國立臺灣大學理學院化學系

碩士論文

Department of Chemistry

College of Science

National Taiwan University

Master's Thesis

針對雙歧桿菌屬開發基因編輯技術以調控代謝途徑：破壞限制修飾系統促進基因編輯效率

Development of CRISPR-Cas9 Base Editing Tools for Metabolically Engineerable *Bifidobacteria*: Disruption of Restriction-Modification Systems Facilitates Gene Editing Efficiency

林泓君

Hung-Chun Lin

指導教授：徐丞志 博士

Advisor: Cheng-Chih Hsu, Ph.D.

中華民國 113 年 6 月

June 2024

doi:10.6342/NTU202404344

國立臺灣大學碩士學位論文

口試委員會審定書

NTU Master Thesis

Oral Defense Approval Form

針對雙歧桿菌屬開發基因編輯技術以調控代謝途徑：

破壞限制修飾系統促進基因編輯效率

Development of CRISPR/Cas9 Base Editing Tools for Metabolically Engineerable *Bifidobacteria*: Disruption of Restriction-Modification Systems Facilitates Gene Editing Efficiency

本論文係 林泓君 君（學號 R11223205 ）在國立臺灣大學化學系完成之碩士學位論文，於民國 113 年 6 月 17 日承下列考試委員審查通過及口試及格，特此證明。

口試委員 (Committee Members)：

_____ （簽名 Signature）

（指導教授 Advisor）

_____ _____

_____ _____

系主任、所長 _____ （簽章）

(Dept./Institute Chair's Signature and Seal)

doi:10.6342/NTU202404344

誌謝

很謝謝在我念碩士班的這段期間給予我指導、建議以及任何幫助的大家，沒有你們的幫忙就不會有現在的研究成果，我也無法順利畢業。首先謝謝我的指導教授徐丞志老師，在入學前半年讓我先進實驗室學習實驗技術以及適應環境，謝謝老師願意放手讓我自己思考題目的方向、建立獨立思考還有培養解決問題的能力，也謝謝老師二話不說就買 PCR 機器，讓這個題目能進行的更順利。我也很欣賞老師的報告能力和演講的台風，總是能抓住大家的目光，最後祝老師事業有成、研究順利。

我要特別謝謝張明姿老師，總是不間斷的與我們討論實驗結果與後續的研究方向，Manuscript 也一字一句的跟我來回修改，真的學到了很多，與你們合作很愉快，希望我的 Broken English 沒有讓你感到不禮貌；琬錡學姊，謝謝你從題目開始的第一天努力到最後一天，同時進行很多個計劃辛苦了，你總是能用最快的速度提供給我 Plasmids 使用，讓這個計畫得以快速推進；謝謝雅貞學姊和孟潔在研究初期建立了電穿孔以及 mCherry 的實驗方法，為此計畫奠定了基礎；在這裡我也要特別謝謝科技的進步，發展出大型語言模型 ChatGPT，能幫忙修正文法上的錯誤，對非母語的人們來說真的節省了很多英文寫作的時間。

接著是實驗室的大家，容我來一個一個道謝:雅貞學姊，謝謝你帶我做實驗、跟我討論實驗方向、跟著你一起上機學習做 Proteomics，從你身上學到了做實驗的嚴謹；Kevin，謝謝你在我剛進實驗室時就帶我做 LC-QqQ，學到了很多相關的實驗設計的方法和技巧，讓我在後來能盡情的使用這個好用的儀器；興翔，很佩服你每次都能在 Group meeting 上提出很關鍵的問題來幫助大家，祝你找到理想的 Post Doc；柏辰，謝謝你三不五時關心我、願意跟我討論實驗，還有帶我上機。洪鈺，抱歉每次去生物房做實驗都嚇到你；Laura，很羨慕你的語言能力，祝你們投稿跟畢業都順利；怡歆，謝謝你盡心處理環安衛的事情，恭喜你通過資格考。

doi:10.6342/NTU202404344

接著是同屆的昀辰和宇萱，雖然我們做的題目差距非常大，能互相討論的部分不多，但很謝謝你們常常陪我去買午餐，還有回答我各種疑難雜症，像是系上規定或流程等等，祝你們未來都能找到理想的工作或研究實驗室。智翔，你很好聊，也常常幫忙實驗室的大小事；宸宣，從生物分析化學到乖乖再到每次的畢業卡片，都被你的美術天分震懾到；偉凱，除了專精 ML 跟理論計算之外，你是個很懂得生活的人，祝你們未來一年能順利畢業。姿吟和婉靖，祝你們接下來碩士班的路程順利；欣蓓學姊，謝謝你常常來關心我的研究還有人生規劃，在 Amuc_1100 計畫中扮演著溝通的橋樑，祝你仕途順利；謝謝梅梅總是幫忙訂購耗材還有維護生物房的環境，跟行政相關的事情，找你幫忙很讓人放心。

最後謝謝媽媽讓我無後顧之憂地完成學業，也在找工作與人生的道路上給予我建議；謝謝安茹總是聽我講報告、跟我討論新點子，祝你明年順利畢業，也謝謝你們忍受與包容我的壞脾氣。謝謝口試委員們徐丞志老師、張明姿老師、朱忠瀚老師以及林建達老師給予我指導，提出了我之前沒想過的一些觀點。最後也謝謝自己選擇來台大化學所，讓自己的未來多了些選擇。沒有大家給予我的幫助與支持，就不會有現在的我，謝謝，並再次祝福所有人未來一帆風順。

doi:10.6342/NTU202404344

摘要

腸道微生物群中的雙歧桿菌屬(*Bifidobacterium*)被視為益生菌，與人類健康息息相關，不僅能夠改善腸道健康，還能增強免疫力。為了深入探討雙歧桿菌是如何影響人類宿主、釐清是否有特定分子在此扮演重要角色，針對雙歧桿菌特定基因和代謝路徑的調控至關重要。然而，雙歧桿菌內限制修飾(RM)系統會將外來未甲基化的 DNA 切除，造成轉形(transformation)效率大幅降低，導致難以在雙歧桿菌屬中開發基因工程技術。本篇論文中，我們成功在雙歧桿菌屬中建立了一系列 CRISPR-Cas9 基因鹼基編輯器(cytosine base editor systems, cBEST)，藉由精準調控鹼基編輯器和 sgRNA 的表現量來編輯雙歧桿菌屬中特定的基因。為了解決轉形效率差的問題，我們在龍根菌(*Bfidobacterium longum* NCIMB 8809)以及青春雙歧桿菌(*Bifidobacterium adolescentis* DSM 20083)中利用 cBEST 技術消除限制酶或甲基化特定 DNA 序列來躲避限制修飾系統，我們成功將轉型效率提高至三十萬倍，而基因編輯成功率則提高至 100%。接著，為了驗證 cBEST 方法能精準調控特定的代謝路徑，我們透過分析不同變異株的甲硫胺酸(methionine)循環路徑，了解 DNA 甲基化與甲硫腺苷(5'-methylthioadenosine)調控的關係。我們也成功在不同雙歧桿菌屬中編輯膽鹽水解酶(*bsh*)，並透過 LC-MS/MS 定量分析，證明我們能精準調控雙歧桿菌屬的膽酸代謝路徑。整體而言，我們成功利用 CRISPR-Cas9 基因編輯技術在雙歧桿菌屬中精準調控代謝途徑，並提供策略去躲避限制修飾系統，我們相信此方法必能成為研究腸道微生物強而有力的工具。

關鍵字: 基因編輯、CRISPR-Cas9、雙歧桿菌、限制修飾系統、DNA 甲基化、甲硫腺苷、膽鹽水解酶

doi:10.6342/NTU202404344

Abstract

Intestinal microbiota members of the *Bifidobacterium* genus are increasingly recognized as probiotic potential and therapeutics applications to improve human health. However, the paucity of genetic tools and the prevalence of restriction modification (RM) systems limit the genetic manipulation in *Bifidobacterium*. In this study, we established a series of CRISPR/Cas9 cytosine base editor systems (cBESTs) with fine tuning base editor and sgRNA expression that are portable across multiple *Bifidobacteria*. We showed that bypassing RM systems by either eliminating restriction endonucleases or matching DNA methylation patterns in *B. longum* NCIMB 8809 and *B. adolescentis* DSM 20083 significantly improved both transformation (up to 300,000-fold) and editing efficiencies (up to 100%). Furthermore, we successfully manipulated methionine cycle pathway to elucidate the correlation between DNA methylation and 5'-methylthioadenosine regulation through *MetK*-deficient and RM-disrupted strains. Finally, we achieved the genetic knockout of the conserved bile salt hydrolase (*bsh*) gene in *B. longum* NCIMB 8809, *B. longum* DSM 20219 and *B. infantis* DSM 20088, effectively deactivating bile acids deconjugation. In summary, the ability to efficiently engineer *Bifidobacterium* genomes will definitely open new avenues for research and applications towards improving human health.

doi:10.6342/NTU202404344

Keywords: Genome editing, CRISPR-Cas9, *Bifidobacterium*, restriction-modification systems, DNA methylation, 5’-methylthioadenosine, bile salt hydrolase

doi:10.6342/NTU202404344

目次

doi:10.6342/NTU202404344

doi:10.6342/NTU202404344

圖次

doi:10.6342/NTU202404344

doi:10.6342/NTU202404344

doi:10.6342/NTU202404344

表次

doi:10.6342/NTU202404344

This thesis is a collaboration project with Professor Mingzi M. Zhang from Institute of Molecular and Genomic Medicine, National Health Research Institutes. The manuscript, entitled Highly Efficient DSB-Free CRISPR-Cas9 Base Editing in Bifidobacterium with Bypass of Restriction Modification Systems, is currently being under review at ACS Synthetic Biology, and the thesis author is the primary author of the manuscript.

Chapter 1. Introduction

1-1. *Bifidobacterium*

Advances in multi-omics analyses and bioinformatics have shed light on the crucial roles of human gut microbes in shaping human physiology and diseases.[1, 2] Maintaining the composition of gut microbiota, such as probiotics, is important for improving human health. A particular genus known as *Bifidobacterium* stands out as probiotics with numerous beneficial effects on human health, including the regulation of immune response, maintenance of gut barrier function, metabolic homeostasis, and shaping of gut microbiota composition at different life stages.[3-6]

Through analyzing the metabolomic profiles of gut microbes by mass spectrometry (MS), we are starting to appreciate the role of gut microbial metabolism in relation to human health and diseases and discover novel bioactive compounds with therapeutic potential.[2, 7] Our previous work indicated that *Bifidobacterium longum* NCIMB 8809 can protect against metabolic disorders through its metabolite 5'-methylthioadenosine (MTA) derived from methionine,[8] which was identified by LC-MS/MS analysis.

doi:10.6342/NTU202404344

Understanding the intricate cross-talk across *Bifidobacteria*, the host and other members of the gut microbiota require precise genetic modifications of individual genes, metabolic pathways and microbial strains. Conducting research with these engineered strains to validate hypotheses generated from clinical data could unlock the potential of *Bifidobacteria* as probiotics and therapeutics in both food and pharmaceutical industries.[5, 9] However, the limited accessibility to genetic tools in *Bifidobacteria* poses a huge challenge in conducting such mechanistic studies.[10, 11]

1-2. CRISPR-Based Genome Editing Methods for Bacteria

The adaptation of the clustered regularly interspaced short palindromic repeat (CRISPR) and CRISPR-associated (Cas) proteins stands out as revolutionary milestone in biotechnology by accomplishing precise genome editing with remarkable accessibility across viruses, prokaryotic and eukaryotic organisms.[12, 13] It requires only a single guide RNA (sgRNA) to target specific gene locus and a programmable Cas protein, allowing deletions, insertions and precise point mutations.[14] A well-established method employing cytidine deaminase (APOBEC-1) fused with Cas9 (D10A) nickase from *Streptococcus pyogenes* enables cytosine (C) to thymine (T) mutagenesis without inducing double-

doi:10.6342/NTU202404344

strand DNA breaks (DSBs) (**Figure 1.1**).[15, 16] Augmenting this fusion protein with uracil DNA glycosylase inhibitor (UGI) has shown the enhancement of C-to-T editing efficiency.[17]

While CRISPR-based genome editing has been successfully developed and applied to specific genus of gut microbiota such as *Clostridium*, *Bacillus*, *Lactobacilli* and *Bacteroides*,[18-21] only few genomic engineering methods have been reported for *Bifidobacteria*.[10] One reported genome editing method involves in utilizing the characterized endogenous CRISPR-Cas system in *Bifidobacterium animalis subsp. lactis* DSM 10140.[22] Although the applicability of endogenous CRISPR-Cas system is more suitable for bacteria, some *Bifidobacteria* genome do not contain CRISPR-Cas system, and its diversity may lead to unpredictable editing outcomes. An alternative solution is the delivery of the well-established CRISPR-Cas9 cytosine base editing systems into *Bifidobacterium* for precise genetic manipulation without DSBs. Despite the potential advantages, challenges arising from thick peptidoglycan layers and active restriction-modification (RM) systems result in extremely low transformation efficiency, which creates obstacles in accessing synthetic biology tools for *Bifidobacteria*.[23, 24]

doi:10.6342/NTU202404344

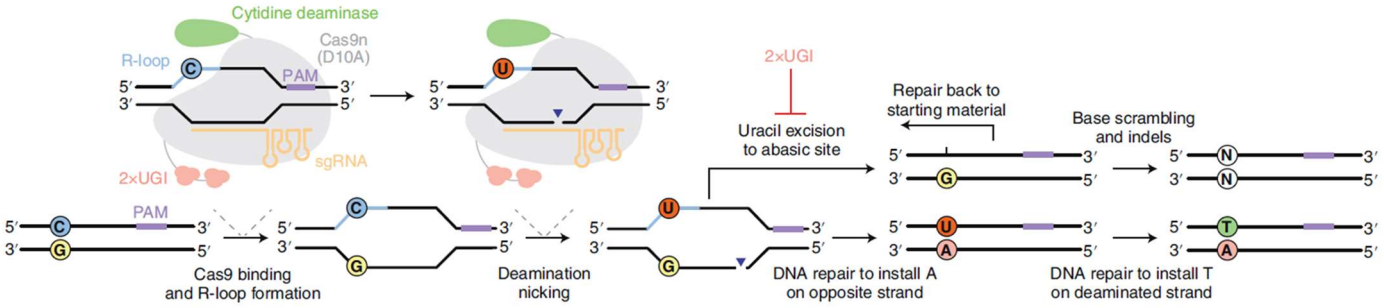


Figure 1.1. Molecular mechanism of cytosine base editor introducing C-to-T mutagenesis. Figure reprinted from Anzalone *et al* (2020).[14]

1-3. Restriction Modification Systems

Restriction-modification (RM) systems differentiate foreign DNA from endogenous genomic DNA through two main enzymatic components: modification methyltransferase (MTase) and restriction endonuclease (REase). The MTase protects DNA by adding a methyl group to cytosine or adenine through S-adenosyl-L-methionine (SAM) (**Figure 1.2A**), whereas the REase cleaves the specific sequences lack of methylation. RM systems are categorized based on the protein composition, recognition sequences and other factors. Type I RM systems are complex, comprising three distinct proteins—restriction, modification and specificity subunits—with relatively long and bipartite recognition sequences. Type II RM systems are the most prevalent and well-studied, constituting about 90% of all RM systems. They are simple and characterize by two separate proteins

doi:10.6342/NTU202404344

responsible for restriction and modification, with typically shorter and palindromic recognition sequences. Type III RM systems comprise multiple restriction and modification subunits, with short and non-palindromic recognition site in the MTase. (**Figure 1.2B**).[25, 26] Single molecule real time (SMRT) sequencing has revealed a diverse methylome across *Bifidobacterium breve* strains, highlighting variations in RM systems.[27] Taken together, the highly variable REases cleave the foreign non-methylated DNA, reducing the transformation efficiency and making it challenging to develop genetic tools in *Bifidobacteria*. Hence, we would need some strategies to bypass the native RM systems.

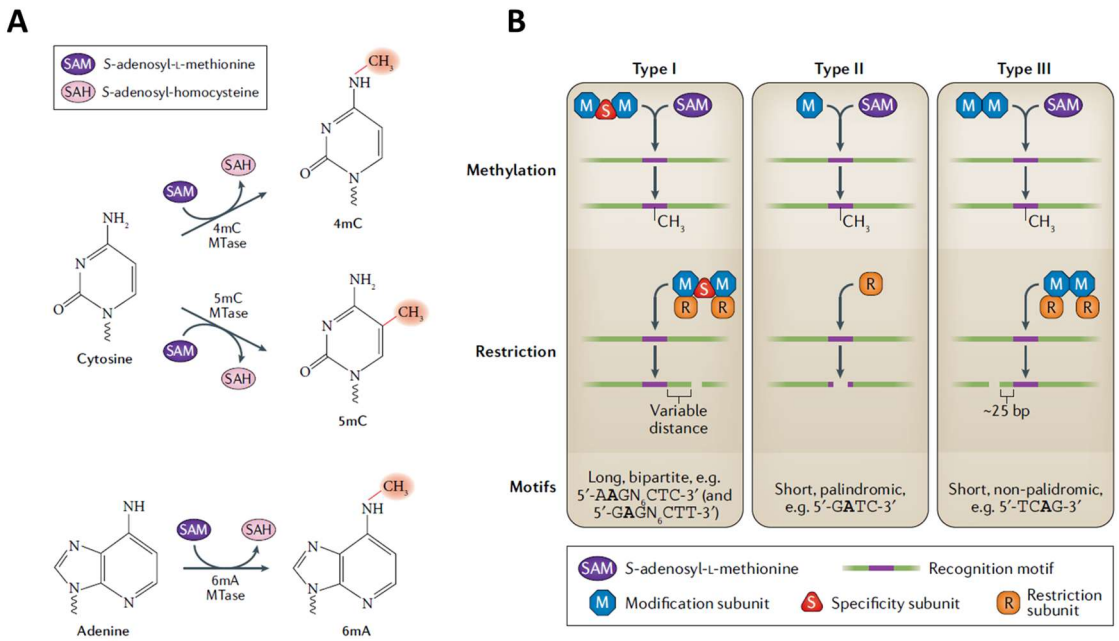


Figure 1.2. DNA methylation and types of RM systems in bacteria. Figure reprinted from Beaulaurier *et al* (2019).[26]

doi:10.6342/NTU202404344

1-4. Research Objectives

CRISPR-based genome editing enables precise manipulation of biosynthetic pathways through activation, interference, knockout or knock-in strategies. If we can overcome the issue of low transformation efficiency caused by RM systems, we will have greater confidence in developing CRISPR-based gene editing tools in *Bifidobacteria*. Furthermore, the synergistic use of LC-MS/MS and CRISPR-based technologies will undoubtedly advance natural product discovery and metabolic regulation, leading to the development of new therapeutics and a deeper understanding of microbial biochemistry as well as gut microbes-host interaction.

In this thesis, we address the challenge of gene editing in *Bifidobacteria* by establishing CRISPR-Cas9 cytosine base editor systems (cBEST). For strains with low transformation efficiencies, we sought to bypass their RM systems through disrupting REases or matching DNA methylation patterns. Furthermore, we showcased the utility of our cBEST by using the identical construct to knockout a conserved bile salt hydrolase (*bsh*) gene and achieve bile acids perturbation in different *Bifidobacterium* strains. Our CRISPR-Cas9 base editing tools provide a novel solution for tuning metabolic pathways in *Bifidobacteria*.

doi:10.6342/NTU202404344

Chapter 2. Materials and Methods

2-1. Chemicals and Reagents

All chemicals and reagents used in this study are listed in **Supplementary Table 6.2**.

2-2. Bacterial Strains and Culturing

Bacterial strains used in this study are listed in **Supplementary Table 6.3**. *Bifidobacterium* strains and *Clostridium scindens* DSM 5676 were propagated in BCRC and cultured in NEOGEN Lactobacilli MRS broth containing 0.05% wt/vol L-cysteine (MRSC) at 37 ºC under anaerobic conditions (80% N_2, 10% CO_2, 10% H_2) using Whitley DG250 anaerobic workstation. Bacterial growth curves were carried out in 20 mL MRSC broth inoculated with 400 μL overnight seed cultures over 24 hours. Aliquots were removed at indicated time points for OD_{600} measurements using a Metertech Model6^+.

2-3. Protospacer Design for Gene Knockout

The protospacer design followed the previous work.[16] Briefly, we identified Gln codons (CAA and CAG) on the forward strand or Trp (AGG) codons on the reverse strand

doi:10.6342/NTU202404344

within the theoretical editing window, situated 11 to 17 nucleotides upstream from PAM (NGG). Next, a BLAST search of 12nt+NGG was conducted to check a single perfect match in whole genome to minimize potential off-target effects. A minimum of 2-3 protospacers were selected per targeted gene to ensure successful gene editing. List of protospacers used in this study are listed in **Supplementary Table 6.4**.

2-4. Electroporation of *Bifidobacterium* Strains

Culture media was first optimized for *B. longum* NCIMB 8809 to improve transformation efficiency (**Figure 2.1**). Cysteine and sucrose are essential for *Bifidobacterium* cultures and sodium chloride can be used as cell wall weakening agent to improve transformation efficiency,[28] but the bacterial growth may slightly slow down.

Overnight seed cultures in MRSC broth were diluted 30-fold in MRSC broth supplemented with 0.2 M sucrose and 0.2 M NaCl and incubated anaerobically until OD_{600} reached 0.3~0.4. Other *Bifidobacterium* species were cultured in MRSC broth supplemented with 0.2 M sucrose. After bacterial culturing, bacteria were chilled on ice for 15 min, washed with cold 10% glycerol twice and resuspended in 1/75 of the original culture volume of 10% glycerol. 0.1 mL bacteria were mixed with plasmids (1.5 μg

doi:10.6342/NTU202404344

cBEST plasmids, 500 ng non-methylated pMGC-mCherry and pMGC-Cas9n plasmids, or 250 ng host-methylated plasmids) and chilled on ice for 15 min. For direct transformation of Golden Gate assembly reaction, 5 μL of the Golden Gate assembly mixture was added to 0.1 mL of bacteria. Upon transfer to a pre-chilled 0.1 cm-cuvette, cells were electroporated at 1.5 kV/cm, 5 ms using MicroPulser Electroporator (Bio-Rad) and recovered immediately with 0.9 mL MRSC broth for 2.5 hours at 37 ºC anaerobically. The bacteria were plated on MRSC agar containing 10 μg/mL of chloramphenicol and incubated at 37 ºC anaerobically for 2-3 days.

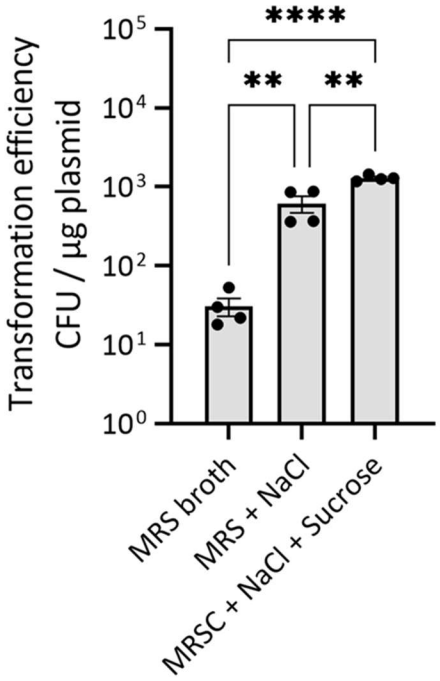
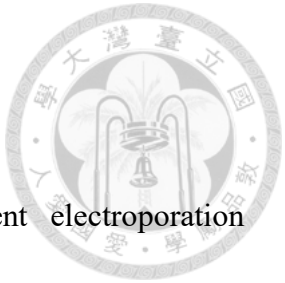


Figure 2.1. Medium optimization for competent cell preparation via pMGC-mCherry. (n=4, four biological replicates; **P < 0.01, ****P < 0.0001)

doi:10.6342/NTU202404344



2-5. Determination of Transformation Efficiencies

Transformation efficiencies were determined by independent electroporation experiments with consistent plasmid quantities. Colony forming units (CFU) were counted on chloramphenicol resistance plates following appropriate dilution to obtain between 20 and 1000 CFUs per plate. Transformation efficiencies were finally calculated as CFU per μg of plasmids.

2-6. Validation of Genome Editing and Determination of Genome Editing Efficiencies as well as Gene Editing Windows

All gene editing plasmids were transformed into *Bifidobacterium* and plated on 10 μg/mL chloramphenicol selection MRSC agar plates. For editing of wild type strains, all transformants were screened by PCR and Sanger sequencing, with two independent experiments. For editing in REase knockout strains, *B. infantis* DSM 20088 or using methylated editing plasmids, five colonies from each plate were randomly picked for PCR and sequencing due to their high transformation efficiencies (>10^3 CFU/μg plasmid). Editing efficiencies were calculated as [number of edited colonies / number of sequenced colonies]. “0” means no transformant obtained. For editing windows, sequences of

doi:10.6342/NTU202404344

protospacers with position numbering are displayed with PAM at positions 0 to -2. C-to-T conversion rate for each shaded nucleotide was calculated as [number of C-to-T edits / number of sequenced colonies]. Plasmids used in this study are listed in **Supplementary Table 6.5**. Primers used in this study for PCR and sequencing are listed in **Supplementary Table 6.6**.

2-7. Plasmid Curing from *Bifidobacterium* Transformants

To cure cBEST editing plasmids from *Bifidobacterium* transformants, chloramphenicol-resistant colonies were streaked onto non-selective MRSC agar plates. Every two days, single colonies were re-streaked onto new non-selective MRSC plates. Typically, after 5 passages on non-selective MRSC plate, all of colonies were cured of their plasmids and available for functional studies.

2-8. mCherry Reporter Assay for Promoter Screening and Characterization

To screen and characterize promoter strength, we conducted experiments in biological triplicates over three separate days. Each triplicate was averaged and plotted as three value to represent. Chloramphenicol-resistant *Bifidobacterium* transformants

doi:10.6342/NTU202404344

containing plasmids driving mCherry expression from indicated promoters were inoculated into MRSC broth supplemented with 10 μg/mL chloramphenicol. Overnight cultures were diluted 1:50 into 1 mL MRSC broth with 10 μg/mL chloramphenicol and grown anaerobically at 37 ºC for 5 or 24 h to obtain exponential or stationary phase cultures, respectively. Cells were harvested, washed with PBS twice, and resuspended in 400 μL PBS. Both mCherry fluorescence (Ex: 554 nm; Em: 610 nm) and culture absorbance (600 nm) were measured using a Synergy H1 Microplate Reader from BioTek. Relative fluorescence unit (RFU) was defined as the ratio of fluorescence to absorbance.

2-9. Bacterial Plate Cultures for Monitoring Methionine Cycle Pathway

For the analysis of methionine cycle pathway, 100 μL of overnight cultures were plated on MRSC agar plate. After 18 h incubation, bacteria were scraped and extracted with 0.6 mL 80% LC-grade methanol. Samples were homogenized by ultrasonication (25 W, 20 kHz, 1 min) and then centrifuged for 13000 rpm 15 min. For untargeted LC-QE analysis, supernatants were directly transferred to autosampler vials for LC-MS/MS analysis. For targeted LC-QqQ analysis, 50 μL supernatants were mixed with 50 μL internal standards (100 ppb methionine-^{13}C, d3 and 100 ppb 5'-methylthioadenosine-d3)

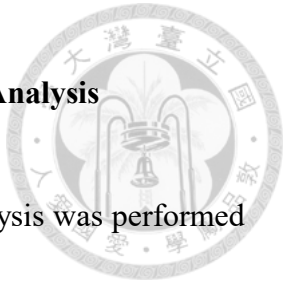
doi:10.6342/NTU202404344

and transferred to autosampler vials for LC-MS/MS analysis.

2-10. Metabolomics Using LC-QE Analysis for Bacterial Plate Cultures

For metabolomics data, LC-MS/MS analysis was performed by a Dionex U3000 UPLC system coupled with a Thermo Scientific Q Exactive Plus equipped with heated electrospray ionization (HESI). 3 μL of sample was injected and separated using Acquity HSS T3 column (2.1×100 mm, 1.7 μm) at 40 ºC. Mobile phase A was 0.1% formic acid in deionized water while mobile phase B was 0.1% formic acid in acetonitrile. The elution separation gradient was as follows: 5% B for 1 min, linear increase to 95% B at 8 min, held at 95% B for 3 min, decreased linearly to 5% B at 12 min and held for another 3 min. Mass spectrometer parameters were also as follows: ionization voltage 3.5 kV (positive mode), capillary temperature 250 ºC and sheath gas flow rate of 25 μL/min. The scan method was operated in Top10 data-dependent acquisition (ddMS2) with a normalized collision energy of 30.

doi:10.6342/NTU202404344



2-11. Targeted Methionine Derived Metabolites Using LC-QqQ Analysis

For targeted five methionine-dericed metabolites, LC-QqQ analysis was performed by an ExionLC AC system coupled with a SCIEX Triple Quad 5500. 3 μL of sample was injected and separated using Acquity UPLC BSH C18 column (2.1×100 mm, 1.7 μm) at 40 ºC. Mobile phase A was 0.1% formic acid in deionized water while mobile phase B was 0.1% formic acid in acetonitrile. The elution gradient for separation was as follows: 5% B for 1 min, increased linearly to 95% B at 5 min, held at 95% B for 3 min, decreased linearly to 5% B at 9 min and held for another 3 min.

The multiple reaction monitoring method for individual compounds is shown in **Figure 2.2**. Peak areas for methionine (MET) and homocysteine (homoCys) were normalized to the peak area of methionine-^{13}C, d3, while peak areas for S-adenosyl-L-methionine (SAM), S-adenosyl-L-homocysteine (SAH) and 5'-methylthioadenosine (MTA) were normalized to the peak area of 5'-methylthioadenosine-d3. We chose bacterial plate cultures for LC-QqQ analysis to minimize noise and improve comparability, as the presence of these metabolites in MRSC broth may introduce interference (**Figure 2.3**).

doi:10.6342/NTU202404344

compound	Q1 (Da)	Q3 (Da)	time (ms)	DP (V)	EP (V)	CE (V)	CXP (V)
MET	150	56	50	50	14	21	7
		104				12	7
		133				13	9
MET-IS (13C,d3)	154	108	50	61	15	16	9
		55				22	7
		137				14	10
MTA	298	136	50	48	6	25	11
		119				70	8
		61				67	9
MTA-IS (d3)	301	136	50	50	6	25	11
		119				67	9
		64				51	8
SAM	399	250	50	70	10	19	9
		298				17	12
		136				33	9
SAH	385	136	50	80	10	23	12
		88				67	5
		250				19	10
homoCys	136	90	50	50	14	17	6
		56				22	8
		118				11	8

Figure 2.2. QqQ multiple reaction monitoring (MRM) method for methionine-derived metabolites. (A) MRM tuning parameters containing parent ion (Q1), daughter ion (Q3), ion dwell time, declustering potential (DP), entrance potential (EP), collision energy (CE) and cell exit potential (CXP).

doi:10.6342/NTU202404344

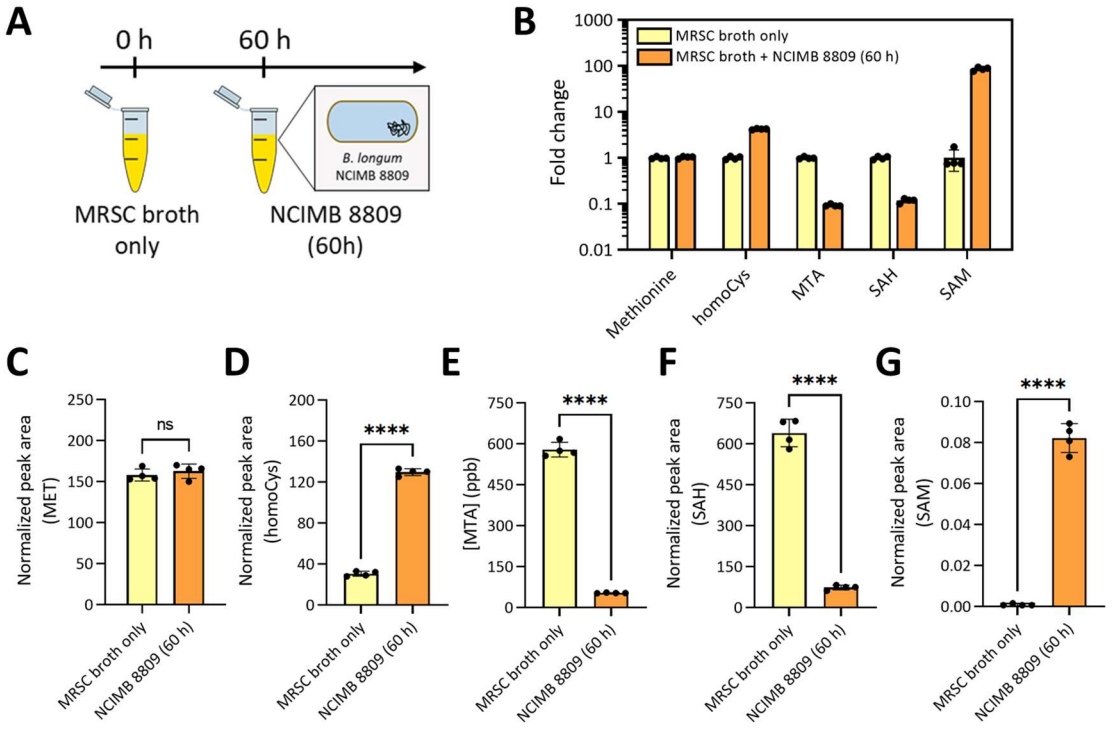


Figure 2.3. Evaluation of interference in MRSC broth. (A) Bacterial broth cultures. *B. longum* NCIMB 8809 was incubated for 60 h. (B) Fold change compared to the MRSC broth only. (C-G) Comparison of normalized peak area for each compound between MRSC broth and bacterial cultures. (n=4; error bars, s.e.m.; ns, not significant, ****P < 0.0001)

doi:10.6342/NTU202404344

2-12. Bacterial Cultures and Quantification Methods for Bile Acids Production

10 μL of overnight *Bifidobacterium* cultures were inoculated into 490 μL MRSC broth supplemented with 100 μM taurocholic acid (TCA). For bacterial co-cultures experiments, overnight three *Bifidobacterium* strains and *Clostridium scindens* DSM 5676 cultures were mixed at the ratio of 1:1:1:3 and inoculated into TCA-supplemented MRSC broth. After 24 h incubation, 600 μL of ice-cold 100% LC-grade methanol spiked-in 2 ppm cholic acid-d4 was mixed with 150 μL cultures thoroughly. Samples were homogenized by ultrasonication (25 W, 20 kHz) for 1 min and then centrifuged for 13000 rpm 10 min at 4 ºC. Supernatants were diluted another five times into 80% LC-grade methanol and transferred to autosampler vials for quantitative LC-QqQ analysis. Standard mixtures with a concentration range from 1 to 200 μM were prepared following the same protocol.

Quantitative LC-QqQ was performed by an ExionLC AC system coupled to a SCIEX Triple Quad 5500. 3 μL of sample was injected and separated using Acquity UPLC BSH C18 column (2.1×100 mm, 1.7 μm) at 40 ºC. Mobile phase A was 1 mM ammonium acetate in deionized water while mobile phase B was 95% acetonitrile with 5% A. The elution gradient for separation was as follows: 30% B for 1 min, increased linearly to 100% B at 8 min, held at 100% B for 3 min, decreased linearly to 30% B at 12

doi:10.6342/NTU202404344

min and held for another 3 min. The multiple reaction monitoring method for individual compounds and the calibration curves are showed in **Figure 2.4**. The function of DCA production by *C. scindens* DSM 5676 was tested in 100 μM cholic acid-supplemented MRSC broth cultures (**Figure 2.5**).

A

compound	RT (min)	Q1 (Da)	Q3 (Da)	time (ms)	DP (V)	EP (V)	CE (V)	CXP (V)
TCA	3.4	533	462	50	62	10	33	18
			516				17	22
			498				19	21
CA	4.18	426	355	50	75	9	29	14
			373				21	15
			91				98	10
CA-d4	4.16	430	359	50	87	9	28	14
			377				21	15
DCA	5.53	410	357	50	68	11	18	14
			375				15	15
			339				23	15

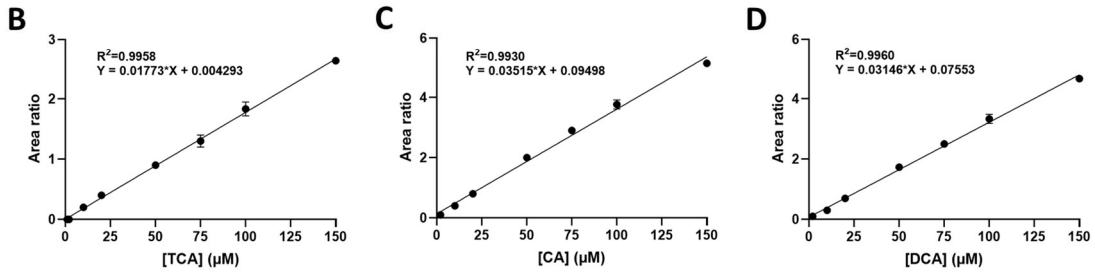


Figure 2.4. QqQ multiple reaction monitoring (MRM) method and the calibration curves. (A) MRM tuning parameters containing Q1, Q3, ion dwell time, DP, EP, CE and CXP. Calibration curves and their equation as well as R^2 value for (B) TCA, (C) CA and (D) DCA. Area ratio were defined by the peak area normalized by peak area of CA-d4 internal standard.

doi:10.6342/NTU202404344

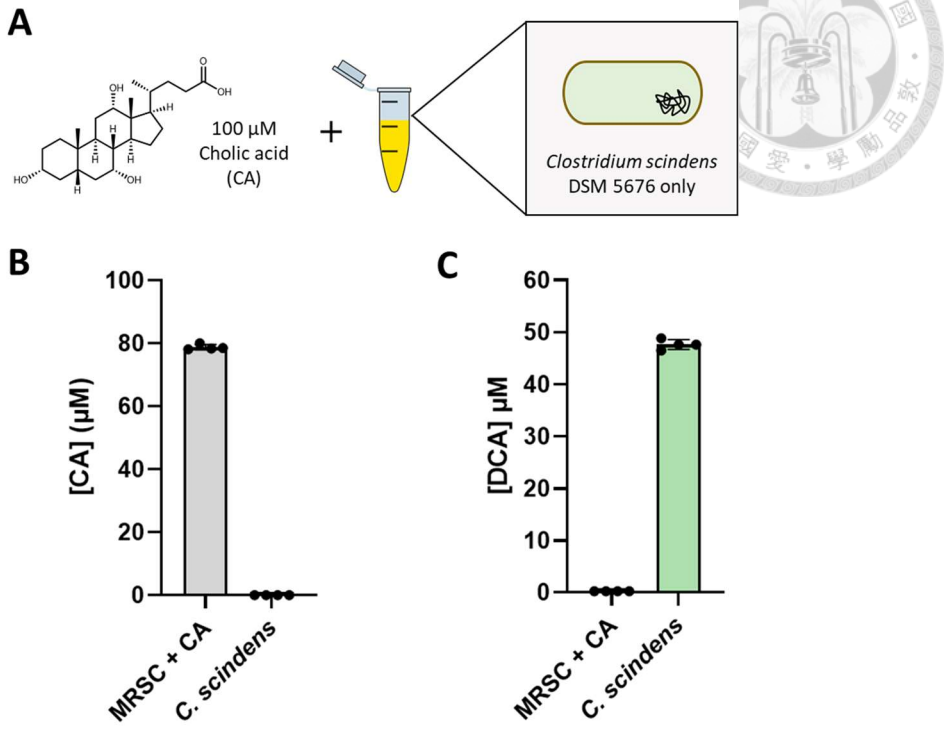


Figure 2.5. Evaluation for DCA production of *C. scindens* DSM 5676. (A) Experimental design for stimulating DCA production: *C. scindens* DSM 5676 was cultured in MRSC broth supplemented with 100 µM CA for 24 h. (B and C) Quantification of CA and DCA metabolites using LC-QqQ analysis (n=4; error bars, standard deviation).

doi:10.6342/NTU202404344

2-13. Statistical Analysis

All statistical analyses were performed in GraphPad Prism 9. For the analysis of mCherry expression levels, unpaired two-tailed Student's T-test was conducted for comparison between exponential and stationary cultures. For the analysis of transformation efficiencies, one-way ANOVA followed by Tukey's post hoc test was employed for multiple comparisons or unpaired two-tailed Student's T-test was used for comparison between two group. For analysis of metabolic data, one-way ANOVA followed by Tukey's post hoc test was employed for multiple comparisons. Asterisks represent significant differences in P-values (*P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001) and ns means no statistical significance.

doi:10.6342/NTU202404344

All plasmids and promoters used in this thesis were provided by Professor Mingzi M. Zhang and her group member, Wan-Chi Hsiao. Their kind collaboration is essential to the success of this study.

Chapter 3. Results

3-1. Promoter Characterization in *Bifidobacteria*

Precise gene expression is crucial for effective microbial bioengineering. However, only few promoters were characterized and used for heterologous gene expression in *Bifidobacteria*.[10, 11] Establishing promoters of varied strengths and minimally-sized synthetic promoters will be crucial to building synthetic tools or genome editing toolkit for *Bifidobacterium* spp..

To identify promoters for the construction of genome editing systems, we first characterized the relative strengths of selected native and synthetic promoters in *B. longum* NCIMB 8809 (**Figure 3.1A**). All constructs were derived from the pMGC-mCherry plasmid (**Figure 3.1B**), which uses the established promoter in *Bifidobacterium bifidum* S17 (P_{gap}) to drive mCherry expression.[29] Constitutive promoters from *Streptomyces griseus* such as P3 and P6[30] yielded detectable mCherry expression in *B. longum* NCIMB 8809, albeit at significantly lower levels compared to P_{gap}. Among the synthetic promoters, the tetracycline-inducible tcp830 promoter (P_{tcp830})[31] demonstrated

doi:10.6342/NTU202404344

the highest expression level. The synthetic *kasO** promoter (P_{kasO*})[32] and its variant with shorter spacer between the -10 and -35 regions ($P_{kasO*17}$)[33] (**Figure 3.1C**) displayed relatively moderate expression levels, similar to P_{gap}. With the exception of P_{gap} and P_{tcp830}, the other promoters yielded consistent mCherry expression during exponential and stationary growth. Taken together, we successfully identified promoters of varied strengths over a 26-fold range in *B. longum* NCIMB 8809, which will be useful for controlling the expression of heterologous genes in *Bifidobacteria*.

To examine the portability of these promoters, we transformed these mCherry reporter plasmids into *B. adolescentis* DSM 20083. While the overall mCherry fluorescence was higher, the general trend of weak and strong promoters was similar to that observed in *B. longum* NCIMB 8809 (**Supplementary Figure 6.1A**). We also observed variable P_{gap} strengths using the identical pMGC-mCherry plasmid across *Bifidobacterium* spp. (**Supplementary Figure 6.1B**).

doi:10.6342/NTU202404344

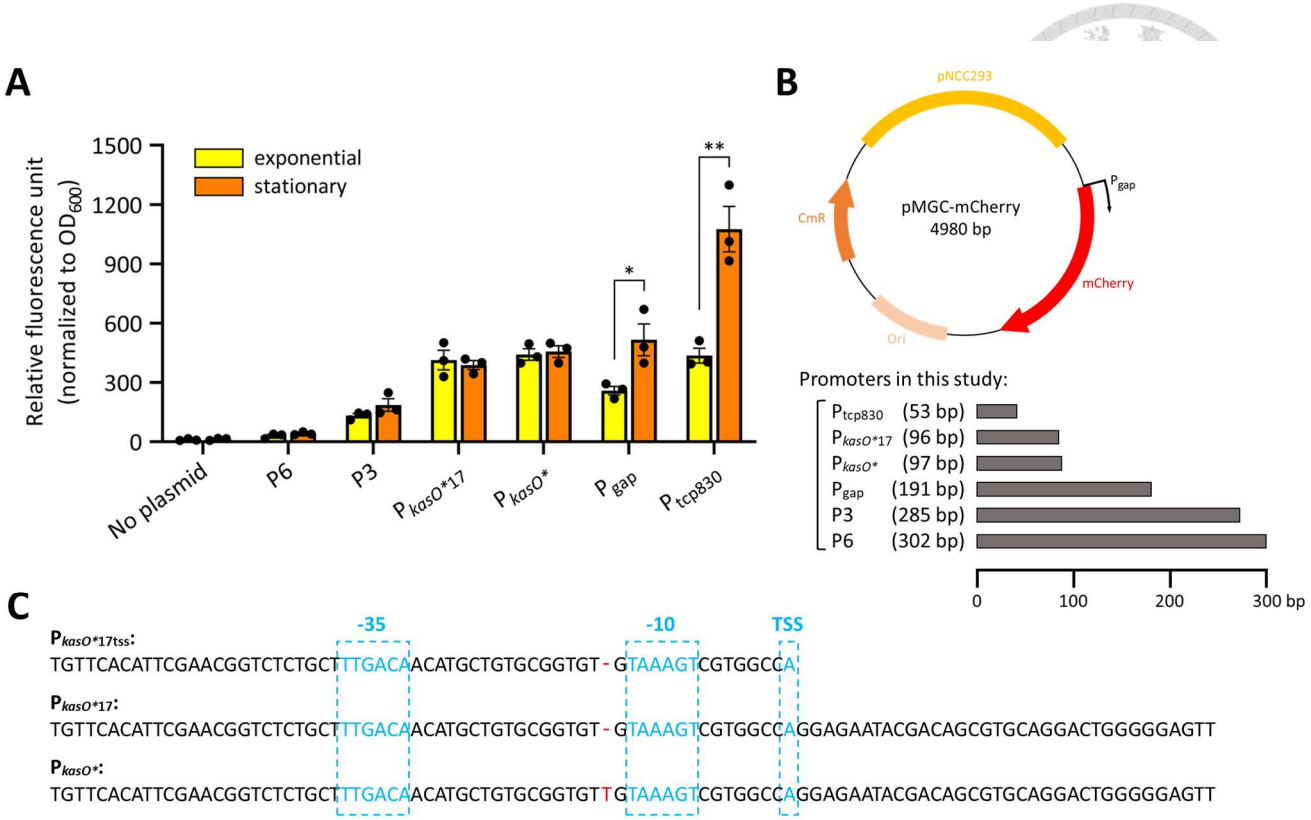


Figure 3.1. Evaluation of promoter strength in *B. longum* NCIMB 8809. (A) Relative expression levels of various promoters in exponential and stationary phase cultures (n=3; error bars, s.e.m.; *P < 0.05, **P < 0.01). (B) Genetic map of pMGC-mCherry plasmid with different promoters used in this study. (C) Sequence alignment of P_{kasO*} and its variants, $P_{kasO*17}$ and $P_{kasO*17tss}$. The blue dashed boxes indicated the -10, -35 regions and transcription start site (TSS). $P_{kasO*17tss}$ is truncated at the TSS site for sgRNA expression.

doi:10.6342/NTU202404344

3-2. Cytosine Base Editor Systems (cBESTs) in *B. longum* NCIMB 8809

Based on the promoters we characterized above, we designed and constructed four sets of base editor constructs to be tested in *B. longum* NCIMB 8809. To precisely locate the genomic locus for base editing, strong and moderate promoters (P_{tcp830} and $P_{kasO*17tss}$) were chosen to drive the sgRNA expression. Considering the influence of Cas9 expression levels on toxicity and genome editing efficiency in other organisms,[34, 35] relatively moderate and weak promoters ($P_{kasO*17}$, P_{kasO*} and P3) were tested to drive the expression of the base editor (spCas9 (D10A) fused with APOBEC1 and UGI (**Figure 3.2A**)). These cytosine Base Editor Systems (cBESTs) are compatible with Golden Gate assembly, which enabled easy cloning of sgRNA sequences in a single step. To edit the *SpeE* gene, a candidate gene involved in the production of 5'-methylthioadenosine (MTA), which has shown potential protection against obesity in our previous study,[8] we designed four different protospacers for each cBEST construct to test, yielding a total of sixteen constructs (**Figure 3.2B**).

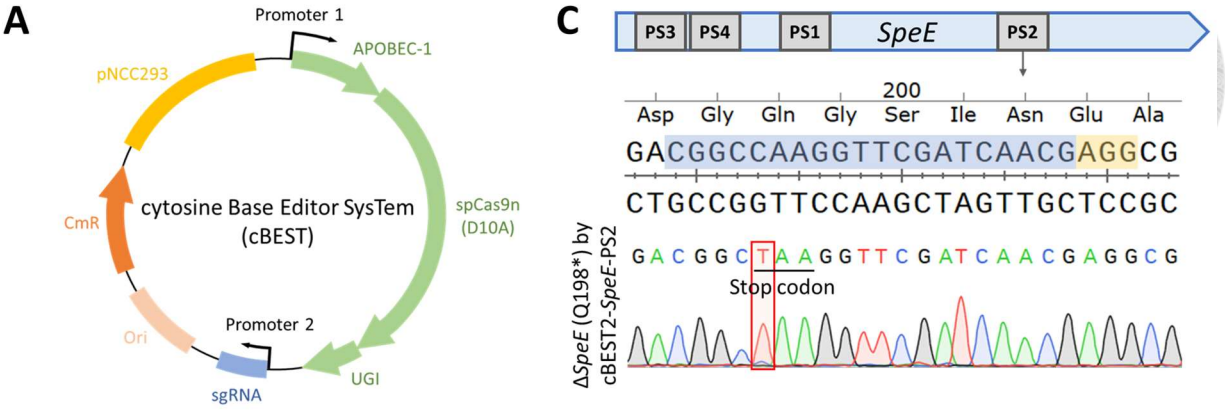
cBEST constructs targeting the *SpeE* gene loci were transformed into *B. longum* NCIMB 8809 and gene editing efficiencies were determined as the proportion of base-edited transformants to the total sequenced transformants in each agar plate (**Figure 3.2B**). Sanger sequencing confirmed the desired C-to-T edits within the predicted editing

doi:10.6342/NTU202404344

window of 11-17 nucleotides distal of the PAM sequence (**Figure 3.2C**, **Supplementary Figure 6.2**). Among the cBEST constructs, cBEST4 exhibited the highest editing efficiencies (up to 100% in each experiment) compared to cBEST2 and cBEST3, while cBEST1 did not make gene mutation. Notably, editing efficiency may also depend on the protospacers used. In the case of *SpeE*-PS2, cBEST2 (25%) outperformed cBEST4 (0%). These highlighted the crucial role of fine-tuning base editor as well as sgRNA expression levels and the usage of promoters to achieve efficient gene editing.

Next, we conducted LC-MS/MS metabolic analysis to evaluate the impact of the Δ*SpeE* (Q198*) strain on MTA production (**Figure 3.3**). Methionine, S-adenosyl-L-methionine (SAM) and MTA were successfully detected, with confident MS/MS fragmentation (**Figure 3.3 E and F**). Surprisingly, no regulation of MTA levels was observed in the Δ*SpeE* (Q198*) strain (**Figure 3.3D**). The intermediate product dcSAM, the substrate for *SpeE* to produce MTA, was not detected. Additionally, we did not find a gene responsible for converting SAM to dcSAM in the publicly available whole genome sequence of *B. longum* NCIMB 8809. This evidence suggests the existence of alternative, unreported enzymes that may be involved in MTA production through different pathways.

doi:10.6342/NTU202404344



B

Host	Target gene	Protospacer	**cBEST1** Promoter 1: $P_{kasO*17}$ Promoter 2: $P_{kasO*17tss}$	**cBEST2** Promoter 1: P_{kasO*} Promoter 2: $P_{kasO*17tss}$	**cBEST3** Promoter 1: P3 Promoter 2: $P_{kasO*17tss}$	**cBEST4** Promoter 1: P3 Promoter 2: P_{tcp830}
B. longum NCIMB 8809	*SpeE*	PS1	0/2, 0/6	2/4, 1/2	0/1, 0	1/1, 0
		PS2	0/1, 0/6	2/5, 1/7	0, 0/1	0/2, 0
		PS3	0, 0/3	0/3, 0	1/4, 0	0, 1/1
		PS4	0, 0/2	0/1, 0	0, 0	4/4, 0

Figure 3.2. Construction and validation of four cBEST constructs for *B. longum* NCIMB 8809. (A) Genetic map of cBEST construct containing both sgRNA and the base editor. (B) Editing efficiencies of cBEST1 to cBEST4 targeting the *SpeE* gene with indicated promoters driving sgRNA and the base editor expression, along with different protospacers (PS) and their corresponding PAM sequences in each row. (C) Sequencing results for precise introduction of a stop codon (black line) due to C-to-T edit (red frame) using the cBEST2-*SpeE*-PS2 plasmid.

doi:10.6342/NTU202404344

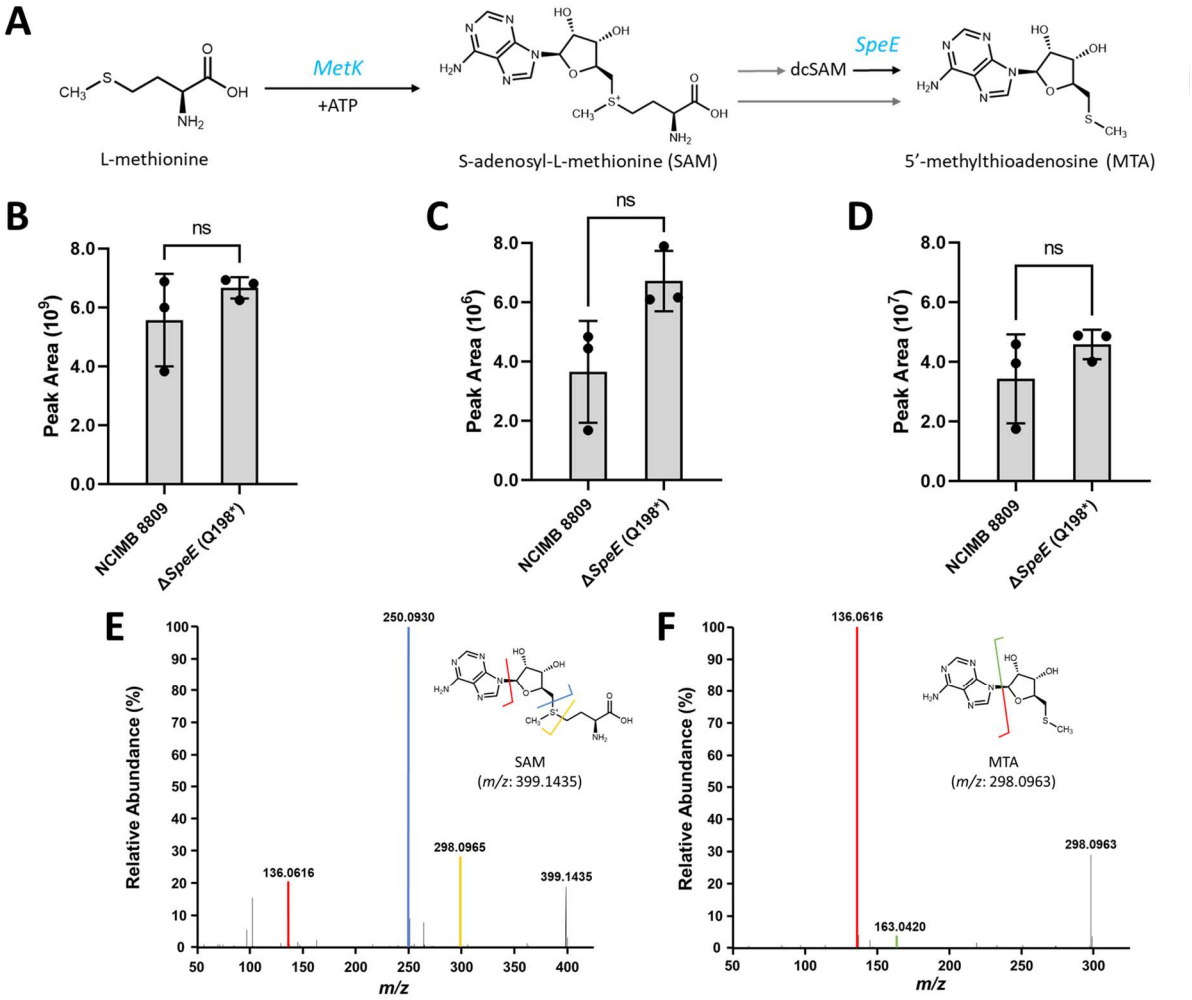


Figure 3.3. Evaluation of MTA regulation in the Δ*SpeE* (Q198*) strain. (A) Biosynthetic pathway of MTA with the indicated genes searched from KEGG database. LC-MS/MS analysis depicting peak areas for (B) L-methionine, (C) S-adenosyl-L-methionine and (D) 5'-methylthioadenosine in both wild-type and the Δ*SpeE* (Q198*) strain. dcSAM was not detected. (n=3, biological triplicates; error bars, standard deviation; ns, not significant) (E and F) Annotation of MS/MS fragmentation patterns for SAM and MTA.

doi:10.6342/NTU202404344

3-3. Enhanced Transformation and Gene Editing Efficiencies in RM-Disrupted Strains Derived from *B. longum* NCIMB 8809 and *B. adolescentis* DSM 20088

Although we successfully established cBEST editing tools for *B. longum* NCIMB 8809, we noticed the extremely low transformation efficiencies (< 10 CFU/μg plasmids) obtained with cBEST plasmids and hypothesized that RM systems may constitute a significant barrier to the DNA delivery into *Bifidobacteria* other than Cas-associated toxicity.

To determine if the contribution of RM systems towards transformation and gene editing efficiencies, we sought to disrupt its native RM systems. Previous studies had assigned the existing RM systems and their predicted recognition motifs in *B. longum* NCIMB 8809.[24] We targeted three predicted REase genes including *HsdR*, *EcoRII_0606* and *EcoRII_0983* in *B. longum* NCIMB 8809 using cBEST2 constructs. Consequently, we obtained five REase knockout strains, comprising three individual genes knockout, a REase double knockout (the Δ*0606*Δ*0983* strain) and a REase triple knockout (**Figure 3.4 A-C**).

We next evaluated whether disrupting REases affects bacterial growth and antibiotics resistance. No significant differences were observed in bacterial growth curves

doi:10.6342/NTU202404344

across all the REase knockout strains (**Figure 3.4D, Supplementary Figure 6.3A**). In the chloramphenicol sensitivity test, there was no inhibition in bacterial growth at 0.5 μg/mL chloramphenicol, while 2.5 μg/mL chloramphenicol completely inhibited bacterial growth, which were all consistent with the wild type (**Figure 3.4E, Supplementary Figure 6.3B**). This indicated the successful loss of editing plasmids in all the REase knockout strains, showing the potential to knockout other interest genes and conduct further functional studies.

Subsequently, we compared transformation efficiencies across all the REase knockout strains using pMGC-mCherry (**Figure 3.5A**) and pMGC-Cas9n plasmids (**Figure 3.5B**). With pMGC-mCherry, we found that, except for the Δ*HsdR* strain, transformation efficiency dramatically increased two to three orders of magnitude (**Figure 3.5C**). The lack of transformation efficiency improvement in the Δ*HsdR* strain might be attributed to the absence of *HsdR* recognition sequence (5'-GATN$_5$TGCC-3') in pMGC-mCherry (**Supplementary Table 6.7**). With pMGC-Cas9n, each knockout strain exhibited varying extent of transformation efficiency improvement. The Δ*0606*Δ*0983*Δ*HsdR* strain showed the most significant enhancement, reaching 300,000-fold compared to the wild type (**Figure 3.5D**). These results revealed the major role of the RM systems in restricting foreign DNA, especially large plasmids such as pMGC-

doi:10.6342/NTU202404344

Cas9n, in *B. longum* NCIMB 8809.

Additionally, we evaluated gene editing efficiencies by employing the cBEST2-*SpeE*-PS1 plasmid across these REase knockout strains. With the exception again of the Δ*HsdR* strain, nearly all REase knockout strains exhibited 100% gene editing efficiency (**Figure 3.6**). Furthermore, we assessed gene editing efficiency in the Δ*0606*Δ*0983*Δ*HsdR* strain using four cBEST-*SpeE*-PS1 plasmids. The use of cBEST2, cBEST3 and cBEST4 constructs all achieved 100% gene editing efficiency (**Figure 3.6**). Consistent editing sites were observed at positions 16 and 17 within the *SpeE*-PS1 protospacer (**Figure 3.6, Supplementary Figure 6.4**). These data highlight the profound impact of restriction endonucleases on both transformation and gene editing processes. The case of the Δ*HsdR* strain might suggest potential decoupling of plasmid transformation and genome editing or indicate that its transformation efficiency is not high enough to maintain plasmids integrity.

To determine if disruption of REases could enhance both transformation and gene editing efficiencies in other *Bifidobacterium* sp., we compared wild type *Bifidobacterium adolescentis* DSM 20083 to a REase-disrupted Δ*Sau3AI* (Q260*) strain (**Supplementary Figure 6.5A**). As expected, we observed significant improvements in transformation efficiency, with 86-fold enhancement using pMGC-mCherry and 565-fold enhancement

doi:10.6342/NTU202404344

using pMGC-Cas9n in the Δ*Sau3AI* (Q260*) strain (**Figure 3.7A**). Additionally, the Δ*Sau3AI* (Q260*) strain exhibited 100% gene editing efficiency in cBEST3 and cBEST4 constructs targeting *Sau3AI*-PS1 (**Figure 3.7B**) and consistent editing sites at position 13 within the *Sau3AI*-PS1 protospacer (**Figure 3.7C, Supplementary Figure 6.5 B-E**), similar to the REase knockout strains observed in *B. longum* NCIMB 8809. Notably, we failed to obtain any edited transformants with cBEST1 and cBEST2, suggesting that base editor expression levels may be too high for *B. adolescentis* DSM 20083 (**Supplementary Figure 6.1A**).

In conclusion, native RM systems constitute a major barrier to transformation and genome editing in *Bifidobacterium*, and that disruption or bypass of RM systems can significantly enhance genome editing of these genetically recalcitrant bacteria.

doi:10.6342/NTU202404344

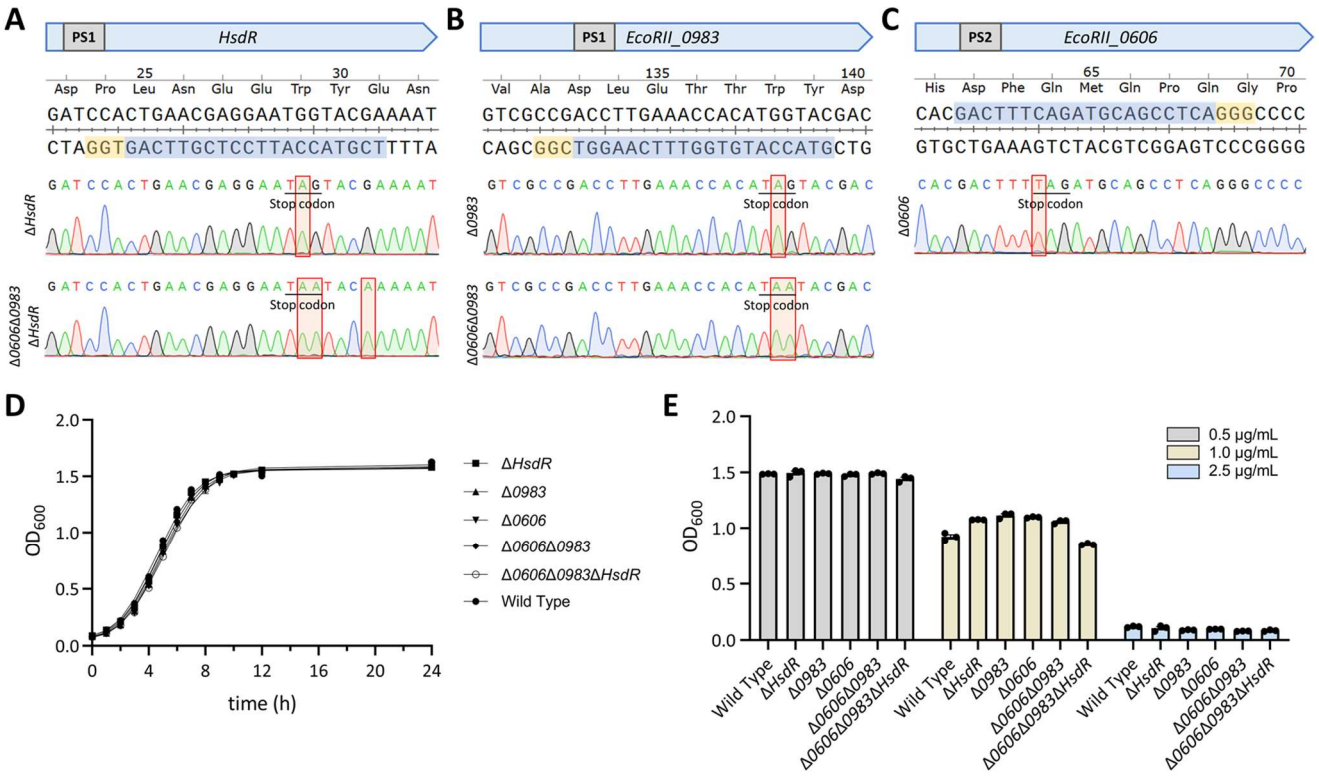


Figure 3.4. Obtaining five REase knockout strains derived from *B. longum* NCIMB 8809 with their chloramphenicol sensitivity test and bacterial growth curves. (A to C) Sequencing results confirmed precise introduction of stop codon (black line) due to successful C-to-T edits (red frame). (D) Bacterial growth in MRSC broth over 24 h. Bacterial log phase cultures were monitored hourly. (n=3, biological triplicates) (E) Antibiotics sensitivity test conducted at the indicated chloramphenicol concentrations for 24 h. (n=3, biological triplicates)

doi:10.6342/NTU202404344

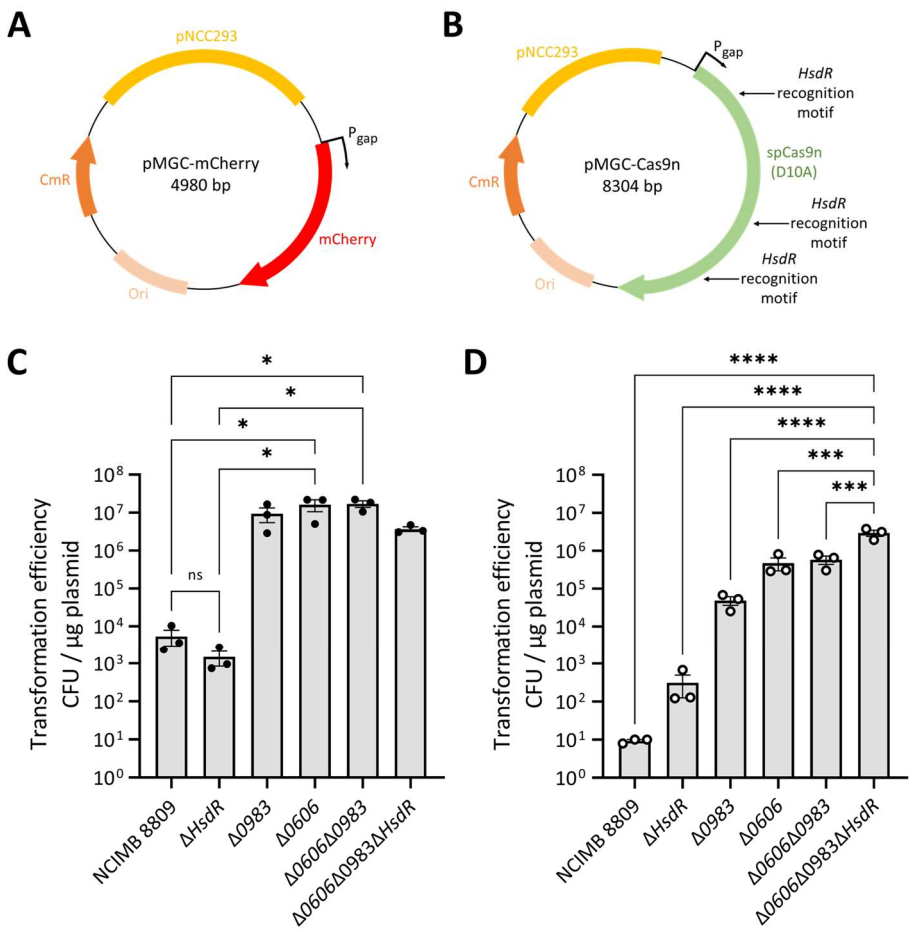


Figure 3.5. Assessment of all REase knockout strains derived from *B. longum* NCIMB 8809 through transformation efficiencies. Genetic map of (A) pMGC-mCherry and (B) pMGC-Cas9n plasmids. Comparison of transformation efficiencies across all REase knockout strains using (C) pMGC-mCherry and (D) pMGC-Cas9n plasmids. (n=3, biological triplicates; error bars, s.e.m.; ns, not significant, *P < 0.05, ***P < 0.001, ****P < 0.0001)

doi:10.6342/NTU202404344

Host	Plasmid	Editing efficiency
Δ0606Δ0983ΔHsdR	cBEST1-*SpeE*-PS1	3/5
NCIMB 8809	cBEST2-*SpeE*-PS1	2/4, 1/2
ΔHsdR		0/5
Δ0983		5/5
Δ0606		4/5
Δ0606Δ0983		5/5
Δ0606Δ0983ΔHsdR		5/5
Δ0606Δ0983ΔHsdR	cBEST3-*SpeE*-PS1	5/5
Δ0606Δ0983ΔHsdR	cBEST4-*SpeE*-PS1	5/5

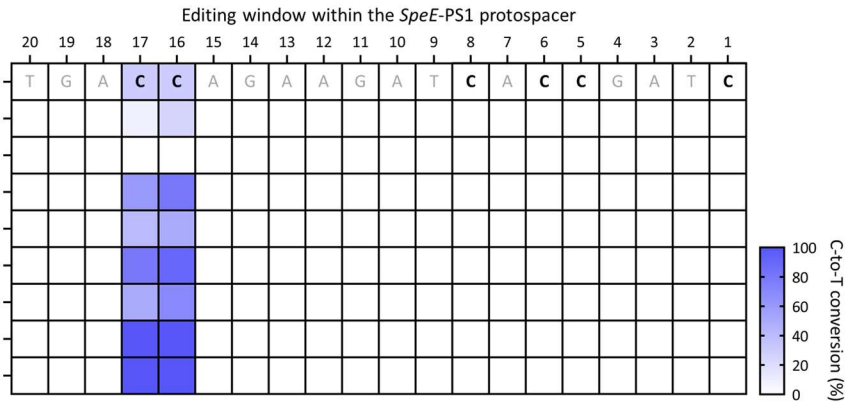
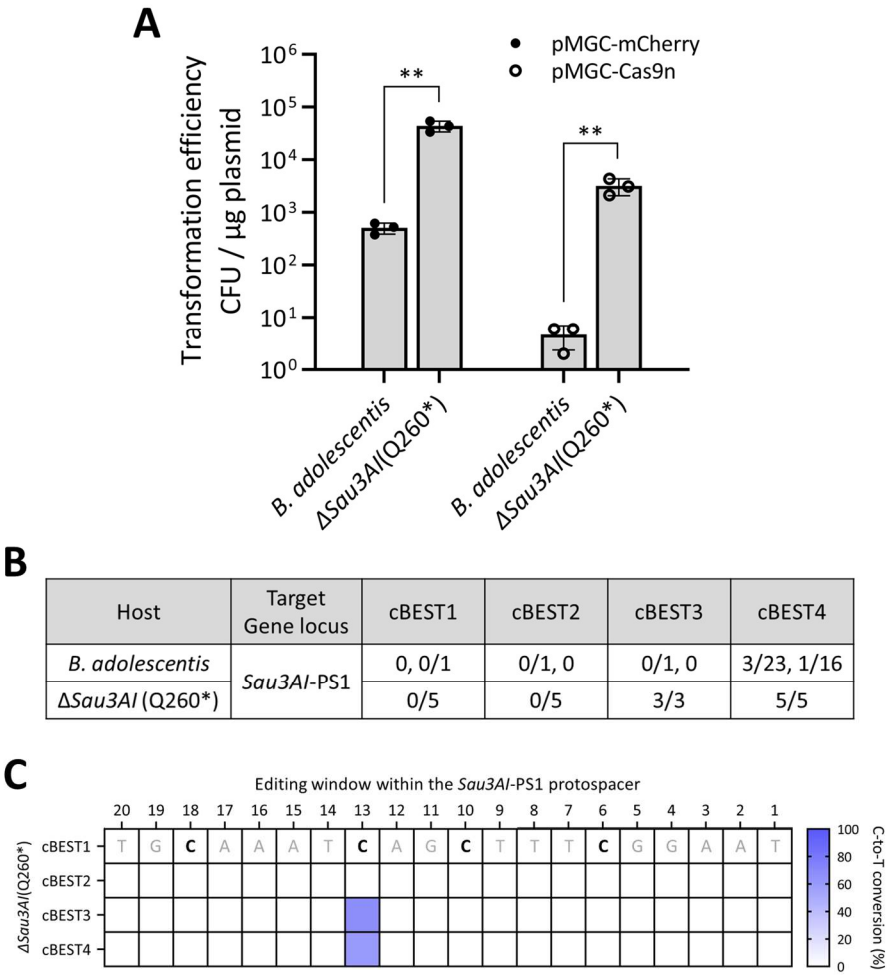


Figure 3.6. Assessment of all REase knockout strains derived from *B. longum* NCIMB 8809 through editing efficiencies and their editing windows. The cBEST2-*SpeE*-PS1 plasmid was transformed into all REase knockout strains to compare editing efficiencies, while four cBEST-*SpeE*-PS1 plasmids were transformed into the *Δ0606Δ0983ΔHsdR* strain.

doi:10.6342/NTU202404344



B

Host	Target Gene locus	cBEST1	cBEST2	cBEST3	cBEST4
B. adolescentis	*Sau3AI*-PS1	0, 0/1	0/1, 0	0/1, 0	3/23, 1/16
Δ*Sau3AI* (Q260*)		0/5	0/5	3/3	5/5

Figure 3.7. Assessment of the Δ*Sau3AI* (Q260*) strain derived from *B. adolescentis* DSM 20083 through transformation and editing efficiencies. (A) Comparison of transformation efficiencies between wild type and the Δ*Sau3AI* (Q260*) strain using both pMGC-mCherry and pMGC-Cas9n plasmids. (n=3, biological triplicates; error bars, s.e.m.; **P < 0.01) (B) Evaluation of editing efficiencies targeting the *Sau3AI*-PS1 locus. (C) Editing window for cBEST constructs targeting the *Sau3AI*-PS1 locus in the Δ*Sau3AI* (Q260*) strain.

doi:10.6342/NTU202404344

3-4. Monitoring Methionine Cycle Pathway in RM-Disrupted Strains

Although we failed to regulate MTA abundance in the Δ*SpeE* strains, untargeted LC-MS/MS analysis revealed significant upregulation of MTA in the Δ*0606* (Q64*) strain (**Supplementary Figure 6.6**). Therefore, RM systems might involve in MTA regulation. DNA methyltransferase (DNMT or MTase) utilizes SAM as a methyl donor to methylate genomic DNA, generating S-adenosyl-L-homocysteine (SAH) as a byproduct (**Figure 3.8 A and B**). Conversely, REases cleave non-methylated DNA. The *MetK* gene is essential for RM systems because it ensures sufficient supply of SAM, which is crucial for DNA methylation. Without functional *MetK*, bacteria may fail to methylate their DNA correctly, leading to self-restriction and cell death. However, in REase-deficient strains, the requirement for DNA methylation diminishes.

We postulated that as REase activity decreases, DNMT activity may also decline, prompting a shift in metabolic pathway toward MTA synthesis as a compensatory mechanism to balance SAM abundance. To verify this hypothesis, we evaluated the REase knockout strains and further targeted the *MetK* gene to engineer the methionine cycle. Using the cBEST4 construct, we attempted to knockout the *MetK* gene across all REase knockout strains. As expected, the Δ*0606*Δ*0983*Δ*HsdR* strain, which has completely inactivated REases, successfully edited the *MetK* gene. Surprisingly, we also

doi:10.6342/NTU202404344

achieved *MetK* knockouts in the Δ*0606*Δ*0983*, Δ*0606* and Δ*0983* strains (**Figure 3.8C**). Sequencing results confirmed the precise introduction of a stop codon in the *MetK* gene (**Figure 3.8D, Supplementary Figure 6.7**).

We subsequently monitored the regulations of methionine cycle pathway in all these engineered strains, including Δ*SpeE, MetK*-deficient and RM-disrupted strains, using targeted LC-QqQ analysis (**Figure 3.9**). In the Δ*SpeE* strain (grey bar), compared to wild type *B. longum* NCIMB 8809, all metabolite levels remained unchanged, confirming again that the *SpeE* gene does not contribute to MTA production (**Figure 3.9, Supplementary Figure 6.8C**). In *MetK*-deficient strains (pink bar), SAM, MTA and SAH were significantly downregulated due to SAM depletion (**Figure 3.9 B-D**). In contrast, all RM-disrupted strains (blue bar) significantly increase MTA abundance, supporting our hypothesis that decreased DNA methylation leads to compensatory MTA production from SAM (**Figure 3.9C**).

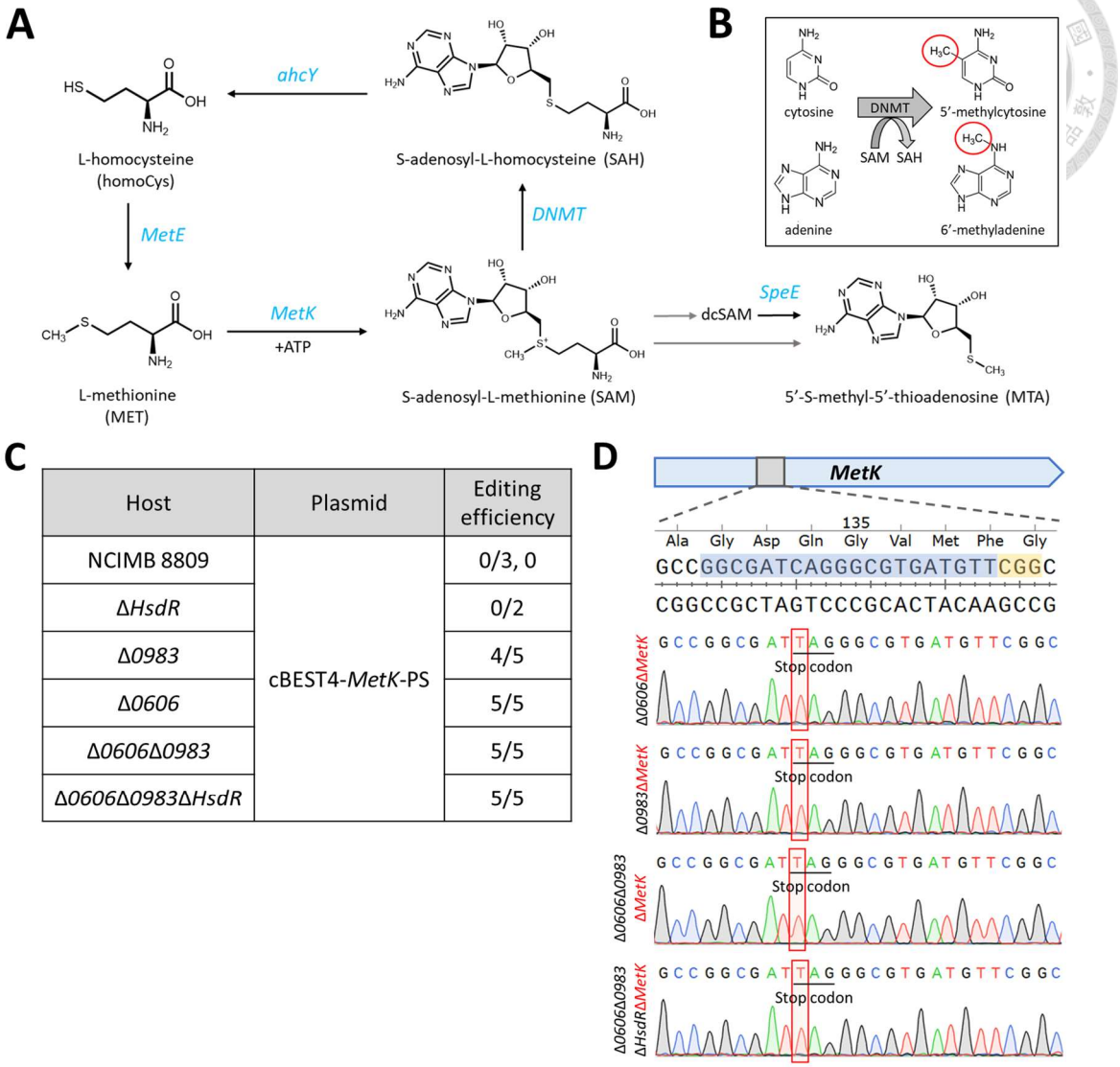
Successful *MetK* knockout in RM-disrupted strains could be attributed to several factors. One possible reason is the activation of alternative pathways or metabolic compensation mechanisms that utilize other methyl donors to maintain DNA methylation, allowing these bacteria to survive without SAM. Additionally, selective pressure may induce genetic changes, causing MTases or residual REases to be turned-off in SAM-

doi:10.6342/NTU202404344

deficient strains to avoid self-restriction. These compensatory mechanisms could allow the bacteria to persist despite the loss of *MetK* gene and SAM metabolite. The inability to edit the *MetK* gene in the Δ*HsdR* strain might be its fully functional Type II RM system, which strongly requires SAM to methylate specific DNA sequences (5'-CCWGG-3') (**Supplementary Table 6.7**). Further investigation into the genetic regulation of these strains could provide deeper insights into their survival strategies under SAM-deficient conditions.

Through engineering multiple strains to precisely manipulate methionine cycle pathway, we elucidated the role of DNA methylation in native RM systems. This study also identified strains with upregulated and downregulated MTA, which may serve as valuable models strains for functional studies related to MTA metabolism and host-microbial interactions.

doi:10.6342/NTU202404344



Host	Plasmid	Editing efficiency
NCIMB 8809	cBEST4-*MetK*-PS	0/3, 0
Δ*HsdR*		0/2
Δ*0983*		4/5
Δ*0606*		5/5
Δ*0606*Δ*0983*		5/5
Δ*0606*Δ*0983*Δ*HsdR*		5/5

Figure 3.8. Successful *MetK* knockouts in RM-disrupted strains. (A) Methionine cycle pathway with the indicated genes searched from KEGG database. (B) DNA methylation through DNMT (DNA methyltransferase, MTase). (C) Comparison of editing efficiencies using the cBEST4-*MetK*-PS plasmid across RM-disrupted strains. (D) Sequencing results of four successful *MetK* knockout strains.

doi:10.6342/NTU202404344

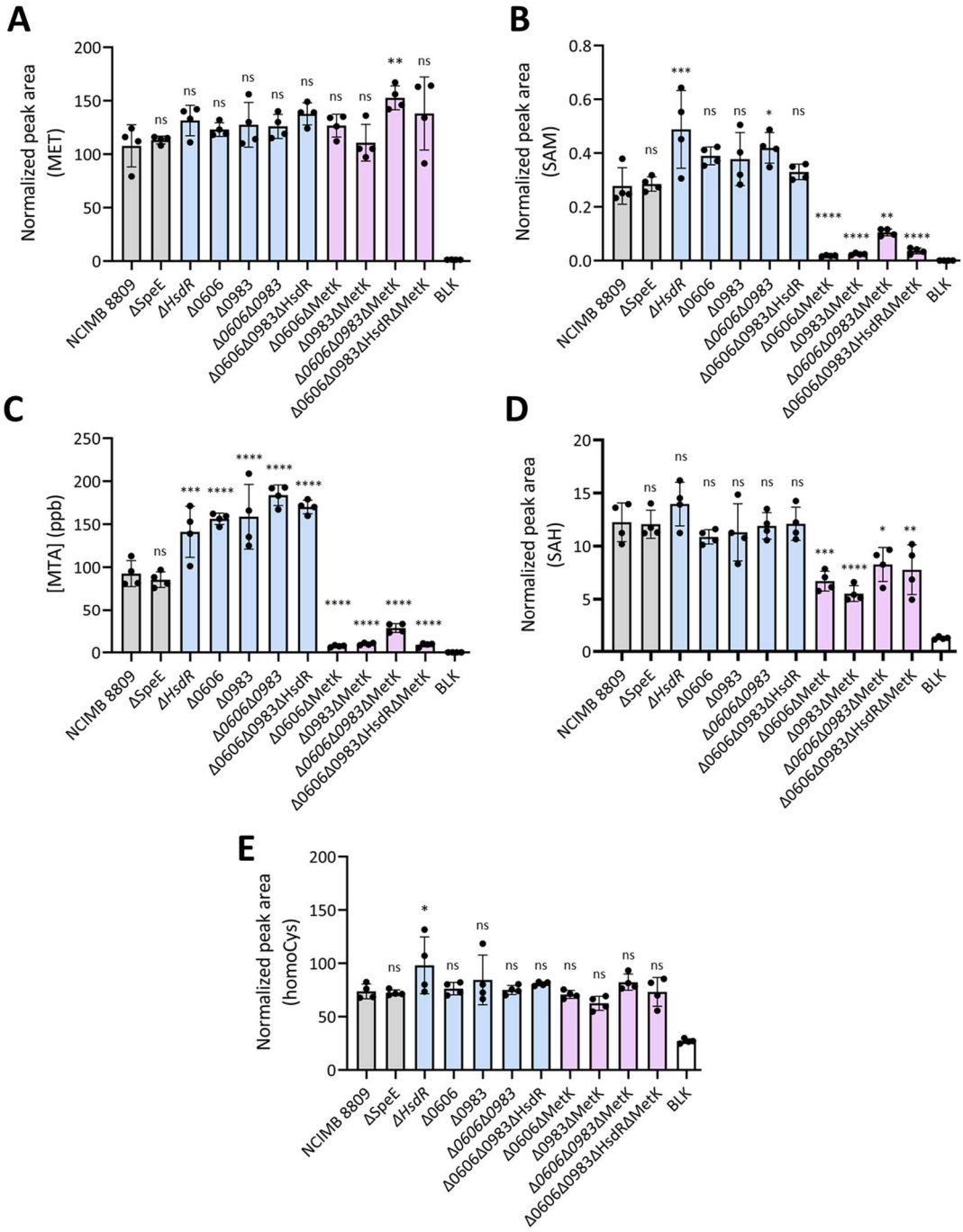


Figure 3.9. Targeted LC-MS/MS analysis for the regulations of methionine cycle pathway. Normalized peak areas for each indicated metabolite across multiple engineered *B. longum* NCIMB 8809 strains. (n=4, biological replicates; error bars, standard deviation; ns, not significant, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001; BLK, blank control)

doi:10.6342/NTU202404344

3-5. Unique RM Systems across *Bifidobacterium* Strains

The RM-disrupted strains demonstrated extremely high transformation efficiency and retained the machinery necessary for DNA methylation. By passaging pMGC-Cas9n plasmids through the Δ*0606*Δ*0983*Δ*HsdR* strain of *B. longum* NCIMB 8809 and the Δ*Sau3AI* (Q260*) strain of *B. adolecentis* DSM 20083, we should match their DNA methylation patterns to circumvent the native RM systems in the parent strains (**Figure 3.10A**). As expected, plasmids methylated by the Δ*0606*Δ*0983*Δ*HsdR* strain improved the transformation efficiency of *B. longum* NCIMB 8809 by 1700-fold, while those methylated by the Δ*Sau3AI* (Q260*) strain improved the transformation efficiency of *B. adolescentis* DSM 20083 by 400-fold, indicating successful evasion of the native RM systems (**Figure 3.10B**).

However, these methylated plasmids reduced transformation efficiency in other *Bifidobacterium* species (**Figure 3.10B**). These outcomes might be attributed to undesired methylation patterns. Even within the same species, the pMGC-Cas9n plasmid methylated by the *B. longum* NCIMB 8809 Δ*0606*Δ*0983*Δ*HsdR* strain failed to improve transformation efficiency in *B. longum* DSM 20219, suggesting the strain-specific RM systems across the *Bifidobacterium* genus.

doi:10.6342/NTU202404344

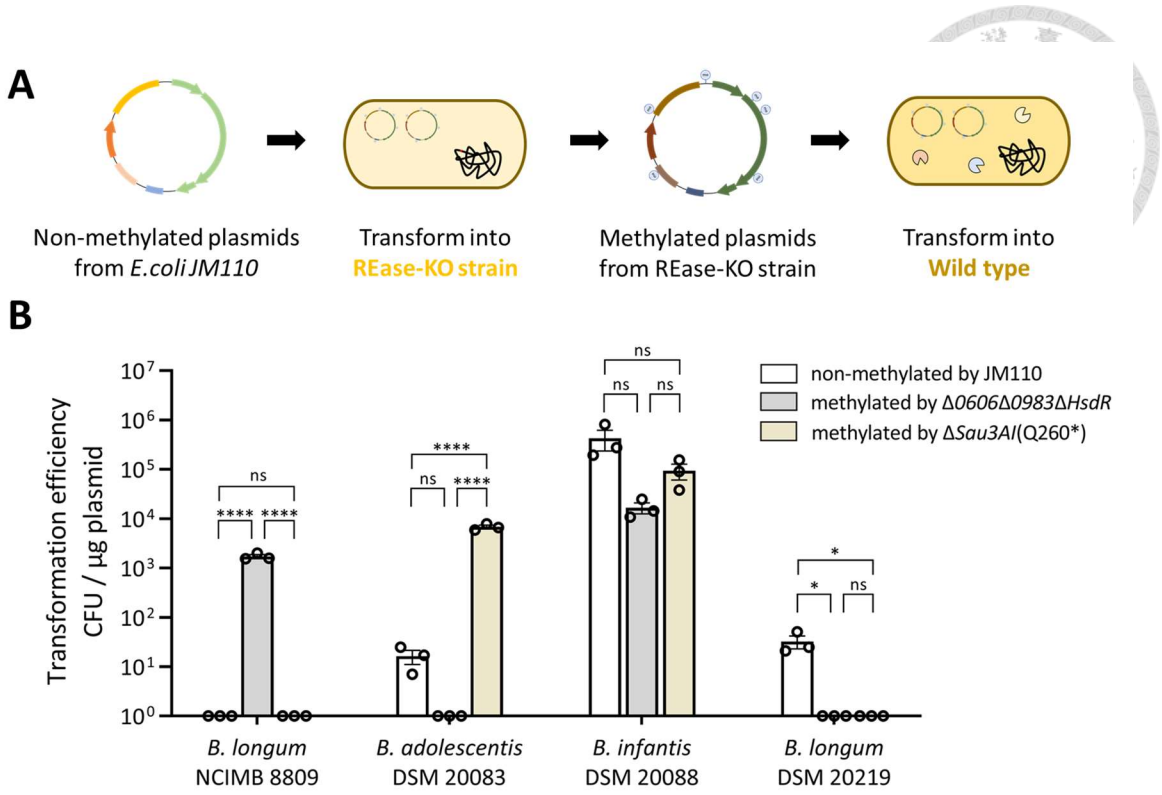


Figure 3.10. Comparison of transformation efficiency using non-methylated or methylated pMGC-Cas9n plasmids. (A) Experimental workflow: pMGC-Cas9n methylated by the Δ*0606*Δ*0983*Δ*HsdR* strain or the Δ*Sau3AI* (Q260*) strain were considered to match the methylation patterns of *B. longum* NCIMB 8809 or *B. adolescentis* DSM 20083, providing the ability to escape from RM systems. (B) Comparison of transformation efficiency using non-methylated or host-methylated pMGC-Cas9n plasmids across *B. longum* NCIMB 8809, *B. adolescentis* DSM 20083, *B. infantis* DSM 20088 and *B. longum* DSM 20219. (n=3, biological triplicates; error bars, s.e.m., ns, not significant, *P < 0.05, ****P < 0.0001)

doi:10.6342/NTU202404344

3-6. Highly Efficient and Streamlined Genome Editing in RM-Disrupted Strains

Based on 100% gene editing efficiency in REase knockout strains and the significant improvement of transformation efficiency using methylated plasmids, we hypothesized that host-methylated plasmids with ability of evading RM systems may facilitate genome editing. To evaluate this hypothesis, we targeted the bile salt hydrolase (*bsh*) gene for validation.

Bile acids serve as key metabolites that govern host fat emulsification, glucose and lipid homeostasis, cholesterol metabolism and the composition of gut microbiota.[36, 37] Imbalances in the bile acids pool have been implicated in several diseases, including *Clostridioides difficile* infection (CDI), inflammatory bowel disease (IBD) and metabolic syndrome.[38, 39] Bile salt hydrolase, only found in few bacteria such as *Bifidobacterium*, *Lactobacillus* and *Bacteroides*, is involved in catalyzing the deconjugation of glycine- or taurine-conjugated bile acids, which initiates the production of secondary bile acids and balances the bile acids pool in the gut (**Figure 3.11**). Recent studies have identified that bile salt hydrolase produces amino acid-conjugated bile acids from primary bile acids.[36, 40, 41] In our preliminary data, we also detected glutamic acid-conjugated cholic acid and other biotransformations, such as the addition of an ethyl group to deoxycholic acid, produced by *B. longum* NCIMB 8809 (**Supplementary Figure 6.9**). Given the critical

doi:10.6342/NTU202404344

role of this multifunctional enzyme, we believe that obtaining *bsh* mutant strains can help related studies for further understanding of how gut microbiota and bile acids regulation impact on human physiology.

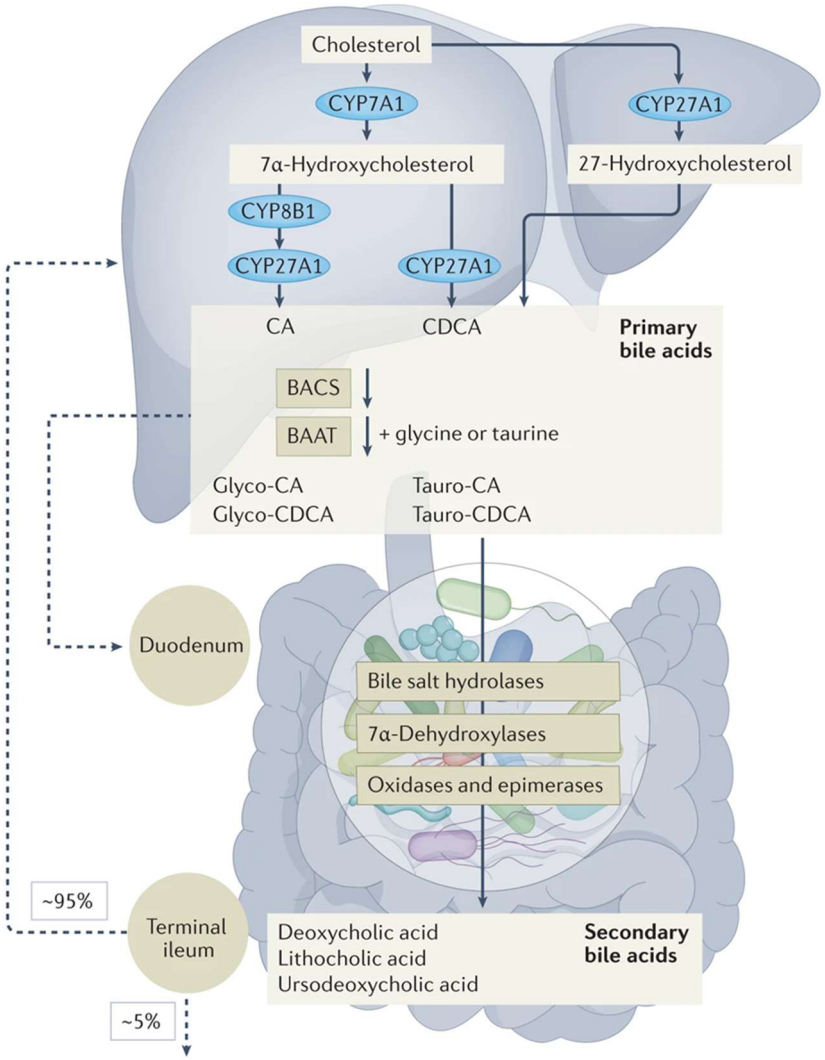


Figure 3.11. Bile acids circulation in humans. Primary bile acids are initially synthesized in the liver and then metabolized by gut microbiota. This process begins with the deconjugation of primary bile acids, mediated by bile salt hydrolases, and proceeds through several steps, ultimately leading to the production of secondary bile acids. Figure reprinted from Collins *et al* (2023)[39].

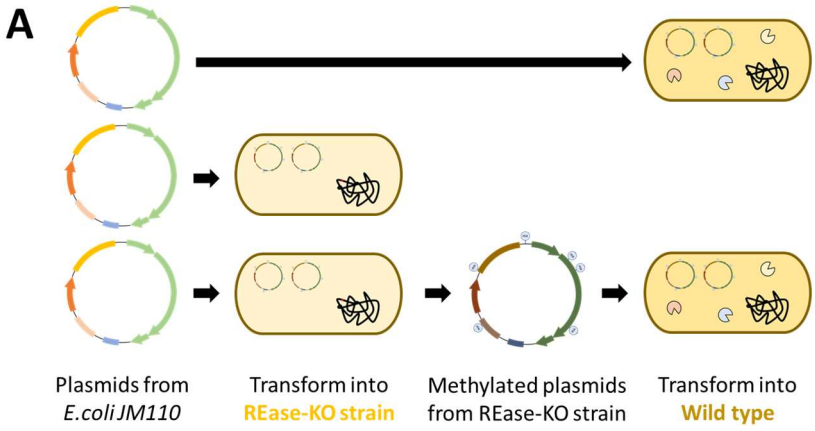
doi:10.6342/NTU202404344

Therefore, we compared three strategies involving different competent cells and the use of non-methylated/methylated plasmids to edit the *bsh* gene using the cBEST2 construct with three distinct protospacers (**Figure 3.12**). As expected, non-methylated plasmids yielded few transformants and gene mutations in the wild type strain, whereas the cBEST2-*bsh*-PS1 plasmid achieved 100% editing efficiency in the Δ*0606*Δ*0983*Δ*HsdR* strain (**Figure 3.12B**). Notably, the use of methylated plasmids significantly increased transformation efficiency (data not shown), with all constructs achieving 100% gene editing efficiency in the wild type, regardless of the protospacers used (**Figure 3.12B**). Detailed editing patterns and their corresponding editing windows are shown in **Supplementary Figure 6.10**.

Given the high transformation efficiency and intact DNA methylation machinery in the RM-disrupted *B. longum* NCIMB 8809 Δ*0606*Δ*0983*Δ*HsdR* strain, we sought to streamline the cloning and genome editing process (**Figure 3.13**). For constructing the cBEST4-*bsh*-PS1 editing plasmid, direct transformation of the all-on-one Golden-Gate assembly mixture into the Δ*0606*Δ*0983*Δ*HsdR* strain resulted in 100% editing efficiency (**Figure 3.13**; **Supplementary Figure 6.11A**), demonstrating simultaneous amplification and methylation without the use of *E. coli*, saving three days in plasmids preparation. Host-methylated cBEST editing plasmids extracted from the Δ*0606*Δ*0983*Δ*HsdR* strain

doi:10.6342/NTU202404344

were consistently edited the target gene locus in the wild type *B. longum* NCIMB 8809 strain (**Figure 3.13**; **Supplementary Figure 6.11B**). No transformants were obtained with this workflow using the parent *B. longum* NCIMB 8809 strain (**Figure 3.13**). Overall, we demonstrated a streamlined genome editing approach using RM-disrupted *Bifidobacterium* strains.



B

Strategy	*bsh*-PS1	*bsh*-PS2	*bsh*-PS3
NCIMB 8809 non-methylated	0, 0/1	0/2, 0	1/3, 3/3
Δ0606Δ0983ΔHsdR non-methylated	5/5	0/5	3/5
NCIMB 8809 methylated	5/5	5/5	5/5

Figure 3.12. Comparison of gene editing efficiencies across three strategies for RM systems evasion. (A) Plasmids were methylated by REase knockout strains to match the methylation patterns of the parent host. The methylated plasmids were transformed into the wild type. (B) Comparison of editing efficiencies using the cBEST2 editing plasmids through the three strategies.

doi:10.6342/NTU202404344

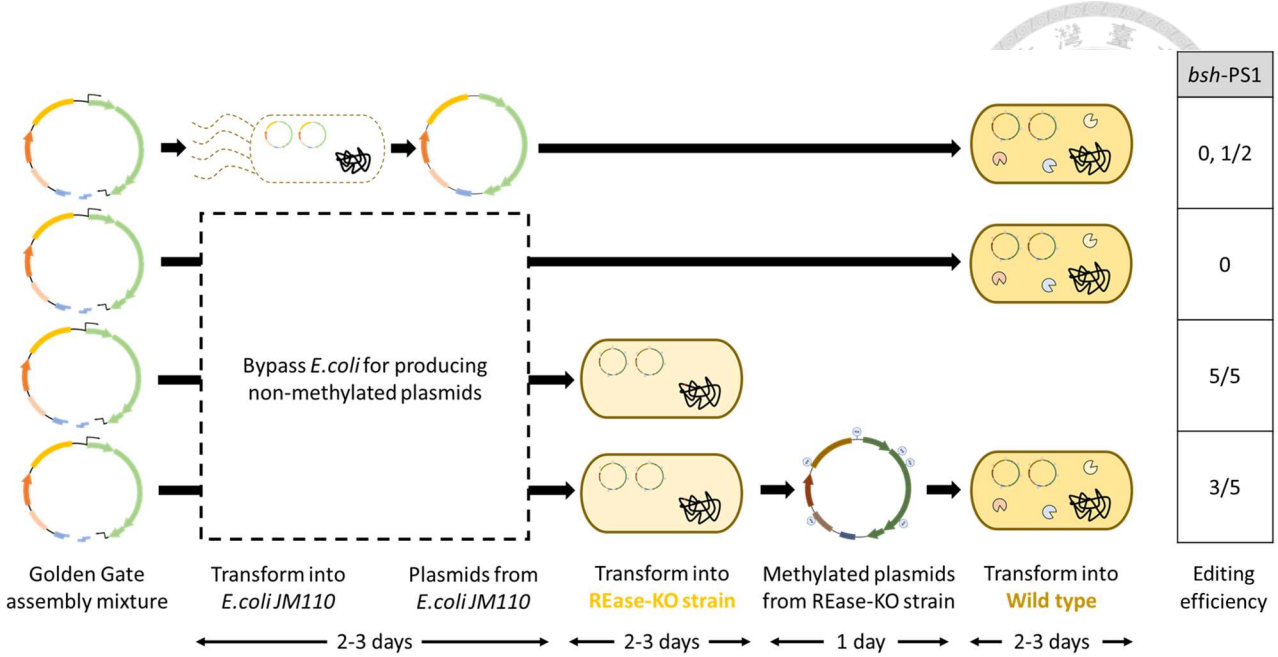


Figure 3.13. Comparison of gene editing efficiencies for streamlined genome editing workflow. Strategy for direct methylation and amplification through REase-knockout strains. One-pot Golden Gate assembly mixture for constructing the cBEST4-*bsh*-PS1 plasmid was directly transformed into the Δ*0606*Δ*0983*Δ*HsdR* strain to evaluate gene editing efficiency. The intact methylated plasmid was then extracted and re-transformed into the wild type *B. longum* NCIMB 8809.

doi:10.6342/NTU202404344

3-7. Identical cBEST Constructs for Metabolic Regulations in *Bifidobacterium* spp.

To determine if the cBEST constructs work in multiple *Bifidobacterium* species and strains other than *B. longum* NCIMB 8809 and *B. adolescentis* DSM 20083, we leveraged on the highly conserved *bsh* gene in *Bifidobacteria*, particularly 100% sequence identity at the *bsh*-PS1 genomic locus in strains *B. longum* NCIMB 8809, *B. infantis* DSM 20088 and *B. longum* DSM 20219. We were able to target *bsh*-PS1 using the identical cBEST editing plasmids. We cloned *bsh*-PS1 into cBEST2, cBEST3 and cBEST4 backbones and assessed gene editing efficiency in these strains. With the cBEST2-*bsh*-PS1 plasmid, only *B. longum* DSM 20219 successfully made C-to-T edits (**Figure 3.14A, Supplementary Figure 6.13**). Conversely, the cBEST3-*bsh*-PS1 and the cBEST4-*bsh*-PS1 plasmid yielded 100% gene editing efficiency with successful introduction of a stop codon in *B. infantis* DSM 20088 and *B. longum* DSM 20219 (**Figure 3.14 A, Supplementary Figure 6.13**). For the differences in editing windows across different strains (**Figure 3.14B**), it is possibly due to the variations in the base editor expression levels.

Next, we quantified taurocholic acid (TCA), cholic acid (CA) and deoxycholic acid (DCA) using LC-MS/MS analysis to evaluate if the BSH function is deactivated in those Δ*bsh* strains during individual strain cultures. As expected, only the wild type NCIMB 8809 and the Δ*0606*Δ*0983*Δ*HsdR* strain consumed TCA and produced CA when cultured

doi:10.6342/NTU202404344

in TCA-supplemented MRSC broth. In contrast, the Δ*bsh* and the Δ*0606*Δ*0983*Δ*HsdR*Δ*bsh* strains were unable to metabolize TCA and CA was undetectable (**Figure 3.15C**). We observed the same bile acids regulations in *bsh* knockouts of *B. infantis* DSM 20088 and *B. longum* DSM 20219 (**Figure 3.15C**).

Additionally, we tried to simulate the process in the human gut where primary bile acids are metabolized by bile salt hydrolase and then converted into secondary bile acids by the *bai* operon (**Figure 3.11 and Figure 3.15A**). Certain *Clostridium* species, such as *Clostridium scindens* DSM 5676, possess 7α-dihydroxylation activity, enabling the conversion of CA to DCA via the bile acids-inducible (*bai*) operon (**Figure 3.15A**). Therefore, we further co-cultured *Bifidobacterium* with *C. scindens* DSM 5676 (**Figure 3.15D**). In *C. scindens* DSM 5676 co-cultured with wild type *Bifidobacteria*, DCA production from TCA was successfully observed. In contrast, no DCA was detected upon co-culture of the *bsh*-deficient strains. (**Figure 3.15E**)

doi:10.6342/NTU202404344

A

Species	Strains	Target gene locus	cBEST2	cBEST3	cBEST4
B. longum	NCIMB 8809	*bsh* - PS1	0, 0/1	0/1, 1/1	0, 1/2
	DSM 20219		1/4, 6/8	2/2, 3/3	2/2, 2/2
B. infantis	DSM 20088		0/5, 0/5	5/5, 5/5	5/5, 5/5

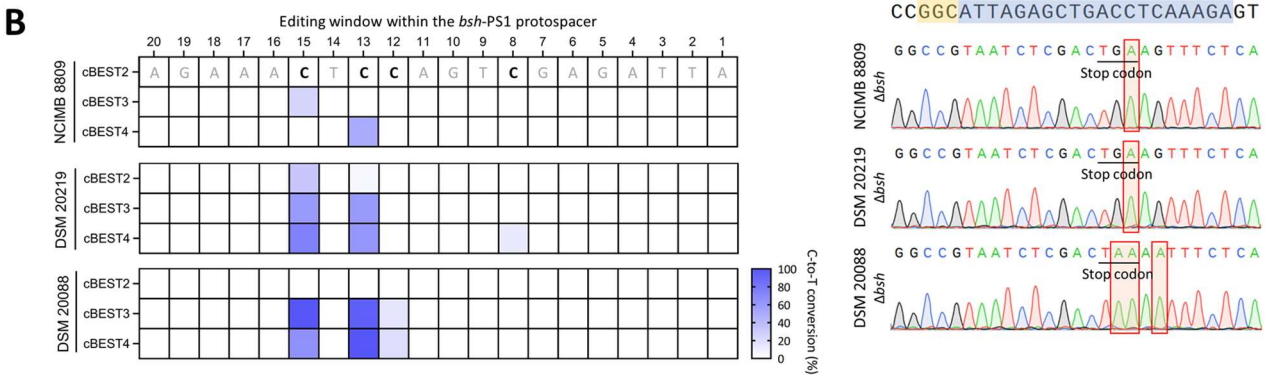


Figure 3.14. Application of cBEST constructs to edit the *bsh* gene in *Bifidobacterium* sp.. (A) Comparison of editing efficiencies using three cBEST-*bsh*-PS1 editing plasmids across *Bifidobacterium* strains. (B) Editing windows for indicated conditions. (C) Sequencing results of three successful Δ*bsh* strains using the identical cBEST4-*bsh*-PS1 editing plasmid.

doi:10.6342/NTU202404344

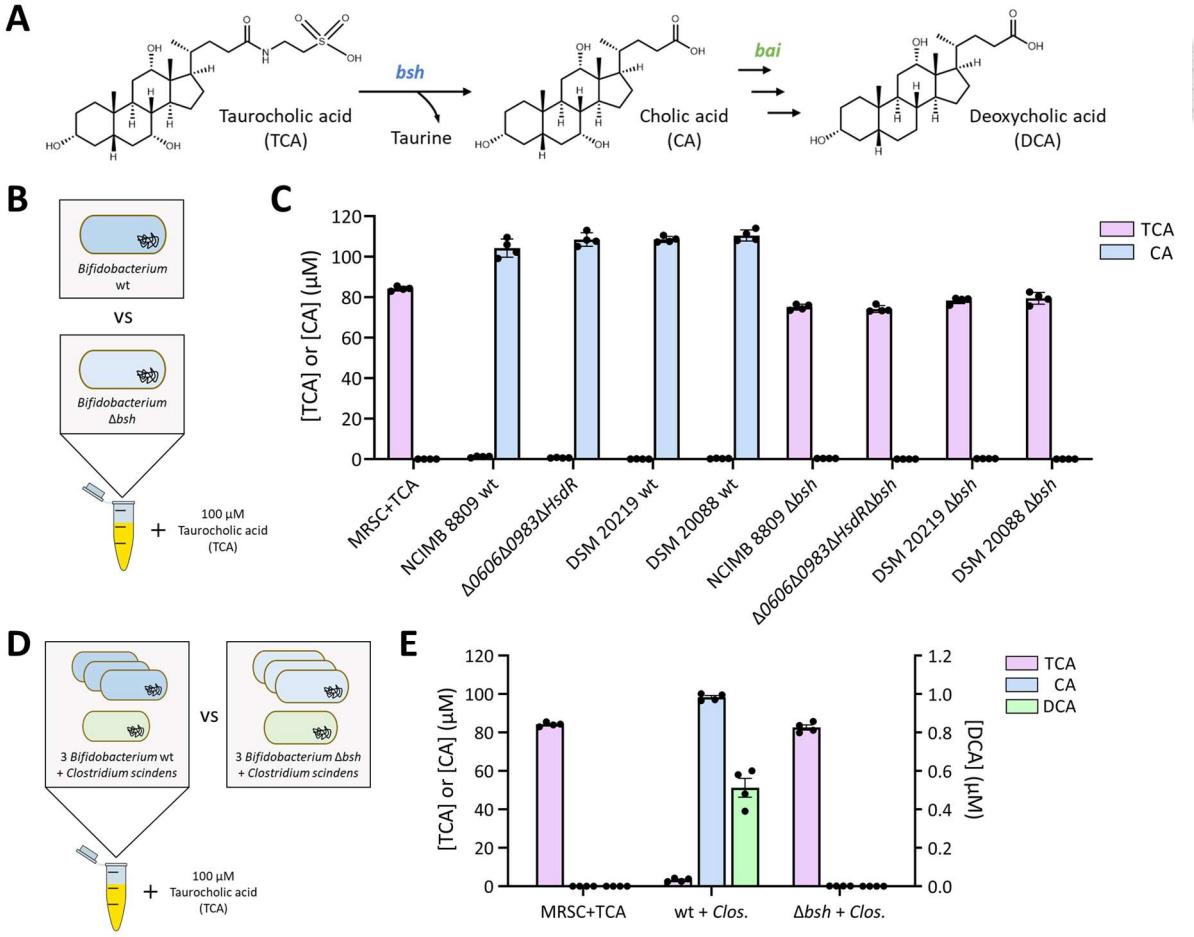


Figure 3.15. Quantitative LC-MS/MS analysis for bile acids regulations. (A) Biosynthesis pathway from TCA to DCA by gut microbiota. Bile salt hydrolase (*bsh*) deconjugates TCA into CA, and bile acid-inducible operon (*bai*) convert CA into DCA. (B) Individual wild type or Δ*bsh* strains cultures in TCA-supplemented MRSC broth for 24 h. (C) Quantification of TCA and CA metabolites in indicated *Bifidobacterium* wild type and Δ*bsh* strains. (D) Three *Bifidobacterium* wild type or Δ*bsh* strains with *Clostridium scindens* DSM 5676 (*Clos.*) co-cultures in TCA-supplemented MRSC broth for 24 h. (E) Quantification of TCA, CA and DCA metabolites in two co-cultures. (n=4, biological replicates; error bars, s.e.m.)

doi:10.6342/NTU202404344

Chapter 4. Discussion & Conclusion

We successfully developed a set of cBEST constructs for efficient genome editing in *Bifidobacteria* spp.. Our findings underscore the importance of fine tuning sgRNA and the base editor expression to strike a balance between efficient genetic modifications and cell toxicity. Notably, cBEST4 produced fewer transformants than cBEST2 (8 colonies vs 20 colonies by four individual *SpeE* protospacers), which potentially attributed to the strong tcp830 promoter driving sgRNA. Furthermore, the varying expression levels observed in pMGC-mCherry across various *Bifidobacterium* species indicated the diversity in promoter strength (**Supplementary Figure 6.1B**). This indicates the necessity of optimizing promoter combinations or characterizing new synthetic promoters, especially if our cBEST constructs fail to produce the desired mutant strains.

Bioengineering tools based on homologous recombination systems were developed before the advent of CRISPR-based methods. However, these required high transformation efficiency and extensive screening for successful recombinants.[42, 43] With the precision of CRISPR/Cas in inducing DSBs, the homologous arms can readily facilitate genome repair by introducing the insertion/deletion to achieve the gene mutation.[22] Several CRISPR loci of *Bifidobacterium* strains have been characterized,[27, 44, 45] and the successful use of the endogenous CRISPR-Cas system to edit genome in *B.*

doi:10.6342/NTU202404344

animalis subsp. lactis DSM 10140, *B. animalis* AR688 and *B. breve* strains holds promise for the utility of CRISPR-based genome editing techniques.[22, 45, 46] Although the use of endogenous CRISPR system might be more compatible with bacteria and solve the transformation efficiency issue, the diversity of CRISPR systems might lead to undesired mutations and the DSBs cause cell toxicity. Therefore, we established the programmable cBEST systems to avoid DSBs and allow precise C-to-T mutagenesis. The approach may be theoretically compatible with the entire genus of *Bifidobacterium* if the expression level of both sgRNA and the base editor can be well-controlled and the transformation efficiency is high enough.

Inaccessibility of developing genetic tools for *Bifidobacteria* is due to that the tenacious and highly diverse RM systems cause extremely low transformation efficiency. Several strategies have been explored to escape RM barriers and promote bioengineering accessibility. Techniques such as plasmid artificial modification method,[47] introduction of the MTases into *E. coli* to match the methylation patterns, and the Artificial SyngenicDNA approach,[48, 49] mutating the RM systems recognition sequences in the plasmids, have shown significant improvement in transformation efficiency. The successful demonstration of our tools enables the precise gene knockout of restriction endonucleases in *B. longum* NCIMB 8809 and *B. adolescentis* DSM 20083, providing a

doi:10.6342/NTU202404344

significant enhancement in both transformation and gene editing efficiencies. Likewise, methylated plasmids derived from the Δ*0606*Δ*0983*Δ*HsdR* strain or the Δ*Sau3AI* (Q260*) strain for matching methylation patterns demonstrated the same effects, which is consistent in previous work about RM-silent CRISPR/Cas9 toolkit for *Staphylococcus aureus*.[49] Evasion of RM systems for higher transformation efficiency allows more sophisticated genome editing techniques, such as multiplex editing.[50] The approach with a single plasmid targeting several interested genes could facilitate the study of entire metabolic pathways or biosynthetic gene clusters (BGCs).[51]

Finally, the successful generation of *bsh* knockout strains in *B. longum* NCIMB 8809, *B. longum* DSM 20219 and *B. infantis* DSM 20088 demonstrates the effectiveness and applicability of our base editing tools to the entire genus of *Bifidobacterium*. Utilizing these metabolically engineered strains in appropriate animal models will enable more comprehensive investigations into the complex interplay between host-microbiota interactions.[18, 39] These will hold promise for a better understanding of the impact on gut health and its potential effects on hosts.

In conclusion, our CRISPR/Cas9 genome editing tools have been established and held promise to apply to the entire genus of *Bifidobacterium*. Looking ahead, the ability to precisely manipulate *Bifidobacterium* genomes opens up numerous avenues for

doi:10.6342/NTU202404344

research and therapeutic development. Future studies could focus on elucidating other metabolic pathways influenced by gut microbiota, developing engineered strains with enhanced probiotic properties or higher yield of natural products, creating animal models for diseases linked to gut microbiota imbalances. These advancements pave the way for innovative approaches in both fundamental research and clinical applications, offering significant potential for improving human health.

doi:10.6342/NTU202404344

Chapter 5. References

1. Li, D.; Wang, P.; Wang, P.; Hu, X.; Chen, F., The gut microbiota: A treasure for human health. *Biotechnol. Adv.* **2016,** *34* (7), 1210-1224.

2. Fan, Y.; Pedersen, O., Gut microbiota in human metabolic health and disease. *Nat. Rev. Microbiol.* **2021,** *19* (1), 55-71.

3. Bottacini, F.; van Sinderen, D.; Ventura, M., Omics of bifidobacteria: research and insights into their health-promoting activities. *Biochem. J.* **2017,** *474* (24), 4137-4152.

4. Wong, C. B.; Odamaki, T.; Xiao, J. Z., Insights into the reason of Human-Residential Bifidobacteria (HRB) being the natural inhabitants of the human gut and their potential health-promoting benefits. *FEMS Microbiol. Rev.* **2020,** *44* (3), 369-385.

5. Derrien, M.; Turroni, F.; Ventura, M.; van Sinderen, D., Insights into endogenous *Bifidobacterium* species in the human gut microbiota during adulthood. *Trends Microbiol.* **2022,** *30* (10), 940-947.

6. Henrick, B. M.; Rodriguez, L.; Lakshmikanth, T.; Pou, C.; Henckel, E.; Arzoomand, A.; Olin, A.; Wang, J.; Mikes, J.; Tan, Z.; Chen, Y.; Ehrlich, A. M.; Bernhardsson, A. K.; Mugabo, C. H.; Ambrosiani, Y.; Gustafsson, A.; Chew, S.; Brown, H. K.; Prambs, J.; Bohlin, K.; Mitchell, R. D.; Underwood, M. A.; Smilowitz, J. T.; German, J. B.; Frese, S. A.; Brodin, P., Bifidobacteria-mediated immune system imprinting early in life. *Cell* **2021,** *184* (15), 3884-3898.e11.

7. Koppel, N.; Maini Rekdal, V.; Balskus, E. P., Chemical transformation of xenobiotics by the human gut microbiota. *Science* **2017,** *356* (6344).

8. Lyu, Q.; Chen, R. A.; Chuang, H. L.; Zou, H. B.; Liu, L.; Sung, L. K.; Liu, P. Y.; Wu, H. Y.; Chang, H. Y.; Cheng, W. J.; Wu, W. K.; Wu, M. S.; Hsu, C. C., *Bifidobacterium* alleviate metabolic disorders via converting methionine to 5'-methylthioadenosine. *Gut Microbes* **2024,** *16* (1), 2300847.

9. He, B. L.; Xiong, Y.; Hu, T. G.; Zong, M. H.; Wu, H., *Bifidobacterium* spp. as functional foods: A review of current status, challenges, and strategies. *Crit. Rev. Food Sci. Nutr.* **2023,** *63* (26), 8048-8065.

10. Zuo, F.; Chen, S.; Marcotte, H., Engineer probiotic bifidobacteria for food and biomedical applications - Current status and future prospective. *Biotechnol. Adv.* **2020,**

doi:10.6342/NTU202404344

45, 107654.

11. Zuo, F.; Marcotte, H., Advancing mechanistic understanding and bioengineering of probiotic lactobacilli and bifidobacteria by genome editing. *Curr. Opin. Biotechnol.* **2021,** *70*, 75-82.

12. Knott, G. J.; Doudna, J. A., CRISPR-Cas guides the future of genetic engineering. *Science* **2018,** *361* (6405), 866-869.

13. Ramachandran, G.; Bikard, D., Editing the microbiome the CRISPR way. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2019,** *374* (1772), 20180103.

14. Anzalone, A. V.; Koblan, L. W.; Liu, D. R., Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* **2020,** *38* (7), 824-844.

15. Banno, S.; Nishida, K.; Arazoe, T.; Mitsunobu, H.; Kondo, A., Deaminase-mediated multiplex genome editing in *Escherichia coli*. *Nat. Microbiol.* **2018,** *3* (4), 423-429.

16. Tong, Y.; Whitford, C. M.; Robertsen, H. L.; Blin, K.; Jørgensen, T. S.; Klitgaard, A. K.; Gren, T.; Jiang, X.; Weber, T.; Lee, S. Y., Highly efficient DSB-free base editing for streptomycetes with CRISPR-BEST. *Proc. Natl. Acad. Sci. USA* **2019,** *116* (41), 20366-20375.

17. Wang, L.; Xue, W.; Yan, L.; Li, X.; Wei, J.; Chen, M.; Wu, J.; Yang, B.; Yang, L.; Chen, J., Enhanced base editing by co-expression of free uracil DNA glycosylase inhibitor. *Cell Res.* **2017,** *27* (10), 1289-1292.

18. 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.; Longman, R. S.; Artis, D.; Guo, C. J., Genetic manipulation of gut microbes enables single-gene interrogation in a complex microbiome. *Cell* **2022,** *185* (3), 547-562.e22.

19. Wang, Y.; Wang, D.; Wang, X.; Tao, H.; Feng, E.; Zhu, L.; Pan, C.; Wang, B.; Liu, C.; Liu, X.; Wang, H., Highly Efficient Genome Engineering in *Bacillus anthracis* and *Bacillus cereus* Using the CRISPR/Cas9 System. *Front. Microbiol.* **2019,** *10*, 1932.

20. Hidalgo-Cantabrana, C.; Goh, Y. J.; Pan, M.; Sanozky-Dawes, R.; Barrangou, R., Genome editing using the endogenous type I CRISPR-Cas system in *Lactobacillus*

doi:10.6342/NTU202404344

crispatus. *Proc. Natl. Acad. Sci. USA* **2019,** *116* (32), 15774-15783.

21. Zheng, L.; Tan, Y.; Hu, Y.; Shen, J.; Qu, Z.; Chen, X.; Ho, C. L.; Leung, E. L.; Zhao, W.; Dai, L., CRISPR/Cas-Based Genome Editing for Human Gut Commensal *Bacteroides* Species. *ACS Synth. Biol.* **2022,** *11* (1), 464-472.

22. Pan, M.; Morovic, W.; Hidalgo-Cantabrana, C.; Roberts, A.; Walden, K. K. O.; Goh, Y. J.; Barrangou, R., Genomic and epigenetic landscapes drive CRISPR-based genome editing in *Bifidobacterium*. *Proc. Natl. Acad. Sci. USA* **2022,** *119* (30).

23. O'Connell Motherway, M.; O'Driscoll, J.; Fitzgerald, G. F.; Van Sinderen, D., Overcoming the restriction barrier to plasmid transformation and targeted mutagenesis in *Bifidobacterium breve* UCC2003. *Microb. Biotechnol.* **2009,** *2* (3), 321-32.

24. O'Callaghan, A.; Bottacini, F.; O'Connell Motherway, M.; van Sinderen, D., Pangenome analysis of *Bifidobacterium longum* and site-directed mutagenesis through by-pass of restriction-modification systems. *BMC Genomics* **2015,** *16*, 832.

25. Roberts, R. J.; Belfort, M.; Bestor, T.; Bhagwat, A. S.; Bickle, T. A.; Bitinaite, J.; Blumenthal, R. M.; Degtyarev, S.; Dryden, D. T.; Dybvig, K.; Firman, K.; Gromova, E. S.; Gumport, R. I.; Halford, S. E.; Hattman, S.; Heitman, J.; Hornby, D. P.; Janulaitis, A.; Jeltsch, A.; Josephsen, J.; Kiss, A.; Klaenhammer, T. R.; Kobayashi, I.; Kong, H.; Krüger, D. H.; Lacks, S.; Marinus, M. G.; Miyahara, M.; Morgan, R. D.; Murray, N. E.; Nagaraja, V.; Piekarowicz, A.; Pingoud, A.; Raleigh, E.; Rao, D. N.; Reich, N.; Repin, V. E.; Selker, E. U.; Shaw, P. C.; Stein, D. C.; Stoddard, B. L.; Szybalski, W.; Trautner, T. A.; Van Etten, J. L.; Vitor, J. M.; Wilson, G. G.; Xu, S. Y., A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.* **2003,** *31* (7), 1805-12.

26. Beaulaurier, J.; Schadt, E. E.; Fang, G., Deciphering bacterial epigenomes using modern sequencing technologies. *Nat. Rev. Genet.* **2019,** *20* (3), 157-172.

27. Bottacini, F.; Morrissey, R.; Roberts, R. J.; James, K.; van Breen, J.; Egan, M.; Lambert, J.; van Limpt, K.; Knol, J.; Motherway, M. O.; van Sinderen, D., Comparative genome and methylome analysis reveals restriction/modification system diversity in the gut commensal *Bifidobacterium breve*. *Nucleic Acids Res.* **2018,** *46* (4), 1860-1877.

doi:10.6342/NTU202404344

28. Park, M. J.; Park, M. S.; Ji, G. E., Improvement of electroporation-mediated transformation efficiency for a *Bifidobacterium* strain to a reproducibly high level. *J. Microbiol. Methods* **2019,** *159*, 112-119.

29. Grimm, V.; Gleinser, M.; Neu, C.; Zhurina, D.; Riedel, C. U., Expression of fluorescent proteins in bifidobacteria for analysis of host-microbe interactions. *Appl. Environ. Microbiol.* **2014,** *80* (9), 2842-50.

30. Shao, Z.; Rao, G.; Li, C.; Abil, Z.; Luo, Y.; Zhao, H., Refactoring the silent spectinabilin gene cluster using a plug-and-play scaffold. *ACS Synth. Biol.* **2013,** *2* (11), 662-9.

31. Rodríguez-García, A.; Combes, P.; Pérez-Redondo, R.; Smith, M. C.; Smith, M. C., Natural and synthetic tetracycline-inducible promoters for use in the antibiotic-producing bacteria *Streptomyces*. *Nucleic Acids Res.* **2005,** *33* (9), e87.

32. Wang, W.; Li, X.; Wang, J.; Xiang, S.; Feng, X.; Yang, K., An engineered strong promoter for streptomycetes. *Appl. Environ. Microbiol.* **2013,** *79* (14), 4484-92.

33. Bai, C.; Zhang, Y.; Zhao, X.; Hu, Y.; Xiang, S.; Miao, J.; Lou, C.; Zhang, L., Exploiting a precise design of universal synthetic modular regulatory elements to unlock the microbial natural products in *Streptomyces*. *Proc. Natl. Acad. Sci. USA* **2015,** *112* (39), 12181-6.

34. Zhao, Y.; Li, G.; Chen, Y.; Lu, Y., Challenges and Advances in Genome Editing Technologies in *Streptomyces*. *Biomolecules* **2020,** *10* (5).

35. Arroyo-Olarte, R. D.; Bravo Rodríguez, R.; Morales-Ríos, E., Genome Editing in Bacteria: CRISPR-Cas and Beyond. *Microorganisms* **2021,** *9* (4).

36. Gentry, E. C.; Collins, S. L.; Panitchpakdi, M.; Belda-Ferre, P.; Stewart, A. K.; Carrillo Terrazas, M.; Lu, H. H.; Zuffa, S.; Yan, T.; Avila-Pacheco, J.; Plichta, D. R.; Aron, A. T.; Wang, M.; Jarmusch, A. K.; Hao, F.; Syrkin-Nikolau, M.; Vlamakis, H.; Ananthakrishnan, A. N.; Boland, B. S.; Hemperly, A.; Vande Casteele, N.; Gonzalez, F. J.; Clish, C. B.; Xavier, R. J.; Chu, H.; Baker, E. S.; Patterson, A. D.; Knight, R.; Siegel, D.; Dorrestein, P. C., Reverse metabolomics for the discovery of chemical structures from humans. *Nature* **2024,** *626* (7998), 419-426.

37. Foley, M. H.; O'Flaherty, S.; Barrangou, R.; Theriot, C. M., Bile salt hydrolases:

doi:10.6342/NTU202404344

Gatekeepers of bile acid metabolism and host-microbiome crosstalk in the gastrointestinal tract. *PLoS Pathog.* **2019,** *15* (3), e1007581.

38. Wahlström, A.; Sayin, S. I.; Marschall, H. U.; Bäckhed, F., Intestinal Crosstalk between Bile Acids and Microbiota and Its Impact on Host Metabolism. *Cell Metab.* **2016,** *24* (1), 41-50.

39. Collins, S. L.; Stine, J. G.; Bisanz, J. E.; Okafor, C. D.; Patterson, A. D., Bile acids and the gut microbiota: metabolic interactions and impacts on disease. *Nat. Rev. Microbiol.* **2023,** *21* (4), 236-247.

40. Quinn, R. A.; Melnik, A. V.; Vrbanac, A.; Fu, T.; Patras, K. A.; Christy, M. P.; Bodai, Z.; Belda-Ferre, P.; Tripathi, A.; Chung, L. K.; Downes, M.; Welch, R. D.; Quinn, M.; Humphrey, G.; Panitchpakdi, M.; Weldon, K. C.; Aksenov, A.; da Silva, R.; Avila-Pacheco, J.; Clish, C.; Bae, S.; Mallick, H.; Franzosa, E. A.; Lloyd-Price, J.; Bussell, R.; Thron, T.; Nelson, A. T.; Wang, M.; Leszczynski, E.; Vargas, F.; Gauglitz, J. M.; Meehan, M. J.; Gentry, E.; Arthur, T. D.; Komor, A. C.; Poulsen, O.; Boland, B. S.; Chang, J. T.; Sandborn, W. J.; Lim, M.; Garg, N.; Lumeng, J. C.; Xavier, R. J.; Kazmierczak, B. I.; Jain, R.; Egan, M.; Rhee, K. E.; Ferguson, D.; Raffatellu, M.; Vlamakis, H.; Haddad, G. G.; Siegel, D.; Huttenhower, C.; Mazmanian, S. K.; Evans, R. M.; Nizet, V.; Knight, R.; Dorrestein, P. C., Global chemical effects of the microbiome include new bile-acid conjugations. *Nature* **2020,** *579* (7797), 123-129.

41. Rimal, B.; Collins, S. L.; Tanes, C. E.; Rocha, E. R.; Granda, M. A.; Solanki, S.; Hoque, N. J.; Gentry, E. C.; Koo, I.; Reilly, E. R.; Hao, F.; Paudel, D.; Singh, V.; Yan, T.; Kim, M. S.; Bittinger, K.; Zackular, J. P.; Krausz, K. W.; Desai, D.; Amin, S.; Coleman, J. P.; Shah, Y. M.; Bisanz, J. E.; Gonzalez, F. J.; Vanden Heuvel, J. P.; Wu, G. D.; Zemel, B. S.; Dorrestein, P. C.; Weinert, E. E.; Patterson, A. D., Bile salt hydrolase catalyses formation of amine-conjugated bile acids. *Nature* **2024,** *626* (8000), 859-863.

42. Hirayama, Y.; Sakanaka, M.; Fukuma, H.; Murayama, H.; Kano, Y.; Fukiya, S.; Yokota, A., Development of a double-crossover markerless gene deletion system in *Bifidobacterium longum*: functional analysis of the α-galactosidase gene for raffinose assimilation. *Appl. Environ. Microbiol.* **2012,** *78* (14), 4984-94.

doi:10.6342/NTU202404344

43. Zuo, F.; Zeng, Z.; Hammarström, L.; Marcotte, H., Inducible Plasmid Self-Destruction (IPSD) Assisted Genome Engineering in Lactobacilli and Bifidobacteria. *ACS Synth. Biol.* **2019,** *8* (8), 1723-1729.

44. Hidalgo-Cantabrana, C.; Crawley, A. B.; Sanchez, B.; Barrangou, R., Characterization and Exploitation of CRISPR Loci in *Bifidobacterium longum*. *Front. Microbiol.* **2017,** *8*, 1851.

45. Han, X.; Chang, L.; Chen, H.; Zhao, J.; Tian, F.; Ross, R. P.; Stanton, C.; van Sinderen, D.; Chen, W.; Yang, B., Harnessing the endogenous Type I-C CRISPR-Cas system for genome editing in *Bifidobacterium breve*. *Appl. Environ. Microbiol.* **2024**, e0207423.

46. Li, J.; Song, X.; Xiong, Z.; Wang, G.; Xia, Y.; Yang, Y.; Ai, L., Establishment of CRISPR-Cas9 system in *Bifidobacteria animalis* AR668. *Microb. Cell Fact.* **2023,** *22* (1), 112.

47. Yasui, K.; Kano, Y.; Tanaka, K.; Watanabe, K.; Shimizu-Kadota, M.; Yoshikawa, H.; Suzuki, T., Improvement of bacterial transformation efficiency using plasmid artificial modification. *Nucleic Acids Res.* **2009,** *37* (1), e3.

48. Johnston, C. D.; Cotton, S. L.; Rittling, S. R.; Starr, J. R.; Borisy, G. G.; Dewhirst, F. E.; Lemon, K. P., Systematic evasion of the restriction-modification barrier in bacteria. *Proc. Natl. Acad. Sci. USA* **2019,** *116* (23), 11454-11459.

49. Yang, P.; Yang, J.; Lin, T.; Liu, Q.; Yin, Y.; Chen, D.; Yang, S., Efficient Genome Editing in Most *Staphylococcus aureus* by Using the Restriction-Modification System Silent CRISPR-Cas9 Toolkit. *ACS Synth. Biol.* **2023,** *12* (11), 3340-3351.

50. McCarty, N. S.; Graham, A. E.; Studená, L.; Ledesma-Amaro, R., Multiplexed CRISPR technologies for gene editing and transcriptional regulation. *Nat. Commun.* **2020,** *11* (1), 1281.

51. Volke, D. C.; Martino, R. A.; Kozaeva, E.; Smania, A. M.; Nikel, P. I., Modular (de)construction of complex bacterial phenotypes by CRISPR/nCas9-assisted, multiplex cytidine base-editing. *Nat. Commun.* **2022,** *13* (1), 3026.

52. Roberts, R. J.; Vincze, T.; Posfai, J.; Macelis, D., REBASE--a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.* **2015,** *43* (Database issue), D298-9.

doi:10.6342/NTU202404344

53. Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W. T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C. C.; Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C. C.; Yang, Y. L.; Humpf, H. U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; P, C. A. B.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P. M.; Phapale, P.; Nothias, L. F.; Alexandrov, T.; Litaudon, M.; Wolfender, J. L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D. T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. O.; Pogliano, K.; Linington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N., Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016,** *34* (8), 828-837.

doi:10.6342/NTU202404344

Chapter 6. Supporting Information

Supplementary Table 6.1. Abbreviations in this study.

Abbreviation	Full name
bsh	bile salt hydrolase
C	cytosine
CA	cholic acid
Cas	CRISPR-associated protein
cBEST	cytosine base editor system
CDI	*Clostridioides difficile* infection
CE	collision energy
CFU	colony forming units
CRISPR	clustered regularly interspaced short palindromic repeats
CXP	cell exit potential
DCA	deoxycholic acid
dcSAM	S-adenosyl-L-methioninamine (decarboxylated SAM)
DP	declustering potential
DSB	double strand break
EP	entrance potential
HESI	heated electrospray ionization
homoCys	L-homocysteine
IBD	inflmmatory bowel disease
MET	L-methionine
MetK	S-adenosylmethionine synthase
MRS	de Man, Rogosa, Sharpe
MRSC	MRS broth with 0.05% wt/vol L-cysteine
MS	mass spectrometry
MTA	5'-S-methyl-5'-thioadenosine
MTase, DNMT	DNA methyltransferase
ORF	open reading frame
PAM	protospacer adjacent motif
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PS	protospacer
Q1	parent ion
Q3	daughter ion
QqQ	triple quadrupole mass spectrometer (SCIEX 5500)
REase	restriction endonuclease
RFU	relative fluorescence unit
RM	restriction modification
RT	retention time
SAH	S-adenosyl-L-homocysteine
SAM	S-adenosyl-L-methionine
SMRT	single-molecule real-time
SpeE	spermine/spermidine synthase
T	thymine
TCA	taurocholic acid
UGI	uracil glycosylase inhibitor
UPLC	ultraperformance liquid chromatography

doi:10.6342/NTU202404344

Supplementary Table 6.2. Chemicals used in this study.

Chemicals	Manufacturer	Cas/LOT No.
5'-methylthioadenosine	Chem Scene	2457-80-9
5'-Deoxy-5'-(methylthio)adenosine-d3	Toronto Research Chemicals	D242602
Acetonitile	JT Baker	75-05-8
AGAR, Bacteriological	Acumedia	111750A
Ammonium acetate	Sigma Aldrich	631618
Chloramphenicol	Sigma Aldrich	56-75-7
Cholic acid	Sigma Aldrich	81-25-4
Cholic acid-d4	Cambridge Isotope Laboratories	116380-66-6
Deoxycholic acid	Tokyo Chemical Industry	302-95-4
Difco Agar Noble	BD	6271644
Ethanol	Honeywell	64-17-5
Formic acid	ACROS	64-18-6
Glycerol	BioShop	3A81066
L-Cysteine	Bioshop	52-90-4
L-homocysteine	Sigma Aldrich	6027-13-0
L-methionine	Tokyo Chemical Industry	63-68-3
L-methionine-^{13}C,d3	Cambridge Isotope Laboratories	73488-65-0
Lactobacilli MRS broth	NEOGEN	US113980G
Methanol	DUKSAN	67-56-1
PBS Buffer 10X	BIOMAN	PBS101000
S-adenosyl-L-homocysteine	Sigma Aldrich	979-92-0
S-adenosyl-L-methionine	Sigma Aldrich	86867-01-8
Sodium chloride	Cyrusbioscience	1140111
Sucrose	JT Baker	57-50-1
TAE Buffer 50X	BIOMAN	TAE501000
Taurocholic acid	Sigma Aldrich	345909-26-4
YEATaq II DNA polymerase	Yeastern Biotech	FYT601-500U

doi:10.6342/NTU202404344

Supplementary Table 6.3. Strains used in this study.

Strains	Description	source
E.coli		
DH10β	For routine plasmids maintenance and cloning	NEB
JM110	For unmethylated (Dam- Dcm-) plasmid expression	ATCC
B. longum subsp. longum		
NCIMB 8809	isolate from Nursling stools	BCRC
Δ*SpeE* (Q117*)	NCIMB 8809 mutant in *SpeE* (Q117*) using cBEST2-*SpeE* -PS1 plasmid	This study
Δ*SpeE* (Q198*)	NCIMB 8809 mutant in *SpeE* (Q198*) using cBEST2-*SpeE* -PS2 plasmid	This study
Δ*HsdR*	NCIMB 8809 mutant in *HsdR* (W29*) using methylated cBEST2-*HsdR* -PS1 plasmid	This study
Δ*0983*	NCIMB 8809 mutant in *EcoRII_0983* (W138*) using cBEST2-*0983* -PS1 plasmid	This study
Δ*0983* Δ*MetK*	Δ*EcoRII_0983* mutant in *MetK* (Q134*) using cBEST4-*MetK* -PS plasmid	This study
Δ*0606*	NCIMB 8809 mutant in *EcoRII_0606* (Q64*) using cBEST2-*0606* -PS2 plasmid	This study
Δ*0606* Δ*MetK*	Δ*EcoRII_0606* mutant in *MetK(Q134*)* using cBEST4-*MetK* -PS plasmid	This study
Δ*0606* Δ*0983*	Δ*EcoRII_0606* mutant in *EcoRII_0983* (W138*) using cBEST2-*0983* -PS1 plasmid	This study
Δ*0606* Δ*0983* Δ*MetK*	The Δ*0606* Δ*0983* strain mutant in *MetK* (Q134*) using cBEST4-*MetK* -PS plasmid	This study
Δ*0606* Δ*0983* Δ*HsdR*	The Δ*0606* Δ*0983* strain mutant in *HsdR* (W29*) using cBEST2-*HsdR* -PS1 plasmid	This study
Δ*0606* Δ*0983* Δ*HsdR* Δ*bsh*	The Δ*0606* Δ*0983* Δ*HsdR* strain mutant in *bsh* (W22*) using cBEST2-bsh-PS1 plasmid	This study
Δ*0606* Δ*0983* Δ*HsdR* Δ*MetK*	The Δ*0606* Δ*0983*Δ*HsdR* strain mutant in *MetK* (Q134*) using cBEST4-*MetK* -PS plasmid	This study
Δ*bsh*	NCIMB 8809 mutant in *bsh* (W22*) using methylated cBEST2-*bsh* -PS1 plasmid	This study
DSM 20219	type strain, isolate from adult intestine	BCRC
Δ*bsh*	DSM 20219 mutant in *bsh* (W22*) using cBEST4-*bsh* -PS1 plasmid	This study
B. adolescentis		
DSM 20083	type strain, isolate from adult intestine	BCRC
Δ*Sau3AI* (Q260*)	spontaneous mutant in *Sau3AI* (Q260*)	This study
B. longum subsp. infantis		
DSM 20088	type strain, isolate from infant intestine	BCRC
Δ*bsh*	DSM 20088 mutant in *bsh* (W22*) using cBEST4-*bsh* -PS1 plasmid	This study
Clostridium scindens		
DSM 5676	type strain, isolate from human faeces	BCRC

doi:10.6342/NTU202404344

Supplementary Table 6.4. Protospacers used in this study.

Bacterial strain	Target gene	Genbank code	PS code	Sequence (5'-3')	PAM (NGG)
B. longum NCIMB 8809	*SpeE*	ALO72771.1	PS1	TGACCAGAAGATCACCGATC	TGG
			PS2	CGGCCAAGGTTCGATCAACG	AGG
			PS3	GGTTTCCCAGAACGCGAAGC	CGG
			PS4	TGGACTCGCCCTCATAGGCC	AGG
	EcoRII_0606	ALO72279.1	PS2	GACTTTCAGATGCAGCCTCA	GGG
	EcoRII_0983	ALO72654.1	PS1	GTACCATGTGGTTTCAAGGT	CGG
	HsdR	ALO73026.1	PS1	TCGTACCATTCCTCGTTCAG	TGG
	MetK	ALO73020.1	PS	GGCGATCAGGGCGTGATGTT	CGG
	bsh	ALO72466.1	PS1	AGAAACTCCAGTCGAGATTA	CGG
			PS2	GCTCGTCTCCCAGATTGTGC	CGG
			PS3	CTCGTCTCCCAGATTGTGCC	GGG
B. adolescentis DSM 20083	*Sau3AI*	BAF40013.1	PS1	TGCAAATCAGCTTTCGGAAT	CGG
B. infantis DSM 20088	*bsh*	ACJ52536.1	PS1	AGAAACTCCAGTCGAGATTA	CGG
B. longum DSM 20219	*bsh*	-	PS1	AGAAACTCCAGTCGAGATTA	CGG

Note. Fwd: Forward strand of target gene; Rev: Reverse strand of target gene.

The sequence of *bsh*-PS1 is 100% conserved in *B. longum* NCIMB 8809, *B. infantis* DSM 20088, *B. breve* DSM 20213 and *B. longum* DSM 20219.

doi:10.6342/NTU202404344

Supplementary Table 6.5. Plasmids used in this study.

Construct	Description	Source
pMGC-mCherry	To examine the portability of various promoters	(29)
P3-mCherry		This study
P6-mCherry		This study
P_{tcp830}-mCherry		This study
P_{kasO*}-mCherry		This study
$P_{kasO*17}$-mCherry		This study
pMGC-Cas9n	To evaluate the transformation efficiency	This study
cBEST1-SpeE-PS1	To evaluate the promoters combination for cBEST and to engineer the *SpeE* gene in NCIMB 8809	This study
cBEST1-SpeE-PS2		This study
cBEST1-SpeE-PS3		This study
cBEST1-SpeE-PS4		This study
cBEST2-SpeE-PS1		This study
cBEST2-SpeE-PS2		This study
cBEST2-SpeE-PS3		This study
cBEST2-SpeE-PS4		This study
cBEST3-SpeE-PS1		This study
cBEST3-SpeE-PS2		This study
cBEST3-SpeE-PS3		This study
cBEST3-SpeE-PS4		This study
cBEST4-SpeE-PS1		This study
cBEST4-SpeE-PS2		This study
cBEST4-SpeE-PS3		This study
cBEST4-SpeE-PS4		This study
cBEST2-0983-PS1	To construct the RM-disrupted strains in NCIMB 8809	This study
cBEST2-0606-PS2		This study
cBEST2-HsdR-PS1		This study
cBEST4-MetK-PS	To construct the MetK-deficient strains in NCIMB 8809	This study
cBEST2-bsh-PS1	To construct the bsh-deficient strains in Bifidobacterium spp.	This study
cBEST2-bsh-PS2		This study
cBEST2-bsh-PS3		This study
cBEST3-bsh-PS1		This study
cBEST4-bsh-PS1		This study
cBEST1-Sau3AI-PS1	To construct the RM-disrupted strains in DSM 20083	This study
cBEST2-Sau3AI-PS1		This study
cBEST3-Sau3AI-PS1		This study
cBEST4-Sau3AI-PS1		This study

Note. All the plasmids used in this study were kindly made by Wan-Chi Hsiao and provided us to test.

doi:10.6342/NTU202404344

Supplementary Table 6.6. Primers used for PCR and sequencing in this study.

Primer names	Sequence(5'-3')
SpeE -PCR-F	CATGTCCGCTGATCTACTTCATCTCG
SpeE -PCR-R	TTGTCGTCGGTGAGAATAGTCGACG
SpeE -seq	TCTTGGCGAACAGTTCACGGT
MetK -PCR-F	GTTCCTGGTCTTCGGCGAAGTCAC
MetK -PCR-R	CGTAGGTGTCCACGATGATCTTGCGG
MetK -seq	ATATTGCGACGTGCAGTCCAAGGTC
bsh -PCR-F	GTGCAAGCCAACCAAAGCGATGG
bsh -PCR-R	CGTTGCTGACGCACATGTAGTTGC
bsh- seq	GCACATCGACGTCATCATGATGCAC
HsdR -PCR-F	ATGGCTGTTTCCAAAGTCGAATCTCG
HsdR- PCR-R	GCCGTCAATCATGGTGCTCCAT
HsdR -seq	CACGAACATGCTCGGTATCTGCTTC
0606 -PCR-F	TGCCAGACTCAACACTCCTGTAG
0606 -PCR-R	ATAGCTGCGCTTGTCACCATCC
0606 -Seq	AACAATCTCCCGGTACGGCA
0983 -PCR-F	ATGCAGGCGGATGGAGAAGA
0983 -PCR-R	CAGGAAATCGGGTTTCTTCCCG
0983 -Seq	GTATTCGATCACCACGTCGTC
Sau3AI -PCR-F	CATGGTGTACTACTACGATGACCGC
Sau3AI -PCR-R	CCAGGAAGTAAATTGGTTCCTCGTCC
Sau3AI -seq	CCTCCAACAAGACGGTCATTACCG

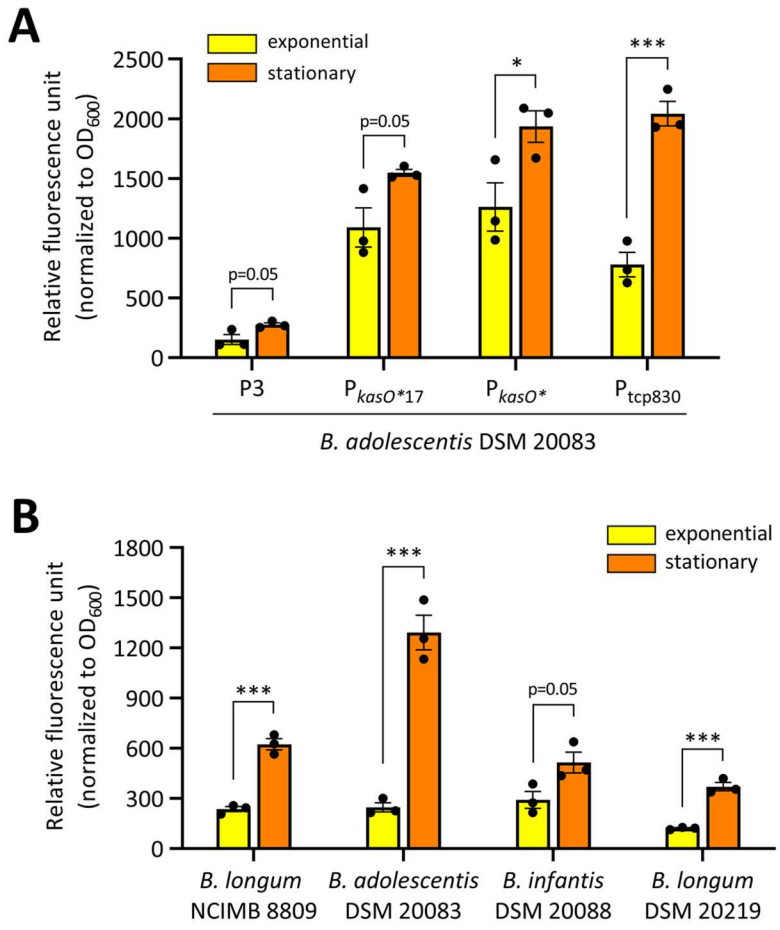
doi:10.6342/NTU202404344

Supplementary Table 6.7. List of RM systems searched from REBASE[52].

Strain	REase	MTase	RM type	Predicted recognition motif	Sites on pMGC-mCherry	Sites on pMGC-Cas9n
B. longum NCIMB 8809	*HsdR*	*HsdM*	I	5'-GATN$_5$TGCC-3'	0	3
	MvaI, EcoRII	*606*	II	5'-CCWGG-3'	10	34
	-	*958*	II	5'-TCGGCCGA-3'	0	0
B. adolescentis DSM 20083	1229, *Sau3AI*	1233	II	5'-GATC-3'	15	56
	1281, *Rkpn2KI*	1283	II	5'-CCNGG-3'	18	55
B. infantis DSM 20088	0091	0092	II	5'-GGCGCC-3'	2	7
	289R	289M	II	5'-CTGCAG-3'	2	3
	1146	1145	II	5'-GTCGAC-3'	1	7
	-	1196	II	5'-GAATTC-3'	1	1
	-	1215, 1346	II	5'-AAGCTT-3'	1	0
	1324	1324	II	5'-GAGGAC-3'	6	8

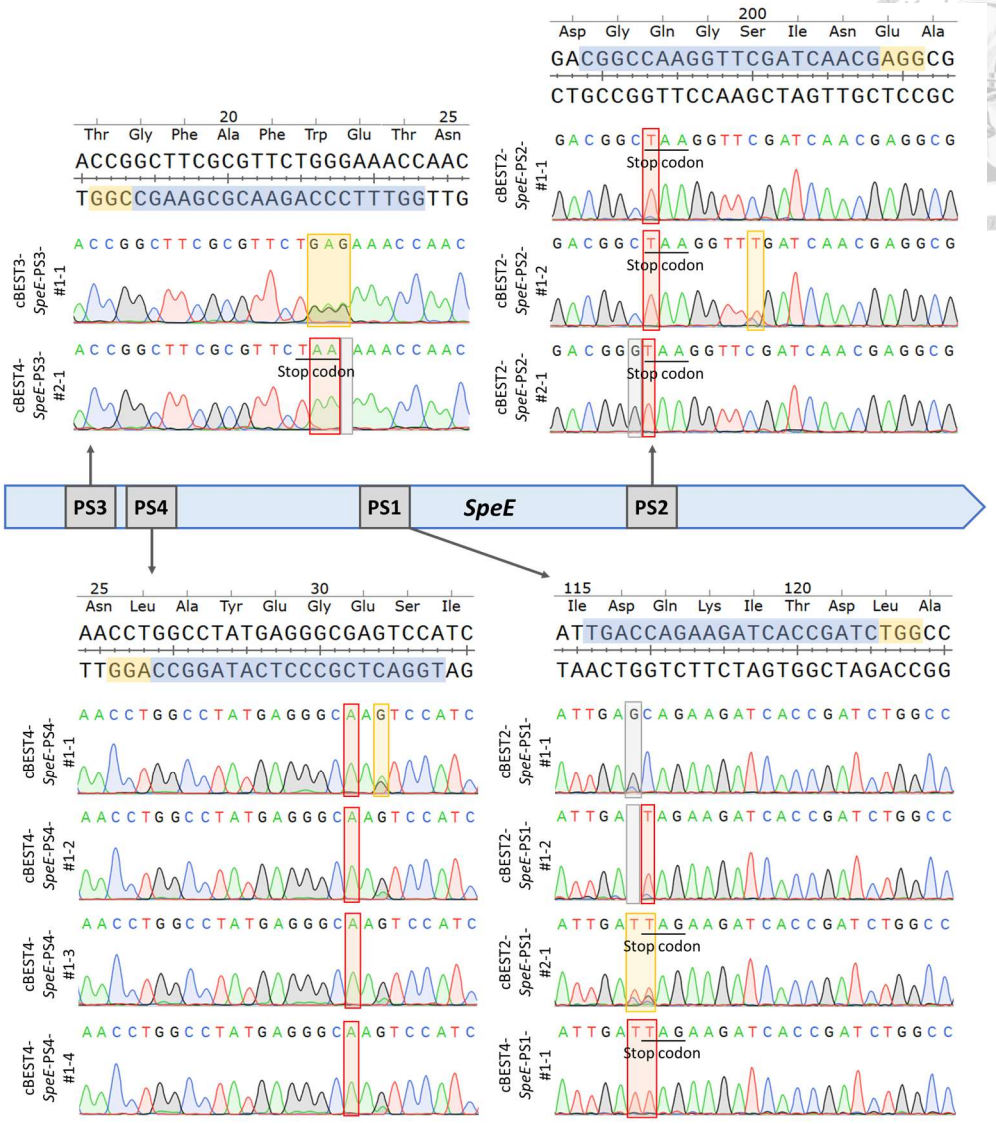
Note. *B. longum* DSM 20219 is not listed since there is no whole genome data uploaded to any database.

doi:10.6342/NTU202404344



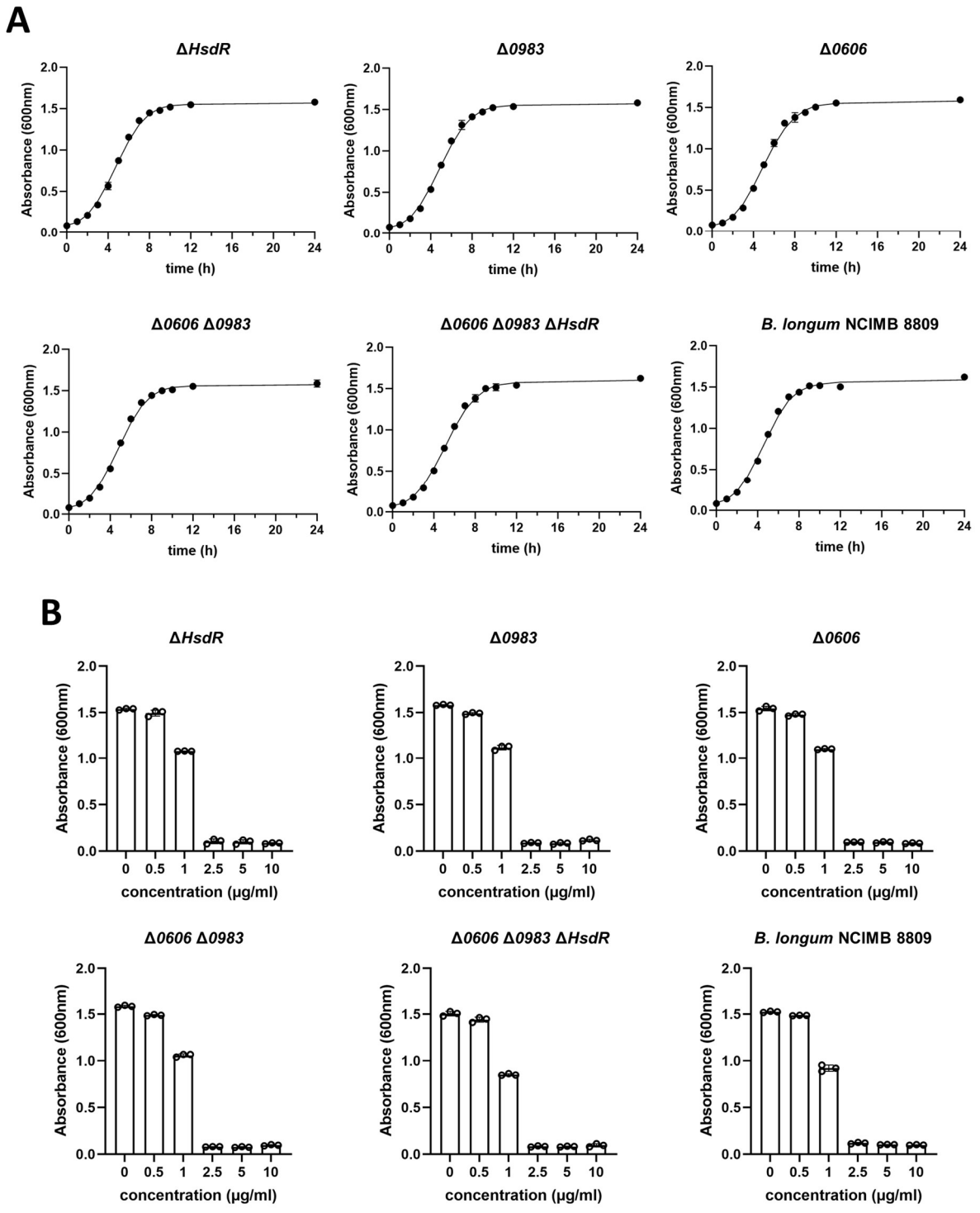
Supplementary Figure 6.1. Relative promoters strength across *Bifidobacterium* sp.. (A) Relative expression levels of indicated promoters in *B. adolescentis* DSM 20083 during exponential and stationary phase cultures. (B) Relative mcherry fluorescence levels in exponential and stationary phase cultures of *Bifidobacterium* sp. transformed with pMGC-mCherry. (n=3; error bars, s.e.m.; *P < 0.05, ***P < 0.001)

doi:10.6342/NTU202404344



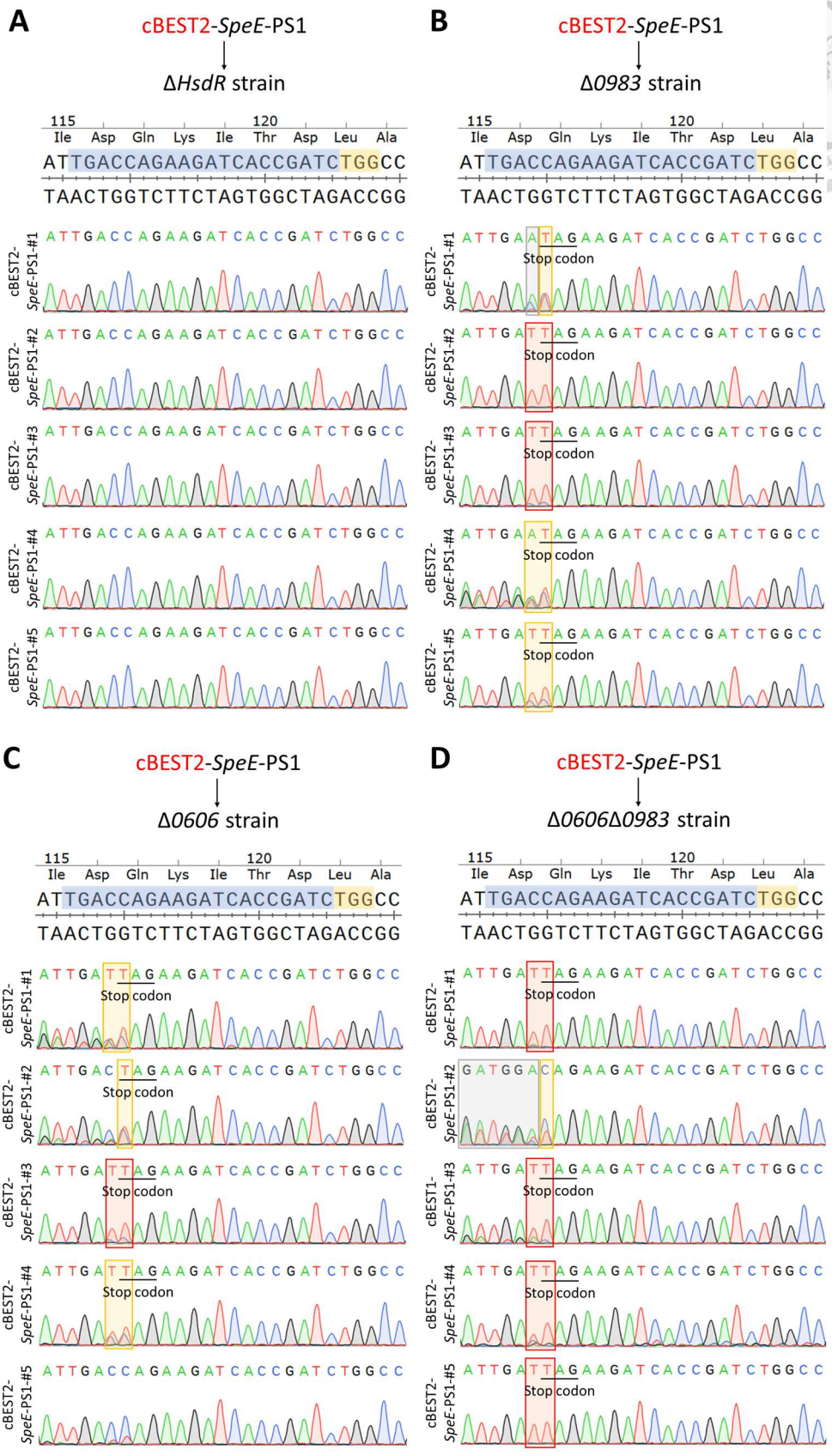
Supplementary Figure 6.2. Successful gene editing within four protospacers of the *SpeE* gene in *B. longum* NCIMB 8809. Sequencing results confirmed precise introduction of stop codon (black line) due to successful C-to-T edits (red frame) in four indicated protospacers. Protospacers 1 and 2 targeted the forward strand, while protospacers 3 and 4 targeted the reverse strand. The top row displays the wild type sequences for reference.

doi:10.6342/NTU202404344

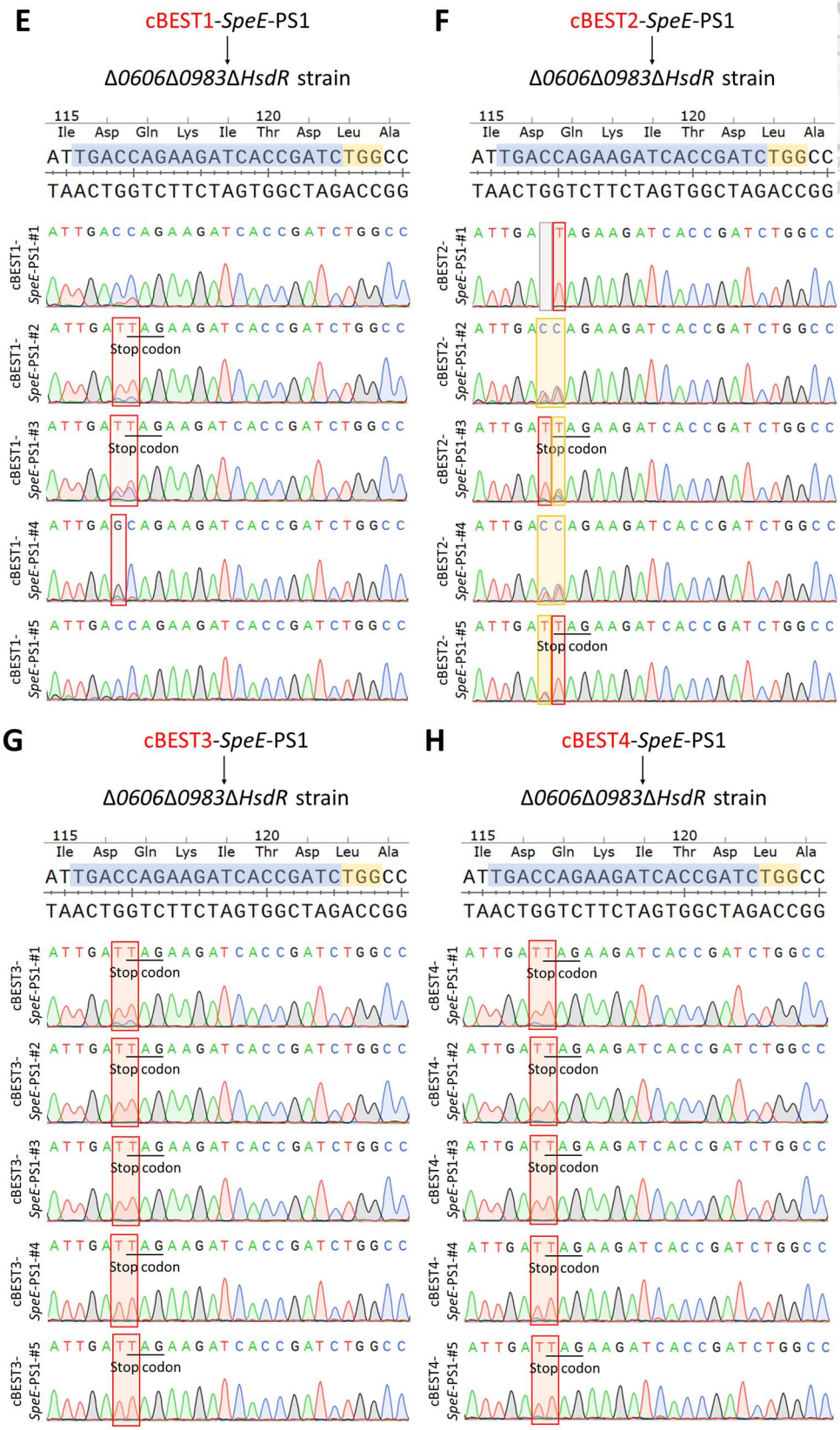


Supplementary Figure 6.3. Bacterial growth curves and chloramphenicol sensitivity test for all REase knockout strains derived from *B. longum* NCIMB 8809. (n=3, error bars, standard deviation)

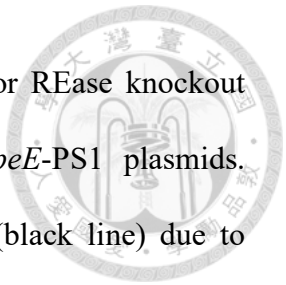
doi:10.6342/NTU202404344



doi:10.6342/NTU202404344

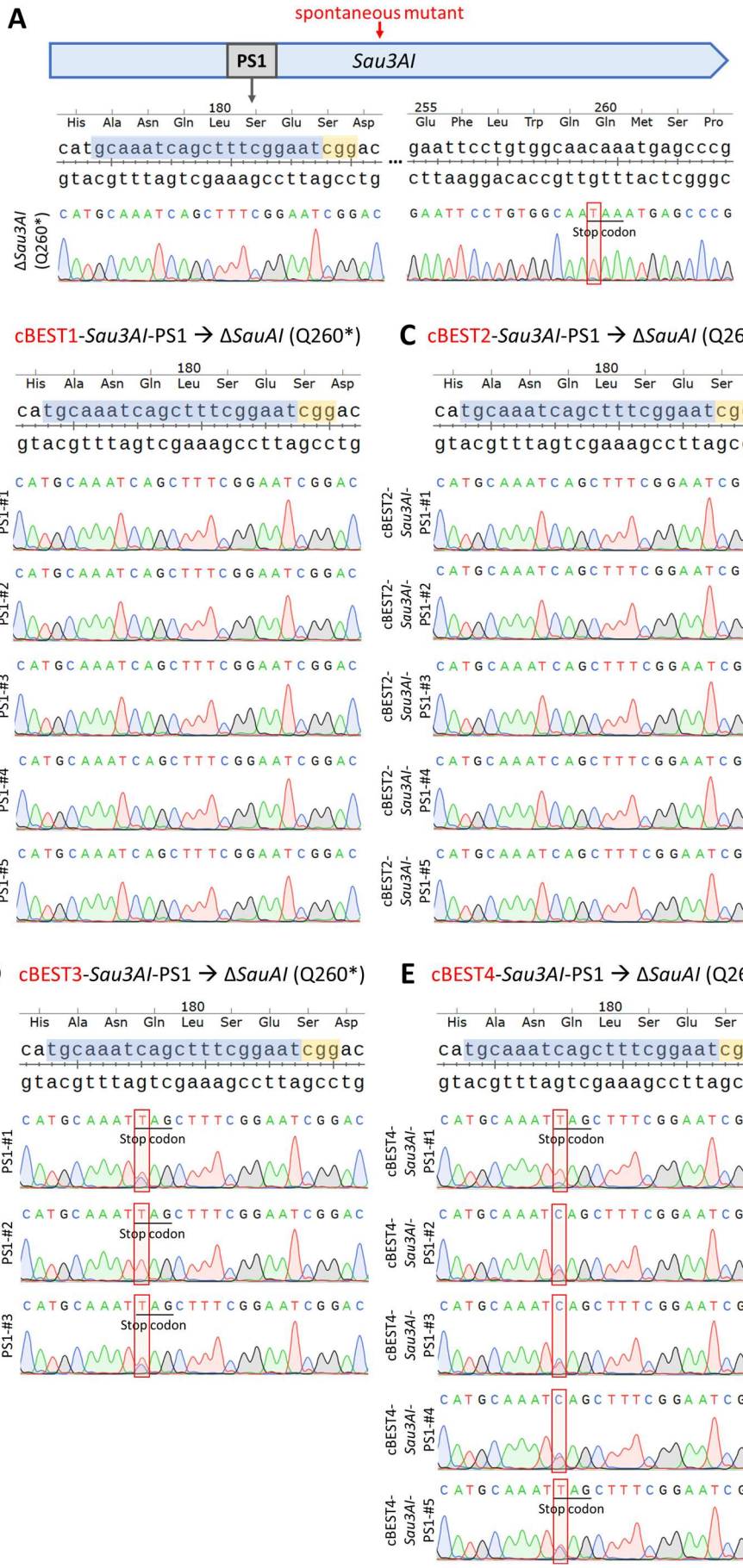


doi:10.6342/NTU202404344



Supplementary Figure 6.4. Evaluation of gene editing patterns for REase knockout strains derived from *B. longum* NCIMB 8809 using cBEST-*SpeE*-PS1 plasmids. Sequencing results confirmed precise introduction of stop codon (black line) due to successful C-to-T edits (red frame) in the *SpeE*-PS1 protospacer. The orange frames represent mixed edited sites and the grey frames represent undesired point mutation or deletion. The top row displays the wild type sequences for reference. (A to E) Sequencing details for transforming cBEST2-*SpeE*-PS1 plasmid across REase knockout strains. (E to H) Sequencing details for transforming four cBEST-*SpeE*-PS1 plasmids into the Δ*0606*Δ*0983*Δ*HsdR* strain.

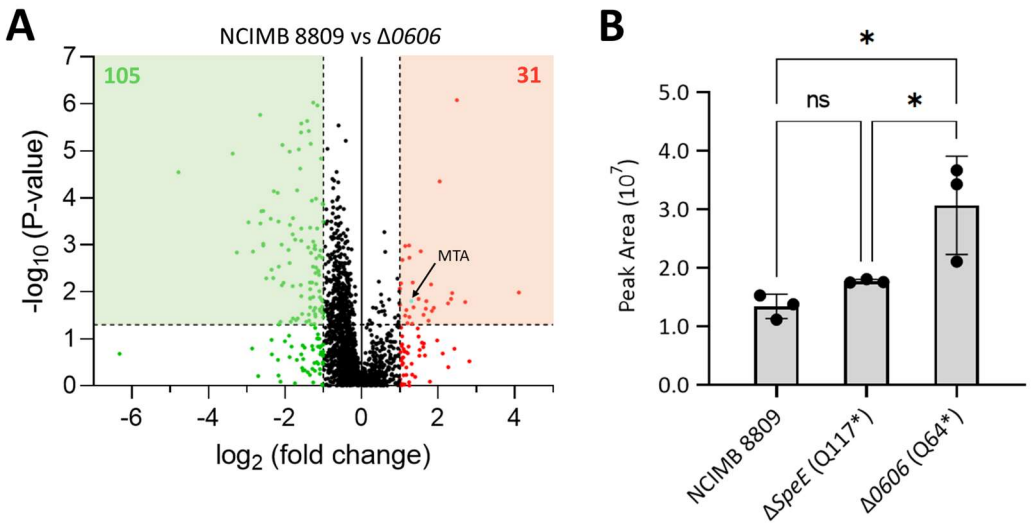
doi:10.6342/NTU202404344



doi:10.6342/NTU202404344

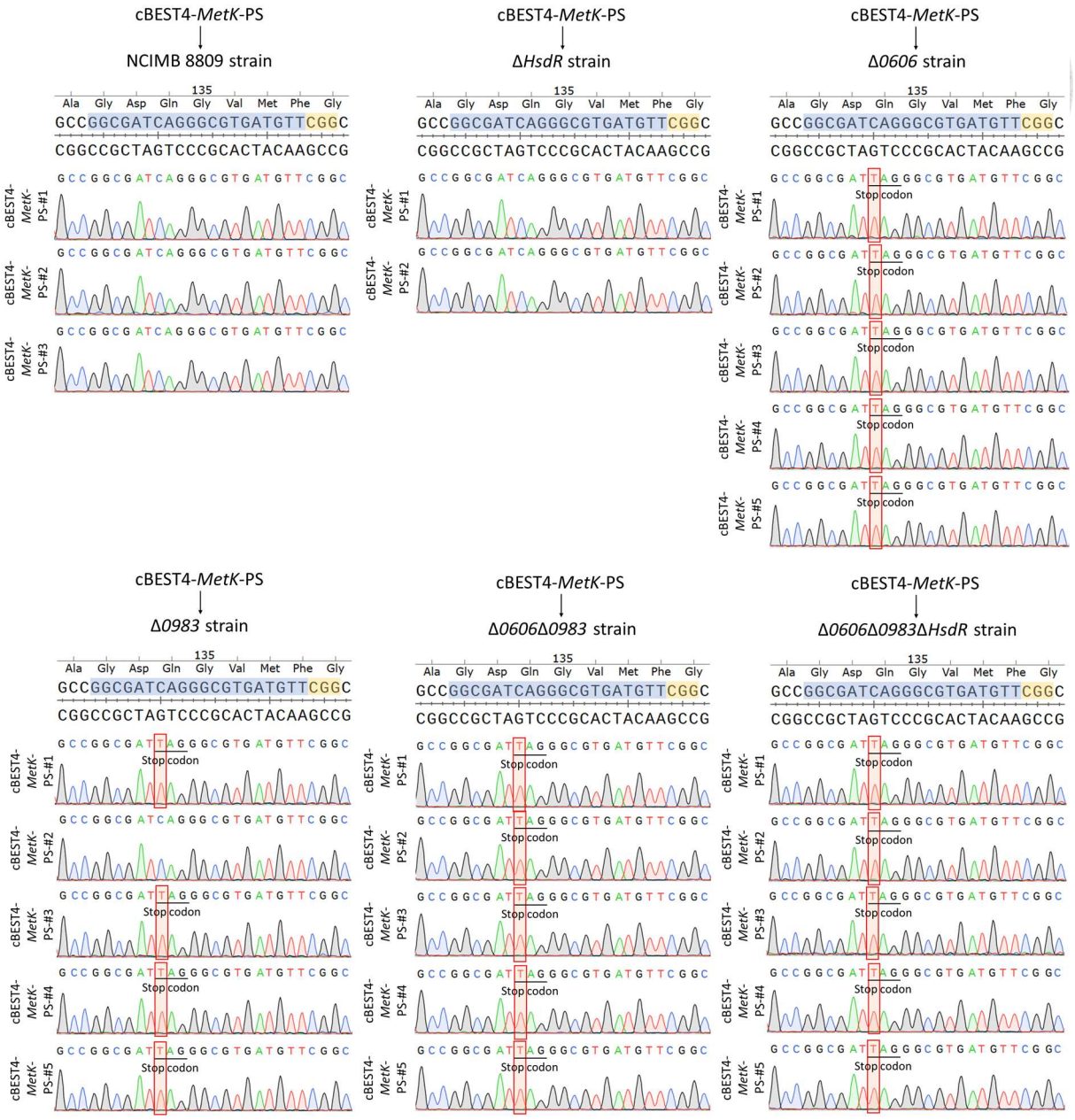
Supplementary Figure 6.5. Gene editing patterns using four cBEST-*Sau3AI*-PS1 plasmids in the Δ*Sau3AI* (Q260*) strain. (A) Sequencing result for the spontaneous mutant Δ*Sau3AI* (Q260*) strain. (B-E) Sequencing results confirmed precise introduction of stop codon (black line) due to successful C-to-T edits (red frame) in the *Sau3AI*-PS1 protospacer. The orange frames represent mixed edited sites. The top row displays the wild type sequences for reference. For (D), only three transformants were obtained on the agar plate for sequencing.

doi:10.6342/NTU202404344



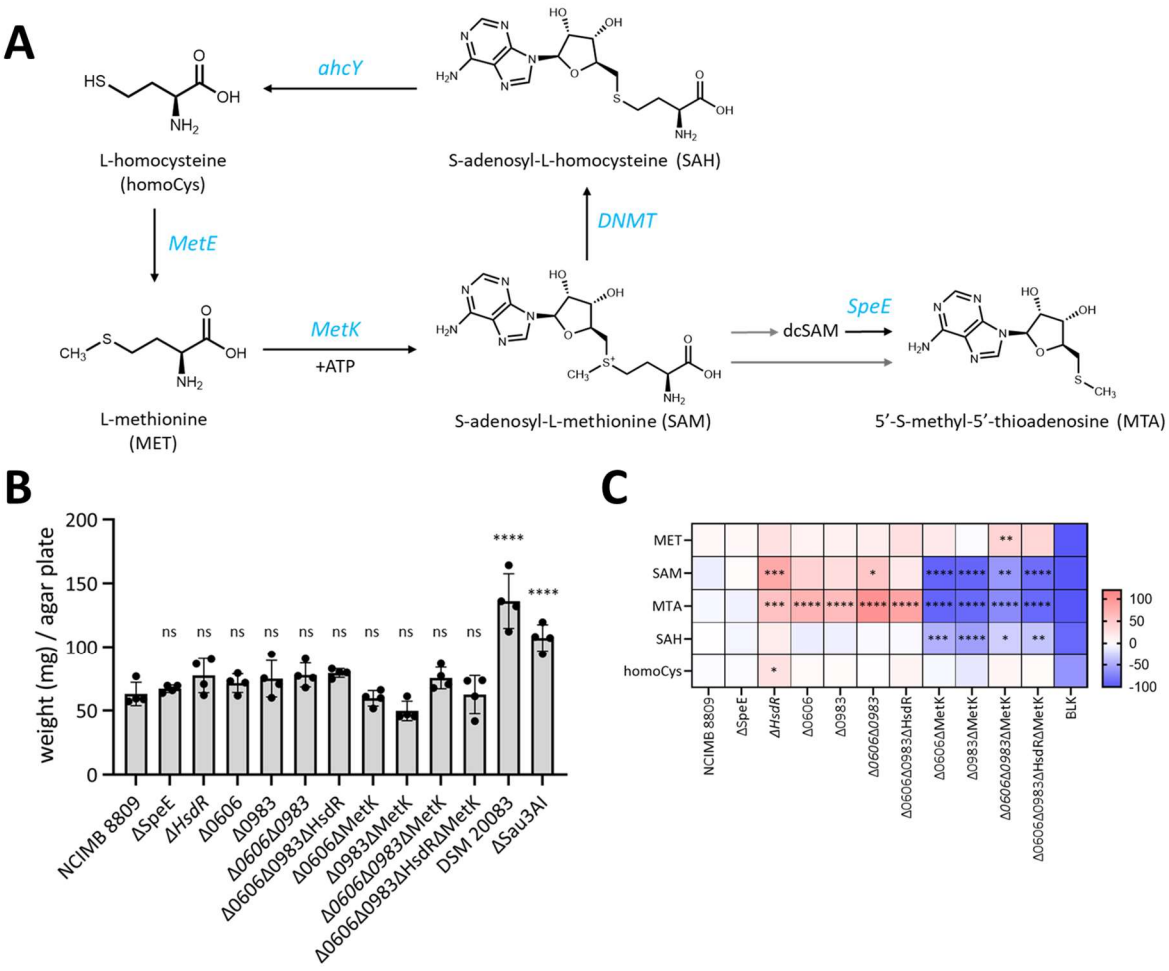
Supplementary Figure 6.6. Metabolomic analysis for engineered *B. longum* NCIMB 8809 strains. (A) Volcano plot analysis of bacterial metabolites from the Δ*0606* (Q64*) strain and *B. longum* NCIMB 8809. Data points are colored by fold change: red, up-regulated; green, down-regulated in the Δ*0606* (Q64*) strain. (B) LC-MS/MS analysis depicting peak areas for MTA in wild-type NCIMB 8809, the Δ*SpeE* (Q117*) strain and the Δ*0606* (Q64*) strain. (n=3, biological triplicates; error bars, standard deviation.; ns, not significant, *P < 0.05)

doi:10.6342/NTU202404344



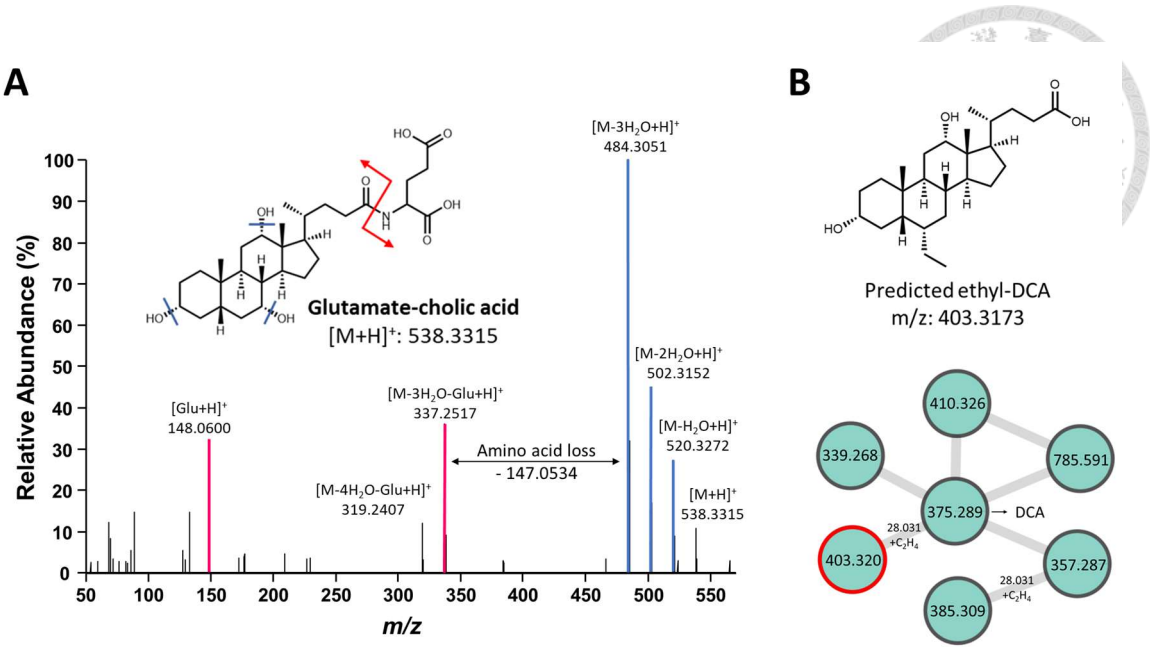
Supplementary Figure 6.7. Gene editing patterns using the cBEST4-*MetK*-PS plasmid in the multiple RM-disrupted strains. Sequencing results confirmed precise introduction of a stop codon (black line) due to successful C-to-T edits (red frame) in the *MetK*-PS protospacer.

doi:10.6342/NTU202404344



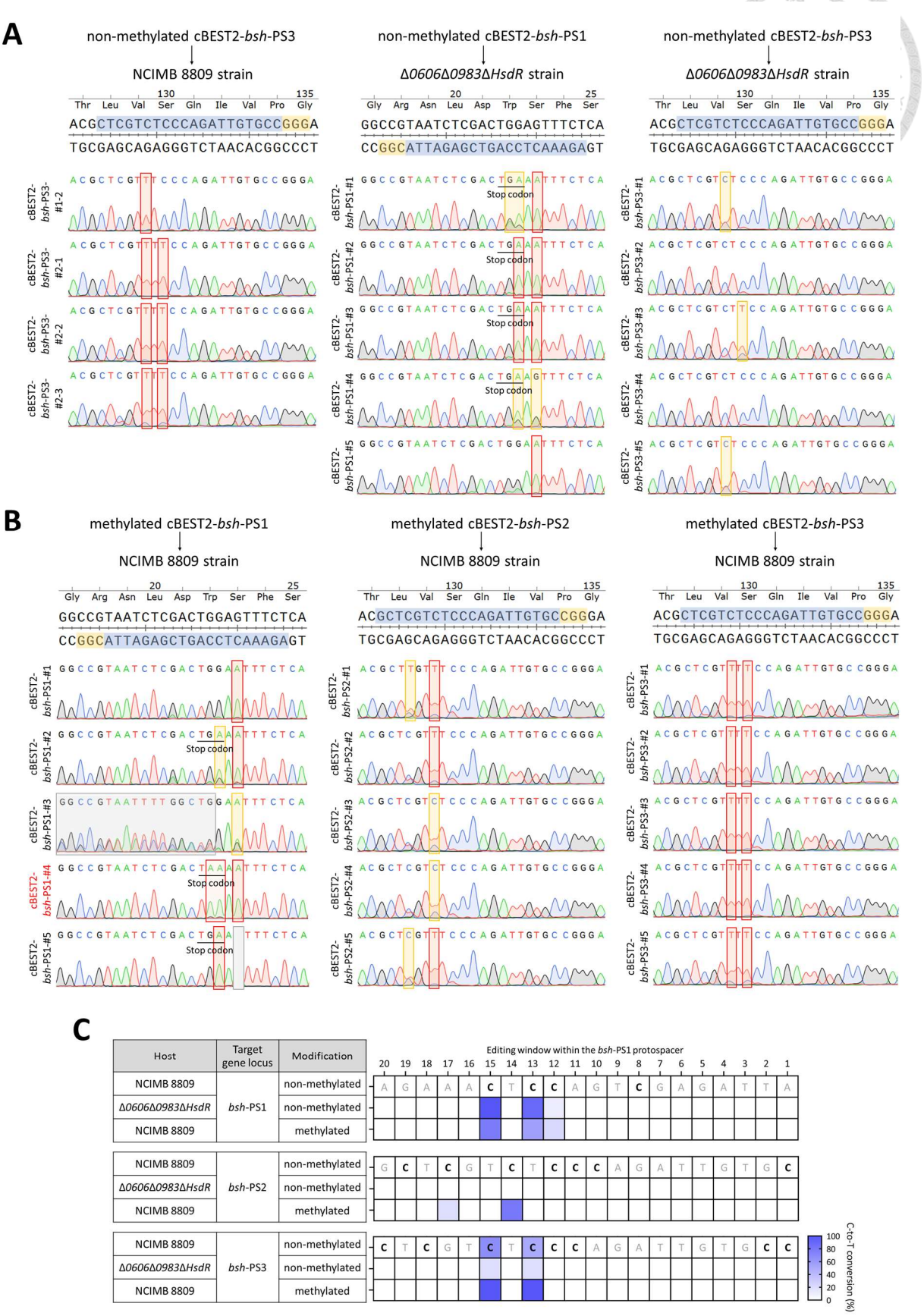
Supplementary Figure 6.8. Summary for methionine cycle pathway using targeted LC-MS/MS analysis. (A) Biosynthesis pathway of methionine cycle pathways. (B) Bacterial dry weight scraped from each plate after 18 h incubation. (C) Heat map for relative abundance of each metabolite across multiple engineered *B. longum* NCIMB 8809 strains. (n=4, biological replicates; ns, not significant, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001)

doi:10.6342/NTU202404344

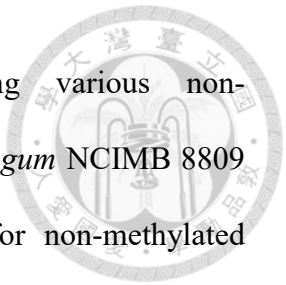


Supplementary Figure 6.9. Exploring microbiome-producing natural products using LC-MS/MS and GNPS[53]. (A) Annotation of MS/MS fragmentation patterns for glutamate-cholic acid. *B. longum* NCIMB 8809 was cultured on MRSC agar plate supplemented with 100 μM CA and then followed the metabolomic analysis (2-10. Metabolomics Using LC-QE Analysis for Bacterial Plate Cultures). (B) Using molecular network (GNPS) to explore DCA-modified metabolites. *B. longum* NCIMB 8809 was cultured on MRSC agar plate supplemented with 100 μM DCA and then followed the metabolomic analysis (2-10. Metabolomics Using LC-QE Analysis for Bacterial Plate Cultures).

doi:10.6342/NTU202404344

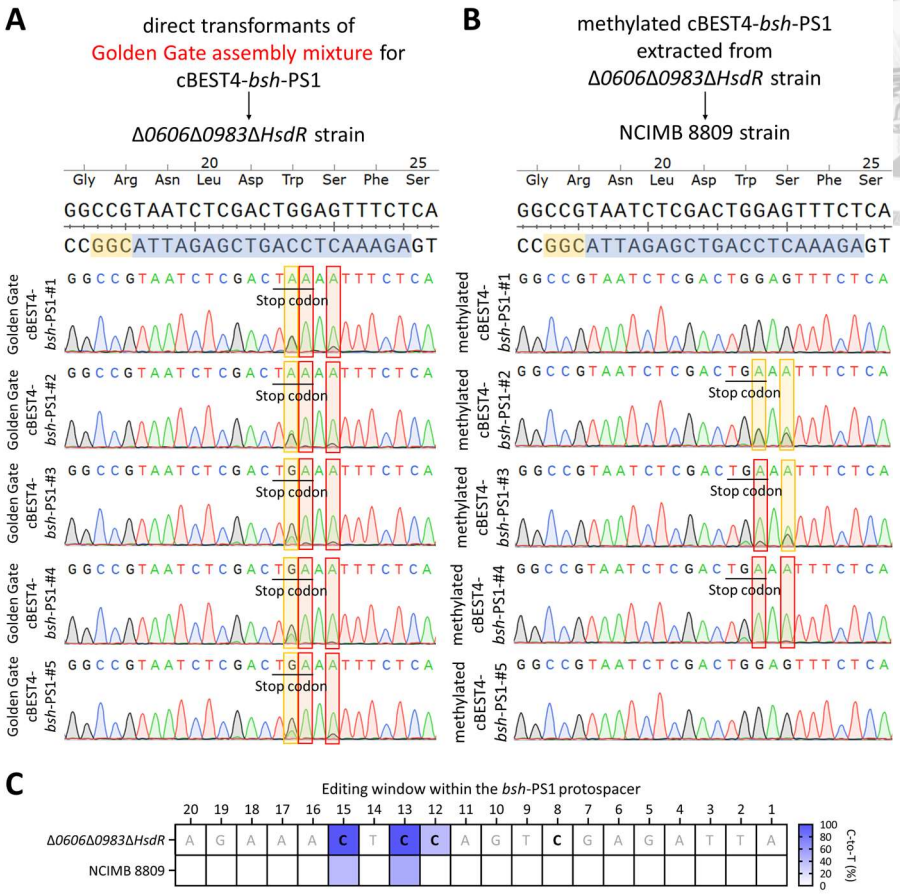


doi:10.6342/NTU202404344



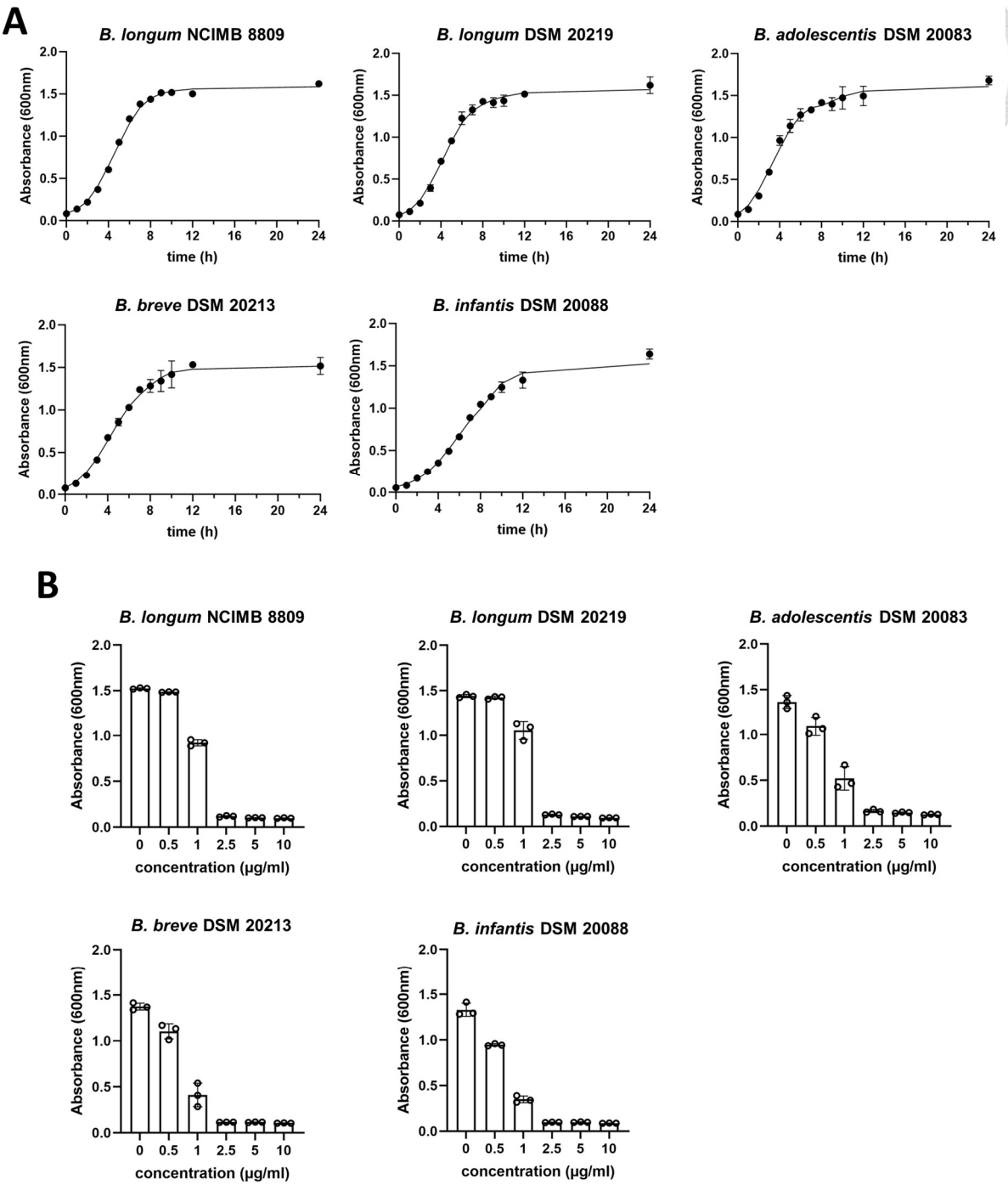
Supplementary Figure 6.10. Gene editing patterns using various non-methylated/methylated cBEST2-*bsh*-PS plasmids in wild type *B. longum* NCIMB 8809 and the Δ*0606*Δ*0983*Δ*HsdR* strain. (A) Gene editing patterns for non-methylated cBEST2-*bsh*-PS plasmids into NCIMB 8809 or the Δ*0606*Δ*0983*Δ*HsdR* strain. (B) Gene editing patterns for methylated cBEST2-*bsh*-PS plasmids into NCIMB 8809. Sequencing results confirmed introduction of stop codon (black line) due to successful C-to-T edits (red frame) in three indicated protospacers. The orange frames represent mixed edited sites and the grey frames represent undesired point mutation or deletion. (C) Editing windows for indicated conditions.

doi:10.6342/NTU202404344



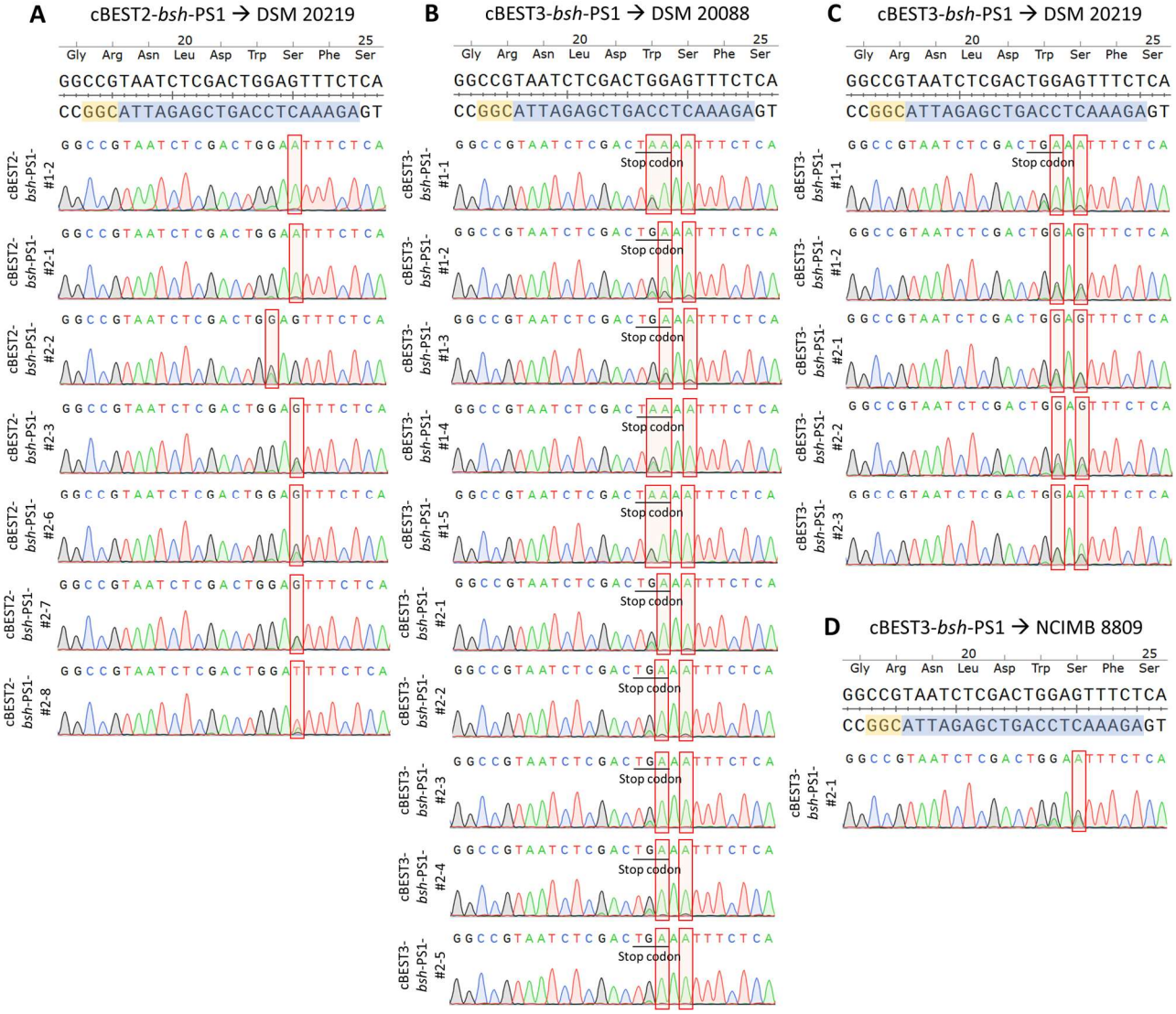
Supplementary Figure 6.11. Gene editing patterns using the two methods in Figure 3.13. Sequencing results confirmed introduction of stop codon (black line) due to successful C-to-T edits (red frame) in three indicated protospacers. The orange frames represent mixed edited. (A) Gene editing patterns of Golden Gate assembly mixture for constructing cBEST4-*bsh*-PS1 plasmid through the Δ*0606*Δ*0983*Δ*HsdR* strain. (B) Gene editing patterns for methylated cBEST4-*bsh*-PS1 plasmid transforming into *B. longum* NCIMB 8809. (C) Editing window for the two methods.

doi:10.6342/NTU202404344

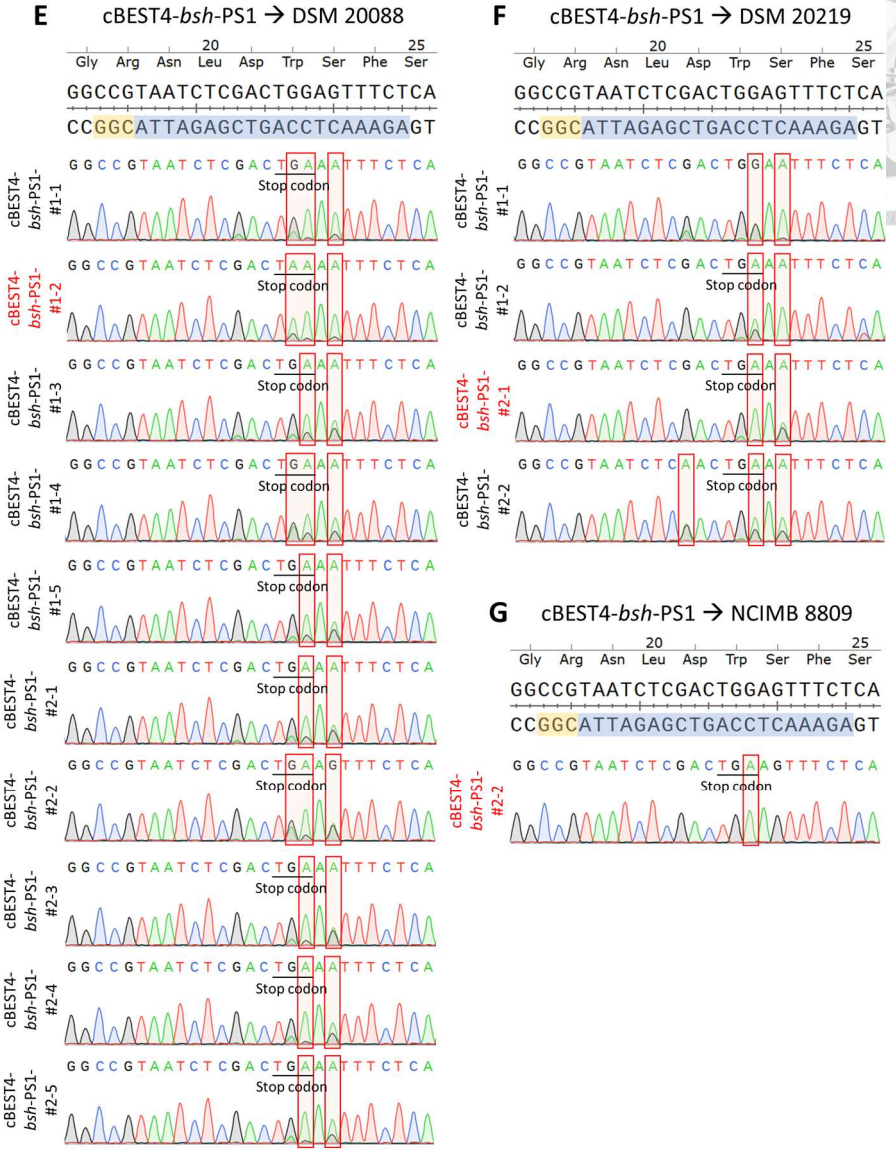


Supplementary Figure 6.12. Bacterial growth curves and chloramphenicol sensitivity test for multiple *Bifidobacterium* species and strains. (n=3, error bars, standard deviation)

doi:10.6342/NTU202404344



doi:10.6342/NTU202404344



Supplementary Figure 6.13. Gene editing patterns for various *Bifidobacterium* strains using three cBEST-*bsh*-PS1 plasmids in Figure 3.14B. Sequencing results confirmed precise introduction of stop codon (black line) due to successful C-to-T edits (red frame). The orange frames represent mixed edited sites. Gene editing patterns for (A) cBEST2-*bsh*-PS1 plasmid transforming into DSM 20219, (B to D) cBEST3-*bsh*-PS1 plasmid transforming into three strains and (E to G) cBEST4-*bsh*-PS1 plasmid transforming into three strains. Strains using scarlet letters were used for functional studies after plasmid curing.

doi:10.6342/NTU202404344