min and washed with DM29 and resuspended in DM29 to an OD_{600} of approximately 1. Cultures were then inoculated into fresh media supplemented with different glycans and cell membrane treated glycan media to an initial OD_{600} of 0.05 and incubated at 37 °C without shaking in a 96-well plate (Greiner Bio-One). Cell growth was monitored by plate reader (Tecan F200).

Measurements of fructose

The fructose in BU conditioned inulin media was measured using a High Sensitivity Fructose Assay Kit (Sigma). BU was grown anaerobically at 37ºC in ABB media for 12-16 hr. Cells were harvest by centrifuge at 3,200 x g for 10 min and washed with DM29 and inoculated into DM29 media supplemented with 5 g L^{-1} inulin for 16 hr. Next, cell pellets were collected with centrifugation at 3,200 x g for 10 min and washed twice with DM29. The washed cell pellets were resuspended into same volume of DM29-inulin and incubated at 37°C for 3 hr. The supernatants were collected with centrifugation at 3,200 x g for 10 min and filtered using a 0.2 µm filter (Whatman) prior to fructose measurement. The fructose concentrations in control solutions were also measured (5 g L^{-1} inulin solution and DM29-inulin media).

Growth characterization of AC with fructose as the primary carbon source

AC was grown anaerobically in ABB media at 37 ºC for 16-24 hr, and passaged in the same media with a dilution of 20-fold until they reached exponential phase (OD_{600}~1.0). Cells were harvested with centrifugation at 3,200 x g for 10 min and washed with DM29 and resuspended into DM29 again to OD_{600} of approximately 1. The cultures were then inoculated into DM29-fructose (5 g L^{-1}, Sigma) and DM29-glucose (5 g L^{-1}) media in a 96-well plate (Greiner Bio-One) to an initial OD_{600} of 0.05. Cell growth was monitored using a Tecan F200 plate reader.

QUANTIFICATION AND STATISTICAL ANALYSIS

Bioinformatic analysis of species and barcoded strain abundances

For the 16S rDNA gene sequencing data analysis, we used previously described custom scripts in Python 3.7 and aligned to a reference database of V3-V4 16S rRNA gene sequences as previously described (Venturelli et al., 2018; Zhang et al., 2014). Relative abundance was calculated as the read count mapped to each species divided by the total number of reads for each condition. The absolute abundance of each species was calculated by multiplying the relative abundance determined by NGS sequencing by the OD_{600} measurement for each sample. For the sequencing analysis of barcode-tagged strains, paired end reads were first merged using PEAR (Paired-End reAd mergeR) v0.9.10 (Zhang et al., 2014) after which barcodes were extracted by searching for exact matches of the immediate upstream and downstream sequences within the reads. Barcodes with less than 100% match were discarded. The relative abundance of each barcoded strain was calculated as the number of reads mapped to each barcode divided by the total reads that mapped to each condition. Absolute abundance of each mutant was calculated by multiplying the relative abundance by the OD_{600} measurement of each condition.

Transcriptional profiling of BU

The BU WT and genome modified mutants were grown at 37 ºC anaerobically in ABB media for 12-16 hr. Cultures were then diluted by 20-fold into ABB media and then grown at 37 ºC anaerobically until the culture reached exponential phase (OD_{600} ~1.0). The cell pellets were collected by centrifugation at 3,200 x g for 10 min and washed with BMM-C media. Next, the cell pellets were resuspended into BMM-C and then inoculated into 5 mL BMM, BMM-xyloglucan (5 g L^{-1} xyloglucan, Megazyme) or BMM-type II media (5 g L^{-1} type II mucin, Sigma) to an initial OD_{600} of 0.05. Cells were harvested for total RNA extraction when the OD_{600} reached 0.6-0.8. RNA was extracted using the RNeasy Mini Kit (Qiagen) and genomic DNA was digested using the RNase-Free DNase Set (Qiagen). The RNA integrity number (RIN, an algorithm for assigning integrity values to RNA measurements) was measured by Agilent TapeStation 4150 with the Agilent High Sensitivity RNA ScreenTape. The samples were then processed by GENEWIZ (NJ, USA) by performing rRNA depletion, cDNA library preparation and sequencing. rRNA depletion was performed by using Ribozero rRNA Removal Kit (Illumina). The NEBNext Ultra RNA Library Prep Kit (NEB) was used for RNA sequencing library preparation. The sequencing libraries were sequenced with an Illumina HiSeq instrument. Image analysis and base calling were conducted by the HiSeq Control Software (HCS). Raw sequence data generated from Illumina HiSeq was converted into FASTQ files and de-multiplexed using Illumina's bcl2fastq 2.17 software. One mismatch was allowed for index sequence identification.

The compressed FASTQ files were quality checked using the FastQC tool v0.11.8 (Andrews, 2010). Packages from the BBTools suite v38.42 (Bushnell, 2014) including BBDuk, BBSplit, and BBMap were used to filter high quality reads, trim adapters using built-in adapter reference file, remove rRNA reads, and map sequences to the reference genome (*B. uniformis* DSM 6597). The featureCounts package v1.6.4 (Liao et al., 2014) from the SubRead suite was used for read mapping to gene features and quantifying raw counts for each transcript. The DESeq2 Bioconductor library v4.0.3 (Love et al., 2014) was used in R v4.0.4 to normalize read counts across samples and quantify differential gene expression using a negative binomial generalized linear models with apeglm shrinkage estimator (Zhu et al., 2019). The scripts for analysis of RNA-seq data are available from Zenodo (doi:10.5281/zenodo.5722360)

Real-time quantitative reverse transcription PCR (qRT-PCR) of BU *in vitro* cultures

The BU WT and genome modified mutants were grown at 37ºC anaerobically in ABB media for 12-16 hr. Cultures were diluted by 20-fold into ABB media and then grown at 37ºC anaerobically until the culture reached exponential phase (OD_{600} ~1.0). Cell pellets were collected by centrifugation at 3,200 x g for 10 min and then washed with BMM-C media. The washed cell pellets were then resuspended into BMM-C and then inoculated into 5 mL BMM or BMM-xyloglucan (5 g L^{-1}) media to an initial OD_{600} of 0.05. Cells were harvested for total RNA extraction when the OD_{600} reached 0.6-0.8. Total RNA was extracted with RNeasy Mini Kit (Qiagen) and treated with DNase I (Thermo Fisher Scientific) to remove the genomic DNA. We performed cDNA synthesis with 0.5-1 µg of total RNA using the iScript Select cDNA Synthesis Kit (Bio-Rad Laboratories). We performed Real Time quantitative PCR (RT-qPCR) on the Bio-Rad CFX connect Real-Time PCR instrument with SYBR™ Green PCR Master Mix (Thermo Fisher Scientific). We computed the fold changes of the target genes by normalizing to the two reference genes 16S rRNA gene and *gyrA* (*BACUNI_RS15400*, encoding DNA gyrase A) using the geometric mean (Jo Vandesompele, 2002). We computed the fold change using the equation 2^x where $x = (Ct_{control} - \sqrt{Ct_{16Sc}\, Ct_{gyrAc}}) - (Ct_{mutant} - \sqrt{Ct_{16Sm}\, Ct_{gyrAm}})$ where $Ct_{control}$, Ct_{mutant}, Ct_{16Sc}, Ct_{gyrAc}, Ct_{16Sm}, Ct_{gyrAm} denote the Ct value of the gene in Δ*tryP-24*, 16S rRNA gene in Δ*tryP-24*, *gyrA* in Δ*tryP-24*, target gene of the PUL mutant strain, 16S rRNA gene of the PUL mutant strain and *gyrA* in the PUL mutant strain, respectively. We used three technical replicates per gene for PCR.

Bioinformatic analysis of PULs in human gut microbiome datasets

Two human gut microbiome datasets, which contained 154,723 metagenome-assembled genomes (MAGs) from 9,428 human gut microbiomes (Pasolli et al., 2019) and 92,143 MAGs from 11,850 human gut microbiomes (Almeida et al., 2019), were used to find PULs (including *PUL11*, *PUL12*, *PUL17*, *PUL18*, *PUL2*, *PUL37* and *PUL43*) of BU. All MAGs were annotated by Prodigal v2.6.3 (Hyatt et al., 2010) and DIAMOND BLASTP v0.9.28.129 (Buchfink et al., 2015) was used to find hits to reference proteins within each PUL with settings of '-k 1 -e 1e-5 --query-cover 25 --id 50' (protein sequence hit searching criteria of maximum number of target sequence set as 1, e-value cutoff set as 1e-5, query coverage cutoff set as 25%, and sequence identity cutoff set as 50%). Annotation by dbCAN2 (Zhang et al., 2018) was used to identify glycoside hydrolases (GHs) from the DIAMOND BLASTP hits. We then used three criteria to assign PUL positive hits: 1) satisfies the requirement of essential genes for the function of each PUL; 2) GH annotation of the essential gene was consistent with the PUL reference; 3) essential genes of each PUL were within a gene array of size less than 30 genes (**Figure S5**)

For all MAGs, we used CheckM v1.0.11 (Parks et al., 2015) to evaluate genome completeness and to assign the total MAG dataset into subsets with > 50%, > 60%, > 70%, > 80%, and > 90% genome completeness. The finer taxonomic information of MAGs within Bacteroidetes was parsed by GTDB-Tk v0.1.3 (Chaumeil et al., 2019) with default settings. The abundance ratios of PULs in all MAGs, only BU MAGs, and other non-BU MAGs were calculated from the resulted PUL positive hit table. The cooccurrence ratios of PULs in all MAGs was also parsed out accordingly. Results of abundance ratios and cooccurrence ratios that were calculated from a series of genome completeness subsets were combined and visualized together.

Logistic growth and generalized Lotka-Volterra (gLV) models

In monoculture experiments (**Figure 1**), we use the logistic growth model to describe population growth dynamics. The logistic growth model for species i takes the following form:

$$\frac{dx_i}{dt} = r_i x_i \left(1 - \frac{x_i}{k_i}\right),$$

where x_i is the absolute abundance of species i, parameter r_i is its maximum growth rate, and k_i is its steady-state abundance.

The fold change in steady-state abundance for the logistic growth model was determined by the ratio of the inferred parameter k of the BU ΔPUL strain by the inferred parameter k for BU Δ*tryP-24* in each media condition. The fold change in growth rate for the logistic growth model was determined by the ratio of the inferred parameter r of the BU ΔPUL strain by the inferred parameter r for BU Δ*tryP-24* in each media condition.

In co-culture experiments (**Figure 5**) of species 1 (BU WT or one of its mutants) with species 2 (one of the butyrate producers), we use the generalized Lotka-Volterra (gLV) model to describe growth dynamics and inter-species interactions. Specifically, the gLV model can be written as the following ordinary differential equation:

$$\frac{dx_1}{dt} = x_1(\mu_1 + a_{11}x_1 + a_{12}x_2),$$

$$\frac{dx_2}{dt} = x_2(\mu_2 + a_{21}x_1 + a_{22}x_2),$$

where, for $i, j \in \{1,2\}$, x_i is the absolute abundance of species i, the non-negative parameter μ_i describes the basal growth rate of species i, a_{ij} is the inter-species interaction coefficient that quantifies how the abundance of species j modifies the growth rate of species i, and a_{ii} is the intra-species interaction coefficient that quantifies how abundance of species i affects its own growth rate. The gLV model has been used before to describe inter-species interactions in complex microbial communities and to predict their emerging community dynamics (Venturelli et al., 2018).

Parameter inference for logistic growth and gLV models

To infer the growth and interaction parameters in both the logistic growth and the gLV models, we use an adaptive Markov Chain Monte Carlo (MCMC) method (Haario et al., 2001), which we describe here in detail. The custom code is available from Zenodo (doi:10.5281/zenodo.5722360). Let θ be the parameters that needs to be inferred from one of the models, for the logistic growth model, we have $\theta = (r_1, \dots, r_n, k_1, \dots, k_n)$ and for the gLV model $\theta = (\mu_1, \dots, \mu_n, a_{11}, \dots, a_{nn})$, where in either case, n is the total number of species of interest. A unique θ is inferred for each medium type. For each medium type, we performed a set of experiments, which includes monoculture of all species of interest and, in some cases, co-culture of BU and butyrate producers. For the *q*-th experiment, we took *m* time-series abundance measurements for three biological replicates with mean $\mathbf{x}_q = (\mathbf{x}_{q1}, \dots \mathbf{x}_{qm})$ and standard deviation $\sigma_q = (\sigma_{q1}, \dots, \sigma_{qm})$, where $\mathbf{x}_{qi}$ and σ_{qi} are both n-dimensional vectors for all species. Given these observations in all *p* experiments: $\mathbf{x} = (\mathbf{x}_1, \dots \mathbf{x}_p)$ and $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_p)$, the posterior distribution of θ, which we denote by $P(\theta|\mathbf{x}, \boldsymbol{\sigma})$, is found using the adaptive MCMC method.

In particular, we assume that uncertainty for the *k*-th measurement in the *q*-th experiment is modeled by an additive and independent noise, which is distributed according to $N(0,\sigma_{qk}^2)$. Given a fixed θ, we first simulate the logistic growth/gLV model for each experiment *q* to obtain the model predicted abundance $\bar{\mathbf{x}}_{qk}(\theta)$ at every instant *k*. The likelihood to observe the sequence of abundance measurements $\mathbf{x}$ can then be computed as:

$$P(\mathbf{x}|\theta; \boldsymbol{\sigma}) = \prod_{q=1}^{p} \prod_{k=1}^{m} f(\mathbf{x}_{qk} - \bar{\mathbf{x}}_{qk}(\theta); \sigma_{qk}),$$

where $f(\cdot; \sigma_{qk})$ is the probability density function for the normal distribution $N(0,\sigma_{qk}^2)$. The posterior probability can then be described according to Bayes rule as $P(\theta|\mathbf{x}) \propto P(\mathbf{x}|\theta; \boldsymbol{\sigma})P(\theta)$, where $P(\theta)$ is the prior parameter distribution. In this paper, for all cases except xyloglucan co-culture, we chose uniform priors for the parameters. Specifically, the priors for all growth rates μ_i are $U(0,2)$, the priors for all inter-species interaction coefficients a_{ij} are $U(-2.5,2.5)$, and the priors for all intra-species interaction coeffects a_{ii} are $U(-2.5,0)$. The boundaries for these distributions are chosen to be sufficiently large to contain similar logistic growth/gLV parameters identified in the literature (Venturelli et al., 2018) and to ensure positive, bounded growth when simulating monoculture experiments. A normal prior distribution was used to infer parameters in the xyloglucan co-culture experiments, and parameters of this distribution are listed in **Supplementary Data 2**. Since logistic growth models cannot capture cell growth in death phase,

if the OD_{600} of the *k*-th measurement in a monoculture experiment *q* drops more than 20% from that of the (*k*-1)-th measurement, then x_{qk} is not used for parameter inference. Data points excluded for parameter inference were indicated with an empty circle in **Figure S7**. In monoculture, many Δ*PUL* strains did not grow (OD_{600}<0.08 for all time) in media supplemented with glycans. The maximum growth rates (r_i) for these Δ*PUL* strains in the respective mediums are set to 0. Similarly, for Δ*PUL*/butyrate producer coculture experiments, no inference was made on the interaction coefficients between the two species in conditions that did not display growth, and thus the coefficients were set to 0.

An adaptive, symmetric, random-walk Metropolis MCMC algorithm (Haario et al., 2001) is then used to draw samples from the posterior distribution $P(\theta|\mathbf{x})$. Specifically, given the current sample $\theta^{(l)}$ at step *l* of the Markov chain, the proposed sample for step (*l*+1) is $\theta^{(l+1)}= \theta^{(l)}+\delta^{(l)}$, where $\delta^{(l)}$ is drawn randomly from a normal distribution. The algorithm is adaptive in the sense that the covariance of this normal distribution is given by $\alpha \cdot \gamma_l^2$, where γ_l^2 is the covariance of $\theta^{(1)},\ldots, \theta^{(l)}$ and α is a positive parameter. In this paper, depending on the carbon source, parameter α is either chosen to be 0.1or 0.5. The posterior probability $P(\theta^{(l+1)}|\mathbf{x})$ of the proposed sample $\theta^{(l+1)}$ is then computed, and the proposed sample is accepted with probably 1 if $\frac{P\left(\theta^{(l+1)}\middle|\mathbf{x}\right)}{P\left(\theta^{(l)}\middle|\mathbf{x}\right)} > 1$, and it is accepted with probability β if $\frac{P\left(\theta^{(l+1)}\middle|\mathbf{x}\right)}{P\left(\theta^{(l)}\middle|\mathbf{x}\right)} = \beta \leq 1$.

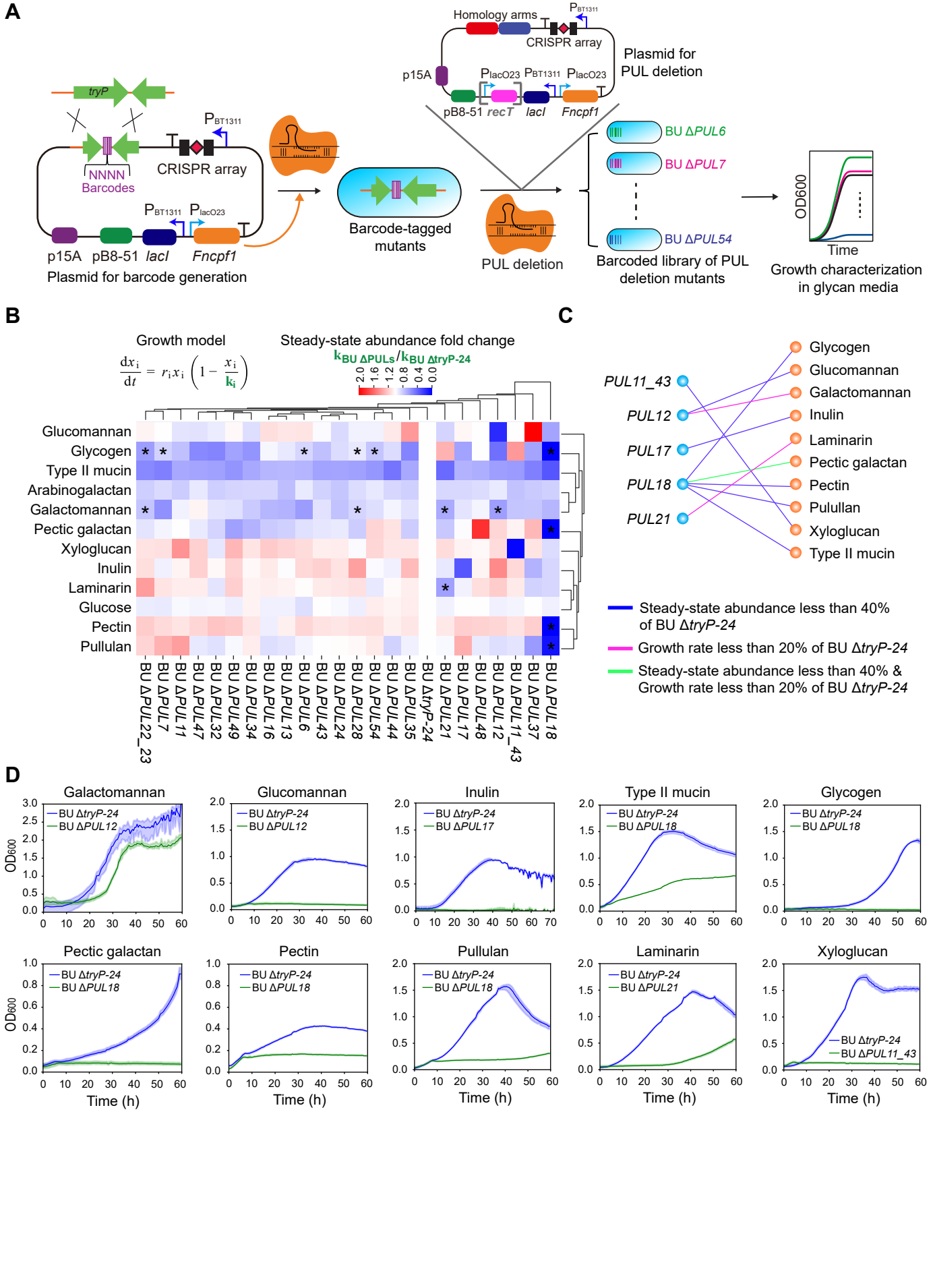
The algorithm described above was implemented using custom code in MATLAB R2016a (The MathWorks, Inc., Natick, MA, USA), where the logistic growth/ gLV models are solved using variable step solver ode23s. For each carbon source, we collected at least 300,000 MCMC samples after a burn-in period of at least 100,000 samples. The Gelman-Rubin potential scale reduction factor (PSRF) was used to evaluate convergence of the posterior distributions, where a PSRF closer to 1 indicates better convergence. We found that out of 186 parameters, 82% of them have PSRFs less than 1.2, and the median of PSRF is 1.03, indicating that the parameters drawn from MCMC had converged to the posterior parameter distribution. The medians and the coefficient of variations (CVs) of the marginal distributions of the identified parameters are summarized in **Supplementary Data 2**. The median of the parameter distributions was used to simulate the temporal abundance trajectory in **Figure 5** and **Figure S7**.

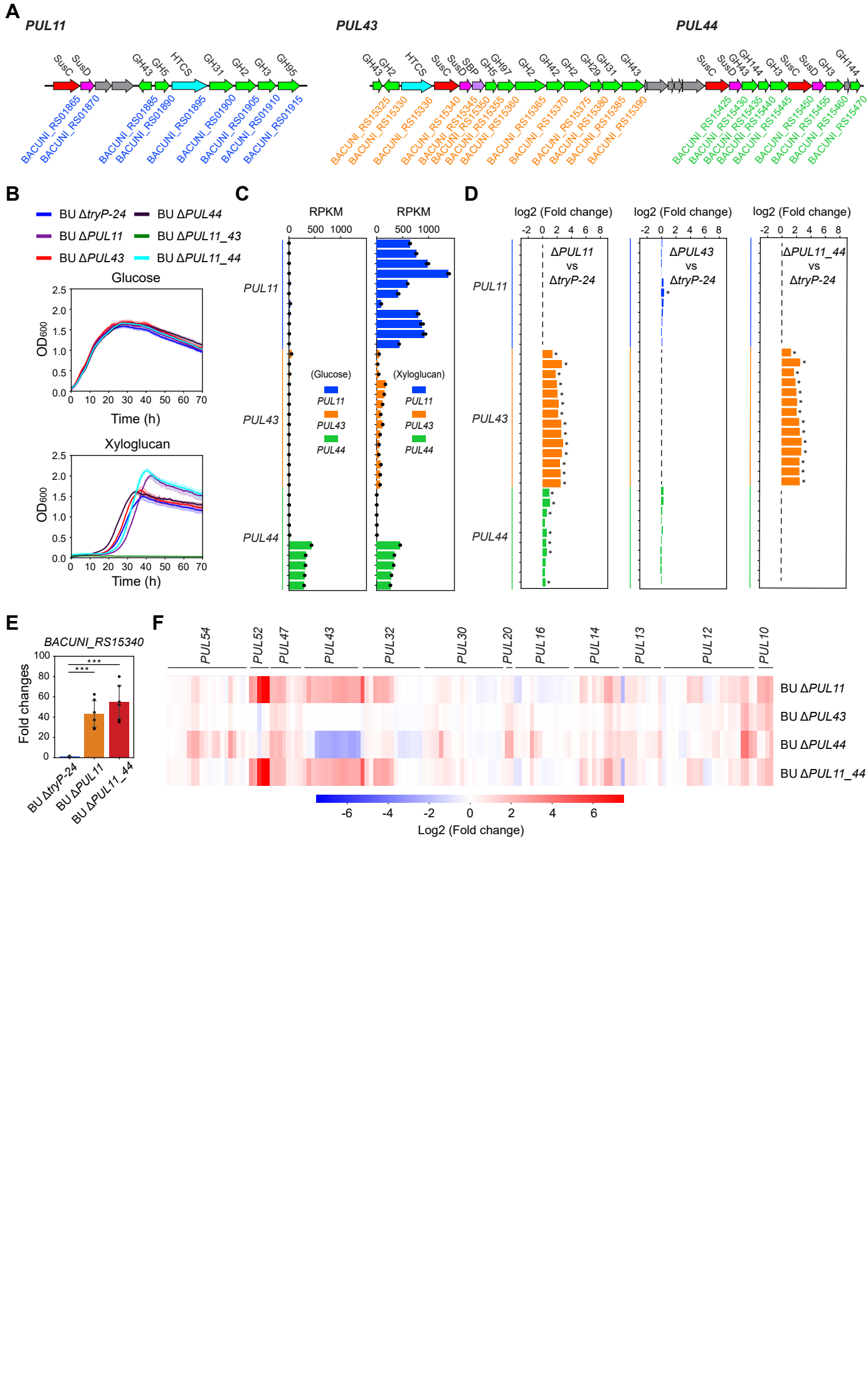
Regression model for predicting butyrate concentration

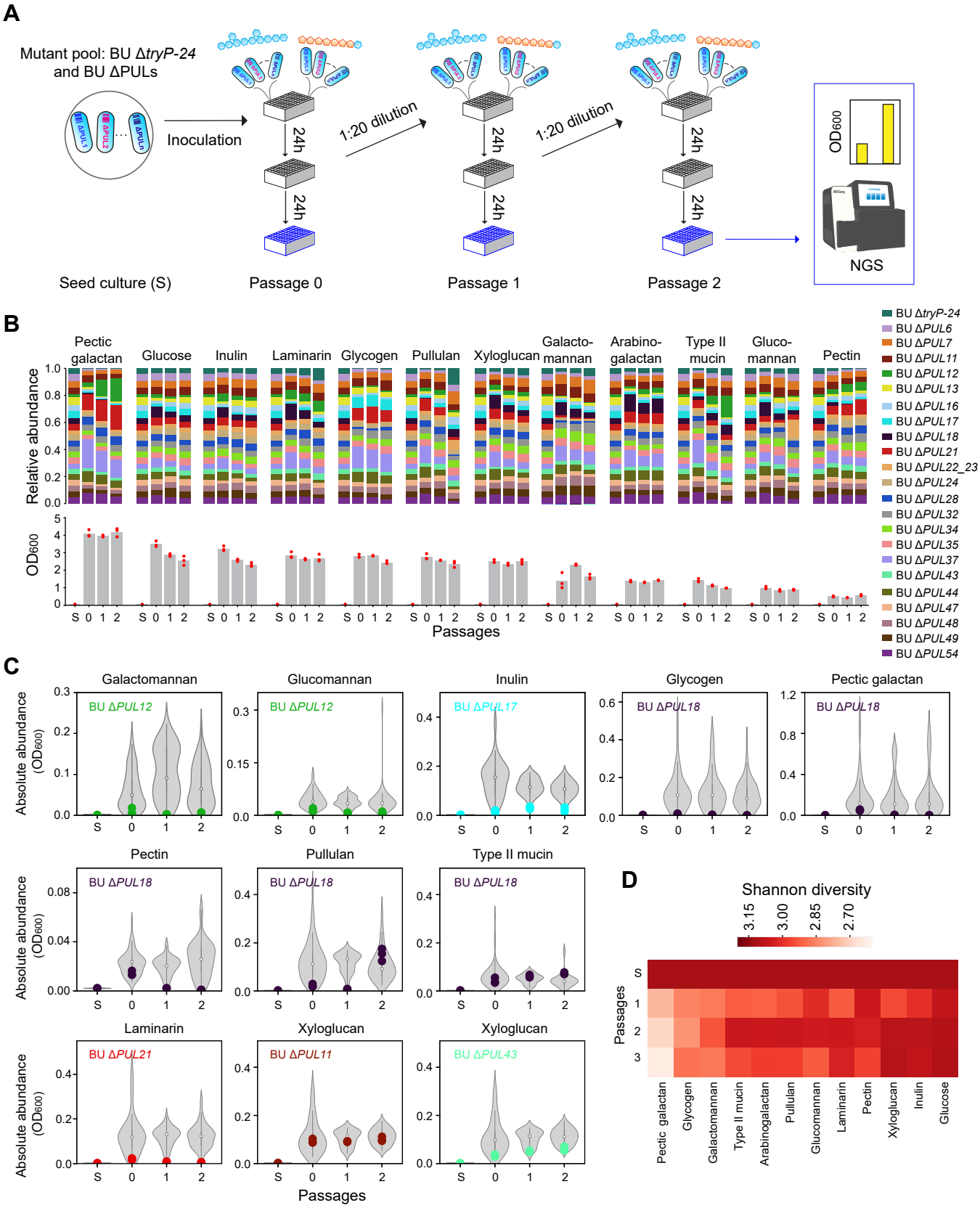
A previous study (Clark et al., 2021) has shown that end point butyrate concentration in a microbial community in batch culture can be predicted by the absolute abundance of butyrate producers. Inspired by this, we propose the following linear regression model for end point butyrate concentration:

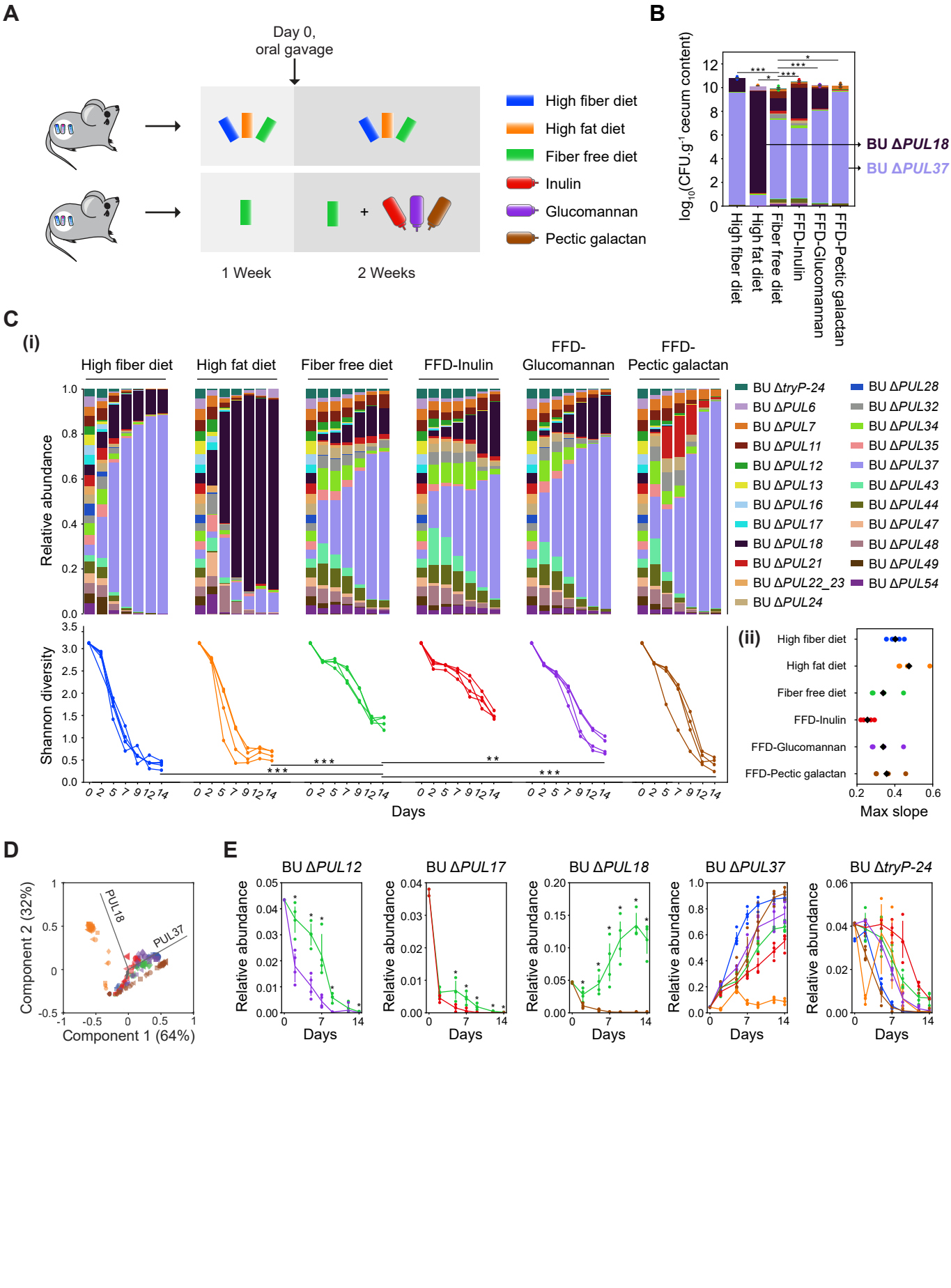
$$B = \sum_i k_i x_i,$$

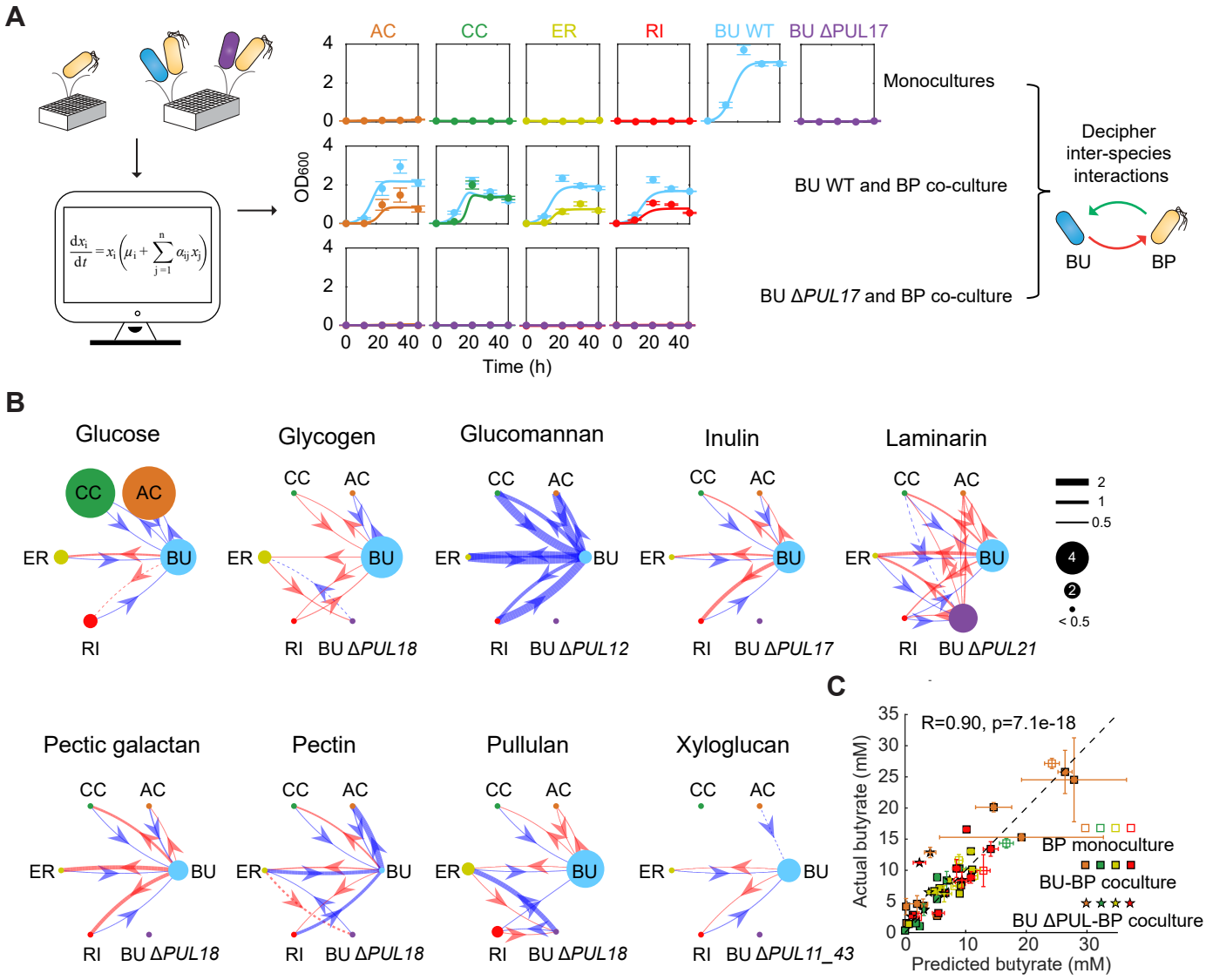
where B is the predicted butyrate concentration, x_i represents end point butyrate producer abundances, where i=AC, CC, ER, or RI, and k_i is a constant parameter. A major assumption in this model is that the parameter k_i is independent of carbon source in growth media, the absence/presence of BU, and the absence/presence of ΔPUL strains. The inferred parameters k_i for AC, CC, ER and RI were 19.2, 3.99, 8.31 and 19.35 mM OD_{600}^{-1} respectively. The parameters k_i were obtained by performing least square regression on all monoculture and coculture experiments with a butyrate producer. Linear regression is performed using MATLAB function fitlm with intercept constrained to be 0. The custom code is available from Zenodo (doi:10.5281/zenodo.5722360).

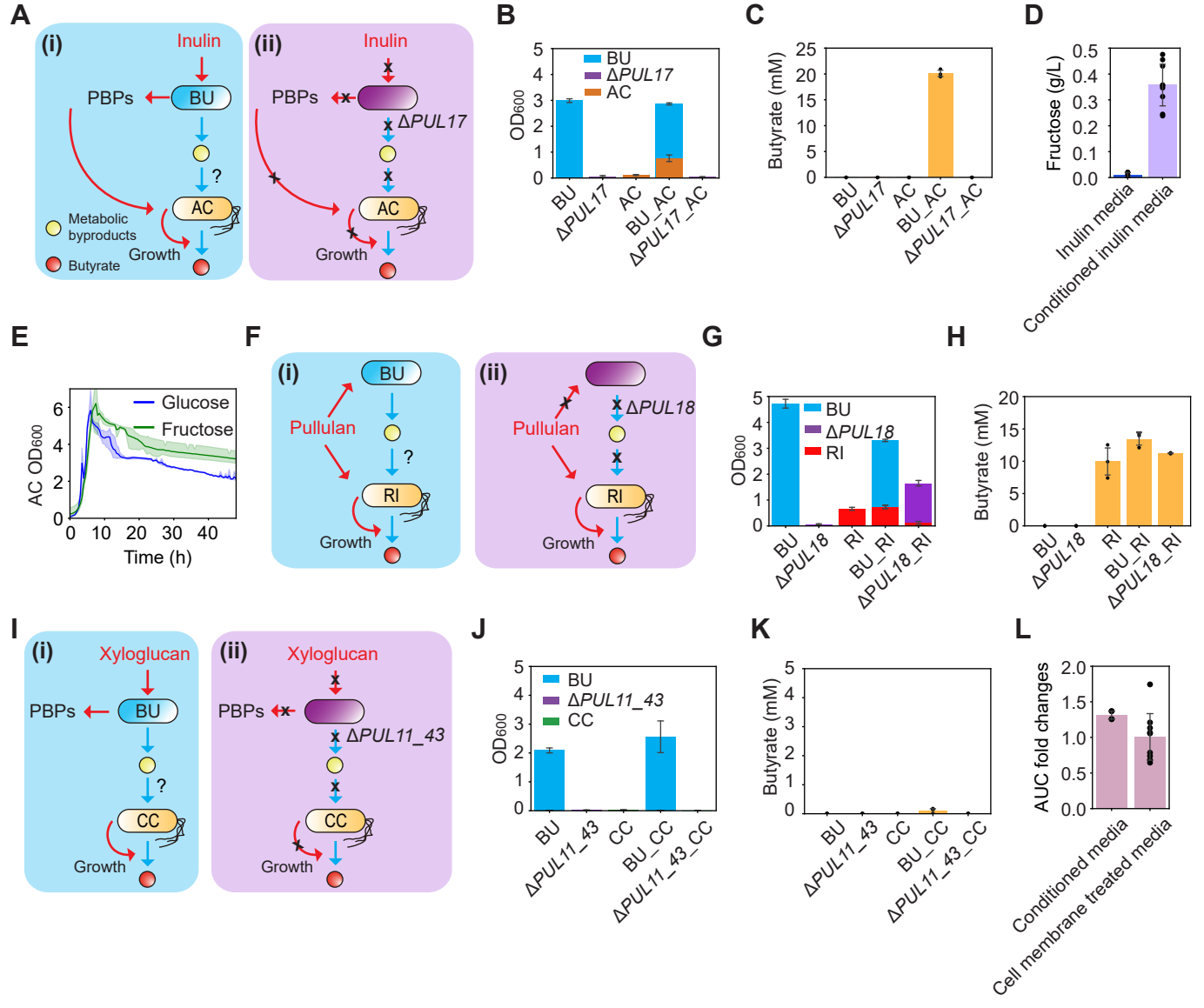

A
tryP
NNNN Barcodes
CRISPR array
Plasmid for barcode generation
p15A pB8-51 lacI Fncpf1
Barcode-tagged mutants
Homology arms
CRISPR array
Plasmid for PUL deletion
pB8-51 recT lacI Fncpf1
PUL deletion
BU ΔPUL6
BU ΔPUL7
BU ΔPUL54
Barcoded library of PUL deletion mutants
OD600
Time
Growth characterization in glycan media
B
Growth model
Steady-state abundance fold change
Glucomannan
Glycogen
Type II mucin
Arabinogalactan
Galactomannan
Pectic galactan
Xyloglucan
Inulin
Laminarin
Glucose
Pectin
Pullulan
BU ΔPUL22_23
BU ΔPUL7
BU ΔPUL11
BU ΔPUL47
BU ΔPUL32
BU ΔPUL49
BU ΔPUL34
BU ΔPUL16
BU ΔPUL13
BU ΔPUL6
BU ΔPUL43
BU ΔPUL24
BU ΔPUL28
BU ΔPUL54
BU ΔPUL44
BU ΔPUL35
BU ΔtryP-24
BU ΔPUL21
BU ΔPUL17
BU ΔPUL48
BU ΔPUL12
BU ΔPUL11_43
BU ΔPUL37
BU ΔPUL18
C
PUL11_43
PUL12
PUL17
PUL18
PUL21
Glycogen
Glucomannan
Galactomannan
Inulin
Laminarin
Pectic galactan
Pectin
Pullulan
Xyloglucan
Type II mucin
Steady-state abundance less than 40% of BU ΔtryP-24
Growth rate less than 20% of BU ΔtryP-24
Steady-state abundance less than 40% & Growth rate less than 20% of BU ΔtryP-24
D
Galactomannan
Glucomannan
Inulin
Type II mucin
Glycogen
Pectic galactan
Pectin
Pullulan
Laminarin
Xyloglucan
OD600
Time (h)
BU ΔtryP-24
BU ΔPUL12
BU ΔPUL17
BU ΔPUL18
BU ΔPUL21
BU ΔPUL11_43

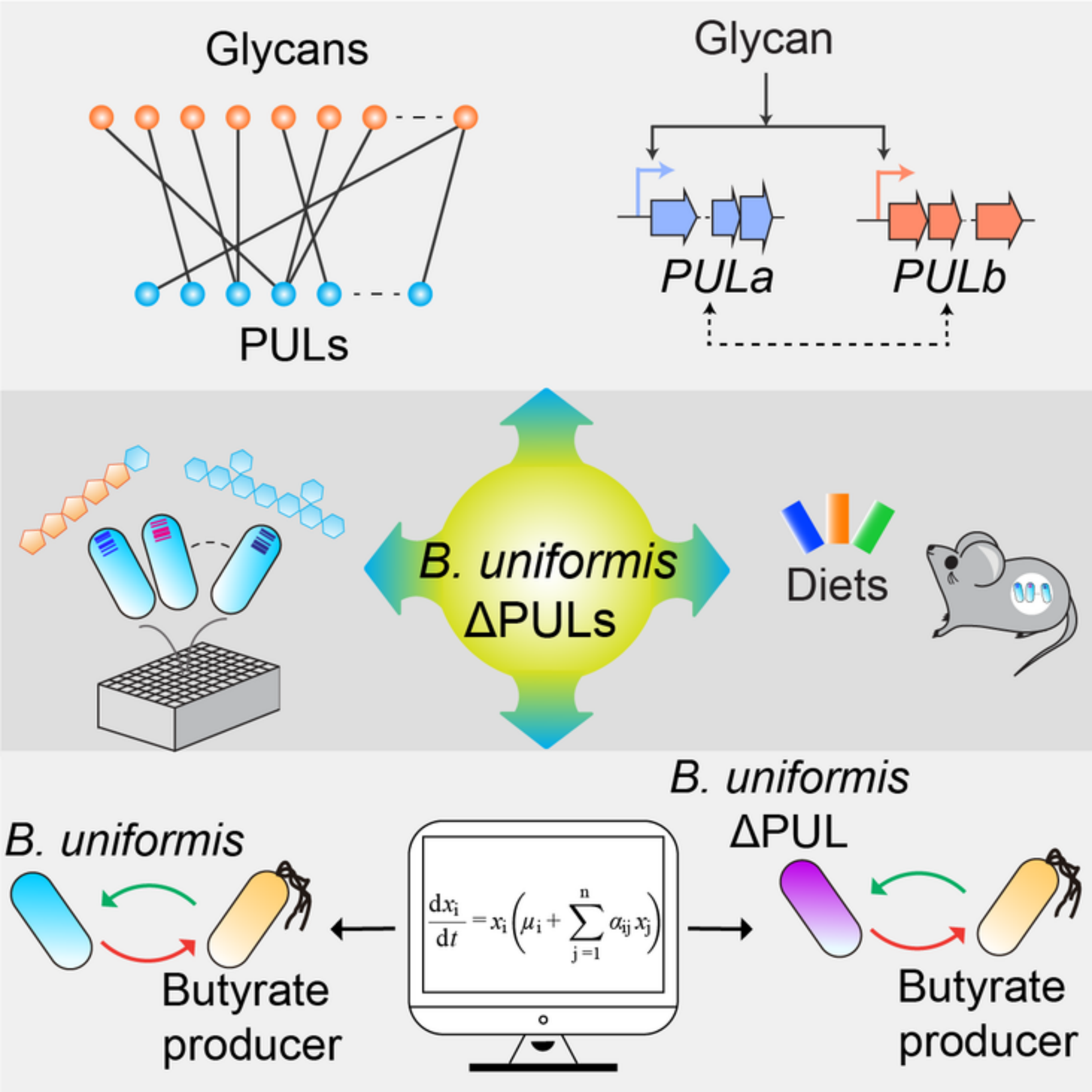

A
PUL11
PUL43
PUL44
SusC SusD GH43 GH5 HTCS GH31 GH2 GH3 GH95
BACUNI_RS01865 BACUNI_RS01870 BACUNI_RS01885 BACUNI_RS01890 BACUNI_RS01895 BACUNI_RS01900 BACUNI_RS01905 BACUNI_RS01910 BACUNI_RS01915
GH43 GH2 HTCS SusC SusD SBP GH5 GH97 GH2 GH42 GH2 GH29 GH31 GH43
BACUNI_RS15325 BACUNI_RS15330 BACUNI_RS15336 BACUNI_RS15340 BACUNI_RS15345 BACUNI_RS15350 BACUNI_RS15355 BACUNI_RS15360 BACUNI_RS15365 BACUNI_RS15370 BACUNI_RS15375 BACUNI_RS15380 BACUNI_RS15385 BACUNI_RS15390
SusC SusD GH43 GH144 GH3 SusC SusD GH3 GH144
BACUNI_RS15425 BACUNI_RS15430 BACUNI_RS15435 BACUNI_RS15440 BACUNI_RS15445 BACUNI_RS15450 BACUNI_RS15455 BACUNI_RS15460 BACUNI_RS15470
B
BU ΔtryP-24
BU ΔPUL44
BU ΔPUL11
BU ΔPUL11_43
BU ΔPUL43
BU ΔPUL11_44
Glucose
OD600
Time (h)
Xyloglucan
C
RPKM
(Glucose)
(Xyloglucan)
PUL11
PUL43
PUL44
D
log2 (Fold change)
ΔPUL11 vs ΔtryP-24
ΔPUL43 vs ΔtryP-24
ΔPUL11_44 vs ΔtryP-24
E
BACUNI_RS15340
Fold changes
BU ΔtryP-24
BU ΔPUL11
BU ΔPUL11_44
F
PUL54 PUL52 PUL47 PUL43 PUL32 PUL30 PUL20 PUL16 PUL14 PUL13 PUL12 PUL10
BU ΔPUL11
BU ΔPUL43
BU ΔPUL44
BU ΔPUL11_44
Log2 (Fold change)


A
Mutant pool: BU ΔtryP-24 and BU ΔPULs
Inoculation
24h
1:20 dilution
Seed culture (S)
Passage 0
Passage 1
Passage 2
OD600
NGS
B
Pectic galactan
Glucose
Inulin
Laminarin
Glycogen
Pullulan
Xyloglucan
Galacto-mannan
Arabino-galactan
Type II mucin
Gluco-mannan
Pectin
Relative abundance
OD600
Passages
BU ΔtryP-24
BU ΔPUL6
BU ΔPUL7
BU ΔPUL11
BU ΔPUL12
BU ΔPUL13
BU ΔPUL16
BU ΔPUL17
BU ΔPUL18
BU ΔPUL21
BU ΔPUL22_23
BU ΔPUL24
BU ΔPUL28
BU ΔPUL32
BU ΔPUL34
BU ΔPUL35
BU ΔPUL37
BU ΔPUL43
BU ΔPUL44
BU ΔPUL47
BU ΔPUL48
BU ΔPUL49
BU ΔPUL54
C
Galactomannan
BU ΔPUL12
Glucomannan
BU ΔPUL12
Inulin
BU ΔPUL17
Glycogen
BU ΔPUL18
Pectic galactan
BU ΔPUL18
Pectin
BU ΔPUL18
Pullulan
BU ΔPUL18
Type II mucin
BU ΔPUL18
Laminarin
BU ΔPUL21
Xyloglucan
BU ΔPUL11
Xyloglucan
BU ΔPUL43
Absolute abundance (OD600)
Passages
D
Shannon diversity
3.15
3.00
2.85
2.70
Passages
Pectic galactan
Glycogen
Galactomannan
Type II mucin
Arabinogalactan
Pullulan
Glucomannan
Laminarin
Pectin
Xyloglucan
Inulin
Glucose


A
Day 0,
oral gavage
1 Week
2 Weeks
High fiber diet
High fat diet
Fiber free diet
Inulin
Glucomannan
Pectic galactan
B
log10(CFU.g-1 cecum content)
BU ΔPUL18
BU ΔPUL37
High fiber diet
High fat diet
Fiber free diet
FFD-Inulin
FFD-Glucomannan
FFD-Pectic galactan
C
(i)
Relative abundance
Shannon diversity
Days
BU ΔtryP-24
BU ΔPUL6
BU ΔPUL7
BU ΔPUL11
BU ΔPUL12
BU ΔPUL13
BU ΔPUL16
BU ΔPUL17
BU ΔPUL18
BU ΔPUL21
BU ΔPUL22_23
BU ΔPUL24
BU ΔPUL28
BU ΔPUL32
BU ΔPUL34
BU ΔPUL35
BU ΔPUL37
BU ΔPUL43
BU ΔPUL44
BU ΔPUL47
BU ΔPUL48
BU ΔPUL49
BU ΔPUL54
(ii)
Max slope
D
Component 2 (32%)
Component 1 (64%)
PUL18
PUL37
E
BU ΔPUL12
BU ΔPUL17
BU ΔPUL18
BU ΔPUL37
BU ΔtryP-24
Relative abundance
Days


A
AC
CC
ER
RI
BU WT
BU ΔPUL17
Monocultures
BU WT and BP co-culture
BU ΔPUL17 and BP co-culture
Decipher inter-species interactions
BU
BP
OD600
Time (h)
B
Glucose
Glycogen
Glucomannan
Inulin
Laminarin
Pectic galactan
Pectin
Pullulan
Xyloglucan
BU ΔPUL18
BU ΔPUL12
BU ΔPUL17
BU ΔPUL21
BU ΔPUL11_43
C
R=0.90, p=7.1e-18
Actual butyrate (mM)
Predicted butyrate (mM)
BP monoculture
BU-BP coculture
BU ΔPUL-BP coculture


A
(i)
Inulin
PBPs
BU
?
AC
Metabolic byproducts
Growth
Butyrate
(ii)
Inulin
PBPs
ΔPUL17
AC
Growth
B
OD600
BU
ΔPUL17
AC
BU
ΔPUL17
AC
BU_AC
ΔPUL17_AC
C
Butyrate (mM)
BU
ΔPUL17
AC
BU_AC
ΔPUL17_AC
D
Fructose (g/L)
Inulin media
Conditioned inulin media
E
AC OD600
Glucose
Fructose
Time (h)
F
(i)
BU
Pullulan
?
RI
Growth
(ii)
ΔPUL18
Pullulan
RI
Growth
G
OD600
BU
ΔPUL18
RI
BU
ΔPUL18
RI
BU_RI
ΔPUL18_RI
H
Butyrate (mM)
BU
ΔPUL18
RI
BU_RI
ΔPUL18_RI
I
(i)
Xyloglucan
PBPs
BU
?
CC
Growth
(ii)
Xyloglucan
PBPs
ΔPUL11_43
CC
Growth
J
OD600
BU
ΔPUL11_43
CC
BU
ΔPUL11_43
CC
BU_CC
ΔPUL11_43_CC
K
Butyrate (mM)
BU
ΔPUL11_43
CC
BU_CC
ΔPUL11_43_CC
L
AUC fold changes
Conditioned media
Cell membrane treated media


Glycans
PULs
Glycan
PULa
PULb
B. uniformis
ΔPULs
Diets
B. uniformis
B. uniformis
ΔPUL
$\frac{dx_i}{dt} = x_i\left(\mu_i + \sum_{j=1}^{n} \alpha_{ij} x_j\right)$
Butyrate
producer
Butyrate
producer