Synthetic and Systems Biotechnology 9 (2024) 577–585

Contents lists available at ScienceDirect

Synthetic and Systems Biotechnology

journal homepage: www.keaipublishing.com/en/journals/synthetic-and-systems-biotechnology

Original Research Article

Facilitating stable gene integration expression and copy number amplification in *Bacillus subtilis* through a reversible homologous recombination switch

Haoyu Guo [a,b,c,d], Rongzhen Tian [a,b,c,d], Yaokang Wu [a,b,c,d], Xueqin Lv [a,b,c,d], Jianghua Li [a,b], Long Liu [a,b], Guocheng Du [a,b], Jian Chen [a,b,c,d], Yanfeng Liu [a,b,c,d,*]

[a] Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi, 214122, China
[b] Science Center for Future Foods, Jiangnan University, Wuxi, 214122, China
[c] Engineering Research Center of Ministry of Education on Food Synthetic Biotechnology, Jiangnan University, Wuxi, 214122, China
[d] Jiangsu Province Engineering Research Center of Food Synthetic Biotechnology, Jiangnan University, Wuxi, 214122, China

ARTICLE INFO

Keywords:
Bacillus subtilis
Stabilized gene amplification system
Homologous recombination
recA

ABSTRACT

Strengthening the expression level of integrated genes on the genome is crucial for consistently expressing key enzymes in microbial cell factories for efficient bioproduction in synthetic biology. In comparison to plasmid-based multi-copy expression, the utilization of chromosomal multi-copy genes offers increased stability of expression level, diminishes the metabolic burden on host cells, and enhances overall genetic stability. In this study, we developed the "**BacAmp**", a stabilized gene integration expression and copy number amplification system for high-level expression in *Bacillus subtilis*, which was achieved by employing a combination of repressor and non-natural amino acids (ncAA)-dependent expression system to create a reversible switch to control the key gene *recA* for homologous recombination. When the reversible switch is turned on, genome editing and gene amplification can be achieved. Subsequently, the reversible switch was turned off therefore stabilizing the gene copy number. The stabilized gene amplification system marked by green fluorescent protein, achieved a 3-fold increase in gene expression by gene amplification and maintained the average gene copy number at 10 after 110 generations. When we implemented the gene amplification system for the regulation of *N*-acetylneuraminic acid (NeuAc) synthesis, the copy number of the critical gene increased to an average of 7.7, which yielded a 1.3-fold NeuAc titer. Our research provides a new avenue for gene expression in synthetic biology and can be applied in metabolic engineering in *B. subtilis*.

1. Introduction

Synthetic biology achieves the biosynthesis of target products by reallocating intracellular resources, fundamentally depending on gene expression. To produce proteins, enzymes, or metabolites at an economically viable scale, it's imperative to ensure that these genes are expressed at sufficiently high level [1–3]. The utilization of plasmids, which generally contain multiple copies, facilitates the achievement of elevated levels of target gene expression. However, plasmid-based systems suffer from segregational instability, where the plasmids may not be evenly distributed to daughter cells during cell division, leading to some cells without the plasmid. Structural instability including mutations or rearrangements in the plasmid can lead to incorrect or inefficient protein production. Additionally, allele segregation issues arise when non-productive plasmids outcompete productive ones, leading to a population of cells that do not contribute to the desired product yield. These factors collectively result in reduced productivity and reliability, undermining the feasibility of large-scale production [4–7]. Therefore, integrating and expressing the target gene in the genome is preferred for industrial-scale fermentation. Generally, the primary method to enhance gene expression levels involves swapping genetic regulatory elements, such as promoters and ribosome binding sites. However,

Peer review under responsibility of KeAi Communications Co., Ltd.
* Corresponding author. Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi, 214122, China.
E-mail address: yanfengliu@jiangnan.edu.cn (Y. Liu).

https://doi.org/10.1016/j.synbio.2024.04.010
Received 23 January 2024; Received in revised form 9 April 2024; Accepted 9 April 2024
Available online 16 April 2024

single-copy expression strategies are usually limited, titrating the copy number of the gene is a complementary strategy to improve expression [7–10].

Bacillus subtilis, as a typical food safety industrial model strain, has been widely used in biosynthesis such as vitamin B2 and *N*-acetylglucosamine [11–16]. Currently, the methods for constructing multiple copies of genes in *B. subtilis* have been developed including 16S ribosomal DNA-Mediated Multi-Gene Integration, chemically inducible chromosomal gene amplification, CRISPR-associated transposases mediated multicopy chromosomal integration [17–20]. The chemically inducible chromosomal gene amplification has been also widely applied in other microorganisms. Under antibiotic stress or other selective pressures, microbes can rapidly adapt through RecA-mediated homologous recombination, which induces the formation of consecutive copies on the chromosome. In this process, RecA, as a principal member of the "strand exchange protein" enzyme family, plays a crucial role [21, 22]. The RecA protein catalyzes the pairing and strand exchange between homologous DNA molecules and is central to homologous recombination. Once a sufficient number of target genes are obtained, the recombinase gene *recA* is typically knocked out to prevent homologous recombination that leads to a reduction in gene copy number [7, 18]. However, knocking out *recA* results in the loss of DNA repair

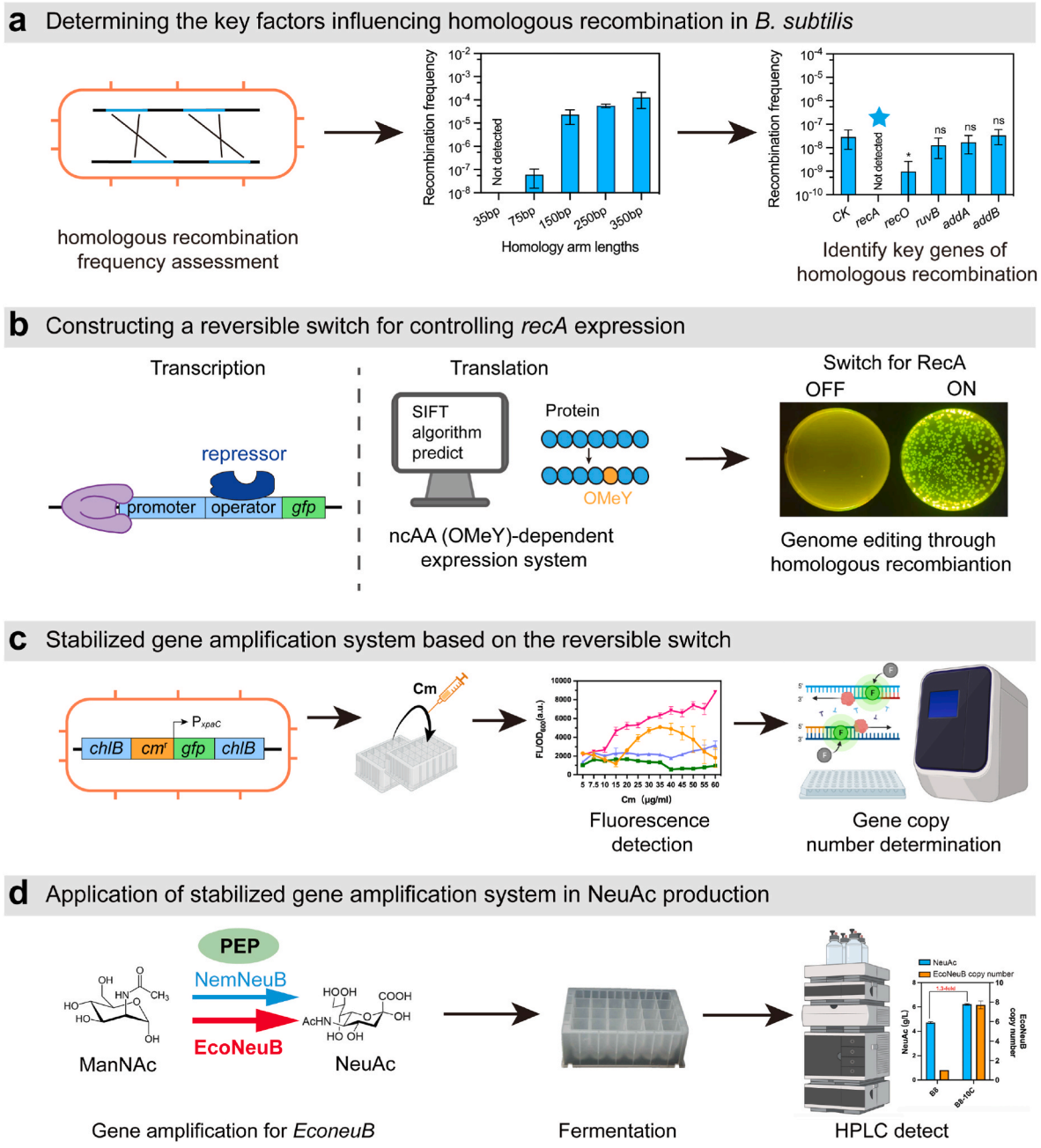

Fig. 1. Facilitating stable gene integration expression and copy number amplification in *B. subtilis* through a reversible homologous recombination switch. (a) A homologous recombination assessment strain was constructed to determine the influence of homology arm length and different genes. (b) A reversible switch with low leakage was constructed to control the homologous recombination in *B. subtilis* by inducing or repressing *recA*. (c) Combine the gene amplification with the reversible homologous recombination switch. When homologous recombination (RecA) is turned on, gene amplification can be followed to create multiple gen copies. After that, recombination was turned off to stabilize the multicopy gene that had been formed. (d) Apply the stabilized gene amplification system in NeuAc synthesis, and the multiple copies of EcoNeuB were made to improve sialic acid production.

capabilities and eliminates the possibility of homologous recombination-based genome editing. Without *recA*, techniques relying on homologous recombination for genome engineering become ineffective, significantly diminishing the ability to manipulate and improve the microbial genetic structure for enhanced functions or characteristics. This represents a major impediment to research and industrial applications that require continuous genetic improvement and adaptation. In *B. subtilis*, homologous recombination is a complex process involving the coordinated action of multiple genes. While the *recA* gene is a central player, several other genes, including *ruvAB*, *recO*, and *addAB*, also play crucial roles in influencing and facilitating homologous recombination [23–26].

In this study, a stable multicopy gene amplification system "**BacAmp**" has been built in *B. subtilis* by utilizing a reversible homologous recombination switch (Fig. 1). First, we constructed a test strain to assess the key factors influencing homologous recombination in *B. subtilis*: the length of the homology arms and the impact of specific genes, particularly *recA*. Additionally, we developed a reversible switch dependent on repressor and non-natural amino acids (ncAA)-dependent expression system to control the *recA* gene. Subsequently, the function of the stabilized gene amplification system was tested by green fluorescent protein, achieved a 3-fold increase in gene expression by gene amplification, and maintained the average gene copy number at 10 after 110 generations. Ultimately, our application of the gene amplification system to modulate *N*-acetylneuraminic acid production led to an increase in the *neuB* gene's copy number sourced from *Escherichia coli* to an average of 7.7, resulting in a 30 % rise in the NeuAc titer.

2. Materials and methods

2.1. Plasmids and strains

All plasmids were constructed by Gibson assembly with pHT01 as backbone in Top10. For genome editing in *B. subtills*, a homologous recombination box with 1000 bp homology arms was used, and the *Cre-lox* system was used to eliminate the antibiotic resistance gene. The following antibiotics were used for the selection: 50 μg/ml kanamycin, 50 μg/ml spectinomycin, 60 μg/ml zeocin, and 5 μg/ml chloramphenicol.

For the TAG integration selection, we used the SIFT algorithm to predict the residue tolerance of the proteins (http://sift.jcvi.org). The SIFT algorithm can anticipate amino acid substitutions with minimal impact on protein function codons by using sequence homology. For the GFP, we choose the sites closest to the translation start codon (ATG) to facilitate genome editing. For the RecA, we choose the 54th (TAT) amino acid position to TAG. This selection was made because the amino acid at this position happened to be tyrosine, and SIFT calculations indicated that mutations at this site had the least impact on the protein. This design aims to minimize the impact of non-natural amino acid insertion on *recA*.

All the strains and plasmids used in this study are listed in Supplementary Table 1. All sequences of promoters and gene expression cassettes used in this study are listed in the **Supplementary sequence**.

2.2. Strains cultivation

For cell culture, Luria-Bertani (LB) medium (10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl) was used unless noted otherwise.

For the fermentation process in this study, a 24-well plate was utilized, with conditions set at 37 °C and 220 rpm for a duration of 72 h. For NeuAc biosynthesis, Bacillus-Fermentation (BF) medium (12 g/L yeast extract, 6 g/L tryptone, 6 g/L $(NH4)_2SO_4$, 12.5 g/L $K_2HPO_4{\cdot}4H_2O$, 2.5 g/L KH_2PO_4, 3 g/L $MgSO_4{\cdot}7H_2O$, and 60 g/L glucose, pH 7.0, adjusted using NH_4OH) was used.

The working concentrations of different inducers used in this study are as follows: 30 g/L for gluconate, 0.5 mM for isopropyl-β-D-thiogalactoside (IPTG), 1 mM for O-methyl-L-tyrosine (OMeY), 10 mM for mannose, and 15 % for xylose.

2.3. Gene amplification protocol

To obtain the amplification of the construct, the colonies were inoculated into 1 mL of LB medium with the working concentration of chloramphenicol in the 24-well plate and incubated for 12 h at 37 °C, 220 rpm. Next, 1 μL of culture was inoculated into 1 mL of fresh LB media with twice the working concentration. Keep passaging until the desired gene copy number is obtained.

2.4. Gene copy number determination

Quantitative real-time PCR was used to determine the gene copy number. In the context of continuous passage culture, bacterial solutions were utilized as templates to ascertain the average copy number. To identify strains with a high copy number for NeuAc biosynthesis, single colonies were streaked from bacterial passaging solution. Subsequently, a single colony was selected as a template for analysis. First, we use the lysis buffer (code no. 9164, Takara) to dilute the template. The control template consisted of a fusion of the reference gene DNA fragment and the target gene. The qPCR experiments were performed using QuantStudio 3 Real-Time PCR Instruments. All qPCR primer sequences are provided in Supplementary Table 2.

2.5. Homologous recombination frequency analysis

The frequency of homologous recombination was detected through Fluctuation AnaLysis CalculatOR (FALCOR) [27]. First, the test strain was cultured in 1 mL of fresh LB medium for 12 h at 37 °C and 220 rpm. Then 1 μL of culture was inoculated into 1 mL of fresh LB media for 12 h at 37 °C and 220 rpm. For test strains with inducible control of gene, add the appropriate working concentration of the inducer to the 1 mL of fresh LB. Next, 100 μL of culture was spread onto a kanamycin-resistant agar plate. After culturing the plate for 24 h, the number of colonies on the plate was counted. The total colony in the culture is calculated based on OD_{600}, with a conversion factor of 1×10^8 colonies per OD1. Next, FALCOR was used to estimate frequency m, and the analysis method used was Ma–Sandri–Sarkar maximum likelihood method. The number of r was counted from the observed plates and the total number of viable cells N was calculated from the dilution factor. At least six biological replicates were performed for each strain.

2.6. Fluorescence detection

Cytation 3 Multi-Mode Reader (BIOTEK) was used to detect the fluorescence intensities, and the excitation and emission wavelengths were 488 nm and 523 nm, respectively. Cell mass was assessed by OD_{600}. Gen5 CHS 2.06 was used to collect the biomass and fluorescence intensity data. When calculating the biological fluorescence intensity, both the fluorescence and biomass values need to be subtracted by the corresponding blank controls.

2.7. Liner fragment transformation for genome editing capability

To test the ability of bacteria to perform genomic editing through homologous recombination, cells were first cultured in 1 mL of LB media for 8 h. Next, 200 μL culture was inoculated into fresh 800 μL LB medium, to which 3 % xylose and the is added. After 2 h of induction, 2500 ng linear DNA fragments were added to the culture and transformed for another 2 h. Finally, the bacteria were spread on antibiotic-resistant plates. In the whole process, the corresponding inducer for *recA* expression needs to be added to the medium.

2.8. NeuAc determination

For the determination of NeuAc and ManNAc, High-performance liquid chromatography (HPLC; Agilent 1200 Series) was used with an ultraviolet absorption detector (210 nm) and an Aminex HPX-87H column (300 × 7.8 mm; Bio-Rad). The mobile phase consisted of 10 mM H_2SO_4 with a flow rate of 0.5 mL/min at 40 °C during the analysis. An Agilent OpenLAB Control Panel was used to collect the HPLC data.

2.9. Fed-batch fermentation in 3-L fermenter

The fermentation medium consists of 24 g/L yeast extract, 24 g/L tryptone, 8 g/L $(NH_4)_2SO_4$, 13 g/L $K_2HPO_4{\cdot}4H_2O$, 3 g/L KH_2PO_4, 3.5 g/L $MgSO_4{\cdot}7H_2O$, 2 g/L NaCl, 12 g/L urea, and 30 g/L initial glucose. After streaking the plate, incubate at 37 °C for 12 h. Then, pick a single colony to inoculate into 2 mL of LB medium and culture at 37 °C, 220 rpm for 10 h. Following this, transfer 1 mL of the culture into 50 mL of BF medium and incubate at 37 °C, 220 rpm for 12 h to reach an OD_{600} of 15. First, colonies were picked from the plates and inoculated into 2 mL of LB medium and culture at 37 °C, 220 rpm for 10 h. Following this, transfer 1 mL of the culture into 50 mL of BF medium and incubate at 37 °C, 220 rpm for 12 h to reach an OD_{600} of 15. Next, 50 mL of the seed was inoculated into a 3-L fermenter with an initial 1.45 L of fermentation medium. Keep pH at 6.8 using 15 % NH_4OH. The temperature was maintained at 37 °C, and the aeration rate was maintained at 2.5 v.v.m. Set the stirring and supplement feeding linked to DO automatically at 20 %, using 700 g/L glucose as the carbon source supplement during fermentation. During the fermentation period, take samples to determine glucose concentration, OD_{600}, NeuAc production, and ManNAc production.

3. Results and discussion

3.1. Determining the key factors influencing homologous recombination in B. subtilis

We first developed a system to assess homologous recombination in *B. subtilis*. Initially, the kanamycin resistance gene, featuring a 29-base pair (bp) deletion, was integrated into the genome. Subsequently, the 29-base pair deletion was introduced at a different genomic site, flanked by homologous arms of varying lengths (35 bp to 350 bp). This allowed the strain to spontaneously obtain kanamycin resistance through homologous recombination (Fig. 2a). The frequency of homologous recombination was determined by comparing colony counts on kanamycin-resistant plates with the actual bacterial count in the culture [28,29]. Notably, when the length of the homologous arm was 35 bp, no colonies with kanamycin resistance were observed on the plates (Fig. 2b). This observation is consistent with prior reports suggesting that *B. subtilis* requires a minimum homologous arm length of 70 bp for effective homologous recombination [30]. However, a notable increase in recombination frequency was observed once homologous arms exceeded 150 bp. It is also noteworthy that while the probability of homologous recombination improves with arm lengths from 150 bp to 350 bp, the upward trend in recombination frequency appears to plateau (Fig. 2b).

Numerous studies have indicated that *recA* is a key gene involved in

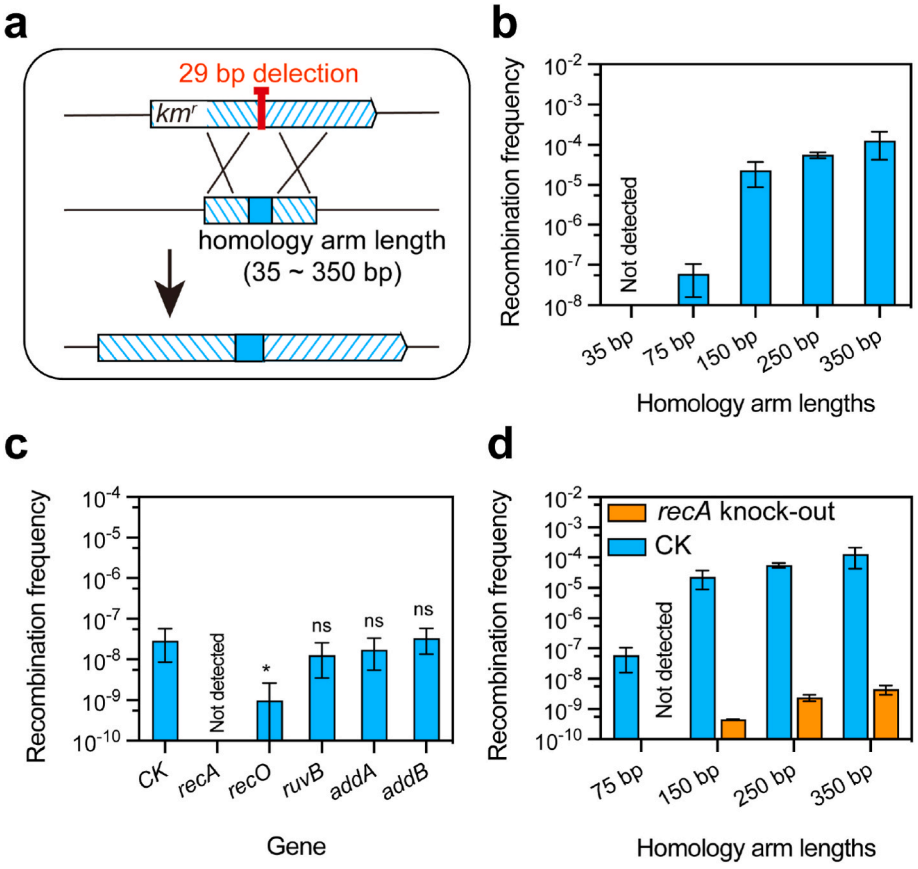

Fig. 2. Determining the key factors influencing homologous recombination in *B. subtilis* (a) The homologous recombination frequency assessment system enabled the obtain of the kanamycin resistance gene through homologous recombination. (b) Assessing the recombination frequency under different lengths of homology arms. (c) Determining the influence of knocking out different genes to the recombination frequencies with homology arms of 75 bp. Significance (P value) was evaluated by a two-sided *t*-test. *, $P < 0.05$; NS, not significant ($P \geq 0.05$). (d) Assessing the recombination frequency under different lengths of homologous arms after the knockout of *recA*.

homologous recombination in *B. subtilis*, serving as a fundamental member of the "strand exchange protein" enzyme family [31–35]. Additionally, other genes associated with homologous recombination include *recO*, *recF*, *recH*, *recP*, *recX*, *recG*, *scbC*, *scbD*, *ruvB*, *addA*, and *addB* [23–26,36]. We selected a subset of these genes and individually knocked them out in the test strain with a homology arm length of 75 bp. The assessment of recombination frequencies after each knockout revealed that *recA* deletion resulted in the absence of single colonies on kanamycin-resistant plates, indicating a drastic reduction in recombination frequency and a significant impact on the cell's ability to undergo homologous recombination (Fig. 2c). In contrast, the knockout of other genes showed minimal effects on recombination frequency. Furthermore, when *recA* was knocked out in strains integrated with 150–350 bp homology arms, the recombination frequency decreased by more than 10^4 (Fig. 2d).

3.2. Constructing a reversible switch for controlling recA expression

Previous chemically induced chromosomal adaptive evolution involved forming multicopy genes and then knocking out the key gene *recA* responsible for homologous recombination to establish stable gene copy numbers [7]. However, knocking out *recA* also renders the strain incapable of genomic editing through homologous recombination. This limitation poses a challenge in further genetic manipulations that rely on the homologous recombination mechanism for targeted genome modifications. Therefore, we aimed to design a reversible switch to regulate the expression of the *recA* gene. Turning off the switch maintains the stability of gene copy numbers, while turning it on retains the ability of the strain to undergo genome editing through homologous recombination. To ensure the rigorousness of the switch, we explored the combination of repressor-based transcription regulation system and non-canonical amino acids (ncAA)-dependent translation regulation system (Fig. 3a).

The repressor can bind to the operator and inhibit gene expression by blocking the attachment of RNA polymerase to the promoter, thereby preventing the transcription of the genes downstream of the operator [37]. The ncAA-dependent expression system is according to the method described by Tian et al. [38]. The aminoacyl-tRNA synthetase (aaRS) and mutate tRNA for ncAA O-methyl-L-tyrosine (OMeY) insertions were incorporated into the *B. subtilis* and an amber stop codon (TAG) was incorporated into the gene expression framework, therefore the expression of the gene requires OMeY (Fig. 3a). Green Fluorescent Protein (GFP) is utilized as a reporter to measure the activation strength of reversible switch when switched on and to assess their leakage levels when turned off. Here, the gene expression strengths (ON state) and leakage levels (OFF state) of four promoters and four repressor candidates, as well as a ncAA-dependent system, were evaluated. The results show that the promoters P_{graC} and P_{srfAA}, when paired with respective repressors, resulted in higher gene expression strengths, albeit with elevated levels of gene leakage. Conversely, the expression strengths for the mannose and xylose inducible promoters P_{man} and P_{xylA} showed relatively subdued expression levels. To reduce gene leakage while preserving adequate expression strength, we combined the P_{srfAA} promoter with the GntR repressor within a ncAA-dependent system, establishing a dual reversible switch. In OFF state gene expression leakage is minimal (only 1.2 %), while there is a substantial increase in gene expression (86.5-fold) in ON state (Fig. 3b). Subsequently, the promoter of GFP is replaced with an IPTG (Isopropyl β-D-1-thiogalactopyranoside) inducible promoter, which is combined with ncAA-dependent system to construct the reversible switch (Fig. 3c). We also utilize the lac operon from *Escherichia coli* combined with ncAA-dependent expression system. Remarkably, upon activation with IPTG and OMeY, the gene expression intensity reached an 85-fold increase (Fig. 3c).

Next, we applied the reversible switch to the *recA* gene in *B. subtilis*. The results from the fragment transformation indicate that the addition of the inducer allows for genome editing through homologous recombination while the absence of induction leads to a lack of RecA protein, preventing genome editing (Fig. 3d). To evaluate the genome editing efficiency through homologous recombination upon activation of the reversible switch, we conducted a series of transformation experiments using linear donor DNA fragments. The single-stranded DNA-annealing protein from *Collinsella stercoris* phage (CspRecT) has been proven to significantly enhance the efficiency of single-locus genome editing in *E. coli* by promoting effective pairing and annealing between oligonucleotides and target DNA at the replication fork [39]. To explore whether the combined use of oligo mediated recombineering can enhance the efficiency of genome integration, we also incorporate CspRecT into *B. subtilis*. Additionally, we compare their genome editing abilities under different inducible reversible switches. The results demonstrated that using LacI combined with a non-canonical amino acid-dependent expression system as a reversible switch to control RecA expression achieves the highest efficiency of genome editing through homologous recombination (Fig. 3e). Subsequently, in the HR-250 strain (the homologous recombination frequency test strain, 250 bp homology arm), we employed different repressors combined with the ncAA-dependent expression system to the *recA* gene. This allowed us to assess the homologous recombination frequency under both activation and deactivation of the reversible switch, with the BSXC strain and its RecA knockout counterpart serving as controls. It was observed that all three reversible switches decreased recombination frequencies upon activation. Remarkably, the P_{grac} promoter, combined with the LacI repressor and ncAA-dependent expression system, exhibited a lower recombination frequency in the OFF state, outperforming that in the *recA* knockout strain (Fig. 3f).

3.3. Stabilized gene amplification system based on the reversible switch

The chemically inducible chromosomal gene amplification is driven by homologous recombination. Under the pressure of antibiotics, when the homologous regions on the chromosome crossover between the leading and lagging strand, one of the two daughter cells obtains a gene copy and has a growth advantage, and the other one will be gradually eliminated [7,18,40,41](Fig. 4a). We first tested the feasibility of chemically induced chromosomal adaptive evolution in wild-type *B. subtilis*, by taking the homology arm sequence derived from *Synechocystis* sp. PCC6803 [7]. A chloramphenicol resistance gene and a GFP gene between two homology arms of 900 bp each were inserted into three different genomic sites. The cells were cultured in LB media with increasing concentrations of chloramphenicol, therefore cells with multiple copies of the chloramphenicol resistance gene and GFP gene had a greater growth advantage. Variations were found in gene copy numbers and fluorescence intensity at different genomic sites (Fig. 4b). As chloramphenicol concentration rises, there's an initial increase in GFP copy number and fluorescence at site-A, which then decreases beyond 40 μg/mL (Fig. 4b). At site-C, despite the GFP copy number increased, the fluorescence intensity remained low (Fig. 4b). At the site-B, an increase in chloramphenicol concentration was associated with a continuous rise in GFP gene copy number, reaching an average peak of 9.3. Notably, at a chloramphenicol concentration tenfold higher than the baseline, the fluorescence intensity showed a 3.8-fold increase compared to the initial level. Site B shows the highest gene amplification copy number, which may be related to its proximity to the replication origin, resulting in longer unzipping time during DNA replication. Site A, farther from the replication origin, has a relatively lower gene copy number. For site C, despite having a higher gene copy number but consistently low fluorescence intensity, possibly indicating a cellular burden beyond a certain copy number threshold, leading to resource allocation to cell growth over protein expression. Subsequently, we optimized the transfer time during the amplification process, with results indicating higher bioluminescence levels at a 12-h transfer time (Fig. 4c).

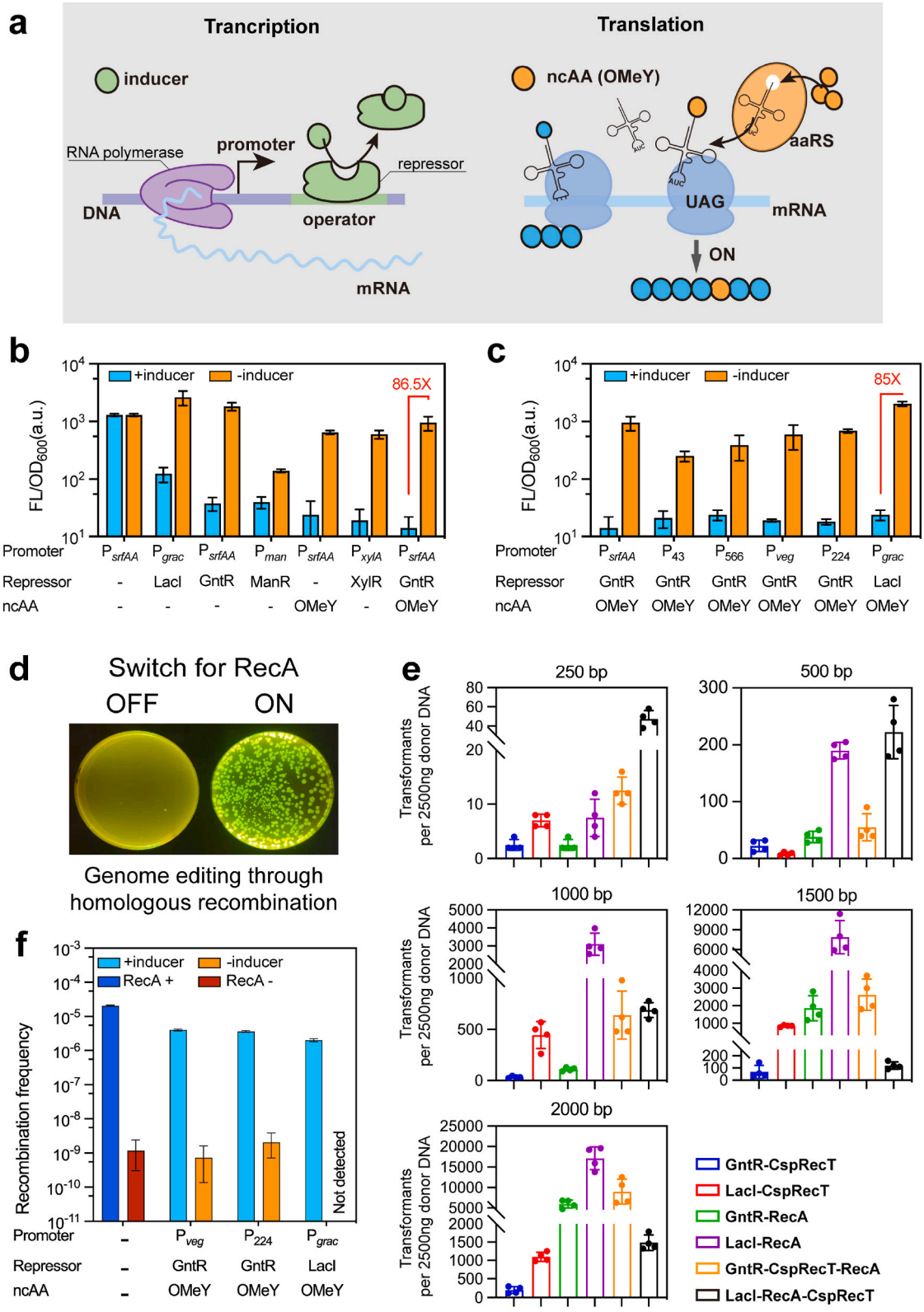

Fig. 3. Constructing a reversible switch for controlling recA expression for genome editing (a) Incorporating repressor and ncAA-dependent expression system to construct reversible switch with minimal leakage. (b) Assessing the fluorescence expression strength during activation and the leakage level during deactivation under different reversible switches. (c) Assessing the fluorescence expression strength during activation and the leakage level while combining the ncAA-dependent expression system and different repressors. (d) Transformants are obtained through genome editing via homologous recombination when the inducer is present (right), while homologous recombination and genome editing are ineffective without the inducer and transformants are not produced (left). (e) Evaluating the genome editing level of strains when homologous recombination is activated. The homologous arm of the transformed fragment is between 250 and 2000 bp. (f) Testing the recombination frequency of different reversible switches when induction expression is turned off in strains used for homologous recombination frequency testing (homologous arm length = 250 bp).

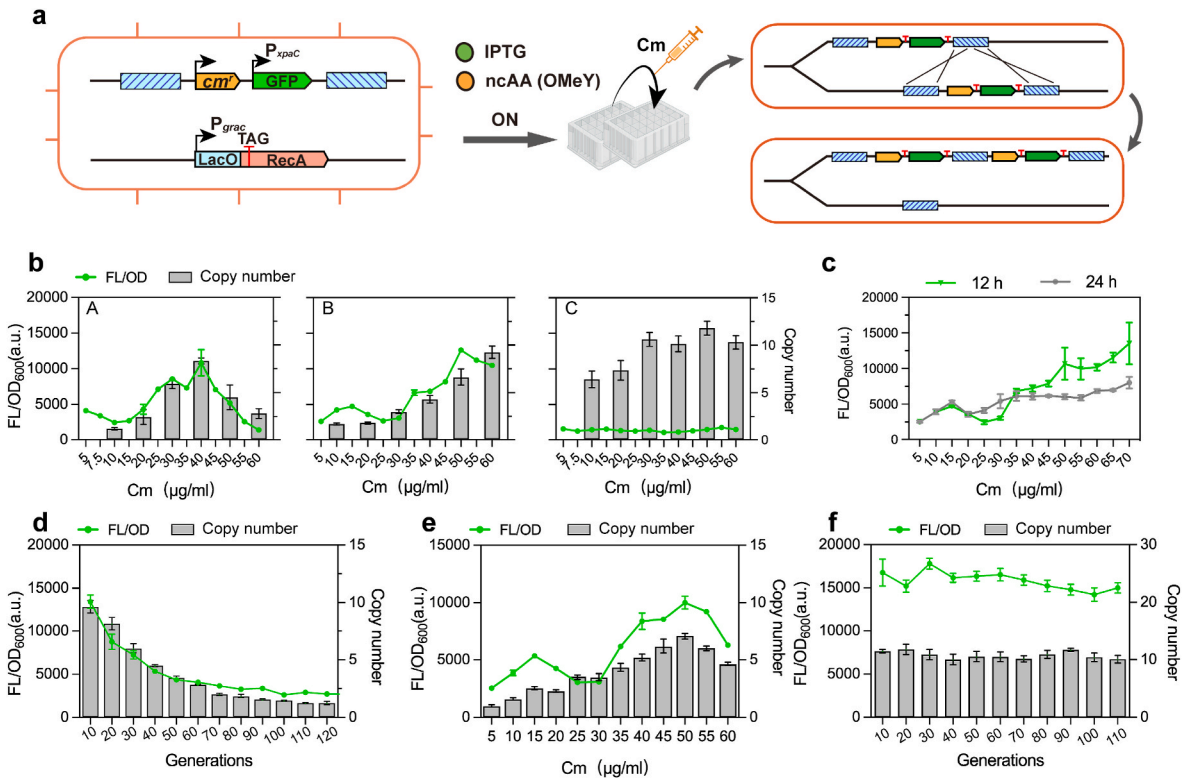

Fig. 4. Constructing stabilized gene integration expression and copy number amplification system. (a) Constructing a system capable of forming stable multi-copy gene. (b) Amplifying the *gfp* gene at three selected sites: Site A (*perR*), Site B (*mscL*) and siteC (*trpP*) in the genome of the wild-type strain (BSXC) to create multi-copy genes. (c) Optimizing the transfer time during the passaging process. (d) The wild-type *B. subtilis*, after generating multi-copy genes, was passaged continuously in an antibiotic-free medium, resulting in a decrease in both gene copy number and fluorescence intensity. (e) Amplifying *gfp* gene at site B in the strain whose *recA* was regulated by LacI repressor and ncAA-dependent expression. (f) Turning off the reversible switch of *recA* to stabilize the multiple copies of the gene.

To evaluate the genomic stability of *B. subtilis* endowed with multiple gene copies, the strain was passaged several generations without chloramphenicol, and a trend was observed that both the number of gene copies and the bioluminescence decreased (Fig. 4d). This suggests that without the selective pressure of antibiotics, the additional gene copies are unstable and may be lost, indicating a tendency toward genomic unstable with more gene copies (Fig. 4d).

Since different genomic sites have already been experimented with in the wild-type *B. subtilis*, we selected the gene loci with the best effects to implement a stable gene amplification multicopy system. Gene amplification was induced in the recombinant regulated strain by adding the inducer to trigger *recA* gene expression. During serial passage with progressively increasing antibiotic concentrations, both the GFP gene copy number and bioluminescence increased. However, upon reaching tenfold the initial antibiotic concentration, a decrease in both gene copy number and bioluminescence was observed (Fig. 4e). Next, to assess genomic multicopy genes in strains regulated for homologous recombination, the reversible switch was employed to the *recA* of the wild-type *B. subtilis* strains that had multicopy gene through amplification (BSXC-GFPb). The strains were then passaged in an antibiotic-free medium to observe the stability of the multicopy gene without selection pressure. After 110 generations, the average gene copy number remained at 10, and the bioluminescence was consistently stable, which rules out the possibility of instability of the target genes generated by the leakage of RecA (Fig. 4f). After securing stable gene copy numbers, it was observed that the strains possess the capacity for further gene editing upon the active of RecA. This indicates a potential for dynamic control of genome editing processes, contingent on the inducible reactivation. Notably, the induction of genome editing should be conducted under antibiotic pressure to prevent a reduction in gene copy number due to recombination events that occur with the formation of multicopy genes.

3.4. Application of stabilized gene amplification system in N-acetylneuraminic acid production

To test the application of **BacAmp** in synthetic biology, we apply this to the biosynthesis of *N*-acetylneuraminic acid (NeuAc) in *B. subtilis*. As one of the most important sialic acid compounds, NeuAc is widely used as an intermediate in nutritional chemistry and medicine. It plays an important role in maintaining the health of human brain function in the elderly, promoting the brain development of infants, enhancing the immune capacity of the body, and so on. In addition, it can also be used as an intermediate in anti-cancer drugs [42,43].

Currently, the biosynthesis of NeuAc in *B. subtilis* primarily involves three pathways [42]. For the NeuC pathway chosen in this study, the final reaction catalyzed by the NeuB (*N*-acetylneuraminic acid synthase) is the rate-limiting step (Fig. 5a). Previous research has demonstrated that in the engineered strain *B. subtilis* with NemNeuB from *Neisseria meningitidis*, co-introducing EcoNeuB from *E. coli* controlled by an IPTG-inducible promoter on plasmids can increase the titer of NeuAc [44]. Given that IPTG-inducible promoters may lead to high levels of gene expression, having multiple IPTG-inducible EcoNeuB genes can impose significant metabolic stress on the cells [45]. Moreover, when the inducer is added, different cells exhibit varying levels of induction [46]. Cells with lower induction strength, which experience less metabolic burden, tend to have a growth advantage. Consequently, during subsequent fermentation, strains with higher production intensity may be outcompeted and lost, leading to variability in production outcome [46]. Furthermore, since IPTG was already utilized for the induction of *recA* in this study, EcoNeuB was replaced with a constitutive promoter commonly used in *B. subtilis*. Based on a bacterial strain that produced

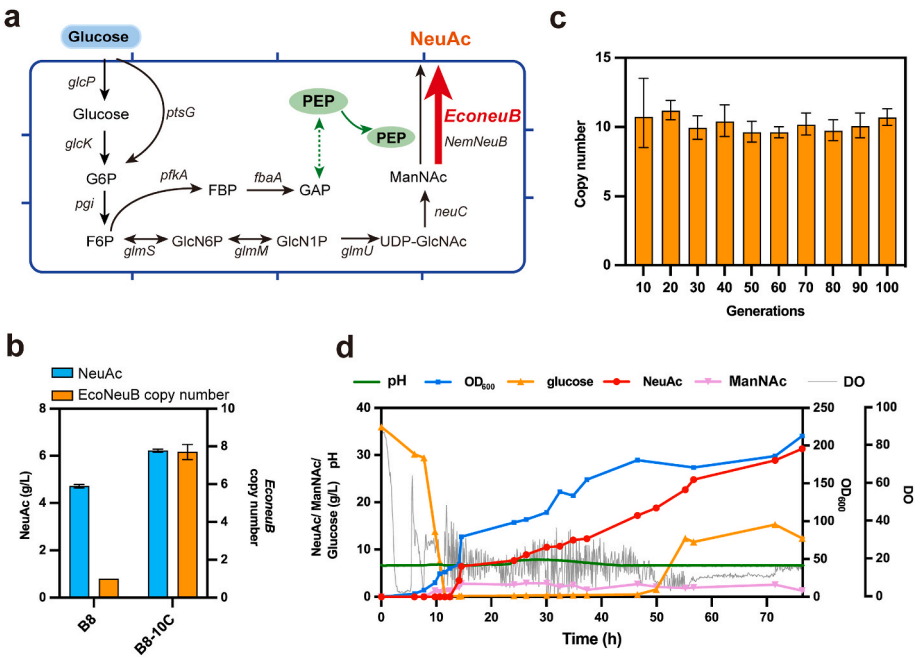

Fig. 5. Application of stabilized gene integration expression and copy number amplification system in NeuAc production. (a) The NeuC pathway in *B. subtilis* for the synthesis of NeuAc, where *neuB* has selected EcoNeuB from *Escherichia coli* and NemNeuB from *Neisseria meningitidis*, with gene multicopy of *EconeuB*. G6P, Glucose-6P; F6P, Fructose-6P; FBP, fructose-1,6-bisphosphate; GAP, glyceraldehyde 3-phosphate; PEP, phosphoenolpyruvic acid; GlcN6P, glucosamine-6P; GlcN1P, glucosamine-1P; UDP–GlcNAc, UDP-*N*-acetylglucosamine; ManNAc, *N*-acetylmannosamine; NeuAc, *N*-acetylneuraminic acid. *glmS*, encoding glutamine-fructose-6-phosphate aminotransferase; *glmM*, encoding phosphoglucosamine mutase; *glmU*, encoding glucosamine-1-phosphate acetyltransferase; *neuC*, UDP-*N*-acetylglucosamine-2-epimerase; *neuB*, encoding *N*-acetylneuraminic acid synthase. (b) The titer of *N*-acetylneuraminic acid and the gene copy number of *EconeuB* after gene amplification. (c) The stability of *EconeuB* copy number. (d) Fed-batch fermentation of strain B8–10C in a 3-L fermenter.

4.7 g/L of NeuAc, we further engineered the strain by amplifying the gene *EconeuB* originates from *E. coli*. Ultimately, through gene amplification and increasing the antibiotic concentration to ten times the original level, the gene copy number reached 7.7. When fermentation proceeded without antibiotics, the yield of NeuAc reached 6.3 g/L, a 30 % increment compared to the control (Fig. 5b). Next, we tested the stability of the multiple gene copies, a single colony with 11 gene copy numbers was passaged without chloramphenicol, the result showed that after 100 generations of passaging without antibiotics, the gene copy number remained stable at 11 copies (Fig. 5c). Subsequently, fed-batch fermentation was performed in a 3-L fermenter for the strain with 11 *EconeuB* copy, and the NeuAc titer and productivity reached 31.3 g/L and 0.41 g/L/h, respectively (Fig. 5d). At the end of the fermentation, the average gene copy number of *EconeuB* in the fermentation broth remained at 11. This is currently the highest yield of de novo synthesis in *B. subtilis* through plasmid-free fermentation.

4. Conclusion

In this study, we developed a stabilized gene amplification in *B. subtilis* through a reversible homologous recombination switch. The determination of recombination frequency validated the effect of homology arm length and crucial *recA* gene on homologous recombination in *B. subtilis*. Subsequently, a reversible switch based on repressor and ncAA-dependent expression system was developed to ensure stringent control over gene activation and silencing. The implementation of the reversible switch involving the gene *recA* subsequently enabled stabilized gene amplification. The system we developed in this study can optimize gene expression by regulating the copy number of key genes, enabling precise control over metabolic pathways. This tool can be used for the efficient production of valuable biomolecules, such as pharmaceuticals, industrial enzymes, and biofuels.

Notes

The authors declare no competing financial interest.

Data availability statement

All sequence maps are deposited in Zenodo (10.5281/zenodo.10554169).

CRediT authorship contribution statement

Haoyu Guo: Data curation, Methodology, Software, Writing – original draft. **Rongzhen Tian:** Data curation, Methodology, Writing – original draft. **Yaokang Wu:** Data curation, Writing – review & editing. **Xueqin Lv:** Data curation, Writing – review & editing. **Jianghua Li:** Data curation, Writing – review & editing. **Long Liu:** Conceptualization, Writing – review & editing. **Guocheng Du:** Conceptualization, Writing – review & editing. **Jian Chen:** Supervision, Writing – review & editing. **Yanfeng Liu:** Conceptualization, Supervision, Writing – review & editing.

Declaration of competing Interest

The authors declare no competing financial interest.

Acknowledgements

This study was financially supported by the National Key Research and Development Program of China (2020YFA0908300), the National Natural Science Foundation of China (32222069, 32172349), the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (32021005), and the Natural Science Foundation of Jiangsu Province (BK20202002).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.synbio.2024.04.010.

References

[1] Choi KR, Jang WD, Yang D, Cho JS, Park D, Lee SY. Systems metabolic engineering strategies: integrating systems and synthetic biology with metabolic engineering. Trends Biotechnol 2019;37(8):817–37. https://doi.org/10.1016/j.tibtech.2019.01.003.

[2] Nielsen J, Keasling JD. Engineering cellular metabolism. Cell 2016;164(6):1185–97. https://doi.org/10.1016/j.cell.2016.02.004.

[3] Lu H, Villada JC, Lee PKH. Modular metabolic engineering for biobased chemical production. Trends Biotechnol 2019;37(2):152–66. https://doi.org/10.1016/j.tibtech.2018.07.003.

[4] Wang Z-K, Gong J-S, Qin J, Li H, Lu Z-M, Shi J-S, et al. Improving the intensity of integrated expression for microbial production. ACS Synth Biol 2021;10(11):2796–807. https://doi.org/10.1021/acssynbio.1c00334.

[5] Arbel-Groissman M, Menuhin-Gruman I, Naki D, Bergman S, Tuller T. Fighting the battle against evolution: designing genetically modified organisms for evolutionary stability. Trends Biotechnol 2023;41(12):1518–31. https://doi.org/10.1016/j.tibtech.2023.06.008.

[6] Bang HB, Son J, Kim SC, Jeong KJ. Systematic metabolic engineering of *Escherichia coli* for the enhanced production of cinnamaldehyde. Metab Eng 2023;76:63–74. https://doi.org/10.1016/j.ymben.2023.01.006.

[7] Tyo KEJ, Ajikumar PK, Stephanopoulos G. Stabilized gene duplication enables long-term selection-free heterologous pathway expression. Nat Biotechnol 2009;27(8):760–5. https://doi.org/10.1038/nbt.1555.

[8] Nadler F, Bracharz F, Kabisch J. Copyswitch—in vivo optimization of gene copy numbers for heterologous gene expression in *Bacillus subtilis*. Front Bioeng Biotechnol 2019;6. https://doi.org/10.3389/fbioe.2018.00207.

[9] Peng B, Esquirol L, Lu Z, Shen Q, Cheah LC, Howard CB, et al. An in vivo gene amplification system for high level expression in *Saccharomyces cerevisiae*. Nat Commun 2022;13(1):2895. https://doi.org/10.1038/s41467-022-30529-8.

[10] Kachroo AH, Jayaram M, Rowley PA. Metabolic engineering without plasmids. Nat Biotechnol 2009;27(8):729–31. https://doi.org/10.1038/nbt0809-729.

[11] Liu Y, Liu L, Li J, Du G, Chen J. Synthetic biology toolbox and chassis development in *Bacillus subtilis*. Trends Biotechnol 2019;37(5):548–62. https://doi.org/10.1016/j.tibtech.2018.10.005.

[12] Liu Y, Zhang Q, Qi X, Gao H, Wang M, Guan H, et al. Metabolic engineering of *Bacillus subtilis* for riboflavin production: a review. Microorganisms 2023;11(1):164. https://doi.org/10.3390/microorganisms11010164.

[13] Schwechheimer SK, Park EY, Revuelta JL, Becker J, Wittmann C. Biotechnology of riboflavin. Appl Microbiol Biotechnol 2016;100(5):2107–19. https://doi.org/10.1007/s00253-015-7256-z.

[14] Liu Y, Link H, Liu L, Du G, Chen J, Sauer U. A dynamic pathway analysis approach reveals a limiting futile cycle in *N*-acetylglucosamine overproducing *Bacillus subtilis*. Nat Commun 2016;7(1):11933. https://doi.org/10.1038/ncomms11933.

[15] Lv X, Zhang C, Cui S, Xu X, Wang L, Li J, et al. Assembly of pathway enzymes by engineering functional membrane microdomain components for improved *N*-acetylglucosamine synthesis in *Bacillus subtilis*. Metab Eng 2020;61:96–105. https://doi.org/10.1016/j.ymben.2020.05.011.

[16] Gu Y, Lv X, Liu Y, Li J, Du G, Chen J, et al. Synthetic redesign of central carbon and redox metabolism for high yield production of *N*-acetylglucosamine in *Bacillus subtilis*. Metab Eng 2019;51:59–69. https://doi.org/10.1016/j.ymben.2018.10.002.

[17] Zhang X, Xu Z, Liu S, Qian K, Xu M, Yang T, et al. Improving the production of salt-tolerant glutaminase by integrating multiple copies of mglu into the protease and 16S rDNA genes of *Bacillus subtilis* 168. Molecules 2019;24(3):592. https://doi.org/10.3390/molecules24030592.

[18] Albertini AM, Galizzi A. Amplification of a chromosomal region in *Bacillus subtilis*. J Bacteriol 1985;162(3):1203–11. https://doi.org/10.1128/jb.162.3.1203-1211.1985.

[19] Yang S, Zhu J, Zhou X, Zhang J, Li Q, Bian F, et al. RNA-guided DNA transposition in corynebacterium glutamicum and *Bacillus subtilis*. ACS Synth Biol 2023;12(7):2198–202. https://doi.org/10.1021/acssynbio.3c00193.

[20] Zhang Y, Sun X, Wang Q, Xu J, Dong F, Yang S, et al. Multicopy chromosomal integration using CRISPR-associated transposases. ACS Synth Biol 2020;9(8):1998–2008. https://doi.org/10.1021/acssynbio.0c00073.

[21] Mazina OM, Mazin AV. Chapter eleven-Reconstituting the 4-strand DNA strand exchange. Methods in Enzymology. Academic Press; 2018. p. 285–305.

[22] Roca AI, Cox MM. RecA protein: structure, function, and role in recombinational DNA repair. Progress in Nucleic Acid Research and Molecular Biology. Academic Press; 1997. p. 129–223. https://doi.org/10.1016/s0079-6603(08)61005-3.

[23] Alonso JC, Stiege AC, Lüder G. Genetic recombination in *Bacillus subtilis* 168: effect of *recN, recF, recH* and *addAB* mutations on DNA repair and recombination. Mol Gen Genet MGG 1993;239(1):129–36. https://doi.org/10.1007/BF00281611.

[24] Fernández S, Kobayashi Y, Ogasawara N, Alonso JC. Analysis of the *Bacillus subtilis recO* gene: RecO forms part of the RecFLOR function. Mol Gen Genet MGG 1999;261(3):567–73. https://doi.org/10.1007/s004380051002.

[25] Haijema BJ, Noback M, Hesseling A, Kooistra J, Venema G, Meima R. Replacement of the lysine residue in the consensus ATP-binding sequence of the AddA subunit of AddAB drastically affects chromosomal recombination in transformation and transduction of *Bacillus subtilis*. Mol Microbiol 1996;21(5):989–99. https://doi.org/10.1046/j.1365-2958.1996.601424.x.

[26] Hiom K, West SC. Branch migration during homologous recombination: assembly of a RuvAB-holliday junction complex in vitro. Cell 1995;80(5):787–93. https://doi.org/10.1016/0092-8674(95)90357-7.

[27] Hall BM, Ma C-X, Liang P, Singh KK. Fluctuation anaLysis calculatOR: a web tool for the determination of mutation rate using Luria–Delbrück fluctuation analysis. Bioinformatics 2009;25(12):1564–5. https://doi.org/10.1093/bioinformatics/btp253.

[28] Alonso JC, Lüder G, Trautner TA. Intramolecular homologous recombination in *Bacillus subtilis* 168. Mol Gen Genet MGG 1992;236(1):60–4. https://doi.org/10.1007/BF00279643.

[29] Säll T, Nilsson NO. The robustness of recombination frequency estimates in intercrosses with dominant markers. Genetics 1994;137(2):589–96. https://doi.org/10.1093/genetics/137.2.589.

[30] Khasanov FK, Zvingila DJ, Zainullin AA, Prozorov AA, Bashkirov VI. Homologous recombination between plasmid and chromosomal DNA in *Bacillus subtilis* requires approximately 70 bp of homology. Mol Gen Genet MGG 1992;234(3):494–7. https://doi.org/10.1007/BF00538711.

[31] Fernández S, Ayora S, Alonso JC. *Bacillus subtilis* homologous recombination: genes and products. Res Microbiol 2000;151(6):481–6. https://doi.org/10.1016/S0923-2508(00)00165-0.

[32] Romero H, Serrano E, Hernández-Tamayo R, Carrasco B, Cárdenas PP, Ayora S, et al. *Bacillus subtilis* RarA acts as a positive RecA accessory protein. Front Microbiol 2020;11:92. https://doi.org/10.3389/fmicb.2020.00092.

[33] Carrasco B, Manfredi C, Ayora S, Alonso JC. *Bacillus subtilis* SsbA and dATP regulate RecA nucleation onto single-stranded DNA. DNA Repair 2008;7(6):990–6. https://doi.org/10.1016/j.dnarep.2008.03.019.

[34] Sciochetti SA, Blakely GW, Piggot PJ. Growth phase variation in cell and nucleoid morphology in a *Bacillus subtilis recA* mutant. J Bacteriol 2001;183(9):2963–8. https://doi.org/10.1128/jb.183.9.2963-2968.2001.

[35] Yasbin RE, Stranathan M, Bayles KW. The *recE(A)*+ gene of *B. subtilis* and its gene product: further characterization of this universal protein. Biochimie 1991;73(2):245–50. https://doi.org/10.1016/0300-9084(91)90209-J.

[36] Canas C, Carrasco B, Ayora S, Alonso JC. The RecU Holliday junction resolvase acts at early stages of homologous recombination. Nucleic Acids Res 2008;36(16):5242–9. https://doi.org/10.1093/nar/gkn500.

[37] Wu RAY, Bahl CP, Narang SA. Lactose operator–repressor interaction. Current topics in cellular regulation. Academic Press; 1978. p. 137–78. https://doi.org/10.1016/B978-0-12-152813-3.50009-4.

[38] Tian R, Liu Y, Cao Y, Zhang Z, Li J, Liu L, et al. Titrating bacterial growth and chemical biosynthesis for efficient *N*-acetylglucosamine and *N*-acetylneuraminic acid bioproduction. Nat Commun 2020;11(1):5078. https://doi.org/10.1038/s41467-020-18960-1.

[39] Wannier TM, Nyerges A, Kuchwara HM, Czikkely M, Balogh D, Filsinger GT, et al. Improved bacterial recombineering by parallelized protein discovery. Proc Natl Acad Sci USA 2020;117(24):13689–98. https://doi.org/10.1073/pnas.2001588117.

[40] Reams AB, Roth JR. Mechanisms of gene duplication and amplification. Cold Spring Harbor Perspect Biol 2015;7(2):a016592. https://doi.org/10.1101/cshperspect.a016592.

[41] Zhang J. Evolution by gene duplication: an update. Trends Ecol Evol 2003;18(6):292–8. https://doi.org/10.1016/S0169-5347(03)00033-8.

[42] Zhang X, Wang C, Lv X, Liu L, Li J, Du G, et al. Engineering of synthetic multiplexed pathways for high-Level *N* -acetylneuraminic acid bioproduction. J Agric Food Chem 2021;69(49):14868–77. https://doi.org/10.1021/acs.jafc.1c06017.

[43] Zhang X, Cao Y, Liu Y, Liu L, Li J, Du G, et al. Development and optimization of *N*-acetylneuraminic acid biosensors in *Bacillus subtilis*. Biotechnol Appl Biochem 2020;67(4):693–705. https://doi.org/10.1002/bab.1942.

[44] Guo H, Tian R, Wang C, Zhao R, Lv X, Liu L, et al. Improved *N*-acetylneuraminic acid bioproduction by optimizing pathway for reducing intermediate accumulation. Food Bioengineering 2022;1(3–4):205–11. https://doi.org/10.1002/fbe2.12030.

[45] Gomes L, Monteiro G, Mergulhão F. The Impact of IPTG induction on plasmid stability and heterologous protein expression by *Escherichia coli* biofilms. Int J Mol Sci 2020;21(2):576. https://doi.org/10.3390/ijms21020576.

[46] Azaman SNA, Ramanan RN, Tan JS, Rahim RA, Abdullah MP, Ariff AB. Screening for the optimal induction parameters for periplasmic producing interferon-α2b in *Escherichia coli*. Afr J Biotechnol 2010;9(38):6345–54.