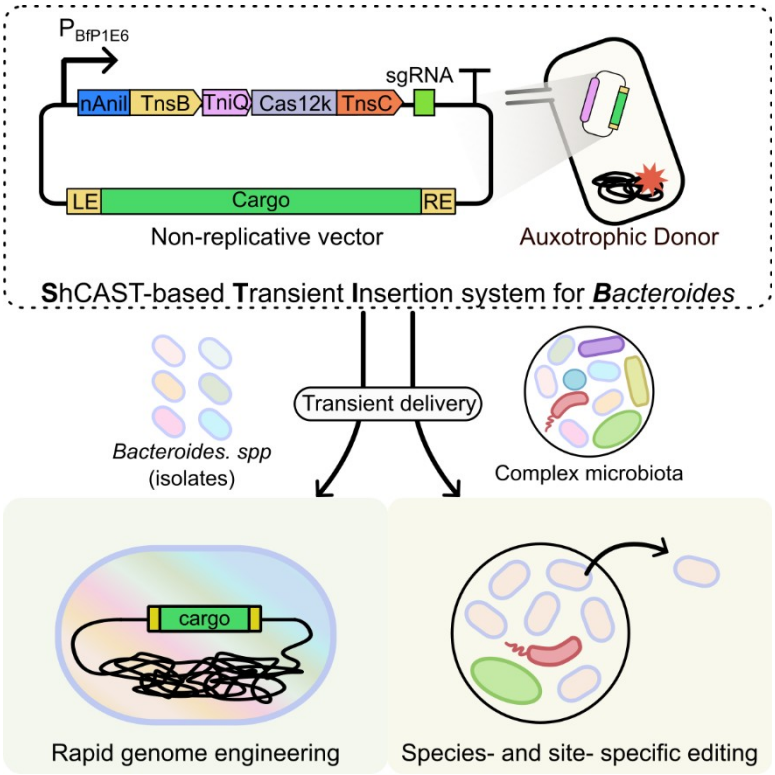
Targeted genomic editing of human gut *Bacteroides* species based on CRISPR-associated transposases

Yucan Hu[1,2,†], Qingmei Li[1,†], Yingying Li[1], Yu Zeng[1], Linggang Zheng[1], Juntao Shen[1], Xiang Gao[3], Guo-Ping Zhao[1,4,5], Wei Zhao[1,2,*], Lei Dai[1,2,6,*]

[1] State Key Laboratory of Quantitative Synthetic Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
[2] University of Chinese Academy of Sciences, Beijing 100049, China.
[3] State Key Laboratory of Microbial Technology, Shandong University
[4] CAS Key Laboratory of Synthetic Biology, CAS Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai, China
[5] State Key Lab of Genetic Engineering and Institutes of Biomedical Sciences, Department of Microbiology and Microbial Engineering, School of Life Sciences, Fudan University, Shanghai, China
[6] Lead contact
† : These authors contributed equally to this work.
* To whom correspondence should be addressed: Email: lei.dai@siat.ac.cn (L.D.), wei.zhao1@siat.ac.cn (W.Z.)

Graphical abstract



Summary

Bacteroides are a group of highly abundant bacteria in the gut and play critical roles in human health and diseases, while most of them are non-model microbes and genetically cumbersome. The development of a widely applicable editing tool for *Bacteroides* is much needed for the manipulation of human gut microbiome. In this study, we develop STIB (ShCAST-based Transient Insertion system for *Bacteroides*), a novel genome editing tool derived from CRISPR-associated transposases that enable rapid and site-specific insertions independent of homologous recombination. By fusing a nicking homing endonuclease to the transposase and an ATPase to Cas12k, we systematically optimize STIB to minimize plasmid cointegration and to achieve >97% on-target insertion. We demonstrate that STIB exhibits broad applicability across different genomic loci in distinct *Bacteroides*, including some non-model species. Finally, we apply STIB to achieve species- and site-specific editing of multiple *Bacteroides* species in a complex synthetic gut microbiota. Overall, our study provides a novel genome editing tool for *Bacteroides* species and expands the toolbox for functional investigation and engineering of the human microbiome.

Introduction

Bacteroides are dominant members of the human gut microbiome and known for their critical roles in modulation of host immunity and metabolism, digestion of polysaccharides, and microbial interactions [1–5]. Previous studies have highlighted their extensive genomic plasticity [6] and stable colonization [7], making them attractive models for studying microbiota-host interactions and for engineering living therapeutics [8–10]. Recent investigations have elucidated molecular mechanisms underlying *Bacteroides*-mediated microbiota interactions, often employing the re-introduction of genetically modified isolates into the microbiome [3,11,12].

Since the pioneering works on the synthetic biology of *Bacteroides* [13,14], a growing number of genetic tools have been developed for *Bacteroides* species, including transposases[15–17], integrases [14,18–20], counterselection-based allelic exchange [21–23], CRISPR-Cas system[24,25], etc. Transposases offer high efficiency but lack specificity. Integrases can accommodate large cargo loads, but their insertions are limited to specific sites and lack programmability. Counterselection-based

allelic exchange requires specific mutant genome backgrounds or multiple steps to induce suicide gene expression. Our previous study has developed CRISPR-Cas systems for programmable editing in diverse *Bacteroides* species[24]. Nevertheless, the CRISPR-Cas system relies on DNA double-strand breaks (DSBs)[26] and the subsequent host homologous recombination (HR) mechanism, which application is limited by the low efficiency of HR in many *Bacteroides* strains[27]. In comparison, base editors of *Bacteroides* species[28,29] avoid DSBs but are limited to single-nucleotide edits.

CRISPR-associated transposases (CASTs) rely on transposases and are coupled with CRISPR-Cas effectors to direct RNA-guided transposition[30]. CAST systems enable efficient and HR-independent insertion of up to ten kilobases (kb) of exogenous cargo DNA into target loci[30–39]. The two major well-characterized CAST types, I-F and V-K, exhibit distinct mechanisms and outcomes[40]. Type I-F CAST systems (e.g., INTEGRATE)[33] demonstrate excellent specificity and efficiency, albeit with minor bidirectional insertion byproducts. An all-in-one system based on INTEGRATE, termed “DART”, has been used to target *Escherichia coli* in complex communities[39]. However, the large number of essential genes (3 Cas effectors, 4 transposase-associated proteins) in type I-F CASTs, coupled with their low-temperature-dependent efficiency (optimal at 25°C or 30°C)[33,41], presents challenges for application in evolutionarily distant gut bacteria like *Bacteroides*, which thrive at 37°C. In contrast, type V-K CAST systems (e.g., ShCAST) feature a more compact gene architecture and yield predominantly unidirectional insertions with high efficiency at 37°C[31,41]. Although the original ShCAST system exhibited approximately 50% specificity and ~20% undesirable cointegration[31,42], recent studies have mitigated these issues through engineering in *E. coli* [41]. The advances of CASTs motivated us to develop ShCAST into a versatile, efficient, and specific genome editing tool for inserting kilobase-scale DNA into *Bacteroides* genomes.

Here, we present **S**hCAST-based **T**ransient **I**nsertion system for ***B****acteroides* (STIB), enabling species- and site-specific editing within pure cultures and complex microbial communities. We constructed a ShCAST-based system using a non-replicative vector for transient expression, streamlining the workflow by eliminating plasmid curing while enhancing biosafety in community. Despite initially encountering significant off-target effects, we successfully optimized the ShCAST system to achieve high on-target efficiency (>97%) by fusing a nicking

homing endonuclease (nAnil) to the transposase (TnsB) and an ATPase (TnsC) to Cas12k. Subsequently, utilizing the optimized system (STIB), we achieved efficient and specific genome insertion with large cargo capacity in diverse *Bacteroides* species. We further demonstrated the utility of STIB by inserting a large exogenous inulin utilization gene cluster to expand the metabolic capacity of *B. thetaiotaomicron.* Finally, we accomplished site- and species- specific editing of *Bacteroides* in a complex microbiota of 40 human gut bacterial species, which allowed the enrichment and isolation of the target species.

Results

Construction and Optimization of STIB system

Previous studies in *Proteobacteria* established that ShCAST activity is dependent on high levels of effector expression[34,36]. Inspired by successful adaptations of Type I-F CAST systems through the tuning of regulatory elements[43,44], we sought to maximize ShCAST expression in *B. thetaiotaomicron*. We first benchmarked seven promoter–ribosomal binding site (RBS) combinations using a luciferase reporter system (**Table S1**)[14,45–47]. Consistent with previous reports[13], the P_{BfP1E6}–*RBS8* configuration yielded the highest luminescence (**Figure S1A**). To further refine translation, we tested a series of RBS variants[13,14] under the control of the P_{BfP1E6} promoter, identifying *RBSB6*, *RBSB5*, and *RBSB2* and *RBS8* as the most potent candidates (**Figure S1B**). Inspired by the compact polycistronic operons common in prokaryotes, we codon-optimized the coding sequences (CDS) for the four ShCAST effector proteins, and designed a ploycristronic construct. In this design, P_{BfP1E6} drives the simultaneous expression of all effectors, with the prioritized RBSs arranged in increasing order of strength across the four genes to ensure balanced and robust protein production (**Figure 1A**). Meanwhile, we constructed a replicating *Bacteroides* vector for maintaining expression of the redesigned ShCAST system with a sgRNA targeting the thymidine kinase gene (*tdk*) (**Table S2**) in *Bacteroides* cells. This vector was introduced into *B. thetaiotaomicron* via conjugation (**Figures S1C and S1D**) to disrupt *tdk* through RNA-guided target insertion, and the mutants would grew on 5-fluoro-2'-deoxyuridine (FUdR)-containing plates [21]. We observed transconjugants with successful editing (**Figures S1E and S1F**), but only one of ninety-three colonies in FUdR-plates had tdk-genotypes, suggesting an extremely low editing efficiency

(**Figure S1G**). Nevertheless, these initial results indicated that the ShCAST system could achieve targeted editing in *Bacteroides*.

Subsequently, we engineered a transient ShCAST system by removing the *Bacteroides* replicon and introducing a chloramphenicol resistance gene (*cmr*) as cargo, enabling a streamlined, one-step genome editing within 4 days (**Figure 1B**). Nevertheless, we detected a low efficiency (~0%, total 20 clones) (**Figures S2A and S2B**), high cointegration rate (100%) (**Figures S2A and S2C**) and poor specificity (1.46%) (**Figure S2D**) in this transient ShCAST system. These findings indicated that the ShCAST system required further optimization to reduce cointegration and enhance both efficiency and specificity.

Previous research in *E. coli* reported that fusing a 5' nicking homing endonuclease (nAnil) to transposase TnsB, termed HELIX, can reduce cointegration and improve specificity[41]. Meanwhile, decreasing AAA^+ ATPase (TnsC) expression or fusing CRISPR effector (Cas12k) with TnsC or target-site-associated factor (TniQ) were shown to enhance specificity [41, 48]. We therefore initially incorporated nAnil and its recognition sequence into our transient vector, creating the ShHELIX construct (**Figure 1C**). Subsequent evaluation revealed a significant reduction in cointegration frequency (Tn-seq: from 99.5% to 8.1%; PCR: from 95.8% to 7.8%), a substantial increase in specificity (from 1.3% to 22%), and a modest improvement in efficiency (from 0% to 7.2%) (**Figure 1D**). Besides, we observed an A/T-rich motif at off-target sites (**Figure S3A**), as previous reported [48]. Inspired by these findings, we designed two approaches to further enhance efficiency and specificity based on the ShHELIX construct. First, we calculated the relative expression strength (RBSB2 used as the base line) according to previous reports[14] and predictions from the RBS Calculator[49]. And these RBSs may reduce the expression range of TnsC by ~100- to ~2800-fold by substituting RBS with different strengths, generating psh63, C53 and C56 (**Figures 1C and S3B**). Second, we fused TnsC or TniQ to Cas12k to generate constructs psh43 and psh53, respectively (**Figure 1C**). All described ShHELIX-based systems were evaluated in parallel transposition experiments. Subsequent colony PCR and Tn-seq analyses revealed that all constructs maintained low cointegration rates (**Figures 1D and S3C**; ShHELIX vs ShCAST: $p<0.01$, psh43 vs ShHELIX: no significant difference.). Moreover, the editing efficiencies of both psh43 and psh63 systems were above 80% and did not differ significantly (**Figures 1D and S3D**; $p=0.373$). Meanwhile, the specificity of psh43 (97.1%)

surpassed that of psh63 (86.3%) (**Figures 1D and S4A**). To mitigate potential site bias, we compared the efficiencies of psh43, psh53, and psh63 at two different target sites (*Bt02* and *Bt03*, **Table S2**). The results demonstrated that psh43 exhibited the significant higher efficiency across the new tested sites (*Bt02* and *Bt03*) (**Figure S4B**, two-way ANOVA: p=0.038).

Collectively, our observations suggest that the psh43 system achieves a rapid, specific and efficient genomic insertion in *Bacteroides*. This optimized transient system, ShCAST-based Transient Insertion system for *Bacteroides,* was designated as STIB.

Systematic characterization of STIB in *B. thetaiotaomicron*

To comprehensively characterize STIB performance across the genome, we next evaluated STIB at ten distinct genomic target sites in *B. thetaiotaomicron*, including the three previously described sites, spanning both ORFs and intergenic regions (**Table S2**). Following transposition experiments, we observed editing efficiencies ranging from 29% to 91% across these sites (**Figure 2A**). We also assessed the impact of cargo size on editing efficiency to determine whether STIB could facilitate large fragment insertions, as reported in *E. coli* [31,41]. STIB vectors carrying cargos of varying sizes (2.5, 5.5, and 8.4 kb) were introduced into *B. thetaiotaomicron* at the *Bt01* locus. Efficiency was evaluated by droplet digital PCR on all transconjugants. Consistent with expectations, high editing efficiencies (>85%) were observed irrespective of cargo size (**Figure 2B**). To further investigate insertion orientation, we comparatively profiled the insertion patterns at the *Bt01* site created by STIB, ShCAST, and ShHELIX. Tn-seq analysis revealed that STIB, like ShCAST, maintained a strong unidirectional insertion preference, with the T-LR (transposon left-end proximity to target site) orientation exceeding 99% (**Figure 2C**). Sanger sequencing of a *Bt01*-targeted colony revealed the *cmr* cargo insertion 61 bp downstream of the PAM, accompanied by the expected 5 bp target site duplications [31] (**Figure 2D**). To extensively examine the editing window, we surveyed insertion positions within the anticipated editing window (100 bp downstream of the PAM site) in all *Bt01* transconjugants. 98.2% of targeted insertions were located between 60 and 66 bp downstream of the PAM site (**Figure 2E**). A similar insertion distribution pattern was observed at the *Bt02* target site, where 77.8% of targeted insertions also localized between 60 and 66 bp downstream of the PAM (**Figure 2F**).

Collectively, these data demonstrate that STIB enables precise and efficient genomic editing and maintains a unidirectional insertion preference.

Leveraging the expedited site-specific editing workflow of STIB, which surpasses other *Bacteroides* editing methods in speed, we designed a proof-of-concept experiment for sequential gene editing (**Figure 2G**). Initially, a STIB vector targeting *Bt01* and carrying the *cmr* was constructed and introduced into *B. thetaiotaomicron* for round 1 editing (~3 days). Concurrently, a second STIB vector targeting *Bt02* with an erythromycin resistance gene (*ermG*) was prepared. Following round 1, a successfully edited colony was used as the chassis for round 2 editing at *Bt02* using the second vector (~4 days). Colony PCR confirmed high editing efficiencies, with 90% of clones edited in round 1 and 71% in round 2 (**Figure 2H**). Successful sequential editing was further validated by growth assays: while wild-type *B. thetaiotaomicron* was inhibited by chloramphenicol, edited strains grew unaffected. Ultimately, only clones subjected to both editing rounds proliferated in BHIS medium containing chloramphenicol and erythromycin (**Figure S5**). This entire sequential two-gene editing procedure was accomplished within one week, a 2- to 3-fold reduction in timeframe compared to traditional inducible promoter-based methods [22,24]. These results demonstrate that STIB facilitates rapid and sequential editing with high efficiency in *Bacteroides*.

Genomic editing of different *Bacteroides* species by STIB

We next assessed the activity of STIB in other gut *Bacteroides* species. Five representative species were chosen, comprising four type strains (*B. ovatus* ATCC8483, *B. vulgatus* ATCC8482, *B. fragilis* NCTC9343, *B. uniformis* ATCC8492) and two non-model strains (*B. salyersiae* DA1247 and *B. fragilis* BSC606), the latter of which were isolated from different healthy people stools. In *B. ovatus*, targeting four intergenic sites (*Bo01–Bo04*) and one ORF (*Bo05*) yielded editing efficiencies exceeding 85% expect at *Bo02* (30%). In non-model *B. salyersiae*, efficiencies ranged from 6% to 100% across five intergenic sites. In *B. vulgatus*, efficiency varied depending on the target site, with 0% at *Bv02* and *Bv03* but exceeding 93% at other sites (**Figure 2I**). We then hypothesized that a conserved 16S rDNA region could function as a universal target across different taxa. Targeting conserved 16S rDNA sequences in *B. fragilis* and *B. uniformis* (**Figure S6A**) yielded editing efficiencies exceeding 90% across all three tested strains, as confirmed by PCR (**Figures S6B and S6C**). These findings underscore

STIB's capacity for robust genome editing across diverse *Bacteroides* species, encompassing non-model strains. While highly efficient, the editing efficacy can be influenced by target site selection.

Targeted insertion of gene-of-interest in *Bacteroides* species

To demonstrate STIB's utility for genome engineering, we introduced genes-of-interest at designated target sites in *B. thetaiotaomicron* and *B. ovatus*. Gain-of-function was validated by inserting a green fluorescent protein (GFP) gene into the *Bt01* locus of *B. thetaiotaomicron* and the *Bo04* locus of *B. ovatus*. Robust GFP expression was detected in the transconjugants (**Figures 3A and 3B**). Subsequently, insertion of the *cmr* gene at the *Bt02* site in *B. thetaiotaomicron* and *Bo04* in *B. ovatus*. We observed significantly enhanced growth of these mutants compared to the wild type in BHIS medium supplemented with 25 μg/mL chloramphenicol (**Figures 3C and 3D**). These results suggest that STIB can effectively insert genes-of-interest to specific genomic sites in *Bacteroides* species.

We further explored STIB's potential for metabolic engineering by constructing an 8.4 kb cargo. This cargo contained an inulin utilization module and a *cmr* cassette. The inulin utilization module[23] comprises three core genes from the *B. ovatus* encoding an outer membrane SusC-like transporter (*BACOVA_04505*), an inulin-binding surface SusD-like lipoprotein (*BACOVA_04504*), and a periplasmic glycoside hydrolase (*BACOVA_04501*), driven by the constitutive promoter P_{BT1311} (**Figure 3E**). Following transposition and chloramphenicol selection, colony PCR confirmed 75% editing efficiency, using two specific primer pairs flanking the cargo insertion gene and the adjacent genomic regions (**Figure 3F**). Three isolated edits harboring the inulin utilization module (designated Bt-IMGC) were cultivation for 4 days, and exhibited significantly enhanced growth of the mutants in minimal inulin medium (M9S-inulin) compared to the wild type (**Figure 3G**). Moreover, after conjugation and cultivation of all transconjugants in M9S-inulin medium without antibiotics, daily PCR analysis of genomic DNA revealed a progressive enrichment of these engineered cells within the population (**Figure 3H**). Our findings suggest that the STIB can be effectively employed to expand the metabolic capabilities of *Bacteroides* chassis.

Species- and site-specific editing of different *Bacteroides* within a synthetic gut bacterial community

Finally, we applied STIB for precise genome editing of *Bacteroides* within a complex synthetic community consisting of common taxa from the human gut microbiota [50]. We assembled a synthetic microbial community (synCom) comprising 40 representative human gut bacterial species spanning six phyla (**Table S3**). To evaluate the specificity of our system, this synCom included not only the target species but also six non-target *Bacteroides* species. To ensure biocontainment, we engineered an auxotrophic *E. coli* donor strain, S17DMK, by deleting the essential *dapA* gene (dihydrodipicolinate synthetase) in S17-1 λ pir (**Figure S7A**). Comparative analysis demonstrated that S17DMK exhibited significantly higher conjugation efficiency than the commonly used WM3064 donor ($p=0.0137$, Mann-Whitney U test), prompting its use in all subsequent microbiome editing experiments (**Figures S7B and S7C**).

With this experimental framework, we deployed STIB constructs targeting unique genomic loci (*Bt01*, *Bo03*, and *Bv01*) to selectively edit and enrich for *B. thetaiotaomicron*, *B. ovatus*, and *B. vulgatus*, respectively, within the synCom (**Figure 4A**). Following conjugation and positive selection with chloramphenicol, all experimental groups showed significantly enhanced growth compared to negative controls (**Figure S8A**). Metagenomic sequencing was then performed to quantify the enrichment and assess editing specificity at a community-wide level. This analysis revealed a highly selective enrichment for each targeted species: the relative abundance of *B. thetaiotaomicron* increased substantially from 0.3% to 98% (**Figure 4B**), *B. ovatus* from 0.9% to 86% (**Figure 4C**), and *B. vulgatus* from 0.3% to 69% (**Figure 4D**). This corresponds to >100 fold enrichment for targeted *Bacteroides* species (**Figure 4E**).

We isolated the *Bacteroides* mutants on selective BHIS agar. Subsequent PCR analysis and 16S rDNA Sanger sequencing confirmed both the precise site-specific insertions (**Figures 4F, 4G and 4H**) and the correct taxonomic identity of the isolates (**Table S4**). To evaluate STIB's on-target fidelity within the complex community, we performed a mapping analysis of transposon-genome junction reads (**Methods**). This analysis confirmed extremely low cointegration (**Figure S8B**) and exceptional on-target specificity (100%) at the intended genomic loci for all targeted species (**Figures 4I-K**). Furthermore, the insertions were highly precise, consistently localizing to a narrow window between 56 and 63 bp downstream of the PAM site (**Figures 4I-K**).

Collectively, these results demonstrate that STIB enables high-fidelity, species- and site-specific genome editing of *Bacteroides* directly within a complex microbial community. The ability to dramatically enrich and subsequently isolate targeted strains, even those starting from low abundance, highlights STIB's potential as a powerful tool for sequence-guided functional engineering of rare *Bacteroides* targets in their native ecosystems.

Discussion

Bacteroides, as dominant members of the human gut microbiome, making their precise manipulation essential for understanding and engineering the gut microbiome. However, current tools are largely developed for isolates and often depend on cellular HR capabilities. Herein, we developed STIB, a programmable tool that enables site-specific insertion of large genetic cargo across multiple *Bacteroides* species and complex communities. STIB is derived from the ShCAST system, known for efficient, HR-independent genome insertion in *Proteobacteria* [38,51]. However, initial expression of ShCAST in *Bacteroides* resulted in very low specificity and efficiency, coupled with high cointegration ratio – performance significantly inferior to that observed in *E. coli*. Optimization of STIB through the incorporation of nAnil transitioned the system to a "cut-and-paste" mechanism to minimize cointegration, while the fusion of TnsC to Cas12k significantly enhanced targeting specificity and efficiency (**Figure 1**). Comparative analyses across multiple target sites in diverse *Bacteroides*, and using varied cargo size (2.5, 5.5 and 8.4 kb), confirmed STIB's consistently high efficiency (**Figures 2A, B, I; Figure S6**).

Compared with previous genetic tools for *Bacteroides* species, STIB offers simplicity, achieving genomic insertion of large DNA fragments at any desired sites with high specificity (**Table S5**). Tyrosine integrases also enable genomic insertions, however, the insertions are limited to attachment sites and accompanied by co-integration of plasmid backbones.[18] Unlike HR-based methods or those employing smaller transposases, STIB leverages a complex transposome, inherently conferring superior efficiency, particularly when delivering large cargo. STIB utilizes a readily programmable sgRNA enabling high specificity and programmability. The STIB vector architecture eliminates the cumbersome requirement for replicon replacement when adapting the system to non-model *Bacteroides*. Moreover, the transient nature of the STIB vector obviates the need for inducible expression systems and plasmid curing

procedures. As a direct result of these innovations, STIB reduces the time investment in precise genetic manipulation, and significantly streamlines the workflow compared to traditional inducible protocols.

STIB extends previous community editing tools developed for *E. coli* [39,52,53] to the gut commensal *Bacteroides*. It streamlines the targeted enrichment and subsequent isolation of *Bacteroides*, even those present at low relative abundance (~0.3%) within the complex microbiota (**Figure 4**). This capability offers some advantages over traditional labor-intensive isolation methods. We employed a chloramphenicol resistance gene as a selection marker given the intrinsic sensitivity of most *Bacteroides* strains to chloramphenicol [54]. Moreover, we demonstrated a “niche expansion” strategy using an inulin utilization cluster[23], analogous to the porphyran loci used in previous studies[55,56]. Notably, these metabolic loci are often kilobases in size, presenting a challenge for size-limited methods [26]. STIB, by leveraging the large cargo capacity of transposons, is ideal to deliver these payloads, in microbiome engineering.

Nevertheless, STIB has several limitations. STIB inherits constraints typical of CAST systems, such as target immunity interference[57], which can hinder sequential insertions in close proximity. Furthermore, we observed variable efficiency at certain recalcitrant target sites (**Figures 2A and 2I**). Addressing these limitations may focus on integrating advanced sgRNA design algorithms[58] for strain-level resolution or co-expressing activators, such as ribosomal protein S15 or the Pir protein, to enhance system stability[39].

Despite recent advances in metagenomic engineering of complex microbial communities[45], conjugation-based community editing tools still face the challenge of delivery efficiency. Delivery efficiency is inherently limited by low conjugation frequencies(10^{-6})[60], especially between evolutionarily distant species[61]. Recent advancements in high-efficiency conjugation machinery[62,63] or phage-based delivery[52,53,64–66] provide novel delivery strategies with high efficiency and specificity. Building upon the resource of human gut bacteriophages[67,68] and genome editing tools, engineered phage could offer a compelling approach to achieve high-efficiency, targeted microbiome editing *in situ*.

STIB serves as a modular, transient-expression platform that bridges foundational toolkits with *in situ* engineering strategies. Its large cargo capacity facilitates the stable insertion of diverse functional modules, ranging from

CRISPRi circuits[14] for spatiotemporal gene regulation[69] to biosensing devices for real-time diagnostics[70–72]. STIB can be coupled with CRISPR-based biocontainment systems that strictly restrict horizontal gene transfer within complex environments[73]. By combining advanced delivery vectors with STIB's ability to deploy these complex genetic programs, this platform provides a powerful foundation for precision microbiome engineering.

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Lei Dai (lei.dai@sait.ac.cn).

Materials availability

The STIB plasmid from this study is available on WeKwikGene(accession number 0002259).

Data and code availability

The raw NGS data and source code in this study have been deposited in Zenodo (identifier: 10.5281/zenodo.14880538).

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Acknowledgments

We thank Zepeng Qu, Lu Wu and Yonggang Yang for providing bacterial strains and for their contributions to strain cultivation. We thank Weixun Li for valuable discussions. We also acknowledge the use of Gemini for language editing and polishing. The authors reviewed and revised the content and retain full responsibility for the content presented in this paper. This work was supported by the National Key R&D Program of China (2022YFA0912900), the National Natural Science Foundation of China (32471494 and U22A20540 to W. Z.), the National Natural Science Foundation of China (32101176 to Y.L.) and State Key Laboratory of Microbial Technology Open Projects Fund (Project NO. M2024-06).

Author contributions

Conceptualization: Lei Dai, Wei Zhao

Methodology: Yucan Hu, Qingmei Li, Yu Zeng, Yingying Li.

Investigation: Lei Dai, Wei Zhao, Yucan Hu

Visualization: Yucan Hu, Qingmei Li

Supervision: Lei Dai, Wei Zhao

Writing—original draft: Yucan Hu, Qingmei Li

Writing—review & editing: Yucan Hu, Qingmei Li, Yingying Li, Linggang Zheng, Juntao Shen, Xiang Gao, Guo-Ping Zhao, Wei Zhao, Lei Dai.

Declaration of Interests

The authors declare no competing interests.

Declaration of interests

The authors have filed a patent on STIB.

STAR Methods

Materials and Methods

Experimental Model and Study Participant Details

Bacterial strains

E. coli strain S17-1 λpir was used for plasmid construction and conjugation in pure culture. For editing complex microbial communities, the mutant *E. coli* donor strain S17DMK was employed. Conjugation efficiency of S17DMK was benchmarked against *E. coli* WM3064. Recipient strains for STIB-mediated genome editing included *B. thetaiotaomicron* VPI-5482, *B. ovatus* ATCC 8483, *B. salyersiae* DA1247, *B. vulgatus* ATCC 8482, *B. fragilis* NCTC 9343 and BSC606, and *B. uniformis* ATCC 8492. A synthetic community (synCom) comprising 40 human gut bacterial strains was used to model complex gut microbiota. All strains (**Table S3**) were cultivated anaerobically at 37°C.

The *E. coli* strain S17-1 λ pir was cultured in LB medium. WM3064 was cultured in LB medium with 60 μg/mL diaminopimelic acid (DAP). The mutate *E. coli* strain S17DMK was cultured in LB medium with 60 μg/mL DAP and 50 μg/mL Kanamycin.

Wild type strains of *Bacteroides* in **Table S3** (*B. thetaiotaomicron*, *B. ovatus*, *B. fragilis*, *B. salyersiae* and *B. vulgatus*) were grown in Brain Heart Infusion medium (BHI, Oxoid Cat#CM1135) with Vitamin K3 (final concentration with 0.5 mg/L) and hemin (final concentration with 5 mg/L), which called BHIS. Bacteria were cultured in anaerobic chamber (Coy, specify model; 85% N_2, 10% H_2 and 5% CO_2). The mutate *B. thetaiotaomicron* for sequential editing was cultured in BHIS with chloramphenicol or/and erythromycin. In addition, the mutate Bt-IMGC would culture in M9S (*Bacteroides* minimal medium) with 0.5%(w/v) inulin (Orafti® HP, ~94.5%, average DP >= 23) in a special need.

Method Details

Plasmid construction

All plasmids and primers used in this study were listed in supplemental **Table S6 and S7**, separately.

The reported ShCAST system was operon redesigned and codon optimized using OPTIMIZER[74] (genomes.urv.es), and then was synthesized by BGI Genomics Co., Ltd. In addition, nAnil was ordered from Beijing Tsingke Biotech Co., Ltd. The design details were complemented by **Supplementary Document1**.

The ShCAST operon and its corresponding plasmid backbone were synthesized to generate the psh2200 construct, in which transcription of cas12k, the transposase genes, and the sgRNA was driven by the P_{BfP1E6} promoter. To reduce co-integration events, the nAnil nickase was fused to TnsB, generating the psh3300 construct. Subsequently, Cas12K and TnsC were co-expressed as a fusion protein, resulting in the STIB plasmid psh4300. All the new plasmids were generated via Gibson assembly with pEASY®-Basic Seamless Cloning and Assembly Kit (TransGen Biotech Co., Ltd). All sgRNA targeting different sites were replaced into basic plasmids of different systems via Golden Gate with NEBridge® Golden Gate Assembly kits (using BsaI-HFv2) (New England Biolabs, NEB).

Evaluation of expression element strength

Expression elements consisting of various promoter–RBS combinations fused to a synthetic NanoLuc luciferase gene (Beijing Tsingke Biotech Co., Ltd.) were assembled via Gibson Assembly and inserted into the pB025 vector between the erythromycin resistance gene and the origin of replication. Non-essential

elements in pB025 were removed during vector construction. The resulting plasmids were transformed into *E. coli* S17-1 λ pir as the donor strain.

Conjugation was performed by mixing donor (mid-log phase) and recipient *B. thetaiotaomicron* (late-log phase) at a 1:1 ratio. The mixture was pelleted by centrifugation (4,000 × g, 20 min), resuspended in 1 mL of BHIS medium, and spotted onto the center of a BHIS agar plate. After aerobic incubation at 37 °C for 20 h, transconjugants were harvested in 1 mL of medium, serially diluted, and 100 µL was plated onto BHIS agar supplemented with 200 µg/mL gentamicin and 25 µg/mL erythromycin. Colonies were incubated anaerobically for 3 days, and putative transconjugants were verified. Confirmed mutants were restreaked for isolation, inoculated into BHIS liquid culture, and grown overnight. Cultures were then diluted 1:100 and grown for subsequent NanoLuc luciferase assays.

The NanoLuc luciferase assay was performed as previously described[24]. Briefly, *Bacteroides* strains were grown to exponential phase (OD_{600} = 0.3–0.5). 500 µL of culture was harvested by centrifugation at 13,000 g for 5 min, washed with PBS and the result pellets were resuspended in 50 µL of 1× BugBuster (Novagen). Cells were lysed by incubation at 26 °C for 10 min. Cellular debris was removed by centrifugation at 21,130 × *g* for 10 min at 4 °C. Fifty microliters of the resulting supernatant was mixed with an equal volume of NanoLuc Reaction Buffer (Promega) and incubated at 26 °C for 10 min to initiate luminescence. Luminescence and the corresponding OD_{600} of the original culture (200 µL) were measured using a microplate reader, and luciferase activity was expressed as the luminescence-to-OD_{600} ratio.

For each independent experiment, these ratios were normalized to the sum of all group values within that experiment (i.e., divided by the total sum), yielding relative proportions that sum to 1 per experiment. Non-parametric analyses (Kruskal-Wallis with Dunn's post-hoc tests and Bonferroni correction) were employed due to violated normality and variance homogeneity assumptions. Significance threshold: $p < 0.05$.

Guide RNA design

All guide RNAs (gRNAs) were designed with 23-nucleotide (nt) spacers and a Cas12k-compatible protospacer adjacent motif (PAM: 5′-GTN-3′). For *Bacteroides thetaiotaomicron* transposition assays, gRNAs were designed to direct transposon insertion into the first third to two-thirds of target genes. For

other *Bacteroides* species, neutral genomic sites were identified using TargetFinder[75] (https://github.com/ECBCgit/targetFinder). Sequences from these neutral sites were analyzed to extract 23-nt protospacers immediately downstream of a conserved PAM motif (5′-GTN-3′), which were selected as candidate gRNAs. Potential off-target effects were evaluated using Cas-OFFinder v2.4[76], allowing a maximum of 5 mismatches across the genome.

Transposition experiment in *Bacteroides*

The specific vectors were transformed into *E. coli* S17-1 λ pir as the donor strain. Overnight donor cultures were diluted 1:100 in LB medium supplemented with 100 μg/mL ampicillin and incubated aerobically at 37°C for 3 hours to reach an OD_{600} of 0.2–0.5. Recipient *Bacteroides* cultures (50 μL) were inoculated into 5 mL BHIS medium and grown anaerobically at 37°C until reaching an OD_{600} of 0.6–0.8. Donor cells were washed twice with PBS via centrifugation (4,000 × *g*, 10 min), mixed with recipient cultures at a 1:1 ratio (1 mL/1 mL), pelleted (4,000 × *g*, 20 min), and resuspended in 100 μL BHIS medium. The cell mixture was spotted onto BHIS-agar plates supplemented with 8% horse blood and performed conjugation at 37°C under aerobic or anaerobic conditions for 16–24 hours. Mating lawns were harvested, resuspended in 1 mL BHIS medium, serially diluted (e.g., to 10^{-1}, 10^{-2}, and 10^{-3}), and then 100 μL of each bacterial suspension was plated on BHIS-agar plates containing 200 μg/mL gentamicin and species-specific chloramphenicol concentrations (25 μg/mL for *B. thetaiotaomicron* and *B. ovatus*; 12.5 μg/mL for other *Bacteroides* spp.). After 2–3 days of anaerobic growth, 10–20 colonies were selected for PCR-based verification of editing efficiency. For *B. thetaiotaomicron*, all remaining colonies were pooled for Tn-seq and ddPCR analysis.

Counterselection assay based on FUdR

sgRNA26 was designed for targeting *Bt01* site and assembled into the plasmid of psh2200 and transformed into S17-1 λ pir. Then the transposition experiment was performed as previous described with small modification. After conjugation, the collected mating lawns were streaked on BHI-agar plates with 200 μg/mL FUdR, 200 μg/mL gentamicin and 25 μg/mL chloramphenicol. Selected clones were inoculated in liquid broth and cultured overnight. Subsequently, bacterial cultures were streaked onto selective agar plates containing 200 μg/mL FUdR and 200 μg/mL gentamicin. Meanwhile, wide type *B.thetaiotaomicron* was also

streaked on the other half of the plate. Cultured 2~3 days at 37℃ in anaerobic condition.

Evaluation of cointegration and efficiency based on PCR

PCR reactions were conducted in 10 μL volumes to screen colonies selected after transposition experiments. Briefly, individual clones were inoculated into BHIS liquid medium and cultured overnight in an anaerobic incubator. Subsequently, 20 μL of the cultured cells were mixed with 180 μL of sterilized deionized water to serve as the PCR template. The reactions were performed using 2× ProTaq Master Mix (dye plus) (Accurate Biology). Two primer pairs (**Table S7**) were utilized to assess insertion efficiency and cointegration. The first pair, flanking the upstream and downstream regions of the target site, was used to evaluate insertion efficiency. The second pair, specific to the vector backbone, was employed to assess cointegration. Primers were used at a concentration of 0.5 μM, with wild-type *B.thetaiotaomicron* serving as a negative control. PCR amplification was carried out for 32 cycles, and the resulting amplicons were resolved by electrophoresis on 1% agarose gels stained with SYBR Safe (Thermo Fisher Scientific). One-way ANOVA was used for single-factor analyses and two-way ANOVA was used for two-factor analyses, with Tukey's post-hoc test applied when significant main effects were observed.

Modified Tn-seq library preparation

The colonies grown from BHI-agar plates plus 25 μg/mL chloramphenicol after transposition experiment were scraped and DNA extracted by Ezup Column Bacteria Genomic DNA Purification Kit (Sangon Biotech) or QIAamp PowerFecal Pro DNA Kit (QIAGEN). The library preparation followed protocol for NEBNext Ultra II FS DNA Library Prep Kit for Illumina (NEB) with some modifications. 200 ng gDNA from each biological replicate (n=3) was mixed. Total 600 ng gDNA was used as input and incubated for 8 minutes at 37℃ for fragmentation. After adapter ligation, the adapter-ligated fragments were extracted using 1× Hieff NGS DNA Selection Beads (YEASEN). Next, genome-transposon junctions were enriched by PCR with NEBNext Ultra II Q5 Master Mix, using adapter- and transposon-LE (left end)- specific primers (**Table S7**). Cycling conditions were: 98 ℃ for 30 s; 25 cycles of 98 ℃ for 10 s, 65 ℃ for 75 s; and one cycle of 65 ℃ for 5 min. The PCR product that underwent purification with 0.9× selection beads was used as template into a second, six-cycle PCR, for adding dual index. Then, a

further 0.6× /0.2× size selection was used and the library was eluted in 30 μL ddH_2O. Finally, libraries were quantified by Qubit Fluorimeter and submitted to sequencing on Novaseq 6000 platform, with a target sequencing depth of 2 Gb raw data for each sample.

Analysis of Tn-seq data

The Tn-seq data analysis was performed using a custom shell script and a bioinformatics pipeline based on the BBMap suite (v39.01). The pipeline begins with quality filtering of raw paired-end reads using bbduk.sh, retaining only reads with a minimum Phred score of 20. Filtered reads were then merged using bbmerge-auto.sh with the vstrict model to ensure high accuracy, followed by adapter removal. The merged reads were enriched for specific junction sequences (agataatttgtcactgtaca) using bbduk.sh, with unmatched reads discarded.

Cointegration events were identified by retaining reads containing the plasmid backbone sequence (ccagatgtcaacacagctac). The counts of cointegration and genome insertion reads were recorded using the nohup command. The cointegration ratio can be further estimated with the following formula:

(1)

In formula, represents the ratio of cointegration in sample i. represents the number of reads containing plasmid backbone sequence. represents the number of reads without plasmid backbone.

Reads associated with cointegration were excluded from further analysis. The remaining reads were trimmed to remove the LE sequence and filtered to a minimum length of 60 bp. These reads were further trimmed to a fixed length of 60 bp using bioawk (https://github.com/lh3/bioawk) and mapped to the reference genome (NC_004663.1) using bbmap.sh with strict parameters. Unique reads were extracted, and those mapping to the reference genome were isolated for downstream analysis. Off-target reads were identified based on their position relative to the target window (100 bp from the end of PAM. e.g., target window of *BT01*: 2855853~2855953), and their reverse complements were extracted for sequence logo analysis using the ggseqlogo package[77] in R (v4.3.2). Insertion sites and their frequencies were quantified, and the orientation of reads within

the target region was analyzed to complete the pipeline. Finally, all data were further processed, analyzed and visualized in R (v4.3.2) and inkscape (v1.4.2)[78].

Evaluation of editing efficiency via digital droplet PCR

In this study, *BT1763* was utilized as the internal reference gene, while a LE-junction fragment served as the target gene for detection. Droplet digital PCR (ddPCR) reactions were prepared in a 20 μL volume containing 0.5 ng of genomic DNA from transposition experiment, 500 nM of each primer, 400 nM of specific probes (see **Table S7**), and ddPCR supermix (Shenzhen Biorain Technology Co., Ltd). The reaction mixture was loaded into C4 chips and processed using the DropXpert S6 system (Shenzhen Biorain Technology Co., Ltd). The automated workflow included droplet generation, thermocycling, and quantification. The thermal cycling protocol consisted of an initial step at 50°C for 10 minutes and 95°C for 3 minutes, followed by 40 cycles of 95°C for 15 seconds and 58°C for 30 seconds. All samples were analyzed in parallel, and the efficiencies were calculated by (junction/reference) × 100%. One-way ANOVA was used.

Sequential editing of *Bacteroides thetaiotaomicron*

The round 1 editing used *E. coli* S17-1 λ pir cells containing psh4326C as donor and performed transposition experiment in *B. thetaiotaomicron* as previous described with some modification. The seed donor strain was cultured overnight, and then diluted into fresh LB medium with 1:100 culturing at 37℃ with 220 rpm until OD_{600} around 0.4. Meanwhile, seed of *B. thetaiotaomicron* cultured overnight from single clone was also diluted into fresh BHI medium with 1:100 at 37℃ in anaerobic glove box until OD_{600} around 0.6. The donor was collected and washed twice with PBS, and the 2 mL *B. thetaiotaomicron* was added to resuspend the donor pellet softly. After that, centrifuge with 4,000 rpm for 15 minutes at room temperature and discarding supernatant. Finally, 100 μL BHIS medium was used for resuspension and drop on the center of the BHIS-blood agar plate. Conjugation was achieved by culturing at 37 ℃ in aerobic condition for 24 hours. All bacterial lawn was scraped off and resuspended with 1 mL BHIS liquid medium for preparation the gradient dilution at 10^{-1}, 10^{-2} and 10^{-3}, and then 100 μL bacterial suspension was plated on BHIS agar with chloramphenicol at 25 μg/mL for selection the edited *B. thetaiotaomicron* at 37 ℃ in anaerobic condition for 48 hours. Clones were picked into BHI liquid medium with 25 μg/mL chloramphenicol and cultured overnight. PCR was performed with specific

primers (**Table S7**) to check the clones edited at the target site of *Bt01*. Round 2 editing was performed with the *B. thetaiotaomicron* edited as the recipient strain and *E. coli* S17-1 λ pir cells containing psh4339E as the donor targeting at *Bt02* site, following the same protocol as described above.

Measurement of growth curve and fluorescence

To measure the growth of bacteria, 2 µL of overnight bacterial cultures were combined with 200 µL of fresh BHIS medium supplemented with different antibiotics (25 µg/mL erythromycin, 25 µg/mL chloramphenicol) and transferred into wells of a 96-well microplate (Corning, Cat#3599, NY). The microplate was covered with a lid and placed in a microplate reader (Epoch2, BioTek Ins, U.S.A.) to monitor the optical density (OD) at 600 nm. The plates were maintained at 37°C and subjected to continuous shaking at 355 rpm. OD measurements were taken every 15 minutes for 24 hours. For experiments using M9S-inulin medium, OD readings were recorded every hour for 4 days.

To quantify fluorescence intensity in *Bacteroides*, 2 µL of overnight bacterial cultures were resuspended in 1 or 2 mL of phosphate-buffered saline (PBS) and exposed to oxygen at 37°C for at least 1 hour. Fluorescent signals were analyzed using a flow cytometer (CytoFLEX S, Beckman Coulter, USA) equipped with the FITC channel (excitation: 488 nm, emission: 525/40 nm). Data acquisition and analysis were performed using CytExpert 2.3 software (Beckman Coulter, USA). Samples were processed with a flow rate of 10 µL/min, and a minimum of 10,000 events per sample were recorded to ensure statistical robustness. Gating strategies were applied to exclude debris and aggregates based on forward scatter (FSC) and side scatter (SSC) parameters.

Construction and validation of auxotrophic donor strain S17DMK

The auxotrophic donor strain S17DMK was derived from *E. coli* S17-1 λ pir through targeted modification of the *dapA* gene using the λ-Red recombineering system [79]. Specifically, a synthetic cassette encoding constitutively expressed mScarlet-I and a kanamycin resistance gene (*kanR*) was constructed. This cassette was flanked by ~1000 bp homology arms corresponding to the regions upstream and downstream of the *dapA* gene in the S17-1 λ pir genome, designed to replace *dapA* with the mScarlet-*kanR* marker (Δ*dapA*::mScarlet-*kanR*).

For recombineering, 200 ng of the mScarlet-*kanR* cassette DNA was electroporated into competent S17-pSC101-BAD-gbaA-tet cells. Following electroporation, cells were allowed to recover in 700 µL of LB medium at 37°C for 1 hour before being plated onto LB agar supplemented with kanamycin and diaminopimelic acid (DAP). Kanamycin-resistant colonies were subsequently screened by PCR to verify correct genomic integration of the mScarlet-*kanR* cassette and deletion of *dapA*.

The delivery efficiency of the newly constructed S17DMK donor strain was compared to that of the commonly used auxotrophic donor strain, WM3064. Transposition experiments were performed as described. Briefly, donor and recipient strains were prepared and mixed, and conjugation was allowed to proceed. Following conjugation, appropriate serial dilutions of the mating mixture were plated onto selective media to enumerate transconjugants. Colony-forming units (CFUs) were counted to determine and compare the conjugation efficiencies of S17DMK and WM3064. Group comparisons were performed using Mann-Whitney U test after confirming violations of normality and homogeneity of variance assumptions.

Competition between wild-type and mutant *B. thetaiotaomicron* strains

E. coli S17DMK containing the plasmid psh4326-cmr-inulinase as the donor and *B. thetaiotaomicron* as the recipient. Conjugation was performed as descripted above in enrichment of edits by antibiotic in complex microbiota. Some differences were the lawn suspending with and transferred to M9S plus 0.5% (w/v) inulin, and the start OD_{600} of outgrowth was no less than 0.1. Furthermore, cultures were maintained over a 4-day period with daily collection of time-series samples (including day 0 to establish baseline conditions). These samples were cryopreserved at -80°C for subsequent genomic DNA extraction and PCR analysis.

Genome editing of *Bacteroides* species in a synthetic gut bacterial community

Briefly, the synthetic community (synCom) consisted of 40 species (**Table S3**), all of which were sensitive to chloramphenicol at a concentration of 50 µg/mL. Each species was cultured individually and harvested at an equal volume when the OD_{600} reached 0.3. Five distinct microbial communities were constructed by combining equal volumes of these cultures. The communities were preserved by

freezing at -80°C. To assemble the synCom, 50 µL of each seed community was inoculated into 10 mL of BHIS medium and incubated for 1~2 days in an anaerobic chamber. The resulting synCom was then preserved at -80°C.

For editing of *B. thetaiotaomicron* in synCom, the plasmid psh4326C targeting *Bt01* was transformed into *E. coli* S17DMK as the donor strain, cultured aerobically in LB medium supplemented with DAP (60 µg/mL) and ampicillin (100 µg/mL) until reaching OD_{600} 0.2–0.5, washed twice with PBS, and mixed with recipient synthetic communities (pre-cultured anaerobically in BHIS medium for 30–40 hours) at a 1:1 ratio (2 OD_{600}×mL each). The mixture was pelleted, resuspended in BHIS medium, spotted on a 0.45 or 0.22 µm Millipore filter membrane (Sangon Biotech) over BHIS-agar with DAP, and incubated anaerobically at 37°C for 36–40 hours. Conjugation lawns were resuspended in BHIS medium containing chloramphenicol (50 µg/mL), diluted 1:100 for antibiotic enrichment, and after anaerobic growth to obvious turbidity visible, subcultures were analyzed via shallow metagenomic sequencing and streaked onto BHIS-agar with chloramphenicol (25 µg/mL) to isolate colonies for PCR verification of target sites.

For editing of *B. ovatus* and *B. vulgatus*, 100 µL of synCom stocks and 1.5 µL of target species stocks were first mixed and the mixture was inoculated into 10 mL of BHIS medium. Cultivation proceeded at 37°C for two days under anaerobic conditions. Subsequently, cultures were harvested to prepare recipients: 2 OD_{600}×mL of culture for *B. ovatus* editing and 4 OD_{600}×mL for *B. vulgatus* editing. The donor plasmids psh4365 (targeting *Bo03*) and psh4376 (targeting *Bv01*) were used for editing *B. ovatus* and *B. vulgatus*, respectively, following the procedure described above, except that the chloramphenicol concentration during antibiotic enrichment was 25 µg/mL.

Metagenomic library preparation and analysis

The Hieff NGS® OnePot Pro DNA Library Prep Kit V2® (Yeasen, Cat#12195ES96) was used for library preparation, following the manufacturer's instructions. The resulting library DNA was cleaned up and size-selected with Hieff NGS® DNA Selection Beads (Yeasen, Cat#12601ES56), and quantified using the dsDNA High Sensitivity kit on a Qubit (Thermo Fisher). Libraries were further pooled together at equal molar ratios, and the purity and library length distribution were assessed using Bioanalyzer High Sensitivity DNA Kit (Agilent).

Sequencing was performed on the DipSeq (150 bp paired-end reads; BGI Genomics Co., Ltd. , and BerryGenomics Co., Ltd.), with a target sequencing depth of 1-5 Gbs raw data for each sample.

Taxonomic profiling was performed using a custom shell script to automate the analysis of sequencing data with Kraken2 (v2.1.2)[80] and Bracken (v2.6.1)[81]. Initially, raw sequencing files were preprocessed using fastp (v0.23.2)[82] for quality control. Taxonomic classification was then conducted using Kraken2 with a synthetic community (synCom) reference database containing 40 genomes and 30 threads for parallel processing. Species-level abundance estimation was subsequently performed using Bracken, with a read length of 150 bp and 16 threads for computation. The Bracken output was aggregated and formatted, incorporating metadata such as sample name, taxonomy ID, taxonomy level, Kraken-assigned reads, added reads, new estimated reads, relative abundance (RA), and group information.

In parallel, off-target analysis was carried out following the Tn-seq pipeline, with the mapping reference database replaced by the complete genome of the synthetic community (synCom), and increased the length of junction sequence matched (25 nt, sequence: ACGACAGATAATTTGTCACTGTACA) for avoiding false positive mapping. The all results were manually inspected to determine whether off-target effects occurred in unintended genomes.

Quantification and Statistical Analysis

Quantification, statistical analysis, and data visualization were performed in R using aplot[83], ggplot2, lubridate[84], reshape2[85], stringr (v1.5.2), ggridges (v0.5.7), ggsignif[86], ggalluvial[87], RColorBrewer (v1.1-3), dplyr (v1.1.4), tidyr (v1.3.1), ggseqlogo (v0.2)[77], patchwork (v1.3.2), and emmeans (v2.0.0). Statistical tests were conducted using a custom R function. Normality was evaluated by the Shapiro–Wilk test, and homogeneity of variance was assessed using F-tests or Levene’s test. When assumptions were met, parametric tests (two-sided Student’s t-test or one-way ANOVA with Tukey’s post hoc test) were applied; otherwise, non-parametric tests (Mann–Whitney U test or Kruskal–Wallis test with Dunn’s post hoc test and Bonferroni correction) were used.

For Figure S3D, Figure 1, Figure 2B, and Figure 2I, one-way ANOVA was applied despite limited sample sizes. Given the small number of observations, assumption tests have limited statistical power; therefore, a consistent parametric framework was used to facilitate comparison across conditions.

References

1. Wexler, A.G., and Goodman, A.L. (2017). An insider's perspective: Bacteroides as a window into the microbiome. Nat Microbiol *2*, 1–11. https://doi.org/10.1038/nmicrobiol.2017.26.

2. Zafar, H., and Saier, M.H. (2021). Gut *Bacteroides* species in health and disease. Gut Microbes *13*, 1848158. https://doi.org/10.1080/19490976.2020.1848158.

3. Wang, K., Zhang, Z., Hang, J., Liu, J., Guo, F., Ding, Y., Li, M., Nie, Q., Lin, J., Zhuo, Y., et al. (2023). Microbial-host-isozyme analyses reveal microbial DPP4 as a potential antidiabetic target. Science *381*, eadd5787. https://doi.org/10.1126/science.add5787.

4. Jiang, K., Li, W., Tong, M., Xu, J., Chen, Z., Yang, Y., Zang, Y., Jiao, X., Liu, C., Lim, B., et al. (2023). Bacteroides fragilis ubiquitin homologue drives intraspecies bacterial competition in the gut microbiome. Nat Microbiol, 1–15. https://doi.org/10.1038/s41564-023-01541-5.

5. Qu, Z., Liu, H., Yang, J., Zheng, L., Huang, J., Wang, Z., Xie, C., Zuo, W., Xia, X., Sun, L., et al. (2025). Selective utilization of medicinal polysaccharides by human gut Bacteroides and Parabacteroides species. Nat Commun *16*, 638. https://doi.org/10.1038/s41467-025-55845-7.

6. Zhang, Z.J., Cole, C.G., Coyne, M.J., Lin, H., Dylla, N., Smith, R.C., Pappas, T.E., Townson, S.A., Laliwala, N., Waligurski, E., et al. (2024). Comprehensive analyses of a large human gut Bacteroidales culture collection reveal species- and strain-level diversity and evolution. Cell Host & Microbe. https://doi.org/10.1016/j.chom.2024.08.016.

7. Lee, S.M., Donaldson, G.P., Mikulski, Z., Boyajian, S., Ley, K., and Mazmanian, S.K. (2013). Bacterial colonization factors control specificity and stability of the gut microbiota. Nature *501*, 426–429. https://doi.org/10.1038/nature12447.

8. Lai, Y., Hayashi, N., and Lu, T.K. (2022). Engineering the human gut commensal *Bacteroides thetaiotaomicron* with synthetic biology. Current Opinion in Chemical Biology *70*, 102178. https://doi.org/10.1016/j.cbpa.2022.102178.

9. Farrar, M.D., Whitehead, T.R., Lan, J., Dilger, P., Thorpe, R., Holland, K.T., and Carding, S.R. (2005). Engineering of the gut commensal bacterium Bacteroides ovatus to produce and secrete biologically active murine interleukin-2 in response to xylan. Journal of Applied Microbiology *98*, 1191–1197. https://doi.org/10.1111/j.1365-2672.2005.02565.x.

10. Hamady, Z.Z.R., Scott, N., Farrar, M.D., Lodge, J.P.A., Holland, K.T., Whitehead, T., and Carding, S.R. (2010). 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. https://doi.org/10.1136/gut.2008.176131.

11. Jin, W.-B., Li, T.-T., Huo, D., Qu, S., Li, X.V., Arifuzzaman, M., Lima, S.F., Shi, H.-Q., Wang, A., Putzel, G.G., et al. (2022). Genetic manipulation of gut microbes enables single-gene interrogation in a complex microbiome. Cell *185*, 547-562.e22. https://doi.org/10.1016/j.cell.2021.12.035.

12. Chen, B., Sun, L., Zeng, G., Shen, Z., Wang, K., Yin, L., Xu, F., Wang, P., Ding, Y., Nie, Q., et al. (2022). Gut bacteria alleviate smoking-related NASH by degrading gut nicotine. Nature *610*, 562–568. https://doi.org/10.1038/s41586-022-05299-4.

13. Whitaker, W.R., Shepherd, E.S., and Sonnenburg, J.L. (2017). Tunable Expression Tools Enable Single-Cell Strain Distinction in the Gut Microbiome. Cell *169*, 538-546 e12. https://doi.org/10.1016/j.cell.2017.03.041.

14. Mimee, M., Tucker, A.C., Voigt, C.A., and Lu, T.K. (2015). Programming a Human Commensal Bacterium, Bacteroides thetaiotaomicron, to Sense and Respond to Stimuli in the Murine Gut Microbiota. Cell Systems *1*, 62–71. https://doi.org/10.1016/j.cels.2015.06.001.

15. Goodman, A.L., McNulty, N.P., Zhao, Y., Leip, D., Mitra, R.D., Lozupone, C.A., Knight, R., and Gordon, J.I. (2009). Identifying Genetic Determinants Needed to Establish a Human Gut Symbiont in Its Habitat. Cell Host & Microbe *6*, 279–289. https://doi.org/10.1016/j.chom.2009.08.003.

16. Liu, H., Shiver, A.L., Price, M.N., Carlson, H.K., Trotter, V.V., Chen, Y., Escalante, V., Ray, J., Hern, K.E., Petzold, C.J., et al. (2021). Functional genetics of human gut commensal Bacteroides thetaiotaomicron reveals metabolic requirements for growth across environments. Cell Reports *34*, 108789. https://doi.org/10.1016/j.celrep.2021.108789.

17. Wu, M., McNulty, N.P., Rodionov, D.A., Khoroshkin, M.S., Griffin, N.W., Cheng, J., Latreille, P., Kerstetter, R.A., Terrapon, N., Henrissat, B., et al. (2015). Genetic determinants of in vivo fitness and diet responsiveness in multiple human gut Bacteroides. Science *350*, aac5992. https://doi.org/10.1126/science.aac5992.

18. Wang, J., Shoemaker, N.B., Wang, G.R., and Salyers, A.A. (2000). Characterization of a Bacteroides mobilizable transposon, NBU2, which carries a functional lincomycin resistance gene. J Bacteriol *182*, 3559–3571. https://doi.org/10.1128/JB.182.12.3559-3571.2000.

19. Rajeev, L., Salyers, A.A., and Gardner, J.F. (2006). Characterization of the integrase of NBU1, a Bacteroides mobilizable transposon. Mol

Microbiol *61*, 978–990. https://doi.org/10.1111/j.1365-2958.2006.05282.x.

20. Taketani, M., Zhang, J., Zhang, S., Triassi, A.J., Huang, Y.-J., Griffith, L.G., and Voigt, C.A. (2020). Genetic circuit design automation for the gut resident species Bacteroides thetaiotaomicron. Nat Biotechnol *38*, 962–969. https://doi.org/10.1038/s41587-020-0468-5.

21. Koropatkin, N.M., Martens, E.C., Gordon, J.I., and Smith, T.J. (2008). Starch Catabolism by a Prominent Human Gut Symbiont Is Directed by the Recognition of Amylose Helices. Structure *16*, 1105–1115. https://doi.org/10.1016/j.str.2008.03.017.

22. Bencivenga-Barry, N.A., Lim, B., Herrera, C.M., Trent, M.S., and Goodman, A.L. (2020). Genetic Manipulation of Wild Human Gut *Bacteroides*. J Bacteriol *202*. https://doi.org/10.1128/JB.00544-19.

23. García-Bayona, L., and Comstock, L.E. (2019). Streamlined Genetic Manipulation of Diverse *Bacteroides* and *Parabacteroides* Isolates from the Human Gut Microbiota. mBio *10*, e01762-19. https://doi.org/10.1128/mBio.01762-19.

24. Zheng, L., Tan, Y., Hu, Y., Shen, J., Qu, Z., Chen, X., Ho, C.L., Leung, E.L.-H., Zhao, W., and Dai, L. (2022). CRISPR/Cas-Based Genome Editing for Human Gut Commensal *Bacteroides* Species. ACS Synth. Biol. *11*, 464–472. https://doi.org/10.1021/acssynbio.1c00543.

25. Zhan, Y., Zheng, L., Shen, J., Hu, Y., Luo, X., and Dai, L. (2025). Development of a miniaturized CRISPR/Cas gene editing tool for human gut Bacteroides. Sheng Wu Gong Cheng Xue Bao *41*, 2360–2372. https://doi.org/10.13345/j.cjb.240959.

26. Ramachandran, G., and Bikard, D. (2019). Editing the microbiome the CRISPR way. Phil. Trans. R. Soc. B *374*, 20180103. https://doi.org/10.1098/rstb.2018.0103.

27. Reuter, M., Parry, F., Dryden, D.T.F., and Blakely, G.W. (2010). Single-molecule imaging of Bacteroides fragilis AddAB reveals the highly processive translocation of a single motor helicase. Nucleic Acids Res *38*, 3721–3731. https://doi.org/10.1093/nar/gkq100.

28. Liang, J., and Tan, Y. (2023). Highly efficient CRISPR-mediated base editing for the gut Bacteroides spp. with pnCasBS-CBE. Biotechnology Journal *18*, 2200504. https://doi.org/10.1002/biot.202200504.

29. Beller, Z.W., Wesener, D.A., Seebeck, T.R., Guruge, J.L., Byrne, A.E., Henrissat, S., Terrapon, N., Henrissat, B., Rodionov, D.A., Osterman, A.L., et al. (2023). Inducible CRISPR-targeted “knockdown” of human gut Bacteroides in gnotobiotic mice discloses glycan utilization strategies. Proceedings of the National Academy of Sciences *120*, e2311422120. https://doi.org/10.1073/pnas.2311422120.

30. Chen, C., Li, Y.-W., Zheng, Y.-Y., Li, X.-J., Wu, N., Guo, Q., Shi, T.-Q., and Huang, H. (2025). Expanding the frontiers of genome engineering: A comprehensive review of CRISPR-associated transposons. Biotechnology Advances *78*, 108481. https://doi.org/10.1016/j.biotechadv.2024.108481.

31. Strecker, J., Ladha, A., Gardner, Z., Schmid-Burgk, J.L., Makarova, K.S., Koonin, E.V., and Zhang, F. (2019). RNA-guided DNA insertion with CRISPR-associated transposases. Science *365*, 48–53. https://doi.org/10.1126/science.aax9181.

32. Klompe, S.E., Vo, P.L.H., Halpin-Healy, T.S., and Sternberg, S.H. (2019). Transposon-encoded CRISPR–Cas systems direct RNA-guided DNA integration. Nature *571*, 219–225. https://doi.org/10.1038/s41586-019-1323-z.

33. Vo, P.L.H., Ronda, C., Klompe, S.E., Chen, E.E., Acree, C., Wang, H.H., and Sternberg, S.H. (2021). CRISPR RNA-guided integrases for high-efficiency, multiplexed bacterial genome engineering. Nat Biotechnol *39*, 480–489. https://doi.org/10.1038/s41587-020-00745-y.

34. Cheng, Z.-H., Wu, J., Liu, J.-Q., Min, D., Liu, D.-F., Li, W.-W., and Yu, H.-Q. (2022). Repurposing CRISPR RNA-guided integrases system for one-step, efficient genomic integration of ultra-long DNA sequences. Nucleic Acids Research *50*, 7739–7750. https://doi.org/10.1093/nar/gkac554.

35. Chen, W., Ren, Z.-H., Tang, N., Chai, G., Zhang, H., Zhang, Y., Ma, J., Wu, Z., Shen, X., Huang, X., et al. (2021). Targeted genetic screening in bacteria with a Cas12k-guided transposase. Cell Reports *36*, 109635. https://doi.org/10.1016/j.celrep.2021.109635.

36. Cui, Y., Dong, H., Tong, B., Wang, H., Chen, X., Liu, G., and Zhang, D. (2022). A versatile Cas12k-based genetic engineering toolkit (C12KGET) for metabolic engineering in genetic manipulation-deprived strains. Nucleic Acids Research *50*, 8961–8973. https://doi.org/10.1093/nar/gkac655.

37. Yang, S., Zhang, Y., Xu, J., Zhang, J., Zhang, J., Yang, J., Jiang, Y., and Yang, S. (2021). Orthogonal CRISPR-associated transposases for parallel and multiplexed chromosomal integration. Nucleic Acids Research *49*, 10192–10202. https://doi.org/10.1093/nar/gkab752.

38. Trujillo Rodríguez, L., Ellington, A.J., Reisch, C.R., and Chevrette, M.G. (2023). CRISPR-Associated Transposase for Targeted Mutagenesis in Diverse Proteobacteria. ACS Synth. Biol. https://doi.org/10.1021/acssynbio.3c00065.

39. Rubin, B.E., Diamond, S., Cress, B.F., Crits-Christoph, A., Lou, Y.C., Borges, A.L., Shivram, H., He, C., Xu, M., Zhou, Z., et al. (2021). Species- and site-specific genome editing in complex bacterial

communities. Nat Microbiol *7*, 34–47. https://doi.org/10.1038/s41564-021-01014-7.

40. Gelsinger, D.R., Vo, P.L.H., Klompe, S.E., Ronda, C., Wang, H.H., and Sternberg, S.H. (2024). Bacterial genome engineering using CRISPR-associated transposases. Nat Protoc, 1–39. https://doi.org/10.1038/s41596-023-00927-3.

41. Tou, C.J., Orr, B., and Kleinstiver, B.P. (2023). Precise cut-and-paste DNA insertion using engineered type V-K CRISPR-associated transposases. Nat Biotechnol, 1–12. https://doi.org/10.1038/s41587-022-01574-x.

42. Strecker, J., Ladha, A., Makarova, K.S., Koonin, E.V., and Zhang, F. (2020). Response to Comment on “RNA-guided DNA insertion with CRISPR-associated transposases.” Science *368*, eabb2920. https://doi.org/10.1126/science.abb2920.

43. Yang, S., Zhu, J., Zhou, X., Zhang, J., Li, Q., Bian, F., Zhu, J., Yan, T., Wang, X., Zhang, Y., et al. (2023). RNA-Guided DNA Transposition in Corynebacterium glutamicum and Bacillus subtilis. ACS Synth. Biol. https://doi.org/10.1021/acssynbio.3c00193.

44. Pechenov, P.Y., Garagulya, D.A., Stanovov, D.S., and Letarov, A.V. (2022). New Effective Method of Lactococcus Genome Editing Using Guide RNA-Directed Transposition. International Journal of Molecular Sciences *23*, 13978. https://doi.org/10.3390/ijms232213978.

45. Ronda, C., Chen, S.P., Cabral, V., Yaung, S.J., and Wang, H.H. (2019). Metagenomic engineering of the mammalian gut microbiome in situ. Nat Methods *16*, 167–170. https://doi.org/10.1038/s41592-018-0301-y.

46. d’Hennezel, E., Abubucker, S., Murphy, L.O., and Cullen, T.W. (2017). Total Lipopolysaccharide from the Human Gut Microbiome Silences Toll-Like Receptor Signaling. mSystems *2*, 10.1128/msystems.00046-17. https://doi.org/10.1128/msystems.00046-17.

47. Mordaka, P.M., and Heap, J.T. (2018). Stringency of Synthetic Promoter Sequences in Clostridium Revealed and Circumvented by Tuning Promoter Library Mutation Rates. ACS Publications. https://doi.org/10.1021/acssynbio.7b00398.

48. George, J.T., Acree, C., Park, J.-U., Kong, M., Wiegand, T., Pignot, Y.L., Kellogg, E.H., Greene, E.C., and Sternberg, S.H. (2023). Mechanism of target site selection by type V-K CRISPR-associated transposases. Science *382*, eadj8543. https://doi.org/10.1126/science.adj8543.

49. Reis, A.C., and Salis, H.M. (2020). An Automated Model Test System for Systematic Development and Improvement of Gene Expression Models. ACS Synthetic Biology. https://doi.org/10.1021/acssynbio.0c00394.

50. Cheng, A.G., Ho, P.-Y., Aranda-Díaz, A., Jain, S., Yu, F.B., Meng, X., Wang, M., Iakiviak, M., Nagashima, K., Zhao, A., et al. (2022). Design, construction, and in vivo augmentation of a complex gut microbiome. Cell *185*, 3617-3636.e19. https://doi.org/10.1016/j.cell.2022.08.003.

51. Yap, Z.L., Rahman, A.S.M.Z., Hogan, A.M., Levin, D.B., and Cardona, S.T. (2024). A CRISPR-Cas-associated transposon system for genome editing in Burkholderia cepacia complex species. Applied and Environmental Microbiology *0*, e00699-24. https://doi.org/10.1128/aem.00699-24.

52. Nethery, M.A., Hidalgo-Cantabrana, C., Roberts, A., and Barrangou, R. (2022). CRISPR-based engineering of phages for in situ bacterial base editing. Proceedings of the National Academy of Sciences *119*, e2206744119. https://doi.org/10.1073/pnas.2206744119.

53. Brödel, A.K., Charpenay, L.H., Galtier, M., Fuche, F.J., Terrasse, R., Poquet, C., Havránek, J., Pignotti, S., Krawczyk, A., Arraou, M., et al. (2024). In situ targeted base editing of bacteria in the mouse gut. Nature, 1–8. https://doi.org/10.1038/s41586-024-07681-w.

54. Wexler, H.M. (2007). Bacteroides: the Good, the Bad, and the Nitty-Gritty. Clin Microbiol Rev *20*, 593–621. https://doi.org/10.1128/CMR.00008-07.

55. Shepherd, E.S., DeLoache, W.C., Pruss, K.M., Whitaker, W.R., and Sonnenburg, J.L. (2018). An exclusive metabolic niche enables strain engraftment in the gut microbiota. Nature *557*, 434–438. https://doi.org/10.1038/s41586-018-0092-4.

56. Whitaker, W.R., Russ, Z.N., Stanley Shepherd, E., Popov, L.M., Louie, A., Lam, K., Zong, D.M., Gill, C.C.C., Gehrig, J.L., Rishi, H.S., et al. (2025). Controlled colonization of the human gut with a genetically engineered microbial therapeutic. Science *389*, 303–308. https://doi.org/10.1126/science.adu8000.

57. Park, J.-U., Tsai, A.W.-L., Mehrotra, E., Petassi, M.T., Hsieh, S.-C., Ke, A., Peters, J.E., and Kellogg, E.H. (2021). Structural basis for target site selection in RNA-guided DNA transposition systems. Science *373*, 768–774. https://doi.org/10.1126/science.abi8976.

58. Rottinghaus, A.G., Vo, S., and Moon, T.S. (2023). Computational design of CRISPR guide RNAs to enable strain-specific control of microbial consortia. Proceedings of the National Academy of Sciences *120*, e2213154120. https://doi.org/10.1073/pnas.2213154120.

59. Querques, I., Schmitz, M., Oberli, S., Chanez, C., and Jinek, M. (2021). Target site selection and remodelling by type V CRISPR-transposon systems. Nature *599*, 497–502. https://doi.org/10.1038/s41586-021-04030-z.

60. Guiney, D.G., Hasegawa, P., and Davis, C.E. (1984). Plasmid transfer from Escherichia coli to Bacteroides fragilis: differential expression of antibiotic resistance phenotypes. Proc. Natl. Acad. Sci. U.S.A. *81*, 7203–7206. https://doi.org/10.1073/pnas.81.22.7203.

61. Tamminen, M., Virta, M., Fani, R., and Fondi, M. (2012). Large-Scale Analysis of Plasmid Relationships through Gene-Sharing Networks. Molecular Biology and Evolution *29*, 1225–1240. https://doi.org/10.1093/molbev/msr292.

62. Neil, K., Allard, N., Roy, P., Grenier, F., Menendez, A., Burrus, V., and Rodrigue, S. (2021). High-efficiency delivery of CRISPR-Cas9 by engineered probiotics enables precise microbiome editing. Molecular Systems Biology *17*. https://doi.org/10.15252/msb.202110335.

63. Neil, K., Allard, N., Grenier, F., Burrus, V., and Rodrigue, S. (2020). Highly efficient gene transfer in the mouse gut microbiota is enabled by the Incl2 conjugative plasmid TP114. Commun Biol *3*, 523. https://doi.org/10.1038/s42003-020-01253-0.

64. Du, J., Meile, S., Baggenstos, J., Jäggi, T., Piffaretti, P., Hunold, L., Matter, C.I., Leitner, L., Kessler, T.M., Loessner, M.J., et al. (2023). Enhancing bacteriophage therapeutics through in situ production and release of heterologous antimicrobial effectors. Nat Commun *14*, 4337. https://doi.org/10.1038/s41467-023-39612-0.

65. Łobocka, M., Dąbrowska, K., and Górski, A. (2021). Engineered Bacteriophage Therapeutics: Rationale, Challenges and Future. BioDrugs *35*, 255–280. https://doi.org/10.1007/s40259-021-00480-z.

66. Hsu, B.B., Plant, I.N., Lyon, L., Anastassacos, F.M., Way, J.C., and Silver, P.A. (2020). In situ reprogramming of gut bacteria by oral delivery. Nat Commun *11*, 5030. https://doi.org/10.1038/s41467-020-18614-2.

67. Shen, J., Zhang, J., Mo, L., Li, Y., Li, Y., Li, C., Kuang, X., Tao, Z., Qu, Z., Wu, L., et al. (2023). Large-scale phage cultivation for commensal human gut bacteria. Cell Host & Microbe *31*, 665-677.e7. https://doi.org/10.1016/j.chom.2023.03.013.

68. Camarillo-Guerrero, L.F., Almeida, A., Rangel-Pineros, G., Finn, R.D., and Lawley, T.D. (2021). Massive expansion of human gut bacteriophage diversity. Cell *184*, 1098-1109.e9. https://doi.org/10.1016/j.cell.2021.01.029.

69. Ronda, C., and Wang, H.H. (2022). Engineering temporal dynamics in microbial communities. Current Opinion in Microbiology *65*, 47–55. https://doi.org/10.1016/j.mib.2021.10.009.

70. Riglar, D.T., Giessen, T.W., Baym, M., Kerns, S.J., Niederhuber, M.J., Bronson, R.T., Kotula, J.W., Gerber, G.K., Way, J.C., and Silver, P.A.

(2017). Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. Nat Biotechnol *35*, 653–658. https://doi.org/10.1038/nbt.3879.

71. Zou, Z.-P., Du, Y., Fang, T.-T., Zhou, Y., and Ye, B.-C. (2023). Biomarker-responsive engineered probiotic diagnoses, records, and ameliorates inflammatory bowel disease in mice. Cell Host & Microbe *31*, 199-212.e5. https://doi.org/10.1016/j.chom.2022.12.004.

72. Schmidt, F., Zimmermann, J., Tanna, T., Farouni, R., Conway, T., Macpherson, A.J., and Platt, R.J. (2022). Noninvasive assessment of gut function using transcriptional recording sentinel cells. Science *376*, eabm6038. https://doi.org/10.1126/science.abm6038.

73. Hayashi, N., Lai, Y., Fuerte-Stone, J., Mimee, M., and Lu, T.K. (2024). Cas9-assisted biological containment of a genetically engineered human commensal bacterium and genetic elements. Nat Commun *15*, 2096. https://doi.org/10.1038/s41467-024-45893-w.

74. Puigbò, P., Guzmán, E., Romeu, A., and Garcia-Vallvé, S. (2007). OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. Nucleic Acids Res *35*, W126–W131. https://doi.org/10.1093/nar/gkm219.

75. Bernhards, C.B., Liem, A.T., Berk, K.L., Roth, P.A., Gibbons, H.S., and Lux, M.W. (2022). Putative Phenotypically Neutral Genomic Insertion Points in Prokaryotes. ACS Synth. Biol. *11*, 1681–1685. https://doi.org/10.1021/acssynbio.1c00531.

76. Bae, S., Park, J., and Kim, J.-S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics *30*, 1473–1475. https://doi.org/10.1093/bioinformatics/btu048.

77. Wagih, O. (2017). ggseqlogo: a versatile R package for drawing sequence logos. Bioinformatics *33*, 3645–3647. https://doi.org/10.1093/bioinformatics/btx469.

78. Inkscape Project (2020). Inkscape. Version 0.92.5.

79. Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proceedings of the National Academy of Sciences *97*, 6640–6645. https://doi.org/10.1073/pnas.120163297.

80. Wood, D.E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. Genome Biology *20*, 257. https://doi.org/10.1186/s13059-019-1891-0.

81. Lu, J., Breitwieser, F.P., Thielen, P., and Salzberg, S.L. (2017). Bracken: estimating species abundance in metagenomics data. PeerJ Comput.

Sci. *3*, e104. https://doi.org/10.7717/peerj-cs.104.

82. Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

83. Shuangbin Xu, Wang, Q., Wen, S., Li, J., He, N., Li, M., Hackl, T., Wang, R., Zeng, D., Wang, S., et al. (2025). aplot: Simplifying the creation of complex graphs to visualize associations across diverse data types. The Innovation *6*, 100958. https://doi.org/10.1016/j.xinn.2025.100958.

84. Grolemund, G., and Wickham, H. (2011). Dates and times made easy with lubridate. Journal of Statistical Software *40*, 1–25.

85. Wickham, H. (2007). Reshaping data with the reshape package. Journal of Statistical Software *21*, 1–20.

86. Constantin, A.-E., and Patil, I. (2021). ggsignif: R package for displaying significance brackets for "ggplot2." PsyArxiv. https://doi.org/10.31234/osf.io/7awm6.

87. Brunson, J.C. (2020). ggalluvial: Layered Grammar for Alluvial Plots. J Open Source Softw *5*, 2017. https://doi.org/10.21105/joss.02017.

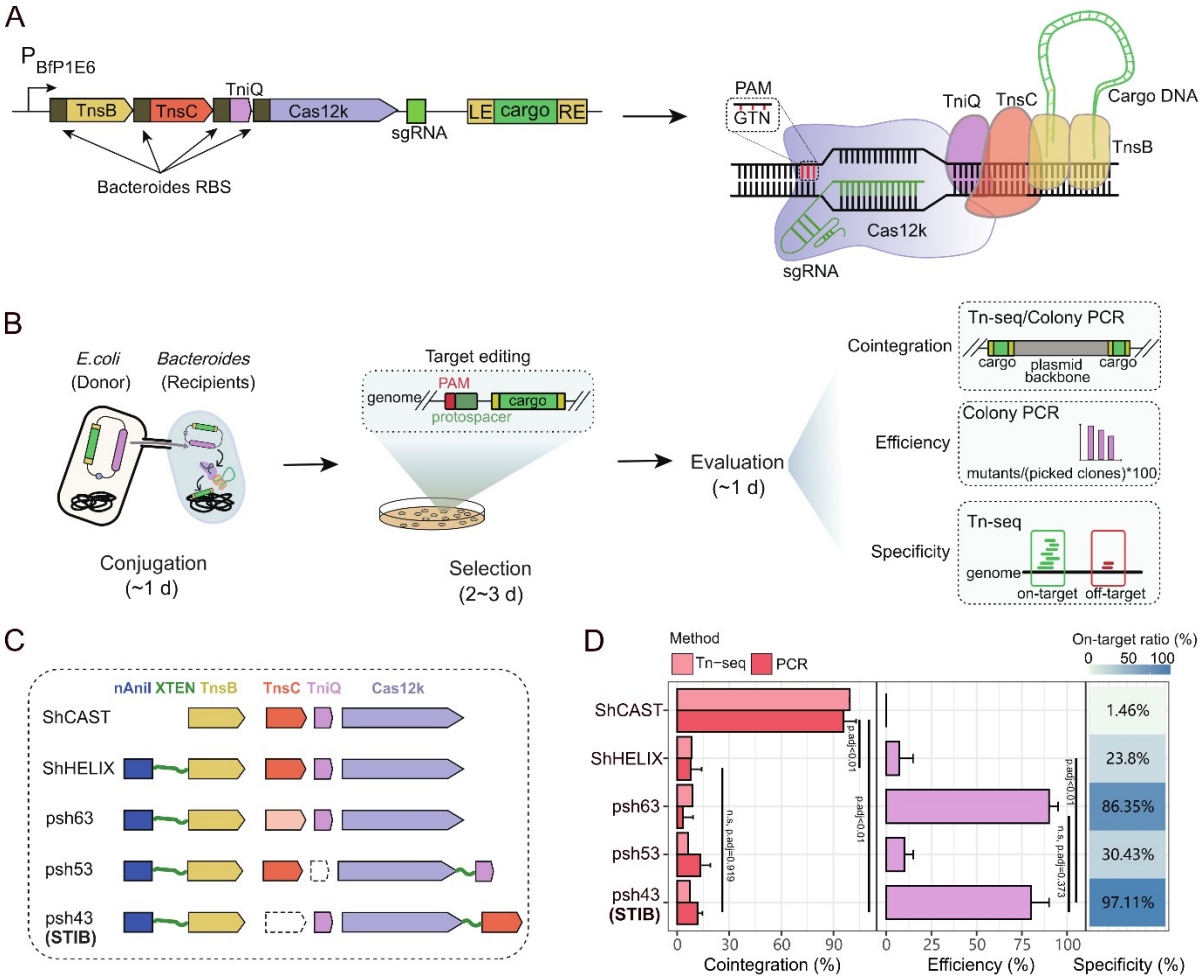


Figure 1. Optimization of ShCAST-based genome editing tools for human gut *Bacteroides* species. (A) Schematic of polycristronic and codon optimized ShCAST system for *Bacteroides.* **(B)** Schematic of genomic transposition experiments workflow. **(C)** Schematic of optimized ShCAST-based genome editing tools for *Bacteroides.* **(D)** Systematic evaluation of different optimized ShCAST system at cointegration, efficiency and specificity in *B.thetaiotaomicron* at *Bt01* site. Error bars represent mean + sd (n = 3 biological replicates). Statistical significance between psh43(STIB) and other systems was assessed ANOVA test and Tukey HSD post-hoc test. The on-target ratio was evaluated by Tn-seq (see Methods, n=3 biological replicates).

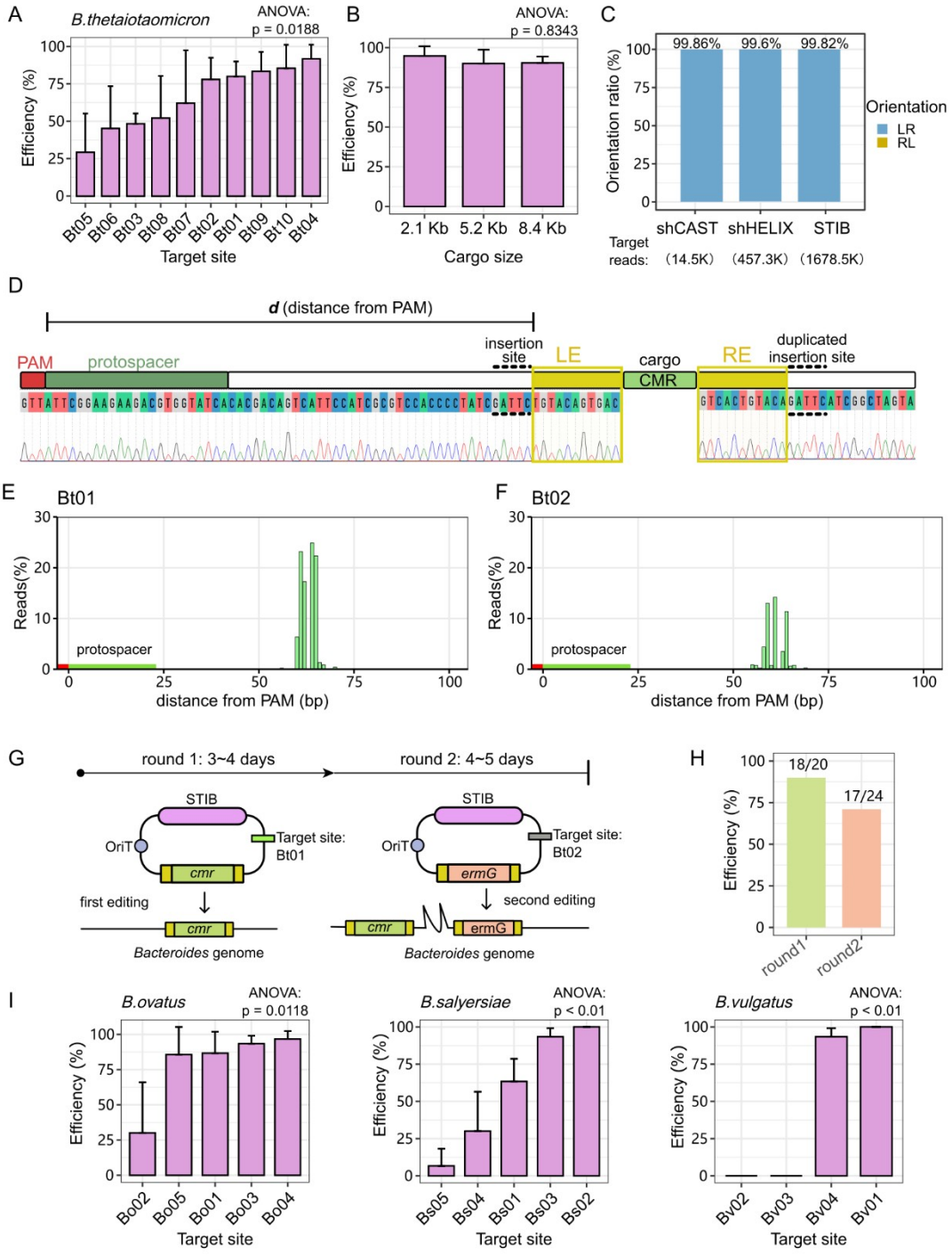


Figure 2. Systematic characterization of ShCAST-based Transient Insertion system for *Bacteroides* species (STIB). (A) Evaluation of editing efficiency at different sites in *B. thetaiotaomicron*, via PCR (data is shown as mean + sd., for n = 3, biological replicates; ANOVA test: p = 0.0188). **(B)** Editing efficiency at *Bt01* site with different cargo sizes assessed by ddPCR. Data are shown as mean + sd. for n = 3 biological replicates. ANOVA test: p = 0.8343. **(C)** Orientation ratio of targeted insertion using ShCAST (14.5 thousand reads), ShHELIX (457.3 thousand reads) and STIB (1,678.5 thousand reads). **(D)** Sanger

chromatogram shows the features of product after STIB targeting *Bt01.* The PAM location is marked as red box. The protospacer location is marked as dark green box. The insertion sites and duplicated insertion sites are denoted by dotted line. Left end (LE) and right end (RE) are framed in yellow box. Cargo is marked as light green box. **(E)** Insertion positions at *Bt01*, identified by Tn-seq. **(F)** Insertion positions at *Bt02*, identified by Tn-seq. **(G)** Schematic of rapid sequential editing of STIB-mediated gene manipulation. **(H)** Quantification of sequential editing efficiency based on PCR. **(I)** Efficiency at different genomic sites in *B.ovatus* ATCC8483*, B. salyersiae* DA1247 (isolate) and *B.vulgatus* ATCC8482. ANOVA test shows $p < 0.05$.

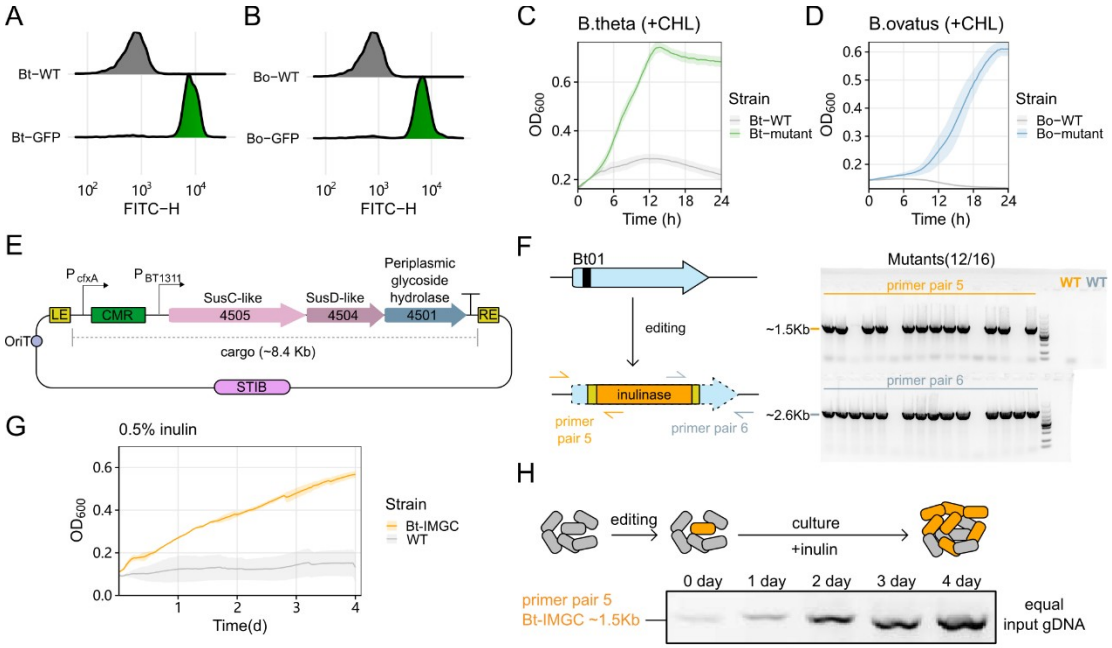


Figure 3. Targeted genomic insertions in *Bacteroides* species enables the enrichment of mutant strains. (A) The expression intensities of exogenous *gfp* gene in *B.thetaiotaomicron*. **(B)** The expression intensities of *gfp* gene in *B.ovatus*. **(C)** The growth curve of wild type (Bt-WT, gray) and *Bt02* mutant (Bt-mutant, green) in BHI plus 25 μg/mL chloramphenicol. **(D)** Growth of *B.ovatus* wild type (Bo-WT, gray) and *B.ovatus* mutant (Bo-mutant, blue) in BHI plus 25 μg/mL chloramphenicol. **(E)** Schematic of inulin utilization cluster from *B.ovatus* ATCC8483. The three inulin utilization genes were controlled by strong constitutive promoter P_{BT1311}. This three-gene cluster was loaded at the rear of *cmr* cassette. **(F)** Colony PCR verification of inulin cluster insertion following STIB editing at *Bt01* site. Primers are indicated by orange and pewter arrows. WT: wild type. **(G)** Growth curve of wild type *B.thetaiotaomicron* (WT, gray) and edits with inulin utilization cluster (*B. thetaiotaomicron* inulin metabolism gene cluster (Bt-IMGC), orange) in M9S-inulin medium. **(H)** Edited bacterial cells with inulin utilization capacity were enriched in population. PCR verification of edits at different time points, using genome-inulinase specific primer pair 5, described as **(F).** The line represents the mean of data, and the shadow represents the s.d. Mean and s.d shown for n = 3 biological replicates in **(C)**, **(D)**, **(G)**.

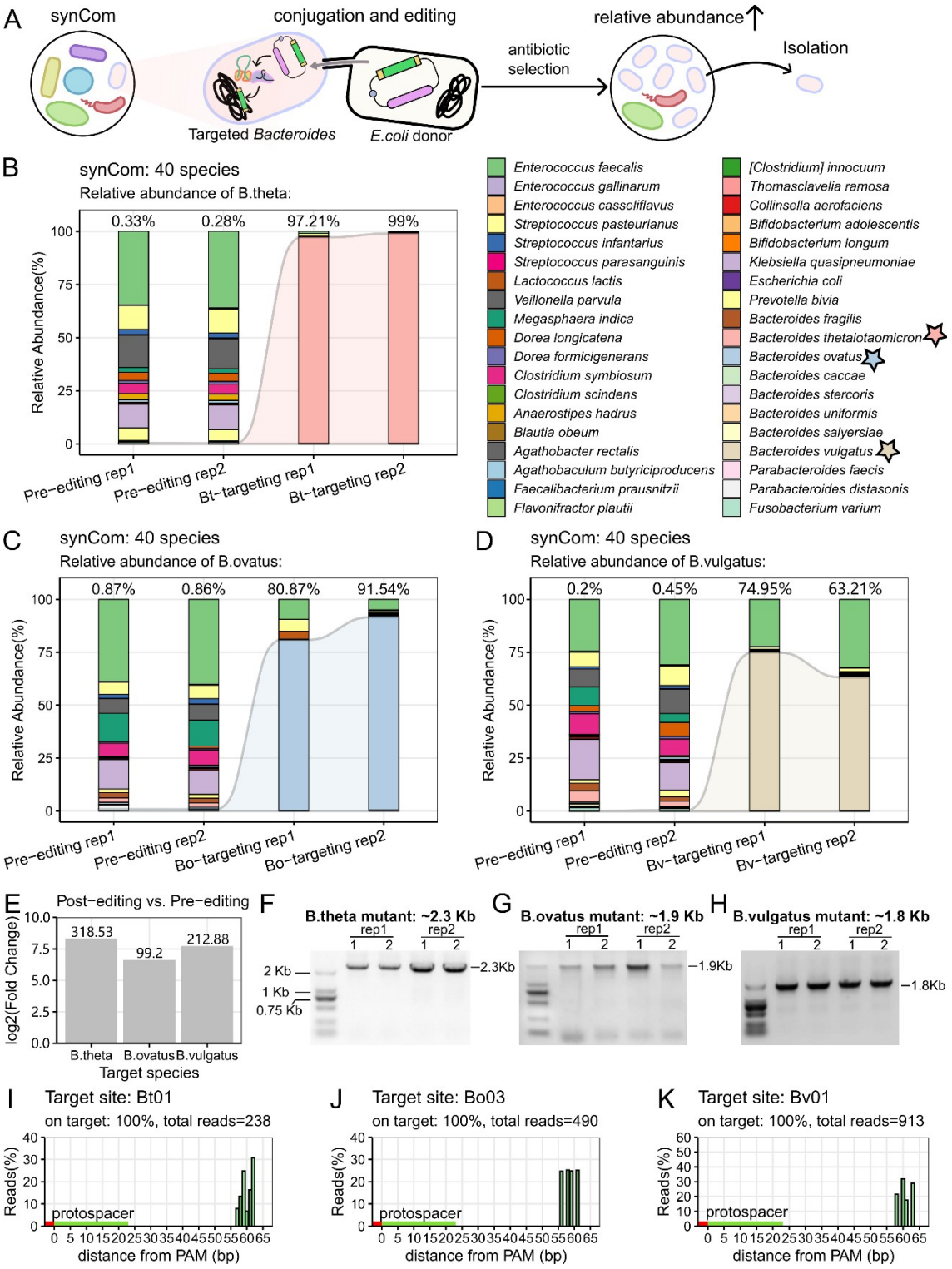


Figure 4. Species- and site- specific editing of *Bacteroides* within a complex gut microbial community. (A) STIB mediates targeted editing in microbiota, followed by antibiotic selection for propagation and isolation of edits. **(B), (C) and (D)** Changes in the community composition of the 40-species synthetic human gut bacterial community before and after STIB editing of target species. The chloramphenicol concentrations used (*B. thetaiotaomicron*: 50

µg/mL, *B. ovatus*: 25 µg/mL, *B. vulgatus*: 25 µg/mL.). The percentage value above each bar denotes the relative abundance of targeted species. n = 2 biological replicates. Targeted species are marked by stars (*B. thetaiotaomicron*: soft red, *B. ovatus*: powder blue, *B. vulgatus*: light beige.). **(E)** Fold change of relative abundance (Post-editing / Pre-editing) for *B. thetaiotaomicron* (318.5×), *B. ovatus* (99.2×), and *B. vulgatus* (212.8×). Values represent mean fold change from two biological replicates. **(F), (G), (H)** PCR verification of target editing in isolates. **(I), (J), (K)** Position of all metagenome-wide insertions located at expected target site (100 bp downstream of PAM). Red short line represents PAM location. Green line represents protospacer.