

OPEN

Engineering dual-glycan responsive expression systems for tunable production of heterologous proteins in *Bacteroides thetaiotaomicron*

Darryl R. Jones^{1,3}, Marshall B. Smith^{1,2,3}, Richard McLean¹, Julie M. Grondin¹,
Carolyn R. Amundsen¹, G. Douglas Inglis^{1,2}, Brent Selinger² & D. Wade Abbott^{1,2*}

Genetically engineering intestinal bacteria, such as *Bacteroides thetaiotaomicron* (*B. theta*), holds potential for creating new classes of biological devices, such as diagnostics or therapeutic delivery systems. Here, we have developed a series of *B. theta* strains that produce functional transgenic enzymes in response to dextran and arabinogalactan, two chemically distinct glycans. Expression systems for single glycan induction, and a novel “dual-glycan” expression system, requiring the presence of both dextran and arabinogalactan, have been developed. In addition, we have created two different chromosomal integration systems and one episomal vector system, compatible with engineered recipient strains, to improve the throughput and flexibility of gene cloning, integration, and expression in *B. theta*. To monitor activity, we have demonstrated the functionality of two different transgenic enzymes: NanoLuc, a luciferase, and BuGH16C, an agarase from the human intestinal bacterium, *Bacteroides uniformis* NP1. Together this expression platform provides a new collection of glycan-responsive tools to improve the strength and fidelity of transgene expression in *B. theta* and provides proof-of-concept for engineering more complex multi-glycan expression systems.

The lower digestive tract of mammals contains diverse and abundant bacterial species, collectively referred to as the distal gut microbiota (DGM)¹. The total collection of genes encoded within the DGM greatly outnumbers the host genome, and encode essential functions, such as digestion of dietary fiber and the synthesis of vitamins and amino acids². In recent years, public and scientific interest in the DGM has intensified as mechanistic relationships between the DGM and host health have begun to be defined³. The DGM is required for intestinal and immune system development and function⁴ and can reduce pathogen colonization by competitive exclusion⁵. Imbalances in community structure, termed dysbiosis⁶, have been linked with chronic health conditions such as diabetes⁷, obesity⁸, and inflammatory bowel disease⁹. These issues have been exacerbated by Western diets, which are low in dietary fiber¹⁰, and the prophylactic and therapeutic use of broad spectrum antibiotics¹¹.

Bacteroides thetaiotaomicron (*B. theta*) is a prominent member of the human DGM. It is detectable in nearly half of healthy Western fecal microbiomes¹² and is one of the most studied members of Bacteroidetes¹³. Members of this phylum are known for their ability to saccharify a wide variety of chemically complex polysaccharides from diverse biological sources, including terrestrial plants², seaweed¹⁴, bacteria¹⁵, fungi¹⁶, and animals¹⁷. Glycan metabolism in Bacteroidetes is genetically encoded within polysaccharide utilization loci (PULs)¹⁸, which are clusters of co-regulated genes typically comprised of TonB-dependent SusC/SusD-like transporter systems, regulatory proteins, and carbohydrate active enzymes (CAZymes) specific for the consumption of discrete substrates. PULs are activated by one of three classes of regulators, including SusR-like, AraC hybrid two-component systems (HTCSs), and extracytoplasmic function (ECF) σ /anti- σ systems^{19,20}. Although these regulatory proteins

¹Lethbridge Research and Development Centre, Agriculture and Agri-Food Canada, Lethbridge, Alberta, T1J 4B1, Canada. ²Department of Biological Sciences, University of Lethbridge, Lethbridge, Alberta, T1K 3M4, Canada.

³These authors contributed equally: Darryl R. Jones and Marshall B. Smith. *email: wade.abbott@canada.ca

operate through different mechanisms they serve an analogous function within the cell, which is to detect products enzymatically released from a substrate and induce PUL expression.

Augmenting the metabolism of intestinal bacteria, such as *B. theta*, with transgenic enzymes and transporters is a promising approach for establishing and reprogramming established DGM structure. The introduction of enzymes with new functions or that operate at faster rates can release more fermentable sugars²¹, reduce complications associated with slow transit of ingesta^{22,23}, and provide a selective advantage in a nutrient competitive ecosystem. Previously, the transfer of a fructan specific SusC/SusD-like transporter endowed the recipient *B. theta* strain with the ability to metabolize a chemically distinct glycan²⁴. Recently, *Bacteroides stercoris* and *B. theta* were engineered to utilize porphyran, a structural polysaccharide found in the cell walls of red algae²⁵. This transfer provided access to a “privileged” nutrient, and selective feeding on porphyran enabled engraftment of the engineered strain into an established microbiota. Such engineering strategies could usher in a new era of personalized medicine and microbiome engineering.

Despite these advances, genetic elements for tunable and tightly controlled expression of heterologous proteins in bacterial strains adapted for long-term host colonization are still required. *Bacteroides* spp. have unique regulatory and conjugation elements^{25–28}, and genetic tools developed for other organisms generally are not tractable in *B. theta*. In this regard, several *Bacteroides*-specific tools have been developed. The pExchange plasmid for homologous recombination has existed since 1991²⁹. Using this counter-selectable system, unmarked chromosomal deletions and insertions can be made with single base pair resolution^{30,31}. Alternatively, pNBU2, a plasmid containing a mobilizable transposon element³² facilitates integration of the entire plasmid indiscriminately into one of two tRNA^{SER} sites³⁰. More recently, constitutive promoter elements and ribosome binding sites (RBS) have been screened and ranked to provide a wide-range of expression levels for heterologous proteins^{31,33}. Episomal and chromosomal inducible gene expression systems also exist, and they can be regulated by anhydrotetracycline, a synthetic chemical inducer³¹, or the polysaccharide α -mannan³⁴. Activation of each of these systems requires the addition of a single inducer for transgene expression.

To increase the fidelity of glycan responsive episomal and chromosomal expression systems in *B. theta* we have altered the natural regulatory mechanisms of PULs by placing them behind dextran (DX)³⁵ and arabinogalactan (AG)²⁰ responsive promoters. In addition, we have engineered a first in-in-class dual-glycan expression platform that is dependent upon the presence of two chemically distinct glycans that induce a minor diauxic growth phenotype in *B. theta*³⁶. This dual-glycan expression system provided improved regulatory control of transgene expression in *B. theta* and enhanced the level of activity of two different reporter enzymes (e.g. a luciferase and agarase) from two different genomic loci and an expression plasmid. Importantly, most of these genetic modifications were not deleterious to bacterial fitness *in vitro*. This study provides unique insights into promoter engineering for improved control of gene expression in *B. theta*, and presents a platform for the future development of other multi-glycan responsive expression systems.

Results

Construction of a chromosomal platform for targeted insertion and expression of transgenes.

In developing a recombinant chromosomal expression system, we sought to recommission the genome space of a characterized PUL. *B. theta* PUL75 (Fig. 1a) is dedicated to the metabolism of homogalacturonan (HG), a pectic glycan comprised of galacturonic acid³⁷. Compared to the wild-type strain (WT), *B. theta* with a deletion of PUL75 (*B. theta*- Δ PUL75; base pairs: 5371714...5408988) was severely restricted in its ability to grow on minimal medium (MM) containing HG as sole carbon source (Fig. 1b;³⁷). The vacated genome space in *B. theta*- Δ PUL75 was targeted for chromosomal integration using a complementary PUL75-5' flank and PUL75-3' flank sequence inserted into the pExchange backbone (Fig. 1c). In addition, a multiple cloning site (MCS) flanked by an upstream RBS, downstream terminator, and inducible or constitutive promoters were placed between the PUL75-5' and PUL75-3' sequences. This vector series, referred to as pINTEGRATE (pINT), enables precise and unmarked insertion of expression cassettes into the chromosome of *B. theta*- Δ PUL75. To complement this vector series we developed a pNBU2 plasmid series in parallel that contained an identical MCS and expression elements³⁰. The pNBU2 vectors undergo discriminatory single-cross over integration of the entire plasmid into the *B. theta* genome at one of two tRNA^{SER} genes. This system enables faster evaluation of chromosomal transgene function from a second genome locus, but the modified strain is left with residual plasmid DNA and selectable markers in its genome.

Validation of *B. theta* promoters for regulation of transgenes.

Three different promoters were selected for insertion into pINT and pNBU2 vectors (Fig. 1c) based upon previous studies. These included a strong constitutive promoter P_{ON} (*bt1311* promoter and the *rpil** RBS³³), and two inducible promoters: the DX responsive “P_{DX}”, which drives expression of the PUL48 *susC*-like gene *bt3090*³⁵ (Fig. 2a); and the AG responsive “P_{AG}”, which drives expression of the PUL5 *susC*-like gene *bt0268*³³ (Fig. 2b). DX and AG were selected because they induce some of the strongest glycan responsive expression in *B. theta*, have distinct chemical structures (Fig. 2a,b), are from diverse biological sources (DX is from a bacterial glycan; AG is a plant cell wall glycan), and are induced by distinct regulatory mechanisms; PUL48 uses a SusR-like system (Fig. 2a) and PUL5 uses a HTCS system (Fig. 2b). Despite performing analogous functions in the cell, the SusR-like and HTCS regulators are structurally unrelated¹⁹. SusR-like proteins are relatively rare, with only five homologs identified in *B. theta*, two of which regulate the PULs involved in the metabolism of chemically distinct glucans¹⁹. The first SusR protein was discovered in the Sus PUL³⁸ and its mechanism remains undefined. HTCSs are the most abundant PUL regulatory proteins in *B. theta*¹⁹. They are fused integral membrane proteins with a periplasmic exposed carbohydrate binding domain that operate through a “scissor blade-like” mechanism³⁹ and a cytoplasmic AraC DNA-binding domain. Following carbohydrate binding, HTCSs undergo a histidine-aspartate phosphorylation cascade and bind DNA as a dimer^{36,40,41}.

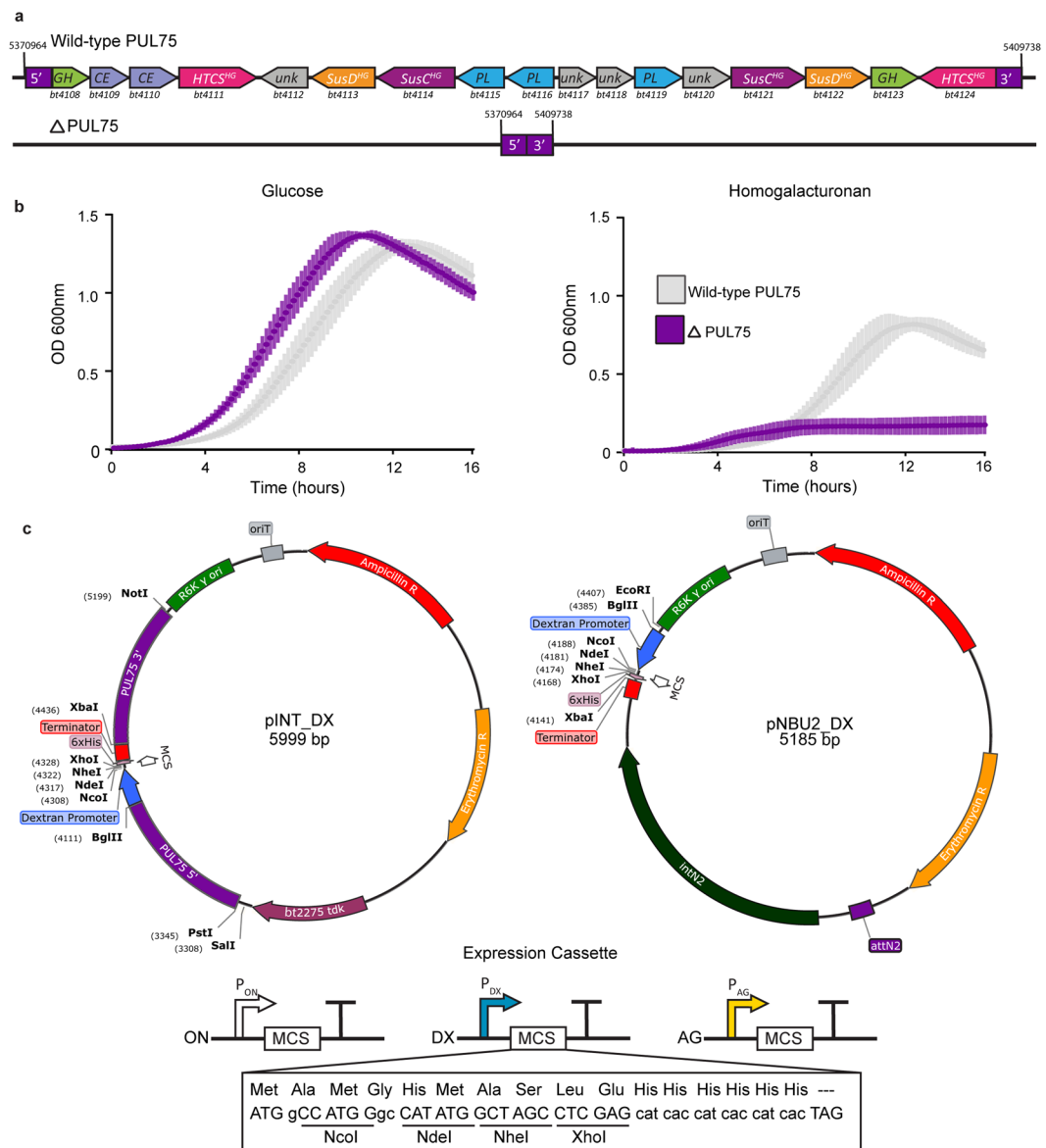


Figure 1. Target genomic locus and integrative expression vectors. **(a)** Schematic of *B. theta* genomic locus PUL75 containing genes required for homogalacturonan (HG) utilization (top). These genes are deleted in the *B. theta* Δ PUL75 strain (bottom). **(b)** Growth profiles of wild-type *B. theta* (grey) and the Δ PUL75 strain (purple) in MM containing either glucose (GLC) or HG as sole carbon source. Growth was measured by OD₆₀₀ readings over the course of 16 h. Values represent the average of three biological replicates. Variation is represented by standard error of the mean (s.e.m.) **(c)** Plasmid maps of the pINT (left) and pNBU2 (right) vectors. The orientation of the DX, AG, and ON, expression cassettes and the sequence of the MCS are indicated.

The ability of each promoter to drive transgene expression was determined using NanoLuc, a luciferase, as a reporter gene³³. To evaluate proximity effects and differential expression platforms, NanoLuc-pINT and NanoLuc-pNBU2 vectors were conjugated into *B. theta*. Luciferase activity assays established that although genome location resulted in statistical differences in luciferase activity under some conditions, they were small when compared to glycan-induced expression (Fig. S1, $P < 0.05$ in 12/40 comparisons). When grown on glucose (GLC), DX, or AG, the P_{ON} constructs consistently produced luminescence at levels that were three orders of magnitude ($P < 0.01$) above the baseline produced by strains without the NanoLuc gene (Fig. 2c), and P_{DX} and P_{AG} constructs responded to DX or AG in a dose dependent fashion, respectively (Fig. 2d). Intriguingly, P_{AG} displayed higher activity in the presence of GLC and AG than with AG alone (Fig. 2e).

Engineering tunable expression of glycan-responsive PUL regulator proteins in *B. theta*. Next, we attempted to increase the heterologous production of NanoLuc in *B. theta*, by modifying the induction status of the PUL48 and PUL5 regulatory proteins within the cell. Typically, regulatory proteins, such as HTCS

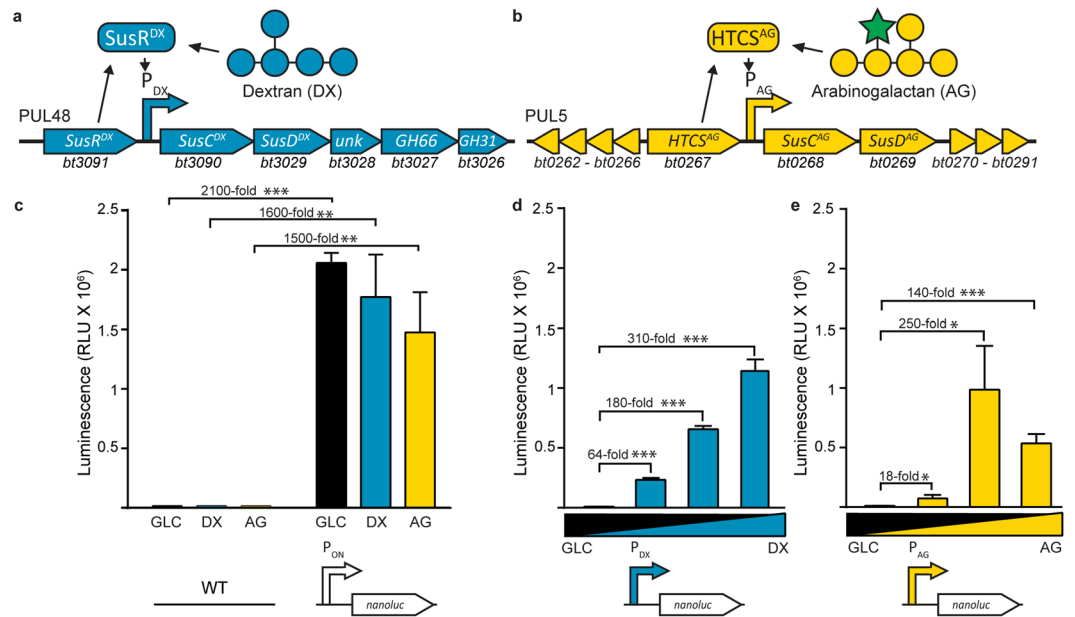


Figure 2. Inducible expression cassettes generate a dose dependent response in NanoLuc activity in *B. theta*. **(a)** Schematic of wild-type regulation by PUL48. P_{DX} is activated by the SusR-like protein SusR^{DX} in the presence of DX. **(b)** Schematic of wild-type AG regulation. P_{AG} from PUL5 is activated by the HTCS protein HTCS^{AG} in response to the presence of AG. Schematic representations of monomeric glucose (GLC; blue circle) the polymer dextran (DX) composed of a chain of GLC; and arabinogalactan (AG), composed of a galactose (GAL, yellow circle) backbone and arabinose side chains (green star). **(c)** NanoLuc activity under constitutive expression with the P_{ON} promoter is measured post induction with GLC (black bar), DX (blue bar), and AG (yellow bar). Luminescence signal for NanoLuc production under the **(d)** DX or **(e)** AG inducible promoters in the presence of increasing concentrations of DX or AG, respectively. Values in **(d-f)** represent the average and s.e.m. for three biological replicates. Asterisks represent levels of significance (**p* < 0.05, ***p* < 0.01, ****p* < 0.001).

and SusR-like homologs, are constitutively expressed at low levels and are not induced along with CAZymes and transport proteins during glycan metabolism³⁵. We reasoned that placing strong constitutive or inducible promoters upstream of the regulator gene in PUL5 and PUL48 may create feedback loops that result in higher levels of expression. Therefore, *B. theta* strains with a series of different promoter architectures were generated (Fig. S2). These included: (1) a constitutive promoter, (2) a positive feedback loop (DX → DX; AG → AG), and (3) alternate forms of hybrid-induction (i.e. DG → AG and AG → DX). (1) The DX constitutive strain was created by placing the P_{ON} *bt1311* promoter³³ upstream of *bt3091*, the PUL48 SusR-like gene (SusR^{DX}), to generate the strain *B. theta*-P_{ON}SusR^{DX}. The AG constitutive strain was created by placing P_{ON} upstream of *bt0267*, the PUL5 HTCS gene (HTCS^{AG}), to generate *B. theta*-P_{ON}HTCS^{AG}. (2) The positive feedback systems were created by placing P_{DX} upstream of SusR^{DX} to create *B. theta*-P_{DX}SusR^{DX}, and P_{AG}³³ upstream of HTCS^{AG} to create *B. theta*-P_{AG}HTCS^{AG}. (3) The hybrid-induction systems were created by inserting P_{AG} in front of SusR^{DX} and P_{DX} in front of HTCS^{AG} to create *B. theta*-P_{AG}SusR^{DX}, and *B. theta*-P_{DX}HTCS^{AG}, respectively. In addition, a third hybrid strain was created, *B. theta*-P_{AG}SusR^{DX} + P_{DX}HTCS^{AG}, which contained both hybrid regulatory modifications (Fig. S2).

To determine the effect of the modifications to the regulatory proteins of PUL5 and PUL48, the growth profiles of each *B. theta* strain were determined (Fig. 3a). Strains with constitutive regulation (*B. theta*-P_{ON}SusR^{DX} and *B. theta*-P_{ON}HTCS^{AG}) and positive feedback regulation both performed similar to wild-type. Each hybrid strain displayed a growth defect when cultured solely on the glycan sensed by the regulatory protein (DX for P_{AG}SusR^{DX}, or AG for P_{DX}HTCS^{AG}, Fig. 3a); however, this defect was mitigated when the strains were cultured on a mixture of DX and AG. This phenotype may be a result of the hybrid strains having a lower basal level of expression when regulated by a glycan-responsive promoter. Whereas on the mixture, the promoters are fully activated and lead to sufficient regulatory protein production to effectively activate their PULs preventing a growth defect.

To examine the effects of the genetic modifications to the regulation of PUL5 and PUL48 on the expression of a transgene, NanoLuc activity was measured in the engineered *B. theta* strains. When placed under SusR^{DX} regulation, the P_{ON} and P_{DX} promoter increased activity approximately three-fold over wild-type in both the pINT (Fig. 3b) and pNBU2 (Fig. S1) platforms (*P* < 0.01). The hybrid strain, *B. theta*-P_{AG}SusR^{DX}, displayed the lowest NanoLuc signal on DX but similar levels on the DX-AG mixture (*P* < 0.001), which is consistent with what was observed in the growth profiles (Fig. 3a). The AG-responsive systems displayed similar patterns, with a lower increase in NanoLuc activity in the P_{ON} and P_{AG} systems (Fig. 3c, *P* < 0.01).

In contrast to the strains with single modifications, the strain with both hybrid regulator systems (*B. theta*-P_{AG}SusR^{DX} + P_{DX}HTCS^{AG}) displayed noticeable increases (*P* < 0.01) in NanoLuc activity (Fig. 3d). This strain produced the highest luminescence values (*P* < 0.001) for chromosomal NanoLuc in this study when

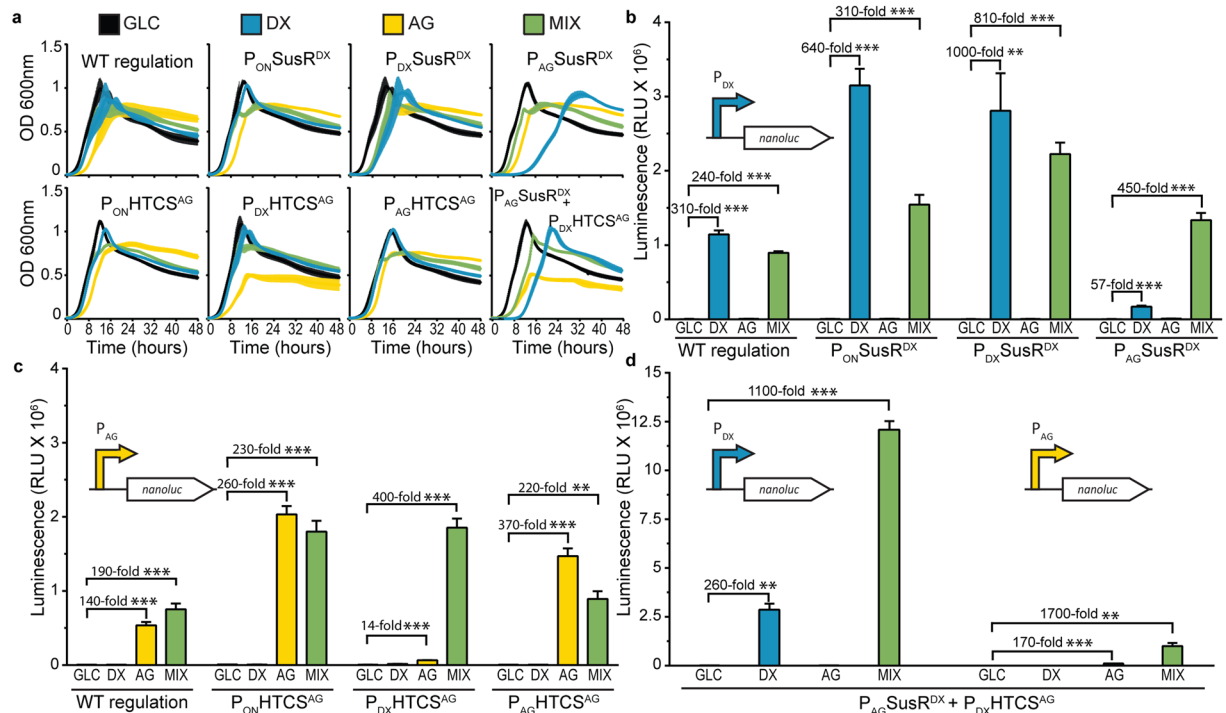


Figure 3. Engineering regulatory elements alters growth profiles, levels of gene activity, and gene induction requirements. **(a)** Growth profiles of WT *B. theta* and engineered regulatory strains on GLC (black), DX (blue), AG (yellow), and a mixture of DX and AG (green). **(b)** Luminescence detected from DX-responsive NanoLuc in strains with engineered *SusR^{DX}* when grown on GLC (black), DX (blue), AG (yellow), and the mixture of DX and AG (green). **(c)** Luminescence detected from AG-responsive NanoLuc in strains with engineered *HTCS^{AG}* when grown on GLC (black), DX (blue), AG (yellow), and the mixture of DX and AG (green). **(d)** Luminescence detected from both DX- and AG-responsive NanoLuc in a strain with engineered *SusR^{DX}* and *HTCS^{AG}* when grown on GLC (black), DX (blue), AG (yellow), and the mixture of DX and AG (green). Error bars in **(a)** represent standard deviation (s.d.) and error bars in **(b–d)** represent s.e.m. for three biological replicates. Asterisks represent levels of significance (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

NanoLuc was placed behind P_{DX} and the strain was cultured on mixed sugars. Although it did not reach the same absolute value for reporter activity, NanoLuc positioned behind P_{AG} exhibited the greatest relative difference compared to baseline luminescence on GLC ($P < 0.01$), owing to the high baseline of P_{DX} NanoLuc. These results suggest that DX responses are preferential over AG responses in *B. theta* for these systems.

Augmenting transgene expression using a *B. theta* episomal expression vector. Plasmid-based expression systems hold several advantages over chromosomal expression systems, including efficient transformation, higher gene copy numbers, and shortened timelines due to lack of requirement for double cross-over selection methods. Several expression vectors compatible with *Bacteroides* spp. have been reported^{31,34}. To compare the potency of the DX and AG regulated expression cassettes developed here with previously described vector expression systems, we created the pEpisomalPromoter (pEP) series of plasmids for *B. theta* (Fig. 4a). The pEP series contain the *mobA* and *repA* genes derived from the *Bacteroides* plasmid pBI143 DNA⁴¹ to allow the vector to persist and replicate independent of the *B. theta* chromosome. Vectors are maintained by erythromycin selection and are equipped with the P_{DX} , P_{AG} , or P_{ON} expression cassettes from pINT to select glycan responsiveness. As a proof of concept, the NanoLuc gene was cloned into a pEP vector under the control of the DX promoter, and expression was induced by culturing the cells with GLC, DX, AG or a mixture of DX/AG. Plasmid-based expression resulted in higher ($P < 0.001$) luminescence values when compared to chromosomal expression (Fig. 4b). Combining plasmid-based expression with reciprocal modifications to hybrid promoter regulation amplified ($P < 0.001$) effects for selective expression when compared with chromosomal NanoLuc expression (Fig. 4c). The *B. theta*- P_{DX} -*SusR^{DX}* and *B. theta*- P_{ON} -*SusR^{DX}* strains transformed with a P_{DX} -pEP vector had similar levels ($P < 0.001$) of luminescence ($60\text{--}75 \times 10^6$ RLU) on DX and DX-AG mix, suggesting DX levels were sufficient to maximize expression in both systems. In comparison, the *B. theta*- P_{AG} -*SusR^{DX}* strain only displayed activation to similar levels when grown on a mixture of DX and AG glycans ($P < 0.001$). This suggests that basal levels of *SusR^{DX}* when grown on pure DX was not enough to drive plasmid expression; however, this effect was overcome by adding AG to the medium (Fig. 4b). An approximately two-fold amplification ($P < 0.001$) of NanoLuc activity over the other strains was observed in the *B. theta*- P_{AG} -*SusR^{DX}* + P_{DX} -*HTCS^{AG}* strain when grown on the AG-DX mix. This suggests that inducing both regulators results in higher levels of expression of heterologous genes using the pEP-plasmid system. *B. theta*- P_{AG} -*SusR^{DX}* + P_{DX} -*HTCS^{AG}* transformed with P_{DX} -pEP had the highest levels of

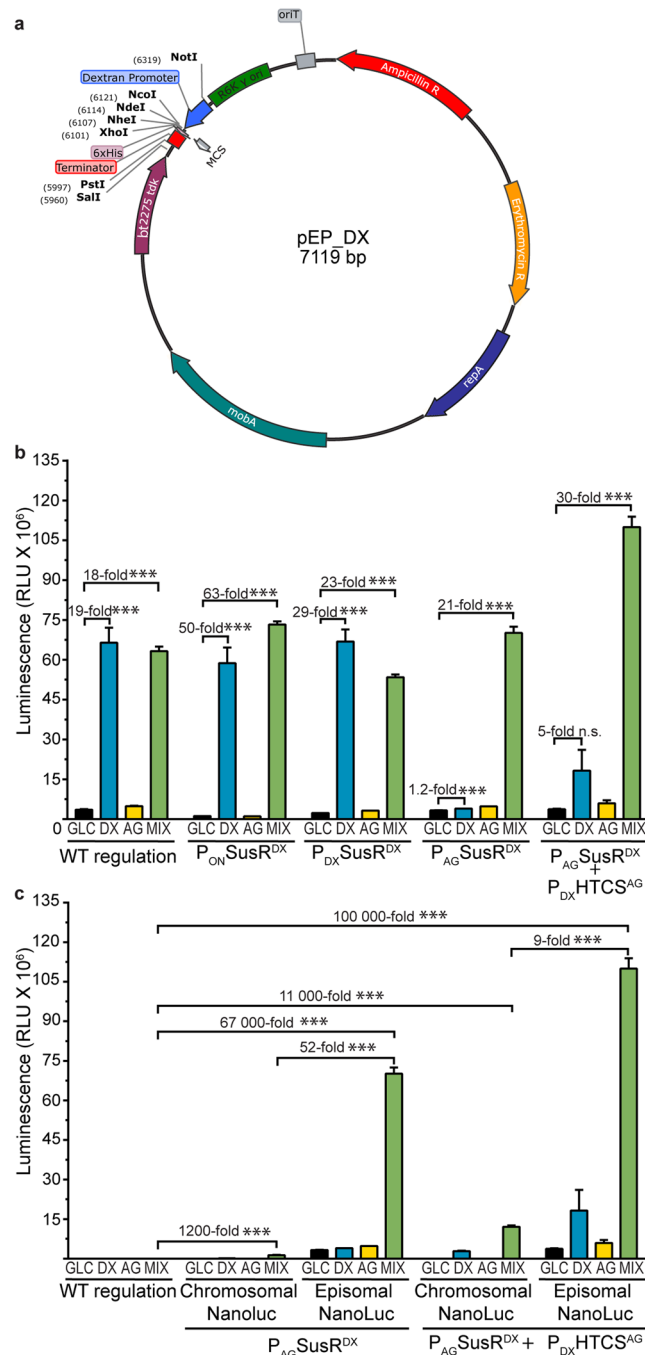


Figure 4. Episomal transgene expression results in higher enzyme activity when compared to chromosomal-based expression. **(a)** Vector map of the extra-chromosomal expression pEP_DX vector. **(b)** Luminescence signal detected from wild-type (WT), $P_{ON}SusR^{DX}$, $P_{DX}SusR^{DX}$, $P_{AG}SusR^{DX}$, and $P_{DX}HTCS^{AG} + P_{AG}SusR^{DX}$ strains containing the NanoLuc under the control of the episomal pEP_DX promoter, in the presence of GLC (black bars), DX (blue bars), AG (yellow bars) or the DX and AG mixture (green bars). **(c)** Comparison of luminescence from engineered regulatory strains when NanoLuc is encoded in a plasmid or in the chromosome. Error bars represent s.e.m. for three biological replicates. Asterisks represent levels of significance (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

reporter activity in this study. Use of the plasmid with this strain culminated in an order of magnitude increase over the chromosomal system and a six-order of magnitude increase over the background luminescence of *B. theta* WT (Fig. 4c).

Heterologous production of an agarase in modified *B. theta* strains. To determine whether the dual-glycan expression systems developed here can be used to produce a heterologous carbohydrate active enzyme, we used the pINT vector to integrate a family 16 glycoside hydrolase from *B. uniformis* NP1 (*BuGH16*)

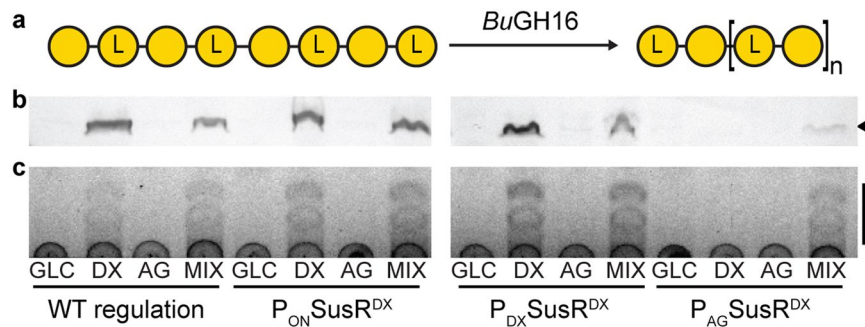


Figure 5. The heterologous *BuGH16* agarase is enzymatically functional in engineered *B. theta* strains. (a) Schematic model of *BuGH16* endo-agarase activity agarose to generate neoagarooligosaccharides (NAOS). NAOS is a repeating disaccharide of L-3,6-anhydro-galactose (L-3,6) and D-galactose, with L-3,6 at the non-reducing end. (b) pINT integrated hexa-histidine tagged *BuGH16* (black arrow) detected by Western blot from induced *SusR^{DX}* strains. (c) Thin-layer chromatography of culture supernatant from induced engineered strains incubated with agarose shows the presence of NAOS (black bar) generated from agarose by active *BuGH16*.

into the *B. theta*-*SusR^{DX}* strain series. *BuGH16* is an endo- β -agarase that cleaves agarose to produce neoagarotetraose as a terminal product⁴² (Fig. 5a). Wild-type *B. theta* does not possess this catalytic activity. *BuGH16*, fused to a N-terminal outer membrane anchoring tag, was cloned into the pINT- P_{DX} vector and integrated into the *B. theta*-*SusR^{DX}* strain. The outer membrane tag enabled whole-cell agarolysis assays to be performed as the enzyme cargo is displayed on the outer surface. For the *B. theta*-P_{ON}*SusR^{DX}* and *B. theta*-P_{DX}*SusR^{DX}* strains, transgenic *BuGH16* protein activity was detected and agarose activity was observed in all conditions containing DX (Fig. 5b). Production of *BuGH16* in the P_{AG}*SusR^{DX}* multi-glycan induction strain only displayed *BuGH16* activity when the cultures were treated with both DX and AG. This pattern is similar to the NanoLuc activity in the dual-glycan pINT (Fig. 3b) and pEP vector (Fig. 4b) systems. Taken together, these results suggest that engineering HTCS and *SusR*-like regulator proteins to respond to mixtures of chemically defined glycan inducers in chromosomal and plasmid-based systems can improve the regulatory fidelity of heterologous enzyme production in *B. theta*.

Discussion

Engineering intestinal bacteria to express heterologous proteins is a promising approach to improve host health. Previous studies with *B. theta* have developed expression cassettes that are constitutive³³, or respond to common glycans^{33,34} or synthetic chemicals³¹. We have expanded upon this palette of genetic parts to develop the pINT, pNBU2, and pEP-based gene expression systems. pINT constructs specifically integrate its genetic cargo into the vacated genome space of *B. theta*- Δ PUL75. This genome space was selected to prevent large increases to the genome size and to repurpose a decommissioned locus known to be involved in metabolism of a glycan. Each expression platform has been designed for convenient subcloning of target genes, tailoring of promoter selectivity, and if desired, altering the targeted integration site within the genome as each DNA segment is flanked by directional restriction sites (Fig. 1c). Three different validated promoters were inserted into these expression cassettes, including P_{ON} (*bt1311*), a strong constitutive promoter; and P_{DX} (*bt3090*) and P_{AG} (*bt0267*) responsive promoters, two selective glycan responsive elements, P_{AG} and P_{DX}, that respond to chemically distinct glycans³⁵ (Figs. 2a,b, S2). Constitutive expression of NanoLuc driven by P_{ON} was 1,500 to 2,000-fold greater than baseline auto-luminescence, on three different carbon sources (Fig. 2c). These results are consistent with previous findings³³, and underpin that this constitutive promoter operates independent of glycan-specific regulatory networks (Fig. 2c). Surprisingly, DX and AG induction of NanoLuc resulted in different activity profiles. DX displays a conventional dose-dependent relationship when mixed with GLC (Fig. 2d); whereas, the AG-responsive promoter displayed optimal NanoLuc activity at 50:50 (GLC:AG), and there was an inhibitory effect observed at when treated solely with AG (Fig. 2e). The AG-responsive promoter used in this study is from PUL5 and was chosen due to previously determined high levels of expression²⁰ and proven activity³³; however, PUL5 is not the sole AG responsive PUL in the *B. theta* genome as PUL65 has also been shown to be involved in AG utilization²⁰. When the PUL65 HTCS is deleted and the PUL5 HTCS is left intact, the mutant strain grows to a higher density (150%) on AG than the WT *in vitro*²⁰. This indicates that PUL65 may have a repressive effect on PUL5 that is alleviated by deleting the PUL65 HTCS, a hypothesis that could explain the repression on nanoLuc observed in this study. These relationships may be exploited to tailor expression profiles using the AG-responsive system.

To explore the potential of augmenting glycan responsive expression levels and providing higher fidelity in controlling regulation, the promoter elements of *SusR^{DX}* and *HTCS^{AG}* were engineered to be constitutively expressed under control of P_{ON}, and in response to DX and AG (Fig. 3). For both *SusR^{DX}* and *HTCS^{AG}*, constitutive expression and positive feedback had the largest effect, boosting NanoLuc activity three-fold and two-fold above WT for DX (Fig. 3b) and AG (Fig. 3c), respectively. We believe this may be due to increases in the cellular pool of regulatory protein available to sense sugar and upregulate target genes. Supporting this, it has previously been observed that increasing the copy number of the canonical *sus* regulatory gene by incorporating it into a plasmid leads to increases in reporter gene activity⁴³. In the case of hybrid-regulation (P_{AG}*SusR^{DX}* and P_{DX}*HTCS^{AG}*), both DX and AG were required for maximal expression. These results suggest that the regulators

are not induced on the single sugars, and when both DX and AG are supplied, the promoters become activated and lead to higher levels of target gene expression. This hybrid effect was more pronounced when both $SusR^{DX}$ and $HTCS^{AG}$ were introduced into the same strain ($B. theta$ - $P_{AG}SusR^{DX} + P_{DX}HTCS^{AG}$) and NanoLuc was induced by DX (Fig. 3d). A similar expression profile was not observed with AG induction, however, suggesting that the expression of AG-associated genes may be deprioritized during DX metabolism and that these hybrid systems are glycan specific. Previously, the impact of multiple glycans pairs, including DX and AG, on the growth phenotype of *B. theta* and HTCS phosphorylation was reported³⁶. Although the diauxic effect varied for individual glycan pairs, a subtle diauxic for DX and AG was observed, supporting that DX is prioritized over AG metabolism³⁵. However, because this growth effect was modest and the growth phenotype on a DX-AG mixture reported here appears to be a blending of the two signal glycan phenotypes rather than a clear diauxic prioritization (Fig. 1a), it suggests that both DX and AG are digested simultaneously. In this regard, other combinatorial or prioritized glycan metabolic cascades^{35,36} could be exploited to engineer diverse spectrums of dual-glycan response systems, thereby providing further plasticity in tuning the expression of transgenic enzymes.

In order to determine if the dual-glycan effects are conserved using expression vectors, we built the pEP vector system, cloned NanoLuc behind P_{DX} , and transformed the $SusR^{DX}$ strain (Fig. 4a). *B. theta- $P_{AG}SusR^{DX}$ produced an approximately 20-fold increase on the mixed glycans compared to expression on pure glycans (Fig. 4b). Encouragingly, this effect was further amplified in the *B. theta- $P_{AG}SusR^{DX} + P_{DX}HTCS^{AG}$ strain (Fig. 3d). The higher luminescence activity observed using pEP-expression is attributed to the potential for higher copy numbers of the expression cassette within the cell. These vectors replicate within the cell independently of the genome as pBI143, the natural plasmid containing the mobilization and replication genes that were inserted into pEP, has been estimated to be maintained at a copy number of 20 per cell⁴¹. Additionally, it may result due to proximal effects of regulation. The *SusR*-like regulatory proteins are integral membrane proteins, which may require proximity to its target genes to maximally activate their expression. Episomal expression vectors may overcome this obstacle by diffusing to the cytoplasmic DNA-binding domain of transmembrane regulatory proteins.**

To demonstrate that the dual-glycan expression system is compatible for a second transgenic enzyme, we replaced NanoLuc with *BuGH16*, an agarase from *B. uniformis* NP1. *BuGH16* cleaves the β -D-galactose-(1,4)-3,6-anhydro-L-galactose bond in agarose⁴² (Fig. 5a). We compared *BuGH16* activity with the P_{DX} promoter in *B. theta* strains with modified $SusR^{DX}$. *BuGH16* production (Fig. 5b) and agarase activity (Fig. 5c) was observed when induced by DX, and DX and AG mixtures in the control, $P_{ON}SusR^{DX}$ and $P_{DX}SusR^{DX}$ strains. Consistent with NanoLuc, dual-glycan regulation was only observed in the $P_{AG}SusR^{DX}$ system when *BuGH16* was induced by a mixture of DX and AG. The observation that the hybrid-systems require two chemically distinct glycans for expression represents a new engineering solution for regulating transgene expression in *B. theta*. Other expression systems requiring two unique molecules for induction have been described, including the single chain tetracycline repressor in *E. coli* that requires both tetracycline and 4-dedimethylamino-anhydrotetracycline⁴⁴, or the classic *lac* operon that requires both lactose and cyclic adenosine monophosphate bound catabolite activator protein for full expression^{45,46}. The use of a dual-glycan expression system reported here, however, has the advantage of being regulated by common dietary glycans.

Conclusion

Gene expression systems with unique induction requirements for transgene expression have been created in *B. theta* by modifying the promoter elements of regulatory genes from two distinct PULs. These systems have been shown to increase expression of two different transgenes spanning a dynamic range of 2,000-fold³⁴ to 100,000-fold³¹, consistent with other studies. Most significantly, our discovery that engineering regulatory gene promoters to be activated by chemically distant glycans (i.e. DX and AG), resulted in a dual-glycan expression system that is dependent on both glycans to maximize gene expression. This property was reproduced using integration plasmids that target two different genomic loci, and from pEP, an in-house designed episomal expression system. The results from this study sets the stage for engineering more complex glycan-directed regulatory engineering and rationally designed bacterial tools that respond to dietary glycans.

Methods

Vector construction. The recipient strain, *B. theta*- Δ PUL75, was previously created by counter selectable homologous recombination³⁷. PUL75 (genome base pairs 5371714...5408988) was targeted for chromosomal deletion inserting complementary flank sequences inserted into the pExchange backbone. To generate the complementary pINTEGRATE vectors (pINT_DX) a 750 bp region upstream of PUL75 (PUL75 5') and the promoter and initiator codon of *bt3090* were amplified from *B. theta* genomic DNA and assembled by overlap PCR to contain a BglII cut site between the two fragments (to facilitate future promoter variants of pINT) with a PstI cut site on the 5' end of the insert and an NcoI site on the 3' end. Concurrently, the terminator from pNBU2³¹ was amplified with a 5' forward primer containing an MCS with 6xHis tag and a 3' reverse XbaI primer. After digest with PstI/NcoI and NcoI/XbaI respectively, these inserts were ligated into pExchange-tdk (pEX-tdk)³¹ that were restricted with PstI and XbaI, and transformed into chemically competent S17-1 λ pir *E. coli* cells. The amplified construct was then digested with XbaI and ligated to a 750 bp PCR fragment corresponding to the downstream region of PUL75 (PUL75 3') with an inserted XbaI site on the 5' end and a NheI (XbaI isocaudomer) site on the 3' end. The construct was confirmed by restriction digest and Sanger sequencing. pINT_AG and pINT_ON were created by replacing the BglII to NcoI fragment in pINT_DX with the promoter and initiator codon from *bt0268* genomic DNA and *bt1311* and the *rpil** RBS from pAT593³³, respectively.

The single-step integrative vector pNBU2_DX was generated by amplifying the Promoter-MCS fragment from pINT_DX and ligating it into a variant of pNBU2 amplified to create BglII and XbaI cut sites between the R6K origin of replication and the terminator of pNBU2³³. The promoter and MCS from pINT_AG and pINT_ON

were amplified with BglII and XbaI, respectively, and cloned into the reciprocal sites of pNBU2_DX to create pNBU2_AG and pNBU2_ON.

To create the episomal expression constructs pEP_DX, pEP_AG, and pEP_ON, the promoter to terminator fragments from pINT_DX, pINT_AG, and pINT_ON were amplified with NotI and PstI restriction enzyme sites, respectively, and cloned into a variant of pEx-tdk containing the *repA* and *mobA* genes (Genbank AAB39963.1 and AAB39964.1, respectively), which allow the plasmid to exist in *B. theta* extra-chromosomally.

The reporter genes NanoLuc⁴⁷ and BuGH16 (*np1_842*) were fused to respective promoters by overlapping PCR and cloned into the BglII/XhoI or NotI/XhoI sites of the respective vectors to create reporter constructs. In order to target BuGH16 to the cell surface and to avoid interference with the translocation of endogenous proteins, the native N-terminal signal peptide of BuGH16 (nucleotides 1–67) was replaced with that of the putative polysaccharide lyase family 6 (PL6) (*bt4116*; bases 1–63) from PUL75, predicted to be expressed on the surface of *B. theta*⁴⁸.

Regulatory gene promoter engineering. pEx-tdk was used to insert one of three promoters in front of the start codon of *bt3091* (*SusR^{DX}*) or *bt0267* (*HTCS^{AG}*). The three promoters were P_{bt1311} (promoter from pAT593³³), P_{bt3090} (the contiguous sequence between the stop codon of *bt3091* and the start codon *bt3090*), or P_{bt0268} (promoter from pMM660³³) and were denoted as P_{ON}, P_{DX}, and P_{AG} respectively. A terminator sequence was appended to the 5' end of each promoter to reduce context dependence. The insertions were targeted by cloning the 750 bp flanks on either side of the desired insertion site into the pEx-tdk vector.

Bacterial conjugations. Donor cultures of *E. coli* strain S17-1λpir were grown in 5 mL lysogeny broth (LB)⁴⁹ with 100 µg/mL ampicillin. Recipient strains of *B. theta* were routinely grown in 5 mL Tryptone Yeast Extract Glucose (TYG)⁵⁰ at 37 °C in an anaerobic atmosphere (85% N₂, 10% CO₂, 5% H₂). Donor and recipient cultures were pelleted by centrifugation and resuspended together in 1 mL of TYG and plated on Supplemented Brain Heart Infusion (BHIS)⁵¹ agar. To allow cell mating to occur, cultures were grown agar side down at 37 °C overnight under aerobic conditions. During cell mating, the vector is transferred from the donor *E. coli* and integrated into the chromosome of recipient *B. theta* through homologous recombination.

After 16–24 h, the resulting biomass was scraped from the plate, suspended in TYG broth and serially diluted. Cell suspensions were plated on BHI agar with 200 µg/mL gentamycin to select against *E. coli*, and 25 µg/mL erythromycin to select against *B. theta* that had not received vector. Pure cultures were prepared from arbitrarily-selected resistant colonies.

pNBU2-based conjugations were screened by colony PCR to confirm vector integration. For the double crossover conjugations using pEx-tdk and pINT, eight resistant colonies were arbitrarily chosen and grown in TYG broth. At this stage, a second recombination event may occur excising the vector backbone and selection markers from the chromosome and retaining the sequence of interest. Each cell culture was plated on BHI agar with 200 µg/mL 5-fluorodeoxyuridine (FUDR) to select against cells that retained the plasmid in their chromosome. Pure cultures were screened by colony PCR. Positive clones underwent genome extraction for sequence confirmation.

Growth curves. Strains of *B. theta* were inoculated from a glycerol stock into 5 mL TYG medium and grown for 16–24 h in an anaerobic atmosphere. TYG cultures were diluted 1 in 50 into prewarmed anaerobic 2 × *Bacteroides* Minimal Medium (MM)⁵². 100 µL of diluted cells were plated in a transparent 96 well plate containing 100 µL of prewarmed anaerobic 1% (w/v) carbohydrate solution to achieve final concentrations of 1X MM and 0.5% (w/v) sugar. Carbohydrates examined were 0.5% (w/v) glucose (GLC), 0.5% homogalacturonan (HG), 0.5% (w/v) dextran (DX), 0.5% (w/v) arabinogalactan (AG), or a mixture of 0.25% (w/v) DX combined with 0.25% (w/v) AG (mix). Plates were sealed with clear, gas permeable membranes, and turbidity was measured in a BioTek Synergy HT plate reader programmed to read absorbance at 600 nm every 10 m for 48 h. Technical triplicates, each being the average of three observations, were tested for each condition. Negative controls consisted of wells containing 100 µL medium and 100 µL 18 MΩcm⁻¹ ultrapure H₂O. The average of the triplicates normalized for media background was plotted with the standard error of the mean.

NanoLuc assays. Dense TYG pre-cultures were created as described above for growth curves. TYG cultures were diluted 100-fold into prewarmed MM with 0.5% carbohydrate solution in borosilicate glass tubes. After 24 h incubation at 37 °C under anaerobic conditions, cells were pelleted by centrifugation and the supernatant was discarded. Cells were lysed by resuspending in 1/10th culture volume of BugBuster® (Millipore Sigma) and incubating at room temperature for 10 min. Phosphate buffered saline (PBS) was used to dilute lysate 20 to 200-fold to ensure the signal was within the dynamic range of the luminescence detector. 30 µL samples were added to an opaque 96 well microplate and combined with 30 µL of freshly prepared NanoGlo reagent (Promega). Luminescence was measured using a Synergy HT Multi-detection plate reader. Technical replicates, each being the average of three observations, were tested for each condition. The luminescence of the negative controls (30 µL PBS + 30 µL NanoGlo) was used to normal signal for each replicate. The detected value corrected for the dilution factor and normalized to the density of the original culture as shown in Eq. (1).

$$\frac{((\text{luminescent reading} - \text{negative control}) \times \text{dilution factor})}{\text{OD}_{600} \text{ of culture}} \quad (1)$$

The technical triplicates were then averaged and plotted with the standard error of the mean.

BuGH16 Agarolysis assays. The *B. theta*-*SusR^{DX}*-*bugh16* strain was culture as described above for the NanoLuc assays; however, cells were not lysed. Whole cells were washed and resuspended in 2X MM before

being incubated with an equal volume of 0.8% (w/v) agarose under aerobic conditions at 37 °C for 24 h. The liquid phase was clarified by centrifugation before heat killing at 100 °C for 10 min. Nine μL of the resulting solution was spotted onto a silica thin layer chromatography (TLC) plate. The TLC plate was run in a 2:1:1 (v/v) solution of butanol: acetic acid: 18 M Ωcm^{-1} ultrapure H_2O . After drying, the plate was stained with one part 0.2% (w/v) ethanol) dihydroxynaphthalene to two parts 3.75:1 ethanol: sulfuric acid solution, and developed by heating at 100 °C for 5 min.

Western blotting. *B.theta*-SusR^{DX}-*bugh16* cells were lysed in 1:45 of the culture volume BugBuster (Millipore). Cell lysate was mixed 1:1 with SDS loading buffer (10 μL of 2 M glycerol, 1.28 M β -mercaptoethanol, 125 mM Tris pH 6.8, 140 mM SDS, 60 μM bromophenol blue) and denatured at 95 °C for 5 min. Samples were run on 15% sodium dodecyl sulfate polyacrylamide gels at 200 V for 1 h in a Novex Mini Cell (Invitrogen). Protein was transferred from gels to a Sequi-Blot polyvinylidene fluoride membrane (BioRad) in an X Cell II Blot Module (Invitrogen) at 30 V for 16 h. Membranes were incubated with blocking buffer (5% (w/v) skim milk powder in a solution of 50 mM tris, 150 mM NaCl and 0.1% (v/v) tween 20 (TBS-T)) for 2 h followed by incubation with a 1:2500 dilution of rabbit anti-6-his IgG conjugated to horse radish peroxidase (Bethyl) in blocking buffer for 6 h. Membranes were washed three times by rinsing in 18 M Ωcm^{-1} ultrapure H_2O and rocking in TBS-T for 5 min. An Opti-4CN (BioRad) substrate kit was used to develop the membranes.

Statistical analysis. To examine whether or not changes in NanoLuc activity between strains and conditions were significant, data were analyzed using GraphPad Prism 7.00. T-tests were used to make specific comparisons between conditions. Each condition had three biological replicates and tests were performed without assuming a consistent standard deviation. *P* values of 0.05, 0.01, and 0.001 were used to define levels of significance.

Data availability

Activity assay data is available upon request.

Received: 3 April 2019; Accepted: 24 September 2019;

Published online: 22 November 2019

References

1. Tropini, C., Earle, K. A., Huang, K. C. & Sonnenburg, J. L. The Gut Microbiome: Connecting Spatial Organization to Function. *Cell Host Microbe* **21**, 433–442, <https://doi.org/10.1016/j.chom.2017.03.010> (2017).
2. Hooper, L. V., Midtvedt, T. & Gordon, J. I. How host-microbial interactions shape the nutrient environment of the mammalian intestine. *Annu. Rev. Nutr.* **22**, 283–307, <https://doi.org/10.1146/annurev.nutr.22.011602.092259> (2002).
3. Larsen, O. F. A. & Claassen, E. The mechanistic link between health and gut microbiota diversity. *Sci Rep* **8**, 2183, <https://doi.org/10.1038/s41598-018-20141-6> (2018).
4. Xu, J. & Gordon, J. I. Honor thy symbionts. *Proceedings of the National Academy of Sciences* **100**, 10452–10459, <https://doi.org/10.1073/pnas.1734063100> (2003).
5. Callaway, T. R. *et al.* Probiotics, prebiotics and competitive exclusion for prophylaxis against bacterial disease. *Animal health research reviews* **9**, 217–225, <https://doi.org/10.1017/s1466252308001540> (2008).
6. Frank, D. N., Zhu, W., Sartor, R. B. & Li, E. Investigating the biological and clinical significance of human dysbioses. *Trends in microbiology* **19**, 427–434, <https://doi.org/10.1016/j.tim.2011.06.005> (2011).
7. Brunkwall, L. & Orho-Melander, M. The gut microbiome as a target for prevention and treatment of hyperglycaemia in type 2 diabetes: from current human evidence to future possibilities. *Diabetologia* **60**, 943–951, <https://doi.org/10.1007/s00125-017-4278-3> (2017).
8. Davis, C. D. The Gut Microbiome and Its Role in Obesity. *Nutrition today* **51**, 167–174, <https://doi.org/10.1097/NT.000000000000167> (2016).
9. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59 (2010).
10. Sonnenburg, E. D. Diet-induced extinctions in the gut microbiota compound over generations. *Nature* **529**, 212–215, <https://doi.org/10.1038/nature16504> (2016).
11. Bartlett, J. G. Clinical practice. *Antibiotic-associated diarrhea*. *N Engl J Med* **346**, 334–339, <https://doi.org/10.1056/NEJMc011603> (2002).
12. Human Microbiome Project, C. *et al.* Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214, <https://doi.org/10.1038/nature11234> (2012).
13. Wexler, A. G. & Goodman, A. L. An insider's perspective: *Bacteroides* as a window into the microbiome. *Nature microbiology* **2**, 17026, <https://doi.org/10.1038/nmicrobiol.2017.26> (2017).
14. Hehemann, J.-H., Boraston, A. B. & Czjzek, M. A sweet new wave: structures and mechanisms of enzymes that digest polysaccharides from marine algae. *Current Opinion in Structural Biology* **28**, 77–86 (2014).
15. Salazar, N., Gueimonde, M., Hernandez-Barranco, A. M., Ruas-Madiedo, P. & de los Reyes-Gavilan, C. G. Exopolysaccharides produced by intestinal *Bifidobacterium* strains act as fermentable substrates for human intestinal bacteria. *Applied and environmental microbiology* **74**, 4737–4745, <https://doi.org/10.1128/aem.00325-08> (2008).
16. Abbott, D. W., Martens, E. C., Gilbert, H. J., Cuskin, F. & Lowe, E. C. Coevolution of yeast mannan digestion: Convergence of the civilized human diet, distal gut microbiome, and host immunity. *Gut microbes* **6**, 334–339, <https://doi.org/10.1080/19490976.2015.1091913> (2015).
17. Bjursell, M. K., Martens, E. C. & Gordon, J. I. Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, *Bacteroides thetaiotaomicron*, to the suckling period. *The Journal of biological chemistry* **281**, 36269–36279, <https://doi.org/10.1074/jbc.M606509200> (2006).
18. Grondin, J. M., Tamura, K., Dejean, G., Abbott, D. W. & Brumer, H. Polysaccharide Utilization Loci: Fueling Microbial Communities. *J Bacteriol* **199**, <https://doi.org/10.1128/JB.00860-16> (2017).
19. Ravcheev, D. A., Godzik, A., Osterman, A. L. & Rodionov, D. A. Polysaccharides utilization in human gut bacterium *Bacteroides thetaiotaomicron*: comparative genomics reconstruction of metabolic and regulatory networks. *BMC Genomics* **14**, 873, <https://doi.org/10.1186/1471-2164-14-873> (2013).
20. Martens, E. C. *et al.* Recognition and Degradation of Plant Cell Wall Polysaccharides by Two Human Gut Symbionts. *PLoS Biol* **9**, e1001221, <https://doi.org/10.1371/journal.pbio.1001221> (2011).
21. Wallace, R. J. Ruminant microbiology, biotechnology, and ruminant nutrition: progress and problems. *Journal of Animal Science* **72**, 2992–3003, <https://doi.org/10.2527/1994.72112992x> (1994).

22. Choct, M. & Annonson, G. Anti-nutritive effect of wheat pentosans in broiler chickens: roles of viscosity and gut microflora. *Br Poult Sci* **33**, 821–834, <https://doi.org/10.1080/00071669208417524> (1992).
23. Hughes, R. J. Relationship between digesta transit time and apparent metabolisable energy value of wheat in chickens. *Br Poult Sci* **49**, 716–720, <https://doi.org/10.1080/00071660802449145> (2008).
24. Joglekar, P. *et al.* Genetic Variation of the SusC/SusD Homologs from a Polysaccharide Utilization Locus Underlies Divergent Fructan Specificities and Functional Adaptation in *Bacteroides thetaiotaomicron* Strains. *mSphere* **3**, <https://doi.org/10.1128/mSphereDirect.00185-18> (2018).
25. Shepherd, E. S., DeLoache, W. C., Pruss, K. M., Whitaker, W. R. & Sonnenburg, J. L. An exclusive metabolic niche enables strain engraftment in the gut microbiota. *Nature* **557**, 434–438, <https://doi.org/10.1038/s41586-018-0092-4> (2018).
26. Bayley, D. P., Rocha, E. R. & Smith, C. J. Analysis of cepA and other *Bacteroides fragilis* genes reveals a unique promoter structure. *FEMS Microbiol Lett* **193**, 149–154, <https://doi.org/10.1111/j.1574-6968.2000.tb09417.x> (2000).
27. Vingadassalom, D. *et al.* An unusual primary sigma factor in the Bacteroidetes phylum. *Mol Microbiol* **56**, 888–902, <https://doi.org/10.1111/j.1365-2958.2005.04590.x> (2005).
28. Mastropaolo, M. D., Thorson, M. L. & Stevens, A. M. Comparison of *Bacteroides thetaiotaomicron* and *Escherichia coli* 16S rRNA gene expression signals. *Microbiology* **155**, 2683–2693, <https://doi.org/10.1099/mic.0.027748-0> (2009).
29. Smith, C. J., Rogers, M. B. & McKee, M. L. Heterologous gene expression in *Bacteroides fragilis*. *Plasmid* **27**, 141–154 (1992).
30. Koropatkin, N. M., Martens, E. C., Gordon, J. I. & Smith, T. J. Starch Catabolism by a Prominent Human Gut Symbiont Is Directed by the Recognition of Amylose Helices. *Structure* **16**, 1105–1115 (2008).
31. Lim, B., Zimmermann, M., Barry, N. A. & Goodman, A. L. Engineered Regulatory Systems Modulate Gene Expression of Human Commensals in the Gut. *Cell* **169**, 547–558.e515, <https://doi.org/10.1016/j.cell.2017.03.045> (2017).
32. Wang, J., Shoemaker, N. B., Wang, G. R. & Salyers, A. A. Characterization of a *Bacteroides mobilizable* transposon, NBU2, which carries a functional lincomycin resistance gene. *Journal of Bacteriology* **182**, 3559–3571, <https://doi.org/10.1128/jb.182.12.3559-3571.2000> (2000).
33. Mimeo, M., Tucker, A. C., Voigt, C. A. & Lu, T. K. Programming a Human Commensal Bacterium, *Bacteroides thetaiotaomicron*, to Sense and Respond to Stimuli in the Murine Gut Microbiota. *Cell Systems* **1**, 62–71, <https://doi.org/10.1016/j.cels.2015.06.001> (2015).
34. Horn, N. *et al.* A Novel Tightly Regulated Gene Expression System for the Human Intestinal Symbiont *Bacteroides thetaiotaomicron*. *Frontiers in Microbiology* **7**, 1080, <https://doi.org/10.3389/fmicb.2016.01080> (2016).
35. Rogers, T. E. *et al.* Dynamic responses of *Bacteroides thetaiotaomicron* during growth on glycan mixtures. *Molecular Microbiology* **88**, 876–890, <https://doi.org/10.1111/mmi.12228> (2013).
36. Schwalm, N. D. III, Townsend, G. E. II & Groisman, E. A. Prioritization of polysaccharide utilization and control of regulator activation in *Bacteroides thetaiotaomicron*. *Mol Microbiol* **104**, 32–45, <https://doi.org/10.1111/mmi.13609> (2017).
37. Luis, A. S. *et al.* Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic *Bacteroides*. *Nature microbiology* **3**, 210–219, <https://doi.org/10.1038/s41564-017-0079-1> (2018).
38. Cho, K. H., Cho, D., Wang, G. R. & Salyers, A. A. New regulatory gene that contributes to control of *Bacteroides thetaiotaomicron* starch utilization genes. *J Bacteriol* **183**, 7198–7205, <https://doi.org/10.1128/JB.183.24.7198-7205.2001> (2001).
39. Lowe, E. C., Basle, A., Czjzek, M., Firbank, S. J. & Bolam, D. N. A scissor blade-like closing mechanism implicated in transmembrane signaling in a *Bacteroides* hybrid two-component system. *Proc Natl Acad Sci USA* **109**, 7298–7303, <https://doi.org/10.1073/pnas.1200479109> (2012).
40. Townsend, G. E. 2nd, Raghavan, V., Zwir, I. & Groisman, E. A. Intramolecular arrangement of sensor and regulator overcomes relaxed specificity in hybrid two-component systems. *Proc Natl Acad Sci USA* **110**, E161–169, <https://doi.org/10.1073/pnas.1212102110> (2013).
41. Smith, C. J., Rollins, L. A. & Parker, A. C. Nucleotide sequence determination and genetic analysis of the *Bacteroides* plasmid, pBI143. *Plasmid* **34**, 211–222, <https://doi.org/10.1006/plas.1995.0007> (1995).
42. Pluvinaige, B. *et al.* Molecular basis of an agarose metabolic pathway acquired by a human intestinal symbiont. *Nature communications* **9**, 1043, <https://doi.org/10.1038/s41467-018-03366-x> (2018).
43. D'Elia, J. N. & Salyers, A. A. Effect of regulatory protein levels on utilization of starch by *Bacteroides thetaiotaomicron*. *J Bacteriol* **178**, 7180–7186, <https://doi.org/10.1128/jb.178.24.7180-7186.1996> (1996).
44. Kamionka, A. *et al.* Induction of single chain tetracycline repressor requires the binding of two inducers. *Nucleic Acids Research* **34**, 3834–3841, <https://doi.org/10.1093/nar/gkl316> (2006).
45. Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* **3**, 318–356 (1961).
46. Malan, T. P., Kolb, A., Buc, H. & McClure, W. R. Mechanism of CRP-cAMP activation of lac operon transcription initiation activation of the P1 promoter. *Journal of Molecular Biology* **180**, 881–909 (1984).
47. Hall, M. P. *et al.* Engineered Luciferase Reporter from a Deep Sea Shrimp Utilizing a Novel Imidazopyrazinone Substrate. *ACS Chemical Biology* **7**, 1848–1857, <https://doi.org/10.1021/cb3002478> (2012).
48. Nielsen, H. Predicting Secretory Proteins with SignalP. *Methods Mol Biol* **1611**, 59–73, https://doi.org/10.1007/978-1-4939-7015-5_6 (2017).
49. Bertani, G. Studies On Lysogenesis I.: The Mode of Phage Liberation by Lysogenic *Escherichia coli*. *Journal of Bacteriology* **62**, 293–300 (1951).
50. Yokoyama, M. T., Carlson, J. R. & Holdeman, L. V. Isolation and characteristics of a skatole-producing *Lactobacillus* sp. from the bovine rumen. *Applied and environmental microbiology* **34**, 837–842 (1977).
51. Bacic, M. K. & Smith, C. J. Laboratory Maintenance and Cultivation of *Bacteroides* Species. *Current protocols in microbiology* **0**(13), 2–24, <https://doi.org/10.1002/9780471729259.mc13c01s9> (2008).
52. Martens, E. C., Chiang, H. C. & Gordon, J. I. Mucosal Glycan Foraging Enhances Fitness and Transmission of a Saccharolytic Human Gut Bacterial Symbiont. *Cell Host & Microbe* **4**, 447–457, <https://doi.org/10.1016/j.chom.2008.09.007> (2008).

Acknowledgements

We wish to thank Eric Martens (University of Michigan) for the kind gift of pExchange integration vector, and Jeffery Smith (East Carolina University) for the gift of pBI143, containing the *repA* and *mobA* genes. This work is supported by Alberta Innovates (Project ID: BIOFS-12-026) and Agriculture and Agri-Food Canada (Project ID: #1679).

Author contributions

D.W.A. and G.D.I. secured the funding. D.R.J. designed and engineered expression vectors, interpreted results, generated figures and assisted to write the paper. M.B.S. engineered regulatory elements, performed growth and activity assays, interpreted results, generated figures and assisted to write the paper. D.R.J. and M.B.S. contributed equally to this study. R.M., J.M.G. and C.R.A. helped design constructs and engineer strains. G.D.I. and B.S. contributed to the design of the study and reporter assays, interpretation of results, and writing of the manuscript. D.W.A. conceived of the study, direct research, helped generate figures, and wrote paper. All authors contributed to the editing of final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-53726-w>.

Correspondence and requests for materials should be addressed to D.W.A.

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

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019