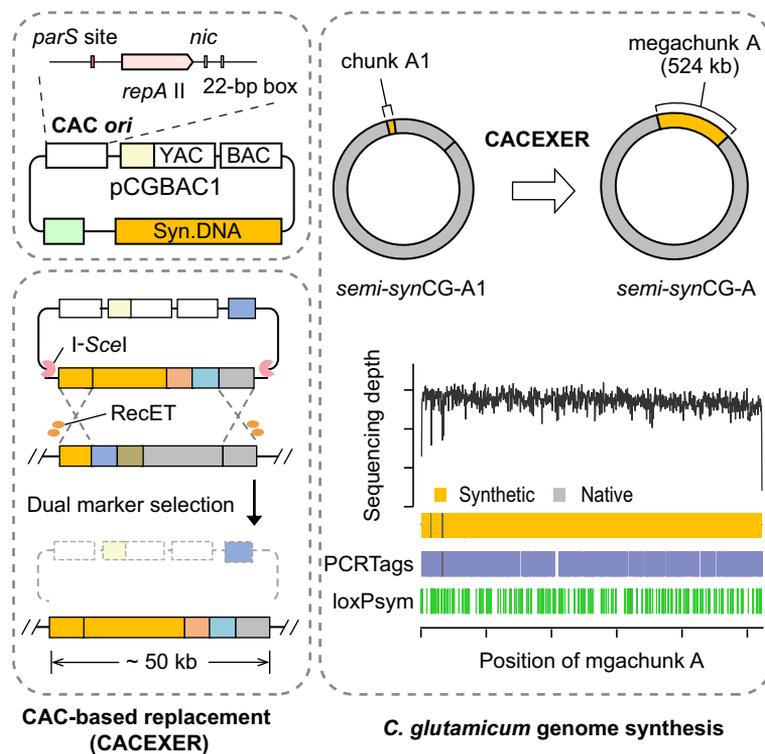


Research Article

Expediting genome synthesis of
Corynebacterium glutamicum with an artificial
chromosome vector

We developed an artificial chromosome plasmid (CAC) for large-scale genome replacement of *Corynebacterium glutamicum*, accelerating the genome synthesis for this organism. This work offers design principles for advancing *de novo* genome design and synthesis for industrially relevant Gram-positive microbes.

Zhanhua Zhang (张展华), Peixiong Hong (洪沛雄), Zebin Li (李泽斌), Baitao Li (李百涛), Tai Chen (陈泰), Yue Shen (沈玥), Xiaofeng Yang (杨晓锋), Yanrui Ye (叶燕锐), Yun Wang (王云), Zhanglin Lin (林章凛)

esyanyruiye@scut.edu.cn (Y. Ye),
wangyun@genomics.cn (Y. Wang)
zhanglinlin@gdut.edu.cn (Z. Lin).

Highlights

An artificial chromosome-like plasmid was developed for *Corynebacterium glutamicum*.

The *C. glutamicum* artificial chromosome facilitates stepwise genome replacements of ~50 kb.

In total, 361 kb of synthetic DNA was integrated into the *C. glutamicum* genome (11%).

Research Article

Expediting genome synthesis of
Corynebacterium glutamicum with an artificial
chromosome vector

Zhanhua Zhang (张展华)¹, Peixiong Hong (洪沛雄)¹, Zebin Li (李泽斌)¹, Baitao Li (李百涛)^{2,3}, Tai Chen (陈泰)⁴, Yue Shen (沈玥)^{2,3,4}, Xiaofeng Yang (杨晓锋)¹, Yanrui Ye (叶燕锐)^{1,*}, Yun Wang (王云)^{2,3,4,*}, and Zhanglin Lin (林章凇)^{1,5,*}

Recent advances in genome synthesis have relied on scalable DNA assembly and delivery, and efficient recombination techniques. While these methods have enabled rapid progress for *Escherichia coli* and yeast, they are often inadequate for other microorganisms. Here, we devised a *Corynebacterium glutamicum* artificial chromosome (CAC), which combines a replicating system from a closely related strain with an innate partitioning system. This CAC vector can efficiently deliver DNA fragments up to 56 kb and maintain stability in *C. glutamicum*. Leveraging the CAC vector, we developed CAC Excision Enhanced Recombination (CACEXER), a streamlined strategy for iterative genome replacements in *C. glutamicum*. Using this approach, we integrated 361 kb (11%) of synthetic DNA into the genome, creating *semi-synCG-A*. This strain paves the way to establish *C. glutamicum* as the third industrial microorganism, alongside *E. coli* and *Saccharomyces cerevisiae*, to undergo large-scale genome synthesis.

Introduction

Genome synthesis is a powerful technique for expansion of genomics knowledge and novel functions. Organisms with redesigned genomes offer opportunities for exploring codon compression [1–4], genome rearrangement [1,2,5,6], and genome simplification [7]. Over the past 15 years, notable progress has been achieved in synthesizing a complete *Mycoplasma mycoides* genome [8,9] and genome-wide **synonymous recoding** (see [Glossary](#)) of *E. coli* genomes, including the completed Syn61 [4] and the ongoing Ec_Syn57 [10,11]. The Synthetic Yeast Genome Project (Sc2.0) has assembled all synthetic *S. cerevisiae* chromosomes, with ongoing efforts to consolidate them into a fully synthetic strain [1,2,12–14]. Synonymous codon substitutions have also been partially introduced in other microorganisms, such as *Salmonella typhimurium* LT2 [15], and the recoded essential genes of *Caulobacter crescentus* have been assembled in an extra plasmid-based copy [16]. Genome synthesis has now been extended to higher organisms [17,18]. The SynMoss project has simplified a portion of a chromosome arm in *Physcomitrium patens* [19] and a project was proposed to rearrange the genome of marine diatoms into 50 similarly sized chromosomes [20]. In addition, the similar techniques have been utilized for partial genome engineering of mammalian cells. The essential amino acid valine biosynthesis pathway was genomically integrated into a Chinese hamster ovary (CHO) cell line [21], and multiplex base editing was used to convert TAG codons into TAA in human HEK293T cells [22].

Among the completed synthetic genome projects, transplantation of whole synthetic genomes has been demonstrated for *Mycoplasma* (1 Mb) [9]. By contrast, genome synthesis for *E. coli*

Technology readiness

This work establishes a bacterial artificial chromosome (BAC)-like vector for *Corynebacterium glutamicum* and introduces the streamlined *C. glutamicum* Artificial Chromosome (CAC) Excision Enhanced Recombination (CACEXER) strategy for *de novo* genome synthesis of this industrially important Gram-positive microbe. To date, we have successfully integrated 361 kb (or 11%) of synthetic DNA into the genome. To further enhance this CAC-based approach and accelerate genome synthesis, we aim to incorporate recent advancements in large-scale DNA assembly and delivery. In addition, genome debugging to mitigate growth defects will be crucial for achieving a fully synthetic *C. glutamicum* genome. Based on our current progress, we propose that this technology has reached a Technology Readiness Level (TRL) between 3 and 4, as defined by NASA.

¹School of Biology and Biological Engineering, South China University of Technology, Guangzhou, Guangdong 510006, China

²BGI Research, Shenzhen, Guangdong 518083, China

³Guangdong Provincial Key Laboratory of Genome Read and Write, BGI-Shenzhen, Shenzhen, Guangdong 518120, China

⁴BGI Research, Changzhou, Jiangsu 213299, China

⁵Present address: School of Biomedicine, Guangdong University of Technology, Guangzhou, Guangdong 510006, China

(4 Mb) and *S. cerevisiae* (12 Mb) requires two foundational techniques: (i) delivery of large synthetic sequences inside the cells; and (ii) efficient recombination for genomic iterative replacements by large synthetic DNA segments. More specifically, for *E. coli*, a **bacterial artificial chromosome (BAC)** was required for delivery of large synthetic sequences (91–136 kb), and the lambda-red recombination system has been effectively utilized in Replicon Excision Enhanced Recombination (REXER) in Syn61 [3]. For yeast, delivery of large synthetic sequences (30–60 kb) by either **yeast artificial chromosome (YAC)** or direct transformation was successful in Sc2.0, and with its long-exploited innate homologous recombination capacity, Switching Auxotrophies Progressively for Integration (SwAP-In) was also developed [2,23–25]. Similar to yeast, the ongoing synthetic moss project relies on the delivery of large synthetic sequences by direct transformation and recombination via its innate capacity [19]. However, these techniques are often lacking in other organisms, including bacteria [26], posing a significant barrier to genome synthesis beyond *E. coli*, yeast, and moss.

C. glutamicum is generally recognized as safe (GRAS) and serves as a vital industrial microorganism extensively used for manufacturing various amino acids, organic acids, and bio-based products [27]. The synthesis of its genome holds promise to accelerate strain improvement and refine this indispensable cell factory with enhanced functionalities. As a proof of concept, we previously developed a genomic iterative replacement method based on **Rac prophage exonuclease-recombinase (RecET)** for *C. glutamicum* [28]. In this pilot study, we successfully redesigned and integrated 55.1 kb of synthetic sequence into a 53.4-kb region of the wild-type genome. Our work demonstrated the feasibility of recoding **PCRTags** and stop codons, decoupling overlapped genes, and inserting **loxPsym sites** in *C. glutamicum* without noticeably compromising its fitness or genome stability [28]. Furthermore, we found that **Synthetic Chromosome Rearrangement and Modification by loxP-mediated Evolution (SCRaMble)** can also function well in *C. glutamicum*, a technique hitherto limited to synthetic yeast chromosomes [2].

However, in the previous study, we had to rely on direct integration of 6–10-kb DNA fragments into *C. glutamicum*, because there was no BAC-like or YAC-like artificial chromosome vector available, which hindered larger-scale genome synthesis for this organism [2,4,29,30]. In the current study, we successfully engineered such a vector, termed **CAC**, capable of delivering large DNA fragments up to 56 kb. In addition, we devised a streamlined strategy for iterative genome replacements by using **meganuclease** I-SceI excision of DNA fragments from the CAC plasmids, followed by RecET-mediated recombination [4,31–33]. This approach significantly expedited the genome synthesis of *C. glutamicum*. We have now successfully replaced additional 361 kb (11%) of the genome. Overall, our work paves the way toward the total genome synthesis of a third industrial microorganism, *C. glutamicum*, alongside *E. coli* and *S. cerevisiae*, and offers insights applicable to genome synthesis in other microorganisms.

Results

Design and testing of an artificial chromosome for *C. glutamicum*

In our preliminary experiments, we found that existing BAC vectors, with or without the addition of the broad host-range RK2 replicon, could not be maintained in *C. glutamicum* (Figure S1A, Tables S1–S3, and Note S1 in the supplemental information online). This was expected, because BAC and RK2 replicons are generally nonfunctional in Gram-positive bacteria [34–37]. The result underscored the need to develop a CAC vector. To this end, we followed the principles of BAC/YAC vectors, which serve three crucial functions: (i) the capacity to harbor large DNA fragments; (ii) the replication of these fragments; (iii) and the proper partitioning of the plasmids into daughter cells [3,38–40]. We drew inspiration from two natural replicons found in large plasmids in species closely related to *C. glutamicum*. One replicon, derived from the 51-kb plasmid pCXC100 of the

*Correspondence: esyanyuie@scut.edu.cn (Y. Ye), wangyun@genomics.cn (Y. Wang), and zhanglinlin@gdut.edu.cn (Z. Lin).

Gram-positive *Leifsonia xyli* subsp. *Cynodontis* [41,42], contains a replication gene *repA* (although clear annotation for an origin of replication is lacking), partitioning genes, *parA* and *parB*, and a direct repeat *parS* site. The other replicon, originated from the 67.8-kb pBL90 plasmid from *Brevibacterium Lactofermentum* DSM 1412 [43,44], harbors a replication gene *repA* II, an origin of replication (*nic*), and a 22-base pair (bp) box motif homologous to the replication start point. Furthermore, this replicon carries partitioning genes *parA* II and *parB*, although clear annotation for the partition site is lacking [44] (Figure S1 and Note S1).

We validated their functionality by separately integrating the *E. coli* replicon and the two replicons into pOK12 [45]. This resulted in the creation of two putative *E. coli*-*C. glutamicum* shuttle vectors, named pOK12CXC100 and pOK12BL90 (Figure S1B). Following electroporation into *C. glutamicum*, pOK12BL90 yielded viable transformants, confirmed through colony PCR targeting the plasmid-borne *repA* II on the pBL90 replicon. By contrast, pOK12CXC100 failed to generate viable transformants, indicating the inefficacy of the replicon from pCXC100 in *C. glutamicum*.

We then assembled the putative *S. cerevisiae*-*E. coli*-*C. glutamicum* shuttle vector pCGBACYT90 by integrating the YAC replicon into the pOK12BL90 backbone. Electroporation of this vector into *C. glutamicum* was followed by verification of transformants through colony PCR (Figure S1B). Next, we incorporated the 55.1-kb synthetic sequence (chunk A1) used in a previous study [28] into pCGBACYT90 to evaluate its capacity to carry large DNA fragments, resulting in pCGBACYT90-A1 (Figure S1C). After plasmid assembly in *S. cerevisiae* and propagation in the intermediate host *E. coli* [33], pCGBACYT90-A1 was introduced into *C. glutamicum* via electroporation, and confirmed by PCR verification. However, restreaking revealed the absence of all junction bands for pCGBACYT90-A1 (Figure S1D), indicating vector instability when loaded with large DNA fragments during passage.

The bacterial **partitioning system (parABS)** is widespread on low copy-number plasmids and chromosomes [46]. We hypothesized that the replicon cloned from pCGBACYT90 may lack a centromere-like *parS* site, causing vector instability during DNA segregation [47]. Given that segregation of the *C. glutamicum* chromosome (3.2 Mb) is driven by its intrinsic parABS system, which includes ten *parS* sites proximal to the origin of replication (*oriC*) [48], we exploited this endogenous parABS system to segregate and partition the artificial chromosome vector. In doing so, the DNA sequence of *parA* II and *parB* on pCGBACYT90-A1 plasmid were replaced with a 16-bp genomic partition site *parS* (5'-tggttcacgtgaaca-3'), resulting in the creation of pCGBAC1-A1 (Figures 1A, S1E, and Note S1). Following assembly and extraction of the pCGBAC1-A1 plasmid (Figure S1F), we electroporated it into *C. glutamicum*. Verification of synthetic PCRTags on the pCGBAC1-A1 plasmid in transformants through colony PCR yielded the expected results (Figure S1G). Upon restreaking of individual colonies of the transformants, the presence of synthetic PCRTags on pCGBAC1-A1 was reaffirmed (Figure 1B), indicating successful delivery and maintenance of the pCGBAC1-A1 plasmid carrying the 55.1-kb chunk A1 in *C. glutamicum*.

The strain with pCGBAC1-A1 was also subjected to passage culture of ~130 generations, and the stability of the plasmid was assessed through both PCRTag assays and next-generation sequencing (NGS). Eight single colonies from the passage culture were selected and examined, revealing that these strains carried all synthetic fragments without detectable duplication or recombination events occurring among the 20 loxPsym sites (Figures 1C, S1G, and Table S4 in the supplemental information online). These results collectively affirm the faithful stability of the CAC vector in carrying large synthetic DNA fragments exceeding 50 kb in size without compromising genome stability.

Glossary

Bacterial artificial chromosome

(BAC): large-capacity cloning vector used to carry and maintain large DNA fragments (typically up to 300 kb) in bacterial cells for genome engineering and synthetic biology applications.

Corynebacterium glutamicum

artificial chromosome (CAC): newly developed vector designed to deliver and maintain large DNA fragments in *C. glutamicum*, enabling iterative genome replacement and large-scale genome synthesis.

LoxPsym site: 34-bp palindromic variant of the loxP site that can undergo recombination in both orientations. The insertion of loxPsym sites into the genome enables genomic rearrangements mediated by Cre recombinase. This approach was first implemented in the synthetic yeast genome project and has since been extended to the synthetic *C. glutamicum* genome project.

Meganuclease: also termed 'homing endonuclease'; site-specific endonucleases that typically recognize long DNA sequences (>14 bp) and induce double-strand breaks.

Partitioning system (parABS):

tripartite system responsible for the proper segregation of low copy-number plasmids and chromosomes during cell division in most bacterial species and several archaea. It typically comprises three key components: ParA, an ATPase that facilitates movement; ParB, a CTPase that binds to the *parS* DNA sequence; and one or more *parS* sites, which serve as centromere-like elements to ensure proper partitioning of genetic material.

PCRTags: short pairs of recoded sequences that are designed *in silico* to selectively distinguish either wild-type or synthetic genome sequences. PCRTags facilitate rapid tracking and verification of synthetic DNA integration using PCR-based assays.

Rac prophage exonuclease-

recombinase (RecET): type of homologous recombination system involving RecE (an exonuclease) and RecT (a recombinase), typically derived from *Escherichia coli* Rac prophage.

Synonymous recoding: modification of the coding sequence of a genome by substituting synonymous codons that encode the same amino acid or translational stop signal.

Genomic replacement by CACEXER

We then engineered a helper plasmid, pXMJ19-RecET-I-SceI, harboring *I-SceI* and *RecET*, conferring chloramphenicol resistance (+3, *cat*), to complement the CAC vector (Figure S2A, Tables S1–S3, and Note S1). We chose *I-SceI* meganuclease over CRISPR/Cas9 to streamline excision of double-stranded (ds)DNA fragments from CAC [3,4] (Figure 2A). The use of *I-SceI* also minimizes cellular burden and circumvents Cas9 cytotoxicity [49–51]. As in the previous study [28], *RecET* was used to promote genomic replacement. We used dual selection markers for augmented recombination, one with -1 (*rpsL*) and $+1$ (*kanR*) downstream of synthetic sequence [28], and -2 (*sacB*, confers sensitivity to sucrose) on the backbone; the other with -2 (*sacB*) and $+2$ (*speR*) downstream of synthetic sequences, and -1 (*rpsL*) on the backbone [4,28,52]. This allowed for concurrent selection of synthetic fragment replacement from the CAC plasmid into the genome, and removal of the negative marker from the targeted genomic site. We named this system CACEXER.

We next used CACEXER to integrate a 53.4-kb synthetic DNA (chunk A4) into the strain *semi-synCG-A3* with the helper plasmid pXMJ19-RecET-I-SceI. Strain *semi-synCG-A3* was constructed by sequential replacements of 6–10-kb DNA fragments into the genome of *semi-synCG-A1* [28], harboring a *rpsL-kanR* cassette at the 3' end of chunk A3 (Figure 2A and Notes S1 and S2 in the supplemental information online). Chunk A4 encompasses the region from *NCgl0032* to *NCgl0077* (Figure 2A). It was computationally designed, divided into 11 minichunks 4–5 kb in length, and synthesized by BGI (Changzhou, China) [28].

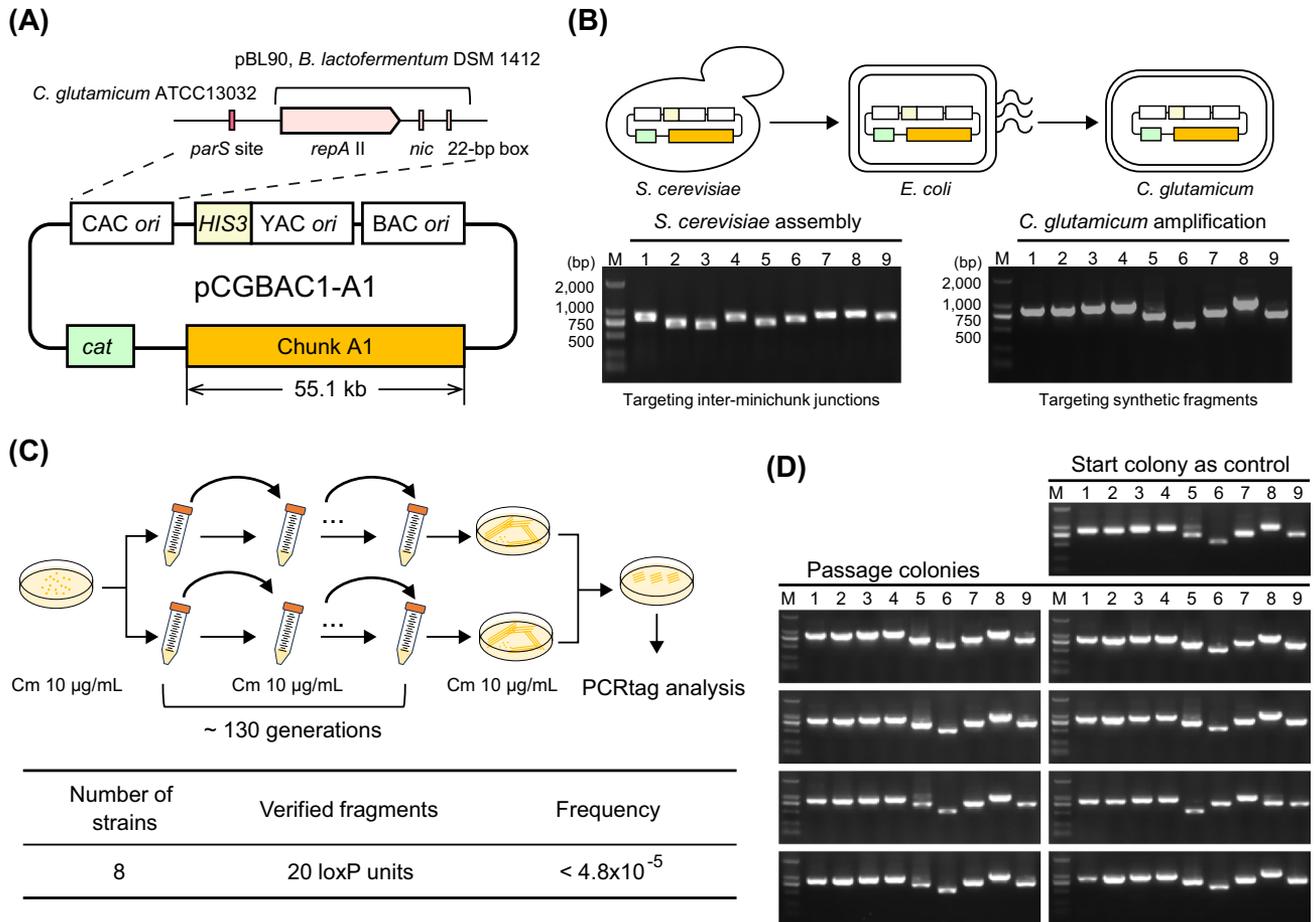
We assembled the CAC plasmid pCGBAC1-A4 in yeast (Figure S2B), amplified it in *E. coli*, and electroporated it into *semi-synCG-A3*. We subsequently induced the expression of the *I-SceI* meganuclease and *RecET* recombination components with arabinose and IPTG (Figures 2A and S2C). Recombinant strains were selected on agar plates supplemented with chloramphenicol, spectinomycin, and streptomycin. We used genotyping to confirm the loss of the double-selection marker (*rpsL-kanR*) from the genome, the loss of the negative selection marker *rpsL* from the vector backbone, and the insertion of the new double-selection marker (*sacB-speR*) into the genome (Figure 2A). We further sequenced the junctions between the DNA fragment integrated by CACEXER and the rest of the genome for the post-CACEXER strains (Figure S2D), and confirmed that the mismatched sequences generated by *I-SceI* excision from the CAC plasmid were removed and that the replacement was scarless (Figure S2E). We analyzed the expression of *RecE*, *RecT*, and *I-SceI* via western blotting using a modified construct in which *RecE* and *RecT* were tagged at the N and C termini with HA tags, and *I-SceI* was C-terminally tagged with a FLAG tag (Figure S2F and Note S1). All three proteins were successfully expressed 4 h post induction (Figure S2G,H). The expression levels of *RecT* and *I-SceI* were relatively higher than that of *RecE*, an observation consistent with the inherently moderate expression levels of exonucleases, as previously reported [50,53]. It has been suggested that overexpression of *RecT* relative to *RecE* enhances recombination efficiency [53].

Upon PCRTag analysis, we observed that 68 out of the 70 synthetic PCRTags in this chunk were successfully integrated, while two wild-type PCRTags were not substituted as designed (Tables S5–S7 in the supplemental information online). To ensure the complete integration of chunk A4, we divided A4 into two subchunks and assembled two CAC plasmids accordingly: pCGBAC1-A4a, which carried a 29-kb synthetic fragment (from *NCgl0032* to *NCgl0059*) with the *sacB-speR* double-selection marker, and pCGBAC1-A4b, which contained the remaining 25-kb synthetic fragment (from *NCgl0059* to *NCgl0077*) with the *rpsL-kanR* marker (Table S5). We electroporated pCGBAC1-A4a into *semi-synCG-A3* and induced replacement. Selection and genotyping were performed as described above (Figures 2B and S2I). This time, we

Synthetic Chromosome Rearrangement and Modification by loxP-mediated Evolution (SCRaMbLE):

genome engineering approach that utilizes Cre-lox recombination to generate structural variations in synthetic chromosomes, allowing for rapid evolution and functional optimization of engineered genomes.

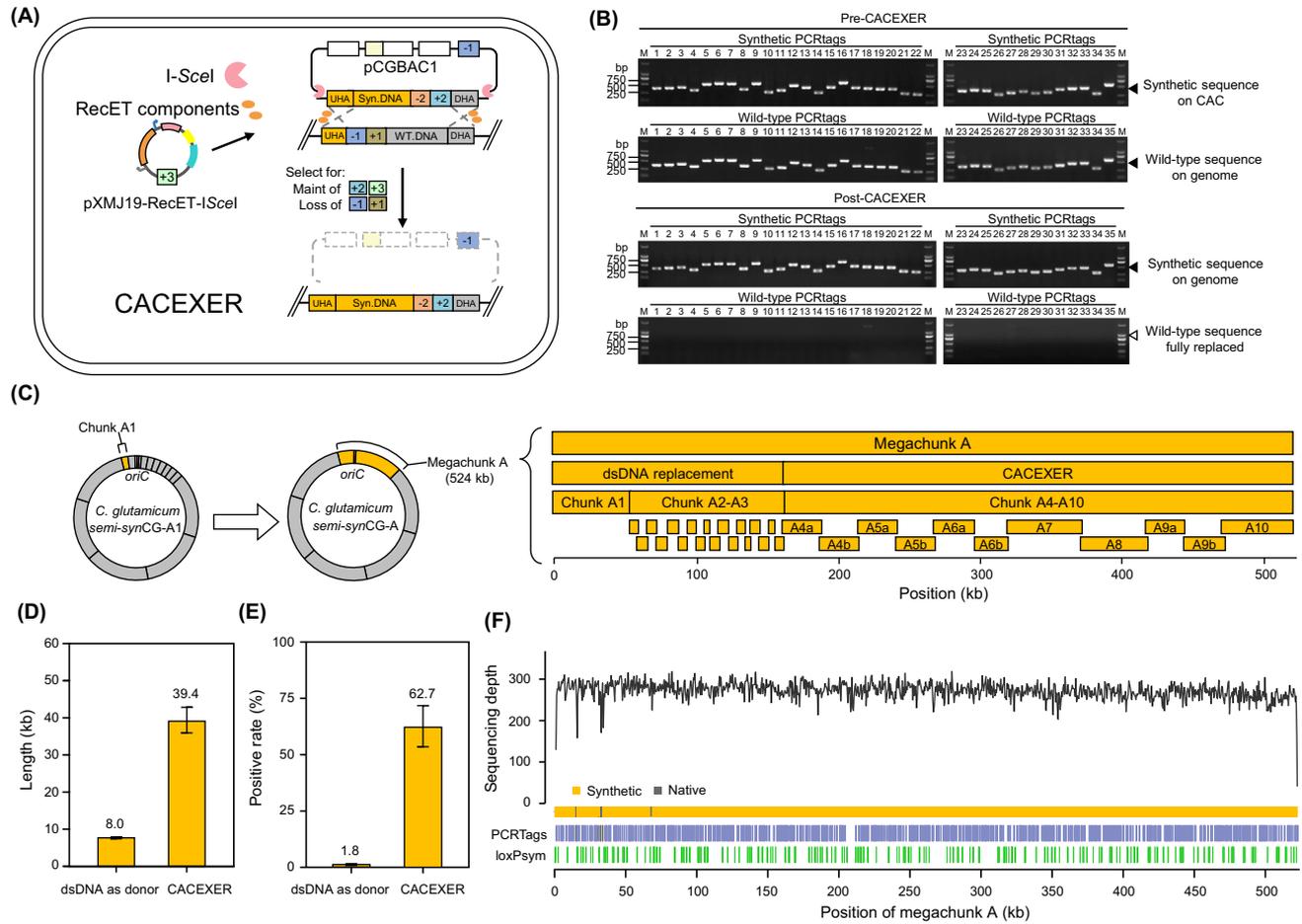
Yeast artificial chromosome (YAC): large-capacity cloning vector used to carry and maintain large DNA fragments (typically up to Mb scale) in yeast cells.



Trends in Biotechnology

Figure 1. Design, construction, and characterization of a *Corynebacterium glutamicum* artificial chromosome (CAC) plasmid pCGBAC1-A1. (A) Schematic of the pCGBAC1-A1 plasmid. The pCGBAC1 replicon contains a yeast artificial chromosome (YAC) origin and a *HIS3* (pale yellow) marker, a bacterial artificial chromosome (BAC) origin and a *cat* gene (light green), for maintenance in *Saccharomyces cerevisiae* and *Escherichia coli*, and the tailored CAC origin comprising a replication gene *repA II*, an origin of replication (*nic*), and a 22-base pair (bp) box motif homologous to replication start point (light pink) from *Brevibacterium lactofermentum* DSM 1412 pBL90 plasmid, as well as the *parS* site from the genome (red). The 55.1-kb synthetic fragments (bright yellow) were incorporated to generate the pCGBAC1-A1 plasmid. (B) Assembly and delivery of the pCGBAC1-A1 plasmid among *S. cerevisiae*, *E. coli*, and *C. glutamicum*. The pCGBAC1-A1 plasmid was assembled by homologous recombination in *S. cerevisiae* MYA3666 and verified by colony PCR spanning each inter-minichunk junction. After extraction from yeast, pCGBAC1-A1 was electroporated into *E. coli* NEB 10-beta for amplification. Following extraction from *E. coli*, it was electroporated into *C. glutamicum*, and verification was performed using selected synthetic PCRTags and primers targeting the junctions between synthetic sequence and the CAC backbone. Identical PCR results were obtained across at least three repetitions for each assembly and delivery step; see also Tables S1 and S2 in the supplemental information online. (C) Stability of the pCGBAC1-A1 plasmid after ~130 generations. PCRTag assay and next-generation sequencing (NGS) of *C. glutamicum* harboring pCGBAC1-A1 was performed. Eight strains were also verified by NGS and no losses were observed. The maximum estimation of loxPsym loss frequency on pCGBAC1-A1 was 4.8×10^{-5} per generation; see also Table S4 in the supplemental information online. (D) Colony PCR analysis for pCGBAC1-A1 passage. The starting colony was used as a positive control. Lane M, DL2000 DNA Marker (TaKaRa, China); lanes 1–9, PCR products using the same primers for synthetic fragment analysis in (B); see also Table S2.

additionally restreaked the strains onto selection plates for overnight growth. Two colonies were then selected and verified by Sanger sequencing and PCRTag analysis, confirming complete replacement of the wild-type genomic sequence with synthetic chunk A4a in both colonies. This resulted in the creation of strain *semi-synCG-A4a* (Figures 2B and S2I). Subsequently, we assembled and electroporated the pCGBAC1-A4b into *semi-synCG-A4a*. This time, three out of 15 colonies selected after restreaking were confirmed to contain the correct synthetic chunk A4 by Sanger sequencing and PCRTag analysis. Thus, the two-step strategy accomplished the full integration of synthetic chunk A4 into the genome.



Trends in Biotechnology

Figure 2. Scarless genomic integration of synthetic DNA into the *Corynebacterium glutamicum* genome using *C. glutamicum* artificial chromosome (CAC) plasmids CAC Excision Enhanced Recombination (CACEXER). (A) Strategy for integrating synthetic DNA into *C. glutamicum* genome using CACEXER. The starting strain contained the double-selection marker *rpsL-kanR* at the 3' end (-1/+1 is shown) and the pXMJ19-RecET-I-SceI plasmid conferring chloramphenicol resistance (+3, *cat*, light green). The CAC plasmid delivered the synthetic DNA fragments into cells via electroporation. The transformants were selected on agar plates with +2 and +3. The helper plasmid expressed I-SceI and RecET components upon induction of arabinose and IPTG. Successful replacement of genomic DNA with synthetic DNA was identified by the gain of +2 and the loss of -1, indicating successful integration and loss of the CAC backbone, respectively. The double-selection markers used in different steps of CACEXER were -1 (*rpsL*, aqua blue), +1 (*kanR*, khaki), -2 (*sacB*, orange), and +2 (*speR*, light blue). (B) PCRtag analysis verification pre and post CACEXER. Left panel: CACEXER results for chunk A4a. Right panel: CACEXER results for chunk A4b; see also Tables S5 and S7 in the supplemental information online. (C) Assembly workflow of *semi-synCG-A*. Starting from *semi-synCG-A1*, an additional 18 steps using a double-stranded (ds)DNA fragment of integrations and 11 steps of CACEXER was performed. The position of the replication of the replication origin *oriC* (red bar) is indicated; see also Figure S6 and Table S5 in the supplemental information online. (D, E) Comparison of the length (D) and positive rate (E). The positive rate in (E) is judged by clones with the correct recombination of markers. Source data for (D, E) are in Table S5 and Data S1 in the supplemental information online. Error bars represent standard error of the mean (SEM). (F) Next-generation sequencing (NGS) coverage map for megachunk A verification. Read coverage (y axis) is plotted against the starting base pair (x axis). Except for eight PCRTags (seven in chunk A1 [28] and one in chunk A2; see also Figure S7B in the supplemental information online) indicated by drops in sequencing depth, the megachunk A incorporated the designed features without evidence of genome duplications, insertion sequence (IS), or transposon mobility. The vertical lines indicate the positions of PCRTags (blue) and loxP sites (green). Abbreviations: maint., maintenance; WT, wild-type.

Subsequent iterative CACEXER for genome replacement

We conducted an additional nine rounds of iterative CACEXER based on *semi-synCG-A4*, successfully integrating an additional 308 kb of genome sequence labeled as chunks A5–A10 (Figures 2C, S3, S4, and Table S5 in the supplemental information online). In each strain, the double-selection marker *rpsL-kanR* or *sacB-speR* from the previous round served as a template for the next round of CACEXER. For chunks A6–A8, and A10, we were able to completely replace

the targeted genomic regions using CACEXER. However, complete integration of chunk A5 and chunk A9 proved challenging, leading us to split each into two 25-kb subchunks to accomplish the integration (Figure 2C and Table S5).

Judging from both phenotyping and genotyping, the rate of correct recombinants ranged from 22.2% to 100% in eight of the nine iterative CACEXER steps, with the exception of the integration of chunk A7, which had a rate of only 0.84% (Figure S4A and Table S5). We examined the underlying cause for this low rate and observed that, while *sacB*-mediated negative selection reliably facilitated integration for 30-kb fragments, it exhibited poor performance for integration of 50-kb fragments (Table S5), indicating a possible inadequate supply of levansucrase for 50-kb integration (Figure S5). In addition, Sanger sequencing revealed that several failed integrations stemmed from unexpected insertions by the genomic *ISCG1* transposase [54] in AT-rich loci within the coding sequence of *sacB* (Figure S5B). To address these issues, we used a strong constitutive promoter P_{tuf} for *sacB* [28,55,56], and optimized the codons of the gene (labeled '*sacB*^{CG}'). We also introduced an additional negative selection marker, *pheS** (-4, *pheS*^{T262A_A309G}; confers sensitivity to 4-chloro-DL-phenylalanine) [4,57]. The combined *pheS**-*sacB*^{CG}-mediated negative selection achieved a 100% rate of recombinants for the 55.7-kb chunk A7 (Table S5 and Note S1).

We conclude that iterative CACEXER can expedite genomic replacement of DNA fragments up to 56 kb, offering simplified operations and achieving a moderate success rate for replacement.

Characteristics of the semi-synthetic *C. glutamicum* genome

After completing the iterative CACEXER replacement, we deleted the double-selection marker (Figure S6A in the supplemental information online). In addition, we successfully decoupled a pair of overlapping genes, which was not achieved in the initial attempt (Figure S6B), establishing the *semi-synCG-A* strain. The sequence of *semi-synCG-A* comprises chunk A1 [28] and chunks A2–A10 (Figure 2C and Figure S7 in the supplemental information online). In total, 519 097 bp of the genomic region of *C. glutamicum* ATCC13032 were substituted with 523 620 bp, altering 96.8% (455/470) of genes compared with the wild-type. This strain contained various designed features, including recoding of 790 PCRTags, swapping of 157 stop codons from TAG to TAA, decoupling of 80 pairs of overlapped genes, insertion of 178 loxPsym sites, and deletion of three annotated insertion sequence (IS) elements: *tnp2f* (*NCgl0179*), *tnp16a* (*NCgl0235*), and *tnp17a* (*NCgl0348*), each longer than 500 bp. Notably, the number of decoupled gene pairs within megachunk A alone (80) was comparable with the total number of decoupled genes in the whole genomic synthesis of *E. coli* syn61 (91) [4].

NGS sequencing of the *semi-synCG-A* genome confirmed all aforementioned features, and showed no evidence of genome duplications, IS, or transposon mobility (Figure 2F). Nineteen point mutations were observed in the 361-kb A4–A10 chunks replaced by CACEXER, with a mutation rate (5×10^{-5}) lower than that of dsDNA replacement in chunks A1–A3 (2×10^{-4} , Table S6) [28]. Genomic stability of *semi-synCG-A* without Cre expression was also assessed. Both PCRTag analysis and whole-genome sequencing demonstrated stability over 130 generations of nonselective growth in three independent lineages, with no loss of loxPsym sites and no additional mutations or genome rearrangements (Figure S8 in the supplemental information online). This affirmed the faithful genome integrity.

During the construction of *semi-synCG-A*, double-selection markers were inserted into the coding sequences of hypothetical genes (*NCgl0059*, *NCgl0077*, *NCgl0099*, *NCgl0124*, *NCgl0153*, *NCgl0177*, and *NCgl0296*) in seven out of 11 CACEXER steps without impacting organism

survival. However, we observed slow growth of strains when the selection marker was inserted in intergenic regions upstream of either *parA* (*NCgl2989*, ATPase involved in chromosome partitioning) or *NCgl0006* (a hypothetical gene), which are proximal to the chromosomal origin of replication (*oriC*). In addition, the inefficient replacement of chunk A9 suggests that interrupting *ushA* (*NCgl0322*, a putative 5-nucleotidase precursor [58]) is not tolerated. With 23% of the genes in *C. glutamicum* still lacking complete functional annotation, genome synthesis provides an opportunity to better understand these unannotated regions.

Phenotyping of the *semi-synCG-A*

To assess the impact of modifications on the cell morphology of the *semi-synCG-A* strain, we conducted a comparative analysis with the wild-type strain ATCC13032 using scanning electron microscopy (SEM). Interestingly, the cell shape of the *semi-synCG-A* remained unaffected, while we noted a longer cell length compared with ATCC13032, consistent with our previous observations [28] (Figure 3A).

The fitness of *semi-synCG-A* was evaluated through serial dilution and microplate liquid growth assays under various conditions, using the wild-type strain ATCC13032 as a reference (Figure 3B, Figure S9, Table S8, and Data S1 in the supplemental information online). Slower growth of *semi-synCG-A* strain was consistently observed on BHIS plates at different temperatures (30°C, 25°C, and 37°C), under different pH (6.0 and 8.0), and under osmotic and oxidative stresses (Figure S9). In BHIS medium (pH 7.5) at 30°C, the doubling times for the *semi-synCG-A* strain and ATCC13032 were 3.45 h and 2.32 h, respectively ($P < 0.001$, *t*-test). The final optical density (OD)₆₀₀ (at 24 h) was 0.64 ± 0.00 for *semi-synCG-A*, and 0.58 ± 0.01 for ATCC13032 ($P < 0.05$, *t*-test), respectively. In BHIS medium (pH 7.5) at 25°C, the doubling time for *semi-synCG-A* was 1.42-fold longer than for ATCC13032 (5.81 ± 0.34 h versus 4.10 ± 0.14 h, $P < 0.001$, *t*-test), while the final OD₆₀₀ for *semi-synCG-A* was 1.30-fold higher (0.65 ± 0.03 versus 0.50 ± 0.01 , $P < 0.001$, *t*-test). In BHIS medium (pH 7.5) at 37°C, the doubling time for *semi-synCG-A* strain increased by 1.56-fold (4.96 ± 0.07 h versus 3.18 ± 0.10 h, $P < 0.001$, *t*-test), and the final OD₆₀₀ decreased by 51% (0.27 ± 0.00 versus 0.55 ± 0.03 , $P < 0.001$, *t*-test). In BHIS medium with 1 M sorbitol (pH 7.5) at 30°C, the doubling time for *semi-synCG-A* strain and ATCC13032 was 4.63 ± 0.11 h and 3.73 ± 0.24 h ($P < 0.001$, *t*-test), respectively, with no significant difference in the final OD₆₀₀ (0.53 ± 0.04 versus 0.56 ± 0.04 , $P = 0.256$, *t*-test). In BHIS medium with 2 M sorbitol (pH 7.5) at 30°C, the doubling time for *semi-synCG-A* and ATCC13032 was 4.10 ± 0.22 h and 5.09 ± 0.20 h, respectively ($P < 0.001$, *t*-test), with no significant difference in the final OD₆₀₀ (0.64 ± 0.03 versus 0.69 ± 0.01 , $P < 0.05$, *t*-test). In BHIS medium (pH 6.0) at 30°C, *semi-synCG-A* exhibited a 1.47-fold longer doubling time compared with ATCC13032 (5.00 ± 0.27 h versus 3.41 ± 0.09 h, $P < 0.01$, *t*-test), with no significant difference in the final OD₆₀₀ (0.28 ± 0.00 versus 0.27 ± 0.00 , $P = 0.51$, *t*-test). In BHIS medium (pH 8.0) at 30°C, the doubling time increased by 2.28-fold (8.07 ± 0.38 h versus 3.53 ± 0.25 h, $P < 0.001$, *t*-test), and the final OD₆₀₀ decreased by 34% (0.89 ± 0.02 versus 1.51 ± 0.03 , $P < 0.001$, *t*-test). In BHIS medium under oxidative stress (H₂O₂) at 30°C, *semi-synCG-A* exhibited a 1.16-fold longer doubling time than that of ATCC13032 (4.24 ± 0.36 h versus 3.67 ± 0.12 h, $P < 0.05$, *t*-test) and a 30% decrease in the final OD₆₀₀ (0.45 ± 0.02 versus 0.65 ± 0.02 , $P < 0.001$, *t*-test). On CGXII plates and in CGXII medium (pH 7.0, at 30°C), *semi-synCG-A* also consistently exhibited slower growth. In CGXII medium, the doubling time of *semi-synCG-A* was 1.30-fold longer (5.33 ± 0.22 h versus 2.28 ± 0.03 h, $P < 0.001$, *t*-test), and the final OD₆₀₀ decreased significantly (0.14 ± 0.01 versus 0.76 ± 0.02 , $P < 0.001$, *t*-test). The causes of the observed growth discrepancies require further investigation (see Discussion). These results align with the findings from the fully recoded *E. coli* strain Syn61, which exhibited doubling times 1.3- to 2.5-fold longer than that of the starting strain MDS42 across the tested conditions [4].

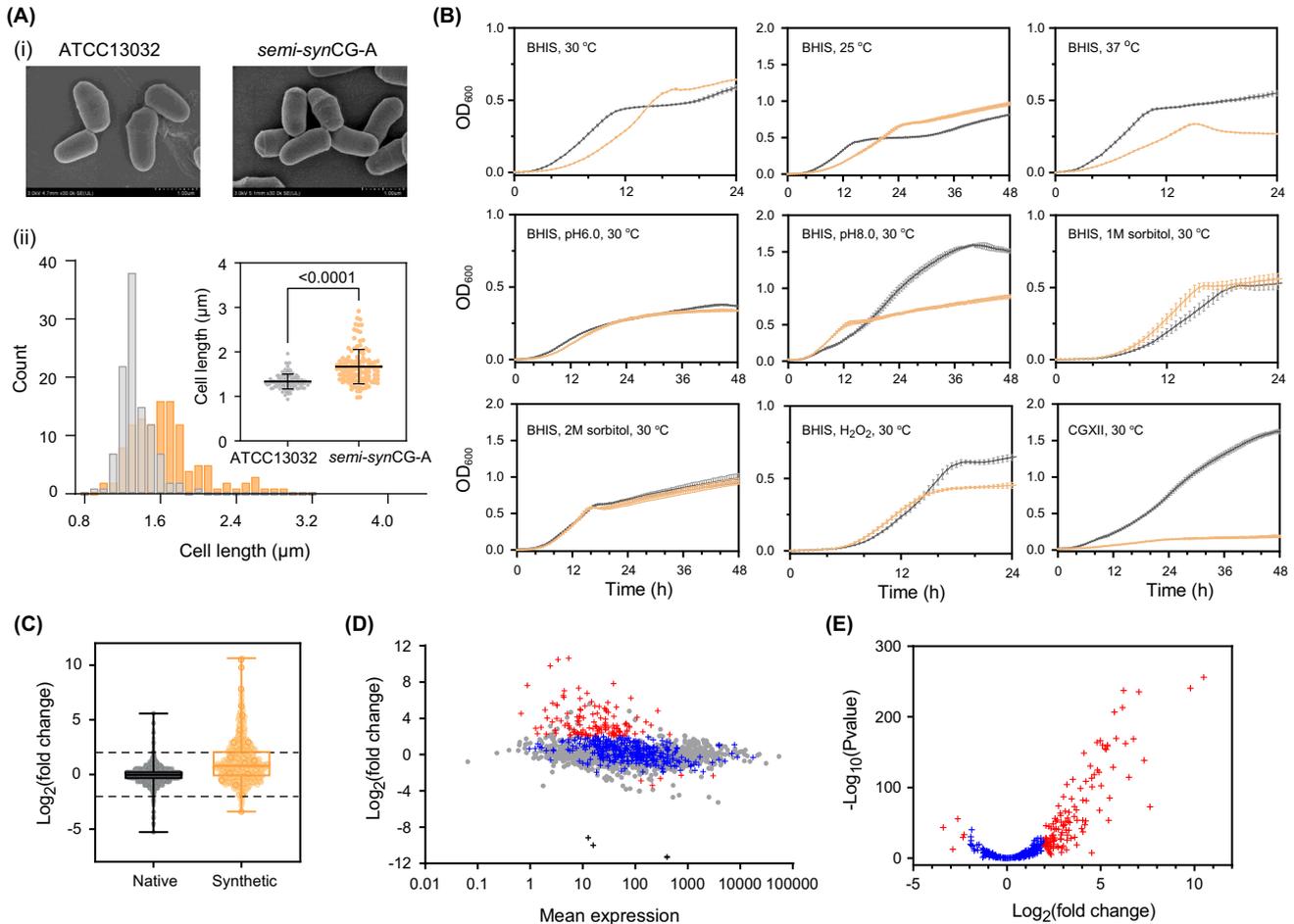


Figure 3. Comparative analysis of *semi-synCG-A* and ATCC13032 strains. (A) (i) Representative images of cell morphology of *semi-synCG-A* (yellow) and ATCC13032 (grey) observed under a cold-field emission scanning electron microscope. Scale bar: 1,00 μm . (ii) Cell lengths quantified from microscopy images (*semi-synCG-A*, length = $1.67 \pm 0.39 \mu\text{m}$, $n = 119$; ATCC13032, length = $1.34 \pm 0.17 \mu\text{m}$, $n = 108$, $P < 0.0001$, Mann–Whitney test); see also Data S1 in the supplemental information online. (B) Growth curves of *semi-synCG-A* and ATCC13032 under various conditions. The *semi-synCG-A* and ATCC13032 strains were grown in microplates for 24 or 48 h. The optical density (OD_{600}) values were measured in BHIS medium at different temperatures (30°C , 25°C , and 37°C), under different pH conditions (6.0 and 8.0), and under osmotic (1 M and 2 M sorbitol) and oxidative stresses (H_2O_2). Growth in CGXII medium was also measured, where a major growth defect was observed. Error bars are standard deviations ($N \geq 3$); see also Table S8 and Data S1 in the supplemental information online. (C–E) Differential gene expression analysis in *semi-synCG-A*. (C) Relative expression of genes within the native genome region (native, grey) and within megachunk A region (synthetic, yellow) in *semi-synCG-A* compared with ATCC13032. Dots represent relative expression levels of specific genes. Fold changes represent the expression rate of each gene in *semi-synCG-A* relative to wild-type ATCC13032 (*semi-synCG-A*, 1.16; ATCC13032, -0.06 ; $P < 0.001$, two-tailed t -test). (D,E), Black crossings represent deleted insertion sequence (IS) elements. Gray circles represent genes within the native genome region. Red crosses represent genes differentially expressed [absolute \log_2 (fold change) > 2 , $P_{\text{adj}} < 0.01$] within megachunk A, whereas blue crosses represent genes that are not differentially expressed within megachunk A [absolute \log_2 (fold change) ≤ 2 , or $P_{\text{adj}} < 0.01$]. Mean expression is indicated as transcripts per million (TPM). (E), P -value distribution of differentially expressed genes in megachunk A; see also Table S9 in the supplemental information online.

Transcriptomic profiling of *semi-synCG-A*

To determine whether gene expression was impacted by the incorporation of synthetic design features, whole-transcriptome sequencing (RNA-sequencing; RNA-seq) was performed to quantify and compare gene expression levels during the exponential phase in *C. glutamicum* ATCC13032 and *semi-synCG-A* (Figure 3C). Excluding the three deleted IS elements, the average transcript levels for genes in synthetic megachunk A increased, and the distribution of \log_2 (fold change) for all measured genes shifted upward accordingly (Figure 3C). In total, 27.2%

(128/470) of the recoded genes were significantly upregulated, and 1.1% (5/470) were significantly downregulated (>fourfold change, $P_{\text{adj}} < 0.01$) (Figure 3D,E and Table S9 in the supplemental information online). Smaller changes were observed by Church and colleagues, where 12.5% of the recoded genes (26/208) in *E. coli* with synonymous codon swaps in five 50-kb segments were significantly upregulated, with only one gene (0.5%) downregulated more than fourfold [10]. The extensive modifications affecting 96.8% (455 out of 470) of genes in our study likely contributed to the higher percentage of gene upregulation observed compared with the wild-type genome.

Twelve genes were more than 60-fold upregulated, six of which were in the phenylpropanoid degradation (*phd*) gene cluster (*phdT*, *phdA*, *phdB*, *phdC*, *phdD*, and *phdE*), with increases ranging from 65-fold to 1420-fold. This cluster encodes the pathway for utilization of phenylpropanoids in *C. glutamicum* [59]. Another four genes were *NCgl0058* (putative membrane protein), *NCgl0065* (putative transmembrane transport protein), *NCgl0020* (putative protease with chaperone function), and *crnT* (*NCgl0074*, creatinine transporter). The remaining *NCgl2945* and *NCgl2946* are two genes with small, hypothetical open reading frames [60]. Except for *NCgl0284* (*phdE*), all significantly upregulated genes incorporated at least one synthetic design feature.

The most significantly downregulated gene in megachunk A was *ioT2* (*NCgl2953*, a myo-inositol transporter [61]), with an 11-fold decrease in transcription. The other four genes were *oxiB* (*NCgl0168*, putative oxidoreductase), *NCgl0145* (putative glyoxalase), *NCgl2888* (putative membrane protein), and *NCgl0354* (putative acetyl transferase) [58].

A hybrid assembly using Illumina and PacBio data was also performed, which confirmed that no additional copies of synthetic fragments were detected and that the semi-synthetic genome was intact. Thus, there was no change in genome topology (Figure S10A in the supplemental information online). We also analyzed the methylation patterns in the synthetic 524-kb region of *semi-synCG-A*, finding that the methylation motifs were similar for both two strains (Figure S10B). We further examined methylation levels in the 200-bp upstream regions of the start codon for each of the 467 genes in both strains. Statistical analysis revealed no significant correlation between methylation changes and transcriptional levels (χ^2 test, $P < 0.05$; Figure S10C,D).

Discussion

We have developed an efficient approach, CACEXER, for iterative genomic replacement of up to 56 kb per step in *C. glutamicum*. This method leverages a newly developed CAC vector, joining BAC and YAC as a third versatile platform for large DNA manipulations for microorganisms. The CAC design principles are applicable for creating new shuttle vectors and accelerating genomic synthesis for other organisms. Furthermore, this CAC vector shows significant potential for incorporating exogenous genes as extrachromosomal functional modules [62–64], or for expressing large biosynthetic gene clusters (BGCs) with high GC content [65], which allows for complex functional acquisition and facilitates metabolic engineering.

Given that synthetic DNA fragments ranging from 30 to 136 kb were used for the genome synthesis of *E. coli* and yeast [2,4,10,25], we targeted a modest fragment size of 50 kb in this study, which is five times higher than the 10-kb fragments used in our previous work [28]. This increase in fragment size significantly accelerated the overall genome synthesis process for *C. glutamicum*. In addition, we preliminarily assessed the upper capacity of our CAC vector (pCGBAC1) and demonstrated its ability to carry an 80-kb synthetic DNA fragment (Figure S11 and Note S3 in the supplemental information online). Although the full replacement capacity with this larger construct remains untested, we will evaluate this potential in the next phase of genome synthesis.

During the course of our work, Chin and colleagues described a continuous genome synthesis (CGS) method for rapid integration of 500-kb synthetic DNA into the *E. coli* genome in only 10 days by using universal spacers on BAC and conjugation transfer [66]. We anticipate that these advances can be adapted to accelerate *C. glutamicum* genome synthesis. Furthermore, Cas9-mediated genomic excision of wild-type counterparts was exploited to enhance genome replacement efficiency during the construction of the *E. coli* Syn61 genome [4]. We envision a similar approach for *C. glutamicum* using I-SceI.

We detected significant transcriptomic changes and growth defects (to a lesser extent) in *semi-synCG-A*. We hypothesize that these changes likely result from both intragenic and intergenic effects caused by recoding, decoupling of overlapping genes, and insertion of loxPsym sites. Recent multi-omics co-profiling of recoded *E. coli* (Ec_Syn57) revealed that synonymous codon replacements can disrupt intragenic sequences, introducing transcriptional and translational noise, which results in fitness defects [11]. Similar fitness defects and decreased mRNA stability due to loxPsym site insertions were also reported in the Sc2.0 project [12]. One promising approach to address this issue of fitness defects was outlined in a recent preprint [11]. The authors divided the *E. coli* genome into 11 sections (150–451 kb), and used BAC carrying synthetic DNA sequences of ~50 kb to troubleshoot each section via adaptive laboratory evolution. These sections were then merged into a single, fully synthesized genome. We plan to adopt a similar strategy in future studies.

Concluding remarks

Since the first report of *de novo* genome synthesis ~15 years ago for *M. mycoides*, [8,9], remarkable progress has been achieved in genome synthesis for *E. coli* [4,10,11], *S. cerevisiae* [1,2,12], and more recently, higher organisms, such as *P. patens* [19]. However, a greater scale of genome synthesis is currently constrained by two foundational obstacles: (i) the efficient delivery of large synthetic sequences into target organisms; and (ii) the efficient recombination for genomic iterative replacements by these sequences.

C. glutamicum is widely used for manufacturing amino acids, organic acids, and bio-based chemicals. The development of both the CAC vector and the streamlined CACEXER system represents a critical step toward accelerating *de novo* genome synthesis for *C. glutamicum* (see [Outstanding questions](#)). This advance paves the way for creating redesigned genomes, enabling the development of more efficient cell factories for various bioproducts. In the short term, the CAC-based CACEXER system provides a powerful tool for inserting or replacing large DNA fragments in the *C. glutamicum* genome. This capability holds great potential for significantly advancing genome engineering of *C. glutamicum*, and optimizing strains for industrial production. Thus, our study provides a useful example for generalizing *de novo* genome design and synthesis across various industrially relevant microbes, particularly Gram-positive bacteria.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- METHOD DETAILS
 - Strains and growth conditions
 - Molecular cloning methods
 - Construction of shuttle vectors
 - Construction of helper plasmid
 - Assembly of CAC plasmids in yeast
 - Assembly of chunk A4-A10 in CAC plasmids

Outstanding questions

What is the upper delivery and replacement capacity for the CAC vector when used in conjunction with the CACEXER system?

How can the continuous genome synthesis approach developed by Chin and colleagues be adapted to facilitate the rapid integration of much larger synthetic DNA segments into the *Corynebacterium glutamicum* genome?

How can the SynOMICS strategy used by Church and colleagues be utilized to efficiently identify DNA fragments that cause fitness defects and correct these fragments in the synthetic *C. glutamicum* genome?

- Propagation and verification of CAC plasmids in *E. coli*
- Electroporation of CAC plasmids into *C. glutamicum*
- Iterative genome replacement for chunks A4-A10 using the CACEXER method
- Protein expression assays via SDS-PAGE and western blotting
- PCRTag analysis and extra isolation of the strain after recombination
- Stability analysis of CAC plasmid and synthetic genome
- Next-generation sequencing of *C. glutamicum* strains
- Scanning electron microscopy
- Serial dilution assays and growth curve assays
- RNA-seq analysis
- Hybrid assembly and epigenetic analyses by Illumina and SMRT sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to, and will be fulfilled by, the lead contact, Zhanglin Lin (zhanglinlin@gdut.edu.cn).

Materials availability

The *semi-syn*CG-A strain has been deposited with the China General Microbiological Culture Collection Centre (CGMCC; <https://cgmcc.net/>) under accession number CGMCC 1.65040. Requests for the generated plasmids and strains in this study should be directed to the lead contact, Zhanglin Lin (zhanglinlin@gdut.edu.cn).

Data and code availability

The data that support the findings of this study have been deposited into CNGB Sequence Archive (CNSA) of the China National GeneBank DataBase (CNGBdb) with accession number CNP0005795.

Author contributions

Conceptualization, Z. Lin, Y.Y., and Y.W.; funding and resources, Z. Lin, X.Y., Y.Y., and Y.S.; methodology: Z. Lin, Y.W., Y.Y., and Z.Z.; data production: Z.Z., P.H., and Z. Li; data analysis, investigation, and visualization, Z.Z., Y.W., B.L., Z. Li, P.H., T.C., Y.S., X.Y., Y.Y., and Z. Lin; supervision, Z. Lin, X.Y., Y.Y., and Y.W.; writing – original draft, Z.Z.; writing – review and editing, Z. Lin, Y.W., X.Y., and Y.Y.

Acknowledgments

We thank Guang Zeng, Minmin Zhong, Xiaowei Chen, and Xiaoyu Xie for preliminary experiments, and Yuan Huang and Liang Chen for protein expression assays. We thank GCATbio Co., Ltd and China National Genebank for support for DNA synthesis, bioinformatic analysis, and data storage. This study was supported by the National Key R&D Program of China (2018YFA0901000, 2022YFC2104800, and 2018YFA0900100), National Natural Science Foundation of China (No. 32322047), Natural Science Foundation of Guangdong Province of China (2024A1515030121, and 2023A1515012643), and Jiangsu Provincial Department of Science and Technology (No.BM2023009).

Declaration of interests

Z.Lin, Y.W., Y.Y., Z.Z., X.Y., P.H., Z. Li, T.C., and Y.S. have filed a patent application (CN2025101231482) to the China National Intellectual Property Administration for the work described in this study.

Declaration of generative Ai and AI-assisted technologies in the writing process

During the preparation of this work, we used ChatGPT solely to improve the readability of the manuscript draft. After using ChatGPT, we reviewed and edited the language as needed and take full responsibility for the content of the manuscript.

Supplemental information

Supplemental information associated with this article can be found online at <https://doi.org/10.1016/j.tibtech.2025.02.019>.

References

- Dymond, J.S. *et al.* (2011) Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature* 477, 471–476
- Richardson, S.M. *et al.* (2017) Design of a synthetic yeast genome. *Science* 355, 1040–1044
- Wang, K. *et al.* (2016) Defining synonymous codon compression schemes by genome recoding. *Nature* 539, 59–64
- Fredens, J. *et al.* (2019) Total synthesis of *Escherichia coli* with a recoded genome. *Nature* 569, 514–518
- Shen, Y. *et al.* (2016) SCRaMbLE generates designed combinatorial stochastic diversity in synthetic chromosomes. *Genome Res.* 26, 36–49
- Zhou, S. *et al.* (2023) Dynamics of synthetic yeast chromosome evolution shaped by hierarchical chromatin organization. *Natl. Sci. Rev.* 10, nwad073
- Hutchison 3rd, C.A. *et al.* (2016) Design and synthesis of a minimal bacterial genome. *Science* 351, aad6253
- Gibson, D.G. *et al.* (2008) Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* 319, 1215–1220
- Gibson, D.G. *et al.* (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* 329, 52–56
- Ostrov, N. *et al.* (2016) Design, synthesis, and testing toward a 57-codon genome. *Science* 353, 819–822
- Nyerges, A. *et al.* (2024) Synthetic genomes unveil the effects of synonymous recoding. *bioRxiv*, Published online June 14, 2024. <https://doi.org/10.1101/2024.06.16.599206>
- Zhao, Y. *et al.* (2023) Debugging and consolidating multiple synthetic chromosomes reveals combinatorial genetic interactions. *Cell* 186, 5220–5236
- Schindler, D. *et al.* (2023) Design, construction, and functional characterization of a tRNA neochromosome in yeast. *Cell* 186, 5237–5253
- Schindler, D. *et al.* (2024) Methodological advances enabled by the construction of a synthetic yeast genome. *Cell Rep. Methods* 4, 100761
- Lau, Y.H. *et al.* (2017) Large-scale recoding of a bacterial genome by iterative recombining of synthetic DNA. *Nucleic Acids Res.* 45, 6971–6980
- Venet, J.E. *et al.* (2019) Chemical synthesis rewriting of a bacterial genome to achieve design flexibility and biological functionality. *Proc. Natl. Acad. Sci. U. S. A.* 116, 8070–8079
- Bai, S. *et al.* (2024) Advances on transfer and maintenance of large DNA in bacteria, fungi, and mammalian cells. *Biotechnol. Adv.* 76, 108421
- James, J.S. *et al.* (2024) The design and engineering of synthetic genomes. *Nat. Rev. Genet.*, Published online November 6, 2024. <https://doi.org/10.1038/s41576-024-00786-y>
- Chen, L.-G. *et al.* (2024) A designer synthetic chromosome fragment functions in moss. *Nat. Plants* 10, 228–239
- Pampuch, M. *et al.* (2022) Towards synthetic diatoms: the *Phaeodactylum tricornutum* Pt-syn 1.0 project. *Curr. Opin. Green Sustain. Chem.* 35, 100611
- Trolle, J. *et al.* (2022) Resurrecting essential amino acid biosynthesis in mammalian cells. *eLife* 11, e72847
- Chen, Y. *et al.* (2022) Multiplex base editing to convert TAG into TAA codons in the human genome. *Nat. Commun.* 13, 4482
- Shen, Y. *et al.* (2017) Deep functional analysis of synII, a 770-kilobase synthetic yeast chromosome. *Science* 355, eaaf4791
- Xie, Z.-X. *et al.* (2017) “Perfect” designer chromosome V and behavior of a ring derivative. *Science* 355, eaaf4704
- Zhang, W. *et al.* (2023) Manipulating the 3D organization of the largest synthetic yeast chromosome. *Mol. Cell* 83, 4424–4437
- Asin-Garcia, E. *et al.* (2023) Metagenomics harvested genus-specific single-stranded DNA-annealing proteins improve and expand recombining in *Pseudomonas* species. *Nucleic Acids Res.* 51, 12522–12536
- Chai, M. *et al.* (2021) Synthetic biology toolkits and metabolic engineering applied in *Corynebacterium glutamicum* for biomanufacturing. *ACS Synth. Biol.* 10, 3237–3250
- Ye, Y. *et al.* (2022) Genomic iterative replacements of large synthetic DNA fragments in *Corynebacterium glutamicum*. *ACS Synth. Biol.* 11, 1588–1599
- Mitchell, L.A. *et al.* (2021) De novo assembly and delivery to mouse cells of a 101 kb functional human gene. *Genetics* 218, iyab038
- Venter, J.C. *et al.* (2022) Synthetic chromosomes, genomes, viruses, and cells. *Cell* 185, 2708–2724
- Suzuki, N. *et al.* (2005) Multiple large segment deletion method for *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 69, 151–161
- Wu, M. *et al.* (2020) Homing endonuclease I-SceI-mediated *Corynebacterium glutamicum* ATCC 13032 genome engineering. *Appl. Microbiol. Biotechnol.* 104, 3597–3609
- Robertson, W.E. *et al.* (2021) Creating custom synthetic genomes in *Escherichia coli* with REXER and GENESIS. *Nat. Protoc.* 16, 2345–2380
- Handelsman, J. *et al.* (2002) Cloning the metagenome: culture-independent access to the diversity and functions of the uncultivated microbial world. *Methods Microbiol.* 33, 241–255
- Hain, T. *et al.* (2008) Novel bacterial artificial chromosome vector pUVBBAC for use in studies of the functional genomics of *Listeria* spp. *Appl. Environ. Microbiol.* 74, 1892–1901
- Stalker, D.M. *et al.* (1981) Nucleotide sequence of the region of the origin of replication of the broad host range plasmid RK2. *Mol. Gen. Genet.* 181, 8–12
- Kolatka, K. *et al.* (2010) Replication and partitioning of the broad-host-range plasmid RK2. *Plasmid* 64, 119–134
- Kouprina, N. *et al.* (2006) Selective isolation of large chromosomal regions by transformation-associated recombination cloning for structural and functional analysis of mammalian genomes. In *YAC Protocols* (MacKenzie, A., ed.), pp. 85–101, Humana Press
- Shizuya, H. *et al.* (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl. Acad. Sci. U. S. A.* 89, 8794–8797
- Mukai, T. *et al.* (2020) Overcoming the challenges of megabase-sized plasmid construction in *Escherichia coli*. *ACS Synth. Biol.* 9, 1315–1327
- Li, T.Y. *et al.* (2004) Characterization of the replicon of a 51-kb native plasmid from the gram-positive bacterium *Leifsonia xyli* subsp. *cyndontis*. *FEMS Microbiol. Lett.* 236, 33–39
- Yin, P. *et al.* (2006) A Type I_B ParB protein involved in plasmid partitioning in a gram-positive bacterium. *J. Bacteriol.* 188, 8103–8108
- Amador, E. *et al.* (2000) A *Brevibacterium lactofermentum* 16S rRNA gene used as target site for homologous recombination. *FEMS Microbiol. Lett.* 185, 199–204
- Novikov, A.D. *et al.* (2016) Nucleotide sequence and structural analysis of cryptic plasmid pBL90 from *Brevibacterium lactofermentum*. *Russ. J. Genet.* 52, 1131–1136
- Jeffrey, V. and Joachim, M. (1991) New pUC-derived cloning vectors with different selectable markers and DNA replication origins. *Gene* 100, 189–194
- Debaugny, R.E. *et al.* (2018) A conserved mechanism drives partition complex assembly on bacterial chromosomes and plasmids. *Mol. Syst. Biol.* 14, e8516
- Roberts, D.M. (2023) A new role for monomeric ParA/Soj in chromosome dynamics in *Bacillus subtilis*. *MicrobiologyOpen* 12, e1344
- Böhm, K. *et al.* (2020) Chromosome organization by a conserved condensin-ParB system in the actinobacterium *Corynebacterium glutamicum*. *Nat. Commun.* 11, 1485
- Liu, J. *et al.* (2017) Development of a CRISPR/Cas9 genome editing toolbox for *Corynebacterium glutamicum*. *Microb. Cell Factories* 16, 205
- Cho, J.S. *et al.* (2017) CRISPR/Cas9-coupled recombining for metabolic engineering of *Corynebacterium glutamicum*. *Metab. Eng.* 42, 157–167
- Wang, B. *et al.* (2018) A RecET-assisted CRISPR-Cas9 genome editing in *Corynebacterium glutamicum*. *Microb. Cell Factories* 17, 63
- Baumgart, M. *et al.* (2013) Construction of a prophage-free variant of *Corynebacterium glutamicum* ATCC 13032 for use as a platform strain for basic research and industrial biotechnology. *Appl. Environ. Microbiol.* 79, 6006–6015

53. Muyrers, J.P. *et al.* (2000) RecE/RecT and Redalpha/Redbeta initiate double-stranded break repair by specifically interacting with their respective partners. *Genes Dev.* 14, 1971–1982
54. Linder, M. *et al.* (2021) Construction of an IS-free *Corynebacterium glutamicum* ATCC13032 chassis strain and random mutagenesis using the endogenous *ISCg1* transposase. *Front. Bioeng. Biotechnol.* 9, 751334
55. Tan, Y. *et al.* (2012) Construction of a novel sacB-based system for marker-free gene deletion in *Corynebacterium glutamicum*. *Plasmid* 67, 44–52
56. Wang, T. *et al.* (2019) An update of the suicide plasmid-mediated genome editing system in *Corynebacterium glutamicum*. *Microb. Biotechnol.* 12, 907–919
57. Kim, G.Y. *et al.* (2023) Synthetic biology tools for engineering *Corynebacterium glutamicum*. *Comput. Struct. Biotechnol. J.* 21, 1955–1965
58. Parise, M.T.D. *et al.* (2020) CoryneRegNet 7, the reference database and analysis platform for corynebacterial gene regulatory networks. *Sci. Data* 7, 142
59. Kallscheuer, N. *et al.* (2016) Identification of the *phd* gene cluster responsible for phenylpropanoid utilization in *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 100, 1871–1881
60. Baumgart, M. *et al.* (2018) *Corynebacterium glutamicum* chassis C1*: building and testing a novel platform host for synthetic biology and industrial biotechnology. *ACS Synth. Biol.* 7, 132–144
61. Bäumchen, C. *et al.* (2009) Myo-inositol facilitators Iot1 and Iot2 enhance d-mannitol formation from d-fructose in *Corynebacterium glutamicum*. *FEMS Microbiol. Lett.* 290, 227–235
62. Postma, E.D. *et al.* (2022) Modular, synthetic chromosomes as new tools for large scale engineering of metabolism. *Metab. Eng.* 72, 1–13
63. Ma, Y. *et al.* (2024) Convenient synthesis and delivery of a megabase-scale designer accessory chromosome empower biosynthetic capacity. *Cell Res.* 34, 309–322
64. Marques, F. *et al.* (2020) Engineering *Corynebacterium glutamicum* with a comprehensive genomic library and phage-based vectors. *Metab. Eng.* 62, 221–234
65. Liang, M. *et al.* (2022) Activating cryptic biosynthetic gene cluster through a CRISPR-Cas12a-mediated direct cloning approach. *Nucleic Acids Res.* 50, 3581–3592
66. Zürcher, J.F. *et al.* (2023) Continuous synthesis of *E. coli* genome sections and Mb-scale human DNA assembly. *Nature* 619, 555–562
67. Langmead, B. *et al.* (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25
68. McKenna, A. *et al.* (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303
69. Li, H. *et al.* (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079
70. Trapnell, C. *et al.* (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111
71. Anders, S. and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol.* 11, R106
72. Schäfer, A. *et al.* (1994) Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* 145, 69–73
73. Lin, Z. *et al.* (2021) Efficient genome editing for *Pseudomonas aeruginosa* using CRISPR-Cas12a. *Gene* 790, 145693
74. Jiang, Y. *et al.* (2015) Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl. Environ. Microbiol.* 81, 2506–2514
75. Keilhauer, C. *et al.* (1993) Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the *ilvB-ilvN-ilvC* operon. *J. Bacteriol.* 175, 5595–5603
76. Krzywinski, M. *et al.* (2009) Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645

STAR★METHODS

KEY RESOURCES TABLE

| Reagent or resource | Source | Identifier |
|---|-----------------------------|---|
| Antibodies | | |
| Anti-HA-HRP | Cell Signaling Technologies | Cat# 2999S |
| anti-FLAG-HRP (DYKDDDDK Tag) | Cell Signaling Technologies | Cat# 86861S |
| Bacterial and virus strains | | |
| <i>E. coli</i> DH5 α | Tsingke Biotech | Cat# TSC-C01 |
| <i>E. coli</i> NEB 10-beta | Biomed | Cat# BC401-01 |
| <i>S. cerevisiae</i> MYA3666 | Dr. Lin laboratory | N/A |
| <i>C. glutamicum</i> ATCC13032 | Dr. Lin laboratory | N/A |
| <i>C. glutamicum</i> semi-synCG-A1 | Ye <i>et al.</i> [28] | N/A |
| For a list of bacterial strains used in this study, see Table S1 in the supplemental information online | This paper | N/A |
| Chemicals, peptides, and recombinant proteins | | |
| His minus medium | FunGenome | Cat# YGM003A-4 |
| Q5 High-Fidelity DNA polymerase | New England Biolabs | Cat# M0491L |
| KOD One Mix DNA polymerase | Toyobo | Cat# KMM-201 |
| KOD FX polymerase | Toyobo | Cat# KFX-101 |
| PBS (pH7.4) | Gibco | Cat# 10010023 |
| Critical commercial assays | | |
| HiPure Gel Pure DNA Micro kit | Magen Biotech | Cat# D2110-03 |
| NucleoSpin Gel and PCR Clean-up kit | Macherey-Nagel | Cat# 740609.50 |
| TIANprep Mini Plasmid Kit | Tiangen | Cat# DP103-03 |
| Frozen-EZ Yeast Transformation II kit | Zymo Research | Cat# T2001 |
| TIANamp Yeast DNA kit | Tiangen | Cat# DP307-02 |
| NucleoBond BAC100 kit | Macherey-Nagel | Cat# 740579.10 |
| TIANamp Bacteria DNA kit | Tiangen | Cat# DP302-02 |
| Quick-Load 1 kb Extended DNA Ladder | New England Biolabs | Cat# N3239S |
| DL 15,000 DNA Marker | Takara | Cat# 3582A |
| Broad Multi Color Pre-Stained Protein Standard | Genscript | Cat# M00624 |
| Deposited data | | |
| NGS sequencing data | This study | https://db.cngb.org/cnsa/ ; Accession number: CNP0005795 |
| RNA-seq data | This study | https://db.cngb.org/cnsa/ ; Accession number: CNP0005795 |
| PacBio and Illumina data | This study | https://db.cngb.org/cnsa/ ; Accession number: CNP0005795 |
| Synthetic megachunk A sequence files | This study | https://db.cngb.org/cnsa/ ; Accession number: CNP0005795 |

(continued on next page)

(continued)

| Reagent or resource | Source | Identifier |
|--|-----------------------------|---|
| CAC plasmid files | This study | https://db.cngb.org/cnsa/ ; Accession number: CNP0005795 |
| Oligonucleotides | | |
| For a list of primers used in this study, see Table S2 in the supplemental information online | This paper | N/A |
| For a list of PCRTAG primers used in this study, see Table S7 in the supplemental information online | This paper | N/A |
| Recombinant DNA | | |
| For a list of plasmids used in this study, see Table S3 in the supplemental information | This study | N/A |
| Software and algorithms | | |
| Bowtie2 v2.2.5 | Langmead <i>et al.</i> [67] | https://bowtie-bio.sourceforge.net/bowtie2/index.shtml |
| GATK3.8 v2.7 | McKenna <i>et al.</i> [68] | https://github.com/broadinstitute/gatk/releases |
| SAMtools v0.1.19 | Li <i>et al.</i> [69] | https://github.com/samtools/samtools |
| TopHat v2.1.1 | Trapnell <i>et al.</i> [70] | http://ccb.jhu.edu/software/tophat |
| DSeq2 v1.30.1 | Anders <i>et al.</i> [71] | https://bioconductor.org/packages/release/bioc/html/DSeq2.html |

METHOD DETAILS

Strains and growth conditions

E. coli DH5 α was used as cloning host, while *E. coli* NEB 10-beta was used for CAC plasmid propagation. *C. glutamicum* strain semi-synCG-A1 was used as the starting strain for genome modifications. The strains used in this study are listed in Table S1 in the supplemental information online.

E. coli strains were cultivated at 37°C in lysogeny broth (LB) medium. *C. glutamicum* strains were cultured at 30°C in BHIS medium for genome replacements and in Neural complex medium (NCM) for preparation of electrocompetent cells [28]. Antibiotic and chemicals concentrations were as followed: 34 μ g/ml chloramphenicol, 50 μ g/ml kanamycin, 50 μ g/ml spectinomycin and 5 μ g/ml tetracycline for *E. coli*; 10 μ g/ml chloramphenicol, 20 μ g/ml streptomycin, 20 μ g/ml kanamycin, 150 μ g/ml spectinomycin, 1.5 μ g/ml tetracycline, 10-20% (w/v) sucrose, and 1.25 mM 4-chloro-DL-phenylalanine for *C. glutamicum*.

Yeast assemblies were performed in *S. cerevisiae* strain VL6-48 (ATCC no. MYA3666). Yeast cells were cultivated at 30°C in either yeast extract peptone dextrose (YPD) medium (10 g/l yeast extract, 20 g/l peptone, and 20 g/l glucose) or SD-His medium (Synthetic dropout medium lacking Histidine, with addition of 20 g/l glucose; FunGenome, China).

Molecular cloning methods

PCR amplification reactions were performed using Q5 High-Fidelity DNA Polymerase (New England Biolabs) or KOD One Mix DNA Polymerase (Toyobo, Japan). Restriction enzymes and Gibson assembly components were purchased from New England Biolabs. DNA products were purified using the HiPure Gel Pure DNA Micro Kit (Magen Biotech, China) or Nucleospin Gel and PCR Clean-up kit (Macherey-Nagel, Germany). Plasmids were isolated from overnight cultures using the TIANprep Mini Plasmid Kit (Tiangen, China) according to the manufacturer's instructions unless otherwise noted. Oligonucleotide synthesis and Sanger sequencing were outsourced to Sangon Biotech (Shanghai, China), RuiBiotech (Beijing, China), or Tianyi Huiyuan (Wuhan, China).

Construction of shuttle vectors

The plasmid pRSII313 was purchased from Addgene (Cat. 35449). The plasmids pBeloBAC11 and pK18*mobsacB* [72] were purchased from Miaoling Biology (Wuhan, China). The RK2 *oriV* and *trfA* genes were originated from pCas12a- λ Red [73]. The 5-kb replicative sequence from pCXC100 (GenBank No. AY380839.1, bases 1-4992) and the plasmid pOK12BL90 containing a second

replicative sequence from pBL90 (GenBank No. KU306397.1, bases 53509-57768) were respectively synthesized by Generay Biotech (Shanghai, China). The pOK12CXC100 and pCGBACYT90 were constructed via Gibson assembly, respectively. The YAC origin and the *HIS3* marker were amplified from pRSII313. The BAC origin and the *cat* marker were amplified from pBeloBAC11. The RP4 *oriT* fragment was amplified from pK18*mobsacB*. The pBL90 replicon II fragment was amplified from pBL90. These four fragments were used for assembly of pCGBACYT90 via Gibson assembly. The pCGBAC1 was generated by inserting *parS* site from the genome of *C. glutamicum* ATCC13032 (5'-tggttcacgtgaaca-3') into the original pCGBACYT90, replacing the DNA sequence of *parA* II and *parB*. Primer sequences are listed in Table S2 in the supplemental information online. Plasmids used in this study are listed in Table S3 in the supplemental information online.

Construction of helper plasmid

The mitochondrial intron-encoded endonuclease from *S. cerevisiae* I-SceI gene (UniProt entry P03882) was codon-optimized and synthesized by Sangon Biotech (Shanghai, China). The P_{BAD} promoter was amplified from pCas [74], which is a gift from Dr Sheng Yang. For constructing pXMJ19-RecET-I-SceI, the fragment P_{BAD} -I-SceI was cloned into the plasmid pXMJ19-recET [28] via Gibson assembly. Primer sequences are listed in Table S2.

Assembly of CAC plasmids in yeast

The CAC plasmids were assembled using yeast transformation-associated recombination (TAR) [33]. Specifically, for pCGBAC1-A1 construction, we labelled the 55.1 kb redesigned synthetic sequence [28] as chunk A1. The CAC plasmid was assembled from seven PCR products amplified from template plasmids, plus two amplified CAC backbone fragments using pCGBAC1 as template. Yeast transformation was performed using Frozen-EZ Yeast Transformation II kit (Zymo Research) according to manufacturer's instructions. 200 ng of each fragment were added to 100 μ l of competent cells in a 1.5 mL microcentrifuge tube and the mixture was incubated in 30°C for 2 hours before plating on SD-His media. Colonies were picked, restreaked on fresh SD-His plate and grown for 1-2 day at 30 °C. Assembly was verified via colony PCR across all the junctions. Single colonies were lysed in 40 μ l of 0.1% (w/v) NaOH, at 95 °C for 30 min, and the supernatant was used for PCR with KOD FX polymerase (Toyobo, Japan). The PCR conditions were: 94°C for 3 min; 30 cycles of 98°C for 10 seconds, 60°C for 30 seconds, 68°C for 1 minute per kb; followed by a final extension at 68°C for 5 minutes. Verified colonies were inoculated for CAC plasmid extraction. Details for the constructing CAC plasmids for CACEXER are described in the section 'Assembly of chunk A4-A10 in CAC plasmid'.

Assembly of chunk A4-A10 in CAC plasmids

The entire genome, including chunk A1, was computationally redesigned [28], resulting in the division of the designer synthetic sequence into six megachunks (megachunk A-F, 495-525 kb each). In this study, megachunk A was divided into nine chunks (A2-A10), each 52-56 kb in size. Chunks A3-A10 were further subdivided into 90 minichunks of 4-5 kb, each overlapping by 60-66 bp. Chunk A2 was subdivided into 97 segments of approximately 600 bp, each overlapping by 35-40 bp. The synthesis of 97 fragments of chunk A2 and the 90 minichunks of chunks A3-A10 was outsourced to BGI.

The assembly of the CAC plasmids, each containing synthetic DNA either around 25 or 50 kb in length, was as followed. The following positive and negative selection markers were used: *rpsL* (-1, streptomycin sensitivity), *kanR* (+1, kanamycin resistance), *sacB* (-2, conferring sucrose sensitivity), and *speR* (+2, spectinomycin resistance). The *rpsL*, *kanR* and *speR* genes has been used previously [28]. The *sacB* gene was amplified from the plasmid pK18*mobsacB*. The codon-optimized *sacB* gene and *pheS** (*pheS*^{T262A_A309G}, -4, conferring 4-chloro-DL-phenylalanine sensitivity) gene were synthesized by RuiBiotech (Beijing, China; Supplemental Note S1). The double-selection cassette $P_{tur}pheS^*-P_{tur}sacB^{Cg}-SpeR$ was obtained by overlapping PCR.

The design of the sequence flanking the synthetic DNA was as followed. On the 5' side, the synthetic DNA was flanked by a 500-1000 bp upstream homologous arm (UHA) for replacement in *C. glutamicum*, and one I-SceI excision site. On the 3' side, the synthetic DNA was flanked by a double-selection cassette *rpsL-KanR*, *sacB-SpeR*, or $P_{tur}pheS^*-P_{tur}sacB^{Cg}-SpeR$, a 500-1000 bp downstream homologous arm (DHA), and another I-SceI excision site. Additionally, a different negative selection marker was inserted into the CAC plasmid backbone for enhanced vector curing.

For assembling a complete chunk or two subchunks, 5-12 fragments of minichunk synthetic DNA, and one pre-constructed fragment containing the double-selection marker and the DHA, were generated by PCR amplification or enzyme digestion. Each

fragment contained 60–80 bp of homology to its adjacent fragments. The CAC plasmid backbone was amplified from pCGBAC1 in two fragments. These fragments were assembled via TAR as described in the ‘Assembly of CAC plasmids in yeast’ section. Specifically, chunks A4, A5, A6 and A9 were assembled either as a complete chunk or as two subchunks, while chunks A7, A8 and A10 were assembled as complete chunks.

Propagation and verification of CAC plasmids in *E. coli*

The CAC plasmids were propagated in *E. coli* NEB 10-beta [33]. Total DNA from a 10 mL *S. cerevisiae* culture was extracted using the TIANamp Yeast DNA Kit (Tiangen, China) following manufacturer’s instructions. The extracted DNA was diluted to approximately 50 ng/μl in Milli-Q H₂O, and 10 μl was electroporated into 100 μl of competent *E. coli* NEB 10-beta cells using a Gene Pulser Xcell (Bio-Rad, USA) set at 2.4kV, 25μF, 200Ω. After electroporation, cells were recovered for 1 hour at 37 °C before plating on selective plates. Single colonies were restreaked and verified via colony PCR at all CAC plasmid junctions. Verified colonies were inoculated in 5 ml LB with appropriate antibiotics, and then transferred to 100–400 ml LB with antibiotics and grown at 37°C until OD₆₀₀ exceeded 2.0. The CAC plasmids were purified using the NucleoBond BAC 100 kit (Macherey-Nagel, Germany) according to the manufacturer’s instructions (Maxi) [40]. Purified DNA was dissolved in MilliQ-H₂O and subjected to enzyme digestion for verification and quantification. Specifically, digested bands were visualized on a 0.6% agarose gel for size determination and approximate DNA mass estimation, using Quick-Load 1 kb Extend DNA Ladder (New England Biolabs) and DL 15,000 DNA Marker (TaKaRa, China) was used as size markers. Electrophoretic analysis of band intensities was performed using ImageJ (NIH).

Electroporation of CAC plasmids into *C. glutamicum*

Electrocompetent cells were prepared as previously described [28]. For efficient CAC electroporation into *C. glutamicum*, more than 100 ng (up to 10 μl) of CAC was electroporated into 100 μl of electrocompetent *C. glutamicum* cells at a final voltage of 2500 V and a pulse time of 5.0 ms. Cells were immediately transferred into 900 μl of pre-warmed BHIS medium and heat-shocked at 46°C for 6 minutes. Subsequently, the cells were incubated at 220 rpm at 30°C for 1.5–4.5 h. Finally, cells were spread on BHIS agar plates containing chloramphenicol and either kanamycin or spectinomycin. The transformants were restreaked, grown overnight on agar plates with corresponding antibiotics, and screened via colony PCR for all the synthetic PCRTags on the synthetic DNA. PCRTag primers and their expected amplicon lengths are listed in Table S7 in the supplemental information online.

Iterative genome replacement for chunks A4–A10 using the CACEXER method

Strain *semi-synCG-A3* was used as the starting strain. Details on construction of *semi-synCG-A3* are in Supplemental Note S2. Genome replacement of chunks A4–A10 was performed using the CACEXER method. The assembled CAC plasmids were electroporated into recipient *C. glutamicum* cells. Cells harboring CAC plasmids with verifications of junctions were inoculated into 10 mL BHIS containing relevant antibiotics and cultivated overnight at 30°C. Then, 1% of the overnight culture was inoculated into 100 mL BHIS medium with antibiotics. When OD₆₀₀ reached 0.3, 0.5% (w/v) L-arabinose (inducing I-SceI) and 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG, inducing RecET) were supplemented, and the culture was incubated for 4 hours. After centrifugation, the cells were resuspended in sterile Milli-Q H₂O and spread on BHIS plates with selection for the helper plasmid (+3, chloramphenicol), the positive selection marker adjacent to synthetic sequence (+1, kanamycin or +2, spectinomycin) and agents selecting against the negative marker both on the genome and on the CAC backbone (-2, sucrose or -1, streptomycin). After 2–3 days of incubation at 30°C, up to 500 single colonies were restreaked on BHIS plates with relevant antibiotics for phenotyping and cultivated overnight.

Protein expression assays via SDS-PAGE and western blotting

The plasmids pXMJ19-tag-RecET-I-SceI and pXMJ19-control were constructed using pXMJ19-RecET-I-SceI as the template. The *semi-synCG-A3* strain carrying pCGBAC1-A4 was electroporated with pXMJ19-RecET-I-SceI, pXMJ19-tag-RecET-I-SceI and pXMJ19-control, respectively, and the resulting strains were used for protein expression assays. After 4 hours of induction, cell pellets were harvested and re-suspended in phosphate-buffered saline (PBS, pH7.4, Gibco). Equal amounts of cells (OD₆₀₀=2.5) were sonicated on ice using an Ultrasonic crusher (Scientz JY92-IIIN, Ningbo, China), followed by centrifugation at 15,000 *g* and 4°C for 20 minutes. The supernatants were diluted and mixed with 6× protein loading buffer (TransGene, Beijing, China) and boiled for 10 min. For Western blotting, total protein samples were separated by 9% or 13% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), followed by electrophoretic transfer to nitrocellulose membranes on ice, at a constant voltage of 80V for 150 minutes. The membranes were stained with Ponceau S, blocked with 5% milk in Tris-buffered saline with Tween 20 (TBST) for 2 h on a shaker at room temperature. Protein signals were detected using Western ECL Substrate (Bio-Rad). The

following antibodies were used: anti-HA-HRP (Cell Signaling Technologies, #2999S) and anti-FLAG-HRP (DYKDDDDK Tag, Cell Signaling Technologies, #86861S).

PCRTag analysis and extra isolation of the strain after recombination

The single colonies with correct antibiotic phenotype were picked and subjected to colony PCR as previously described [28]. The colony PCR targeted the 5' genomic integration sites to verify the loss of the upstream genomic double-selection marker, and 3' genomic integration sites to verify the gain of the downstream genomic double-selection marker, and a junction on CAC backbone to verify its loss. Colony PCR products were sequenced via Sanger sequencing to confirm the correct integration residual excision sites on the genome. Only correctly verified clones were inoculated into BHIS, cultivated overnight, and restreaked on BHIS plates with chloramphenicol. Single colonies on the restreaked plate were randomly picked for PCRTag analysis. This additional round of overnight culture and PCRTag analysis was performed to 1) eliminate residual CAC plasmid backbone, preventing false positive transformants in the subsequent electroporation, and 2) avoid 'patchwork' incomplete replacement of PCRTags or the undesired cutting and/or reannealing of the CAC plasmid [28,33,66]. PCRTag primers and their expected amplicon lengths are listed in Table S7.

Stability analysis of CAC plasmid and synthetic genome

The strain *C. glutamicum* ATCC13032 with CAC plasmid pCGBAC1-A1 was streaked on BHIS plate added with chloramphenicol and the *semi-synCG-A* was streaked on BHIS plate added with streptomycin for 2 days at 30°C. Three independent single colonies from each strain were selected for successive passage for approximately 130 generations. These were then plated on BHIS plates with corresponding antibiotics for 2 days, followed by restreaking of three single colonies from each initial isolate on fresh BHIS plates with corresponding antibiotics overnight. PCRTag analysis was performed on the isolated single colonies. For the CAC plasmid pCGBAC1-A1, 14 PCRTags (2 PCRTags in each of seven subchunks) were selected to access the integrity of synthetic fragments in the absence of SCRaMbLE. For genome stability testing of *semi-synCG-A*, 22 pairs of PCRTags (one pair located in every 25 kb) were chosen to detect the loss of different segments in the absence of SCRaMbLE. Total DNA extractions were performed as described earlier, and paired-end whole genome sequencing was performed by GenePlus (Beijing, China).

Next-generation sequencing of *C. glutamicum* strains

Partial synthetic *C. glutamicum* strains confirmed by PCRTag analysis were subjected to next-generation sequencing (NGS). Total DNA was extracted using the TIANamp Bacteria DNA kit (Tiangen, China) following the manufacturer's instructions. Library construction and paired-end whole genome sequencing were performed by Sangon Biotech (Shanghai, China) or GenePlus (Beijing, China) using DNBSEQ-T7 platform. The sequencing reads QC, data processing and analysis were performed as described previously [23]. The sequencing reads were mapped to a reference sequence of the synthetic *C. glutamicum* genome using Bowtie2 with default parameters [67]. Both GATK3.8 [68] and SAMtools [69] pipelines were used to identify the variants.

Scanning electron microscopy

The preparation of the samples of *C. glutamicum* strains ATCC13032 and the *semi-synCG-A* was performed as described [28]. Cell morphology was imaged using a cold-field emission scanning electron microscope SU8000 (SEM; Hitachi, Japan).

Serial dilution assays and growth curve assays

C. glutamicum strains ATCC13032 and the *semi-synCG-A* were cultured overnight in BHIS medium at 30°C, collected by centrifugation, and resuspended in PBS (pH7.4). Following 1:10 serial dilutions, serial dilution assays were performed in different agar plates and incubated for 1-2 days as previously described [28]: BHIS (0.5 M sorbitol, pH7.5) at 30°C, 25°C, or 37°C; BHI with 1 M and 2 M of sorbitol at 30°C (testing for osmotic stress); BHIS adjusted to pH 6.0 and 8.0 using HCl and NaOH (testing for acid or alkali stress) at 30°C; BHIS at 30°C, where 1 mM hydrogen peroxide (testing for oxidative stress) was added to overnight cultures to treat the cells for three hours; CGXII agar plate (pH7.0) at 30°C [75]. Growth curve assays in liquid medium were performed as described [28]. Overnight cultures were diluted to an initial OD₆₀₀ of ~0.05 in 300 µl liquid medium and grown at 30°C in 100-well Honeycomb microplates using a Bioscreen C automated turbidimeter (Oy Growth Curves Ab Ltd) for continuous monitoring of OD₆₀₀ for 24-48 h. The growth conditions included: BHIS medium (0.5 M sorbitol, pH7.5) at 30°C, 25°C, or 37°C; BHI medium with 1 M and 2 M of sorbitol at 30°C; BHIS medium adjusted to pH 6.0 and 8.0 using HCl and NaOH at 30°C; BHIS medium with 1 mM hydrogen peroxide at 30°C, CGXII medium (pH 7.0) at 30°C. The growth curve assays were repeated three times, and the data were analyzed and plotted using GraphPad Prism 9.

RNA-seq analysis

C. glutamicum strains ATCC13032 and *semi-synCG-A* were cultured overnight in BHIS medium at 220 rpm at 30°C. The cultures were then inoculated into 100 ml BHIS medium and grown to an OD₆₀₀ of 0.8. Cells were harvested by centrifugation, washed twice with cold PBS (pH7.4), and then snap-frozen in liquid nitrogen. RNA extraction, library construction, and sequencing (RNA-seq) were outsourced to Novogene (Beijing, China) using the Illumina NovaSeq X Plus platform. RNA-seq was performed in biological triplicates for each strain. The sequencing reads QC, data processing and analysis were performed as described previously [23]. The alignment and quantification of the reads were analyzed by TopHat [70], while the differential expression of genes was examined with DEseq2 [71].

Hybrid assembly and epigenetic analyses by Illumina and SMRT sequencing

Genomic DNA extraction, library construction and sequencing of both ATCC13032 and *semi-synCG-A* were outsourced to Novogene (Beijing, China) using PacBio Sequel IIe and Illumina NovaSeq X Plus platforms. Low-quality reads were filtered using SMRT Link v8.0 and the remaining reads were assembled into a circular genome without gaps using Canu (<https://github.com/marbl/canu/>, version 2.0). Genome-wide base modification analyses were performed by ipdSummary of SMRTv13.0, and methyltransferase motifs analyses were performed using motifMaker embedded in the SMRT Link v13.0 (<https://www.pacb.com/smart-link/>) with filtering quality score ≥ 30. Circular genome map was generated using Circos, version 0.64 [76], to show the distribution of methylation sites. Methylation levels in 200-bp upstream of the start codon of genes in the megachunk A were analyzed in both strains. Chi-square tests were performed to assess the relationship between methylation changes and significant transcriptional alterations.

QUANTIFICATION AND STATISTICAL ANALYSIS

Various statistical tests were employed to calculate *P* values, as indicated in the text, figure legends, or STAR METHODSs section, where appropriate. In general, results were considered statistically significant when $P < 0.001$, unless stated otherwise. Specifically, in transcriptional analysis, the genes were assessed for statistical significance if absolute log₂ (Fold Change) > 2 and adjusted *P* value < 0.01.