

# Cas9-assisted biological containment of a genetically engineered human commensal bacterium and genetic elements

Received: 7 November 2021

Accepted: 7 February 2024

Published online: 07 March 2024

 Check for updatesNaoki Hayashi<sup>1,2,10</sup>, Yong Lai<sup>3,4,5,10</sup>, Jay Fuerte-Stone<sup>6</sup>, Mark Mimeo<sup>6,7</sup> ✉ & Timothy K. Lu<sup>1,3,4,8,9</sup> ✉

Sophisticated gene circuits built by synthetic biology can enable bacteria to sense their environment and respond predictably. Engineered biosensing bacteria outfitted with such circuits can potentially probe the human gut microbiome to prevent, diagnose, or treat disease. To provide robust biocontainment for engineered bacteria, we devised a Cas9-assisted auxotrophic biocontainment system combining thymidine auxotrophy, an Engineered Riboregulator (ER) for controlled gene expression, and a CRISPR Device (CD). The CD prevents the engineered bacteria from acquiring *thyA* via horizontal gene transfer, which would disrupt the biocontainment system, and inhibits the spread of genetic elements by killing bacteria harboring the gene cassette. This system tunably controlled gene expression in the human gut commensal bacterium *Bacteroides thetaiotaomicron*, prevented escape from thymidine auxotrophy, and blocked transgene dissemination. These capabilities were validated in vitro and in vivo. This biocontainment system exemplifies a powerful strategy for bringing genetically engineered microorganisms safely into biomedicine.

The microbiome has recently become a point of focus for biomedical research. Altered microbial communities have been associated with a multitude of diseases, including inflammatory bowel diseases<sup>1,2</sup>, liver disease<sup>3</sup>, metabolic disorders<sup>4</sup>, and cancer<sup>5</sup>, as well as responses to COVID-19<sup>6</sup> and other infections<sup>7</sup>. Synthetic biology offers powerful tools for manipulating the microbiome and elucidating the functions of its various components. Genetically engineered bacteria and sophisticated genetic circuits provide the means to investigate the exchange of biological signals between human and microbial cells and to eventually gain control over processes that can shield the body from diseases or overcome diseases at their onset.

Stringent regulations limit the use of genetically engineered bacteria to prevent them from being unintentionally released into the environment. The potential benefit of these organisms can be realized only once we have found ways to comply with these regulations. Thus, before the tools of synthetic biology can be applied, for example, to produce a useful protein in situ<sup>8</sup>, to stimulate immunity in response to cancer<sup>9</sup>, or to destroy a harmful virus in a host<sup>10</sup>, technologies are needed to contain the genetically modified organisms and the genetic material they carry, which could find its way into other cells and alter the microbiome in unpredictable ways.

<sup>1</sup>Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. <sup>2</sup>JSR-Keio University Medical and Chemical Innovation Center (JKIC), JSR Corp., 35 Shinanomachi, Shinjuku, Tokyo 160-8582, Japan. <sup>3</sup>Synthetic Biology Group, MIT Synthetic Biology Center, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA. <sup>4</sup>Research Laboratory of Electronics, MIT, Cambridge, MA 02139, USA. <sup>5</sup>Department of Chemical and Biological Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR. <sup>6</sup>Department of Microbiology, The University of Chicago, Chicago, IL 60637, USA. <sup>7</sup>Pritzker School of Molecular Engineering, The University of Chicago, Chicago, IL 60637, USA. <sup>8</sup>Broad Institute, Cambridge, MA 02139, USA. <sup>9</sup>Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA 02139, USA. <sup>10</sup>These authors contributed equally: Naoki Hayashi, Yong Lai. ✉ e-mail: [mmimeo@uchicago.edu](mailto:mmimeo@uchicago.edu); [tim@lugroup.org](mailto:tim@lugroup.org)

Many forms of biological containment have been devised for uncontrolled environments that do not require human monitoring or input<sup>11,12</sup>. There are four main strategies for accomplishing biological containment: auxotrophy, kill switches, genetic firewalls, and physical containment. Auxotrophy takes advantage of engineered auxotrophic organisms, which are unable to synthesize an essential compound or to utilize an environmentally available compound required for their survival. The lack of the integral compound induces lethality outside the uncontrolled area<sup>13</sup>. The kill switch employs the conditional production of a toxic molecule whose gene expression is tightly controlled by an environmentally responsive element. Once engineered bacteria with the kill switch come into an undesirable environment, the switch is turned on and the bacteria are killed by the toxic molecule<sup>14</sup>. Genetic firewalls use alternative genetic codes in bacteria for non-standard amino acids (NSAAs), enabling genetic isolation by preventing horizontal gene transfer (HGT). The growth of the bacteria harboring the genetic firewall thus depends on the use of the NSAAs, such that the firewall sets up robust and orthogonal barriers between the engineered organisms and the environment<sup>15,16</sup>. The artificial genetic code, incorporating NSAAs, makes the standard amino acids from natural microorganisms nonfunctional in the engineered bacteria. Additionally, as bacteriophage infection can lead to the HGT of bacterial genes, a genetic firewall has been created by the deletion of cellular transfer RNAs and release factor 1<sup>17</sup>. Physical containment involves trapping the engineered bacteria in a space with well-designed materials, such as a multilayer hydrogel capsule. The bacteria cannot escape but still have access to nutrients and chemical inputs from their outer environment<sup>18</sup>. These four strategies prevent the engineered strains and the genetic elements from escaping from controlled environments.

Auxotrophy, one of the most promising approaches, has already been applied to develop genetically modified therapeutic bacteria as pharmaceutical agents<sup>8</sup>, such as the biocontainment of *Lactococcus lactis* by the removal of essential genes, such as *thyA*<sup>19</sup>. The *thyA* gene encodes thymidylate synthase, an enzyme required for the synthesis of deoxythymidine monophosphate, which is converted to deoxythymidine triphosphate essential for DNA synthesis. Deficiency of *thyA* causes thymineless death unless thymine or thymidine is added to the growth media<sup>19</sup>. However, the thymidine-auxotrophic system has the potential risk that, depending on the genetic characteristics of the surrounding microbial community, the engineered organisms might acquire the essential gene from other bacteria by HGT<sup>20</sup> and escape their containment. Additionally, the dissemination of synthetic gene circuits to wild-type (WT) strains in nature is still a concern, even though gene flow barriers have been developed by introducing toxin-antitoxin gene systems onto artificial plasmids and bacterial chromosomes<sup>21</sup>. Genes on a plasmid are frequently transferred among bacteria; such transfer is a major cause of the prevalence of antibiotic-resistant bacteria<sup>22</sup>. Existing containment systems can be greatly improved to prevent genetically engineered microorganisms and synthetic genes from spreading.

*Bacteroides thetaiotaomicron* is an attractive cellular chassis for therapies in the human gut microbiota<sup>23</sup>. This species stably colonizes the mammalian gut and is prevalent and abundant in the human gastrointestinal tract<sup>24,25</sup>. It has a broad ability to interact with its host as well as various other microbes because it ferments polysaccharides and glycans<sup>26</sup> and metabolizes bile salts<sup>27</sup>. *Bacteroides* spp. have been reported as an effective treatment for colitis<sup>28,29</sup>, and genetic tools have already been developed to manipulate these cells to sense, record, and respond to a stimulus in the mammalian intestine<sup>30–32</sup>. However, the many potential applications of *B. thetaiotaomicron* have been stymied by the absence of a robust biocontainment system.

We have developed a biocontainment system in *B. thetaiotaomicron* that combines thymidine auxotrophy, an Engineered Riboregulator (ER) for controlled gene expression by the intended bacteria,

and a CRISPR Device (CD) to achieve the reliable containment of engineered strains. This system can degrade an essential gene that could be transferred into the engineered bacteria from other bacteria and can also prevent engineered genetic circuits from disseminating to environmental bacteria. In addition to demonstrating the functionality of this biocontainment system, we provide validation of its evolutionary stability *in vitro* and *in vivo*.

We used a two-step methodology to construct a genetically modified, human commensal thymidine-auxotrophic bacterium bearing the functions mentioned above by introducing a gene cassette into the bacterial genome via recombination. This method, which can also be applied to a range of human gut-associated bacteria, enables the biological containment system to function stably with the controlled expression of a gene of interest (GOI), thus expanding future clinical applications of genetically engineered bacteria.

## Results

### Biological containment design

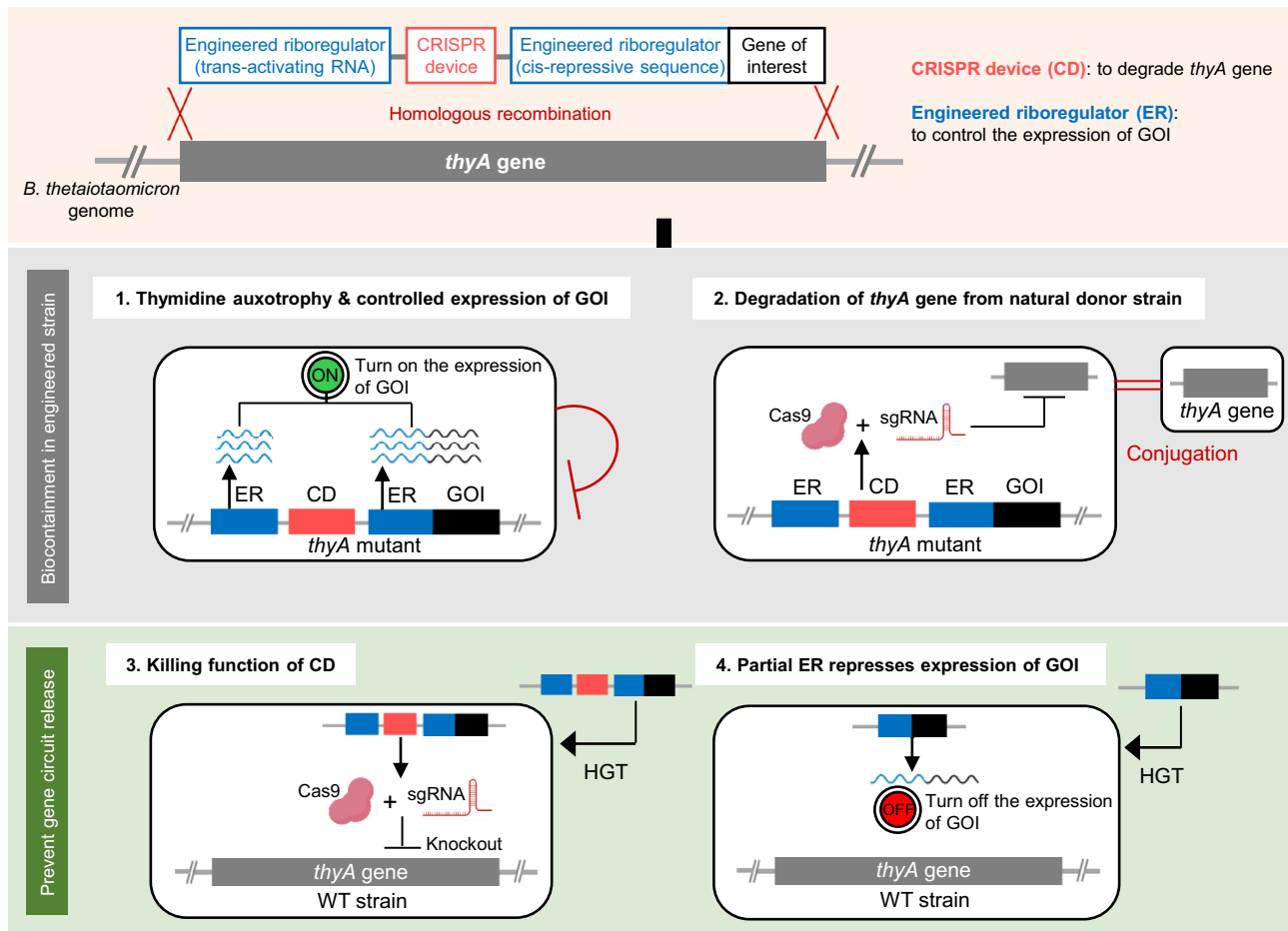
We developed a Cas9-assisted biological containment system that combines thymidine auxotrophy, an ER, and a CD to achieve controlled gene expression and biocontainment in engineered bacteria, and to prevent the release of the genetic circuit to other bacteria in the same environment (Fig. 1). The ER controls the level of gene expression of a GOI that is naturally repressed in all bacteria other than the engineered strain. Only the genetically modified strain can express the GOI; this strain has an intact ER, composed of trans-activating RNA (taRNA) and cis-repressed mRNA (crRNA) containing cis-repressive sequences (CRs) complementary to the taRNA. Without any taRNA, the crRNA solely represses the expression of the GOI<sup>33</sup>, which could be introduced by genetic transduction via bacteriophages. The level of gene expression is controllable by optimization of the ER structure. The CD has been reported as a sequence-specific bactericidal device that induces double-stranded breaks in a target DNA sequence<sup>34</sup>. This system can prevent both the acquisition of an essential gene (*thyA* gene) from environmental bacteria by degrading the gene, and the dissemination of transgenes to environmental bacteria by killing the bacteria harboring the artificial gene cassette with the CD. The containment system provides controlled expression, reducing the potential risk of loss of auxotrophy and dissemination of transgenes.

The containment system was generated through recombination of plasmids bearing the ER and the CD (Supplementary Fig. 1) into the genome of *B. thetaiotaomicron*. Upstream and downstream sequences of the *thyA* gene were introduced flanking the gene cassette, such that double crossover recombination would lead to the integration of the ER and CD and the deletion of the *thyA* gene. The taRNA, SpCas9 gene, CR, and NanoLuc gene as a reporter of the GOI were introduced to the WT *B. thetaiotaomicron* VPI-5482 strain as the first step, and single guide RNA (sgRNA) targeting the *thyA* gene was integrated as the second step (Supplementary Fig. 2a). We then evaluated the characteristics of NanoLuc-producing genetically modified *B. thetaiotaomicron* with the containment system.

### Repression and control of gene expression by the Engineered Riboregulator (ER)

The ER is composed of two components: the crRNA and the taRNA. The secondary structure in the 5' untranslated region (UTR) of the crRNA, which has a cis-repressive sequence (CR), represses the expression of the downstream gene by blocking access to the ribosome binding site (RBS). Sequences of taRNA containing a trans-activating sequence (TA) were optimized to hybridize to the crRNA and expose the RBS, initiating the translation of the NanoLuc reporter gene (Fig. 2a).

We first sought to build the crRNA structure to highly repress the expression of NanoLuc by inserting four types of CRs: crRCN, crR7N, crR10N or crR12N, into the RBS sequence (Supplementary Table 1). Previously, crRNAs had been constructed in *Escherichia coli* that



**Fig. 1 | Approach for biocontainment of the genetically modified strain and genetic elements.** Engineered Riboregulator (ER), CRISPR Device (CD), and Gene of Interest (GOI) are integrated into the genome of *B. thetaiotaomicron* by recombination of the sequence encoding these elements with the *thyA* gene. This sequence plays a role in biocontainment as follows: 1. Thymidine auxotrophy prevents the engineered strain from growing in the absence of thymidine. The ER

controls the expression of the GOI. 2. The CD degrades the *thyA* gene from the donor strain, maintaining thymidine auxotrophy. 3. The released CD kills the wild-type (WT) strain, preventing the dissemination of the gene circuit by knocking out the essential *thyA* gene in the genome. 4. When released to the WT strain, the partial ER represses the expression of the GOI. HGT horizontal gene transfer.

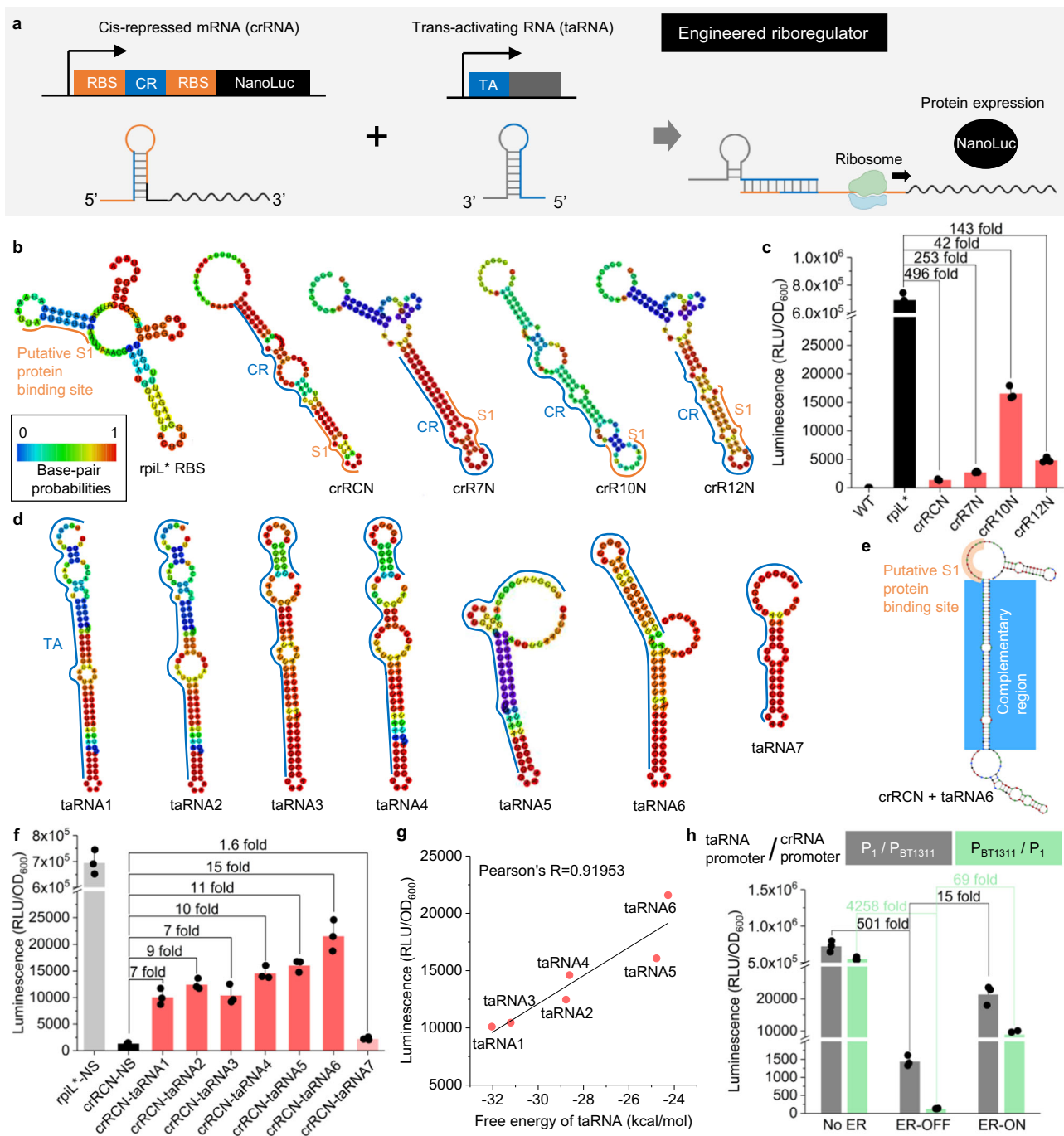
blocked the purine-rich RBS region upstream of the start codon by the formation of secondary structures<sup>33</sup>. Unlike *E. coli*, in which the homology of the Shine-Dalgarno sequence to the 16S rRNA largely dictates the strength of translation, the strength of gene expression in *Bacteroides* correlates poorly with the level of RBS complementarity to the 16S rRNA<sup>35,36</sup>. The RBS of *Bacteroides* tends to be AT-rich and more sensitive to secondary structure than that of *E. coli*<sup>31</sup>. The putative S1 protein binding site in the RBS was reported to have a strong effect on the strength of the RBS<sup>37</sup>. Therefore, we originally targeted this region to make crRNA structures. To design the ER, we used RNAfold WebServer, a server designed for the analysis of nucleic acid systems<sup>38</sup>. The generated RNAfold-predicted structures of crRNA showed various states of the putative S1 protein binding site with base-pair probabilities ranging from zero to one (Fig. 2b).

To evaluate the degree of repression of the GOI by crRNA, we tested the intensity of luminescence through enzymatic activity of NanoLuc. All strains except for the WT strain had luciferase activity (Fig. 2c). However, the activities of those which carried crRCN or crR7N were significantly repressed: 496-fold or 253-fold, respectively, compared to the strain without a CR (rpiL\*). This repression would be caused by the secondary structures of crRNA which block the putative S1 protein binding site, and their higher base-pair probabilities. The rpiL\* strain showed much higher output than crR10N even though the putative S1 protein binding site was blocked. Compared to crR10N

(green color), the secondary structure of rpiL\* (blue color) had the lowest probability of base-pairing (Fig. 2b), and it is possible that the rpiL\* might form a different open structure inside a cell.

Seven types of taRNA were then designed, by introducing inner loops and bulges to destabilize the structures, so that they had different Minimum Free Energy (MFE) and length complementary to crRNA (Supplementary Table 2). We thought that, in addition to the MFE, the mismatches would contribute to resistance to RNase III cleavage of RNA duplexes<sup>33</sup>. The structures of taRNA were analyzed using RNAfold WebServer (Fig. 2d), and the outputs of NanoLuc were measured. NUPACK web server was also utilized for the crRNA/taRNA complex. As shown in Fig. 2e and Supplementary Fig. 3a, crRNA/taRNA complexes had exposed putative S1 protein binding sites. crRCN and crR7N significantly repressed the expression of NanoLuc in the absence of any taRNA, while with taRNA, the NanoLuc activity of the strain that had taRNA6 (crRNA-taRNA6) significantly increased (15 fold; Fig. 2f). The stability of the taRNA affected gene expression, given that the MFE of taRNAs highly correlated with the outputs whereas the stability of the crRNA/taRNA complex did not show any correlation with gene expression levels (Fig. 2g; Supplementary Fig. 3b).

The outputs of this ER system are likely to increase at temperatures higher than the physiological one because the MFE of both taRNAs and crRNAs would rise, and vice versa (outputs would decrease at lower-than-physiological temperatures). In addition to the



**Fig. 2 | Engineered riboregulator (ER) system used to control the expression of the gene of interest in genetically modified *Bacteroides thetaioaomicron*.**

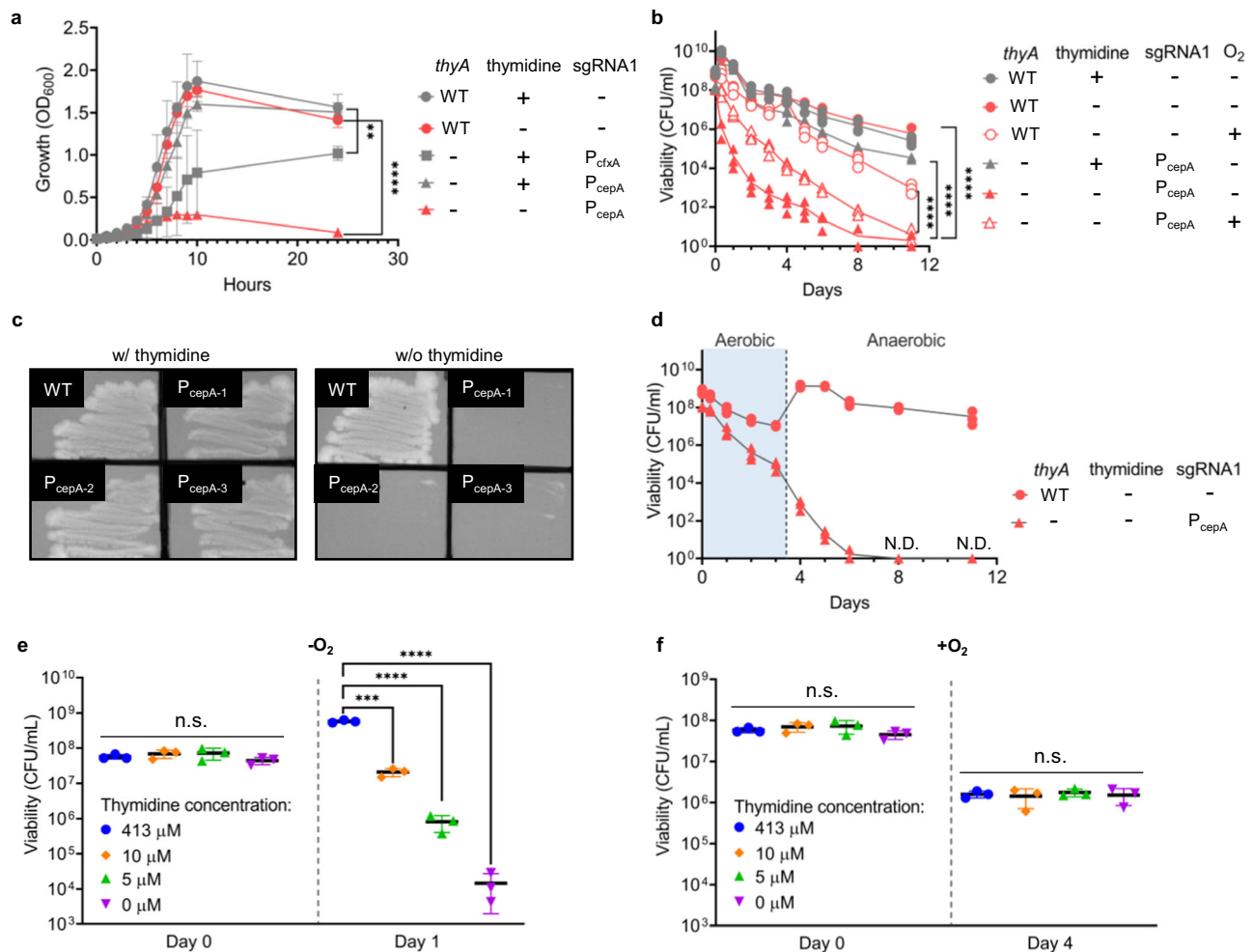
**a** Schematic of ER system. CR and TA represent cis-repressive and trans-activating sequences, respectively. **b** RNAfold-WebServer-predicted structures of cis-repressed mRNA (crRNA) variants. **c** Results of cis repression of crRNA variants. The luminescence values were normalized to OD<sub>600</sub> of the cell suspensions. Fold changes were calculated on the basis of the value with no CR (rpiL\*) and non-sense taRNA. **d** RNAfold-WebServer-predicted structures of trans-activating RNA (taRNA) variants. **e** NUPACK-predicted structure of crRNA (crRCN)/taRNA6 complex. **f** Outputs of NanoLuc for taRNA variants. NS represents non-sense taRNA.

**g** Correlation between outputs and free energy of folding for taRNA variants: taRNA1 to taRNA6. **h** ON/OFF outputs with P<sub>BT1311</sub> and P<sub>1</sub> promoters switched between crRNA and taRNA. Gray columns represent original promoters (P<sub>1</sub>/P<sub>BT1311</sub>) whereas green ones show switched promoters (P<sub>BT1311</sub>/P<sub>1</sub>). ER-OFF, with crRCN and NS; ER-ON, with crRCN and taRNA6. **(c, f, h)** Dots are three biological replicates made on different days and columns are the values of the mean. Error bars represent the standard deviations of the three biological replicates. **g** Dots show the luminescent values of the mean and Minimum Free Energy of taRNA. Source data are provided as a Source Data file.

temperature response of this ER system, a longer complementary sequence than that of 48 bps was required to activate the crRNA. Weaker activation of crRCN by taRNA7 was observed, a taRNA which was designed so that it would have a shorter region of complementarity (26 bps) than the others (48 bps). In contrast, the

resistance of RNA duplexes to RNase III cleavage was not critical to the output, though taRNA3 was built to test resistance by adding mismatches to the RNA/RNA complex.

To further improve the ON/OFF ratio of the gene expression level, we switched the promoters (P<sub>BT1311</sub> or P<sub>1</sub>) controlling taRNA and crRNA



**Fig. 3 | Thymidine auxotrophy of genetically modified *B. thetaiotaomicron*.**

**a** Growth curves of wild-type (WT) and *thyA*-deficient strains with  $P_{cfxA}$ -sgRNA1 and  $P_{cepA}$ -sgRNA1 in the presence (413  $\mu$ M) and absence of thymidine. **b** Cell viability of WT and *thyA*-deficient strain with  $P_{cepA}$ -sgRNA1 in the presence (413  $\mu$ M) and absence of thymidine under anaerobic and aerobic conditions. **c** Restreaked colonies of the *thyA*-deficient strain with  $P_{cepA}$ -sgRNA1 on TYG with thymidine (413  $\mu$ M) and without thymidine, after anaerobic culturing without thymidine for 11 days. **d** Cell viability of WT and *thyA*-deficient strain with  $P_{cepA}$ -sgRNA1 when cells were transferred from aerobic to anaerobic conditions in the absence of thymidine. **e** Cell viability of *thyA*-deficient strain with  $P_{cepA}$ -sgRNA1 at various concentrations of thymidine in the anaerobic condition. **f** Cell viability of the *thyA*-deficient strain

with  $P_{cepA}$ -sgRNA1 at various concentration of thymidine in the aerobic condition. **a** Each dot is the value of the mean, and error bars represent the standard deviations of three biological replicates made on different days (Tukey's multiple comparisons test with one-way analysis of variance on data at 24 h; *thyA* with thymidine and  $P_{cfxA}$ -sgRNA1 vs. *thyA* with thymidine and  $P_{cepA}$ -sgRNA1:  $P = 0.0010$ ,  $**P < 0.01$ ,  $****P < 0.0001$ ). **b–f** Each dot is a biological replicate ( $n = 3$ ), and lines and central bars are the values of the mean. Error bars represent standard deviations (Tukey's multiple comparisons test with one-way analysis of variance on log-transformed data; **e** 413  $\mu$ M vs. 10  $\mu$ M at Day1:  $P = 0.0005$ ,  $***P < 0.001$ ,  $****P < 0.0001$ ). The detection limit is 1 CFU/ml. n.s. represents not significant. N.D. represents no detection. Source data are provided as a Source Data file.

expression (Supplementary Fig. 2a).  $P_{BT1311}$  has been reported to be a stronger promoter than  $P_1$ <sup>31</sup>. We expected gene expression to be repressed more strongly without complementary taRNA (OFF state) if the weaker promoter was placed upstream of crRNA, whereas we expected that the expression level would be retained with taRNA6 (ON state) if the stronger promoter was placed upstream of the taRNA sequence. As shown in Fig. 2h, the strain bearing the switched promoters ( $P_{BT1311}/P_1$ ) had a remarkably improved ON/OFF ratio. This taRNA elicited a  $\sim 70$ -fold change in gene expression between the ON and OFF states.

### Effect of CRISPR Device (CD) on cell growth

We next investigated whether the CD, composed of the SpCas9 gene and sgRNA, affected the cell growth of the genetically modified thymidine-auxotrophic *B. thetaiotaomicron* strain. Two sets of sgRNAs and promoter sequences ( $P_{cfxA}$ -sgRNA1 and  $P_{cepA}$ -sgRNA1) were introduced, via the second double crossover, into the

strain having crRCN and taRNA6 as an ER (Supplementary Fig. 2).  $P_{cfxA}$ -sgRNA1 targets the *thyA* gene with the stronger promoter  $P_{cfxA}$ . In fact, the  $P_{cfxA}$  promoter is approximately 10-fold stronger than the  $P_{cepA}$  promoter<sup>31</sup>.

We next tested the growth curves of the strains in the presence and absence of thymidine. With thymidine, the  $P_{cfxA}$ -sgRNA1 strain grew more slowly than the other strains (Fig. 3a), possibly because the high concentration of sgRNA was harmful to this strain. The high concentration of the sgRNA might have increased the frequency of small-RNA-mediated targeting of mRNA<sup>39,40</sup>. On the other hand, we also observed that with thymidine the  $P_{cepA}$ -sgRNA1 strain grew rapidly, at the rate close to that of the WT strain, but without thymidine this strain grew just a little and then died; the initial growth would have been enabled by the thymidine retained in the cells before the assay began. Overall, high expression of this sgRNA can lead to toxicity and inhibition of cell growth.

## Viability of genetically modified thymidine-auxotrophic *B. thetaiotaomicron*

To validate thymineless death as a method of biological containment for the genetically modified strain, we measured cell viability anaerobically and aerobically. Under anaerobic conditions, the cell concentration of the WT strain increased to over  $1 \times 10^9$  CFU/mL within 8 h and then gradually decreased as a result of lack of nutrients or the accumulation of waste products. Nevertheless, the cell concentration remained higher than  $1 \times 10^5$  CFU/mL for 11 days. Likewise, the  $P_{cepA}$ -sgRNA1 strain with thymidine showed large numbers of living cells throughout the culture (Fig. 3b). On the other hand, we observed an approximately  $10^4$ -fold decrease in CFU/mL of the genetically modified strain without any sgRNA and a  $10^6$ -fold decrease in the  $P_{cepA}$ -sgRNA1 strain 6 days after incubation without thymidine (Supplementary Fig. 4). The CD seemed to help increase the death rate, which might be derived from its toxicity. When the culturing was prolonged up to 11 days, the CFU/mL of  $P_{cepA}$ -sgRNA1 strain decreased by  $10^7$ – $10^8$  (Fig. 3b). We streaked three colonies of the strain obtained at day 11 onto a TYG (trypticase-yeast extract-glucose) agar plate with and without thymidine to determine whether the viable cells were still thymidine auxotrophs. As shown in Fig. 3c, these cells showed thymidine auxotrophy and would die without any growth even if the viability test was extended after day 11.

When the  $P_{cepA}$ -sgRNA1 strain was cultured aerobically, the rate of decrease in viable cells was repressed (Fig. 3b). The cell growth, or DNA replication, of this anaerobe is inhibited in the aerobic state, which would have repressed the thymineless death<sup>41</sup>. This result was consistent with a previous report about *Bacteroides ovatus*<sup>20</sup>. However, our strain decreased in cell numbers more rapidly than the WT strain and did not grow any more, even if it was transferred from the aerobic state to the anaerobic state (Fig. 3b, d).

Viability was tested at various concentrations of thymidine to clarify the concentration dependency of thymineless death (Fig. 3e, f). The decrease in viable cells was not caused by thymineless death in the aerobic state, but possibly by stress from the sgRNA and/or oxygen. On the other hand, under anaerobic conditions, this strain died rapidly at concentrations of thymidine less than 5  $\mu$ M.

The inhibition of thymineless death was also observed when the  $P_{cepA}$ -sgRNA1 strain was subjected to other stressful conditions. To test the effect of varying the carbon source, we measured the growth curve and viability of the engineered *B. thetaiotaomicron* in the presence of monosaccharide (i.e., glucose and galactose), disaccharide (i.e., maltose), and polysaccharide (i.e., starch), with or without thymidine (Supplementary Fig. 5a, b). Engineered *B. thetaiotaomicron* grew rapidly in the presence of thymidine with either a monosaccharide or a disaccharide as the carbon source. On the contrary, no viable cells were detected in the absence of thymidine after 48 h of incubation with these carbon sources (Supplementary Fig. 5b). The growth of the strain was much slower when starch was used as the sole carbon source with thymidine than when a monosaccharide or a disaccharide was used (Supplementary Fig. 5b). With starch as the sole carbon source, after 96 h of incubation in the absence of thymidine significantly lower numbers of viable cells were detected, compared to those with thymidine, (Supplementary Fig. 5b).

We also tested the effect of carbon source starvation on the auxotrophy in the engineered *B. thetaiotaomicron* (Supplementary Fig. 5c). Bacterial cells were cultured in a defined minimal medium (DMM) without glucose after a PBS wash in the presence or absence of thymidine. As expected, the cells stopped growing without a carbon source. Moreover, the viability of strains in the absence of thymidine significantly decreased after 48 and 96 h of incubation, indicating that thymidine auxotrophy persisted in conditions of carbon source starvation.

Considering the important roles and ubiquitous presence of bile acids in the human gut, we tested the effect of crude bile acid

and secondary bile acid (i.e., deoxycholic acid, or DCA) on the auxotrophy of engineered *B. thetaiotaomicron* (Supplementary Fig. 5d, e). The bacteria stopped growing without thymidine (Supplementary Fig. 5d), and no viable cells could be detected after 48 h of incubation in the absence of thymidine in TYG medium supplemented with either crude bile acid or DCA (Supplementary Fig. 5e).

In summary, as the levels of carbon sources and bile acids fluctuate in the gut, their presence or absence can contribute to conditions of stress. Although one carbon source (starch) could delay thymineless death, our engineered strain of *B. thetaiotaomicron* nevertheless maintained its auxotrophy, indicating the robustness and efficiency of our biocontainment.

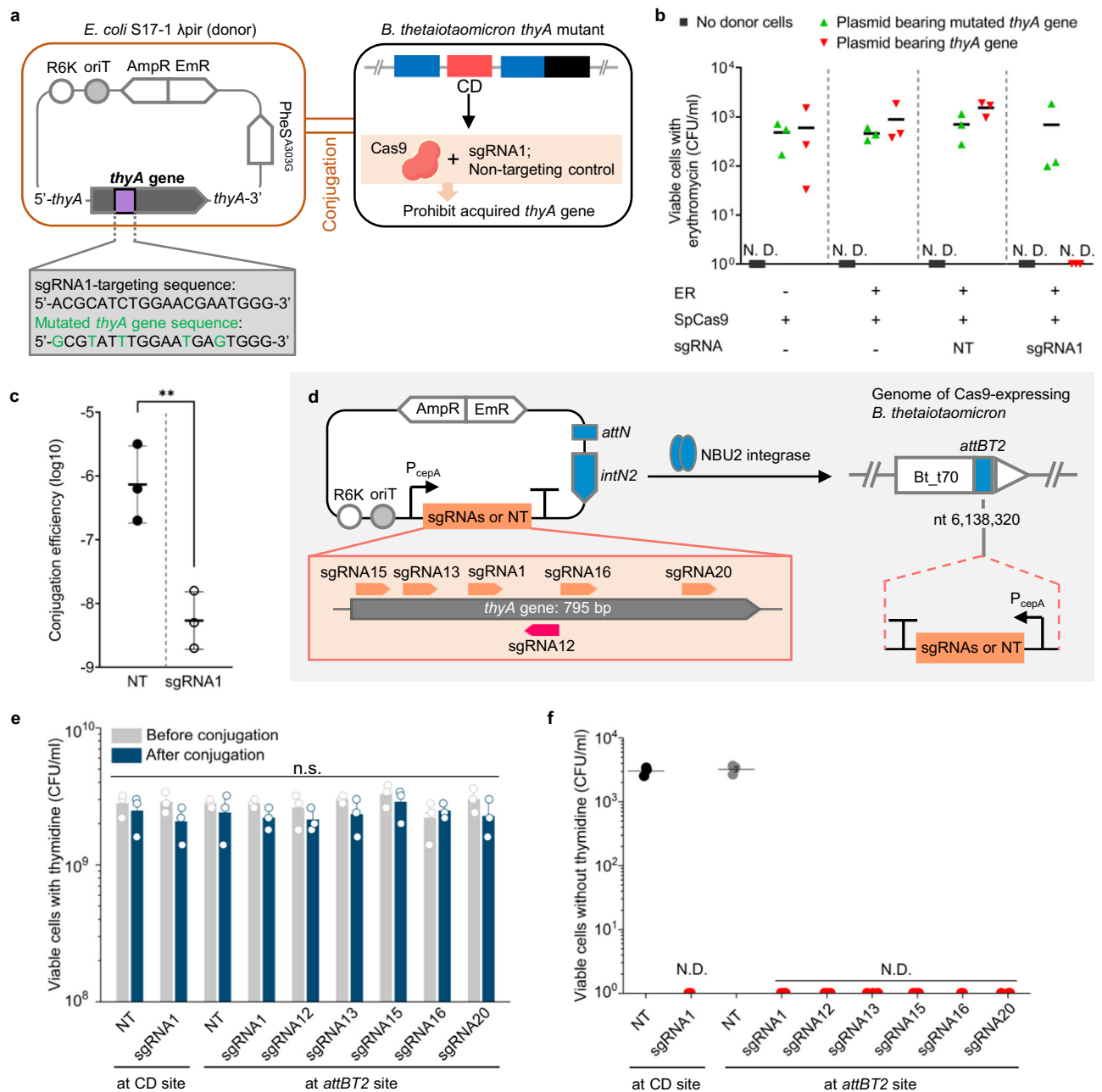
## Blocking acquisition of *thyA* gene by genetically engineered bacteria

Next, we tested whether the CD would prevent the genetically engineered strain from acquiring the *thyA* gene from another bacterial strain. When the *E. coli* S17-1  $\lambda$  pir strain, carrying a plasmid bearing the intact or mutated *thyA* gene, was mixed with the genetically modified *B. thetaiotaomicron* strain, the plasmid transferred to the *Bacteroides* strain by conjugation (Fig. 4a).

To determine whether the CD carried by the genetically modified strains of *B. thetaiotaomicron* would degrade the acquired *thyA* gene, several genetically modified *Bacteroides* strains with different ERs and CDs were constructed and tested for their ability to prevent the acquisition of the *thyA* gene. One strain had neither ER nor sgRNA; others had the ER with or without sgRNAs, including the non-targeting control. Plasmids bearing either an intact or a mutated *thyA* gene were generated by introducing synonymous codons into the sgRNA1-targeting sequence. Filter mating between *B. thetaiotaomicron* and *E. coli* with the plasmids was performed on TYG agar plates in the absence of thymidine.

As shown in Fig. 4b, *B. thetaiotaomicron* transconjugants were detected on BHIS agar plates with erythromycin and thymidine, indicating that these strains had acquired the plasmid-borne mutated *thyA* gene that was not targeted by a CD. Additionally, the strain with the CD having the non-targeting control, which did not degrade the *thyA* gene, had obtained the intact *thyA* gene and grew on the BHIS plates. In contrast, no colonies were observed of the bacterial strain having the CD that targeted the intact, plasmid-borne *thyA* gene. This result demonstrated that the CD can specifically recognize and destroy *thyA*, blocking the effect of HGT of the plasmids. The ability to disable the gene that would otherwise correct the auxotrophy after HGT has occurred is a requirement for safe biological containment.

We also conducted a conjugation experiment on TYG agar plates in the presence of thymidine. Some viable erythromycin-resistant cells having sgRNA1 were detected, in contrast to results obtained in the filter mating without thymidine. These cells would be derived from the higher numbers of viable cells during conjugation compared to the conjugation in the absence of thymidine, which leads to higher numbers of transconjugants. Nevertheless, the conjugation efficiency of the plasmid having an intact *thyA* gene was significantly decreased, at least 156-fold, by the CD (Fig. 4c). The colonies selected with an erythromycin resistance gene were still thymidine auxotrophic and did not have any *thyA* gene since no amplicon was detected by PCR using primers binding to the *thyA* gene (Supplementary Fig. 6a, b). Whole genome sequencing revealed that the backbone of the plasmid bearing the intact *thyA* gene (pNH9223) had integrated downstream of the NanoLuc gene without any integration of *thyA* gene in the transconjugant (Supplementary Fig. 6c; BioSample accession number: SAMN20797097). Therefore, the actual frequency of *thyA* gene transfer would be less than the frequency of erythromycin resistance gene transfer. These results also indicated that the CD could also repress the acquisition of the genes adjacent to the targeted genes.



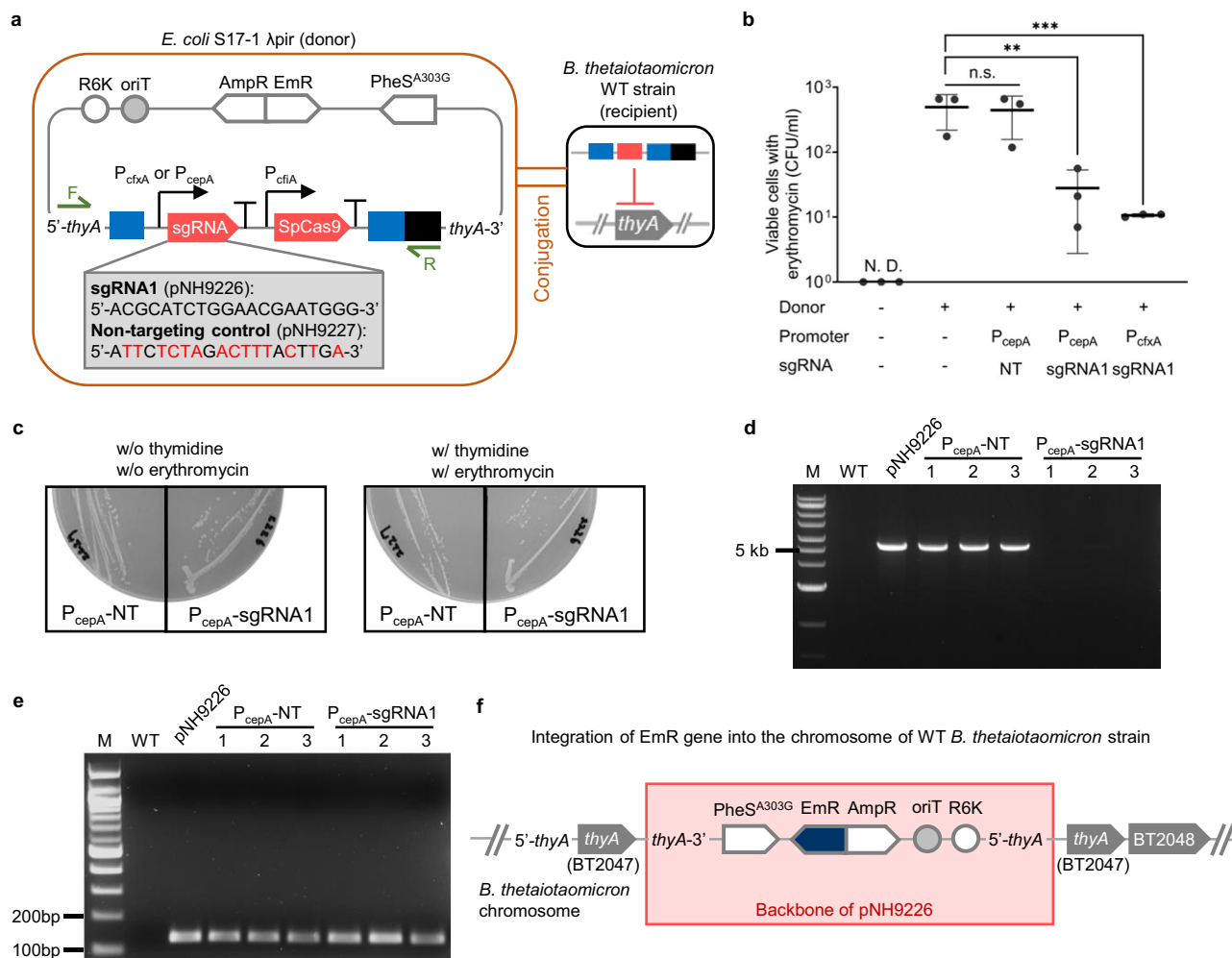
**Fig. 4 | Prevention of acquisition of *thyA* gene by genetically engineered strain.**

**a** Schematic of plasmid construct bearing the *thyA* gene and Horizontal Gene Transfer (HGT) of the *thyA* gene by conjugation from *E. coli* to genetically engineered thymidine-auxotrophic *B. thetaiotaomicron*. The *E. coli* strains have either an intact or a mutated *thyA* gene on their plasmids. The *B. thetaiotaomicron* strains have either the sgRNA1 sequence or a non-targeting control (NT). R6K origin of replication, oriT origin of transfer, AmpR ampicillin resistance cassette, EmR erythromycin resistance cassette, PheS<sup>A303G</sup>, mutated  $\alpha$ -subunit of phenylalanyl-tRNA synthetase gene; *intN2*, tyrosine integrase cassette; *attN*, recombination site with *attBT2* sites on the genome of *B. thetaiotaomicron*. **b** Viable *B. thetaiotaomicron* cells following conjugation on TYG agar plates without thymidine and plating on BHIS agar plates with gentamicin, thymidine, and erythromycin. **c** Conjugation efficiency of *thyA* gene through conjugation on TYG agar plates containing thymidine. The efficiency was calculated by dividing CFU/ml on the BHIS agar plates with

gentamicin, thymidine, and erythromycin by total CFU/ml (two-tailed unpaired Student's t-test on log-transformed data; NT vs. sgRNA1:  $P = 0.0080$ ,  $**P < 0.01$ ). **d** Schematic of integration of a plasmid bearing each of seven types of sgRNA sequences, including NT on the genome of SpCas9-expressing *B. thetaiotaomicron* by pNBU2 technology. **e** Cell viability of *Bacteroides* before and after conjugation between the *E. coli* and genetically engineered thymidine-auxotrophic *B. thetaiotaomicron* with various types of sgRNAs (two-tailed unpaired Student's *t*-test on log-transformed data between before and after conjugation). **f** Viable *B. thetaiotaomicron* when cells were plated on TYG agar plates without thymidine after conjugation on TYG agar plates with thymidine. Three biological replicates were made on the different days. Black bars and gray columns are the values of the mean. Error bars are the standard deviations. The detection limit is 1 CFU/ml. n.s. represents not significant. N.D. represents no detection. Source data are provided as a Source Data file.

Next, we sought to clarify whether the capability to prevent the HGT of the *thyA* gene depends on the targeted sequences in the gene. Genetically modified thymidine-auxotrophic strains bearing the CD were constructed by integration of pNBU2-based plasmids with sgRNAs

and the erythromycin resistance gene after the first recombination<sup>31</sup>; these strains each carried one of the seven types of the sgRNAs (Fig. 4d; Supplementary Table 3) and their promoter P<sub>cepA</sub> at one of two *attBT2* sites located in the 3'-ends of the two tRNA<sup>Ser</sup> genes, BT\_t70 and BT\_t71,



**Fig. 5 | Prevention of dissemination of transgenes by genetically engineered strain.** **a** Schematic of plasmid construct bearing transgenes and Horizontal Gene Transfer (HGT) of this plasmid by conjugation from *E. coli* to wild-type (WT) *B. thetaiotaomicron*. The *E. coli* strains have either sgRNA1 or a non-targeting control (NT) on their plasmids. The *B. thetaiotaomicron* strain has an intact *thyA* gene. R6K, origin of replication; oriT, origin of transfer; AmpR, ampicillin resistance cassette; EmR, erythromycin resistance cassette; PheS<sup>A303G</sup>, mutated  $\alpha$ -subunit of phenylalanyl-tRNA synthetase gene **b**, Viable *B. thetaiotaomicron* when cells were plated on BHIS agar plates with gentamicin, thymidine, and erythromycin after conjugation on TYG agar plates without thymidine. **c** Restreaked colonies on TYG with and without thymidine and erythromycin. **d** Amplicon patterns of three transconjugants isolated from TYG agar plates with thymidine and erythromycin

after conjugation. PCR was performed with primers shown in (a), which are named Seq-*thyA*-F and mmd663 in Supplementary Table 6. **e** PCR amplification of EmR gene fragment (144 bp) in transconjugants with primers named qPCR-EmR-F and qPCR-EmR-R in Supplementary Table 6. **f** Integration of pNH9226 backbone into the chromosome of WT *B. thetaiotaomicron* strain. Three biological replicates were made on the different days. Black bars are the values of the mean. Error bars are the standard deviations (Tukey's multiple comparisons test with one-way analysis of variance on log-transformed data; *E. coli* having P<sub>cepA</sub>-sgRNA1, and *E. coli* having P<sub>cfxA</sub>-sgRNA1 vs. *E. coli* without any promoter or sgRNA:  $P = 0.0028$ , and  $P = 0.0007$ ,  $**p < 0.01$ ,  $***p < 0.001$ ). The detection limit is 1 CFU/ml. n.s. represents not significant. N.D. represents no detection. Source data are provided as a Source Data file.

on the *B. thetaiotaomicron* chromosome. These strains showed thymidine auxotrophy and antibiotic resistance when they were plated on gentamicin-containing TYG agar plates either with thymidine and erythromycin, or without either (Supplementary Fig. 7). *B. thetaiotaomicron* strains with different *thyA*-targeting sgRNAs displayed similar viability in the presence of thymidine before and after conjugation (Fig. 4e). The filter-mating experiment using thymidine auxotrophy as a selective marker revealed that all the strains had the capacity to avoid the HGT of the *thyA* gene (Fig. 4f). The choice of targeted sequence did not seem to highly affect the capability to degrade the *thyA* gene. Furthermore, the CD worked well even though the sgRNAs were integrated at a locus far from the SpCas9 gene. Thus, there seems to be no requirement to target specific sequences of the *thyA* gene in this containment system or to integrate sgRNAs at specific sites, though further research on the effect of the targeted sequence and sgRNA site on disabling the *thyA* gene is still needed.

### Prevention of dissemination of transgenes by genetically engineered bacteria

We evaluated whether the CD could block the HGT of transgenes, such as the CD, ER, and GOI, through conjugation between *E. coli* and *B. thetaiotaomicron*. The donor strain, *E. coli* S17-1  $\lambda$ pir, which carries a plasmid bearing the transgenes (shown in Fig. 5a), was mixed with WT *B. thetaiotaomicron*, the recipient. The plasmids with the CD were expected to degrade the *thyA* gene on the recipient genome so that transconjugants would not survive.

When we selected the transconjugants on BHIS agar plates with gentamicin, thymidine, and erythromycin, the number of transconjugants that had received the plasmid with the CD containing P<sub>cepA</sub> and P<sub>cfxA</sub> was markedly lowered, by approximately 18 and 46 times, respectively, compared to that without either promoter or sgRNA (Fig. 5b). This result indicated that the *thyA* gene on the recipient genome had been destroyed by the CD and that most of the recipients that



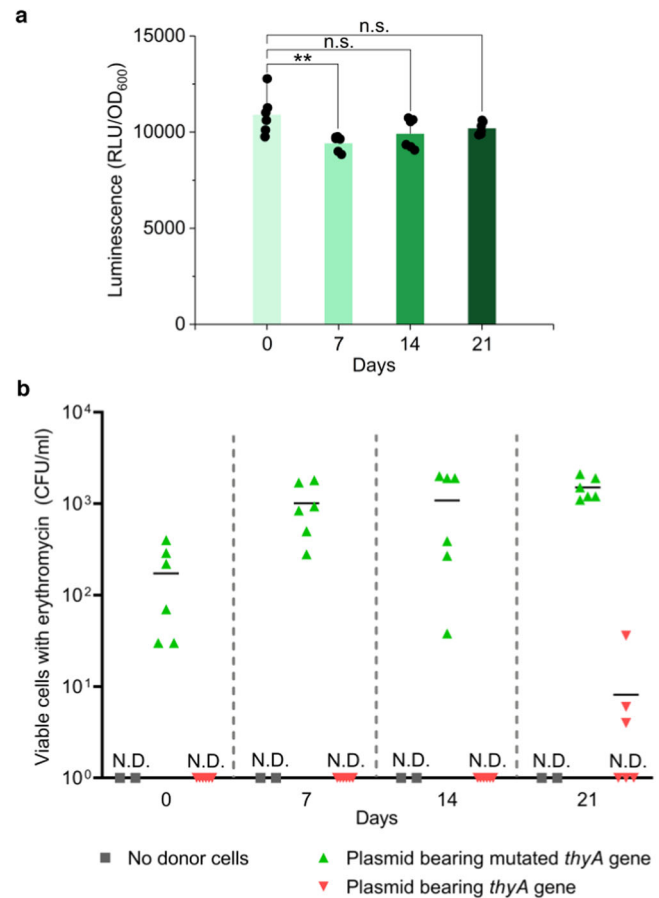
had gotten the plasmid died because their genome was not repaired. The stronger promoters are utilized for the transcription of sgRNA1, the more strictly the HGT is likely to be regulated, as fewer viable cells with  $P_{cfxA}$  were detected on the BHIS agar with erythromycin than cells with  $P_{cepA}$ , though the strength of the promoters did affect cell growth, as described. The appropriate choice of promoter can, therefore, be helpful to avoid the dissemination of the transgenes.

When the colonies of transconjugants were selected with erythromycin and restreaked on TYG in the presence or absence of both thymidine and erythromycin, they showed erythromycin resistance and thymidylate synthase activity regardless of the sgRNAs (Fig. 5c). Thus, the transconjugants with  $P_{cepA}$ -sgRNA1 had acquired the erythromycin resistance gene while retaining the *thyA* gene. To clarify why these transconjugants had an intact *thyA* gene after transfer of the plasmid bearing the erythromycin resistance gene and the CD, the region of transgenes in the transconjugant genomes was amplified by PCR using the primers in Fig. 5a. Interestingly, the DNA purified from transconjugants obtained by conjugation with the *E. coli* strain having  $P_{cepA}$  and sgRNA1 did not have the transgenes, whereas transgenes containing  $P_{cepA}$  and the non-targeting control were detected in the genome of WT *B. thetaiotaomicron* after conjugation (Fig. 5d). As PCR amplification showed, the recipients indeed had acquired the erythromycin resistance gene (Fig. 5e). Whole genome sequencing revealed that only the backbone of the plasmid with the erythromycin resistance gene, but not the CD, was integrated into the genome of the WT recipient strain (Fig. 5f; BioSample accession number: SAMN20797095). The recipient may have received the erythromycin resistance gene and released the transgenes containing the CD because this portion of the transferred plasmid was detrimental to the WT strain. Therefore, the actual frequency of transgenes transferred to the recipient strains would be lower than the value calculated as erythromycin-resistant transconjugants.

### Evolutionary stability of genetic elements

To evaluate stability, we tested how long the ER and CD remained stable during cell growth in vitro. The genetically modified strain with those functions was monocultured up to 21 days and checked for gene expression of NanoLuc by the ER and the prevention of HGT of the *thyA* gene by the CD every 7 days.

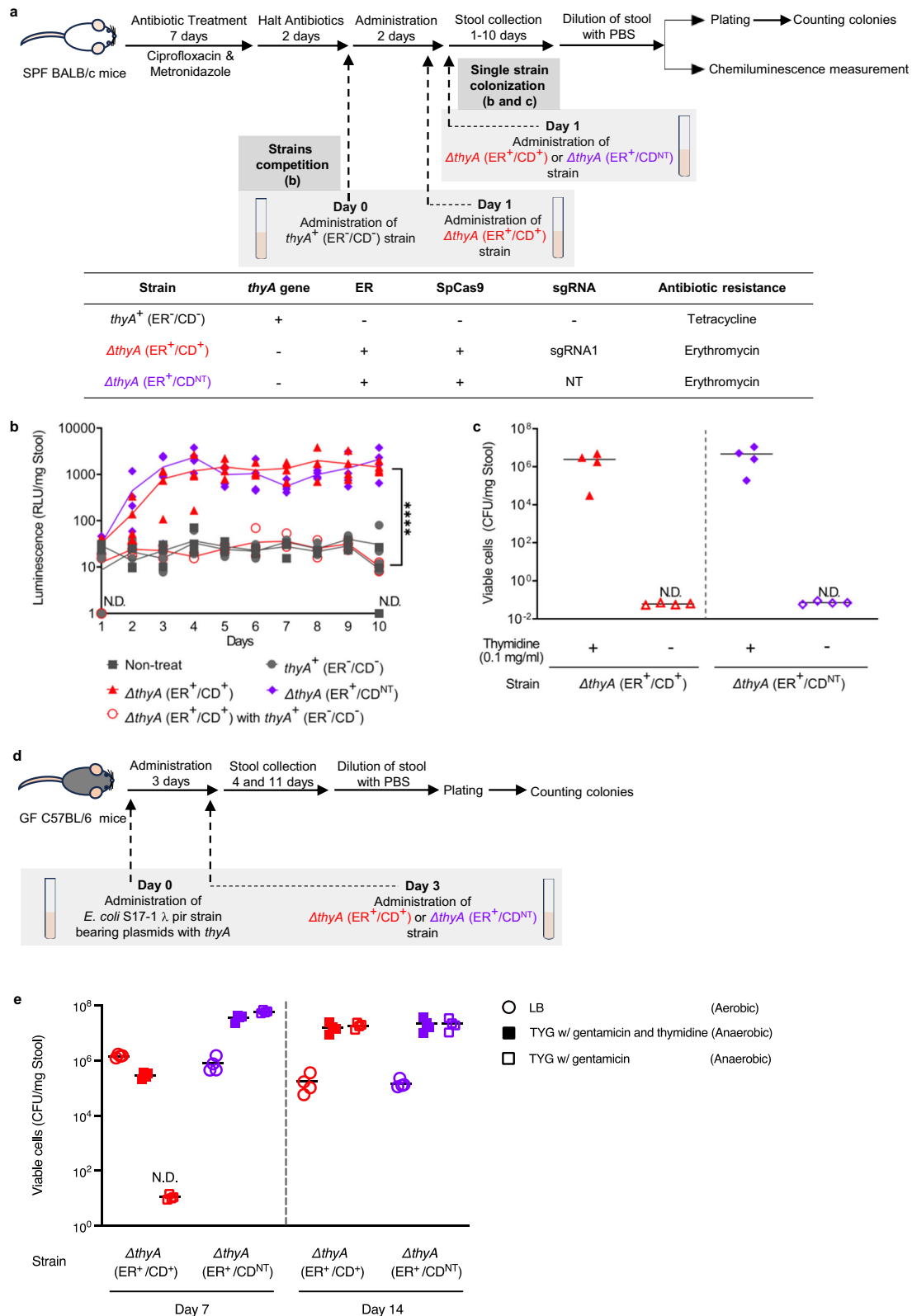
Over the time course of the evolutionary stability test, the expression of NanoLuc was maintained, resulting from the stability of the ER (Fig. 6a). The genetically modified *Bacteroides* strain could receive the plasmid bearing the mutated *thyA* gene, as colonies were detected throughout the period of this experiment. On the other hand, bacterial colonies that had obtained the plasmid having an intact *thyA* gene were not observed for the first 14 days. Viable erythromycin-resistant cells appeared at day 21 (Fig. 6b). Given that there was no erythromycin-resistant colony without any donor cells (the negative control) and there was no detection of the erythromycin resistance gene by PCR in the recipients before conjugation (Supplementary Fig. 8a), we can conclude that the viable cells were not derived from spontaneous mutation during the 21 days of culture. The integration of the plasmid bearing the erythromycin resistance and the intact *thyA* genes was observed by whole genome sequencing (Supplementary Fig. 8b; BioSample accession number: SAMN20797096). Therefore, the CD was stable for at least 14 days under these growth conditions. We assumed that some mutations may have occurred in the sequence of the CD during cell growth and that the CD may have lost functionality after day 14. However, Sanger sequencing revealed no mutation in the region of the CD for thirteen viable colonies grown on BHIS agar with gentamicin, thymidine, and erythromycin. The CD may have lost its functionality via another mechanism, such as a mutation at a locus outside the CD region, though this is still unclear. Overall, this engineered strain was stable for 14 days, which is equivalent to at least 240 cell divisions, corresponding to a  $10^{7.2}$ -fold amplification of the initial inoculant.



**Fig. 6 | In-vivo genetic stability of the genetically engineered strain bearing Engineered Riboregulator (ER) and CRISPR Device (CD).** **a** Outputs of NanoLuc measured every 7 days until day 21. Three biological replicates were made twice. Dots are six biological replicates and columns are the values of the mean. Error bars represent the standard deviations of the six biological replicates (Tukey's multiple comparisons test with one-way analysis of variance on data; day 0 vs. day 7:  $P = 0.0088$ ,  $**P < 0.01$ ). **b** Viable *B. thetaiotaomicron* when cells were plated on BHIS agar plates with thymidine, erythromycin, and gentamicin every 7 days until day 21 after conjugation on TYG agar plates without thymidine. Three biological replicates were made twice. Individual dots represent six biological replicates of recipients, whereas two biological replicates were made for No donor cells (control). Black bars are the values of the mean. The detection limit is 1 CFU/ml. n.s. represents not significant. N.D. represents no detection. Source data are provided as a Source Data file.

### In vivo colonization of the genetically modified *B. thetaiotaomicron* and function of the Engineered Riboregulator (ER), thymidine auxotrophy, and CRISPR Device (CD)

To evaluate whether the genetically modified thymidine-auxotrophic bacteria can colonize mice and whether the ER and thymidine auxotrophy function in vivo, three *B. thetaiotaomicron* strains were administered to specific-pathogen-free (SPF) mice: a *thyA*-deficient strain with the ER and CD,  $\Delta thyA$  (ER<sup>+</sup>/CD<sup>+</sup>); a *thyA*-deficient strain with the ER and the non-targeting CD,  $\Delta thyA$  (ER<sup>+</sup>/CD<sup>NT</sup>); and a *thyA*-intact strain without either, *thyA*<sup>+</sup> (ER/CD<sup>-</sup>). Stools were collected for 10 days to quantify NanoLuc luminescence and viable cells (Fig. 7a). As shown in Fig. 7b, when mice were gavaged with either the  $\Delta thyA$  (ER<sup>+</sup>/CD<sup>+</sup>) or the  $\Delta thyA$  (ER<sup>+</sup>/CD<sup>NT</sup>) strain after 7 days of the antibiotic treatment, the intensity of luminescence increased and then remained stable up to day 10. Thus, the genetically modified thymidine-auxotrophic strains had stably colonized the mice and the ER worked well in vivo. High concentrations of thymidine in the mouse small intestine, which have been reported to be on the order



of several hundred  $\mu\text{M}^{42}$ , may have supported the colonization of the thymidine-auxotrophic strains.

A competition experiment between the  $\Delta\text{thyA}$  (ER<sup>+</sup>/CD<sup>+</sup>) or the *thyA*<sup>+</sup> (ER<sup>-</sup>/CD<sup>-</sup>) strains was also performed. Once the *thyA*-intact strain *thyA*<sup>+</sup> (ER<sup>-</sup>/CD<sup>-</sup>) colonized after the antibiotic treatment, there was inhibition of colonization by  $\Delta\text{thyA}$  (ER<sup>+</sup>/CD<sup>+</sup>), and the output of NanoLuc disappeared. The in vitro growth curve showed that *thyA*<sup>+</sup>

(ER<sup>-</sup>/CD<sup>-</sup>) grew faster than  $\Delta\text{thyA}$  (ER<sup>+</sup>/CD<sup>+</sup>) (Supplementary Fig. 9). The difference in the growth rate, as well as the space occupied by *thyA*<sup>+</sup> (ER<sup>-</sup>/CD<sup>-</sup>) and the competition for nutrients, may have inhibited stable colonization by  $\Delta\text{thyA}$  (ER<sup>+</sup>/CD<sup>+</sup>).

To test the function and stability of thymidine auxotrophy in vivo, viable cells in fecal pellets at day 10 from the  $\Delta\text{thyA}$  (ER<sup>+</sup>/CD<sup>+</sup>)-mono-administered and  $\Delta\text{thyA}$  (ER<sup>+</sup>/CD<sup>NT</sup>)-mono-administered mice were

**Fig. 7 | In vivo colonization, function of Engineered Riboregulator (ER), thymidine auxotrophy and CRISPR Device (CD).** **a** Experimental timeline. Specific-pathogen-free BALB/c mice were treated for 7 days with ciprofloxacin and metronidazole. *B. thetaiotaomicron* strains were administered by oral gavage on Day 0 or Day 1. The thymidine-auxotrophic *B. thetaiotaomicron* strains with ER and SpCas9 have either the sgRNA1 sequence or a non-targeting control (NT). **b** Outputs of NanoLuc in feces of mice (gavaged with *B. thetaiotaomicron* strains as indicated) measured until day 10. Each dot represents a biological replicate, and lines are the values of the mean calculated from the four replicates (Dunnett's T3 multiple comparisons test with Welch's one-way analysis of variance on log-transformed data at day10, \*\*\*\* $P < 0.0001$ ). **c** Viable *B. thetaiotaomicron* when fecal suspensions from mice gavaged with indicated strains without *thyA*<sup>+</sup> (ER<sup>+</sup>/CD<sup>+</sup>) strain were plated on selective agar plates. Black bars are the values of the mean of four biological

replicates. The average detection limits are  $6.0 \times 10^{-2}$  CFU/mg Stool for *ΔthyA* (ER<sup>+</sup>/CD<sup>+</sup>) and  $7.1 \times 10^{-2}$  CFU/mg Stool for *ΔthyA* (ER<sup>+</sup>/CD<sup>NT</sup>), respectively. **d** Experimental timeline to test the function of CD in vivo. Germ-free C57BL/6 mice were gavaged with the *E. coli* S17-1λ pir with the *thyA* gene on a mobilizable plasmid on Day 0, and then with *B. thetaiotaomicron* strains bearing ER and CD on Day 3. The *B. thetaiotaomicron* strains have either the sgRNA1 sequence or a non-targeting control (NT) in the sequence region of CD. **e** Viable *B. thetaiotaomicron* and *E. coli* when fecal pellets collected from mice gavaged with indicated strains on Day 7 and Day 14 were plated on selective agar plates and incubated in aerobic or anaerobic condition. Black bars indicate the values of the mean of four biological replicates ( $n = 4$ ). The average detection limit is  $1.1 \times 10^1$  CFU/mg Stool for *ΔthyA* (ER<sup>+</sup>/CD<sup>+</sup>). N.D. represents no detection. Source data are provided as a Source Data file.

enumerated by plating the fecal homogenate on selective TYG agar plates without thymidine (for cells escaping from thymidine auxotrophy) and TYG agar plates with thymidine (for total viable cells) (Fig. 7a). Unexpectedly, some background growth appeared on these agar plates, but the colony morphology of *B. thetaiotaomicron* differed from that of the background. Therefore, we counted the colonies of the engineered strain by enumerating NanoLuc-positive colonies. Colonies escaping from thymidine auxotrophy were not detected in the fecal pellets (Fig. 7c). The average detection limits of the frequency of escape were  $5.2 \times 10^{-7}$  for *ΔthyA* (ER<sup>+</sup>/CD<sup>+</sup>) and  $1.1 \times 10^{-7}$  for *ΔthyA* (ER<sup>+</sup>/CD<sup>NT</sup>), indicating that the thymidine auxotrophy in the engineered *B. thetaiotaomicron* strains was highly robust.

However, the number of viable thymidine-auxotrophic strains had not decreased after 7 days of the storage of solid feces at ambient temperature in the aerobic condition (Supplementary Fig. 10). This result is partially concordant with observations of *thyA*-deficient *B. ovatus* in the stools of mice<sup>20</sup>, in which thymineless death was repressed in the presence of oxygen. Given that the CD accelerated the death rate in vitro, we expected that our containment system would cause the engineered *Bacteroides* strain to decline, even in feces in the presence of oxygen, more rapidly than noted by ref. 20. The genetically engineered thymidine-auxotrophic strains could have been almost completely dormant without any replication of DNA or transcription of RNA in solid feces. Another possibility is the availability of substances such as nutrients that protected the thymidine-auxotrophic strains of *Bacteroides*.

To evaluate the genetic stability of the ER and CD, we isolated four colonies of the *ΔthyA* (ER<sup>+</sup>/CD<sup>+</sup>) strain from the feces of mice at day 10 and compared the isolates before and after inoculation. There was no significant difference in the luminescence values among the isolates, including controls before inoculation (Supplementary Fig. 11a). When the isolates were conjugated with *E. coli* carrying plasmids having the *thyA* gene, no viable *Bacteroides* cells were detected among the isolates derived from the feces. Thus, the CD remarkably inhibited the HGT of the *thyA* gene (Supplementary Fig. 11b). These results reveal that the ER and CD were stable for at least 10 days in the mouse gastrointestinal tract.

Finally, we validated whether the CD functions to prevent the disruption of thymidine auxotrophy by HGT of the *thyA* gene in vivo. Germ-free (GF) mice were first colonized with *E. coli* containing a mobilizable plasmid bearing the *thyA* gene. Subsequently, thymidine-auxotrophic *B. thetaiotaomicron* strains either with the ER and CD [*ΔthyA* (ER<sup>+</sup>/CD<sup>+</sup>)] or with the ER and non-targeting CD [*ΔthyA* (ER<sup>+</sup>/CD<sup>NT</sup>)] were administered to assess whether there would be plasmid-borne transfer of the *thyA* gene from *E. coli* to *B. thetaiotaomicron* in the mouse gut (Fig. 7d). As expected, we observed colonies of the *ΔthyA* (ER<sup>+</sup>/CD<sup>NT</sup>) strain on TYG agar plates both with and without thymidine when the stools of mice collected on day 7 were plated (Fig. 7e). In contrast, no colony of the *ΔthyA* (ER<sup>+</sup>/CD<sup>+</sup>) strain was detected on the TYG agar plates without thymidine, showing that the thymidine auxotrophy persisted. This result revealed that the CD had

prevented disruption of thymidine auxotrophy in vivo. However, the preventive effect was lost by day 14 (11 days after administration of *Bacteroides* strains), as the thymidine-auxotrophic *ΔthyA* (ER<sup>+</sup>/CD<sup>+</sup>) strain was replaced by a *B. thetaiotaomicron* strain without thymidine auxotrophy. We isolated twelve colonies of the non-auxotrophic strain and performed whole genome sequencing to clarify the mechanism of escape of the thymidine auxotrophy. The sequencing analysis showed that all twelve isolates lacked transgenes and had the *thyA* gene (BioSample accession number: SAMN37130011-SAMN37130022), whereas the ancestral strain contained the transgenes but not the *thyA* gene (BioSample accession number: SAMN37130023). We assume that a double crossover had occurred between the regions flanking the transgenes on the genome and homologous regions flanking the *thyA* gene on the plasmid. Both of the genome and the plasmid have the homologous sequences at the upstream (5'-*thyA*) and downstream (*thyA*-3') of the *thyA* gene, as shown in Supplementary Fig. 2a and Fig. 4a. The transconjugant likely appeared via conjugation between *E. coli* and *B. thetaiotaomicron* after day 7 and the transconjugant with *thyA* grew and outcompeted the thymidine-auxotrophic *ΔthyA* (ER<sup>+</sup>/CD<sup>+</sup>) strain. It is noteworthy that the transgenes we introduced into the WT *B. thetaiotaomicron* were removed upon loss of the CD and acquisition of *thyA*. Therefore, the genetically engineered strain of *B. thetaiotaomicron* with the transgenes was still biocontained even after the CD failed to target the *thyA* gene, as the strain reverted back to the wild-type genotype in the *thyA* locus.

## Discussion

In this study, we describe an auxotrophic biological containment system that combines an ER and a CD. We demonstrate that this system is tunable for the expression of the GOI, that the *B. thetaiotaomicron* host strain does not acquire the *thyA* gene by HGT, and that the dissemination of transgenes is inhibited in this system.

It has been reported that an ER, which was established on plasmids in *E. coli*, could be applied to regulate levels of gene expression and create synthetic ribocomputing devices<sup>43,44</sup>. We expand this technology to a prevalent human commensal, *B. thetaiotaomicron*, by stably introducing the ER onto its genome. In *B. thetaiotaomicron*, the strength of gene expression is not largely affected by the level of RBS complementarity to the 16 S rRNA. This methodology could be applied to other human commensals that, like *Bacteroides*, rely on the strength of the S1 protein binding site for translation. The ER can also be applied as a safeguard when the taRNA and CR are designed for each individual GOI, by having the secondary structures of the crRNA include coding sequences. We created the taRNAs and CRs for the *nanoluc* gene as a model and showed the functionality of the tunable switch in *B. thetaiotaomicron*. The production of NanoLuc was switched on only when its gene was present in the cell with the CR and the corresponding taRNA. Moreover, by taking into consideration the MFE of the secondary structure when designing the taRNAs, an expression range of approximately 10-fold was obtained (Fig. 2f). A larger range of expression in the ON state and more strict suppression in the OFF state

would, however, be required for practical applications. The ER system is likely to be responsive to variations in temperature, considering the MFE. When the genetically engineered *B. thetaiotaomicron* comes out of the body, the production of the protein would not be expected to increase at ambient temperature, because there would be less translation of the GOI under the control of the ER. A deep-learning approach has already been developed to optimize the structure of ERs<sup>45</sup>. A more sophisticated regulatory system could be created by combining the use of the ER as a safeguard and improving the method.

It has been demonstrated that auxotrophy, created by eliminating essential genes, is effective for biological containment<sup>11,12</sup>. Nevertheless, it is possible that the containment system could be breached if the auxotroph were to acquire the essential genes via HGT from natural microorganisms<sup>20</sup>. We have constructed genetically modified thymidine-auxotrophic strains of *B. thetaiotaomicron* and quantified the loss of viable cells in the absence of thymidine. When the cells were cultured for up to 11 days, the CFU/mL of the P<sub>cepA</sub>-sgRNA1 strain decreased by 10<sup>7</sup>–10<sup>8</sup>. Given that the limit recommended by the National Institutes of Health<sup>12</sup> for the escape of engineered microbes or engineered DNA transmission is fewer than 1 cell per 10<sup>8</sup> cells, the performance of this auxotrophic system falls just slightly short of that criterion. In particular, the death rate observed in the aerobic state was slower than that observed in the anaerobic state. The auxotrophic strain survived in the feces of mice in the presence of oxygen. Therefore, for clinical application, our system would require optimization. Essential genes or chassis strains could be chosen so as to create more suitable engineered auxotrophic microorganisms for employing the ER and CD in an uncontrolled environment. Nevertheless, the CD with a promoter of appropriate strength could help to increase the bacterial death rate in the absence of thymidine whereas the engineered microorganisms would grow in the presence of thymidine.

The function of the CD was verified by assessing the decrease in conjugation efficiency through filter mating. It was demonstrated that the CD specifically recognizes a sequence identified by a sgRNA and makes a double-strand break in the sequence. This characteristic would allow us to precisely target an unfavorable gene that could potentially destroy a biological containment system. However, preventing HGT by the CD could potentially be a drawback, because theoretically, numerous variants of the target gene are present in the environment. A preferable candidate as a target sequence would be one chosen from well-conserved sequences in a gene, such as *thyA* and *dapA*, whose deficiency causes auxotrophy, leading to rapid death unless an essential nutrient is supplied<sup>46,47</sup>. Furthermore, introducing multiple sgRNAs would help to cover the wider range of potential variants<sup>34</sup>.

The CD suppressed the frequency of acquisition of the *thyA* gene and transgenes by 1 to 2 orders of magnitude, which seems to be adequate to prevent the acquisition of those genes, considering the conjugation efficiency of *B. ovatus* and the target frequency described above. The conjugation efficiency of this species for the *thyA* gene on its genome has been reported as 0.9–4.2 × 10<sup>-7</sup> while the recommended efficiency is below 1 × 10<sup>-8</sup><sup>20</sup>. If the CD is utilized for gene modification, its efficiency would satisfy this criterion. The CD could be optimized by adjusting the strength of the promoters, as shown in Fig. 5b, though the CD does affect cell growth, so there is a tradeoff. Potentially, the capability could also be improved by using Cas3 and the CasABCDE complex, which is analogous to Cas9 and enhances DNA degradation after the initial cleavage with 3'-to-5' helicase and ssDNA exonuclease activities<sup>34</sup>. Transgenes are unlikely to be expressed by species not closely related to *Bacteroides* because of the specific regulation of gene expression<sup>48–50</sup>, even if the transgenes have transferred. In choosing promoters, we considered both intraspecies HGT and HGT between closely related species. For the CD, we chose active promoters in interspecies populations of *Bacteroides*: P<sub>cepA</sub> or P<sub>cfxA</sub> for the transcription of sgRNA, and P<sub>cfxA</sub> for that of SpCas9<sup>49–52</sup>. If promoters that would function in a broader range of gut bacterial cells are needed

to effectively prevent the dissemination of transgenes, genomic analysis of gut microbiota could be used to obtain information on widely distributed promoters; in fact, some promoters have already been reported to function in a variety of gut bacterial cells<sup>53</sup>. Therefore, functionality can be flexibly tuned to adapt to particular conditions of use. When tested in the mouse gut, the CD prevented the disruption of thymidine auxotrophy for four days. However, the extended experiment showed that the CD was not flawless: there is a possibility that the CD would eventually fail to guard thymidine auxotrophy. The improvements described above would be useful for a more stringent auxotrophic system. In addition, the failure of the CD could be prevented by introducing transgenes at the other sites of *thyA* locus in a *thyA*-deficient strain. This genetic modification avoids loss of the CD which would be caused by genetic exchange between the transgenes and an exogenous *thyA* gene. Nevertheless, this containment system is highly reliable since the genetically engineered strains with both transgenes and *thyA* gene were not detected even when the thymidine auxotrophy did not persist.

The phenotypes introduced by genetic engineering via double crossover, such as the thymidine auxotrophy, the ER, and the CD, were highly robust in vivo as well as in vitro. Those phenotypes in the genetically engineered strains of *B. thetaiotaomicron* were maintained after 10 days of colonization in the mouse gut. The persistent gene expression of Nanoluc regulated by the ER was confirmed in vivo. The CD also efficiently prevented the disruption of thymidine auxotrophy in the gastrointestinal tract by inhibiting the HGT of *thyA* gene from the surrounding environment.

Enzymes synthesized by the human microbiota generate many kinds of metabolites that affect host physiology, such as short chain fatty acids and bile acids<sup>54,55</sup>. Genetically modified microorganisms bearing the related genes, as well as IL-10-producing *Lactococcus lactis* mentioned above, might offer a new therapeutic approach to treat diseases caused by an imbalance of these metabolites. For these organisms to be used effectively, for example, to avoid unintended protein overproduction by the bacteria or the spread of the introduced gene to other bacteria, the expression level of the desired proteins would have to be regulated. Therefore, the framework of this Cas9-assisted biological containment has great potential for the practical use of genetically modified microorganisms in biomedicine and other fields.

We have provided proof-of-concept that the Cas9-assisted thymidine auxotrophic biocontainment system is tunable and useful for simultaneously preventing the dissemination of transgenes and the escape from thymineless death, as well as controlling the level of gene expression. The ON/OFF ratio of the gene expression and the death rate in unfavorable conditions still require improvement to meet the standards for clinical application. One way to achieve a system that can be applied in clinical situations is to use other human commensal bacteria or essential genes, and to design the ER with deep learning. By addressing these challenges, future work can lead to the availability of useful genetically modified microorganisms.

## Methods

### Strains and culture conditions

*B. thetaiotaomicron* VPI-5482 (ATCC 29148) (GenBank: AE015928.1) was used for all evaluation of biological containment of genetically modified bacterial strains and genetic elements. This strain, which is sensitive to erythromycin but resistant to gentamicin, was grown in Brain Heart Infusion (Difco) media with 5 g/L yeast extract (Bacto) (BHI), Trypticase-Yeast extract-Glucose (TYG) media or Defined Minimal Media (DMM), supplemented with 1 mg/L resazurin sodium salt (Alfa Aesar), 10 mg/L hemin (Sigma), 0.5 g/L cysteine hydrochloride (Sigma), and 1 mg/L Vitamin K3 (Sigma)<sup>31</sup>. TYG media contained: 10 g/L trypticase (BBL), 5 g/L yeast extract (Bacto), 1 g/L Na<sub>2</sub>CO<sub>3</sub> (AMERSCO), 2 g/L glucose (Sigma), 80 mM potassium phosphate buffer (pH 7.3), 20 mg/L MgSO<sub>4</sub>·7H<sub>2</sub>O (Sigma), 400 mg/L NaHCO<sub>3</sub>

(Sigma), 80 mg/L NaCl (MACRON), 8 mg/L CaCl<sub>2</sub> (Sigma), and 0.4 mg/L FeSO<sub>4</sub> · 7H<sub>2</sub>O (Sigma). DMM media contained: 1 g/L NH<sub>4</sub>SO<sub>4</sub> (Sigma), 1 g/L Na<sub>2</sub>CO<sub>3</sub>, 80 mM potassium phosphate buffer (pH7.3), 900 mg/L NaCl (MACRON), 20 mg/L CaCl<sub>2</sub> (Sigma), 20 mg/L MgCl<sub>2</sub> · 6H<sub>2</sub>O (VWR), 10 mg/L MnCl<sub>2</sub> · 4H<sub>2</sub>O (Sigma), 10 mg/L CoCl<sub>2</sub> · 6H<sub>2</sub>O (Sigma), 5 µg/L Vitamin B12 (Sigma), 4 mg/L FeSO<sub>4</sub> · 7H<sub>2</sub>O (Sigma), and 5 g/L carbon source. Cultures were grown in an anaerobic chamber with an atmosphere of 2% H<sub>2</sub> balanced with N<sub>2</sub> and CO<sub>2</sub> at 37 °C. All media was pre-reduced overnight in an anaerobic atmosphere before inoculation. Antibiotics, thymidine, carbon sources, bile acids, 4-Chloro-DL-phenylalanine, and agarose were added as appropriate: erythromycin (Sigma, 25 µg/mL), gentamicin sulfate (TEKnova, 200 µg/mL), carbenicillin disodium salt (Gold Biotechnology, 5 µg/mL), tetracycline hydrochloride (IBI Scientific, 5 µg/mL), Polymyxin B sulfate (Merck Millipore, 10 µg/mL), thymidine (Alfa Aesar, 100 µg/mL), galactose (Sigma, 0.5% w/v), maltose (Sigma, 0.3% w/v), starch from potato (Sigma, 0.5% w/v), deoxycholic acid (Sigma, 62.5 µM), sodium cholate (Sigma, 0.5%) as a crude ox bile extract, 4-Chloro-DL-phenylalanine (Sigma, 2 mg/mL), and agarose (LabExpress, 15 g/L).

*E. coli* S17-1 λ pir<sup>56</sup> was utilized for the construction of plasmids and the transformation of *B. thetaiotaomicron* strains through conjugation. *E. coli* S17-1 λ pir was routinely grown in Luria Bertani (LB) broth (LabExpress) supplemented with 100 µg/mL carbenicillin (Gold Biotechnology) for plasmid selection.

### Plasmid construction

All plasmids were constructed with standard cloning procedures. Briefly, DNA fragments were generated by PCR or purchased from Integrated DNA Technologies, Inc. Fragments were assembled by Gibson Assembly Cloning Method into plasmids<sup>57</sup>. Plasmid sequences were confirmed by Sanger sequencing. *E. coli* S17-1 λ pir was transformed with the constructed plasmids by electroporation to propagate the plasmids and for transformation of *B. thetaiotaomicron* through conjugation. Genetic parts, plasmids, and primers used in this study are shown in Supplementary Tables 1–6. The representative plasmids are depicted in Supplementary Fig. 1: pNH9188 was used for the first recombination and pNH9196 was used for the second recombination to construct the genetically engineered strains of *B. thetaiotaomicron*, which were evaluated *in vitro* and *in vivo*.

### Strain construction

Thymidine-auxotrophic genetically modified strains were generated through recombination of plasmids bearing the ER and CD (Supplementary Fig. 1) into the genome of *B. thetaiotaomicron*. The wild-type *B. thetaiotaomicron* VPI-5482 strain was transformed with the plasmids to build genetically modified *thyA*-deficient strains through double crossovers upstream and downstream of the *thyA* locus (Supplementary Fig. 2)<sup>58</sup>. The taRNA sequences, SpCas9 gene, CR, and NanoLuc gene were integrated into the genome of the *B. thetaiotaomicron* by the first recombination, in which the *thyA* gene on the genome was exchanged with transgenes. sgRNA to target the *thyA* gene was integrated by the second recombination (Supplementary Fig. 2a). The plasmids bearing those sequences were used to transform WT *B. thetaiotaomicron* through conjugation with *E. coli* S17-1 λ pir.

Overnight cultures of the WT strain were added to *E. coli* S17-1 λ pir with the plasmids and the mating mixtures were pelleted, resuspended in BHIS medium, spotted onto BHIS agar plates and incubated upright at 37 °C aerobically. After overnight incubation, cells were collected by scraping, resuspended, and plated on BHIS agar plates with gentamicin and erythromycin as selective markers and incubated anaerobically for 2 days at 37 °C. Resultant colonies were re-isolated on the BHIS agar plates with gentamicin and erythromycin.

To obtain the *thyA*-locus-exchanged mutant, each colony was anaerobically cultured for 1 day in 3 mL of BHIS with thymidine without antibiotics. Bacterial cells were collected by centrifugation, washed,

and resuspended in PBS, and the appropriate dilution was plated onto DMM agar plates (0.5% glucose) supplemented with p-Chloro-phenylalanine. The colonies collected from the DMM plates after 3 days of cultivation at 37 °C were streaked on TYG and TYG with thymidine plates to check their sensitivity to thymidine. After a 48-h cultivation, colonies on TYG with thymidine plates were restreaked on TYG with thymidine and TYG with thymidine and erythromycin plates to check their sensitivity to erythromycin. Strains sensitive to thymidine and erythromycin were taken as *thyA*-locus-exchanged mutants.

The genetic exchange was confirmed by PCR using primers which bound to the regions flanking *thyA* or to the sgRNA region. DNAs were purified from both thymidine-auxotrophic and erythromycin-sensitive colonies using the DNeasy Blood & Tissue Kits (Qiagen). After the DNA concentration was determined, PCR was performed using primers encompassing the exchanged site (Supplementary Fig. 2a; Supplementary Table 6). The amplicon sizes were compared with those from the parental strain by agarose gel electrophoresis to confirm that the expected exchange occurred. PCR products of 2885 bp, including *thyA*, were detected in the WT strain. On the other hand, about 8000 bp of the amplicon that was bearing transgenes were detected after the double crossovers were introduced, indicating that the *thyA* gene on the chromosome had been exchanged through the recombination (Supplementary Fig. 2b).

The second recombination, i.e., that between plasmids bearing sgRNAs downstream of promoter P<sub>cepA</sub> or P<sub>cfxA</sub> and the chromosome of genetically modified *B. thetaiotaomicron*, was performed to integrate those sequences. The recombination process was the same as the first recombination except that thymidine was added to all of the media and agar plates. PCR was performed using primers encompassing the exchanged site and the purified DNAs as template. The presence of amplicons was checked by agarose gel electrophoresis to confirm that the expected integration had occurred. After the second recombination, the amplicons were detected at the length of about 6000 bp in strains bearing sgRNAs on the genome, whereas no amplicon was found in the parental strain (Supplementary Fig. 2c).

### Design of engineered riboregulator (ER)

The ERs composed of crRNA and taRNA were designed with RNAfold WebServer, which is used to predict nucleic acid structures<sup>38</sup>. Sequences of 99 nucleotides were simulated for crRNA; the 99-nucleotide sequences were composed of the RBS, CR, and part of the NanoLuc coding sequence, which contains the region considered to contribute to the mRNA folding effect on translation<sup>31,59</sup>. Likewise, structures of 99-nucleotide sequences of taRNA were predicted. The MFE of RNA folding was predicted for each variant using default parameters.

Structures of the crRNA/taRNA complex were also analyzed with NUPACK web server<sup>60</sup>, used to predict the MFE of mRNA structures including rpiL\* and part of the NanoLuc coding sequence in the previous study. NUPACK web server can be applied to multiple complex RNA structures whereas RNAfold WebServer analyzes a single RNA structure<sup>38</sup>. The MFE was calculated using default parameters other than concentrations of the crRNA and taRNA, and maximum complex size. We estimated the concentrations of the crRNA and taRNA to be 100 nM and the maximum complex size to be 2<sup>61</sup>.

### NanoLuc luciferase assay *in vitro*

Overnight cultures in TYG medium with thymidine were diluted 1:100 in fresh medium. The cultures were grown anaerobically at 37 °C to an optical density (OD<sub>600</sub>) of 0.4–0.8, and the cultures were harvested by centrifugation at 4500 × *g* for 10 min, washed once, and resuspended with phosphate buffered saline (PBS). NanoLuc Reaction Buffer (Promega) and cell suspensions were mixed at equal volumes. The luciferase activities were measured with an integration time of 1 s at a gain setting of 100 in BioTek Synergy HI Hybrid Reader 6 min after mixing. The luminescent values were normalized to OD<sub>600</sub> of 300 µL of cell

suspensions in PBS. Fold changes were calculated on the basis of the value with no CR (rpiL\*) and non-sense taRNA, or the value with crRCN and non-sense taRNA.

### Growth curve of genetically modified thymidine-auxotrophic *B. thetaiotaomicron*

Overnight cultures grown in TYG medium or DMM medium (0.5% glucose) with thymidine were harvested by centrifugation at  $3200 \times g$  for 10 min, washed with PBS twice, and resuspended in fresh TYG medium without thymidine or DMM medium without either thymidine or glucose. The cell suspensions were diluted 1:100 in the fresh TYG medium or DMM medium with or without thymidine. For cell growth in various stress conditions, fresh DMM medium with or without thymidine was used to dilute the cell suspensions; this medium included appropriate carbon sources or bile acids, as well as fresh TYG medium with or without thymidine. The cultures were grown anaerobically at 37 °C for 24 h. Samples (300  $\mu$ L) were withdrawn from the cultures for 10 h and at 24 h to monitor growth (OD<sub>600</sub>) in a BioTek Synergy HI Hybrid Reader.

### Viability test of genetically modified thymidine-auxotrophic *B. thetaiotaomicron* in vitro

Overnight cultures grown in TYG or DMM medium (0.5% glucose) with thymidine were centrifuged and washed twice in PBS, and cells were resuspended in fresh TYG without thymidine or in fresh DMM medium without either thymidine and glucose. Cell suspensions were diluted in TYG, TYG medium with bile acids, or DMM media with different carbon sources with or without thymidine, so that the optical density of the cell suspensions at 600 nm (OD<sub>600</sub>) was 0.2. For carbon source starvation, bacterial cells grown in DMM with 0.5% glucose were pelleted (OD<sub>600</sub> = 0.4), washed three times in PBS, resuspended in DMM without glucose in the presence or absence of thymidine, and incubated for up to 100 h. Cultures were incubated anaerobically at 37 °C or aerobically at ambient temperature. Inocula were withdrawn, and colony forming units (CFU) were measured by plating serial dilutions on BHIS with thymidine and gentamicin in an anaerobic chamber. After incubation at 37 °C for more than 4 days, colonies were counted manually and corrected for the dilution factor to obtain CFU/mL.

### Assay for inhibition of the Horizontal Gene Transfer (HGT) of *thyA* gene in vitro

Overnight cultures of genetically modified thymidine-auxotrophic *B. thetaiotaomicron* strains were diluted 1:100 in fresh TYG medium with thymidine, and *E. coli* S17-1  $\lambda$  pir strains bearing plasmids with the *thyA* gene and an erythromycin resistance gene were diluted 1:100 in fresh LB medium with carbenicillin. The cultures were grown anaerobically for *B. thetaiotaomicron* and aerobically for *E. coli* at 37 °C to an optical density of 0.4–0.9 at 600 nm, then harvested by centrifugation at  $3200 \times g$  for 10 min, washed in PBS twice, and resuspended in the fresh TYG medium with or without thymidine. The optical density of the cell suspension was adjusted to 16 with the medium, and *B. thetaiotaomicron* and *E. coli* were mixed at equal volumes. Then, 50 to 100  $\mu$ L of the mixtures were transferred onto a 0.45  $\mu$ m filter disc placed on TYG agar plates, with or without thymidine, and incubated anaerobically at 37 °C overnight. To evaluate conjugation efficiency, cells were washed off the filter and resuspended in BHIS medium with gentamicin and thymidine or in TYG medium with gentamicin. The cell suspension was plated on BHIS agar plates with gentamicin, thymidine, and erythromycin or on TYG agar plates with gentamicin. Total CFU/mL of viable cells after conjugation were determined by plating the cell suspension on BHIS agar plates containing gentamicin and thymidine or TYG agar plates containing gentamicin and thymidine. The colonies were counted manually 4 days after culturing.

DNA was purified from transconjugants grown on BHIS agar plates with gentamicin, thymidine, and erythromycin after conjugation, using the DNeasy Blood & Tissue Kits (Qiagen). PCR was performed with

primers *thyA*-F and *thyA*-R to detect the *thyA* gene. Primers named qPCR-EmR-F and qPCR-EmR-R were used for the EmR gene (Supplementary Table 6). Whole genome sequencing was also performed with the purified DNA to check for the integration of the plasmid bearing the intact *thyA* gene.

### Assay for prevention of the Horizontal Gene transfer (HGT) of transgenes

Overnight cultures of WT *B. thetaiotaomicron* strains were diluted 1:100 in fresh TYG medium with thymidine, and *E. coli* S17-1  $\lambda$  pir strains bearing plasmids with transgenes, such as taRNA, sgRNA, SpCas9, CR and NanoLuc gene, and an erythromycin resistance gene were diluted 1:100 in fresh LB medium with carbenicillin. The cultures were grown anaerobically for *B. thetaiotaomicron* and aerobically for *E. coli* at 37 °C to an optical density of 0.4–0.9 at 600 nm. Cultures were harvested by centrifugation at  $3200 \times g$  for 10 min, washed twice in PBS, and resuspended in fresh TYG medium without thymidine. The optical density of the cell suspension was adjusted to 16 with the medium, and *B. thetaiotaomicron* and *E. coli* were mixed at equal volumes. Then, 50  $\mu$ L of the mixtures were transferred onto a 0.45  $\mu$ m filter disc placed on TYG agar plates without thymidine. Plates were incubated anaerobically at 37 °C overnight. To evaluate the number of cells which had acquired plasmids with the transgenes and the erythromycin resistance gene, cells were washed off the filter and resuspended in the BHIS medium containing gentamicin, thymidine, and erythromycin. The cell suspension was plated on the BHIS agar plates containing the same components. The colonies were counted manually 4 days after culturing.

DNA was purified from transconjugants on BHIS agar plates with gentamicin, thymidine, and erythromycin after conjugation, using the DNeasy Blood & Tissue Kits (Qiagen). PCR was performed with primers Seq-*thyA*-F and mM663 to detect the region including the CD, and qPCR-EmR-F and qPCR-EmR-F for the erythromycin resistance gene (Supplementary Table 6). Whole genome sequencing was also performed with the purified DNA to check the integration of the plasmid bearing the transgenes and the erythromycin resistance gene.

### Whole genome sequencing and assembly

Total DNA of erythromycin-resistant colonies was purified using the DNeasy Blood & Tissue Kits (Qiagen). Genomic DNA library preparation and sequencing were performed at the SeqCenter (Pittsburgh, PA) by using Illumina NextSeq 2000 and NovaSeq 6000 technology. Sample libraries were prepared using the Illumina DNA Prep kit and IDT 10 bp unique dual indices (UDI). Demultiplexing, quality control and adapter trimming were performed with BCL Convert (Version 4.1.5). To analyze genetic alterations by HGT, paired-end sequencing reads (2  $\times$  151 bp) were assembled by SPAdes (Version 3.15.2)<sup>62</sup>. The sequencing reads (2  $\times$  151 bp) were also mapped to the plasmid bearing the *thyA* gene (shown in Fig. 4a) and the plasmid with transgenes (shown in Fig. 5a), using Bowtie 2 (Version 2.4.5)<sup>63</sup>. Raw sequencing data and assembled genomes have been deposited at DDBJ/ENA/GenBank under the accession number: PRJNA754595.

### Evolutionary stability test

The stability of the CD and ER during cell growth was verified by culturing the genetically modified thymidine-auxotrophic *B. thetaiotaomicron* strain with those functions for up to 21 days. The expression of the NanoLuc gene and the ability to block the HGT of the *thyA* gene were checked every 7 days. Overnight cultures in TYG medium with thymidine were diluted 1:100000 in fresh medium and grown anaerobically at 37 °C for 24 h, then grown continuously with dilution into fresh TYG media with thymidine to 10<sup>-5</sup> of optical density at 600 nm every 24  $\pm$  2.5 h<sup>34</sup>. The functioning of the CD and ER was evaluated on samples withdrawn from the passaged culture in accordance with the protocols of the NanoLuc luciferase assay and the assay for frequency

of HGT of the *thyA* gene mentioned above. Samples were grown in continuous culture and measured periodically, as described.

DNA was purified from transconjugants grown on BHIS agar plates with gentamicin, thymidine, and erythromycin after conjugation, using the DNeasy Blood & Tissue Kits (Qiagen). PCR was performed with primers Seq-*thyA*-F and mmd663 to amplify the region including the CD, and qPCR-EmR-F and qPCR-EmR-F to detect the erythromycin resistance gene (Supplementary Table 6). Sanger sequencing of the portion from the P<sub>cepA</sub>-sgRNA1 to the terminator downstream of SpCas9 was performed with the purified amplicons to check if there were any mutations in the CD region. Whole genome sequencing was also performed with the purified DNA to check the integration of the plasmid bearing the intact *thyA* gene.

## Animals

The protocols of SPF animal experiments were approved by the MIT Committee on Animal Care. Female BALB/cj mice (8 weeks) were purchased from The Jackson Laboratory, and four mice in each experimental group were housed together with a 12-h light/ dark cycle at standard room temperature of 20–22 °C at a humidity of 30–70%, and handled in non-sterile conditions. Mice were fed an irradiated mouse diet (ProLab IsoPro RMH3000, LabDiet) and non-sterile water before treatment with antibiotics. Prior to bacterial administration, mice were transferred to clean cages and then gavaged with metronidazole (100 mg/kg) in sterile water every day for 7 days. Over the course of the treatment, they were provided sterile filtered water containing ciprofloxacin hydrochloride (0.625 g/L). Animals were transferred to clean cages and provided fresh medicated water on the third or fourth day of the antibiotic regimen. Two days after the cessation of antibiotic treatment, they were transferred to clean cages and inoculated with *B. thetaiotaomicron* strains ( $-4 \times 10^8$  CFU/mouse) by oral administration. Fecal pellets were collected for the NanoLuc luciferase assay and the viability test of the bacterial strains in mice feces.

All gnotobiotic experiments were performed in accordance with the Guide for the Care and Use of Laboratory Animals (8th ed.) and were approved by the Institutional Animal Care and Use Committee of the University of Chicago. Female germ-free C57BL/6 mice were bred and housed in Trexler-style flexible film isolators (Class Biologically Clean) within polycarbonate mouse cages. Mice were weaned at 21 days of age and were fed an autoclaved, plant-based diet (JL Rat and Mouse/Auto 6 F 5K67, LabDiet). All mice were housed in cages with a 12-h light/ dark cycle at standard room temperature of 20–24 °C at a humidity of 30–70%. 8–12 week-old mice were inoculated with the *E. coli* S17-1  $\lambda$  pir strain bearing plasmids with the *thyA* gene ( $-10^8$  CFU/mouse) and transferred into bioexclusion cages (IsoCage P, Techniplast). After 3 days, genetically engineered *B. thetaiotaomicron* strains ( $-10^8$  CFU/mouse) were administered by oral gavage. Fecal pellets were collected to assay the inhibition of the HGT of the *thyA* gene in the mouse gut.

## NanoLuc luciferase assay in vivo

Fecal pellets were homogenized in PBS with an autoclaved single 5 mm stainless steel bead (Qiagen) using TissueLyser II (Qiagen) at 25 Hz for 2 mins. The concentration of feces was adjusted to less than 27 mg/mL. The fecal suspensions were centrifuged at  $500 \times g$  for 30 s to precipitate debris. The supernatants were diluted 10-fold in PBS. NanoLuc Reaction Buffer (Promega) and supernatants were mixed at equal volumes, and luciferase activities were measured with an integration time of 1 s at a gain setting of 100 in BioTek Synergy HI Hybrid Reader 6 mins after mixing. The luminescent values were normalized to the weight of the stools.

## Viability test of genetically modified thymidine-auxotrophic *B. thetaiotaomicron* in vivo

Fecal pellets were collected and stored in the presence of oxygen at ambient temperature prior to evaluation. Pellets were homogenized in

PBS with a single autoclaved 5 mm stainless steel bead (Qiagen) using TissueLyser II (Qiagen) at 25 Hz for 2 min. The fecal suspensions were centrifuged at  $500 \times g$  for 30 s to precipitate debris.

Supernatants were serially diluted under anaerobic conditions. The dilutions were plated on TYG agar plates either with erythromycin, gentamicin, carbenicillin, and polymyxin B, or with erythromycin, gentamicin, carbenicillin, polymyxin B, and thymidine. Luminescent colonies were counted manually when the NanoLuc Reaction Buffer (Promega) was spotted on the colonies, four days after culturing. The CFU values were normalized to the weight of the stools.

## Assay for inhibition of the Horizontal Gene Transfer (HGT) of *thyA* gene in vivo

Fecal pellets were homogenized in 1 mL of PBS with in a PowerLyzer 24 homogenizer (Qiagen) for 2 min at 2000 rpm, followed by a quick 30 s spin at 300 g to precipitate debris. Supernatants were serially diluted under anaerobic conditions. The dilutions were plated on LB agar plates to measure CFU/mg stool of the *E. coli* strain, on TYG agar plates with gentamicin for the *B. thetaiotaomicron* strains which overcame thymidine auxotrophy, or on TYG agar plates with gentamicin and thymidine for total *B. thetaiotaomicron* strains including thymidine-auxotrophic and non-auxotrophic strains. Colonies were counted manually one day after culturing aerobically on the LB agar plates or two days after culturing anaerobically on the TYG agar plates. The CFU values were normalized to the weight of the stools.

DNA was purified from colonies grown on TYG agar plates with gentamicin, using the DNeasy Blood & Tissue Kits (Qiagen). Whole genome sequencing was performed with the purified DNA to check for the integration of the plasmid bearing the *thyA* gene in the *E. coli* strain.

## Statistics

All data were analyzed using GraphPad Prism 9 (GraphPad Software).

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

Data supporting this study are presented in the main text and Supplementary information, and source data are provided with this paper. Raw sequencing data and assembled genomes of whole-genome sequencing of *Bacteroides thetaiotaomicron* strains have been deposited at DDBJ/ENA/GenBank under the accession number: [PRJNA754595](https://doi.org/10.1038/s41467-024-45893-w). Plasmids used in this study are available from the corresponding author upon request.

## References

- Lloyd-Price, J. et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* **569**, 655–662 (2019).
- Atarashi, K. et al. Ectopic colonization of oral bacteria in the intestine drives TH1 cell induction and inflammation. *Science* **358**, 359–365 (2017).
- Albhaisi, S. A. M., Bajaj, J. S. & Sanyal, A. J. Role of gut microbiota in liver disease. *Am. J. Physiol. Gastrointest. Liver Physiol.* **318**, G84–G98 (2020).
- Fiorucci, S. & Distrutti, E. Bile acid-activated receptors, intestinal microbiota, and the treatment of metabolic disorders. *Trends Mol. Med.* **21**, 702–714 (2015).
- Meng, C., Bai, C., Brown, T. D., Hood, L. E. & Tian, Q. Human gut microbiota and gastrointestinal cancer. *Genom. Proteom. Bioinform.* **16**, 33–49 (2018).
- Zuo, T. et al. Alterations in gut microbiota of patients with COVID-19 during time of hospitalization. *Gastroenterology* **159**, 944–955 (2020).

7. Lai, Y. et al. Illuminating host-mycobacterial interactions with genome-wide CRISPR knockout and CRISPRi screens. *Cell Syst.* **11**, 239–251 (2020).
8. Braat, H. et al. A phase I trial with transgenic bacteria expressing interleukin-10 in Crohn's disease. *Clin. Gastroenterol. Hepatol.* **4**, 754–759 (2006).
9. Chowdhury, S. et al. Programmable bacteria induce durable tumor regression and systemic antitumor immunity. *Nat. Med.* **25**, 1057–1063 (2019).
10. Leonard, S. P. et al. Engineered symbionts activate honey bee immunity and limit pathogens. *Science* **367**, 573–576 (2020).
11. Stirling, F. & Silver, P. A. Controlling the implementation of transgenic microbes: are we ready for what synthetic biology has to offer? *Mol. Cell* **78**, 614–623 (2020).
12. Moe-Behrens, G. H., Davis, R. & Haynes, K. A. Preparing synthetic biology for the world. *Front. Microbiol.* **4**, 5 (2013).
13. Gallagher, R. R., Patel, J. R., Interiano, A. L., Rovner, A. J. & Isaacs, F. J. Multilayered genetic safeguards limit growth of microorganisms to defined environments. *Nucleic Acids Res.* **43**, 1945–1954 (2015).
14. Stirling, F. et al. Rational design of evolutionarily stable microbial kill switches. *Mol. Cell* **68**, 686–697 (2017).
15. Rovner, A. J. et al. Recoded organisms engineered to depend on synthetic amino acids. *Nature* **518**, 89–93 (2015).
16. Mandell, D. J. et al. Biocontainment of genetically modified organisms by synthetic protein design. *Nature* **518**, 55–60 (2015).
17. Robertson, W. E. et al. Sense codon reassignment enables viral resistance and encoded polymer synthesis. *Science* **372**, 1057–1062 (2021).
18. Tang, T. C. et al. Hydrogel-based biocontainment of bacteria for continuous sensing and computation. *Nat. Chem. Biol.* **17**, 724–731 (2021).
19. Steidler, L. et al. Biological containment of genetically modified *Lactococcus lactis* for intestinal delivery of human interleukin 10. *Nat. Biotechnol.* **21**, 785–789 (2003).
20. Wegmann, U., Carvalho, A. L., Stocks, M. & Carding, S. R. Use of genetically modified bacteria for drug delivery in humans: revisiting the safety aspect. *Sci. Rep.* **7**, 2294 (2017).
21. Torres, B., Jaenecke, S., Timmis, K. N., Garcia, J. L. & Diaz, E. A dual lethal system to enhance containment of recombinant microorganisms. *Microbiology (Reading)* **149**, 3595–3601 (2003).
22. Alekshun, M. N. & Levy, S. B. Molecular mechanisms of antibacterial multidrug resistance. *Cell* **128**, 1037–1050 (2007).
23. Claesen, J. & Fischbach, M. A. Synthetic microbes as drug delivery systems. *ACS Synth. Biol.* **4**, 358–364 (2015).
24. Human Microbiome Project, C. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
25. Salyers, A. A. Bacteroides of the human lower intestinal tract. *Annu. Rev. Microbiol.* **38**, 293–313 (1984).
26. Porter, N. T., Luis, A. S. & Martens, E. C. Bacteroides thetaiotaomicron. *Trends Microbiol.* **26**, 966–967 (2018).
27. Yao, L. et al. A selective gut bacterial bile salt hydrolase alters host metabolism. *Elife* **7**, e37182 (2018).
28. Hamady, Z. Z. et al. Xylan-regulated delivery of human keratinocyte growth factor-2 to the inflamed colon by the human anaerobic commensal bacterium *Bacteroides ovatus*. *Gut* **59**, 461–469 (2010).
29. Round, J. L. & Mazmanian, S. K. Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. *Proc. Natl Acad. Sci. USA* **107**, 12204–12209 (2010).
30. Taketani, M. et al. Genetic circuit design automation for the gut resident species *Bacteroides thetaiotaomicron*. *Nat. Biotechnol.* **38**, 962–969 (2020).
31. Mimeo, M., Tucker, A. C., Voigt, C. A. & Lu, T. K. Programming a human commensal bacterium, *bacteroides thetaiotaomicron*, to sense and respond to stimuli in the murine gut microbiota. *Cell Syst.* **1**, 62–71 (2015).
32. Farrar, M. D. et al. Engineering of the gut commensal bacterium *Bacteroides ovatus* to produce and secrete biologically active murine interleukin-2 in response to xylan. *J. Appl. Microbiol.* **98**, 1191–1197 (2005).
33. Isaacs, F. J. et al. Engineered riboregulators enable post-transcriptional control of gene expression. *Nat. Biotechnol.* **22**, 841–847 (2004).
34. Caliendo, B. J. & Voigt, C. A. Targeted DNA degradation using a CRISPR device stably carried in the host genome. *Nat. Commun.* **6**, 6989 (2015).
35. Accetto, T. & Avgustin, G. Inability of *Prevotella bryantii* to form a functional Shine-Dalgarno interaction reflects unique evolution of ribosome binding sites in Bacteroidetes. *PLoS One* **6**, e22914 (2011).
36. Wegmann, U., Horn, N. & Carding, S. R. Defining the bacteroides ribosomal binding site. *Appl. Environ. Microbiol.* **79**, 1980–1989 (2013).
37. Boni, I. V., Isaeva, D. M., Musychenko, M. L. & Tzareva, N. V. Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. *Nucleic Acids Res.* **19**, 155–162 (1991).
38. Gruber, A. R., Lorenz, R., Bernhart, S. H., Neubock, R. & Hofacker, I. L. The Vienna RNA websuite. *Nucleic Acids Res.* **36**, W70–W74 (2008).
39. Kawano, M. et al. Artificial small RNA-mediated growth inhibition in *Escherichia coli* and *Salmonella enterica* serovar Typhimurium. *Biochem. Biophys. Res. Commun.* **521**, 577–583 (2020).
40. Noro, E. et al. Systematic characterization of artificial small RNA-mediated inhibition of *Escherichia coli* growth. *RNA Biol.* **14**, 206–218 (2017).
41. Guzman, E. C. & Martin, C. M. Thymineless death, at the origin. *Front. Microbiol.* **6**, 499 (2015).
42. Vogel-Scheel, J., Alpert, C., Engst, W., Loh, G. & Blaut, M. Requirement of purine and pyrimidine synthesis for colonization of the mouse intestine by *Escherichia coli*. *Appl. Environ. Microbiol.* **76**, 5181–5187 (2010).
43. Callura, J. M., Dwyer, D. J., Isaacs, F. J., Cantor, C. R. & Collins, J. J. Tracking, tuning, and terminating microbial physiology using synthetic riboregulators. *Proc. Natl Acad. Sci. USA* **107**, 15898–15903 (2010).
44. Green, A. A. et al. Complex cellular logic computation using ribocomputing devices. *Nature* **548**, 117–121 (2017).
45. Angenent-Mari, N. M., Garruss, A. S., Soenksen, L. R., Church, G. & Collins, J. J. A deep learning approach to programmable RNA switches. *Nat. Commun.* **11**, 5057 (2020).
46. Isabella, V. M. et al. Development of a synthetic live bacterial therapeutic for the human metabolic disease phenylketonuria. *Nat. Biotechnol.* **36**, 857–864 (2018).
47. Wright, O., Delmans, M., Stan, G. B. & Ellis, T. GeneGuard: a modular plasmid system designed for biosafety. *ACS Synth. Biol.* **4**, 307–316 (2015).
48. Smith, C. J., Rogers, M. B. & McKee, M. L. Heterologous gene expression in *Bacteroides fragilis*. *Plasmid* **27**, 141–154 (1992).
49. Bayley, D. P., Rocha, E. R. & Smith, C. J. Analysis of *cepA* and other *Bacteroides fragilis* genes reveals a unique promoter structure. *FEMS Microbiol. Lett.* **193**, 149–154 (2000).
50. Parker, A. C. & Smith, J. C. Development of an IPTG inducible expression vector adapted for *Bacteroides fragilis*. *Plasmid* **68**, 86–92 (2012).
51. Rogers, M. B., Bennett, T. K., Payne, C. M. & Smith, C. J. Insertional activation of *cepA* leads to high-level beta-lactamase expression in *Bacteroides fragilis* clinical isolates. *J. Bacteriol.* **176**, 4376–4384 (1994).
52. Goto, T., Tanaka, K., Tran, C. M. & Watanabe, K. Complete sequence of pBFUK1, a carbapenemase-harboring mobilizable plasmid from *Bacteroides fragilis*, and distribution of pBFUK1-like plasmids



- among carbapenem-resistant *B. fragilis* clinical isolates. *J. Antibiot.* **66**, 239–242 (2013).
53. 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).
54. Smith, P. M. et al. The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science* **341**, 569–573 (2013).
55. Jia, W., Xie, G. & Jia, W. Bile acid-microbiota crosstalk in gastrointestinal inflammation and carcinogenesis. *Nat. Rev. Gastroenterol. Hepatol.* **15**, 111–128 (2018).
56. Simon, R., Priefer, U. & Pühler, A. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. *Bio/Technol.* **1**, 784–791 (1983).
57. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
58. Kino, Y. et al. Counterselection employing mutated pheS for markerless genetic deletion in *Bacteroides* species. *Anaerobe* **42**, 81–88 (2016).
59. Kudla, G., Murray, A. W., Tollervey, D. & Plotkin, J. B. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324**, 255–258 (2009).
60. Zadeh, J. N. et al. NUPACK: analysis and design of nucleic acid systems. *J. Comput Chem.* **32**, 170–173 (2011).
61. Guet, C. C. et al. Minimally invasive determination of mRNA concentration in single living bacteria. *Nucleic Acids Res.* **36**, e73 (2008).
62. Bankevich, A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput Biol.* **19**, 455–477 (2012).
63. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

## Acknowledgements

We would like to thank all lab members in the Synthetic Biology Group in MIT's Synthetic Biology Center for their great help, especially, Dr. Ky Lowenhaupt, Dr. Isaak E. Mueller, Dr. Giyoung Jung, Dr. Tzu-Chieh Tang, Dr. Kevin M. Yehl and Dr. Maria Eugenia Inda. We thank Karen Pepper for editing the manuscript. We thank Dr. Xiaoqiong Gu for bioinformatics analysis. We also thank the staff at the Koch Institute Swanson Biotechnology Center and the Division of Comparative Medicine at MIT for their assistance for animal experiments. We thank the Duchoissois Family Institute for providing germ-free mice. This study was supported by JSR Corporation in Japan. This work was supported by the National Science Foundation (NSF-CCF-1521925 to T.K.L.), the National Institutes of Health (NIH-5-U01-CA2550554-02 and NIH-50000655-5500001351 to T.K.L., NIGMS R35GM147478 to M.M., and IMSD Program 5R25GM109439-07 and Molecular and Cellular Biology training program T32 GM007183 to J.F.-S.), and the Arnold and Mabel Beckman Foundation through the Beckman Young Investigator Program (M.M.).

## Author contributions

N.H., Y.L., M.M. and T.K.L. conceived and designed this study. N.H., Y.L. and J.F.-S. performed the experiments. N.H., Y.L., J.F.-S., M.M. and T.K.L. analyzed data, discussed the results, and wrote the manuscript.

## Competing interests

T.K.L. is a co-founder of Senti Biosciences, Synlogic, Engine Biosciences, Tango Therapeutics, Corvium, BiomX, Eligo Biosciences, Bota.Bio, and Avendesora. T.K.L. also holds financial interests in nest.bio, Ampliphi, IndieBio, MedicusTek, Quark Biosciences, Personal Genomics, Thyrye, Lextent Bio, MitoLab, Vulcan, Serotiny, and Avendesora. N.H. is an employee of JSR Corporation. Other authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-45893-w>.

**Correspondence** and requests for materials should be addressed to Mark Mimeo or Timothy K. Lu.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

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

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

© The Author(s) 2024