ACS Synthetic Biology

pubs.acs.org/synthbio Letter

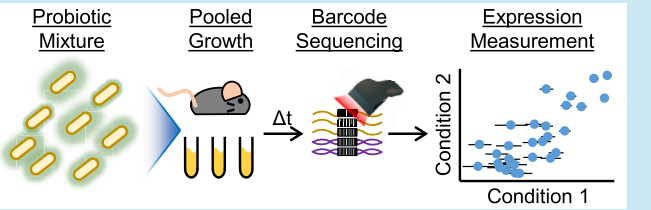
Transcript Barcoding Illuminates the Expression Level of Synthetic Constructs in *E. coli* Nissle Residing in the Mammalian Gut

Nathan Crook, Aura Ferreiro, Zevin Condiotte, and Gautam Dantas*

Cite This: https://dx.doi.org/10.1021/acssynbio.0c00040 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The development of robust engineered probiotic therapies demands accurate knowledge of genetic construct expression in the gut. However, the monetary and ethical costs of testing engineered strains in vertebrate hosts are incompatible with current high-throughput design-build-test cycles. To enable parallel measurement of multiple construct designs, we placed unique DNA barcodes in engineered transcripts and measured barcode abundances via sequencing. In standard curve experiments, the barcode sequences exhibited consistent relationships between input and measured abundances, which allowed us to use transcript barcoding to measure expression levels of 30 GFP-expressing strains of *E. coli* Nissle in parallel. Applying this technology in culture and in the mouse gut, we found that GFP expression in the gut could often be predicted from expression levels in culture, but several strains exhibited gut-specific expression. This work establishes the experimental design parameters and advantages of transcript barcoding to measure the performance of many engineered probiotic designs in mammalian hosts.



As the primary site of nutrient absorption, and a key modulator of host immunity, the gut plays a critical role in human health.[1] Central to this role are the trillions of microbes that inhabit the intestinal lumen, termed the gut microbiota.[2] In return for privileged access to unused dietary material, the microbiota perform a variety of useful functions including producing energy molecules and vitamins, competitively excluding pathogens, and immune signaling.[3,4] As a consequence, certain taxonomic or genetic compositions of the microbiota are associated with, and can often cause, disease states in humans including obesity,[5] inflammation,[6] susceptibility to infection,[7] diabetes,[8] and neurological disorders.[9] With these outcomes in mind, there has been a ballooning interest in altering the ecology and biochemistry of the gut for therapeutic purposes via small drug molecules,[10] fecal microbiota transplants,[11] prebiotics,[12] and probiotics.[13]

Probiotics are defined as live microorganisms that, when administered in adequate amounts, confer a health benefit. Probiotic microbes include lactic acid bacteria, *Escherichia coli* Nissle, and the yeast *Saccharomyces boulardii*. While the mechanisms by which probiotics confer health benefits differ,[14] probiotics share the property that they are generally recognized as safe for human consumption. Coupled with the genetic tractability of many probiotic strains, probiotics have recently been recognized as privileged chassis for genetic engineering to alter the biochemistry of the gut. These next-generation probiotics promise to augment or repair gut function via engineered behaviors including disease detection,[15] biomolecule synthesis,[16] toxin degradation,[17] or programmed cell death.[18] Probiotics are very interesting from a drug delivery perspective due to their massive scalability of manufacture, ease of distribution, and ability to bypass digestive enzymes and deliver their genetic payloads in a site-specific manner. Furthermore, probiotics often achieve high patient compliance[19] and do not permanently colonize,[20] enabling control over drug dosage.

The gut environment, however, imposes barriers in the design of engineered bacterial therapeutics. These include the substantial differences and variability in levels of nutrients, interspecies competition, and immune pressure in the gut compared to more controlled environments like the shake flask or bioreactor. As microbes are known to alter their physiology and metabolic activity in response to changing environmental conditions,[21] we hypothesized that the behavior of engineered probiotics, in particular the transcript abundance of engineered constructs, would be different in the gut than in axenic culture. Further, the expression of synthetic constructs driven by native promoters may be different than the expression of the corresponding native genes due to differences in genomic context or copy number. Precise tuning of gene expression and thus protein concentrations is critical to the efficacy of engineered microbial activity.[22] Taken together, this rationale motivated us to measure the expression of synthetic constructs in gut-resident probiotics.

Received: January 22, 2020
Published: April 23, 2020

Libraries of synthetic constructs typically comprise different expression levels, organization, or copy numbers of one or several genes. The high similarity of expressed sequences between strains precludes pooled expression analysis via RNA-seq. Instead, transcript levels of cloned genes are typically measured by growing multiple replicate cultures per construct, followed by RNA extraction and qPCR. However, for experiments involving many genes, or difficult-to-multiplex culture conditions, this approach becomes problematic. The gut is a good example of these limitations because, like in other applications of engineered strains, many strains must often be tested to find the optimal design. However, modeling the gut often requires complex *in vitro* setups[23] or vertebrate animal models, which are expensive and sometimes unethical to scale to match testing demand. For example, reporter proteins (such as LacZ or anaerobic fluorescent proteins) require a separate experimental group per construct tested. New methods are therefore required to measure the performance of many engineered microbes in nonculture habitats. In response to the challenge of testing many constructs, transcript barcoding has been used to measure the expression of multiple constructs in the same culture.[24−26] These approaches are high throughput and exhibit high precision.[24] However, to our knowledge these approaches have not been used to test gene expression of synthetic constructs in probiotics residing in the gut, and the accuracy of these approaches, as measured using synthetic barcode mixtures of known abundance, has not been tested.

Here we characterize the precision and accuracy of the transcript barcoding approach using synthetic barcode mixtures. We then use this technique to measure in parallel the expression activity of a model 30-member promoter library in the commonly used probiotic *E. coli* Nissle (*EcN*), with barcoded GFP as the carrier transcript. These activities were compared between traditional culture conditions and different sites in the mouse gastrointestinal tract. These experiments illuminated gene expression variability between individuals and different gut sites and establish a platform for future studies seeking to measure the expression of cloned genes in the mammalian gut.

MATERIALS AND METHODS

Strains, Media, Plasmids, and Primers. *EcN* was obtained as a kind gift from Dr. Phillip I. Tarr (Washington University in St. Louis School of Medicine) and authenticated via whole-genome sequencing. Strains were grown in either LB media or modified M9-MOPS media at 37 °C.[27] Agar was added at a concentration of 15 g/L for growth on solid media. For maintenance of the pZE21 plasmid[28] and its derivatives, kanamycin was added to a final concentration of 50 μg/mL after autoclaving. Aerobic growth was performed with shaking (250 rpm) on a Thermo Fisher MaxQ 2000 shaker. Anaerobic growth was performed in a vinyl anaerobic chamber (Type C, Coy Laboratory Products) without shaking.[29] All plasmids used in this study were constructed using Golden Gate Assembly MasterMix from NEB, according to the manufacturer's instructions. Plasmid sequences are provided as Genbank files in Supporting File 1. Upon construction, all plasmids were transformed into chemically competent *EcN*, thus generating biological triplicates. Primer sequences are listed in Supporting File 2.

Chemical Transformation of *EcN*. Chemically competent *EcN* was prepared by growing overnight in 100 mL of LB at 37 °C. This culture was diluted to OD600 0.1−0.2 using cold LB, after which it was grown for an additional 2−3 h at 37 °C, until the culture reached an OD600 of 0.3−0.5. This culture was centrifuged at 3000 × *g* for 5 min at 4 °C. The supernatant was discarded, and cells were resuspended in 1/10 of their original volume in ice cold 100 mM $CaCl_2$ dissolved in a 15% v/v glycerol/water mixture. This cell suspension was kept on ice for 30 min. This cell suspension was again centrifuged at 3000 × *g* for 5 min at 4 °C and resuspended to 1/50 of its original volume in the ice-cold $CaCl_2$/glycerol solution. Fifty-microliter aliquots of this mixture were frozen at −80 °C. To perform transformations, one aliquot of cells per transformation was thawed on ice. About 5 μL of ligation reaction or purified plasmid was added, and the resulting cell/DNA mixture was gently mixed without pipetting. These mixtures were then transferred to a PCR block prechilled to 4 °C. After 30 min, the temperature was raised to 42 °C for 5 min, after which the temperature was returned to 4 °C. After at least 3 min at 4 °C, cells were added to 250 μL of SOC medium in a centrifuge tube and incubated on their sides at 37 °C for 1 h with shaking (250 rpm). After this recovery step, cells were plated onto the appropriate selection media.

Transcriptomic Analysis of *EcN*. *EcN* containing pZE21 was grown overnight in LB media containing kanamycin (LB +Kan) under aerobic conditions. These cultures were used to inoculate triplicate 25 mL cultures in LB+Kan. Aerobic cultures were shaken, but anaerobic cultures were not shaken. At an optical density (600 nm) of 0.3, 5 mL of each culture was centrifuged (3000 × *g* for 5 min at room temperature), resuspended in 0.25 mL of water, placed in an equal volume of DNA/RNA shield, and stored at −80 °C. After overnight growth, 1 mL of each aerobic culture and 3 mL of each anaerobic culture were centrifuged (3000 × *g* for 5 min at room temperature), resuspended in 0.25 mL of water, and placed in an equal volume of DNA/RNA shield. All samples were processed with a ZymoBIOMICS DNA/RNA Miniprep Kit (Zymo) according to the manufacturer's instructions. Ribosomal RNA was depleted, and double-stranded cDNA was synthesized according to previous work.[30] Sequencing libraries were prepared from double-stranded cDNA using the Illumina Nextera Library Preparation Kit. Sequencing was performed on an Illumina NextSeq Sequencer with a High Output Flowcell. All code for analyzing the resulting reads is provided in Supporting File 3. Briefly, reads were trimmed and quality-filtered using trimmomatic (v. 0.33)[31] and mapped to the *EcN* genome using bowtie2 (v. 2.3.5).[32] Read counts for each gene were tabulated using FeatureCounts (v. 1.4.5),[33] and differential expression analysis was performed using ALDEx2 (v. 1.16.0)[34] in R. Read count tables are provided in Supporting File 4.

Flow Cytometry. Strains were grown overnight in LB+Kan media and subcultured into fresh LB+Kan media at an optical density (600 nm) of 0.01. These cultures were grown for 2 h before measuring GFP fluorescence directly on a FACSCalibur flow cytometer (Becton Dickinson). Analysis of FCS files was performed using FlowJo (BD Life Sciences). Sterile medium served as a control to identify noncell particles, which were the only particles not included in the gated, analyzed population.

Generation of Standards for Each Promoter Barcode. Triplicates of each of the 30 *EcN* strains harboring a unique barcoded promoter-GFP plasmid construct were grown overnight in LB+Kan, and plasmid DNA was extracted using a spin miniprep kit (Qiagen), generating three independent extracts of each plasmid. Each plasmid was PCR-amplified for

https://dx.doi.org/10.1021/acssynbio.0c00040
ACS Synth. Biol. XXXX, XXX, XXX−XXX

35 cycles using primers GTGCTGAAGTCAAGTTTGAAGG and AGTTGGAACCTCTTACGTGC using Q5 Hotstart DNA Polymerase 2X Mastermix (NEB) according to the manufacturer's instructions. PCR products were separated on a 2% agarose gel, and the band corresponding to the expected amplicon size (583 bp) was extracted using a Minelute Gel Extraction Kit (Qiagen) according to the manufacturer's instructions, with the exception that all steps were performed at 21 000 × *g*, and buffers PE and EB were allowed to incubate in the column for 5 min before centrifugation. Concentration of the resulting DNA fragments was quantified using the Qubit kit (ThermoFisher). These fragments were then used to assemble barcode mixtures of known relative proportions. An equimolar mix (total concentration of 1 ng/μL) was used to measure barcode-specific bias values for the 1-UMI and 2-UMI methods. For both methods, we additionally constructed four barcode mixes (total concentration of 1 ng/μL) covering a 1000-fold range of barcode abundances to validate these barcode-specific bias values. All input concentrations and read counts for these standard curve experiments are provided in Supporting File 5.

Creation of Strain Mixtures for Culture-Based and Mouse Experiments. Triplicate *EcN* strains maintaining each plasmid were grown overnight in 96-deepwell plates in LB +Kan under aerobic conditions with shaking. The optical density (600 nm) of each culture was measured, and cultures were pooled at equal optical densities and mixed with an equal volume of 30% v/v glycerol. The resulting mixture was divided into 1 mL aliquots and frozen at −80 °C until future use.

Culture-Based *EcN* Growth and 2-UMI (Unique Molecular Identifier) Method for Barcoded Transcript Sequencing. One-hundred microliters of the strain mixture was thawed on ice and used to inoculate triplicate 25 mL *in vitro* cultures in either aerobic or anaerobic environments. Aerobic cultures were shaken, but anaerobic cultures were not shaken. Culture samples were taken at both exponential and stationary growth phases. Specifically, at an optical density of 0.3, 10 mL of each culture was centrifuged (3000 × *g* for 5 min at room temperature), resuspended in 0.5 mL of phosphate-buffered saline, placed in an equal volume of DNA/RNA shield, and stored at −80 °C. After overnight growth, 1 mL of each culture was centrifuged (3000 × *g* for 5 min at room temperature), resuspended in 0.5 mL of phosphate-buffered saline, and placed in an equal volume of DNA/RNA shield. All samples were processed with a ZymoBIOMICS DNA/RNA Miniprep Kit (Zymo) according to manufacturer's instructions. Purified RNA was further treated with DNase I supplied in the RiboPure RNA Purification Kit, Yeast (Ambion) according to manufacturer's instructions. Reverse transcription proceeded with primer TTCGTTTTATTTGCCCGGG using a High Capacity Reverse Transcription Kit (ThermoFisher) according to the manufacturer's instructions. RNase A (1.5 μL, Qiagen) was added to the resulting mixture, and digestion was performed at 37 °C for 1 h. Next, a 3-cycle PCR was performed using Q5 Hotstart DNA polymerase on either the barcode mixes (standard curve), retained culture DNA, or synthesized culture cDNA with primers ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNNNNNNNNNNN-XXXXXX-GGCATGGATGAGCTCTACAAATAA and CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT-NNNNNNNNNNNNNNNNNNN-XXXXXX-TTCGTTTTATTTGCCCGGG (standard curve experiments), or ACACTCTTTCCCTACACGACGCTCTTCCGATCT-NNNNNNNNNNNNNNNNNNN-XXXXXX-TGGCATGGATGAGCTCTAC and CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT-NNNNNNNNNNNNNNNNNNN-XXXXXX- TGAGCCTTTCGTTTTATTTGCC (culture experiments), where each "N" refers to a base that is equally likely to be A, C, T, or G, and "XXXXXX" denotes a sample-specific unique molecular identifier (UMI) (Supporting File 6) This PCR was performed according to the manufacturer's instructions, with 5 μL of template, 1.5 picomoles of each primer, a melting temperature of 65 °C, an extension time of 15 s, and 3 total cycles. The PCR products were purified using Agencourt Ampure XP beads (Beckman Coulter). Five microliters of this product was then used as a template for an enrichment PCR using Q5 HotStart DNA polymerase. This PCR was performed using primers AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT and CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT according to the manufacturer's instructions, with 15 picomoles of each primer, a melting temperature of 72 °C, and 35 total cycles. The expected product (227 bp) was separated on a 3% gel, excised, purified using a Minelute Gel Extraction Kit (Qiagen), and quantified using a Qubit (ThermoFisher). Amplicons from each sample were pooled and analyzed on an Illumina Miniseq sequencer.

Growth of *EcN* in the Mouse Gut and 1-UMI Method for Barcoded Transcript Sequencing. All mouse experiments were approved by the Washington University in Saint Louis School of Medicine Division of Comparative Medicine. Through the duration of the experiment, mice consumed standard chow (LabDiet #5053). Then 10^9 colony forming units of the strain mixture in 100 μL of PBS was delivered to each of 4 C57BL/6 mice (2 male, 2 female, 5 weeks of age) via oral gavage and placed in separate cages. Six hours later, mice were sacrificed via CO_2 asphyxiation, and the contents of the stomachs, proximal small intestines, medial small intestines, ceca, and colons were collected in screw-cap vials containing DNA/RNA Shield (Zymo). These vials were then processed using a ZymoBIOMICS DNA/RNA Miniprep Kit (Zymo) according to the manufacturer's instructions. The resulting RNA was treated with DNase I as above. The RNA was reverse transcribed as above, with the exception that the primer sequence CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT-NNNNNNNNNNNNNNNNN-XXXXXX-TGAGCCTTTCGTTTTATTTGCC was used. RNase A treatment after cDNA synthesis was then performed as above. cDNA was purified using a Minelute PCR Purification Kit (Qiagen). Retained DNA from each sample was subjected to a primer extension reaction using the same primer (but a different sample-specific barcode) using Q5 HotStart DNA polymerase according to the protocol: 98 °C for 15 s, 65 °C for 15 s, and 72 °C for 15 min. The reaction (50 μL total) incorporated 5 μL of template and 150 picomoles of primer CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT-NNNNNNNNNNNNNNNN-XXXXXX-TGAGCCTTTCGTTTTATTTGCC. After this reaction completed, 3 μL of exonuclease I (NEB) was added, and the mixture was incubated for 1 h at 37 °C, followed by 5 min at 98 °C. These reactions were purified using Agencourt Ampure XP beads (Beckman Coulter). Next, a 35-cycle PCR was performed on the purified DNA and cDNA products using Q5 Hotstart DNA polymerase according to the manufacturer's

https://dx.doi.org/10.1021/acssynbio.0c00040
ACS Synth. Biol. XXXX, XXX, XXX−XXX

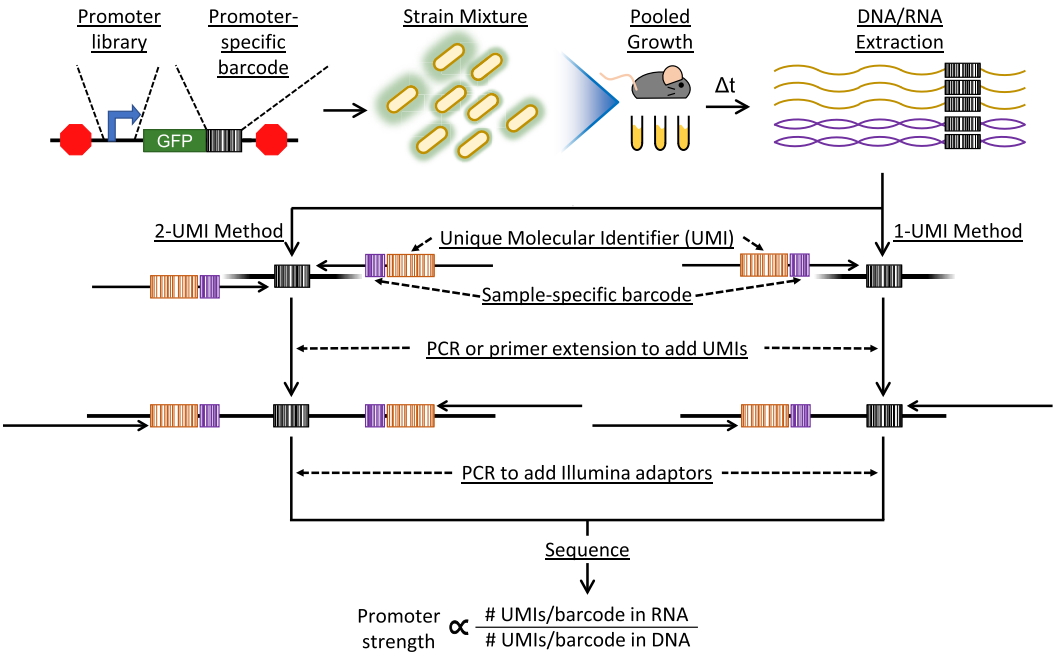


Figure 1. Overview of the transcript barcoding approach. Each promoter of interest was used to drive GFP, which was barcoded at its 3′ end. Strains harboring each expression construct were then mixed and delivered to the appropriate growth context (mice or culture). After, DNA and RNA were extracted, cDNA was synthesized from RNA, and one of two approaches was followed to prepare promoter-specific barcodes for sequencing. In the 2-UMI method (left), barcode-containing sequences were subjected to 3 cycles of PCR with two primers, each containing a different unique molecular identifier (UMI). This 3-cycle PCR was followed by a PCR reaction to add Illumina sequencing adaptors. In the 1-UMI method (right), barcode-containing sequences were subjected to a single primer extension reaction, using a single primer containing a UMI. This reaction was also followed by a PCR reaction to add Illumina sequencing adaptors. The output of both methods was then sequenced, and the number of UMIs identified per promoter barcode in both the cDNA and DNA fractions were counted. The ratio of UMIs/barcode between cDNA and DNA fractions is proportional to promoter strength.

instructions, with an annealing temperature of 72 °C and an extension time of 15 s, with primers AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT and CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT. The expected product (199 bp) was separated on a 3% gel, excised, and purified using a MinElute Gel Extraction Kit, and quantified using Qubit. Amplicons were pooled at equimolar ratios and sequenced on an Illumina Miniseq lane.

Sequencing Analysis. All code for analysis is provided in Supporting File 7. Briefly, reads were first trimmed to the appropriate length using trimmomatic (v. 0.33).[31] Paired-end reads were then merged using usearch (v. 7.0.1090_i86linux64).[35] Merged reads were processed using a custom python script. This script extracted regions of interest including UMI(s), sample barcode, promoter barcode, and flanking regions including GFP and terminator. Reads containing >1 error in GFP and terminator regions were discarded. Reads that contained >1 error in either sample barcode or promoter barcode regions were also discarded. Finally, the number/sequence of unique UMIs (and how many of each UMI) corresponding to each promoter/sample was counted. These UMI sequences and counts were fed into UMI-tools (v. 0.5.0),[36] which accounted for sequencing error in UMI sequences. Finally, the number of unique UMIs for each sample/barcode was counted and written to a text file. All barcode count data for these experiments are provided in Supporting File 8.

Statistical Analysis. Barcode counts, where each barcode was unique to one of 30 promoters, were normalized to generate within-sample relative abundances of the respective constructs. For each sample, the relative abundance of each promoter construct observed in the mRNA-derived cDNA library was divided by the relative abundance observed in the corresponding DNA library. These ratios were transformed by calculating their logarithm base2 (log2 ratio) and serve as an estimate of relative promoter activity within each sample. The log2 ratios were averaged for each promoter across the three replicates of each *in vitro* condition. The *in vivo* log2 ratios were averaged across the replicate mice and the major gut regions of stomach, small intestine (proximal, medial, and distal), and large intestine (cecum and colon). Pearson correlations between culturing conditions and gut sites were calculated using the corrplot (v. 0.84) package in R with Bonferroni correction for multiple hypothesis testing. Hierarchically clustered heatmaps for *in vivo* log2 ratios and pairwise comparisons were generated with the pheatmap (v. 1.0.12) package in R, with Bonferroni correction. Normality of the *in vivo* log2 ratios was assessed with Shapiro-Wilk's test using the *shapiro.test* function in the stats (v. 3.5.1) package in R, and differences in mean log2 ratios between gut sites were assessed by ANOVA using the *anova* function from the same package. Default parameters were used in both cases. Differences in activity between *in vitro* conditions for each promoter, or between the promoters themselves in the gut, were assessed by pairwise Student's *t* tests with Bonferroni correction for multiple hypothesis testing, using the *compare_means* function (ggpubr v. 0.2 package in R). The same function was used to compare activity between gut sites for each promoter, with Benjamini-Hochberg correction for multiple hypothesis testing.

https://dx.doi.org/10.1021/acssynbio.0c00040
ACS Synth. Biol. XXXX, XXX, XXX−XXX

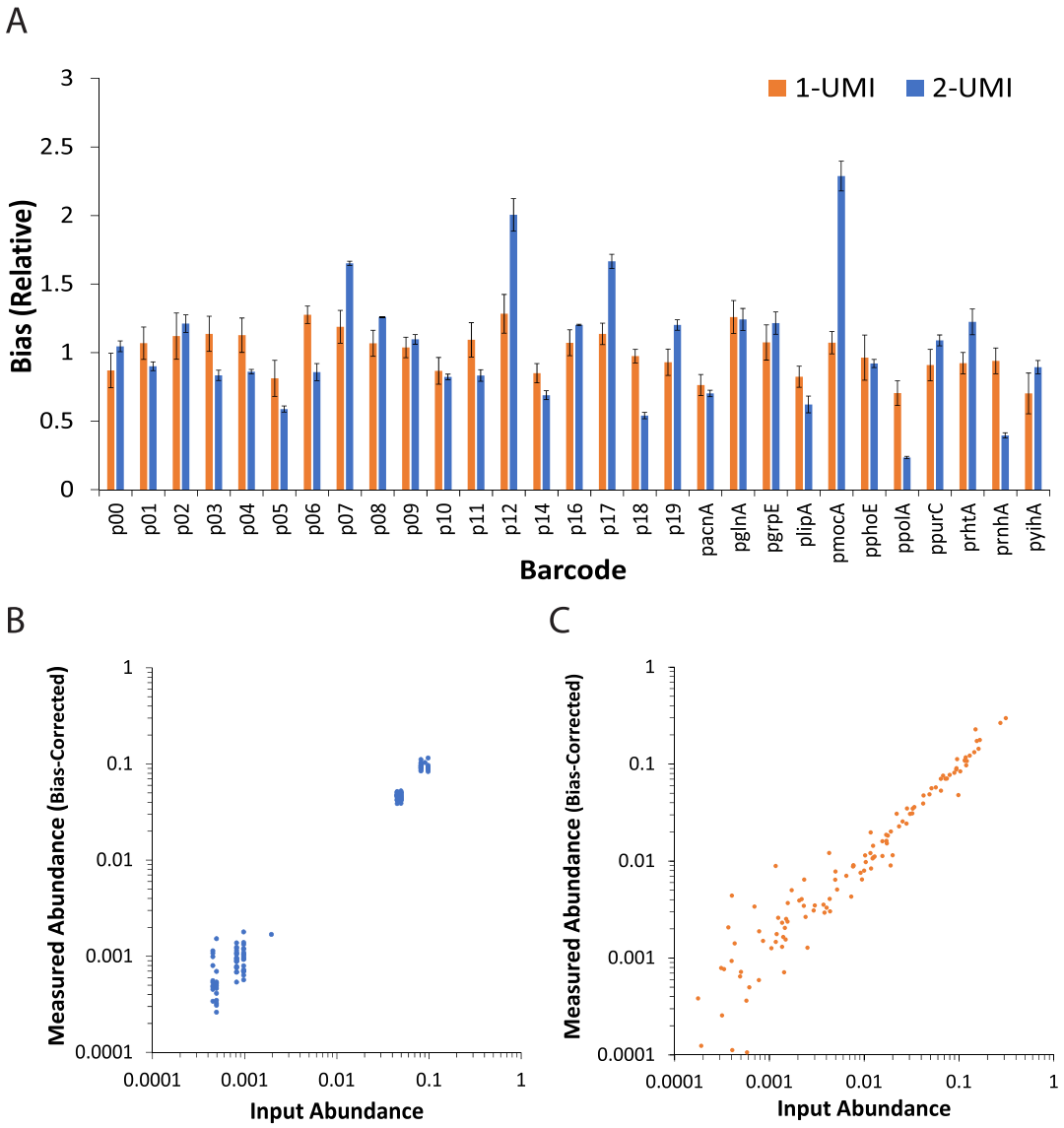


Figure 2. UMI incorporation exhibits reproducible, sequence-specific bias. (A) Relative amounts detected of UMIs incorporated in an equimolar mixture. Error bars indicate the standard deviation observed among three independent experiments. (B) Relationship between input barcode abundance and measured barcode abundance after sequencing and bias correction for the 2-UMI method. (C) Relationship between input barcode abundance and measured barcode abundance after sequencing and bias correction for the 1-UMI method. A different synthetic barcode mixture, which samples the "input abundance" range more densely, was prepared to test the 1-UMI method.

To calculate a mean pooled standard deviation in the context of comparing promoter activities between gut sites, the standard deviation of the log2 ratios was calculated for each promoter in each of the three gut sites. Then, for each promoter p, a pooled standard deviation was calculated for each pair of gut sites i and j as $SD_{pij} = \sqrt{(SD_{pi}^2 + SD_{pj}^2)/2}$. The mean of all such pooled standard deviations was reported. For the power analyses, the *pwr.t.test* function from the pwr (v. 1.2.2) package in R was used. To calculate the number of mice n needed to detect differences of a certain effect size (function parameter d), we scaled the *sig. level* parameter by the number of multiplexed promoters, using a base *sig. level* of 0.05, thereby emulating Bonferroni correction for multiple hypothesis testing. For example, assuming 30 promoters would be tested, the *sig. level* parameter was set to 0.05/30 = 0.00167. In all cases, the power parameter p was set to 0.8.

Linear mixed effects models were fit separately for the log2 ratios in the stomach, small intestine, and large intestine using the lme4 (v. 1.1.21) package in R. In each case, the log2 ratio from the most highly correlated (Pearson) *in vitro* condition for that gut site was the only fixed effect. For all three models, Mouse ID was set as the random effect. Model summary statistics, including marginal and conditional R^2 using the method of Nakagawa et al.,[37] were calculated using the *tab_model* function from the sjPlot (v. 2.6.4) package in R. Expected values were generated using the *fitted* function from the stats (v. 3.5.1) package.

RESULTS

Environmental Context Alters the *EcN* Transcriptome. The gut is known to support varying rates of microbial growth[38] and a radial gradient in oxygen levels.[39] To determine how the transcriptome of *EcN* is modulated by different growth states or oxygen levels, *EcN* was grown in LB media

https://dx.doi.org/10.1021/acssynbio.0c00040
ACS Synth. Biol. XXXX, XXX, XXX−XXX

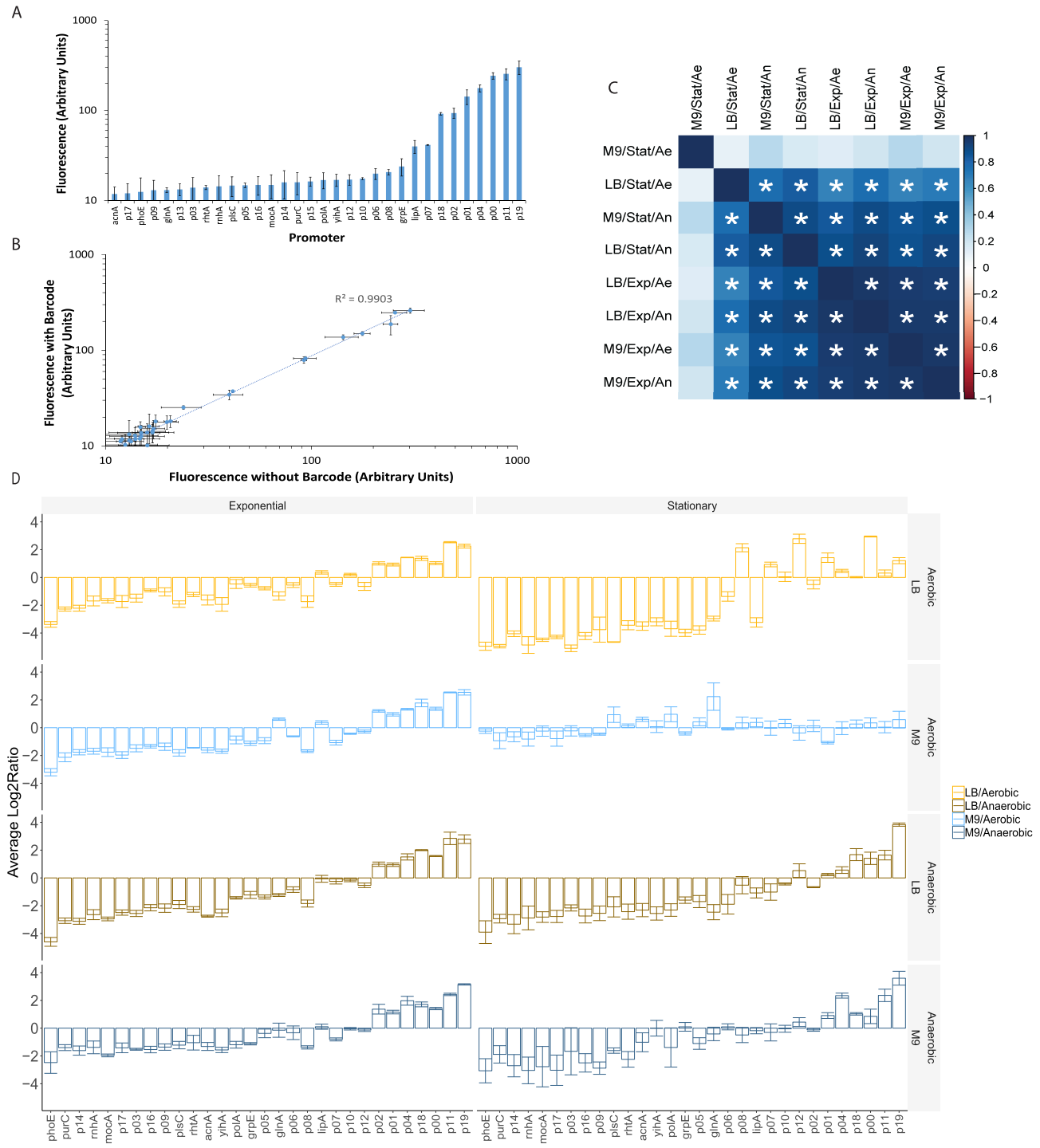


Figure 3. Activities of the promoter library in *EcN* grown in culture. (A) Fluorescence levels exhibited in strains expressing GFP from the indicated promoter. Error bars indicate the standard deviation in mean per-cell fluorescence values observed among 3 biological replicates. (B) Relationship between fluorescence levels exhibited in strains expressing either barcoded or barcode-less GFP across the promoters used in this study. Error bars indicate the standard deviation in mean per-cell fluorescence values observed among 3 biological replicates. (C) Heatmap of pearson correlations between log2 ratios measured in each culture condition. Asterisks indicate significant correlation ($p < 0.05$) after Bonferroni correction for multiple hypothesis testing. Conditions are ordered by average linkage hierarchical clustering. (D) Log2 ratios exhibited in strains grown in the indicated culture condition. Error bars indicate the standard deviation in log2 ratios observed among three biological replicates.

under aerobic or anaerobic conditions, and RNA-seq was performed on RNA extracted from cultures growing under exponential growth or stationary conditions. Principal component analysis of expressed genes revealed substantial clustering by condition (Supporting Figure 1A,B), confirming that gut-relevant conditions indeed cause gene expression changes in *EcN*. RNA-seq also illuminated the expressed regions of *EcN*'s endogenous plasmids pMut1 and pMut2, whose functions are currently unknown (Supporting Figure 1C). On the basis of this analysis, and the known dependence of cloned gene activity on environmental context in other organisms,[40] we hypothesized that cloned genes might also

https://dx.doi.org/10.1021/acssynbio.0c00040
ACS Synth. Biol. XXXX, XXX, XXX−XXX

exhibit altered expression in *EcN* depending on the environment. Moreover, we hypothesized that the expression of cloned genes driven by native promoters might not be predictable based on transcriptomic data of native genes due to differing genomic context and dilution of transcription factors.

Barcodes Enable Accurate Measurements of Nucleic Acid Abundance in Pooled Samples. Because of the high cost of measuring the expression of cloned genes via qPCR in mice, we pursued a transcript barcoding approach.[24−26] An overview of the barcoding and analysis approaches we pursued is shown in Figure 1. Because PCR would be necessary to add sequencing adaptors and prepare enough DNA for sequencing, we were concerned that each barcode sequence might exhibit a different detection frequency, thus skewing our results. To minimize these issues, we adopted two strategies. First, our barcodes were chosen to exhibit a GC content between 40% and 60%. Second, during PCR we appended unique molecular identifiers via one of two strategies. Our first approach, the 2-UMI method, employed a 3-cycle PCR with forward and reverse primers containing 19bp degenerate regions, followed by additional cycles with nondegenerate primers to append sequencing adaptors and increase the concentration of the desired product. Our second approach, the 1-UMI method, employed a single primer extension reaction with a primer containing a 16bp degenerate region, followed by additional cycles with nondegenerate primers to append sequencing adaptors and increase the concentration of the desired product. Specific experimental details are provided in the Materials and Methods.

To measure the precision and accuracy of these methods, we constructed a set of 30 synthetically barcoded *GFP* genes, each with a unique 7 bp barcode at the 3′ end of the gene, directly after the stop codon. We chose to place barcodes at the 3′ end, rather than the 5′ end, of *GFP* to minimize potential changes in gene expression. These barcoded sequences were purified and mixed in equal ratios as described in Supporting File 5, followed by processing in triplicate using either the 1-UMI or 2-UMI method. We found that the observed abundances of UMIs exhibited a consistent, barcode-specific bias, such that certain promoters in independently prepared samples were consistently under- or over-represented across the synthetic mixtures we prepared. The extent of this bias was higher in magnitude (0.24−2.3-fold from expected abundance) for the 2-UMI method than for the 1-UMI method (0.7−1.3-fold from expected abundance) (Figure 2A). Correcting for this bias revealed a tight correlation between input sequence abundance and abundance inferred by sequencing in samples that were not used for the bias calculation (2-UMI method: $r^2 = 0.9716$, $p = 2.7 \times 10^{-113}$, 1-UMI method: $r^2 = 0.9560$, $p = 1.9 \times 10^{-88}$) (Figure 2B,C).

Barcoding Reveals DNA-Normalized mRNA Levels in Culture-Grown *EcN*. To test the efficacy of the barcoding approach, we curated a library of 30 promoters. Twenty of these promoters were taken from a previously published constitutive promoter library, which spanned over a 100-fold range of gene expression,[41] while the remaining 10 were taken from the *EcN* genome. These 10 promoters were chosen based on a recent gut metatranscriptome data set obtained from the fecal matter of healthy individuals.[42] This report detailed the abundance of transcripts mapping to genes from a variety of taxa. Of these, we chose the 10 *E. coli* genes exhibiting the lowest subject-to-subject variation while spanning a wide range of RNA abundance (over 2000-fold) for cloning. This 30-member library was used to drive GFP expression in *EcN*, and the resulting per-cell fluorescence values were measured by flow cytometry (Figure 3A). p19 was the strongest, the *acnA* promoter was the weakest, and the native promoters tended to be weaker than the synthetic promoters. We observed a fairly strong ($r^2 = 0.61$) correlation between the reported strengths of the synthetic promoter library, and the strengths we measured in *EcN* (Supporting Figure 2A). Interestingly, we observed no ($r^2 = 0.0029$) correlation between RNA levels in the gut metatranscriptome and protein levels for the native promoters in culture (Supporting Figure 2B). This discrepancy may be due to imperfect coupling between RNA and protein levels,[43] or it may be due to the differing environmental conditions between the two experiments. We next added the 30 barcodes to these GFP reporters and measured the resulting fluorescence. We found a high correlation between GFP protein levels with and without barcodes, ($r^2 = 0.9903$, $p = 9.1 \times 10^{-32}$), supporting their ability to serve as noninvasive reporters of promoter activity (Figure 3B).

We then applied the 2-UMI method to measure transcript levels in these 30 barcoded strains across 8 conditions, which encompass all combinations of aerobic versus anaerobic, exponential growth versus stationary phase, and LB versus M9 media. We expected that each strain might exhibit a different growth rate, and so we extracted both DNA and RNA from each sample and measured barcode abundance in each. Barcode counts from the DNA and RNA species were normalized within each sample. We then took the ratio of the RNA relative abundances to the corresponding DNA relative abundances to compute RNA expression on a per-DNA basis. The data we report contain the logarithm (base 2) of this ratio (log2 ratio) (Figure 3D). RNA/DNA ratios correlated with GFP protein levels ($r^2 = 0.88$) in aerobic conditions during exponential phase (Supporting Figure 3A), which is the only condition under which GFP protein levels were measured. However, correlating these ratios with the regularized logarithm of read counts for *EcN*'s endogenous genes across all 4 LB media conditions yielded no trend ($r^2 = 0.097$) (Supporting Figure 3B), supporting our focus on cloned genes versus endogenous genes. Except for cells in stationary phase growing in M9 media under aerobic conditions, the log2 ratios were significantly correlated across *in vitro* conditions (Figure 3C). However, some individual promoters exhibited significant condition-dependent gene expression. For example, both mocA and purC exhibited significant differences ($p < 0.05$) in expression in exponential phase compared to stationary phase when grown in LB aerobically, while the activity of p01 was particularly affected by stationary phase after aerobic growth in M9 media (Supporting Figure 5). These results agree with recent reports of condition-dependent expression in cloned genes for other organisms[40] and support the use of transcript barcoding as a high-throughput approach to measure the expression levels of many constructs under several different conditions.

Barcoding Reveals DNA-Normalized RNA Levels in *EcN* Residing in the Mouse Gut. Encouraged by the high throughput of the barcoding method, we asked whether the same approach could reveal differences in expression of the synthetic library when placed in the mouse gut. The 30-member library was grown overnight in LB media and delivered to 4 mice (2 males, 2 females) via oral gavage. Contents of the stomach, small intestine, and large intestine

https://dx.doi.org/10.1021/acssynbio.0c00040
ACS Synth. Biol. XXXX, XXX, XXX−XXX

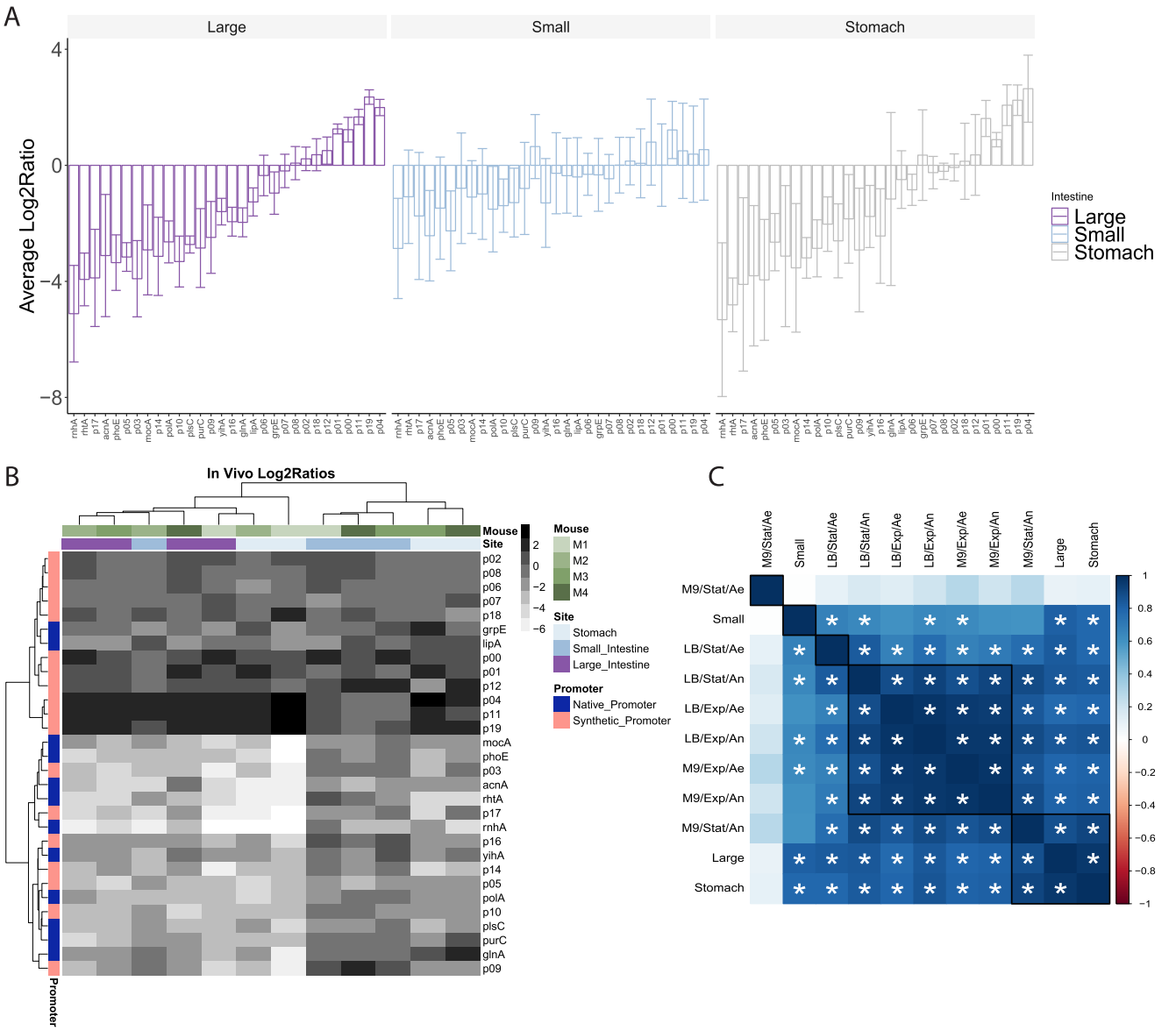


Figure 4. Activities of the promoter library in *EcN* in the gut. (A) Log2 ratios exhibited by the strains detected in the large intestine, small intestine, and stomach. Error bars indicate the standard deviation in mean log2 ratios observed across 4 mice. (B) Heatmap of log2 ratios *in vivo*. Rows and columns are ordered by complete linkage hierarchical clustering. (C) Heatmap of pearson correlation between log2 ratios measured in the *in vitro* culture conditions and the large intestine, small intestine, and stomach. Asterisks indicate significant correlation after Bonferroni correction for multiple hypothesis testing. Conditions are ordered by average linkage hierarchical clustering.

were collected after 6 h, based on data indicating that ingested tracers are present throughout the GI tract at this time.[44] RNA and DNA levels of each barcode were then measured via the 1-UMI approach, due to its lower yield of off-target amplicons than the 2-UMI approach for these samples. While the number of barcoded sequences we recovered was generally lower in the gut than in culture, the distribution of log2 ratios was still approximately normal, supporting the use of statistical tests based on the normal distribution (Supporting Figure 4). This analysis indicated a large range (~8 log2) of gene expression *in vivo* across the promoters, except for in the small intestine which exhibited relatively high intersubject variability and low dynamic range (Figure 4A). Hierarchical clustering revealed that the synthetic and native promoters behaved differently across gut sites, with synthetic promoters generally exhibiting stronger activity than native promoters (Figure 4B). This result was independent of the mouse in which the activities were measured. Furthermore, many pairs of promoters exhibited significantly different gene expression levels, and these differences were more pronounced in the large intestine than in the stomach or small intestine (Supporting Figure 6).

In agreement with the culture-based data, we found that promoter activities were generally correlated across gut and culture conditions (Figure 4C). This correlation indicates that some culture-based conditions could be good predictors of promoter activity in the gut, thereby supporting their use in rapid prototyping of engineered probiotic strains. We found that log2 ratios recovered from cells at stationary phase grown in M9 media under anaerobic conditions were sufficient to predict activities in stomach and large intestine, but no culture condition or combination thereof, even the culture condition most highly correlated with activities in the small intestine (aerobic LB at stationary phase), did well at predicting gene expression in the small intestine (Figure 5, Supporting Figure

https://dx.doi.org/10.1021/acssynbio.0c00040
ACS Synth. Biol. XXXX, XXX, XXX−XXX

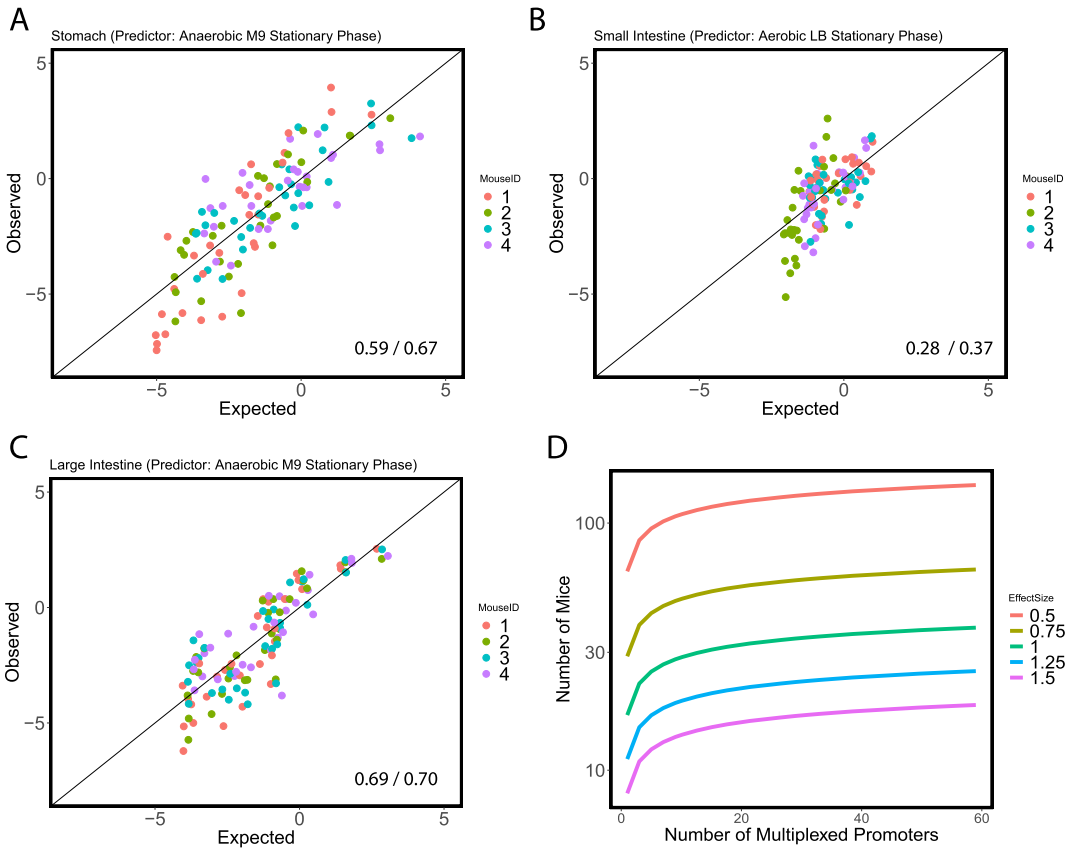


Figure 5. Predicted versus observed *in vivo* log2 ratios and power analysis. Expected versus observed *in vivo* log2 ratios in the (A) stomach, (B) small intestine, and (C) large intestine. Expected values were generated with linear mixed effects models using the most highly correlated *in vitro* condition for each gut site. Model summary statistics are available in Supporting Figure 7. The inset values are the marginal/conditional R^2. (D) Power analysis. Number of mice (vertical axis) required to observe significant differences of a given effect size (legend) for the indicated number of multiplexed promoters (horizontal axis), assuming Bonferroni correction for multiple hypothesis training. Sample sizes were calculated with power = 0.8.

7). In agreement with this, the overall distributions of promoter activities and variances depended significantly on the gut site (Supporting Figure 8), indicating that the activities of some individual promoters might depend on gut site. Indeed, 9/30 promoters exhibited significantly different ($p < 0.05$) activities between gut sites (e.g., *prhtA* exhibited a difference in RNA/DNA ratio of 13-fold between the stomach and the small intestine) (Supporting Table 1).

Our data set enabled us to perform a power analysis for detecting differences between gut sites in promoter activity based upon the interindividual variability we observed. When comparing promoter activities between gut sites, we observed a pooled standard deviation for the log2 ratio averaging 1.2. Therefore, based on a desired statistical power and effect size (difference in means divided by pooled standard deviation), it is possible to compute the number of mice required (Figure 5D). For example, to identify promoters exhibiting an 8-fold change in gene expression (effect size of 3/1.2) between gut sites among a set of 30 promoters at a power of 0.8, 8 mice are required. Conversely, with the 4 mice used in this study, only effect sizes of 4.74, or $2^{(4.74*1.2)} \approx$ 50-fold, changes in expression could be detected among 30 promoters at a power of 0.8. Because the interindividual differences in gene expression we observed are high, the ability to use one group of mice per promoter library, rather than per promoter, makes transcript barcoding a particularly desirable approach to measuring cloned gene expression in the mammalian gut.

DISCUSSION

In this work, we measured the ability of transcript barcoding to report the activities of synthetic constructs in the mammalian gut. We first directly measured the accuracy of this approach using nucleic acids of known concentration. We found that, even when using unique molecular identifiers, separate DNA sequences were unevenly recovered during sequencing library preparation. These effects were substantial, up to 9.7-fold for the commonly used 2-UMI method, and up to 1.8-fold for the commonly used 1-UMI method. However, these biases were consistent for the unique molecular identifiers across input concentrations, allowing for approaches to account for them. First, barcode abundances can be compared as a ratio between conditions to cancel out the bias, which was implemented here as a ratio between barcode abundances in the RNA and DNA pools. As a second approach, multiple barcodes could be used to track each promoter, effectively averaging out the bias. Therefore, it is important that UMI-based studies incorporate one of these two methods, studies that use UMIs to measure the abundance of individual sequences on an absolute basis are predicted to exhibit high technical repeatability (i.e., precision) but low correlation with actual sequence abundance (i.e., accuracy). Alternatively, given more data, it might be possible to train models of UMI incorporation and PCR bias to avoid these issues and improve quantitative accuracy for the myriad

https://dx.doi.org/10.1021/acssynbio.0c00040
ACS Synth. Biol. XXXX, XXX, XXX−XXX

techniques in which these methods are used including next-generation sequencing and vector library construction.

Having validated the accuracy and precision of the barcoding approach, we next applied it to measure gene expression of cloned *EcN* genes in culture and in the mammalian gut. We found that promoter activities were significantly correlated across sites, indicating that, at a broad level, *in vitro* activities can be used to predict promoter activities in the stomach and large intestines. However, we also found evidence that promoters exhibit differing levels of activity in different regions of the intestine. The existence of such differences is understandable, as the abundances of various chemicals (e.g., pH, bile acids, water, short-chain fatty acids, etc.) and other microbial species vary substantially along the gut. These differences could combine to make synthetic construct expression in the gut context-dependent, which is important for several reasons. First, promoters that change activity based on gut site are useful, for example, to design probiotics that change their behaviors upon entering the gut, to deliver biomolecules in a site-specific fashion (e.g., small intestine vs large intestine), or to initiate behaviors upon exiting the gut.[45−47] Second, many biomolecules that are interesting to deliver to the gut (such as drugs) have defined therapeutic windows. The ability to predict promoter activity in different gut sites will enable the design of engineered probiotic therapies with well-defined pharmacokinetics.[47] Third, the performance of some synthetic regulatory circuits depends heavily on the precise expression levels of its regulatory components.[48] Indeed, we were able to use these promoter strength measurements to design probiotics, which expressed a therapeutic protein (*PAL2*) at predictable levels *in vitro* and *in vivo*, which we characterized in previously published work.[49] Engineered *EcN* expressing *PAL2* is currently being investigated in clinical trials as an oral therapy for phenylalanine hydroxylase-deficient phenylketonuria (PKU)[16] and is therefore a clinically relevant example of a therapeutic protein that can be delivered by an engineered probiotic. We previously demonstrated that strains containing p19-PAL2 were able to reduce serum phenylalanine levels up to 50% in a murine model of PKU, while strains containing p08-PAL2 reduced serum phenylalanine to lesser extent, results which agreed with our measurements of p08 and p19 strength using barcoded GFP.[49] This agreement demonstrates the utility of measuring promoter strengths in parallel using barcoded GFP to inform the construction of engineered therapeutic strains. Looking forward, measuring the extent to which promoters change their activity in the gut will allow probiotic engineers to implement more complex decision-making abilities.

In this study, we measured the distributions of cloned gene expression in probiotics residing in the gut. This information is critical to the design of well-powered experiments to measure the performance of synthetic parts in gut-resident probiotics. Moreover, this analysis revealed the extent to which transcript barcoding can reduce the number of mice required to detect differences in promoter activity. Specifically, this approach allows one group of mice per promoter library, rather than per promoter, which is a significant cost and ethical advantage.

Taken together, this work establishes transcript barcoding as a desirable method for measurement of cloned gene expression in organisms residing in habitats that are difficult to model *in vitro*. Looking forward, this approach can be extended to any organism for which genetic tools exist, and any habitat to which engineered microorganisms can be applied. We anticipate that given the current ease with which new synthetic constructs can be constructed or synthesized, transcript barcoding will find increasing use to rapidly test and identify engineered strains exhibiting the desired *in situ* behavior.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.0c00040.

RNA-Seq of *E. coli* Nissle grown; GFP fluorescence data comparisons; DNA/RNA ratio comparison; normality of *in vivo* barcode sequencing data; promoters that exhibit significantly different activities between different *in vitro* conditions; promoters that exhibit significantly different activities from each other within each gut site; details of linear models; distributions of promoter activity and standard deviation between *in vivo* conditions; *p*-values for null hypothesis that promoter activities are the same between gut sites (PDF)

Genbank files of all plasmids used in this study; all primers used in this study; code used to perform RNA-Seq analysis; read counts of *EcN* genes from RNA-Seq analysis; input abundances and read counts for all standard curve barcode sequencing experiments; all barcode sequences used in this study; all code used to perform barcode sequencing analysis; read counts from *in vitro* and *in vivo* barcode sequencing experiments (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

Gautam Dantas − *The Edison Family Center for Genome Sciences & Systems Biology, Department of Pathology and Immunology, and Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri 63110, United States; Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, Missouri 63130, United States;* orcid.org/0000-0003-0455-8370; Email: dantas@wustl.edu

Authors

Nathan Crook − *Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, North Carolina 27606, United States;* orcid.org/0000-0001-6165-1972

Aura Ferreiro − *The Edison Family Center for Genome Sciences & Systems Biology, Washington University School of Medicine, St. Louis, Missouri 63110, United States; Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, Missouri 63130, United States*

Zevin Condiotte − *Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, Missouri 63130, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.0c00040

Funding

This work is supported in part by the NIH Director's New Innovator Award (http://commonfund.nih.gov/newinnovator/), the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK: http://www.niddk.

nih.gov/), the National Institute of General Medical Sciences (NIGMS: http://www.nigms.nih.gov/), the National Institute of Allergy and Infectious Diseases (NIAID: https://www.niaid.nih.gov/), and the Eunice Kennedy Shriver National Institute of Child Health and Human Development (https://www.nichd.nih.gov/) of the National Institutes of Health (NIH) under Award Nos. DP2DK098089, R01GM099538, R01AI123394, and R01HD092414 to G.D. This work was also supported in part by the Kenneth Rainin Foundation Innovation and Breakthrough Awards (13H5) to G.D. N.C. received support from the NIDDK Pediatric Gastroenterology Research Training Program of the NIH, under Award No. T32 DK077653 (Phillip I. Tarr, Principal Investigator). A.F. received support from the Chancellor's Graduate Research Fellowship Program at Washington University in St. Louis. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

Notes

The authors declare no competing financial interest.

Demultiplexed reads for the *EcN* transcriptome have been deposited at NCBI under accession numbers SRR10424477 through SRR10424487. Raw reads for the *in vitro* barcode sequencing experiment have been deposited at NCBI under accession number SRR9062664. Raw reads for the *in vivo* barcode sequencing have been deposited at NCBI under accession number SRR9062663.

■ ACKNOWLEDGMENTS

We would like to acknowledge Jessica Hoisington-Lopez, Eric Martin, and Brian Koebbe for next-generation sequencing and high-throughput computing support at the Edison Family Center for Genome Sciences and Systems Biology at Washington University in St. Louis School of Medicine, and members of the Dantas Lab for helpful discussions relating to work described herein.

■ REFERENCES

(1) Honda, K., and Littman, D. R. (2016) The microbiota in adaptive immune homeostasis and disease. *Nature 535*, 75.

(2) Lloyd-Price, J., Abu-Ali, G., and Huttenhower, C. (2016) The healthy human microbiome. *Genome Med. 8*, 51.

(3) Pickard, J. M., Zeng, M. Y., Caruso, R., and Nunez, G. (2017) Gut microbiota: Role in pathogen colonization, immune responses, and inflammatory disease. *Immunol. Rev. 279*, 70−89.

(4) Sommer, F., Anderson, J. M., Bharti, R., Raes, J., and Rosenstiel, P. (2017) The resilience of the intestinal microbiota influences health and disease. *Nat. Rev. Microbiol. 15*, 630−638.

(5) Walters, W. A., Xu, Z., and Knight, R. (2014) Meta-analyses of human gut microbes associated with obesity and IBD. *FEBS Lett. 588*, 4223−4233.

(6) Halfvarson, J., Brislawn, C. J., Lamendella, R., Vázquez-Baeza, Y., Walters, W. A., Bramer, L. M., D'Amato, M., Bonfiglio, F., McDonald, D., Gonzalez, A., McClure, E. E., Dunklebarger, M. F., Knight, R., and Jansson, J. K. (2017) Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat. Microbiol. 2*, 17004.

(7) Millan, B., Park, H., Hotte, N., Mathieu, O., Burguiere, P., Tompkins, T. A., Kao, D., and Madsen, K. L. (2016) Fecal Microbial Transplants Reduce Antibiotic-resistant Genes in Patients With Recurrent Clostridium difficile Infection. *Clin. Infect. Dis. 62*, 1479−1486.

(8) Tilg, H., and Moschen, A. R. (2014) Microbiota and diabetes: an evolving relationship. *Gut 63*, 1513−1521.

(9) Haran, J. P., Bhattarai, S. K., Foley, S. E., Dutta, P., Ward, D. V., Bucci, V., and McCormick, B. A. (2019) Alzheimer's Disease Microbiome Is Associated with Dysregulation of the Anti-Inflammatory P-Glycoprotein Pathway. *mBio 10*, No. e00632−00619.

(10) Cully, M. (2019) Microbiome therapeutics go small molecule. *Nat. Rev. Drug Discovery 18*, 569.

(11) Gupta, S., Allen-Vercoe, E., and Petrof, E. O. (2016) Fecal microbiota transplantation: in perspective. *Ther. Adv. Gastroenterol. 9*, 229−239.

(12) Makki, K., Deehan, E. C., Walter, J., and Bäckhed, F. (2018) The Impact of Dietary Fiber on Gut Microbiota in Host Health and Disease. *Cell Host Microbe 23*, 705−715.

(13) Ozdemir, T., Fedorec, A. J. H., Danino, T., and Barnes, C. P. (2018) Synthetic Biology and Engineered Live Biotherapeutics: Toward Increasing System Complexity. *Cell Syst. 7*, 5−16.

(14) Ciorba, M. A. (2012) A Gastroenterologist's Guide to Probiotics. *Clin. Gastroenterol. Hepatol. 10*, 960−968.

(15) Lubkowicz, D., Ho, C. L., Hwang, I. Y., Yew, W. S., Lee, Y. S., and Chang, M. W. (2018) Reprogramming Probiotic Lactobacillus reuteri as a Biosensor for Staphylococcus aureus Derived AIP-I Detection. *ACS Synth. Biol. 7*, 1229−1237.

(16) Isabella, V. M., Ha, B. N., Castillo, M. J., Lubkowicz, D. J., Rowe, S. E., Millet, Y. A., Anderson, C. L., Li, N., Fisher, A. B., West, K. A., Reeder, P. J., Momin, M. M., Bergeron, C. G., Guilmain, S. E., Miller, P. F., Kurtz, C. B., and Falb, D. (2018) Development of a synthetic live bacterial therapeutic for the human metabolic disease phenylketonuria. *Nat. Biotechnol. 36*, 857.

(17) Kurtz, C. B., Millet, Y. A., Puurunen, M. K., Perreault, M., Charbonneau, M. R., Isabella, V. M., Kotula, J. W., Antipov, E., Dagon, Y., Denney, W. S., Wagner, D. A., West, K. A., Degar, A. J., Brennan, A. M., and Miller, P. F. (2019) An engineered *E. coli* Nissle improves hyperammonemia and survival in mice and shows dose-dependent exposure in healthy humans. *Sci. Transl. Med. 11*, No. eaau7975.

(18) Gurbatri, C., Coker, C., Hinchliffe, T. E., Lia, I., Castro, S., Treuting, P. M., Arpaia, N., and Danino, T. (2019) Engineered probiotics for local tumor delivery of checkpoint blockade nanobodies. *bioRxiv*, 562785.

(19) Lindsay, K. L., Brennan, L., and McAuliffe, F. M. (2014) Acceptability of and compliance with a probiotic capsule intervention in pregnancy. *Int. J. Gynecol. Obstet. 125*, 279−280.

(20) Zmora, N., Zilberman-Schapira, G., Suez, J., Mor, U., Dori-Bachash, M., Bashiardes, S., Kotler, E., Zur, M., Regev-Lehavi, D., Brik, R. B.-Z., Federici, S., Cohen, Y., Linevsky, R., Rothschild, D., Moor, A. E., Ben-Moshe, S., Harmelin, A., Itzkovitz, S., Maharshak, N., Shibolet, O., Shapiro, H., Pevsner-Fischer, M., Sharon, I., Halpern, Z., Segal, E., and Elinav, E. (2018) Personalized Gut Mucosal Colonization Resistance to Empiric Probiotics Is Associated with Unique Host and Microbiome Features. *Cell 174*, 1388−1405.

(21) Bleuven, C., and Landry, C. R. (2016) Molecular and cellular bases of adaptation to a changing environment in microorganisms. *Proc. R. Soc. London, Ser. B 283*, 20161458.

(22) Arpino, J. A. J., Hancock, E. J., Anderson, J., Barahona, M., Stan, G.-B. V., Papachristodoulou, A., and Polizzi, K. (2013) Tuning the dials of Synthetic Biology. *Microbiology 159*, 1236−1253.

(23) Jalili-Firoozinezhad, S., Gazzaniga, F. S., Calamari, E. L., Camacho, D. M., Fadel, C. W., Bein, A., Swenor, B., Nestor, B., Cronce, M. J., Tovaglieri, A., Levy, O., Gregory, K. E., Breault, D. T., Cabral, J. M. S., Kasper, D. L., Novak, R., and Ingber, D. E. (2019) A complex human gut microbiome cultured in an anaerobic intestine-on-a-chip. *Nat. Biomed. Eng. 3*, 520−531.

(24) Urtecho, G., Tripp, A. D., Insigne, K. D., Kim, H., and Kosuri, S. (2019) Systematic Dissection of Sequence Elements Controlling σ70 Promoters Using a Genomically Encoded Multiplexed Reporter Assay in Escherichia coli. *Biochemistry 58*, 1539−1551.

(25) Fu, G. K., Xu, W., Wilhelmy, J., Mindrinos, M. N., Davis, R. W., Xiao, W., and Fodor, S. P. A. (2014) Molecular indexing enables quantitative targeted RNA sequencing and reveals poor efficiencies in standard library preparations. *Proc. Natl. Acad. Sci. U. S. A. 111*, 1891−1896.

https://dx.doi.org/10.1021/acssynbio.0c00040
ACS Synth. Biol. XXXX, XXX, XXX−XXX

(26) Kivioja, T., Vähärautio, A., Karlsson, K., Bonke, M., Enge, M., Linnarsson, S., and Taipale, J. (2012) Counting absolute numbers of molecules using unique molecular identifiers. *Nat. Methods 9*, 72.
(27) Liu, D., Xiao, Y., Evans, B. S., and Zhang, F. (2015) Negative Feedback Regulation of Fatty Acid Production Based on a Malonyl-CoA Sensor-Actuator. *ACS Synth. Biol. 4*, 132−140.
(28) Lutz, R., and Bujard, H. (1997) Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res. 25*, 1203−1210.
(29) Dowell, V. R., and Hawkins, T. M. (1974) *Laboratory Methods in Anaerobic Bacteriology; CDC Laboratory Manual*; Centers for Disease Control, Atlanta, GA.
(30) Yoneda, A., Wittmann, B. J., King, J. D., Blankenship, R. E., and Dantas, G. (2016) Transcriptomic analysis illuminates genes involved in chlorophyll synthesis after nitrogen starvation in Acaryochloris sp. CCMEE 5410. *Photosynth. Res. 129*, 171−182.
(31) Bolger, A. M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics 30*, 2114−2120.
(32) Langmead, B., and Salzberg, S. L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods 9*, 357.
(33) Liao, Y., Smyth, G. K., and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics 30*, 923−930.
(34) Fernandes, A. D., Macklaim, J. M., Linn, T. G., Reid, G., and Gloor, G. B. (2013) ANOVA-Like Differential Expression (ALDEx) Analysis for Mixed Population RNA-Seq. *PLoS One 8*, No. e67019.
(35) Edgar, R. C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics 26*, 2460−2461.
(36) Smith, T. S., Heger, A., and Sudbery, I. (2017) UMI-tools: Modelling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res. 27*, 491.
(37) Nakagawa, S., Johnson, P., and Schielzeth, H. (2017) The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *J. R. Soc., Interface 14*, 20170213.
(38) Korem, T., Zeevi, D., Suez, J., Weinberger, A., Avnit-Sagi, T., Pompan-Lotan, M., Matot, E., Jona, G., Harmelin, A., Cohen, N., Sirota-Madi, A., Thaiss, C. A., Pevsner-Fischer, M., Sorek, R., Xavier, R., Elinav, E., and Segal, E. (2015) Growth dynamics of gut microbiota in health and disease inferred from single metagenomic samples. *Science 349*, 1101−1106.
(39) Zheng, L., Kelly, C. J., and Colgan, S. P. (2015) Physiologic hypoxia and oxygen homeostasis in the healthy intestine. A Review in the Theme: Cellular Responses to Hypoxia. *Am. J. Physiol.: Cell Physiol. 309*, C350−C360.
(40) Bandiera, L., Furini, S., and Giordano, E. (2016) Phenotypic Variability in Synthetic Biology Applications: Dealing with Noise in Microbial Gene Expression. *Front. Microbiol. 7*, 479 DOI: 10.3389/fmicb.2016.00479.
(41) Anderson, C. (2017) *Anderson Promoter Collection*, iGEM. http://parts.igem.org/Promoters/Catalog/Anderson.
(42) Franzosa, E. A., Morgan, X. C., Segata, N., Waldron, L., Reyes, J., Earl, A. M., Giannoukos, G., Boylan, M. R., Ciulla, D., Gevers, D., Izard, J., Garrett, W. S., Chan, A. T., and Huttenhower, C. (2014) Relating the metatranscriptome and metagenome of the human gut. *Proc. Natl. Acad. Sci. U. S. A. 111*, E2329−E2338.
(43) Vogel, C., and Marcotte, E. M. (2012) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet. 13*, 227.
(44) Padmanabhan, P., Grosse, J., Asad, A. B. M. A., Radda, G. K., and Golay, X. (2013) Gastrointestinal transit measurements in mice with 99mTc-DTPA-labeled activated charcoal using NanoSPECT-CT. *EJNMMI Res. 3*, 60.
(45) Tropini, C., Earle, K. A., Huang, K. C., and Sonnenburg, J. L. (2017) The Gut Microbiome: Connecting Spatial Organization to Function. *Cell Host Microbe 21*, 433−442.
(46) Claesen, J., and Fischbach, M. A. (2015) Synthetic microbes as drug delivery systems. *ACS Synth. Biol. 4*, 358−364.
(47) Jimenez, M., Langer, R., and Traverso, G. (2019) Microbial therapeutics: New opportunities for drug delivery. *J. Exp. Med. 216*, 1005−1009.
(48) Dehli, T., Solem, C., and Jensen, P. R. (2012) Tunable Promoters in Synthetic and Systems Biology, In *Reprogramming Microbial Metabolic Pathways* (Wang, X., Chen, J., and Quinn, P., Eds.) pp 181−201, Springer Netherlands, Dordrecht.
(49) Crook, N., Ferreiro, A., Gasparrini, A. J., Pesesky, M. W., Gibson, M. K., Wang, B., Sun, X., Condiotte, Z., Dobrowolski, S., Peterson, D., and Dantas, G. (2019) Adaptive Strategies of the Candidate Probiotic E. coli Nissle in the Mammalian Gut. *Cell Host Microbe 25*, 499−512.

https://dx.doi.org/10.1021/acssynbio.0c00040
ACS Synth. Biol. XXXX, XXX, XXX−XXX