Expanding the *Bacteroides* synthetic biology toolkit to develop an *in vivo* intestinal malabsorption biosensor.

Giselle McCallum[1*], Juan C. Burckhardt[1*], Jerry He[1], Alice Hong[2], Laurent Potvin-Trottier[3,4,5,§], Carolina Tropini[1,2,6,♦]

[1]Department of Microbiology & Immunology, University of British Columbia, Vancouver, British Columbia, Canada
[2]School of Biomedical Engineering, University of British Columbia, Vancouver, British Columbia, Canada
[3]Department of Biology, Concordia University, Montréal, QC H4B 1R6, Canada
[4]Department of Physics, Concordia University, Montréal, QC H4B 1R6, Canada
[5]Center for Applied Synthetic Biology, Concordia University, Montréal, QC H4B 1R6, Canada
[6]Humans and the Microbiome Program, Canadian Institute for Advanced Research (CIFAR), Toronto, Ontario, Canada

[*] Authors contributed equally
[§]Current address: Biospective Inc, Montréal, QC H3B 2T9, Canada
♦ Corresponding Author: Carolina Tropini
Email: carolina.tropini@ubc.ca

Abstract

The human gut is a highly dynamic physical environment where perturbations—including factors such as acidification, oxygenation, and particle concentration (osmolality)—can influence microbiota composition and contribute to disease states. Understanding gut environmental changes is essential for advancing diagnostic and therapeutic strategies for gut health. However, non-invasive methods for continuous monitoring remain limited. The bacterial gut microbiota represents a powerful platform for continuous, non-invasive biosensing technologies for the gut environment, with genetically tractable commensal species like *Bacteroides thetaiotaomicron* (*B. theta*) emerging as promising hosts for engineered biosensors. However, the availability of genetic tools for precise, modular environmental sensing and reporting control in *B. theta* remains limited. Here, we present an expanded genetic engineering toolkit for *B. theta* that enables precise, fluorescence-based environmental sensing of the gut environment. This toolkit includes (i) three libraries of orthogonally inducible promoters capable of driving fluorescence expression, (ii) a DNA-based system to tune repressor activity in *B. theta*, (iii) a resulting modular transcriptional reporter circuit that integrates native promoter activation with fluorescent outputs, and (iv) characterization of a novel plasmid integration mode in *B. theta*. To demonstrate its utility, we engineered biosensors for gut malabsorption, a condition characterized by increased luminal osmolality. Using identified osmolality-responsive native promoters from *B. theta*, we made biosensors capable of detecting changes in gut physiology through graded fluorescent outputs. These biosensors were validated both *in vitro* and *in vivo* using a murine model of laxative-induced malabsorption, where they enabled near real-time, non-invasive monitoring of single-cell response from fecal samples with sensitivity to subclinical malabsorption levels. By expanding the genetic toolkit for *B. theta* and demonstrating its use in a physiologically relevant context, this approach highlights the potential of engineered gut bacteria as a monitoring platform for diverse gut health applications. This work advances strategies for microbial biosensing and positions gut commensals as key players in next-generation diagnostic methods.

CRediT authorship contribution statement:
GM, JCB: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization
JH: Conceptualization, Methodology, Investigation, Formal analysis

AH: Investigation
LPT: Conceptualization, Methodology, Resources, Writing - Review & Editing, Supervision, Project administration, Funding acquisition
CT: Conceptualization, Methodology, Resources, Writing - Review & Editing, Supervision, Project administration, Funding acquisition

Introduction

The human gut is a dynamic and complex ecological niche characterized by diverse physical environments[1]. These heterogeneous environments are the result of a combination of gradients in physical factors, such as pH, oxygen, and osmolality, that vary significantly throughout the gastrointestinal (GI) tract[2]. Aberrations in the physical environmental factors in the gut can contribute to, or result from, gastrointestinal diseases[2]. Understanding how changing physical parameters relate to disease states is critical for advancing diagnostic and therapeutic approaches to gut health. Currently, continuous, sensitive, and non-invasive monitoring of physical conditions in the gut is not possible due to technological limitations. Existing non-invasive methods, such as stool sample analyses (e.g., Bristol Stool Scoring, fecal calprotectin), are bulk measurements which are not sensitive to subclinical symptoms. Conversely, intestinal biopsies can be highly sensitive to local intestinal damage, but are limited to single time-point measurements and require invasive procedures such as endoscopies and colonoscopies with preparatory steps (i.e., bowel preparation) that alter the interrogated gut environment[1]. These challenges highlight the need for new technologies that can sensitively and continuously assess the gut environment non-invasively.

The bacterial microbiota offers a promising platform for detecting and reporting on the gut environment. Gut bacteria have evolved genetic mechanisms to detect and adjust to the dynamic changes encountered in their intestinal habitat. The transcriptional mechanisms employed by these microbes in response to environmental factors are promising options for developing microbiota-based strategies to monitor changes in gut conditions. Thanks to their small size and high sensitivity, bacteria can detect and respond to changes in environmental conditions along micron-scale gradients that are not detectable by most solid-state diagnostic tools. By engineering bacteria to produce reporter proteins (e.g., fluorescence, luminescence, or color signals) when physiological responses to an environmental cue are activated, biological sensors (biosensors) can be developed to report on the state of the gut environment non-invasively and potentially prior to the development of clinical symptoms.

In the context of the gut, *Bacteroides thetaiotaomicron* (*B. theta*) is a promising chassis for gut biosensors and engineered probiotics. Importantly, unlike *Escherichia coli*, *B. theta* is capable of long-term colonization at high abundance throughout most of the gut, including in intestinal crypts[3–5]. The development of *B. theta* synthetic biology tools has enabled precise gene expression tuning[6–9], genetic memory systems[7], user-controlled colonization and containment strategies[4,10], and constitutive fluorescent protein expression[3], advancing its engineering potential and application as a live probiotic in humans[10]. However, a critical challenge remains: unlike *Escherichia coli*, native *B. theta* promoters cannot drive detectable fluorescent protein expression[3,7,8], limiting the possibility of single-cell interrogation of transcriptional activity *in situ* in this bacterium. Previous studies have relied on luminescence as a reporter of gene expression in *B. theta*[7–9]. However, luminescence is unable to be resolved at the single-cell level and is therefore limited to measuring population-average response. Luminescence reporters also cannot be multiplexed to measure multiple signals and are therefore more challenging to normalize. In contrast, single-cell analysis of transcriptional activity with fluorescence-based methods like flow cytometry or imaging can allow the identification of diverse responses within a population in the gut[11]. Using flow cytometry allows rapid screening of high numbers of cells and could therefore also enable detection of biosensor response from a population with limited colonization levels. Fluorescent tools for gut commensals like *B. theta* could therefore address the need for single-cell measurements of transcriptional activity in the gut.

Here, we addressed the need for sensitive, single-cell, fluorescent-based *B. theta* tools by developing a genetic toolkit for *B. theta* and using it to create a novel fluorescence-based transcriptional reporter system. Specifically, we first engineered and characterized three libraries of repressible promoters capable of driving inducible fluorescent protein expression in *B. theta*, starting from a high-expression constitutive *Bacteroides*-targeting phage promoter[3]. We then identified native promoters upregulated under a variety of common stress conditions encountered in the gut with transcriptomics. Finally, we combined strong inducible promoters from our library with native inducible *B. theta* promoters, a TetR-based inverter module, and a DNA-based system to tune repressor activity[12] to construct a modular biosensor circuit that can translate the induction of a native promoter to changes in cellular fluorescence levels.

To demonstrate the utility of our inverted transcriptional reporter, we constructed and implemented biosensors to detect laxative-driven malabsorption that results in increased intestinal osmolality, using identified polyethylene glycol- and osmolality-responsive native *B. theta* promoters. In the gut, malabsorption is a clinical manifestation of common gastrointestinal pathologies, including inflammatory bowel disease (IBD)[13], celiac disease[14,15], small intestinal bacterial overgrowth (SIBO)[15], food intolerances[15–17], and enteric infections[15,18,19]. Our biosensors responded to both increasing osmolality levels *in vitro* and *in vivo* to laxative-induced malabsorption conditions in mouse models, as measured from single-cell fluorescence levels from fecal and cecal samples. Importantly, this technology enabled continuous detection of malabsorption severity from solid feces, even before overt clinical symptoms manifest. Our modular *B. theta* biosensors provide a platform for long-term, non-invasive, and continuous reporting of environmental conditions in the gut while significantly expanding the genetic engineering toolkit for this emerging model gut bacterium.

Results

Design rationale of a modular, inverted transcriptional reporter for B. theta

The inability of native *B. theta* promoters to drive fluorescent proteins limits the output of existing transcriptional reporters to luciferase or other enzymes[7]. As native promoters are capable of driving repressor expression to a functional level, we sought to build a circuit that links the output of a synthetic repressible promoter to the induction of a native promoter via a repressor-based inverter module (Figure 1A)[20,21]. Existing synthetic promoters have been engineered for *B. theta*, by adding repressor binding sites to constitutive promoters, but these promoters cannot drive fluorophore expression[6–8]. Therefore, to achieve a repressible fluorescence output, we engineered a *Bacteroides*-targeting phage promoter (P_{BfP1E6}) that drives strong constitutive GFP expression in *Bacteroides*[3]. We engineered this promoter to be modulated by three different repressor proteins. In this circuit design, native response to an environmental stimulus thus leads to a reciprocal increase in repressor expression, which decreases GFP expression levels in the cell (Figure 1A). To account for variations in overall cellular gene expression levels across growth phases, this circuit includes a constitutively expressed mScarlet-I3 red fluorescent protein[22] (RFP) as a reference fluorophore. Native promoters can then be inserted modularly into the circuit using Golden Gate assembly[23] to allow for easy construction of multiple biosensor circuits.

Importantly, an inverted transcriptional reporter offers several advantages to direct reporters for biosensor applications. As the reporter is constitutively expressed in the absence of a stimulus, the strong baseline signal ensures that even small changes in repression can be detected, improving sensitivity. Additionally, in the case that native promoter induction is weak, direct reporters can suffer from limited signal amplification resulting in non-detectable reporter expression. In contrast, inverted systems can use repressor

proteins to translate native gene expression to amplified, non-linear decreases in the expression of a reporter, allowing larger detectable dynamic range and clearer distinctions between active and inactive states.

A library of small-molecule inducible promoters facilitates tuneable GFP expression levels in B. theta

As our biosensor design relies on controlling GFP expression indirectly through the expression of a cognate repressor protein (Figure 1A), we sought to engineer a promoter that: i) is tightly regulated by a repressor protein not naturally present in *B. theta*, and ii) has a large dynamic range of expression. To achieve this, we constructed three libraries of promoters regulated by TetR, LacI, or PhlF repressors (consisting of 144, 245, and 131 promoters, respectively) by inserting one or two copies of the corresponding operator sites throughout the P_{BfP1E6} promoter in locations that did not disrupt the UP elements or -7/10 consensus sequence[3] (Figure 1B). Promoters were inserted upstream of a superfolder GFP gene in a Golden Gate-adapted pNBU2 *Bacteroides* conjugation plasmid[3,24]. Constructs were conjugated into *B. theta* and integration into the genome at one of two known integration sites at the 3' end of the serine tRNA genes (attBT2-1 or attBT2-2)[24] was verified with PCR (Methods). To identify promoters with a large dynamic range and low baseline expression, we measured GFP expression from 192 individual colonies from each library by growing them in the presence or absence of the appropriate small molecule inducer (TetR: anhydrotetracycline [aTc][8], LacI: ß-D-1-thiogalactopyranoside [IPTG][7], or PhlF: 2,4-diacetylphloroglucinol [DAPG][25]). Promoter variants in all libraries displayed a wide range of baseline and maximal expression levels, indicating a differing effect of operator site position on repression efficiency and maximum promoter strength (Figure S1A). The parent P_{BfP1E6} promoter did not respond substantially to any inducer, while each library contained multiple promoter variants of interest which showed strong repression and a large dynamic range (i.e., low GFP expression at baseline, and strong induced GFP expression)(Figure S1A). Up to 19 unique functional variants from each library with a range of induction levels were further characterized (Figure 1C) and sequenced (Supplementary Table 1). Based on the two criteria defined above, the best promoter candidates from each library were P_{TetR56}, P_{LacI43}, and $P_{PhlF133}$, showed strong repression, with P_{TetR56}, P_{LacI43}, and $P_{PhlF133}$ displaying 2.4%, 2.9%, and 2.1% of the parent P_{BfP1E6} promoter, respectively, which was only slightly higher than the autofluorescence of wild-type cells (1.6% of P_{BfP1E6}) (Figure 1C). These promoters also displayed a large dynamic range (29.4-, 20.8-, and 11.6-fold change expression, respectively) and strong maximal expression (63.5%, 61.2%, and 40.1% of the parent P_{BfP1E6} promoter, respectively). These expression levels are substantially stronger than strong native promoters, including P_{rpoD} and P_{rRNA}, which only showed up to ~20% of the P_{BfP1E6} promoter (Figure S1B). To further characterize highly repressed promoters from each library, we performed dose-response curves for promoter variants by growing them in a range of concentrations of their respective small molecule inducers. Dose-response curves from promoters in all libraries were fit well with a Hill function typical of small molecule inducible promoters ($r^2 > 0.9$; Supplementary Table 1) and displayed a range of values for maximum expression levels (*Vmax*), induction thresholds (*Km*), and Hill coefficients (*h*) (Figure 1D; Supplementary Table 1).

Given the robust performance of the P_{TetR56}, P_{LacI43}, and $P_{PhlF133}$ inducible promoters, we tested for orthogonality between the induction systems, as this is a critical feature for the construction of complex synthetic circuits[7]. We measured the degree of cross-induction between strong promoters in our libraries by growing several promoter variants from each library in aTc, IPTG, and DAPG, and compared GFP levels to uninduced controls. Promoters were only induced in their cognate inducer (Figure 1E), indicating that our induction systems are highly orthogonal and could be implemented in complex circuits with minimal crosstalk.

Our engineered promoter libraries expand the number of inducible promoters available for *B. theta* and offer a range of induction dynamics for the construction of complex synthetic circuits. To the best of our knowledge, these represent the first inducible promoters that can drive fluorophore expression in *B. theta*,

opening the possibility for synthetic circuits reported by fluorescence in this commensal bacterium. Given its high *Vmax* and tight baseline repression, we selected P_{TetR56} and TetR as the repressible promoter and cognate repressor protein to be implemented in our inverted *B. theta* biosensor circuit.

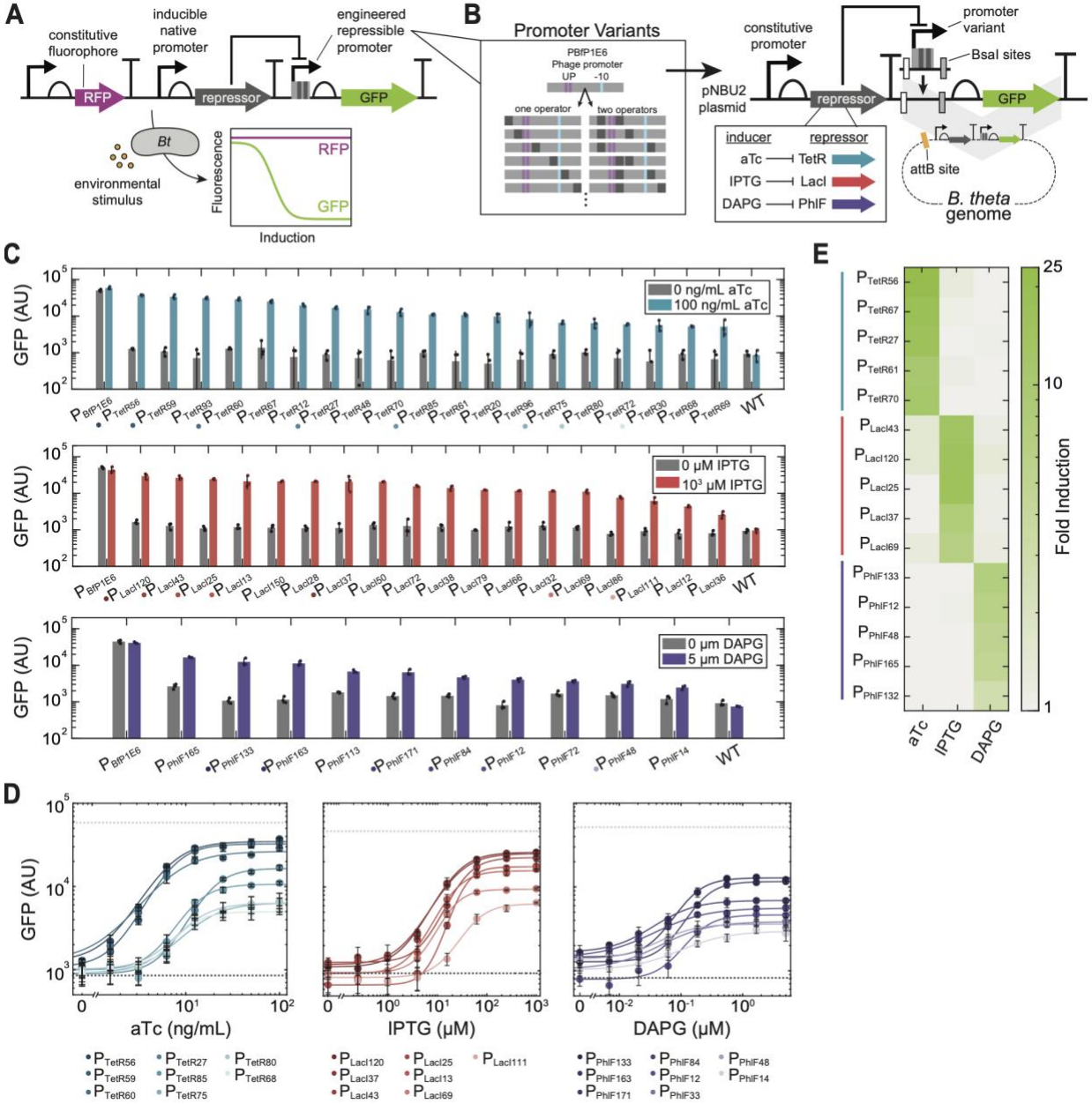
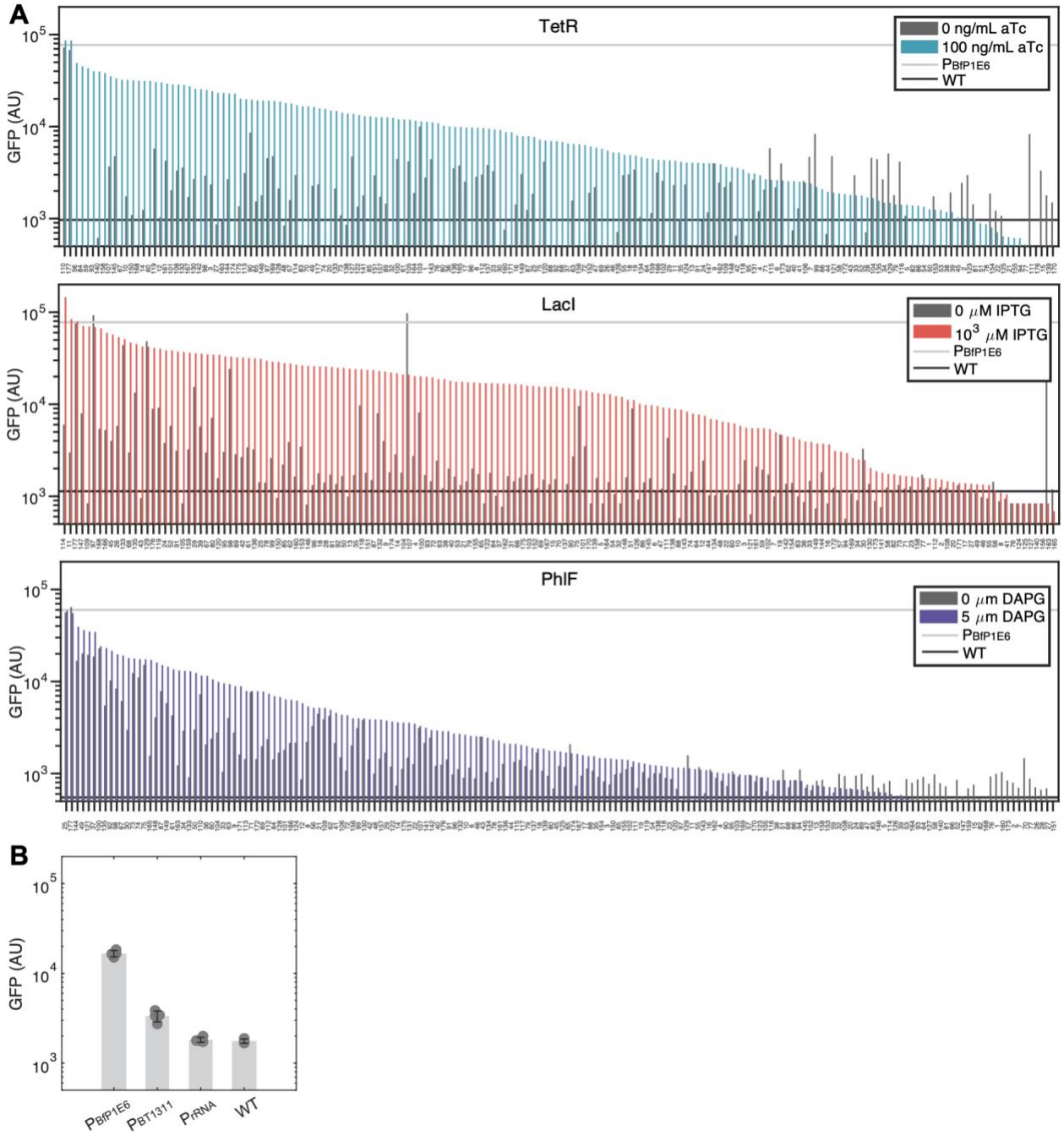


Figure 1. Engineered phage promoters enable repressible expression of GFP in *B. theta.* A) Design of an inverted, fluorescence-based transcriptional reporter circuit for *B. theta (Bt)*, in which GFP expression is repressed by the induction of a native promoter expressing a repressor protein. B) Design and construction of small-molecule inducible promoter libraries in *B. theta*. C) Activity of promoter variants measured by GFP expression in uninduced (base media) and induced conditions (100 ng/mL aTc, 10^3 μM IPTG, or 5 μM DAPG, respectively). D) Dose-response curves for 8 promoters from each library measured by GFP expression. Inducer concentrations were as follows: 2-fold serial dilutions starting at 100 ng/mL aTc; 4-fold serial dilutions starting at 10^3 μM IPTG; and 3-fold serial dilutions starting at 5 μM DAPG. Dark and

light dotted lines represent the WT and Phage promoter GFP levels under maximum induction levels, respectively. Fit lines represent a Hill function fit to dose-response curve data (See Supplementary Table 1 for fit parameters and goodness of fits). E) Orthogonality matrix of strong inducible promoters from each of the three libraries. Promoters were incubated in 100 ng/mL aTc, 10^3 μM IPTG, or 5 μM DAPG, and the fold-change GFP was calculated compared to uninduced conditions.

C,D: Error bars represent the standard deviation from n = 3 biological replicates (unique colonies) of each promoter.



Supplementary Figure 1. Engineered phage promoters enable repressible expression of GFP in *B. theta.* A) Activity of TetR, LacI, and PhlF promoter variants measured by GFP expression in uninduced (base media) and induced conditions (100 ng/mL aTc, 10^3 μM IPTG, or 5 μM DAPG, respectively). Light and dark horizontal lines represent the parent phage promoter and wild type *B. theta*, respectively, in media with an inducer present. B) GFP levels driven by strong native *B. theta* promoters compared to the parent P_{BfP1E6} phage promoter used in our promoter libraries.
B: Error bars represent the standard deviation from n = 3 biological replicates (unique colonies) of each promoter.

Gene expression analysis of B. theta *in physiological stress conditions identifies upregulated promoters that are perturbation-specific*

Having engineered a tunable promoter with a high dynamic range to form the backbone of an inducible reporter system, we next sought to identify native *B. theta* genes that are regulated uniquely and robustly in response to changing environmental conditions. Promoters from these genes could be used to drive TetR expression in our circuit, such that their induction leads to lower expression of GFP from the P_{TetR56} promoter. To this end, we performed transcriptomic profiling of *B. theta* grown *in vitro* under various stress conditions that occur in the gut (pH, oxidative, temperature, and osmotic stress). After identifying promoters that respond specifically to each of these stressors, we then focused our analysis on osmotic stress as a key disrupted factor in multiple intestinal diseases that could be detected with *B. theta* biosensors.

We first performed RNA-seq on *B. theta* cultures grown in complex base medium (~320 mOsm/kg, pH 7, 37 ºC), or in stress conditions (pH 5.8, 42 ºC, or 100 μM H_2O_2). For osmotic stress, we tested response to four osmolytes (sodium chloride [NaCl], sorbitol, lactulose, and polyethylene glycol [PEG]) at two osmolalities representative of intestinal levels during malabsorption (~700 and ~900 mOsm/kg)[26]. These compounds span common types of osmolytes that can be found in the gut from dietary or pharmaceutical sources (i.e., salts, sugar and sugar alcohols, and commercial laxatives, respectively). Overall, most conditions cluster away from the baseline samples in Principal Component Analysis of RNA-seq data (Supplementary Figure 2A), indicating a shift in the gene expression profile of *B. theta* in response to these perturbations. Within osmotic stress conditions, each osmolyte resulted in a unique transcriptome signature (Supplementary Figure 2A), supporting the use of multiple osmolytes to identify genes that are broadly responsive to osmolality.

To identify genes that were significantly upregulated in each stress condition, we performed differential expression analysis for each stress condition compared to base media (Supplementary Table 3). All conditions tested led to significantly increased expression of at least some genes (Figure 2A), and each stress condition showed a set of uniquely upregulated genes (Figure 2B). These uniquely responsive genes may be promising for use in biosensors targeted to detecting specific stress conditions in the gut. We also identified 247 genes that were upregulated in all the conditions tested, which could be of interest in the context of a broad stress sensor. We focused our search on osmolality-responsive genes to develop an intestinal malabsorption biosensor. Specifically, we looked for genes that were not induced under other stress conditions tested, and were activated proportionally to increased concentration of osmolytes. We identified 196 genes that were significantly upregulated in at least one osmolyte at ~900 mOsm/kg (around the highest level of malabsorption measured in the mouse gut)[26,27], but no other stress condition (Figure 2B). Using the reference *B. theta* genome (GCF_000011065.1) to annotate differentially regulated genes, we found the 10 most strongly upregulated genes in at least one osmotic stress condition (Figure 2C; Supplementary Table 2) to be broadly involved in biosynthetic pathways, signal regulation, and transmembrane protein transport. Many of these genes were also upregulated at ~700 mOsm/kg (representative of moderate malabsorption in the gut) and showed a graded induction in response to

increasing osmolality levels in more than one osmolyte, which is a desirable feature for building a broad biosensor for osmotic stress (Figure 2D). Our transcriptomics analyses provide a valuable map of transcriptome response to osmolality and other common stressors in the gut. Altogether, our analyses identified native *B. theta* promoters that respond strongly and specifically to a range of environmental perturbations, providing a useful resource to select gene candidates for the construction of biosensors for environmental parameters in the gut.

As a biosensor ultimately needs to respond in the gut environment, we next sought to identify additional genes that may be responsive to osmotic stress *in vivo*. PEG is a common laxative used to induce malabsorption and has previously been established as a malabsorption and osmotic diarrhea model in mice[26–29]. To characterize the *B. theta* transcriptional response to increased osmolality *in vivo*, we therefore mono-colonized germ-free mice with *B. theta* and treated them with PEG in drinking water for 6 days as previously shown (Figure 2E)[26]. We then measured intestinal osmolality as well as bacterial abundance and performed RNA-seq on the cecal contents of PEG-treated and control mice. PEG treatment resulted in an increased cecal osmolality to ~800 mOsm/kg compared to the 400 mOsm/kg of control mice (Supplementary Figure 2B) and a 10-fold decrease in bacterial abundance (Supplementary Figure 2C). Differential expression analysis from RNA-seq showed that PEG treatment significantly shifted the overall *B. theta* transcriptome (Supplementary Figure 2D; Supplementary Table 4). To find promoter candidates to use in an osmolality biosensor, we identified 7 genes from this dataset that i) were upregulated during PEG treatment *in vivo*, ii) were upregulated in at least one 900 mOsm/kg osmotic stress condition *in vitro*, and iii) were not upregulated in other stress conditions *in vitro*. (Figure 2F; Supplementary Table 2). These genes included outer membrane proteins, hypothetical proteins, a histidine kinase, glycoside hydrolases, and a polysaccharide lyase. Most of these genes showed 2- to 4-fold higher expression during PEG treatment (Figure 2G) with the exception of BT2364, a TonB-dependent transporter that showed a ~32-fold increase during PEG treatment (Figure 2G). Many of these genes were upregulated in PEG both *in vitro* and *in vivo* and some were significantly upregulated during growth in high osmolality induced by other osmolytes (Supplementary Figure 2E). The promoters identified *in vitro* and *in vivo* displayed specific activation in response to osmotic stress and graded induction across a range of osmolality, making them promising candidates for the construction of biosensors that respond precisely to changes in intestinal osmolality.

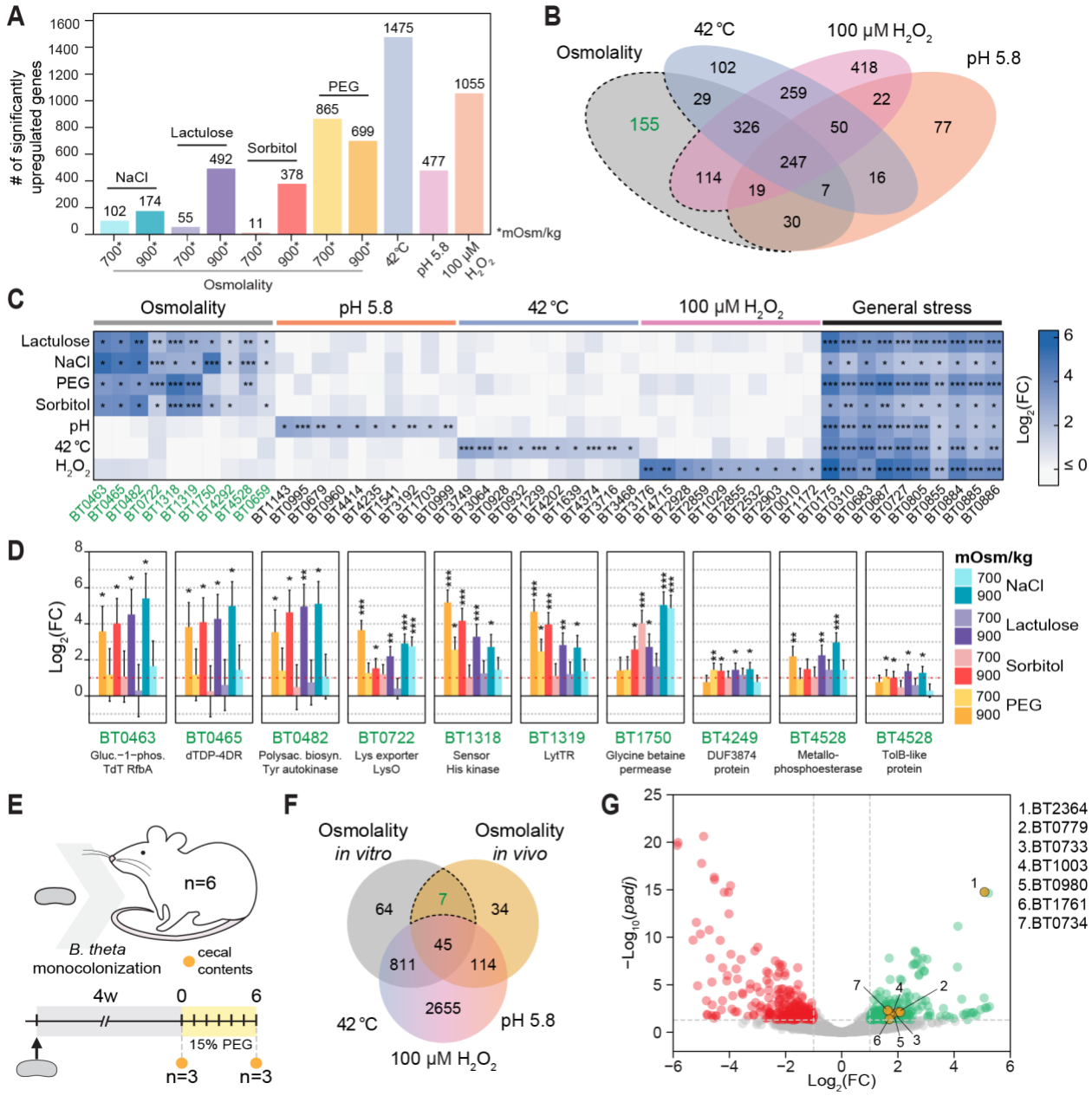
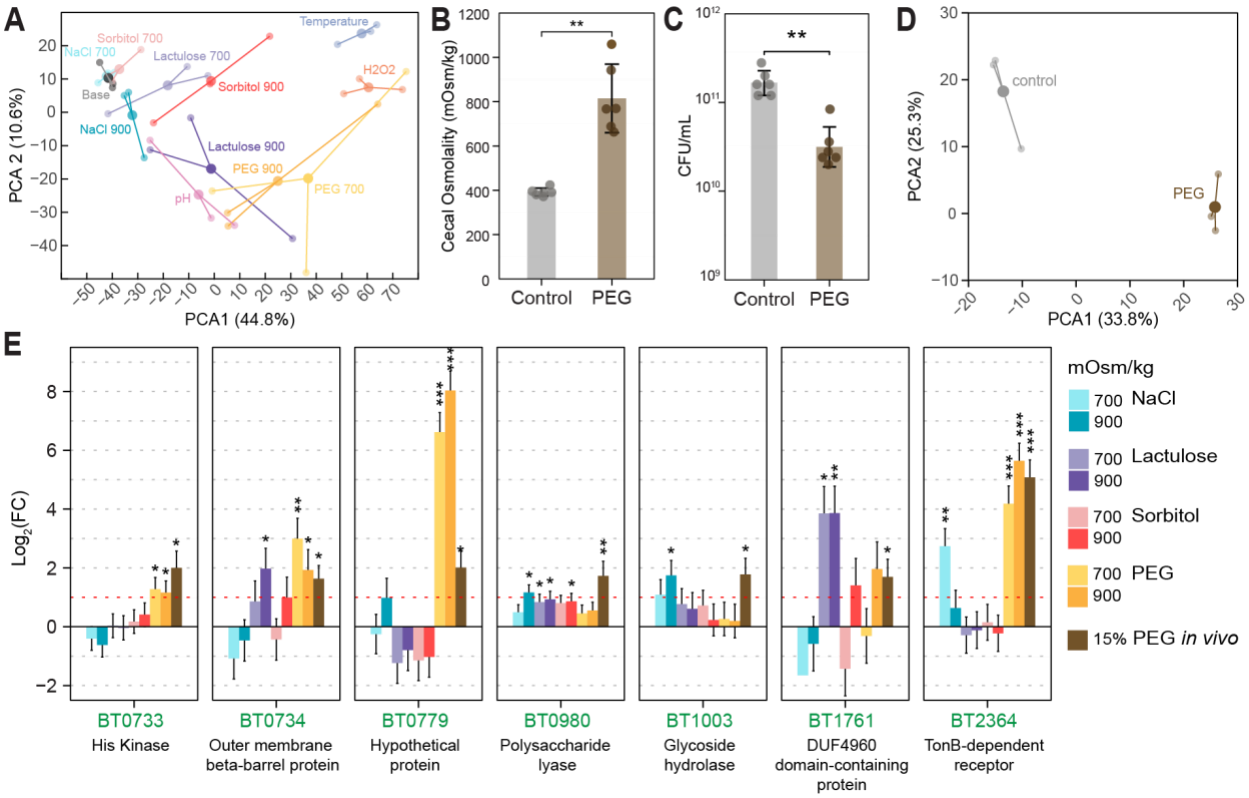


Figure 2. Native *B. theta* promoters respond specifically to physical perturbations *in vitro* and to osmotic stress *in vivo*. A) Significantly upregulated genes compared to base media (*padj* < 0.05 and Log_2(FC)> 1) for every *in vitro* condition. B) Genes significantly upregulated compared to base media that are unique to specific stresses or shared between conditions (responsive to multiple stressors). Dotted line specifies genes unique to osmotic stress that were significantly upregulated in at least one osmolyte at high osmolality (~900 mOsm/kg) C) Log_2(FC) of the 10 most upregulated genes unique to each stress condition (osmolality, pH 5, 42 ºC and 100 μM H_2O_2) and top 10 most upregulated genes present in all conditions (general stress). D) Log_2(FC) of the top 10 genes from osmotic stress conditions at medium (~700 mOsm/kg) and high (~900 mOsm/kg) for all osmolytes tested. E) Experimental schematic for *in vivo* RNA-seq experiment. F) Genes significantly upregulated (*padj* < 0.05 and Log_2(FC) > 1) compared to base media or untreated mice for *in vitro* and *in vivo* experiments, respectively, that are unique or shared between conditions tested. Dotted line highlights genes upregulated in at least one of the high osmolality (900

mOsm/kg) conditions *in vitro*, upregulated *in vivo*, but not upregulated in the general stress conditions. G) Volcano plot showing gene expression changes of all *B. theta* genes from the *in vivo* PEG experiment. Highlighted points are the 7 genes identified in Figure 2F.
D: Error bars represent the standard error from $n = 3$ biological replicates.
C,D: Statistics calculated in DESeq2[30] pipeline (* *padj* < 0.05; ** *padj* < 0.01; *** *padj* < 0.001)



Supplementary Figure 2. The *B. theta* transcriptional landscape is altered by osmolality *in vitro* and *in vivo*. A) Principal component analysis (PCA) of the *B. theta* transcriptome after treatment with NaCl, lactulose, PEG, or sorbitol *in vitro*. Small circles indicate the coordinates of each biological replicate, large circles indicate the mean position of three biological replicates, and lines mark the Euclidean distance from each replicate to the mean (shown in Figure 2A). B) Cecal osmolality of mice treated with 15% PEG for 6 days, or untreated controls. C) Bacterial abundance counts from mice treated with 15% PEG for 6 days, or untreated controls. D) PCA of the *B. theta* transcriptome from mono-colonized mice treated with 15% PEG for 6 days and untreated controls. E) Differential expression analysis of the 7 genes identified that are specifically upregulated in osmotic stress conditions *in vitro* and *in vivo* (highlighted by the dotted line in Fig. 1F). Log_2(FC) of the genes in each *in vitro* osmotic stress condition and following 6 days of treatment with 15% PEG are shown.
B,C: Error bars indicate the standard deviation of $n = 3$ biological replicates. Pairwise ANOVA with Tukey's multiple correction (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$)
E: Error bars indicate the standard error of $n = 3$ biological replicates. Statistics calculated in DESeq2 pipeline (* *padj* < 0.05; ** *padj* < 0.01; *** *padj* < 0.001)

Biosensors incorporating osmolality-specific promoters respond to osmotic stress in vitro

Having identified a selection of osmolality-specific *B. theta* genes, we next aimed to construct candidate osmolality biosensor circuits using promoters from these genes in our inverted transcriptional reporter design (Figure 1A). To screen for functional biosensors, we inserted the best 5 promoter candidates from

genes identified in both our *in vitro* and *in vivo* RNA-seq analyses into our circuit upstream of the *tetR* gene, with the engineered P_{TetR56} promoter driving GFP (Figure 3A). As we ultimately required a fluorescent output that could be distinguished from fecal material, we changed the GFP in our reporter circuit to mGreenLantern (mGL), which is known to be brighter than sfGFP[31,32]. We then grew biosensor strains in base media (~320 mOsm/Kg) and media adjusted to high osmolality (~850 mOsm/kg) with PEG and measured cellular GFP and RFP levels. We calculated reporter activation as the fold decrease in the ratio of GFP to the constitutive RFP control (FC(GFP:RFP)) between base media and osmotic stress conditions. To verify that GFP levels are not decreased by osmotic conditions in the absence of promoter induction, we constructed a control strain in which no promoter is inserted upstream of TetR, and GFP is therefore not repressed (no promoter control [NP]).

Three biosensor candidates incorporating the promoters from the BT2364, *opuA* (BT1749-51), and BT0779 genes showed activation in response to increased osmolality *in vitro* (Figure 3B). As BT2364 is a known SusC/D transporter[33] in carbohydrate utilization and could be affected by dietary intake in addition to osmotic perturbations, we proceeded with the other two candidates. The *opuA* operon is known to be involved in osmotic stress response in many other bacteria[34–36]. BT0779 is a hypothetical protein that has homology to a *Bacteroides* HdeA/HdeB chaperone family protein according to a BLASTP[37] search against the refseq_protein database. In our screen, the NP control showed a minimal response of ~1.2x activation (Figure 3B; Supplementary Figure 3A). The P_{opuA} and P_{BT0779} biosensors showed a 4.3- and 3.0-fold activation in media adjusted with PEG, respectively (Figure 3B). As predicted, this change in GFP:RFP ratio was due to increased repression of GFP with increased osmolality in both biosensors, while RFP levels remain stable across conditions (Supplementary Figure 3A). In contrast, the remaining biosensors tested showed minimal responses similar to the NP control (Figure 3B; Supplementary Figure 3A).

Taking the P_{opuA} and P_{BT0779} biosensors as our most promising candidates, we aimed to further characterize the dose-responses and dynamic range of these reporters. To this end, we grew these strains in media ranging from ~320 to ~900 mOsm/kg, adjusted with PEG, and measured reporter activation. Both biosensors showed repression of GFP leading to graded biosensor activation in response to increasing osmolality, which reaches maximal activation at ~700 mOsm/kg (Figure 3C; Supplementary Figure 3B). The graded activation of the two biosensors to increasing levels of PEG-mediated osmolality *in vitro* suggested that their activity could also detect PEG-induced malabsorption in the gut.

Biosensor dynamic response improves with sequestration of excess repressors

While both the P_{opuA} and P_{BT0779} biosensors responded to increasing osmolality, we noticed that GFP levels were substantially lower than the NP control in baseline media (18.9 and 9.1% of the NP control, respectively) (Supplementary Figure 3B), suggesting that non-zero expression of TetR from the native promoters led to repression of GFP. Since our biosensor response depends on repression of GFP, we reasoned that low baseline GFP levels constrained the dynamic range of our sensors. We therefore hypothesized that sequestering excess TetR from our circuit would increase baseline GFP levels and improve the potential dynamic range of our reporters (Figure 3D). Indeed, previous synthetic circuits incorporating DNA-based repressor-binding "sponges" have been shown to decrease baseline levels of repressor[38] and increase fold change of circuit outputs[12].

To predict the effects of repressor sequestration on the dynamic range of our inverted transcriptional reporter circuit, we simulated our circuit using an ODE-based mathematical model (Methods). The model includes transcription and translation of TetR and GFP, and a TetR 'sponging' term that follows Hill ligand binding dynamics. After fitting model parameters using dose-response curves for the P_{TetR56} promoter (Figure 1D) and the P_{opuA} biosensor (Supplementary Figure 3C), we calculated steady-state GFP levels over a range of induction with varying degrees of sponging (Figure 3E). Including TetR sponging in our model

improved the dynamic range of the system at all induction levels (Figure 3F), though over-sponging resulting in a lack of GFP repression at high induction levels was possible (Figure 3E).

To implement repressor sponging experimentally, we used both chemical and DNA-based sponging methods to sequester TetR in two experiments. We first sequestered TetR chemically using aTc to prevent binding to the TetO operator site, as this method offered a rapid method to evaluate the effect of TetR sequestration without the need for additional cloning. Here, we performed PEG-osmolality dose-response curves as above, but with a range of aTc concentrations supplemented into growth media (Supplementary Figure 3C). In agreement with modeling predictions, TetR sequestration using aTc improved the dynamic range of the biosensors. Specifically, both P_{opuA} and P_{BT0779} biosensors reached a maximal activation (4.9 and 7.8 FC(GFP:RFP), respectively) at the highest osmolality tested (~850 mOsm/kg) with 0.25 ng/mL aTc supplemented into the media (Supplementary Figure 3C). While these results confirm that TetR sequestration improved biosensor performance, a sponging system that does not require small molecule supplementation was preferable for *in vivo* biosensor implementation. Next, we therefore genetically encoded TetR sponging in our biosensor plasmids by including cassettes containing 0, 10, 20, 40, 80, or 160 TetO binding sites ("sponges")[12] and performed PEG-osmolality dose-response curves as above (Figure 3G). With DNA-based repressor sponging, P_{opuA} showed its maximal activation (7.8 FC(GFP:RFP)) at the highest osmolality tested (~880 mOsm/kg) with 20 TetO sites added to the circuit (P_{opuA}+20xTetO)(Figure 3H). P_{BT0779} was more sensitive, reaching its maximal activation (16.1 FC(GFP:RFP)) at ~700 mOsm/kg with 80 TetO binding sites added (P_{BT0779}+80xTetO) (Figure 3H). The difference in optimal sponging levels between biosensors likely stemmed from a higher baseline TetR expression by P_{BT0779} that is reflected in its lower baseline levels of GFP observed in dose response curves (Supplementary Figure 3B). Importantly, the increased dynamic range with TetR sequestration resulted from higher GFP levels at baseline (~320 mOsm/kg), while GFP repression at high osmolality remained strong, as predicted (Supplementary Figure 3D,E).

Having developed biosensors for two promoters (P_{opuA} and P_{BT0779}) with a robust response to increased osmolality in PEG, we set out to verify their response to other common osmolytes. We therefore grew the P_{opuA}+20xTetO and P_{BT0779}+160xTetO biosensors in gradients of the following ionic and non-ionic osmolytes, and compounds predicted to affect osmotic stress: KCl, NaCl, Na_2SO_4, $(NH_4)_2SO_4$, fructose, lactulose, galactose, sucrose, xylose, sorbitol, glycerol, proline, and betaine. The P_{opuA}+20xTetO biosensor responded to most osmolytes tested, demonstrating that this biosensor is broadly responsive to increases in environmental osmolality (Figure 3H). In contrast, the P_{BT0779}+160xTetO biosensor responded strongly to Na_2SO_4 and sucrose, minimally to KCl, NaCl, and $(NH_4)_2SO_4$, and did not respond to other osmolytes (Figure 3H). This may suggest that the previously uncharacterized BT0779 gene responds more specifically to particular causes of osmotic, dehydration, or water structure-related stressors. To ensure that our biosensors do not respond to membrane stress outside of changes to hydration or osmolality, we also tested our P_{opuA}+10xTetO and P_{BT0779}+160xTetO biosensors in increasing concentrations of several membrane damaging agents (bile salts, Triton-X 100, saponin, and SDS), to which we observed no biosensor activation (Supplementary Figure 3F).

Altogether, we showed that TetR sequestration improves the dynamic range of the transcriptional reporter circuit. This optimization strategy provides a toolkit to fine-tune biosensor activity *in vitro*, laying a foundation for their translation into *in vivo* applications.

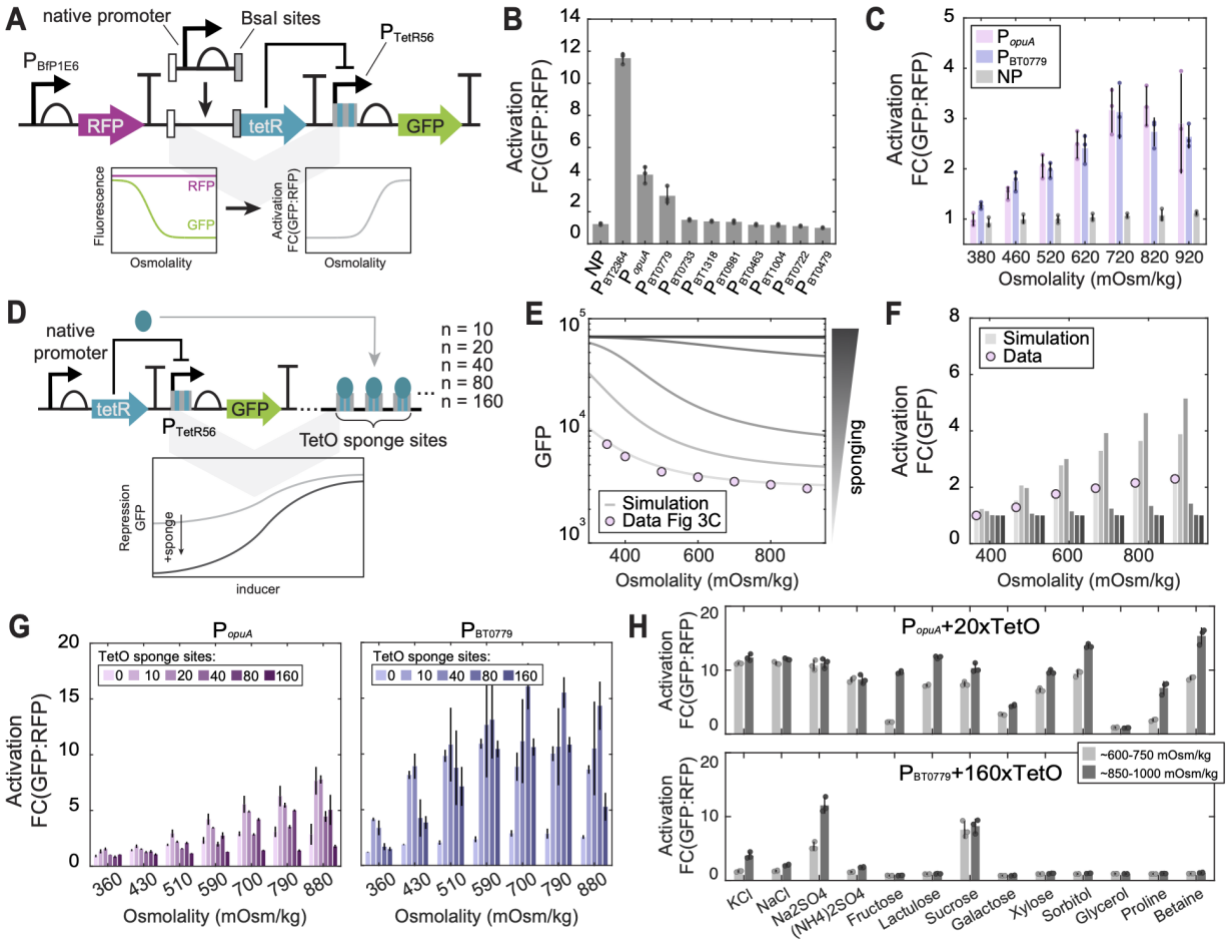
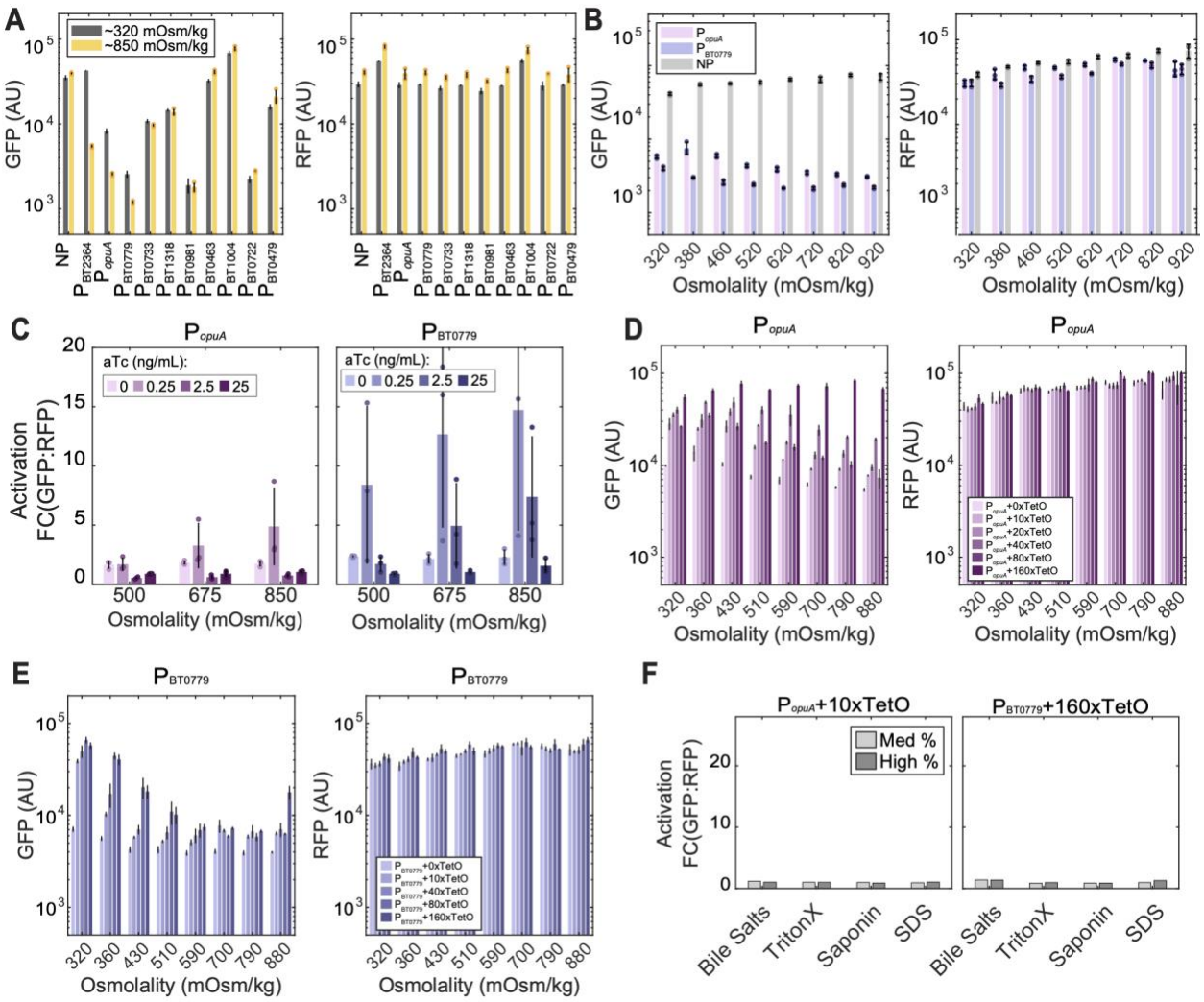


Figure 3. Biosensor response to osmotic stress *in vitro* is enhanced by TetR sequestration. A) A modular Golden Gate entry vector facilitates the assembly of *B. theta* biosensors which can report on the activity of arbitrary native promoters. GFP is downregulated as the native promoter is induced, and a constitutive RFP expressed from P_{BfP1E6} allows for normalization. Biosensor activity is measured as the fold decrease in the ratio GFP to RFP (FC(GFP:RFP)) compared to baseline conditions. B) Induction response of a selection of candidate biosensors using different native promoters or no promoter (NP) grown in baseline media (~320 mOsm/kg) or media adjusted to ~850 mOsm/kg with PEG. C) Dose-response curves for the P_{opuA} and P_{BT0779} biosensors in media adjusted to osmolality with PEG. D) Design of a biosensor with TetR 'sponging', which is predicted to lower baseline repression of GFP. E) Steady-state GFP levels from simulations of the inverted P_{opuA} biosensor circuit with increasing degrees of TetR sponging. Model parameters were fitted to the PEG dose-response curve for the no-sponge P_{opuA} biosensor shown in C (pink circles). F) Fold decrease in steady state GFP levels from simulations shown in E. G) Biosensor response to osmolality with genetically-encoded TetR sequestration. P_{opuA} and P_{BT0779} biosensors with 0, 10, 20, 40, 80, or 160 TetO binding sites present in the circuit were grown in media adjusted to specified osmolalities with PEG. H) Optimized P_{opuA}+20xTetO and P_{BT0779}+160xTetO biosensor response measured in media adjusted to moderate (~600-750 mOsm/kg) or high (~850-1000 mOsm/kg) osmolality with a variety of osmolytes.

B,C,G,H: Error bars represent the standard deviation from n = 3 biological replicates (unique colonies) of each biosensor. Fold decrease in GFP:RFP is calculated relative to those in baseline conditions (~320 mOsm/kg).



Supplementary Figure 3. Sequestration of excess repressor lowers baseline repression of GFP, improving biosensor dynamic range. A) GFP and RFP levels for each biosensor shown in Figure 3B in media adjusted to 320 mOsm/kg or 850 mOsm/kg with PEG. B) P_{opuA} and P_{BT0779} biosensor GFP and RFP levels from experiment shown in Figure 3C in media adjusted to increasing osmolalities with PEG. C) Biosensor response to osmolality with chemical sequestration of TetR via aTc supplementation. P_{opuA} and P_{BT0779} biosensors were grown in media adjusted to specified osmolalities with PEG, and supplemented with 0, 0.25, 2.5, or 25 ng/mL aTc, and GFP:RFP ratios were compared to baseline media (~320 mOsm/kg). D) P_{opuA} biosensor GFP and RFP levels from the TetR sponging experiment in PEG shown in Figure 3G. E) P_{BT0779} biosensor GFP and RFP levels from the TetR sponging experiment in PEG shown in Figure 3G. F) P_{opuA}+10xTetO response measured in media supplemented with membrane-damaging agents at the following concentrations (Med % or High %): Bile salts (0.25% or 0.5% w/v), Triton-X (0.0031% or 0.0063% v/v), Saponin (0.025% or 0.05% w/v), SDS (0.0016% or 0.0031% v/v). For Triton-X and Saponin, concentrations higher than those listed did not allow cell growth.

A,B,C,D: Error bars represent the standard deviation from n = 3 biological replicates (unique colonies) of biosensors in each condition

Identification of an alternative integration site enhances P_{opuA} biosensor performance

The broad and consistent *in vitro* response of our P_{opuA} biosensor to diverse osmolytes (Figure 3H) makes it a promising malabsorption biosensor *in vivo*, where malabsorption can result from a wide range of dietary compounds, pharmaceuticals, or pathologies. To evaluate the performance of our P_{opuA} biosensors under physiological conditions, we tested their response to increased gut osmolality in a murine model of PEG-induced malabsorption. We first colonized germ-free mice with either the P_{opuA}+20xTetO biosensor or the NP control and allowed 1-2 weeks for colonization. After equilibration, we treated mice with 0, 10, or 15% PEG in drinking water for 6 days. Following treatment, we quantified cecal osmolality as a measure of PEG-induced malabsorption, and measured biosensor activation from cecal contents using flow cytometry (Figure 4A). At Day 6, mice colonized with either bacterial strain had average cecal osmolalities of ~380 mOsm/kg without PEG treatment, and significantly increased cecal osmolalities of ~550 mOsm/kg and ~650 mOsm/kg in the 10% and 15% PEG treatment groups, respectively (Figure 4B). This result confirmed that PEG administration effectively induces malabsorption and increases gut osmolality. However, despite these increased osmolality levels, both the P_{opuA}+20xTetO biosensor and the NP control strain showed similar minimal activation levels in all treatment conditions (<1.1-fold activation), when single-cell GFP and RFP levels were measured by flow cytometry (Figure 4C). We reasoned that the lack of P_{opuA}+20xTetO response *in vivo* could be due to the limited cecal osmolality reached; *in vitro*, the activation of this biosensor is less than half maximal at ~600 mOsm/kg and continued to increase up to 880 mOsm/kg (Figure 3G). The osmolality reached with 15% PEG treatments *in vivo* may therefore not have been sufficient to trigger a measurable response.

During previous *in vitro* characterization of the P_{opuA}+20xTetO biosensor, we observed colonies that displayed unusually high activation, which we termed a "high response" variant (P_{opuA}+20xTetO^{HR}). Isolating and further characterizing these colonies revealed that the P_{opuA}+20xTetO^{HR} variant consistently outperformed P_{opuA}+20xTetO, showing up to 1.9X stronger activation and reaching maximal induction at lower osmolality (Figure 4D). As its strong activation at ~600-700 mOsm/kg offered a promising avenue to improve *in vivo* biosensor signal, we further characterized P_{opuA}+20xTetO^{HR} via genomic sequencing and PCR verification. These analyses revealed that P_{opuA}+20xTetO^{HR} colonies had undergone homologous recombination to insert the pNBU2-based biosensor plasmid at the native P_{opuA} locus instead of the canonical attBT2-1 or attBT2-2 insertion sites. In this variant, the entire biosensor plasmid was therefore inserted upstream of the native P_{opuA} promoter without disrupting the *opuA* operon (Figure 4E; Supplementary Figure 4A). Importantly, integration via homologous recombination did not impact strain fitness at any osmolality level in a growth assay, indicating that normal *opuA* function is maintained in these cells (Figure 4F; Supplementary Figures 4B, C). These results demonstrate that, in agreement with studies in *E. coli*[41,42] and other bacteria[43], genomic location can impact circuit dynamics in *B. theta*, and present a novel approach to integrate and optimize the behavior of transcriptional reporter circuits in *Bacteroides*.

Given that the P_{opuA}+20xTetO^{HR} biosensor variant—which was integrated via homologous recombination—showed an increased dynamic change and decreased induction threshold in response to increased osmolality *in vitro* (Figure 4D), we hypothesized that this biosensor could display an improved activation in our PEG-malabsorption mouse model. To test this, we colonized germ-free mice with the P_{opuA}+20xTetO^{HR} strain and treated them with either 15 or 20% PEG in drinking water. During treatment, we collected fecal samples for flow cytometry measurements of biosensor fluorescence on Days 0, 2 (20% PEG group) or 3 (15% PEG group), and 6 of treatment (Figure 4G). At Day 6, mice treated with both 15 and 20% PEG were euthanized and cecal contents were collected for osmolality measurements and flow cytometry analysis. Mice in both treatment groups had an average cecal osmolality of ~660 mOsm/kg (Figure 4H), comparable to the levels previously obtained with P_{opuA}+20xTetO (Figure 4E). We observed that the P_{opuA}+20xTetO^{HR} biosensor showed significant activation by Day 6 of treatment with both 15 and 20% PEG (Figure 4I). Similar final activation levels were reached in the two treatment groups (1.4- and 1.3-fold activation with 15 and 20% PEG treatment, respectively). However, measurements from fecal samples collected at multiple timepoints revealed that the biosensor was activated faster in higher PEG concentrations: in 20% PEG treatment,

biosensors reached a significant 1.2-fold activation by Day 2, whereas in 15% PEG treatment, biosensors had not yet responded on Day 3 (Figure 4I). These results confirm that the P_{opuA}+20xTetO^{HR} effectively detects PEG-induced malabsorption *in vivo* and demonstrate that the genomic integration site plays a crucial role in optimizing biosensor sensitivity and response dynamics. By leveraging targeted integration strategies, this approach enhances the potential for continuous, non-invasive monitoring of the gut environment via fecal sampling.

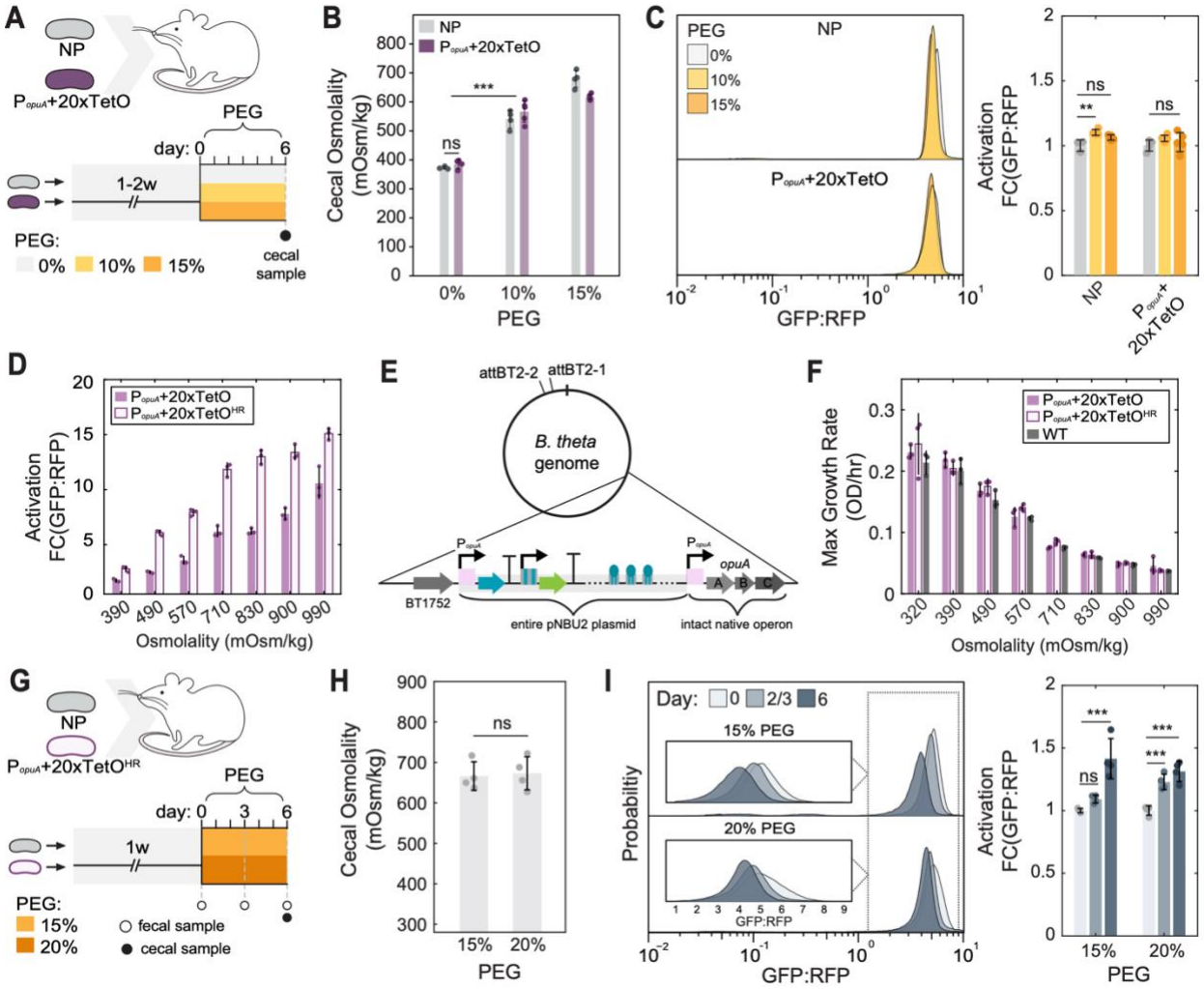
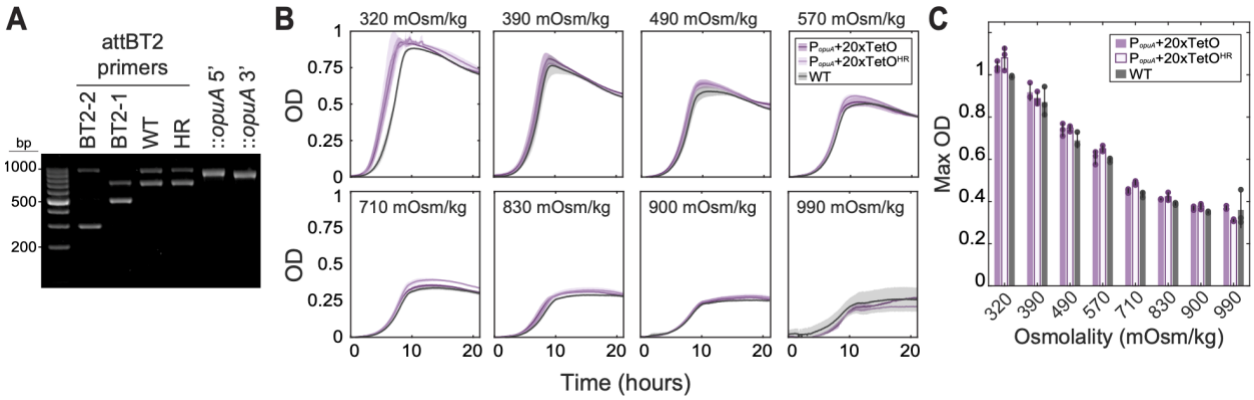


Figure 4. Integration site of transcriptional reporter circuit enhances *B. theta* osmolality sensing *in vitro* and in a malabsorption mouse model. A) Experimental outline of PEG-induced malabsorption mouse model. Mice were mono-colonized with either the NP, or P_{opuA}+20xTetO and treated with 0, 10, or 15% PEG in the drinking water for 6 days. B) Cecal osmolality measurements after 6 days of 0, 10, or 15% PEG treatment. C) *(left*) Single-cell GFP:RFP ratios measured from cecal contents with flow cytometry for 0, 10, and 15% PEG treatments. (*right*) Biosensor activation (fold decrease in the median GFP:RFP ratio) for each biosensor strain from flow cytometry data. D) Activation in response to increasing osmolality adjusted with PEG, measured in isolated biosensors with "high response" (P_{opuA}+20xTetO^{HR}) and standard (P_{opuA}+20xTetO^{HR}) activation. E) Integration of pNBU2 plasmid can occur at the canonical attBT2-1 or attBT2-2 sites, or can undergo homologous recombination in the overlapping promoter sequence without disruption of the native gene. F) Growth rates extracted from growth curves of P_{opuA}+20xTetO and

P_{opuA}+20xTetO^{HR} biosensors, and wild-type *B. theta* in media at increasing osmolality adjusted with PEG indicate that plasmid insertion via homologous recombination does not affect cell fitness. G) Experimental outline of PEG-induced malabsorption mouse model. Mice were mono-colonized with the P_{opuA}+20xTetO^{HR} biosensor and treated with 15 or 20% PEG in the drinking water for 6 days. H) Cecal osmolality measurements after 6 days of PEG treatment. I) (*left*) Single-cell GFP:RFP ratios measured from fecal contents with flow cytometry for Days 0, 2/3, and 6 of PEG. (*right*) Fold decrease in the median GFP:RFP ratio for each biosensor strain from flow cytometry data.
B, C, H, I: Bar values represent the mean value for n = 3 or 4 mice in a treatment group. Error bars represent standard deviation within a group.
C, I: Statistics shown are run using the median GFP:RFP values for each mouse in a treatment group. Error bars represent standard deviation within a treatment group.
Stats: Pairwise ANOVAs with multiple comparisons (* $padj < 0.05$; ** $padj < 0.01$; *** $padj < 0.001$)
D, F: Error bars represent the standard deviation from n = 3 biological replicates (unique colonies) of biosensors in each condition.



Supplementary Figure 4. Integration of transcriptional reporter circuits can occur via homologous recombination without affecting fitness. A) Agarose gel electrophoresis showing (*right*) PCR products from the three possible insertion sites or wild type *B. theta* using our insertion verification primers, and (*left*) PCR products from 2 sets of primers verifying insertion of the P_{opuA}+20xTetO^{HR} biosensor circuits via homologous recombination at the native P_{opuA} genomic site. B) OD_{600} measurements from growth curves for P_{opuA}+20xTetO and P_{opuA}+20xTetO^{HR} biosensors in media adjusted to increasing osmolalities with PEG. C) Max OD_{600} values from growth curves shown in B.
B: Solid lines and shaded error bars represent the mean and standard deviation from n = 3 biological replicates (unique colonies) of biosensors in each condition
C: Error bars represent the standard deviation from n = 3 biological replicates (unique colonies) of biosensors in each condition.

An optimized biosensor circuit reports on osmotic perturbations in an* in vivo *mouse model of laxative-induced malabsorption

While our P_{opuA}+20xTetO^{HR} biosensor effectively responded to the high intestinal osmolalities induced by the consumption of high levels of PEG, a sensor capable of detecting more subtle osmolality shifts would be valuable for diagnosing less severe forms of malabsorption. Given its enhanced sensitivity *in vitro* (Figure 3H), we predicted that our P_{BT0779} biosensor would provide a more sensitive response to PEG-induced malabsorption *in vivo*. First, we tested the response of the biosensor under the high levels of malabsorption tested with the P_{opuA} biosensors. We colonized germ-free mice with either P_{BT0779}+160xTetO or the NP control and allowed 4 weeks for colonization. After equilibration, we treated mice with 0%, 10%, or 15% PEG in drinking water for 6 days to stimulate malabsorption (Figure 5A). Mice treated with 10%

and 15% PEG had average cecal osmolalities of ~550 and 650 mOsm/kg, respectively, with no significant difference in osmolality between mice colonized with NP or P_{BT0779}+160xTetO biosensors (Figure 5B). All PEG-treated groups had significantly higher cecal osmolality than untreated mice (~375 mOsm/kg) (Figure 5B). As expected, the NP control showed no activation after PEG treatment, while P_{BT0779}+160xTetO showed a 55.1- and 48.3-fold decrease in GFP:RFP levels in 10% and 15% PEG-treated mice, respectively, compared to untreated mice and the NP control (Figure 5C). The majority of single P_{BT0779}+160xTetO biosensor cells in both 10% and 15% PEG-treated mice reached similar GFP:RFP levels to a control strain not producing GFP ("OFF"), suggesting that the biosensor reached maximum induction in these conditions (Figure 5C). As observed *in vitro*, this decrease in GFP:RFP levels is the result of decreased GFP levels while RFP levels remained stable (Supplementary Figure 5A). These results indicate that P_{BT0779}+160xTetO is highly sensitive to PEG-induced malabsorption in the gut.

Next, we investigated whether the P_{BT0779}+160xTetO biosensor could detect mild PEG-induced malabsorption by responding to subtle osmolality shifts, even before overt clinical symptoms such as increased fecal water content and decreased stool consistency become apparent. To test this, we colonized germ-free mice with P_{BT0779}+160xTetO and treated them with either 0%, 1%, 2.5%, 5%, or 7.5% PEG for 3 days, after which we collected cecal and fecal samples (Figure 5D). Mean cecal osmolalities reached 366, 384, 422, 479, and 493 mOsm/kg for the 0%, 1%, 2.5%, 5%, or 7.5% treatment groups, respectively (Figure 5E). While cecal osmolality increased gradually with PEG treatment, standard non-invasive measures of malabsorption, such as fecal water content and feces consistency as measured by Bristol Stool Scoring, did not differ between treated and untreated mice, except for 7.5% PEG (Figure 5F,G). The level of malabsorption present would therefore not be detectable by standard clinical methods. The P_{BT0779}+160xTetO biosensor showed strong activation in response to increasing osmolality, measured from both fecal (Figure 5H) and cecal samples (Figure 5I,J). As at higher osmolalities, changes in GFP:RFP ratio are the result of decreased GFP levels while RFP levels remained stable (Supplementary Figure 5B). Biosensors from fecal samples were most activated in mice treated with 5-7.5% PEG, showing an average ~40-fold activation compared to untreated mice (Figure 5H). In individual mice, P_{BT0779}+160xTetO GFP:RFP ratios from cecal samples correlate well with cecal osmolality (Figure 5I), showing this sensor's ability to discern even small differences in osmolality. Notably, GFP:RFP measurements from fecal and cecal samples were strongly linearly correlated ($r^2 = 0.98$; Figure 5J), demonstrating that biosensor response can be reliably measured non-invasively from stool samples. Overall, *in vivo* testing demonstrated that our P_{BT0779}+160xTetO biosensor can reliably detect and report osmolality changes from cecal and fecal samples and is sensitive enough to differentiate malabsorption severity prior to the onset of overt symptoms, underscoring its utility for sensitive, non-invasive gut environmental monitoring.

Taken together, our results provide novel synthetic biology tools for *B. theta* engineering, identify osmolality-responsive promoters, and validate biosensors capable of detecting and reporting malabsorption of relevant pharmaceutical and dietary compounds *in vitro* and *in vivo*. These findings demonstrate the potential of engineered *B. theta* for monitoring of gut physiological states and support the potential impact of microbiota-based diagnostics for gut health.

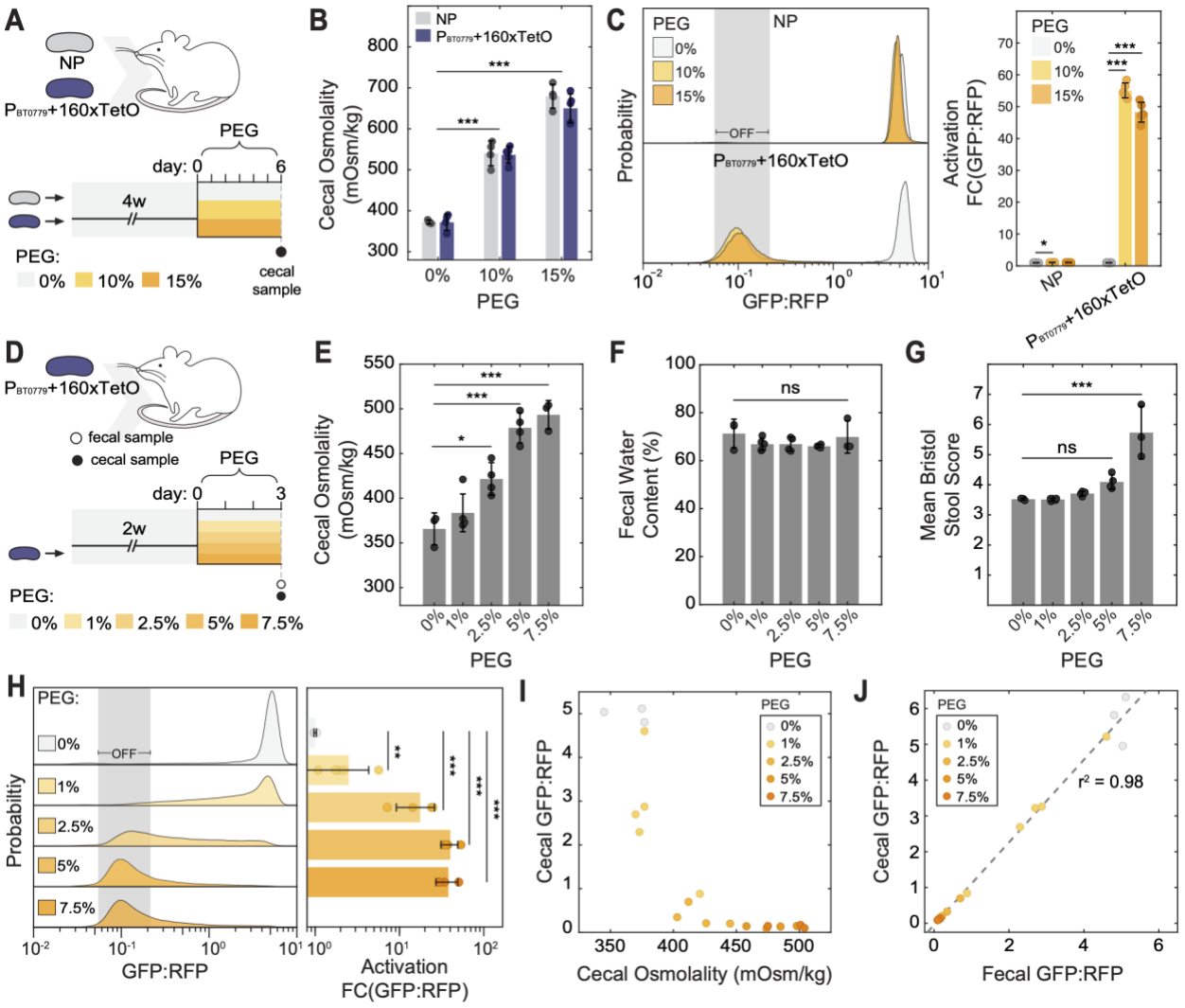
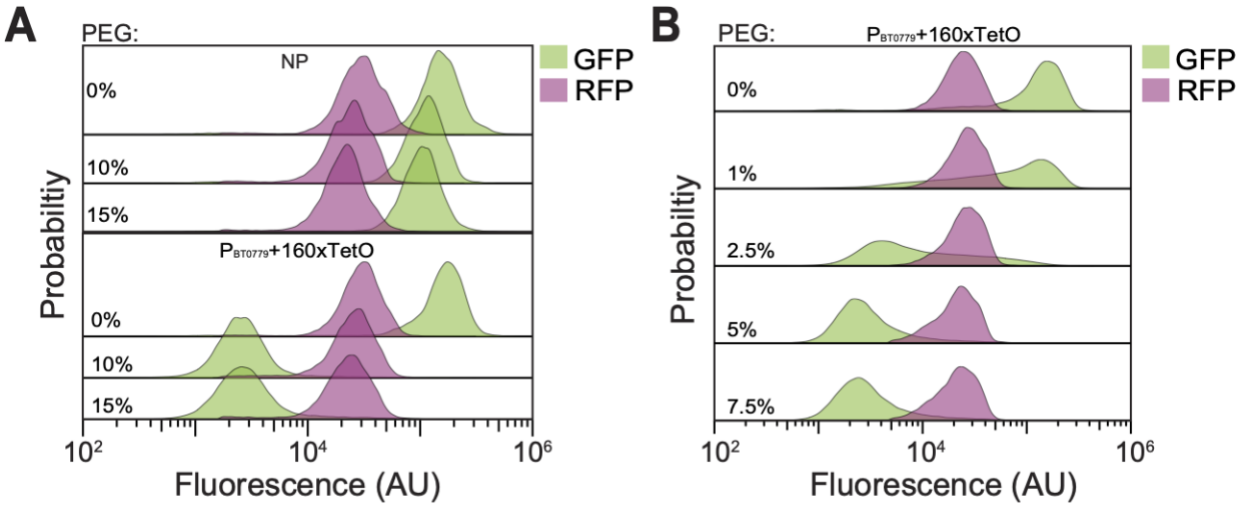


Figure 5. A *B. theta* biosensor can detect laxative-induced malabsorption at subclinical osmolality levels non-invasively. A) Outline of initial biosensor test *in vivo*. Mice were mono-colonized with either the NP control or P_{BT0779}+160xTetO biosensor, and treated with 0%, 10%, or 15% PEG in the drinking water for 6 days. B) Cecal osmolality measurements after 6 days of PEG treatment. C) (*left*) Single-cell GFP:RFP ratios measured via flow cytometry from cecal contents of 3-4 mice in each treatment group. Shaded gray area represents background GFP fluorescence levels as measured from a control strain that produces constant RFP but no GFP. (*right*) Fold decrease in the median GFP:RFP ratio for each biosensor strain from flow cytometry data. Dots indicated median GFP:RFP values for each mouse. Statistics are run on median GFP:RFP values. D) Outline of mouse experiment examining the response of P_{BT0779} in subclinical osmolality levels. Mice were mono-colonized with P_{BT0779}+160xTetO and treated with 0%, 1%, 2.5%, 5%, or 7.5% PEG in the drinking water for 3 days. Cecal and fecal contents were collected after 3 days. E) Cecal osmolality measurements after 3 days of PEG treatment. F) Fecal water content measured after 3 days of PEG treatment. G) Bristol Stool Scores from fecal pellets 3 days after PEG treatment. H) (*left*) Single-cell GFP:RFP ratios measured from fecal contents with flow cytometry. (*right*) Fold decrease in the mean GFP:RFP ratio for each biosensor strain from flow cytometry data. Statistics shown are run on the mean GFP:RFP values for each mouse. I) Median cecal GFP:RFP ratio as a function of cecal osmolality for individual mice in all treatment conditions. J) Median GFP:RFP ratio from fecal and cecal samples. Dotted line shows linear fit, calculated with fit.m function in MATLAB.

B, C, E, F, G, H: Bar values represent the mean value for n = 3 or 4 mice in a treatment group. Error bars represent standard deviation within a group.
Stats: Pairwise ANOVAs with multiple comparisons (* *padj* < 0.05; ** *padj* < 0.01; *** *padj* < 0.001)



Supplementary Figure 5. Biosensor fluorescence measurements are well correlated in fecal and cecal samples A) GFP and RFP measurements taken by flow cytometry from cecal samples of mice colonized with either NP or P_{BT0779}+160xTetO biosensors and treated with either 0, 10, or 15% PEG. B) GFP and RFP measurements taken by flow cytometry from cecal samples of mice colonized with P_{BT0779}+160xTetO biosensor and treated with either 0, 1, 2.5, 5, or 7.5% PEG.

Discussion

In this work, we developed and validated two *B. theta* biosensors that robustly detect environmental osmolality shifts both *in vitro* and in laxative-induced malabsorption models. These biosensors enable sensitive, continuous and non-invasive monitoring via fecal sampling, demonstrating their potential for real-time gut diagnostics. Their development was made possible by the expansion of the *B. theta* synthetic biology toolkit, to which we contributed a library of novel small-molecule-inducible promoters, a DNA-based repressor sponging system adapted from *E. coli*, and a modular fluorescence-based transcriptional reporting circuit. Additionally, we characterized a novel synthetic plasmid insertion mode in *B. theta*, providing insights into how genomic integration site can influence biosensor performance. Beyond the applications presented here, our fluorescence-based *B. theta* genetic tools expand the available parts for building complex synthetic circuits in this commensal bacterium. This advancement not only expands our ability to study *B. theta* biology, but also strengthens its potential as a chassis for microbiota-based diagnostics. To ensure broad accessibility, we have made all plasmids containing all of our characterized promoters, TetR sponge cassettes, biosensor entry vectors, and assembled biosensor circuits available on Addgene (Supplementary Table 6) supporting further research and development in the field.

For effective long-term, continuous monitoring applications, the microbial-based biosensors must achieve stable and abundant gut colonization in addition to detecting and responding to environmental changes in the gut. Historically, *E. coli* has been the preferred chassis for gut bacterial biosensors due to its expansive genetic toolkit and ease of engineering[44,45]. For example, biosensors engineered in probiotic *E. coli* Nissle 1917 and murine commensal *E. coli* NGF-1 strains have been used to detect and report on inflammation[9,46,47] and colorectal tumors[48,49] in the GI tract. Despite these developments, even probiotic *E. coli* strains can only colonize transiently and sparsely following disruption of the native microbiota with antibiotics or intestinal

lavage[50,51]. This limits long-term applications for these biosensors, and risks reporting from a perturbed gut environment. In contrast, *B. theta* is a prevalent and abundant bacterium in the human microbiota, showing ~46% prevalence in human microbiota[52] and composing about 6% of the bacterial microbiota when present[53]. Engineered *B. theta* strains can stably colonize an existing conventional microbiota in mice and humans[4,10], circumventing the need to perturb the native microbiota. Members of the *Bacteroidaceae* family also naturally colonize across gut regions[51,52] and in diverse niches of the colon (e.g., the lumen and outer mucus layer[54], and the intestinal crypts[5], an important feature for detecting divergent responses from spatially distinct populations. Finally, engineering *B. theta* to degrade the exogenous metabolite porphyrin was recently shown to allow tuneable levels of colonization in the gut[4,10], which could give the user control over the survival and abundance of engineered strains. Together, its stable colonization, broad niche occupancy, and potential for tunable population control within the gut environment make *B. theta* a compelling chassis for gut biosensors and engineered genetic circuits.

The ability of the biosensors developed here to produce fluorescence outputs expands the possibilities for multiplexing reporters. Unlike previously used luminescence outputs, the distinct excitation and emission profiles of fluorescent proteins allows reporter signal detection from multiple fluorophores. Combined with the orthogonality of the three repressible promoter libraries presented here, utilizing fluorescence output could therefore enable the construction of multiplexed reporters to process multiple inputs simultaneously. For example, combining the P_{BT0779}+160xTetO and P_{opuA}+20xTetO^{HR} biosensors presented here could combine the sensitivity and robustness of each circuit to form a system capable of detecting and differentiating causes and severity of malabsorption. Combining malabsorption detection with other triggers could also be useful to differentiate gut pathologies with overlapping symptoms based on the gut environment. For example, biosensors that detect both malabsorption and inflammation-related perturbations (e.g., increased levels of reactive oxygen species or nitrate) could aid in differentiating gut pathologies such as celiac disease and IBD that present both of these symptoms, from other dietary intolerances that cause malabsorption alone (e.g., lactose, sorbitol, or fructose intolerances)[15]. Rapid assembly and characterization of such multiplexed biosensors is facilitated by our modular biosensor backbone and parts vectors. Furthermore, our transcriptomics analysis of *B. theta* under different environmental conditions already suggests candidate promoters that could be integrated to make biosensors for a range of environmental triggers (Figure 2C). On a large scale, this system could also be used to assemble a comprehensive reporter library for *B. theta*, that could help us characterize this bacterium's response and adaptation to diverse conditions in the gut and identify promoters that respond to clinically relevant environmental cues.

Single-cell transcriptional analyses have become a powerful tool for understanding gut biogeography, with advanced in droplet-based[56,57], combinatorial barcoding[58–60], fluorescence in situ hybridization (FISH)[11], and transcriptional recording[61,62] methods enabling high-throughput, multiplexed screening that broadly characterizes transcriptional responses to gut environment changes. However, these approaches are experimentally demanding, require extensive sample processing, and/or rely on deep sequencing, limiting their application in rapid biosensing for gut monitoring. Additionally, their untargeted nature can bias detection toward highly expressed genes, potentially overlooking small but environmentally relevant transcriptional changes. In contrast, synthetic biosensing circuits use specific promoters to selectively detect gene activity changes, allowing for targeted measurement of specific responses—even for lowly expressed genes overlooked by untargeted approaches. Furthermore, fluorescence reporters facilitate rapid, single-cell analysis via flow cytometry with minimal sample preparation. For example, measurement of signal from our *B. theta* biosensors took ~3 hours for up to 24 samples, from fecal collection to completion of flow cytometry measurements. Processing time was often limited by the maturation time of the constitutive RFP in our circuit (~2 hours), rather than experimental procedures. In situations where rapid processing is imperative and sensitive reporting of a limited number of transcripts is sufficient (e.g., detection of limited factors for diagnostics), our synthetic biosensor approach allows near real-time signal detection and non-invasive sampling, therefore offering a streamlined option for transcriptional analysis.

The transcriptional reporter system and resulting biosensors presented here can help us interrogate important questions about the gut microbiota and its biogeography in health and disease. The malabsorption detectors characterized in this system could be used to characterize osmotic changes to the gut environment in clinically relevant models affecting malabsorption, intestinal permeability, or hydration, or such as pathogen invasion (e.g., *Trichinella spiralis*[63] or Enteropathogenic *E. coli* (EPEC) infection[64], malabsorptive conditions (e.g., Celiac disease[65], dietary perturbations (e.g., high-fat[66] diet), or inflammatory pathologies (e.g., DSS[67] or immunodeficiency[68] colitis models). Moving towards diagnostics in the human microbiome, gnotobiotic mouse models or gut-on-chip platforms could facilitate the translation of our biosensors to a diagnostic setting without direct administration of these engineered microbes. Along with the novel synthetic biology tools characterized in this work, these tools support the use of *B. theta* in gut environmental monitoring and point to the utility of commensal gut bacteria as diagnostic and therapeutic tools for gut health.

By expanding the synthetic biology toolkit for *B. theta* and demonstrating its application as a biosensor for gut osmolality, this work advances the potential of sensitive microbiota-based diagnostics for real-time, non-invasive gut health monitoring. Beyond malabsorption, the tools and approaches developed here provide a foundation for engineering biosensors to detect the activation of *B. theta* promoters as well as a wide range of gut conditions, including inflammation, pH shifts, and metabolite fluctuations. These innovations establish a foundation for using engineered *Bacteroides* in non-invasive diagnostics and microbiome-based therapeutics.

Methods

Animal Husbandry

Animal experiments at The University of British Columbia were performed in accordance with The University of British Columbia Animal Care Committee. Gnotobiotic mice were generated from germ-free (GF) Swiss Webster colonies. The gnotobiotic mice were maintained in gnotobiotic isolators or ISOcage P-Bioexclusion System (Techniplast) under constant high positive pressure and fed an autoclaved standard diet (Purina Lab Diet 5K67) for the duration of the experiments.
Mice were sacrificed at various time points using carbon dioxide asphyxiation, followed by secondary euthanasia of cervical dislocation.

Bacterial strains and culture conditions

All *E. coli* strains used in this study were grown aerobically at 37°C in LB Miller broth (Fisher BP1426500) or LB Miller plates with 1.5% agar (Fisher BP9744500). Where appropriate, antibiotics were added at the following concentrations: ampicillin/carbenicillin at 100 µg/mL (Fisher BP1760-25/Fisher BP26485); kanamycin at 50 µg/mL (Sigma Aldrich K1377); chloramphenicol at 25 µg/mL (Acros Organics AC227920250). Wild type *B. theta* VPI-5482 (ATCC 29148; GenBank AE015928.1) was used as a base strain for all genetic engineering in *B. theta*. *B. theta* strains were grown in either Brain-Heart Infusion broth (BD B11059) supplemented with 0.5 µg/mL vitamin K1 (Alfa Aesar AAL1057506) and 5 µg/mL hemin chloride (MilliporeSigma 37415GM) (BHIS) or in Tryptone-Yeast extract-Glucose (TYG) media TYG (10 g/L Tryptone [Gibco 211705]; 5 g/L Yeast Extract [Gibco 212750]; 2 g/L Glucose [Fisher BP350500]; 1 M combined K_2HPO_4 and HK_2O_4P, pH 7.2 [Fisher BP362-1 and Fisher BP363-1 combined to pH 7.2], 0.8 % $CaCl_2$ [Fisher BP510-500]; 0.4 mg/mL $FeSO_4$ [Alfa Aesar AAA1517836]; 0.02 g/L $MgSO_4$ [Fisher BP213-1]; 0.4 g/L $NaHCO_3$ [Fisher BP328-500]; 0.08g/L NaCl [Fisher S271]) supplemented with vitamin K1 (0.5 µg/mL), hemin chloride (5 µg/mL) and cysteine hydrochloride (0.5 g/L; BP376-100 Fisher) (TYG). All *B. theta* strains were streaked out in BHIS plates with 1.5% agar. Where appropriate, antibiotics were added: gentamicin (Gen) at 200 µg/mL (Sigma Aldrich G3632-5G); erythromycin (Erm) at 25 µg/mL (Alfa Aesar, AAJ6227909). *B. theta* strains were grown in an anaerobic

chamber (Mandel) under the following atmospheric conditions: 5% CO_2 and 2.4-5% H_2 in N_2 (Gas Mix: Linde BioBlend NI CD5H11U−T). All plates and media used to grow *B. theta* were pre-reduced under anaerobic conditions for at least 12 hours before use. All bacterial cell cultures were stored in 25% glycerol (Fisher G331) at -80°C for long-term storage.

Genetic parts, plasmid, and strain construction

All genetic parts, primers, and plasmids used in the construction of strains are listed in Supplementary Table 7. All assembled plasmids were sequence verified by Oxford Nanopore sequencing (Plasmidsaurus) and have been deposited on Addgene under the accession numbers specific in Supplementary Table 6.

Competent cells and transformations

All cloning and plasmid construction was completed using either *E. coli* S17-1 λpir (ATCC-BAA-2428) or commercially available competent *E. coli* OneShot™ PIR1 cells (Invitrogen C101010). *E. coli* S17-1 λpir were made electrocompetent or chemically competent using the TSS method[69].

TSS chemically competent cells

An overnight culture of *E. coli* S17-1 λpir was sub-cultured at a 1:100 ratio and grown ~2-3 hours until an OD_{600} of 0.4-0.5. Cells were then harvested by centrifugation at 4000 RPM for 10 minutes in a 4°C rotor. Supernatant was removed and cells were resuspended in 10% of the original culture volume of TSS buffer (10% w/v PEG 3350 [Sigma 202444]; 0.03 M $MgCl_2$ [Fisher BP214]; 5% v/v DMSO [Invitrogen D12345] in LB-Miller broth). Cells were dispensed into 100 µL aliquots and stored at -80°C until use.

For transformation, DNA was added to 100 µL thawed cells and the mixture was incubated on ice for 30 minutes. Cells were placed at 42°C for 30 seconds and placed back on ice for 2 minutes. 900 µL SOC recovery media (MP Biomedicals MP113031012) was added, and cells were incubated at 37°C while shaking at 225 RPM for 1hr. 100-300 µL of transformed cells were then plated on LB-agar plates with appropriate antibiotic selection and incubated at 37°C overnight. For DNA assemblies, 5 µL of assembly reaction was used for transformation. For all other transformations, 1µL purified plasmid was added.

Electrocompetent cells

Overnight cultures of *E. coli* S17-1 λpir were diluted at 1:100 into Erlenmeyer flasks containing fresh LB broth and incubated for ~3 hours until they reached OD_{600} of 0.6-1.0. The cells were incubated on ice for 30 minutes. Cells were harvested by centrifugation for 15 minutes at 3500 RPM in a 4°C rotor, supernatant was carefully removed, and cells were resuspended gently in ice-cold sterile water. This process was repeated a second time. Cells were centrifuged, supernatant was removed, and cells were resuspended in ice-cold 10% glycerol solution. Cells were transferred into cold centrifuge tubes and incubated on ice for 30 minutes. Cells were centrifuged again, supernatant was removed, and cells were resuspended in ice-cold 10% glycerol solution. Cells were flash-frozen in liquid N_2 in 100 µL aliquots and stored at -80°C.

For transformation, up to 2.5 µL of purified plasmid or 1 µL of Golden Gate assembly reaction was added to 25 µL of thawed cells and the mixture was incubated on ice for 10 minutes. Cells and DNA were transferred to a cold 2 mm gap electroporation cuvette (BioRad 1652082) and electroporated using a BioRad Gene Pulser Xcell electroporation system at the following settings: Exponential decay pulse, 25 µF, 200 Ω, 2500 V. SOC recovery media was added immediately after electroporation to a volume of 1 mL, and transformed cells were incubated at 37°C for 1 hr, shaking at 225 RPM. 100-300 µL of transformation was plated onto selective LB-agar plates.

Construction genetic parts, plasmids, and strains

Repressible P_{BFP1E6} promoter libraries

Promoter libraries containing TetR, LacI, or PhlF binding sites were ordered as pools of single-stranded oligonucleotides (IDT oPools). All unique promoter sequences were flanked by universal sequences containing primer binding sites for amplification, BsaI recognition sites, and overhang sequences for assembly into the reporter vector. Promoter parts were prepared for Golden Gate assembly by second strand synthesis: briefly, single stranded oligonucleotide libraries were resuspended in nuclease-free water (Fisher BP2484100) and annealed with a universal reverse primer in 1X Phusion HF PCR buffer (NEB B0518). Annealed oligonucleotides were then mixed with Phusion® High-Fidelity DNA Polymerase (NEB M0530) and dNTPs (NEB N0447) and extended at 72°C. Synthesized double-stranded DNA was purified in nuclease-free water using a QIAquick PCR purification kit (Qiagen 28106) following manufacturer protocols.

RBS parts

The strong *B. theta* RBS1P-LP sequence[3] was constructed as a Golden Gate assembly part using assembly PCR. Briefly, two 100 bp oligonucleotides with a 58°C Tm overlap were designed containing the RBS sequence flanked by BsaI recognition sites at each end. Two short outer primers (~58°C Tm) were designed against the 5' and 3' end of the longer oligos. Inner and outer primers were mixed at a 1:10 molar ratio and subjected to rounds of amplification following the manufacturer-recommended PCR protocol for Phusion® High-Fidelity DNA Polymerase. PCR products were purified in nuclease-free water using a QIAquick PCR Purification Kit.

Repressor parts vectors

TetR and LacI repressor coding sequences without BsaI recognition sites were ordered as synthetic gBlocks (IDT). A Golden Gate-compatible PhlF gene was obtained (pR-PhlF[70]; Addgene 49367). Golden Gate parts vectors containing these repressors under *B. theta* promoters were assembled using HiFi DNA assembly using NEBuilder® HiFi DNA Assembly Master Mix (NEB E2621S). Briefly, a ColE1 plasmid backbone[71], *B. theta* promoter (PBfP2E5[3]), RBS (RBS1P-LP[3]), and repressor genes with a terminator were amplified from existing plasmids using Phusion DNA polymerase and assembled using HiFi DNA assembly. BsaI recognition sites were inserted upstream and downstream of the *B. theta* promoter and terminator, respectively, such that the whole expression cassette can be cut out and assembled into a new plasmid as part of a Golden Gate Assembly. After cloning, parts vectors were purified by miniprep (Qiagen 27106) and verified by Oxford Nanopore sequencing (Plasmidsaurus).

Golden Gate-compatible pNBU2 entry vector for repressible promoter screening

Repressor expression cassettes, promoters, and RBS parts described above were assembled into the pNBU2 backbone, which was transferred from *E. coli* to *B. theta* during conjugation and subsequently integrated into the *B. theta* genome in single copy at one of two attBT2 sites. To build a Golden Gate entry vector for osmolality promoters, an existing pNBU2 plasmid (pWW3536[3] [Genbank KY776518]) containing an *sfgfp* gene was adapted for our purposes. In this plasmid, the constitutive promoter driving *sfgfp* was replaced with an mCherry dropout cassette flanked by BsaI recognition sites. The *mCherry* gene and its terminator were PCR amplified from the pWW3515 plasmid[3] (Genbank KY776515), with the forward primer containing an overhang coding a strong *E. coli* promoter (J23100) and RBS[71]. Overhangs also contained BsaI recognition sites. The pNBU2 backbone containing the *sfgfp* gene and its terminator was amplified from the pWW3536 plasmid[3], with primers containing overhangs that overlap the amplified mCherry cassette. PCR products were purified and assembled using HiFi DNA assembly to build an mCherry dropout entry vector, into which parts are assembled into pNBU2 as follows: repressor cassette:BsaI-flanked promoter entry site:RB:GFP:pNBU2 backbone (pEM124; Figure 1B), and colonies containing assembled and background plasmids can be easily distinguished.

Promoter library construction in E. coli

Repressors, cognate promoters, and RBS1P-LP were inserted into the pEM124 pNBU2-GFP backbone via BsaI Golden Gate assembly. Assemblies were transformed into *E. coli* OneShot™ PIR1 cells (Invitrogen

C101010). ~4000 colonies were scraped, resuspended into 10 mL selective LB broth, and stored as glycerol stocks in 200 µL aliquots. To transfer the plasmid library to *E. coli* S17-1 λpir cells for conjugations, one 200 µL aliquot was thawed on ice, diluted 1:100 into selective LB broth, grown for 4 hours at 37°C at 225 RPM, and miniprepped. Purified plasmids were transformed into electrocompetent S17-1 *λ*pir cells and plated on selective LB-agar plates. ~4000 colonies were scraped into 5 mL of selective LB broth and stored at -80°C in 200 µL aliquots (pEM252) to be used for conjugation.

Construction of Golden Gate-compatible biosensor entry vectors

The inverted transcriptional circuit was assembled into the pNBU2 backbone using HiFi assembly as follows: $P_{PBFP1E6}$:RBS:mScarlet-I3:BsaI-flanked entry site:TetR:P_{TetR56}:RBS:mGL:pNBU2 backbone (pEM526; Figure 3A). pDx_mScarlet-I3 was a gift from Dorus Gadella (Addgene plasmid 189757). The mGL gene was ordered as a gBlock (Twist Biosciences) and PCR amplified prior to HiFi assembly. Assemblies were transformed into TSS competent *E. coli* S17-1 λpir cells.

Construction of biosensor circuits

TetR sponging cassettes were adapted for use with Golden Gate assembly into our circuit. Briefly, plasmids containing 10, 20, 40, 80, or 160 TetO binding sites (Addgene 160825-160829) were digested with XbaI (NEB R0145) and XhoI (NEB R0146) according to manufacturer-recommended protocols. The p15A ori backbone of these plasmids was PCR amplified with primers overlapping sponge cassettes, and containing BsaI sites matching the pEM526 entry vector at the upstream end and the native promoter at the downstream end. Sponge cassettes were assembled into these modified backbones with HiFi assembly to make Golden Gate-ready TetR sponging cassettes (EM413, EM414, EM415, EM351, EM352) with overhangs matching the pEM526 entry vector (5'-CCGA-3') at the 5' end and a native promoter part (5'-ATCA-3') at the 3' end. Native promoters and RBSs were amplified from the *B. theta* genome (~500 bp upstream of the CDS), with primers containing BsaI sites and overhangs matching either the pEM526 entry vector (5'-CCGA-3') or sponge cassettes (5'-TGAT-3') at the 5' end and TetR (5'-ACAT-3') at the 3' end. Promoter and TetR sponge parts were assembled into pEM526 via Golden Gate assembly as follows:

10 µL reactions contained:
0.33 µL BsaI HF®v2 (NEB R3733G)
0.33 µL T4 DNA Ligase (NEB M0202M)
1 µL 10x T4 DNA Ligase Buffer (NEB B0202S)
25 fmol Entry Vector Plasmid
50 fmol each additional part
To volume with Nuclease Free water (Millipore Sigma 7732-18-5)

Reactions were cycled as follows:
- 37ºC for 10 minutes
- 25x:
  - 37ºC for 1.5 minutes
  - 16ºC for 3 minutes
- 37ºC for 5 minutes
- 80°C for 10 minutes
- 12°C hold

All assemblies were transformed into TSS competent *E. coli* S17-1 λpir cells. Final biosensor plasmids were conjugated into *B. theta* as below.

B. theta conjugation and integration

Wild type *B. theta* was grown overnight in BHIS broth in anaerobic conditions, then sub-cultured in 5 mL fresh BHIS at 1:50 dilution. *E. coli* S17-1 λpir containing assembled pNBU2 plasmids were grown

overnight in LB-Amp, then sub-cultured in 5 mL fresh LB-Amp. Subcultures for both strains were grown ~4 hours, centrifuged at 4000 RPM for 10 minutes and supernatant was removed. *E. coli* and *B. theta* pellets were serially resuspended in 1 mL of BHIS. 100 µL of resuspension was spot-plated onto BHI+10% Horse Blood (Dalynn Biologicals HH30-500) agar plates and incubated at 37°C aerobically for 14-16 hr. Plates were incubated at 37°C anaerobically for another 14-16 hours. The conjugation mixture was spread onto selective BHIS agar plates with gentamicin and erythromycin and grown anaerobically for two nights. Single colonies were restreaked onto BHIS plates with gentamicin and erythromycin and grown anaerobically for two nights. The genomic insertion site of the plasmid was verified in single colonies by colony PCR (*PCR verification of integration site*). Verified single colonies were picked into liquid BHIS + Erm + Gen and grown overnight. Liquid cultures were mixed at 1:1 ratio with 50% glycerol for long-term storage.

In the case of promoter library conjugations, all single colonies from the first selective BHIS plates (4 plates with ~1000 colonies/plate) were scraped into 10 mL liquid BHIS-Erm-Gen and incubated for 4 hours to select against any remaining donor *E. coli*. This liquid culture was mixed at 1:1 ratio with 50% glycerol. To determine the number of colony-forming units in the glycerol stocks, an aliquot of this mixture was diluted at 1:100, 1:1000, and 1:10000 ratios in BHIS, 100 µL of each dilution was spread onto BHIS-Erm plates and grown overnight, and colonies were counted. The remaining glycerol-culture solution was stored at -80°C as 100 µL aliquots.

PCR verification of integration site

Briefly, a set of primers was designed flanking the attBT2-1 and attBT2-2 insertion sites, with one primer in the pNBU2 backbone such that two bands of specific sizes are produced according to insertion site (Table 1; Supplementary Figure 3E). Colony PCRs were performed with DreamTaq PCR Master Mix (ThermoFisher K1081) according to manufacturer recommended protocols with a total of 1 µM of primers. In the case that biosensor colonies resulting from conjugations produced WT amplicons at both attBT2 sites, integration at the native promoter site were verified using two sets of primers targeting the 5' and 3' end of the plasmid and the flanking genomic regions (Table 2; Supplementary Figure 3E).

Site	***B. theta* F primer (5' - 3')**	***B. theta* R Primer (5' - 3')**	**pNBU2 R Primer (5' - 3')**	**Size with *B. theta* R Primer if WT**	**Size with pNBU2 R if insertion**
attBT2-2	CATGTCTTTATACTTATTATAGGCTGC	GGAATGGATATTTTTATGATGCGG	CCGACTTTACATTCTTGGGTG	691 bp	289 bp
attBT2-1	CGGAGTGATGTTATGACTTACG	CCTCTATTTCCAGTTCACGTTG		911 bp	478 bp

Table 1. Verification PCR for genomic insertion of the pNBU2 plasmid at the canonical attBT2 sites. A combination of 5 primers allow identification of the pNBU2 plasmid insertion site based on amplicon length.

Site	***B. theta* Primer (5' - 3')**	**pNBU2 primer (5' - 3)**	**Size**
opuA 5'	GCAAACGGAGCGAACAATACC	CGCTGTTAATCACTTTACTTTTATCTAATCTGG	800 bp

opuA 3'	CAAACGGTTGATGCAGCG	GTTTGTCGGTGAACGCTCTC	700 bp
BT0779 5'	TCATTAATCTGCCGGTAGCGG	CGCTGTTAATCACTTTACTTTTATCTAATCTGG	900 bp
BT0779 3'	CCTGTGGAGCTTGTGTCTGAG	GTTTGTCGGTGAACGCTCTC	700 bp

Table 2. Verification PCR for genomic insertion of the biosensor plasmids at native promoter sites.

Genomic sequencing and analysis of engineered B. theta

B. theta variants of interest were struck from glycerol stocks onto selective BHIS-Erm plates and grown overnight. Single colonies were picked into fresh TYG and grown overnight. Genomic DNA was extracted using a DNeasy PowerSoil Pro Kit (Qiagen 47014). 1 µg of purified DNA was sequenced using Oxford Nanopore sequencing (Plasmidsaurus). Genome assembly and annotation was performed by the sequencing company (https://plasmidsaurus.com/results_interpretation#bact-assembly).

***In vitro* fluorescence measurements**

Promoter library induction screens and orthogonality testing

Library glycerol stocks were thawed, diluted in fresh selective BHIS-Erm and grown 6 hours at 37ºC anaerobically. Cultures were plated on BHIS-Erm plates to obtain single colonies. 189 single colonies from each library were picked and grown overnight in BHIS-Erm in 96 deep-well plates (Fisher 236600). Single colonies of a strain containing the parent P_{BfP1E6} promoter driving GFP and of the base *B. theta* strain were included in each library as controls. Overnight cultures were diluted at 1:50 into fresh BHIS containing either no inducer or small molecule inducer at the following concentrations: 100 ng/µL anhydrotetracycline hydrochloride (aTc, ThermoScientific J66688.MB); 1000 µM isopropyl β-D-1-thiogalactopyranoside (IPTG, Invitrogen 15529019); or 5 µM 2,4-diacytlphloroglucinol (DAPG, Toronto Research Chemicals D365500). Cultures were grown until they reached an OD_{600} of ~0.3-0.7 (~4-6 hours), removed from anaerobic conditions, harvested by centrifugation for 10 minutes at 4000 RPM, and resuspended twice in PBS (Fisher BP3991). After washing, cells were incubated at room temperature until they had been in aerobic conditions for a total of 2 hours. 200 µL of resuspended cells were transferred to a black, clear bottom 96-well plate (Grenier Bio-One 675096), and OD_{600} and GFP levels (Ex/Em: 485/515 nm; Gain 100) were measured using a Synergy H1 plate reader (Biotek). Normalized GFP was measured as GFP/OD_{600}. For orthogonality tests (Fig 1D), strains from each library were grown as above in base media and each small molecule inducer at the levels specified above. Fluorescence was measured in a plate reader, as above.

B. theta *promoter dose-response curves*

Library variants of interest were struck from glycerol stocks onto selective BHIS-Erm plates and grown overnight. Single colonies were picked into fresh BHIS-Erm and grown overnight. Cultures were diluted 1:50 in 1 mL fresh BHIS supplemented with a small-molecule inducer. Small molecule inducer concentrations were as follows: 2-fold serial dilutions starting at 100 ng/mL aTc; 2-fold serial dilutions starting at 1 mM for IPTG; and 2-fold serial dilutions starting at 5µM for DAPG. The first concentration of all small-molecule inducers shown in Figure 1C is 0 (plain BHIS). Cultures were grown, and GFP was measured as above. Normalized GFP was averaged over three biological replicates.

Biosensor screening and dose-response curves

Biosensor variants of interest were struck from glycerol stocks onto selective BHIS-Erm plates and grown overnight. Single colonies were picked into fresh TYG and grown overnight. Cultures were diluted 1:50 in

1 mL of fresh TYG adjusted to specified osmolalities PEG (Bayer® MiraLAX®). To account for slower growth rates under osmotic stress, cultures at different osmolalities were diluted at staggered time intervals (approximately 10 minutes per 100 mOsm/kg). When appropriate, media was also supplemented with aTc at 0, 0.25, 2.5, or 25 ng/mL. Cultures were grown, and GFP (mGreenLantern; Ex/Em: 485/515 nm; Gain: 90) and RFP (mScarlet-I3; Ex/Em: 570/600 nm; Gain 120) levels were measured. GFP and RFP were averaged over three biological replicates (three independent colonies). GFP and RFP levels were normalized to OD_{600} and to WT levels and changes in the normalized GFP:RFP ratio were measured across osmolalities.

***In vitro* growth curves**

B. theta *growth in osmotic stress conditions in vitro*

Biosensor variants of interest were struck from glycerol stocks onto selective BHIS-Erm plates and grown overnight. Single colonies were picked into fresh TYG and grown overnight. Cultures were diluted 1:10 in 1 mL of fresh TYG and allowed to grow for 4 hours to reach mid exponential phase. Cultures were then diluted again at 1:100 into 200 µL of TYG adjusted to specified osmolalities with PEG (Bayer® MiraLAX®) in an optically clear 96-well plate (Greiner Bio-one 655161) and sealed with a clear seal (Bio-Rad 17010701). Cultures were then grown with continuous shaking at 37 ºC in Biotek H1 synergy plate reader and OD_{600} was measured every 5 minutes for 20 hours.

Growth curve analysis and fitting

Individual OD_{600} curves were fitted with a Gompertz curve using a previously published in-house MATLAB script[36]. Maximum growth rate and OD_{600} were extracted using this code, and plots displaying extracted fits were created with a custom MATLAB script.

***B. theta* transcriptomics**

B. theta *RNA collection from osmotic stress conditions* in vitro

B. theta was struck from glycerol stocks onto BHIS plates and grown overnight. Single colonies were inoculated into pre-reduced TYG and grown overnight. Overnight cultures were then subcultured 1:50 into base TYG medium or TYG supplemented with either NaCl (Fisher S271), PEG (Bayer® MiraLAX®), lactulose (MilliporeSigma 42758450GM), or sorbitol (ACROS Organics AC132730010) to achieve osmolalities of approximately 700 and 900 mOsm/kg, which were confirmed using an Advanced Instruments Osmo1 Single-sample Micro-osmometer. Cultures were incubated until mid-exponential phase, after which samples were collected and preserved with RNAprotect (Qiagen 76506) as per manufacturer instructions. Samples were stored at -80°C until extractions were performed.

B. theta *growth in general stress conditions in vitro*

B. theta was struck from glycerol stocks onto BHIS plates and grown overnight. Single colonies were inoculated into pre-reduced TYG and grown overnight. Overnight cultures were then subcultured 1:50 into the following conditions in TYG: baseline (pH 7.2); temperature stress in baseline media grown at 42 ºC in a 6-well plate; acid stress adjusted to pH 5.8 with 1M HCl (Fisher, SA49); oxidative stress adjusted to 100 µM H_2O_2 (Millipore 216763) by spiking 1M H_2O_2 in at a 1:100 dilution 2 and 4 hours after subculturing to account for the short half-life of H_2O_2 in media. Cultures were incubated at either 37 ºC or 42 ºC until mid-exponential phase, after which samples were collected and preserved with RNAprotect (Qiagen) as per manufacturer instructions. Samples were stored at -80 °C until extractions were done.

In vivo B. theta *transcriptomics during malabsorption*

4-5 week Swiss-Webster germ-free mice were randomly assigned to co-housed cages of the same sex, with a minimum of 3 mice per group. Mice were orally gavaged with 150 µL of overnight cultures of *B. theta* and allowed to equilibrate for 4 weeks. After equilibration, two cages (n=6) of mice were euthanized, and

their cecal contents were collected to quantify bacterial abundance, measure cecal osmolality, and extract RNA. The remaining two cages (n=6) received drinking water supplemented with 15% (w/v) polyethylene glycol (3350, Miralax®) at a final osmolality of 255-265 mOsm/kg for 6 days, as previously described[26]. On day 6 of PEG treatment, mice were euthanized, and their cecal contents were collected. Cecal samples designated for RNA extractions were flash-frozen in dry ice and stored at -80 °C until used.

RNA extraction, library preparation and sequencing of in vitro *cultures and cecal contents*

In vitro and *in vivo* samples stored at -80 °C were thawed on ice, and their RNA was extracted using the RNeasy Mini Kit (Qiagen 74104) or RNeasy PowerMicrobiome Kit (Qiagen 26000-50), respectively, as per manufacturer instructions. RNA samples were depleted of gDNA using the TURBO DNA-free™ Kit (Invitrogen AM1907). RNA quantity and quality were verified by Qubit quantification and Agilent BioAnalyzer prior to NEBNext Ultra II RNA library preparation. Sequencing was performed on the Illumina NextSeq2000 sequencer using a P3 flow cell and 50 bp single-end reads.

Transcriptomics analysis

Transcriptomic sequencing analysis was performed using a modified script of the SAMSA2 pipeline[72], described below. Raw reads were checked for sequencing quality using FastQC91. Following, reads were trimmed using Trimmomatic[73] and ribosomal reads were removed from the data using SortMeRNA[74]. An indexed database was generated for *B. thetaiotaomicron* VPI-5482 using the complete reference genome (GCF_000011065.1) and raw reads were aligned to this database using DIAMOND[75]. The resulting mapped reads were analyzed on DESeq2[30] to perform differential expression (DE) analysis. Differential gene expression was performed using *B. theta* samples grown in base TYG or samples from untreated mice as the reference condition.

Mouse Experiments and *in vivo B. theta* fluorescence measurements

0, 10, and 15% PEG dosage mouse experiments (Figures 4A, 5A)

3-5 week-old Swiss-Webster germ-free mice were randomly assigned to co-housed cages of the same sex. Mice were orally gavaged with 150 µL of overnight cultures of *B. theta* biosensor strains P_{opuA}+20xTetO or P_{BT0779}+160 x tetO, or the No Promoter control strain. Three cages were inoculated per strain (n = 3-5 mice per cage). After a 2-week equilibration period, mice were started on either 0%, 10%, or 15% (w/v) PEG (Bayer® MiraLAX®) in drinking water, at final osmolalities of 0, 104, and 216 mOsm/kg, respectively, for 6 days. Upon completing 6 days, mice were euthanized, and their cecal contents were collected and used for osmolality measurements and flow cytometry analysis.

0-7.5% PEG dosage mouse experiment (Figure 5D)

11-14 week-old Swiss-Webster germ-free mice were randomly assigned to co-housed cages of the same sex, with 3-5 mice per cage. All mice were orally gavaged with 150 µL of overnight cultures of P_{BT0779}+160xTetO. After a 2-week equilibration period, mice were started on either 0%, 1%, 2.5%, 5%, or 7.5% (w/v) PEG (Bayer® MiraLAX®) in drinking water at final osmolalities of 0, 5, 14, 35, and 66 mOsm/kg, respectively, for 3 days. On Day 3, fecal pellets were collected from each mouse and used for flow cytometry and Bristol Stool Scoring[76]. The same fecal pellets were used to measure fecal water content as previously described. Briefly, microfuge tubes were weighed using an analytical balance before fecal pellets were added[77]. The combined weight was recorded, and the samples were dried overnight at 60°C with tube lids open. The final tube weight was then recorded, and fecal water content was calculated as: (fecal weight before drying − fecal weight after drying)/feces weight before drying × 100. On the same day as fecal sample collection, mice were euthanized, and their cecal contents were collected and used for osmolality measurements and flow cytometry analysis.

15 and 20% PEG dosage mouse experiments (Figure 4G)

14-week-old female Swiss-Webster germ-free mice were co-housed in the same cage. Mice (n=5) were orally gavaged with 150 µL of overnight cultures of *B. theta* biosensor strains P_{opuA}+20xTetO^{HR}. After a 1-week equilibration period, mice were started on 20% (w/v) PEG (Bayer® MiraLAX®) in drinking water, at final osmolality of 379 mOsm/kg, for 6 days. Fecal samples were collected on day 0, 2, 4, and 6 of the treatment and processed for flow cytometry. After 6 days of treatment, mice were euthanized, and their cecal contents were collected and used for osmolality measurements and flow cytometry analysis.

Bacterial abundance measurements

Bacterial colony-forming units (CFU) were quantified by sampling cecal contents with 1 µL loops and resuspending in 200 µL of PBS. Ten-fold serial dilutions of the resuspended samples were performed in PBS, and 5 µL of each dilution was spot-plated in BHIS agar. Plates were incubated until single colonies were countable (~24 h).

Cecal osmolality measurements

Cecal contents collected when mice were euthanized were centrifuged at 16,000 x g for 20 minutes at 4°C immediately after collection. The resulting cecal supernatant was used to measure cecal osmolality using an Advanced Instruments Model 3220 osmometer.

Isolation of live bacteria from fecal and cecal contents

Live bacteria isolation protocol was adapted from a previous protocol[78]. One fourth of fecal pellets (~10-20 mg), or ~20-50 mg of cecal contents, were placed in 2 mL microfuge tubes 1-2 h after collection. One stainless steel bead (5.0 mm) was added to each tube together with 1 mL of PBS. The samples were thoroughly homogenized using TissueLyser II (Qiagen) for 10 minutes at 20 Hz. The homogenized samples were then subjected to three iterations of the following steps: vortexed for 15 seconds; centrifuged at 1,000 rpm for 30 seconds at room temperature; 750 µL of supernatant transferred into a fresh 2 mL microcentrifuge tube pre-filled with 750 µL PBS. After the final iteration, tubes were centrifuged at 4,000 x g for 5 minutes at room temperature, the supernatant was discarded, and pellets were resuspended in 1 ml of PBS. Resuspended samples were then diluted 1:10 in PBS in preparation for flow cytometry. Isolations were performed aerobically to allow for fluorescent proteins to mature.

Flow cytometry protocol and analysis

Isolated *B. theta* biosensors from fresh fecal or cecal samples were analyzed on a CytoFLEX LX flow cytometer using the following excitation laser and emission filter combinations for: mGL (ex: 488 nm, em: 525 nm) and mScarlet-I3 (ex: 488 nm, em: 610 nm). Samples were diluted 1:10 or 1:100 in PBS for optimal detection of fluorescence. Single-fluorophore *B. theta* controls (mGL$^+$ mScarlet$^-$ and mGL$^-$ mScarlet$^+$) were grown in TYG, and 1 mL samples were collected at mid-exponential phase (OD_{600} ~0.5), pelleted at 5,000 x g for 5 minutes, resuspended in 1 mL of PBS, and mixed with germ-free fecal sample prepared as explained above. These controls were used to establish gating parameters that distinguish bacterial fluorescence from background fecal fluorescence. Based on the mGL-mScarlet+ gate (mScarlet fluorescence > 4x10 AU), 10,000 events within this gate were recorded for each sample. All samples were recorded at a low-speed flow rate (10 µL/minute). FCS files were imported into FlowJo (v10.10.0), a cell gate was set based on the events that had a high mScarlet$^+$ signal, and then a gate was set at $5x10^3$ A.U. separating mScarlet$^+$ peak from the background fecal or cecal fluorescence. All events within the mScarlet$^+$ gate for each sample were exported, and their fluorescence output values were analyzed and visualized with R using tidyverse (v2.0.0) and ggridges (v0.5.6) packages.

Biosensor Mathematical Model and Parameter Fitting

Model construction

The following ODE-based mathematical models representing 1) P_{TetR56} induction with aTc and 2) biosensor induction with osmolality were constructed in MATLAB (MathWorks) and solved using the ode23.m function.

$$\begin{aligned} \frac{dm_1}{dt} &= \frac{Vmax_1 \cdot A^{h1}}{{km_1}^{h1} + A^{h1}} + \lambda_1 - m_1(\beta_{dil} + \beta_{deg1}) \\ \frac{dP_1}{dt} &= K_1 \cdot m_1 - P_1 \cdot \beta_{dil} \end{aligned} \tag{1}$$

where

$Vmax_1$ is the max transcription rate from P_{TetR56}
A is the concentration of aTc
$h1$ is the Hill coeff. of P_{TetR56}
km_1 is the km of P_{TetR56}
λ_1 is the leaky/baseline expression rate from P_{TetR56}
β_{deg1} is the GFP mRNA depletion rate due to degradation
β_{dil} is the dilution rate due to cell growth

$$
\begin{aligned}
\frac{dm_2}{dt} &= \frac{Vmax_2 \cdot I^{h2}}{{km_2}^{h2} + I^{h2}} + \lambda_1 - m_2(\beta_{dil} + \beta_{deg2}) \\
\frac{dP_2}{dt} &= K_2 \cdot m_2 - P_2 \cdot \beta_{dil} - \frac{S \cdot {P_2}^{hs}}{{km_s}^{hs} + {P_2}^{hs}} \\
\frac{dm_1}{dt} &= \frac{Vmax_1 \cdot {km_1}^{h1}}{{P_2}^{h1} + {km_1}^{h1}} + \lambda_1 - m_1(\beta dil + \beta deg1) \\
\frac{dP_1}{dt} &= K_1 \cdot m_1 - P_1 \cdot \beta_{dil}
\end{aligned}
\tag{2}
$$

where

$Vmax_2$ is the max transcription rate from native promoter
I is the concentration of inducer
$h2$ is the Hill coeff. for the native promoter
km_2 is the km for native promoter
λ_2 is the leaky/baseline expression rate from native promoter
β_{deg2} is the TetR mRNA depletion rate due to degradation
β_{dil} is the dilution rate due to cell growth
K_2 is the translation rate of TetR per mRNA
S is the total number of TetR sponges
km_s is the km for TetR sponging
hs is the Hill coeff. for TetR sponging
$Vmax_1$ is the max transcription rate from P_{TetR56}
km_1 is the km for P_{TetR56}
$h1$ is the Hill coeff. for P_{TetR56}
λ_1 is the leaky/baseline expression rate from P_{TetR56}
β_{deg1} is the GFP mRNA depletion rate due to degradation
K_1 is the translation rate of GFP per mRNA

Hill dynamics were assumed for repressible transcription from the P_{TetR56} and inducible transcription from native *B. theta* promoters. Sponging of TetR follows Hill ligand binding dynamics, with the degree of sponging varying with TetR concentration and number of total binding sites included (*S*). As TetR dimerizes before binding operator sites, the TetR sponging term is multiplied by 2. We assumed that no active degradation of TetR or GFP occurs, as these proteins were not tagged for degradation in our circuit.

The Covariance Matrix Adaptation-Evolution Strategy (CMA-ES) algorithm[79] was employed to fit parameter values for the above models that reflect system behaviour for P_{TetR56} and both P_{opuA} and P_{BT0779} biosensors. Here, CMA-ES works by stochastically exploring the spaces around a given set of estimated parameter values. At each set of parameter values, a cost is calculated as a function of distance between model results with the current parameter set and a user-provided experimental dataset. Over a defined number of iterations, CMA-ES tries to minimize the calculated cost, homing in on a parameter set that represents the closest fit to the experimental data provided.

In this case, initial parameter values were estimated based on biologically relevant transcription and translation rates from *E. coli*[38,80,81], and CMA-ES calculated cost based on the distance between predicted steady-state GFP levels at a given parameter set and experimental GFP measurements from *B. theta* dose response curves. Specifically, CMA-ES (cmaes.m[82]) was first run on Model 1 using the aTc dose response curve (Figure 1D) to fit the following parameters: $\lambda_1, Vmax_1, Km_1, Beta_{deg}, K_1$. Using these estimates as initial input parameters for Model 2, CMA-ES was then run on Model 2 without sponges (S = 0), using

osmolality induction curves from either P_{opuA} or P_{BT0779} biosensors as input data to fit the remaining parameters. In both cases CMA-ES was run 25 times for 500 iterations each. The parameter set with the lowest cost from all runs combined was saved as the optimized parameter set to use for modelling the system with sponging. Initial estimated and optimized parameter values for both biosensors are listed in Supplementary Table 5. Using the parameter sets estimated by CMA-ES for P_{opuA} and P_{BT0779}, Model 2 was then solved over a range of induction levels (I) and repressor sponging (S). Steady-state GFP values from each solution were used to calculate fold change in GFP across conditions.

Statistical Analyses

Statistical analyses performed on given data sets are indicated in the corresponding figure legends. For all figures: * *padj* < 0.05; ** *padj* < 0.01; *** *padj* < 0.001.

Acknowledgments

The authors acknowledge that the land we performed this research on is the traditional, ancestral, and unceded territory of the xwməθkwəy̓əm (Musqueam) Nation. The land it is situated on has always been a place of learning for the Musqueam people, who for millennia have passed on in their culture, history, and traditions from one generation to the next on this site. We encourage others to learn more about the native lands in which they live and work at https://native-land.ca/. The authors acknowledge support from the following: CIFAR / Humans and the Microbiome (FL-001253 Appt 3362, to C.T.); Michael Smith Foundation for Health Research Scholar Award (18239, to C.T.); Canada Foundation for Innovation / Infrastructure Operating Fund (38277); Canada Tier 2 Research Chair, Quantitative Microbiota Biology for Health Applications (CRC-2022-00036, to C.T.); Weston Family Foundation / Weston Family Microbiome Initiative (F21-05234, to C.T., L.P.T.); NSERC Discovery grant (RGPIN-2019-04591, to C.T.); CIFAR Azrieli Global Scholars (GS20-011, to C.T.); NSERC-CREATE SynBioApps Fellowship (to G.M.); NSERC Canada Graduate Scholarships (to G.M.). This work received support from the Centre of Disease Modelling, especially Natalia Caranza, and from ubcFLOW. This research was supported in part through computational resources and services provided by Advanced Research Computing (ARC) at the University of British Columbia. The authors also thank Michael Hunter for critically reading this manuscript and providing feedback, as well as Claire Sie and Apsara Srinivas for their help with animal work.

References

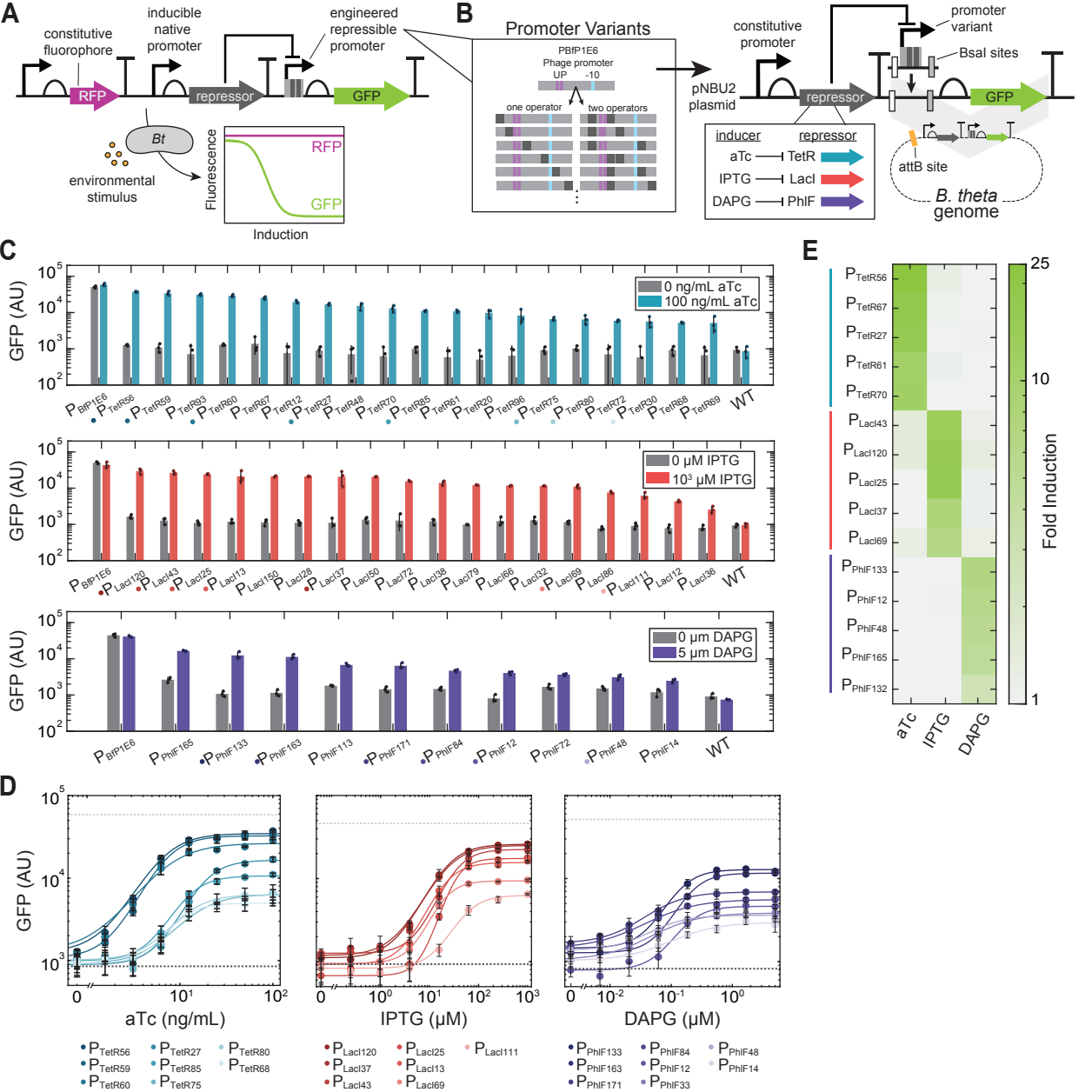
1. Tropini, C., Earle, K. A., Huang, K. C. & Sonnenburg, J. L. The Gut Microbiome: Connecting Spatial Organization to Function. *Cell Host Microbe* **21**, 433–442 (2017).
2. McCallum, G. & Tropini, C. The gut microbiota and its biogeography. *Nat. Rev. Microbiol.* **22**, 105–118 (2024).
3. Whitaker, W. R., Shepherd, E. S. & Sonnenburg, J. L. Tunable Expression Tools Enable Single-Cell Strain Distinction in the Gut Microbiome. *Cell* **169**, 538-546.e12 (2017).
4. 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 (2018).
5. Mondragón-Palomino, O. *et al.* Three-dimensional imaging for the quantification of spatial patterns in microbiota of the intestinal mucosa. *Proc. Natl. Acad. Sci.* **119**, (2022).
6. Parker, A. C. & Jeffrey Smith, C. Development of an IPTG inducible expression vector adapted for Bacteroides fragilis. *Plasmid* **68**, 86–92 (2012).
7. Mimee, 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 Syst.* **1**, 62–71 (2015).
8. 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.e15 (2017).
9. Yang, X. *et al.* A Nitrate/Nitrite Biosensor Designed with an Antiterminator for In Vivo Diagnosis of Colitis Based on Bacteroides thetaiotaomicron. *ACS Synth. Biol.* (2025)

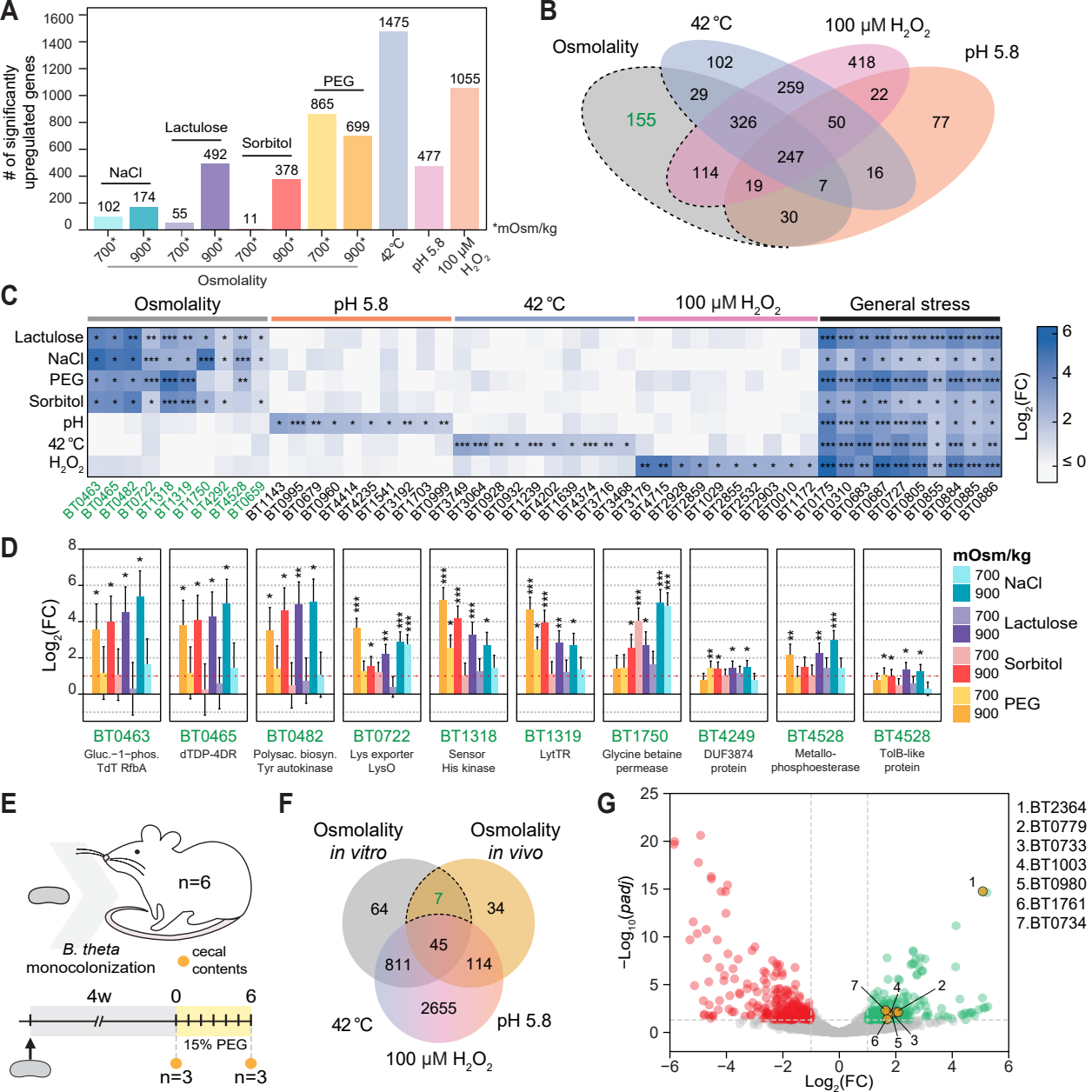
doi:10.1021/acssynbio.4c00602.
10. Whitaker, W. R. *et al.* Controlled colonization of the human gut with a genetically engineered microbial medicine. 2024.10.03.24314621 Preprint at https://doi.org/10.1101/2024.10.03.24314621 (2024).
11. Sarfatis, A., Wang, Y., Twumasi-Ankrah, N. & Moffitt, J. R. Highly multiplexed spatial transcriptomics in bacteria. *Science* **387**, eadr0932 (2025).
12. Wan, X., Pinto, F., Yu, L. & Wang, B. Synthetic protein-binding DNA sponge as a tool to tune gene expression and mitigate protein toxicity. *Nat. Commun.* **11**, (2020).
13. Nolan, J. D., Johnston, I. M. & Walters, J. R. F. Physiology of malabsorption. *Surg. Oxf.* **30**, 268–274 (2012).
14. Owens, S. R. & Greenson, J. K. The pathology of malabsorption: current concepts. *Histopathology* **50**, 64–82 (2007).
15. Wu, T.-T. Malabsorption and Malnutrition Disorders. in *Surgical Pathology of Non-neoplastic Gastrointestinal Diseases* (eds. Zhang, L., Chandan, V. S. & Wu, T.-T.) 191–238 (Springer International Publishing, Cham, 2019). doi:10.1007/978-3-030-15573-5_9.
16. Misselwitz, B. *et al.* Lactose malabsorption and intolerance: pathogenesis, diagnosis and treatment. *United Eur. Gastroenterol. J.* **1**, 151–159 (2013).
17. Jones, H. F., Butler, R. N. & Brooks, D. A. Intestinal fructose transport and malabsorption in humans. *Am. J. Physiol.-Gastrointest. Liver Physiol.* **300**, G202–G206 (2011).
18. Rosenberg, I., Solomons, N. & Schneider, R. Malabsorption associated with diarrhea and intestinal infections. *Am. J. Clin. Nutr.* **30**, 1248–1253 (1977).
19. Viswanathan, V. K., Hodges, K. & Hecht, G. Enteric infection meets intestinal function: how bacterial pathogens cause diarrhoea. *Nat. Rev. Microbiol.* **7**, 110–119 (2009).
20. Pasotti, L., Politi, N., Zucca, S., Angelis, M. G. C. D. & Magni, P. Bottom-Up Engineering of Biological Systems through Standard Bricks: A Modularity Study on Basic Parts and Devices. *PLOS ONE* **7**, e39407 (2012).
21. Gong, X. *et al.* Architecting a transcriptional repressor-based genetic inverter for tryptophan derived pathway regulation in *Escherichia coli*. *Metab. Eng.* **86**, 66–77 (2024).
22. Gadella, T. W. J. *et al.* mScarlet3: a brilliant and fast-maturing red fluorescent protein. *Nat. Methods* **20**, 541–545 (2023).
23. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS ONE* **3**, (2008).
24. 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. *J. Bacteriol.* **182**, 3559–3571 (2000).
25. Schnider-Keel, U. *et al.* Autoinduction of 2,4-Diacetylphloroglucinol Biosynthesis in the Biocontrol Agent Pseudomonas fluorescensCHA0 and Repression by the Bacterial Metabolites Salicylate and Pyoluteorin. *J. Bacteriol.* **182**, 1215–1225 (2000).
26. Tropini, C. *et al.* Transient Osmotic Perturbation Causes Long-Term Alteration to the Gut Microbiota. *Cell* **173**, 1742—1754.e17 (2018).
27. Burckhardt, J. C., Chong, D. H. Y., Pett, N. & Tropini, C. Gut commensal Enterocloster species host inoviruses that are secreted in vitro and in vivo. *Microbiome* **11**, 1–16 (2023).
28. Wrzosek, L. *et al.* Transplantation of human microbiota into conventional mice durably reshapes the gut microbiota. *Sci. Rep.* **8**, 6854 (2018).
29. Clayton, C. A. *et al.* In mouse and gut-on-a-chip models, pre-colonoscopy bowel preparation promotes pathogen colonization of the gut and translocation to other organs. Preprint at https://doi.org/10.1101/2024.11.26.625504 (2024).
30. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
31. Campbell, B. C. *et al.* mGreenLantern: a bright monomeric fluorescent protein with rapid expression and cell filling properties for neuronal imaging. *Proc. Natl. Acad. Sci.* **117**, 30710–30721 (2020).

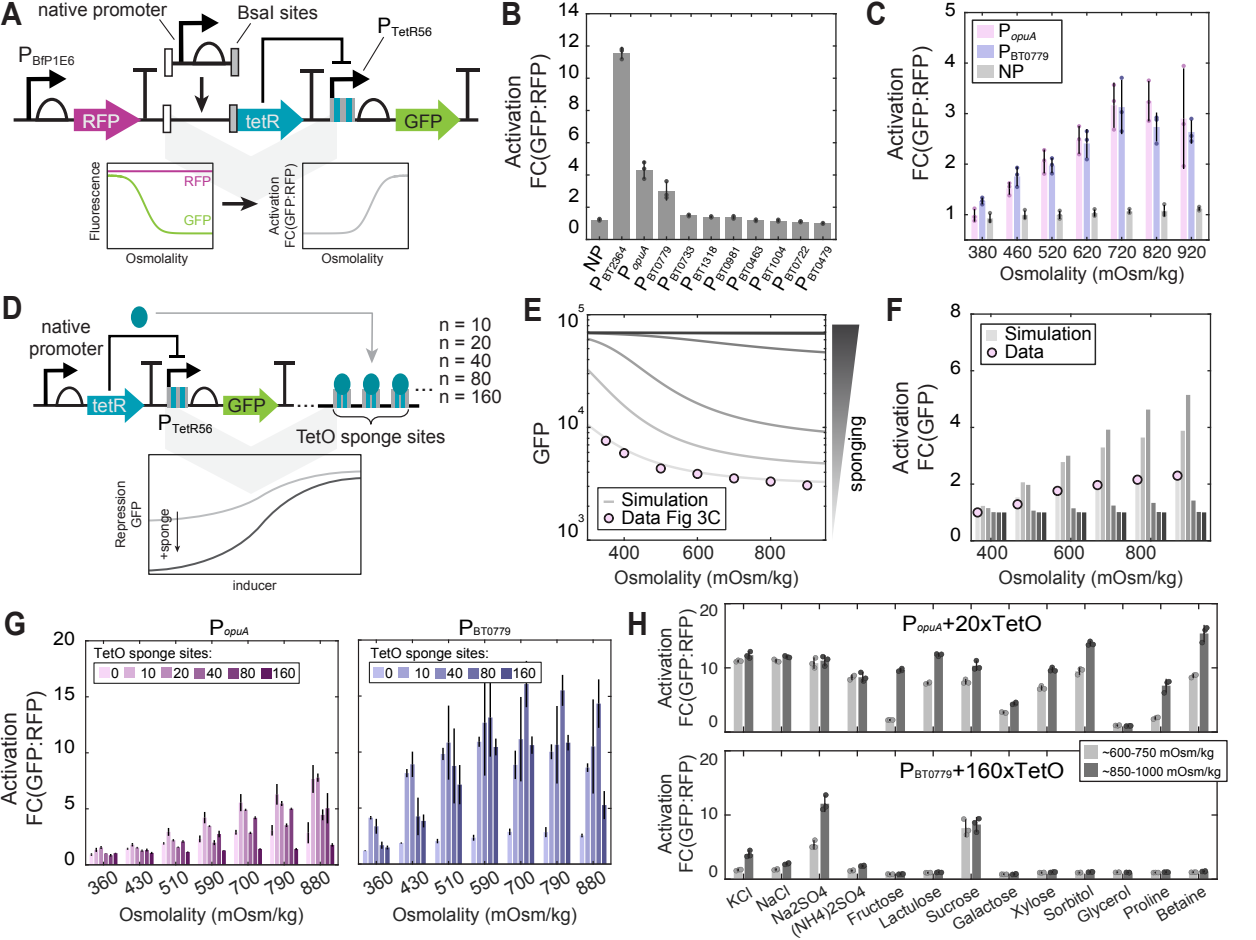
32. Campbell, B. C., Paez-Segala, M. G., Looger, L. L., Petsko, G. A. & Liu, C. F. Chemically stable fluorescent proteins for advanced microscopy. *Nat. Methods* **19**, 1612–1621 (2022).
33. Sonnenburg, J. L., Chen, C. T. L. & Gordon, J. I. Genomic and Metabolic Studies of the Impact of Probiotics on a Model Gut Symbiont and Host. *PLOS Biol.* **4**, e413 (2006).
34. Hoffmann, T. & Bremer, E. Guardians in a stressful world: the Opu family of compatible solute transporters from *Bacillus subtilis*. *Biol. Chem.* **398**, 193–214 (2017).
35. Bremer, E. & Krämer, R. Responses of Microorganisms to Osmotic Stress. *Annu. Rev. Microbiol.* **73**, 313–334 (2019).
36. Ng, K. M. *et al.* Single-strain behavior predicts responses to environmental pH and osmolality in the gut microbiota. *mBio* **0**, e00753-23 (2023).
37. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
38. Potvin-Trottier, L., Lord, N. D., Vinnicombe, G. & Paulsson, J. Synchronous long-term oscillations in a synthetic gene circuit. *Nature* **538**, 514–517 (2016).
39. Cray, J. A., Russell, J. T., Timson, D. J., Singhal, R. S. & Hallsworth, J. E. A universal measure of chaotropicity and kosmotropicity. *Environ. Microbiol.* **15**, 287–296 (2013).
40. P. Gregory, K. *et al.* Understanding specific ion effects and the Hofmeister series. *Phys. Chem. Chem. Phys.* **24**, 12682–12718 (2022).
41. Englaender, J. A. *et al.* Effect of Genomic Integration Location on Heterologous Protein Expression and Metabolic Engineering in E. coli. *ACS Synth. Biol.* **6**, 710–720 (2017).
42. Scholz, S. A. *et al.* High-Resolution Mapping of the Escherichia coli Chromosome Reveals Positions of High and Low Transcription. *Cell Syst.* **8**, 212-225.e9 (2019).
43. Sauer, C. *et al.* Effect of Genome Position on Heterologous Gene Expression in Bacillus subtilis: An Unbiased Analysis. *ACS Synth. Biol.* **5**, 942–947 (2016).
44. Inda, M. E. & Lu, T. K. Microbes as Biosensors. *Annu. Rev. Microbiol.* **74**, 337–359 (2020).
45. Tanna, T., Ramachanderan, R. & Platt, R. J. Engineered bacteria to report gut function: technologies and implementation. *Curr. Opin. Microbiol.* **59**, 24–33 (2021).
46. Riglar, D. T. *et al.* Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. *Nat. Biotechnol.* **35**, 653–658 (2017).
47. Weibel, N. *et al.* Engineering a Novel Probiotic Toolkit in Escherichia coli Nissle 1917 for Sensing and Mitigating Gut Inflammatory Diseases. *ACS Synth. Biol.* **13**, 2376–2390 (2024).
48. Chien, T. *et al.* Enhancing the tropism of bacteria via genetically programmed biosensors. *Nat. Biomed. Eng.* **6**, 94–104 (2022).
49. Gurbatri, C. R. *et al.* Engineering tumor-colonizing E. coli Nissle 1917 for detection and treatment of colorectal neoplasia. *Nat. Commun.* **15**, 646 (2024).
50. Sonnenborn, U. & Schulze, J. The non-pathogenic Escherichia coli strain Nissle 1917 – features of a versatile probiotic. *Microb. Ecol. Health Dis.* **21**, 122–158 (2009).
51. Martinson, J. N. V. & Walk, S. T. Escherichia coli Residency in the Gut of Healthy Human Adults. *EcoSal Plus* **9**, 10.1128/ecosalplus.ESP-0003–2020 (2020).
52. Xu, J. *et al.* A genomic view of the human-Bacteroides thetaiotaomicron symbiosis. *Science* **299**, 2074–2076 (2003).
53. Eckburg, P. B. *et al.* Diversity of the Human Intestinal Microbial Flora. *Science* **308**, 1635–1638 (2005).
54. Vaga, S. *et al.* Compositional and functional differences of the mucosal microbiota along the intestine of healthy individuals. *Sci. Rep.* **10**, 1–12 (2020).
55. Tabula Sapiens Consortium* *et al.* The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. *Science* **376**, eabl4896 (2022).
56. Ma, P. *et al.* Bacterial droplet-based single-cell RNA-seq reveals antibiotic-associated heterogeneous cellular states. *Cell* **186**, 877-891.e14 (2023).
57. Xu, Z. *et al.* Droplet-based high-throughput single microbe RNA sequencing by smRandom-seq. *Nat. Commun.* **14**, 5130 (2023).

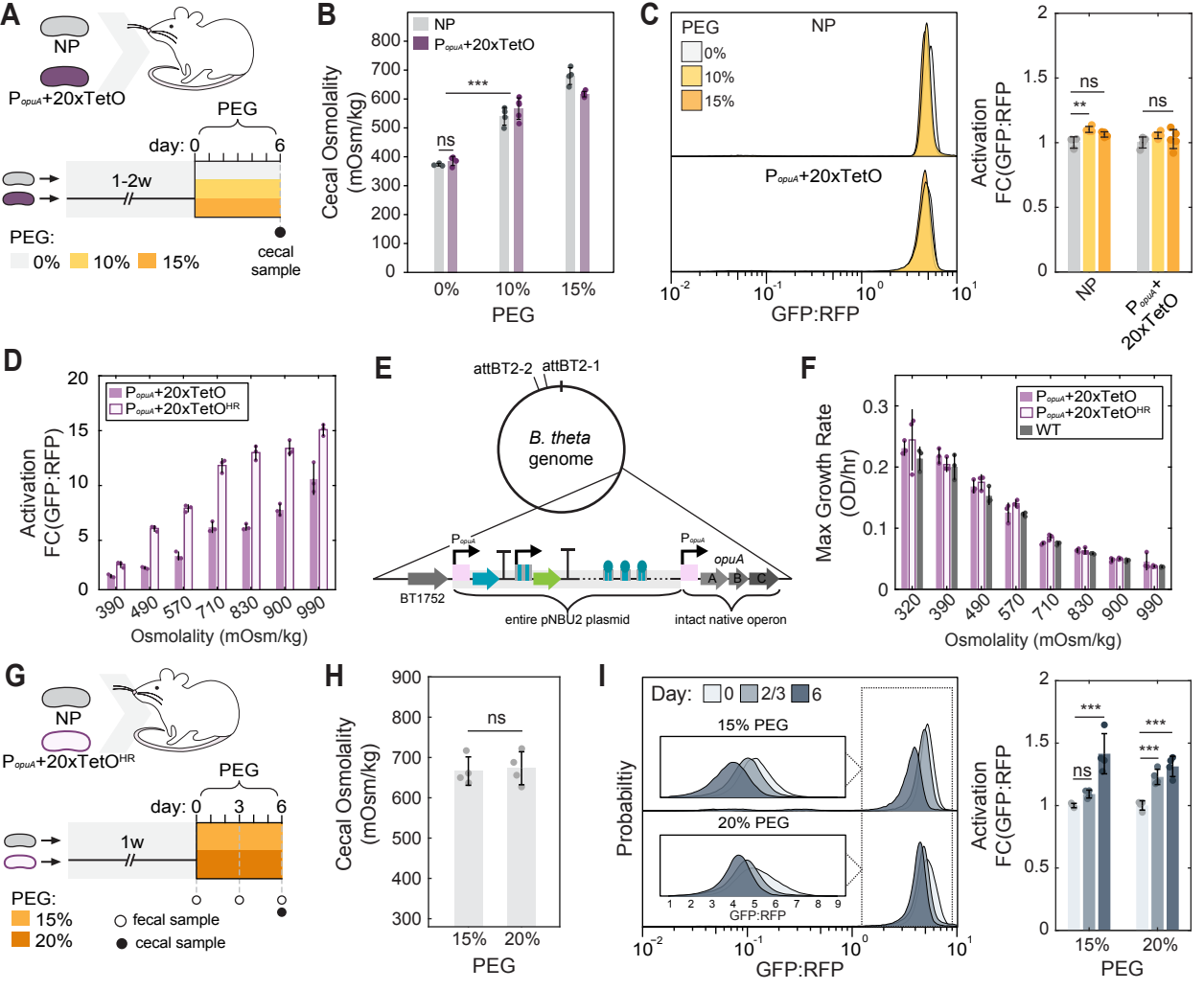
58. Blattman, S. B., Jiang, W., Oikonomou, P. & Tavazoie, S. Prokaryotic single-cell RNA sequencing by in situ combinatorial indexing. *Nat. Microbiol.* **5**, 1192–1201 (2020).
59. Kuchina, A. *et al.* Microbial single-cell RNA sequencing by split-pool barcoding. *Science* **371**, (2021).
60. Gaisser, K. D. *et al.* High-throughput single-cell transcriptomics of bacteria using combinatorial barcoding. *Nat. Protoc.* **19**, 3048–3084 (2024).
61. Tanna, T., Schmidt, F., Cherepkova, M. Y., Okoniewski, M. & Platt, R. J. Recording transcriptional histories using Record-seq. *Nat. Protoc.* **15**, 513–539 (2020).
62. Schmidt, F. *et al.* Noninvasive assessment of gut function using transcriptional recording sentinel cells. *Science* **376**, (2022).
63. Kalia, N. *et al.* Intestinal secretory and absorptive function in Trichinella spiralis mouse model of postinfective gut dysfunction: role of bile acids. *Gut* **57**, 41–49 (2008).
64. Ledwaba, S. E. *et al.* Enteropathogenic Escherichia coli Infection Induces Diarrhea, Intestinal Damage, Metabolic Alterations, and Increased Intestinal Permeability in a Murine Model. *Front. Cell. Infect. Microbiol.* **10**, (2020).
65. Lindfors, K. *et al.* Coeliac disease. *Nat. Rev. Dis. Primer* **5**, 1–18 (2019).
66. Liu, J., Qiao, B., Cai, Y., Tan, Z. & Deng, N. Diarrhea accompanies intestinal inflammation and intestinal mucosal microbiota dysbiosis during fatigue combined with a high-fat diet. *BMC Microbiol.* **23**, 151 (2023).
67. Yang, C. & Merlin, D. Unveiling Colitis: A Journey through the Dextran Sodium Sulfate-induced Model. *Inflamm. Bowel Dis.* **30**, 844–853 (2024).
68. Kennedy, R. J. *et al.* Interleukin 10-deficient colitis: new similarities to human inflammatory bowel disease. *J. Br. Surg.* **87**, 1346–1351 (2000).
69. Chung, C. T., Niemela, S. L. & Miller, R. H. One-step preparation of competent Escherichia coli: Transformation and storage of bacterial cells in the same solution. *Proc Natl Acad Sci USA* **86**, 2172–2175 (1989).
70. Stanton, B. C. *et al.* Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nat. Chem. Biol.* **10**, 99–105 (2014).
71. Moore, S. J. *et al.* EcoFlex: A Multifunctional MoClo Kit for E. coli Synthetic Biology. *ACS Synth Biol* **5**, 1059–1069 (2016).
72. SAMSA2: a standalone metatranscriptome analysis pipeline | BMC Bioinformatics | Full Text. https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-018-2189-z.
73. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
74. Kopylova, E., Noé, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**, 3211–3217 (2012).
75. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
76. Riegler, G. & Esposito, I. Bristol scale stool form. A still valid help in medical practice and clinical research. *Tech. Coloproctology* **5**, 163–164 (2001).
77. Jeong, D. *et al.* Modulation of gut microbiota and increase in fecal water content in mice induced by administration of Lactobacillus kefiranofaciens DN1. *Food Funct.* **8**, 680–686 (2017).
78. Ronda, C., Chen, S. P., Cabral, V., Yaung, S. J. & Wang, H. H. Metagenomic engineering of the mammalian gut microbiome in situ. *Nat. Methods* **16**, 167–170 (2019).
79. Hansen, N. & Ostermeier, A. Completely Derandomized Self-Adaptation in Evolution Strategies. *Evol. Comput.* **9**, 159–195 (2001).
80. McCallum, G. & Potvin-Trottier, L. Using Models to (Re-)Design Synthetic Circuits. in *Methods in Molecular Biology* vol. 2229 91–118 (Humana Press Inc., 2021).
81. Milo, R., Jorgensen, P., Moran, U., Weber, G. & Springer, M. BioNumbers--the database of key numbers in molecular and cell biology. *Nucleic Acids Res.* **38**, D750-3 (2010).
82. Ros, R. & Hansen, N. A Simple Modification in CMA-ES Achieving Linear Time and Space

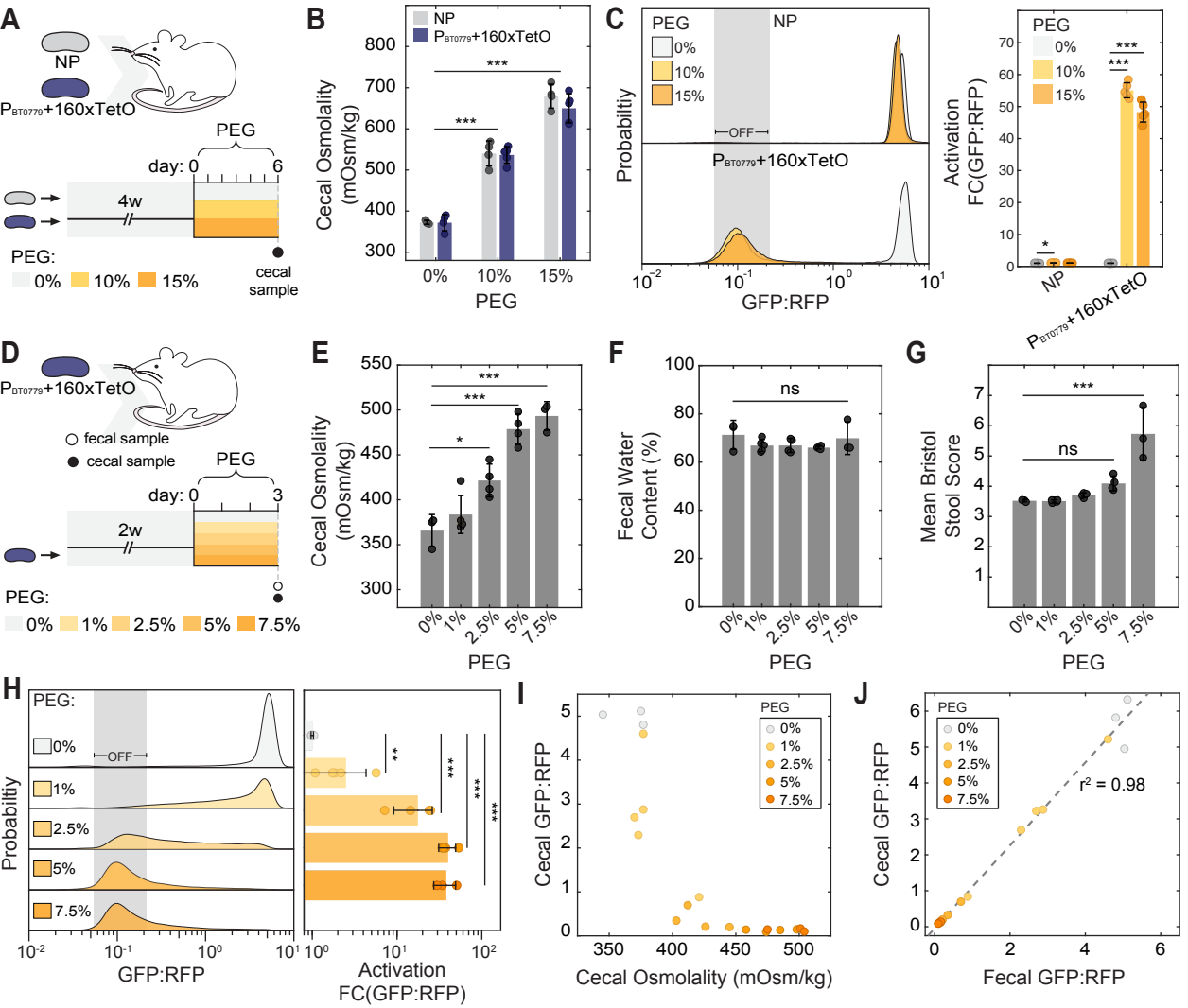
Complexity. in *Parallel Problem Solving from Nature – PPSN X* (eds. Rudolph, G., Jansen, T., Beume, N., Lucas, S. & Poloni, C.) 296–305 (Springer, Berlin, Heidelberg, 2008). doi:10.1007/978-3-540-87700-4_30.

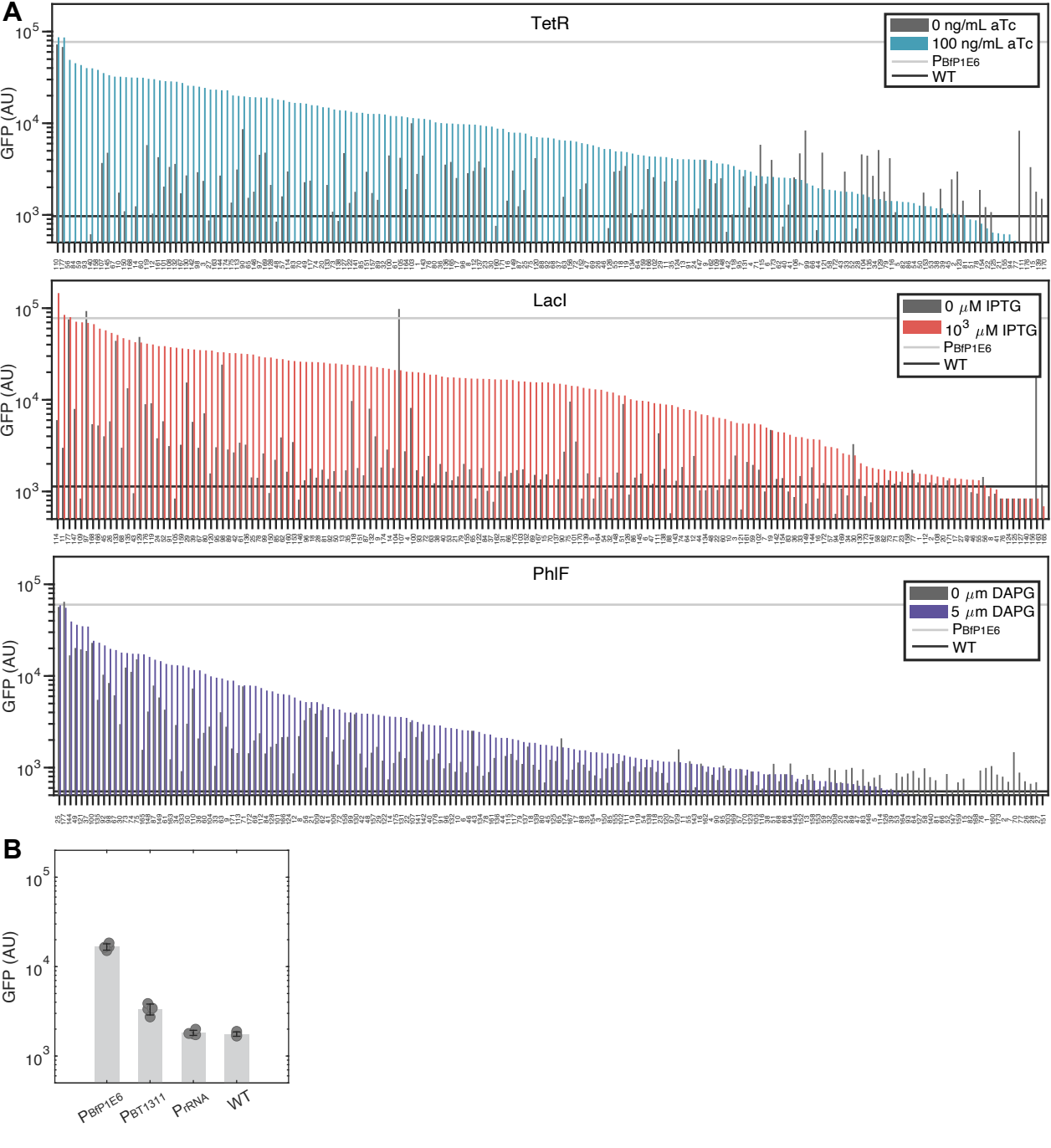

A
constitutive fluorophore
inducible native promoter
engineered repressible promoter
RFP
repressor
GFP
Bt
environmental stimulus
Fluorescence
RFP
GFP
Induction
B
Promoter Variants
PBfP1E6
Phage promoter
UP
-10
one operator
two operators
pNBU2 plasmid
constitutive promoter
repressor
GFP
promoter variant
BsaI sites
attB site
B. theta genome
inducer
repressor
aTc
TetR
IPTG
LacI
DAPG
PhlF
C
GFP (AU)
0 ng/mL aTc
100 ng/mL aTc
0 μM IPTG
10³ μM IPTG
0 μm DAPG
5 μm DAPG
WT
D
GFP (AU)
aTc (ng/mL)
IPTG (μM)
DAPG (μM)
E
Fold Induction
aTc
IPTG
DAPG

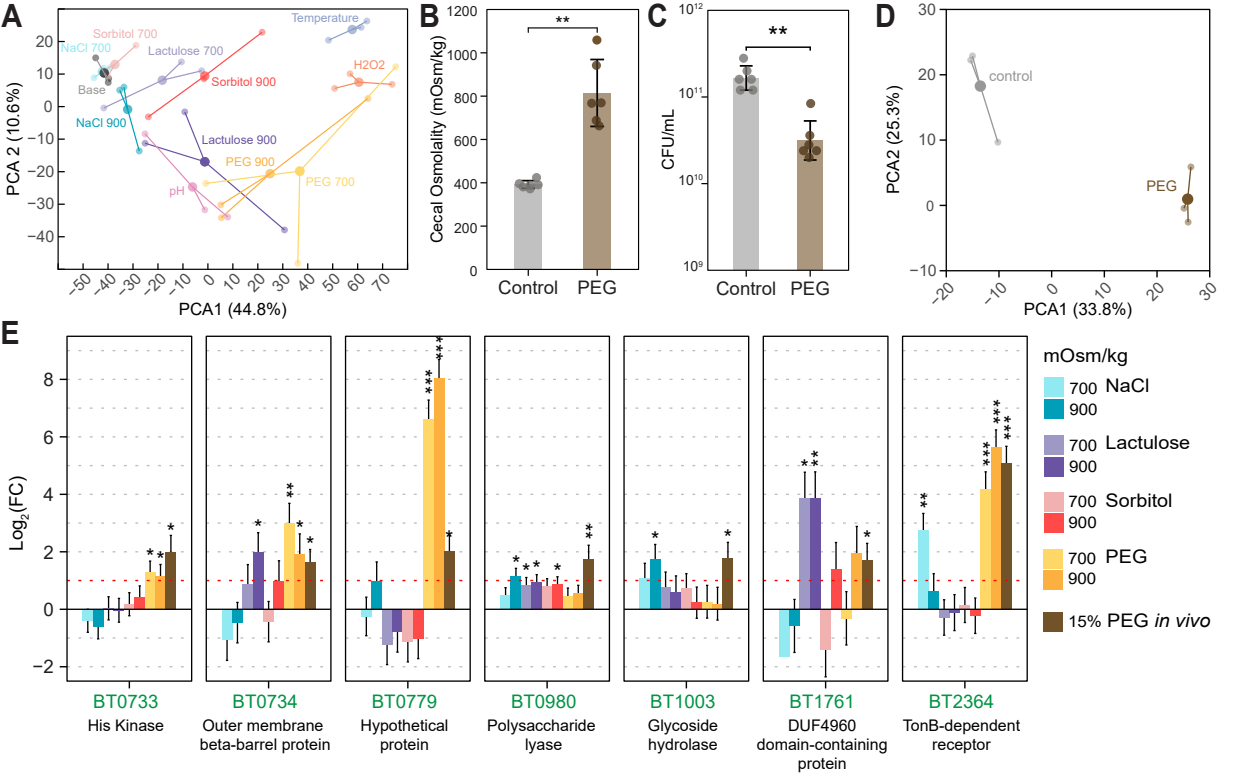

A
# of significantly upregulated genes
NaCl
Lactulose
Sorbitol
PEG
102
174
55
492
11
378
865
699
1475
477
1055
700*
900*
42°C
pH 5.8
100 μM H2O2
Osmolality
*mOsm/kg
B
Osmolality
42 °C
100 μM H2O2
pH 5.8
155
102
418
29
259
22
326
50
77
247
114
16
19
7
30
C
Osmolality
pH 5.8
42 °C
100 μM H2O2
General stress
Lactulose
NaCl
PEG
Sorbitol
pH
42 °C
H2O2
Log2(FC)
D
Log2(FC)
mOsm/kg
700
900
NaCl
Lactulose
Sorbitol
PEG
BT0463
Gluc.−1−phos. TdT RfbA
BT0465
dTDP-4DR
BT0482
Polysac. biosyn. Tyr autokinase
BT0722
Lys exporter LysO
BT1318
Sensor His kinase
BT1319
LytTR
BT1750
Glycine betaine permease
BT4249
DUF3874 protein
BT4528
Metallo-phosphoesterase
BT4528
TolB-like protein
E
n=6
B. theta monocolonization
cecal contents
4w
0
6
15% PEG
n=3
n=3
F
Osmolality in vitro
Osmolality in vivo
64
7
34
45
811
114
2655
42 °C
pH 5.8
100 μM H2O2
G
−Log10(padj)
Log2(FC)
1.BT2364
2.BT0779
3.BT0733
4.BT1003
5.BT0980
6.BT1761
7.BT0734

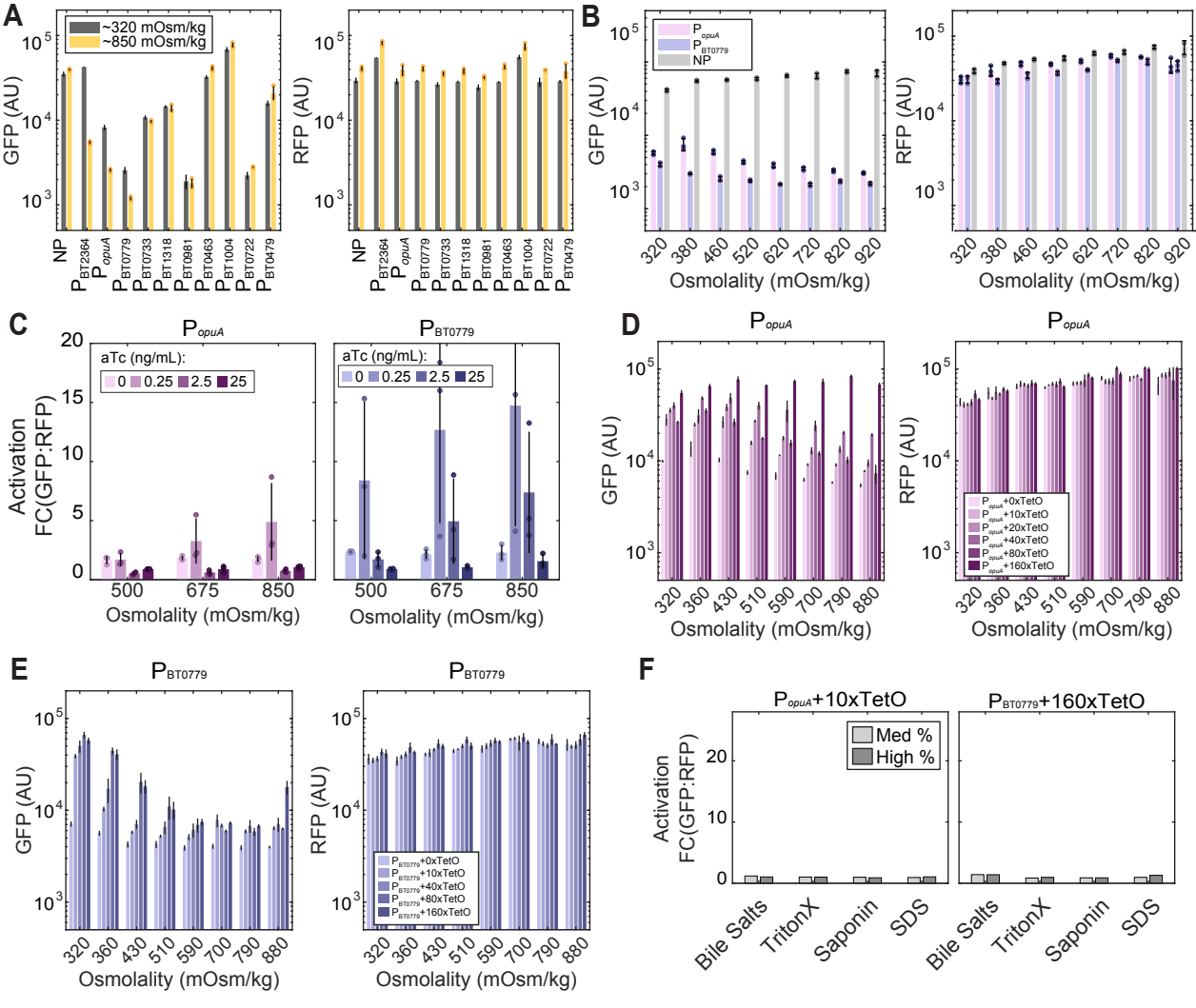

A
native promoter
BsaI sites
P_{BfP1E6}
P_{TetR56}
RFP
tetR
GFP
Fluorescence
RFP
GFP
Osmolality
Activation FC(GFP:RFP)
Osmolality
B
Activation FC(GFP:RFP)
NP
P_{BT2364}
P_{opuA}
P_{BT0779}
P_{BT0753}
P_{BT1318}
P_{BT0981}
P_{BT0453}
P_{BT1004}
P_{BT0722}
P_{BT0479}
C
P_{opuA}
P_{BT0779}
NP
Activation FC(GFP:RFP)
Osmolality (mOsm/kg)
D
native promoter
tetR
P_{TetR56}
GFP
n = 10
n = 20
n = 40
n = 80
n = 160
TetO sponge sites
Repression GFP
+sponge
inducer
E
GFP
sponging
Simulation
Data Fig 3C
Osmolality (mOsm/kg)
F
Simulation
Data
Activation FC(GFP)
Osmolality (mOsm/kg)
G
P_{opuA}
P_{BT0779}
TetO sponge sites:
0 10 20 40 80 160
Activation FC(GFP:RFP)
Osmolality (mOsm/kg)
H
P_{opuA}+20xTetO
P_{BT0779}+160xTetO
~600-750 mOsm/kg
~850-1000 mOsm/kg
Activation FC(GFP:RFP)
KCl
NaCl
Na2SO4
(NH4)2SO4
Fructose
Lactulose
Sucrose
Galactose
Xylose
Sorbitol
Glycerol
Proline
Betaine

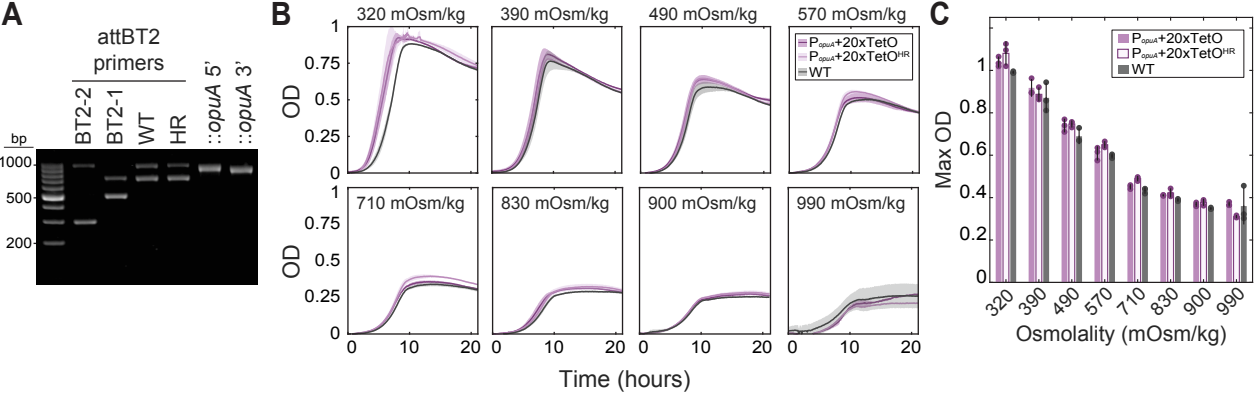

A
NP
PopuA+20xTetO
PEG
day: 0
6
1-2w
PEG:
0%
10%
15%
cecal sample
B
NP
PopuA+20xTetO
Cecal Osmolality (mOsm/kg)
***
ns
0%
10%
15%
PEG
C
PEG
0%
10%
15%
NP
PopuA+20xTetO
GFP:RFP
Activation FC(GFP:RFP)
ns
**
ns
NP
PopuA+ 20xTetO
D
PopuA+20xTetO
PopuA+20xTetOHR
Activation FC(GFP:RFP)
Osmolality (mOsm/kg)
E
attBT2-2
attBT2-1
B. theta genome
PopuA
opuA
A
B
C
BT1752
entire pNBU2 plasmid
intact native operon
F
PopuA+20xTetO
PopuA+20xTetOHR
WT
Max Growth Rate (OD/hr)
Osmolality (mOsm/kg)
G
NP
PopuA+20xTetOHR
PEG
day: 0
3
6
1w
PEG:
15%
20%
fecal sample
cecal sample
H
Cecal Osmolality (mOsm/kg)
ns
15%
20%
PEG
I
Day:
0
2/3
6
15% PEG
20% PEG
Probability
GFP:RFP
Activation FC(GFP:RFP)
***
ns
***
***
15%
20%
PEG

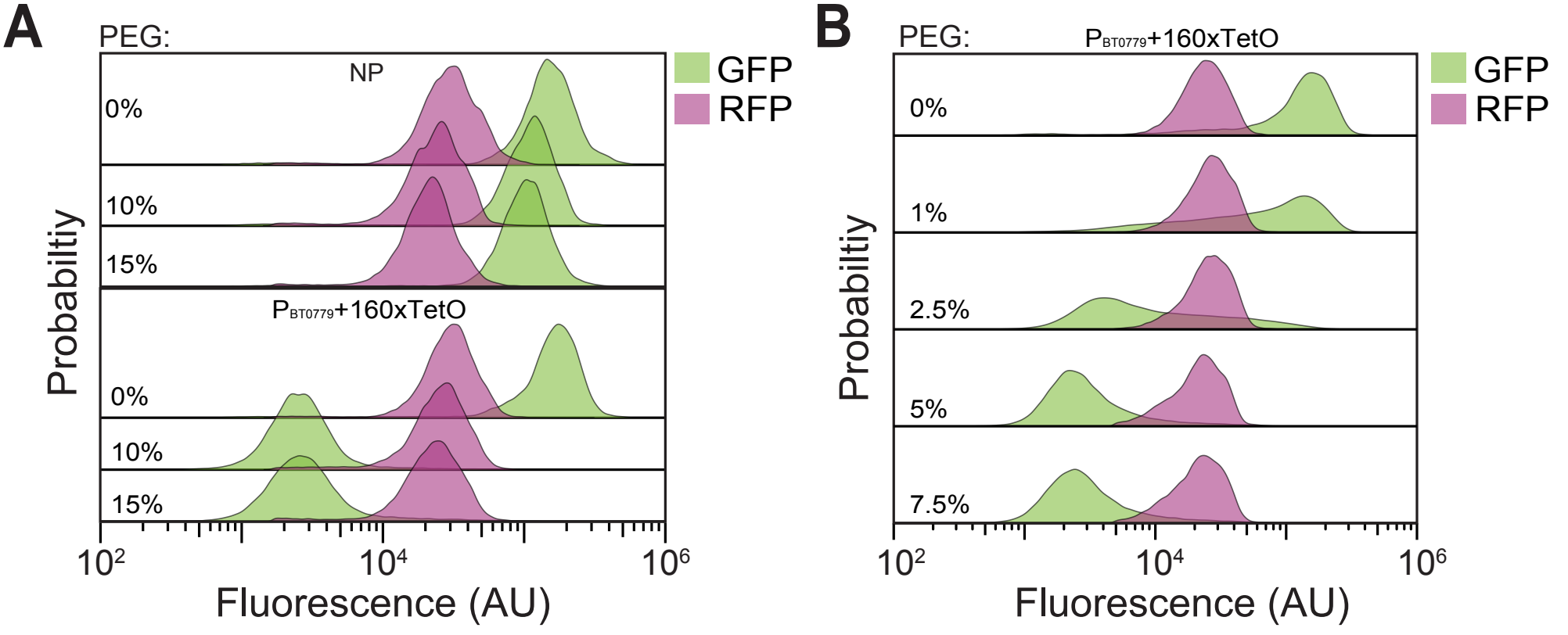

A
NP
PBT0779+160xTetO
PEG
day: 0
6
4w
PEG:
0%
10%
15%
cecal sample
B
NP
PBT0779+160xTetO
Cecal Osmolality (mOsm/kg)
PEG
C
PEG
0%
10%
15%
NP
OFF
PBT0779+160xTetO
Probability
GFP:RFP
Activation FC(GFP:RFP)
D
PBT0779+160xTetO
fecal sample
cecal sample
PEG
day: 0
3
2w
PEG:
0%
1%
2.5%
5%
7.5%
E
Cecal Osmolality (mOsm/kg)
F
Fecal Water Content (%)
ns
G
Mean Bristol Stool Score
H
PEG:
OFF
Probability
GFP:RFP
Activation FC(GFP:RFP)
I
Cecal GFP:RFP
Cecal Osmolality (mOsm/kg)
J
Cecal GFP:RFP
Fecal GFP:RFP
r² = 0.98


A
TetR
0 ng/mL aTc
100 ng/mL aTc
PBfP1E6
WT
GFP (AU)
LacI
0 μM IPTG
10^3 μM IPTG
PBfP1E6
WT
GFP (AU)
PhlF
0 μm DAPG
5 μm DAPG
PBfP1E6
WT
GFP (AU)
B
GFP (AU)
PBfP1E6
PBT1311
PfRNA
WT


A
Sorbitol 700
NaCl 700
Lactulose 700
Temperature
Base
Sorbitol 900
H2O2
NaCl 900
Lactulose 900
PEG 900
PEG 700
pH
PCA 2 (10.6%)
PCA1 (44.8%)
B
Cecal Osmolality (mOsm/kg)
Control
PEG
C
CFU/mL
Control
PEG
D
control
PEG
PCA2 (25.3%)
PCA1 (33.8%)
E
Log2(FC)
mOsm/kg
700 NaCl
900
700 Lactulose
900
700 Sorbitol
900
700 PEG
900
15% PEG in vivo
BT0733
His Kinase
BT0734
Outer membrane beta-barrel protein
BT0779
Hypothetical protein
BT0980
Polysaccharide lyase
BT1003
Glycoside hydrolase
BT1761
DUF4960 domain-containing protein
BT2364
TonB-dependent receptor


A
~320 mOsm/kg
~850 mOsm/kg
GFP (AU)
RFP (AU)
NP
P BT2364
P opuA
P BT0779
P BT0733
P BT1318
P BT0981
P BT0463
P BT1004
P BT0722
P BT0479
B
P opuA
P BT0779
NP
Osmolality (mOsm/kg)
C
P opuA
P BT0779
aTc (ng/mL):
0 0.25 2.5 25
Activation FC(GFP:RFP)
Osmolality (mOsm/kg)
D
P opuA
P opuA+0xTetO
P opuA+10xTetO
P opuA+20xTetO
P opuA+40xTetO
P opuA+80xTetO
P opuA+160xTetO
Osmolality (mOsm/kg)
E
P BT0779
P BT0779+0xTetO
P BT0779+10xTetO
P BT0779+40xTetO
P BT0779+80xTetO
P BT0779+160xTetO
Osmolality (mOsm/kg)
F
P opuA+10xTetO
P BT0779+160xTetO
Med %
High %
Activation FC(GFP:RFP)
Bile Salts
TritonX
Saponin
SDS


A
attBT2 primers
BT2-2
BT2-1
WT
HR
::opuA 5'
::opuA 3'
bp
1000
500
200
B
320 mOsm/kg
390 mOsm/kg
490 mOsm/kg
570 mOsm/kg
710 mOsm/kg
830 mOsm/kg
900 mOsm/kg
990 mOsm/kg
OD
Time (hours)
P opuA+20xTetO
P opuA+20xTetO HR
WT
C
Max OD
Osmolality (mOsm/kg)


A
PEG:
NP
0%
10%
15%
P_{BT0779}+160xTetO
0%
10%
15%
GFP
RFP
Probabiltiy
10^2
10^4
10^6
Fluorescence (AU)
B
PEG:
P_{BT0779}+160xTetO
0%
1%
2.5%
5%
7.5%
GFP
RFP
Probabiltiy
10^2
10^4
10^6
Fluorescence (AU)