

REVIEW ARTICLE

Advances in large DNA fragment assembly for microbial cell factory engineering

Yu Zhang^{1,2,3} | Wenqian Liu^{1,2,3} | Dongyuan Cheng^{2,3} | Jiazhang Lian¹ | Xian Fu^{3,4} | Yue Shen^{3,4}

¹Key Laboratory of Biomass Chemical Engineering of Ministry of Education & State Key Laboratory of Biobased Transportation Fuel Technology, College of Chemical and Biological Engineering, Zhejiang University, Hangzhou, China

²State Key Laboratory of Genome and Multi-omics Technologies, BGI Research, Hangzhou, China

³BGI Research, Changzhou, China

⁴Guangdong Provincial Key Laboratory of Genome Read and Write, BGI Research, Shenzhen, China

Correspondence

Yue Shen.

Email: shenyue@genomics.cn

Funding information

The National Key Research and Development Program of China, Grant/Award Number: 2024YFA0916500; the National Natural Science Foundation of China, Grant/Award Numbers: 32322047, 32471490; Jiangsu Provincial Department of Science and Technology, Grant/Award Number: BM2023009

Abstract

The efficient, rapid, and reliable assembly of DNA fragments is essential for advancing metabolic engineering and synthetic biology. With the rapid advancement of DNA synthesis and assembly technologies, the scale of DNA assembly has expanded from single genes to metabolic pathways and even genomes. Large DNA fragment assembly, in particular, has developed into a pivotal tool for microbial cell factory engineering, significantly contributing to the understanding and construction of biological systems. This review summarizes recent advancements in large DNA fragment assembly methods, encompassing in vitro and in vitro methods for multiple-gene assemblies, as well as genome-scale technologies, including synthetic genome and neochromosome construction. We systematically compare their key features in terms of assembly principle, capacity, and efficiency. Especially, we highlight their applications in microbial cell factory engineering, including heterologous pathway construction, reprogramming host metabolism, and expanding complex biosynthetic networks. Finally, we discuss the challenges and prospects of applying large DNA fragment assembly to advance cell factories. In summary, this review provides a theoretical and technical framework for engineering high-performance microbial cell factories, contributing to the advancement of industrial biomanufacturing.

KEYWORDS

in vitro and in vitro assembly, large DNA fragment assembly, microbial cell factories, neochromosome, synthetic genome

1 | INTRODUCTION

The construction of microbial cell factories has revolutionized industrial biomanufacturing, enabling sustainable production of high-value compounds, including pharmaceuticals, biofuels, and fine chemicals [1]. Building efficient microbial cell factories typically involves: (1) introducing heterologous metabolic pathways larger

than 10 kilobases (kb) (e.g., incorporating dozens of plant-, bacterial-, and fungal-derived genes to produce alkaloids, taxol, and other natural products [2–4]); (2) reprogramming host metabolism at the genome scale (megabase [Mb] level); and (3) expanding complex biosynthetic networks (often requiring DNA constructs spanning hundreds of kb to Mb [5–7]). These requirements highlight the urgent need for advanced

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Quantitative Biology* published by John Wiley & Sons Australia, Ltd on behalf of Higher Education Press.

DNA assembly technologies, especially large DNA fragment (ranging in length from tens of kb to Mb) assembly. This technological capability serves as a foundation for engineering microbial cell factories in the aforementioned aspects.

The large DNA fragment is assembled through a hierarchical splicing strategy using synthetic short oligonucleotides (oligos), which relies heavily on low-cost and large-scale oligo synthesis [8]. Recent advances in DNA synthesis technologies have significantly promoted DNA synthesis quality, throughput, efficiency, and cost reduction. For example, high-density, integrated chip-based DNA synthesis technologies have cut costs and increased throughput exponentially [8–10]. Oligo assembly technologies enable the assembly of short oligos into gene-length double-stranded DNA (dsDNA) (kb level) [11]. Engineered terminal deoxynucleotidyl transferases further enable longer, high-fidelity single-stranded DNA (ssDNA) synthesis [12, 13]. These improvements provide higher-quality and more cost-effective materials for large DNA fragment assembly, thereby substantially advancing the development of DNA assembly methodologies.

Over the past two decades, large DNA fragment assembly has improved in length, accuracy, speed, etc., and a proliferation of novel methods has emerged—spanning in vitro and in vitro assembly, synthetic genomes, and neochromosome (NeoChr) assembly—providing solid technical support for the engineering of microbial cell factories (Figure 1A,B). As summarized in Table 1, traditional and emerging in vitro methods are increasingly excelling in combining standardized and modularized biological parts, transcription units (TU), and multigene pathways [67]. However, due to the instability of in vitro manipulation for large DNA fragments, stable assembly of longer sequences—particularly natural product biosynthetic pathways or biosynthetic gene clusters (BGCs) ranging from 10 kb to Mb—still relies heavily on in vitro assembly approaches [68, 69]. Recent advancements, such as *Escherichia coli* (*E. coli*) genome recoding and the Synthetic Yeast Genome Project (Sc2.0) of *Saccharomyces cerevisiae* (*S. cerevisiae*), have expanded the assembly capability to the genome scale [46, 48], facilitating genome-wide strain engineering. Further innovations, such as the de novo synthesis of supernumerary, orthogonal NeoChrs, push assembly boundaries even further and enable the expansion of metabolic networks [6]. Despite these breakthroughs in large DNA fragment assembly technology, comprehensive reviews on their applications in microbial cell factory development remain limited.

In this review, we systematically summarize recent advances in large DNA fragment assembly technologies, focusing on their applications in microbial cell factory engineering, with particular emphasis on industrial workhorses (e.g., *E. coli* and *S. cerevisiae*).

From in vitro assembly to in vitro assembly and further to synthetic genome assembly and supernumerary designer NeoChr construction, we introduce their key features in terms of assembly principle, capacity, and efficiency and highlight their respective strengths in microbial cell factory engineering. Finally, we discuss the current limitations and future directions in the application of large DNA fragment assembly technology.

2 | IN VITRO DNA ASSEMBLY METHODS FOR HETEROLOGOUS PATHWAY CONSTRUCTION

In vitro DNA assembly has become a central tool for reconstructing heterologous metabolic pathways in microbial systems (Figure 2A). The complexity of large DNA constructs necessitates efficient assembly to maintain stability and improve construction throughput [70]. Restriction digestion-based cloning established standardized frameworks for modular pathway construction, yet its reliance on specific recognition sites limited sequence flexibility. Homology-directed assembly enabled seamless and flexible fragment joining, overcoming sequence limitations. The emergence of programmable nuclease-guided systems and automated platforms has further enhanced precision and scalability, advancing the construction of complex biosynthetic pathways.

2.1 | Restriction digestion-based methods for standardized and modular pathway construction

Restriction digestion-based DNA assembly remains a foundational approach for constructing modular genetic circuits and metabolic pathways. By exploiting the sequence-specific cleavage properties of Type II restriction endonucleases, researchers can systematically join genetic elements in a predictable and reusable manner, enabling standardized design of multigene constructs for microbial cell factories [71]. Early systems, such as the BioBrick assembly (Figure 2B), established a framework for modular cloning (MoClo), defining fixed junction sequences between DNA parts to ensure interoperability and reusability [14]. To overcome limitations associated with the scar sequences and directional constraints of the original standard, BglBrick was developed (Figure 2B). This system employs *Bgl* II and *Bam*HI restriction enzymes, which generate a 6-nucleotide scar encoding a glycine-serine linker, a benign peptide that does not interfere with most protein fusions [16]. This design ensures directional assembly, compatibility with existing BioBrick parts, and stepwise ligation of multiple fragments

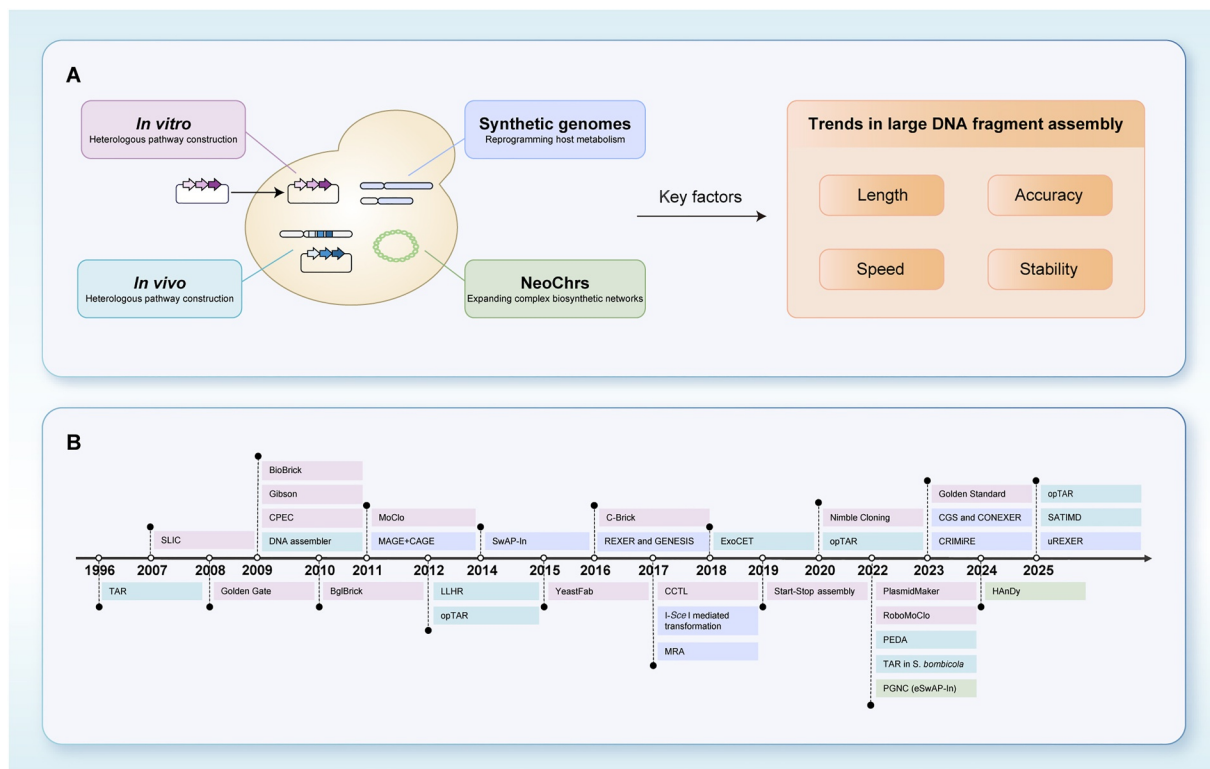


FIGURE 1 Overview of large DNA fragment assembly technologies in microbial cell factories. (A) Large DNA fragment assembly facilitates microbial host engineering in pathway construction, metabolism reprogramming, and network expansion, with current development trends focusing on some key factors. (B) Technological timeline highlights key milestones in large DNA fragment assembly.

without introducing unwanted sequences. Through standardized MoClo, BioBrick-based methods have been applied to the assembly of polyketide synthase (PKS) pathways [15] and the biosynthesis of amino acids and cofactors [72, 73], facilitating the modular engineering of microbial metabolic pathways. However, the reliance on fixed restriction sites limits multigene assembly, as only stepwise two- or three-fragment ligations are possible [16, 67], making BioBrick-based construction time-consuming and less efficient for large or complex pathways.

By contrast, Type IIS enzymes such as *Bsa* I and *BsmB* I cleave DNA outside of their recognition sites to generate user-defined overhang pairs, enabling single-enzyme, directional, and multifragment assembly. It overcomes the fixed-site limitations of Type IIP enzymes, allowing efficient multigene pathway construction in microbial cell factories. As shown in Figure 2B, Golden Gate assembly (GGA) enables one-pot construction of multiple DNA fragments, with successful assembly of up to 11 fragments reported [17, 74]. In practical applications, GGA has enabled the synthesis of complex metabolites in yeast, such as archaeal-eukaryotic hybrid lipids [75] and more than 200 diterpenes [76], showcasing its broad utility and potential in metabolic engineering. Recent advances have optimized overhang sequences to maximize efficiency.

Hamedirad et al. have screened over 200 overhang combinations and demonstrated near 100% assembly efficiency for 3–5 fragment constructs using iBioCAD, underscoring the enhanced accuracy and reliability enabled by overhang optimization [77]. Additionally, Start-Stop assembly relies on *Sap* I to generate 3 bp overhangs encoding start and stop codons, enabling seamless assembly of coding sequences while preserving mRNA secondary structure and ribosome binding site (RBS) activity, supporting high-efficiency multipart assembly and modularity [18]. Complementing these strategies, Expanded Golden Gate takes advantage of *Bsa* I-generated overhangs that are compatible with the multiple cloning site sequences of standard vectors, allowing PCR products to be ligated in a defined orientation and simplifying cloning design [78]. Collectively, these innovations reduce assembly errors, maintain functional integrity of pathway genes, and streamline the construction of heterologous metabolic modules, directly supporting efficient design and testing in microbial cell factories (Figure 2B).

Beyond junction-level optimizations, hierarchical assembly strategies have enabled the construction of increasingly large and complex heterologous pathways. MoClo exemplifies this approach, employing *Bsa* I and *Bpi* I to generate customizable overhangs that allow precise, directional assembly of up to eight DNA

TABLE 1 DNA assembly methods, with a focus on methods with applications in biosynthesis.

Category	Host	Assembly method ^a	Size/ number ^b	Key features	Applications of microbial cell factory engineering
In vitro DNA assembly	No limitation	BioBrick [14, 15]	20 kb/2 fragments	Modular assembly with Type IIP restriction enzymes	Assembly of biological parts, TUs, and multigene pathways
		BglBrick [16]	2 fragments	Reliable: <i>Bgl</i> II/ <i>Bam</i> HI cutters with high efficiency. Methylation-insensitive: Works despite dam/dcm methylation. Benign scar: 6 bp glycine–serine linker.	Enables the construction of gene expression devices, chimeric protein fusions, and targeted genomic integrations in <i>E. coli</i>
		Golden Gate [17]	30 kb	Modular assembly with Type IIS restriction enzymes	Assembly of biological parts, TUs, and multigene pathways
		Start-Stop assembly [18]	15 fragments	Scarless translation junctions; allowing assembly of up to 15 genes from up to 60 parts	Applied for constructing combinatorial carotenoid biosynthetic libraries in <i>E. coli</i>
		MoClo [19, 20]	33 kb/11 TUs	Stepwise, modular construction of multigene pathways; allowing the construction of 11 transcriptional units from 44 individual modules in just three cloning steps	Assembly of biological parts, TUs, and multigene pathways
		YeastFab [21, 22]	6 fragments	Modular DNA assembly framework for constructing multigene pathways in yeast; 11 POT modules enable assembly of a 6-gene pathway; a 151-promoter library enables 3-gene expression optimization	Combinatorial assembly for optimizing β -carotene production
		Golden Standard [23]	20 TUs	6 constitutive + 4 inducible promoters; 4 mono + 2 bicistronic RBSs; 7 N-terminal + 5 C-terminal tags; 2 CDS types; 6 terminators; 15 linkers; 3 transcription factors; 41 host vectors.	Optimizing the biosynthetic pathway for β -carotene production and the expression of recombinant proteins
		RoboMoClo [24]	8 TUs	Robotics-assisted modular cloning	Lycopene biosynthesis in <i>Corynebacterium glutamicum</i>
		Gibson [25]	900 kb	A one-pot, isothermal approach for joining multiple overlapping DNA fragments	Seamlessly constructing synthetic and natural genes, genetic pathways, and entire genomes
		Nimble Cloning [26]	3 fragments	<i>Sfi</i> I/T5 for linearization and overhangs; standardized adapters for primers; more flexible and higher cloning efficiency than Gibson	Supports both PCR products and entry clones
		SLIC [27]	10 fragments	Uses 3'→5' exonuclease to create single-stranded overhangs, enabling annealing of complementary DNA fragments	Multifragment assembly
		CPEC [28]	8.4 kb	Employs a polymerase-mediated extension of overlapping DNA fragments to generate circular, seamless constructs in a single reaction	Assembly and cloning of complex combinatorial libraries and multipart pathways

TABLE 1 (Continued)

Category	Host	Assembly method ^a	Size/ number ^b	Key features	Applications of microbial cell factory engineering
In vivo DNA assembly		C-Brick [29]	2 fragments	Uses Cas12a (Cpf1)-generated 4–5 bp 5' overhangs for directional, modular assembly of multiple DNA fragments; assembly efficiency up to 29.4%, accuracy up to 93.6%	Assembled chromoprotein cassettes for pigment expression in <i>E. coli</i>
		CCTL [30]	-	Combines Cas12a (Cpf1)-assisted cleavage with Taq DNA ligase for high-precision ligation of large DNA fragments; assembly efficiency up to 78% and accuracy 100% in high-GC regions	Actinorhodin production in <i>Streptomyces</i>
		PlasmidMaker [31]	27 kb/11 fragments	101 plasmids (5–27 kb) from up to 11 fragments were assembled across six species with minimal human intervention, handling GC contents up to 77%; accuracy ranging from 18% to 90%	High-throughput assembly of large plasmids with multiple fragments and high GC content
	<i>S. cerevisiae</i>	LLHR [32, 33]	52 kb	The full-length RecET system (LLHR) exhibits 20- and 40-fold higher efficiency than its truncated counterpart and the red system, respectively	Direct cloning of 10 megasynthetase gene clusters from <i>P. luminescens</i> ; identification of a putative colletochlorin BGC in <i>Colletotrichum higginsianum</i>
		ExoCET [34]	106 kb	Direct cloning efficiency has been increased by up to 32-fold, enabling one-step cloning of BGCs larger than 100 kb	Direct cloning of the 106-kb salinomycin BGC from <i>S. albus</i>
		TAR [35]	600 kb	TAR avoids the limitations of in vitro cloning methods in terms of length and cloning BGCs from complex backgrounds	Direct cloning of large human DNA
		opTAR [36]	454 kb	Enhanced stability of assemblies via the addition of yeast origins of replication (the maximum size improved from 142 to 454 kb)	Assembly and stable maintenance of the 454-kb genomic DNA from <i>Synechococcus elongatus</i> PCC 7942
		opTAR [37]	67 kb	Enhanced efficiency of BGC cloning via K1 α selection (from 0.26% to 13%)	Cloning of the 35-kb chelocardin and 67-kb daptomycin biosynthetic gene clusters from <i>Amycolatopsis sulphurea</i> and <i>Streptomyces filamentosus</i>
		opTAR [38]	76 kb	Enhanced efficiency of BGC cloning via hook optimization (a 70% capture rate for corbomycin BGC)	Cloning of the 76-kb corbomycin BGC from <i>Streptomyces</i> sp. WAC01529
		DNA assembler [39–41]	45 kb	Rapid and highly efficient (70%–100%) assembly of multigene pathways	Assembly of three distinct pathways: a 9-kb (3-gene) D-xylose utilization pathway, an 11-kb (5-gene) zeaxanthin biosynthetic pathway, and a combined 19-kb (8-gene) pathway for both functions
					(Continues)

TABLE 1 (Continued)

Category	Host	Assembly method ^a	Size/ number ^b	Key features	Applications of microbial cell factory engineering
Synthetic genome assembly	<i>Aspergillus nidulans</i>	TAR [42]	-	Enhanced efficiency (about 70% fusion rate with 25 bp overlaps) and capacity (one-step assembly of 6 fragments) via NHEJ deficiency	Streamlined construction of self-replicating plasmids lacking bacterial sequences
	<i>Starmarella bombicola</i>	TAR [43]	-	CRISPR/Cas9-mediated one-step in vivo assembly	Assembly of multiple DNA fragments (10 kb) for acid-type sophorolipid production
	<i>Trichoderma reesei</i>	SATIMD [44]	32.7 kb	Targeted integration and stitching efficiencies are close to 100% and over 80%, respectively	Assembly of a 32.7-kb sorbicillinoid BGC (up to 10 different DNA fragments)
	<i>E. coli</i> , <i>Ralstonia eutropha</i> , <i>Pseudomonas putida</i> , <i>Lactobacillus plantarum</i> , and <i>Yarrowia lipolytica</i>	PEDA	6 kb	Assembly efficiency is more than double that of wild-type <i>E. coli</i> lacking T4 ligase and T5 exonuclease	RBS library construction in <i>P. putida</i> and direct pathway assembly into the production strain
	<i>E. coli</i>	MAGE and CAGE [7]	2.3 Mb	MAGE enables scattered, genome-scale editing but introduces a high frequency of off-target mutations; CAGE's reliance on iterative conjugation and selection makes large-fragment assembly a time-consuming process (weeks to months)	Assembly of a UAG stop codon-free <i>E. coli</i> genome for ncAA incorporation and antiviral properties
		REXER and GENESIS [45]	0.5 Mb	Long synthetic DNA is capable of replacing the genome stepwise; the iterative marker-swapping process, utilizing CRISPR/Cas9 and λ -Red recombination, enables precise and efficient assembly; dependence on electroporation to introduce fragments yields low efficiency	Assembly of the recoded genome strain Syn61 for the production of noncanonical heteropolymers and macrocycles and for broad-spectrum phage resistance
		uREXER [46]	0.1 Mb	The streamlined REXER variant uses universal CRISPR/Cas9 spacers to enable more efficient 100 kb fragment replacement	Assembly of the recoded genome strain Syn57
		CGS and CONEXER [47]	0.5–1.1 Mb	Coupling REXER with bacterial conjugation decreases the time for a single integration from 4 days to 1 day; RecA knockout minimizes crossovers between the synthetic DNA and the genome, without the need for sequencing	-
	<i>S. cerevisiae</i>	SwAP-In [48–54]	0.18–0.75 Mb	Coupling yeast HR with dual auxotrophic selection ensures high-fidelity assembly; this stepwise megachunk integration is time-consuming (5 days for single integration); the assembly efficiency decreases as the number of rounds extends	Assembly of synthetic yeast chromosomes synI, synIII, synV, synVI, synIX, synX, and synXIV for conditional tolerance and production improvement

TABLE 1 (Continued)

Category	Host	Assembly method ^a	Size/ number ^b	Key features	Applications of microbial cell factory engineering
NeoChr assembly	<i>S. cerevisiae</i>	I-Sce I-mediated transformation [55–57]	0.77–1.03 Mb	By performing two halves of the assembly in parallel, it is about 50% faster than SwAP-In	Assembly of synthetic yeast chromosomes synII, synVII, and synXIII for production improvement
		MRA [58–61]	0.51–1.45 Mb	Parallel assembly of multiple segments across several haploid strains dramatically reduces the timeline compared to sequential methods; meiotic recombination carries a high risk of crossover with wild-type sequences	Assembly of synthetic yeast chromosomes synIV, synVIII, synXII, and synXVI for conditional tolerance
		CRIMiRE [62, 63]	0.66–1.05 Mb	Using targeted CRISPR/Cas9 cleavage to avoid the random cross-mixing of MRA improves assembly fidelity	Assembly of synthetic yeast chromosomes synXI and synXV
		TAR [64]	100 kb	Successful large-scale assembly (44 fragments, 100-kb NeoChr) with high fidelity (36%)	NeoChrs as an exclusive expression platform for the swapping of the glycolytic pathway
		TAR [5, 65]	100 kb	Higher assembly efficiency of circular chromosomes compared to linear chromosomes (36% vs. 0% faithful to the original design)	The integration of the glycolysis, pentose phosphate, bacterial vanillic acid, and plant anthocyanin synthesis pathways onto NeoChrs enabled the de novo biosynthesis of anthocyanin
		PGNC/eSwAP-In [66]	211 kb	Modular assembly (21 chunks); superior growth rate and stability of circular versus linear chromosomes	PGNC enabled carbon source expansion and enhanced phenotypic diversity via the SCRaMbLE system
		HAnDy [6]	1.024 Mb	High efficiency (about 60%) and superior genetic stability	NeoChrs enhance host metabolism (9 new compounds) and environmental adaptability

^aRepresentative methods and new methods developed in the last 5 years.

^b“–” indicates “data not shown.”

modules simultaneously [19]. Its hierarchical design allows stepwise construction—from basic bioparts to TUs and further to multigene pathways—efficiently building constructs up to 33 kb in just three steps [20]. Similar hierarchical strategies have been applied in GoldenBraid [79] for plants and YeastFab [21, 22] for yeast, facilitating modular pathway construction in more hosts. In addition, the Golden Standard adopts the Standard European Vector Architecture and provides efficient, standardized genetic tools for α -, β -, γ -, and δ -proteobacteria [23]. These platforms allow systematic assembly and combinatorial testing of metabolic modules, supporting rapid exploration of pathway designs and optimization of enzyme expression in different microbial backgrounds.

2.2 | Homologous recombination-based methods for seamless and flexible pathway assembly

Unlike restriction enzyme-based approaches, homologous recombination-based (HR-based) approaches do not rely on predefined recognition sites, allowing precise and seamless joining of multiple fragments regardless of their internal sequences (Figure 2C). Gibson assembly exemplifies this capability, employing T5 exonuclease, DNA polymerase, and DNA ligase in a single isothermal reaction, which enables assembly of multiple fragments, including constructs exceeding 100 kb, in a single step [25, 80]. The seamless nature preserves coding sequences and regulatory elements,

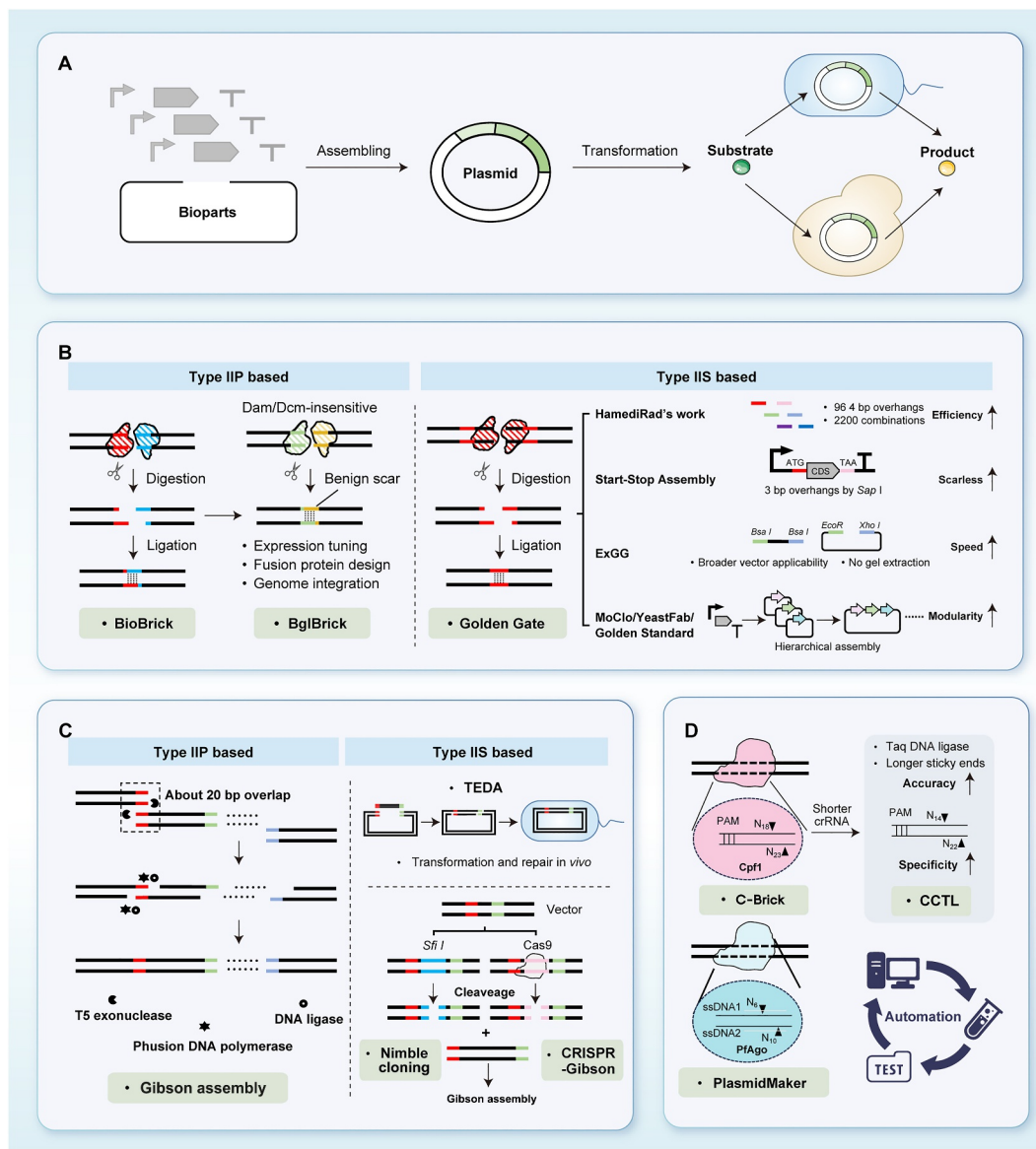


FIGURE 2 Development and applications of in vitro large DNA fragment assembly. (A) In vitro assembly of large DNA fragments enabling heterologous pathway construction in microbial cell factories. Principles and development trends of (B) restriction digestion-based, (C) homologous recombination-based, and (D) programmable nuclease-guided assembly methods.

maintaining the functional integrity of open reading frames, which is an essential feature for constructing heterologous metabolic pathways. Gibson assembly has been successfully applied to the assembly of large human *IGHV* gene fragments [81] and modular genetic circuit construction [82]. T5 exonuclease-dependent assembly only employs T5 exonuclease to generate short single-stranded overhangs that allow homologous DNA fragments to anneal seamlessly. The gaps are subsequently repaired in vivo, enabling simple and efficient multifragment assembly without complex enzyme mixtures [83]. Other HR-based strategies, such as sequence and ligation-independent cloning (SLIC) [27] and circular polymerase extension cloning [28], exploit similar overlap-driven recombination principles,

providing additional options for assembling large DNA constructs.

The flexibility and sequence independence of homology-directed assembly make it particularly suitable for integration with other DNA manipulation strategies. By providing homologous overlaps at fragment ends, HR-based methods can seamlessly join DNA segments generated through diverse upstream processes. This flexibility has been further enhanced through hybrid strategies (Figure 2C). For example, Nimble Cloning, which employs the rare-cutting enzyme *Sfi* I, enables rapid plasmid linearization without extensive fragment domestication [26]. In a second hybrid approach, CRISPR-Gibson hybrid strategies use Cas9 guided by a user-designed single

guide RNA (sgRNA) to generate targeted double-stranded breaks at virtually any desired location in a vector. PCR-amplified inserts with homologous overlaps can then be seamlessly integrated via Gibson assembly [84]. These approaches have been used to construct complex assemblies, including rapamycin polyketide synthase modules [85], a UAS-cDNA/ORF library of over 5500 *Drosophila* genes conserved in humans [86], and entire biosynthetic clusters including fengycin [87]. These combinations leverage the sequence independence of HR-based assembly with the precision of nuclease systems, significantly expanding the design space for reconstructing large biosynthetic clusters and generating high-diversity pathway libraries.

2.3 | Programmable nuclease-guided assembly methods for precision and automation in pathway design

Programmable nucleases have emerged as powerful tools for in vitro DNA assembly, offering unparalleled precision and automation in constructing large DNA fragments (Figure 2D). Precision arises from programmable target recognition and cleavage: Cas12a (Cpf1) guided by crRNA recognizes TTN-type protospacer adjacent motif (PAM) sites and generates defined 4–5 bp 5' overhangs, enabling directional and predictable assembly of multiple fragments, as implemented in the C-Brick system [29]. However, conventional Cpf1 cleaves variably depending on the target sequence and crRNA design, often cutting at multiple sites and producing inconsistent ends that compromise assembly accuracy. As shown in Figure 2D, the Cpf1-assisted cutting and Taq DNA ligase-assisted ligation (CCTL) method employs a 17-nt crRNA to confine cleavage to defined sites, generating predictable 8-nt sticky ends. Combined with Taq DNA ligase's high-temperature, high-specificity ligation, it markedly improves the precision and efficiency of large DNA construct modification [30].

In parallel, *Pyrococcus furiosus* argonaute (PfAgo) is a programmable nuclease that does not rely on PAM sequences. Thermal denaturation of dsDNA allows strand separation, followed by site-specific cleavage using 5'-phosphorylated DNA guides, and reannealing to produce defined sticky-end double-stranded breaks [88]. PfAgo has been incorporated into the fully automated PlasmidMaker platform (Figure 2D), enabling end-to-end plasmid construction from primer design to ligation with high throughput and accuracy [31]. It enables high-throughput assembly of plasmids up to 27 kb and 11 fragments, including constructs with high GC content (up to 77%) and multiple repeated sequences. These programmable nuclease strategies enable precise, predictable assembly of multiple gene modules

and allow automated high-throughput construction, directly facilitating the combinatorial design, rapid iteration, and optimization of heterologous metabolic pathways in microbial cell factories.

3 | IN VIVO ASSEMBLY STRATEGIES FOR THE CONSTRUCTION OF HETEROLOGOUS METABOLIC PATHWAYS

In vitro assembly methods provide high flexibility and efficiency for constructing metabolic pathways; however, their efficiency decreases significantly as the number and size of DNA fragments increase [21, 89]. Large biosynthetic pathways—such as natural product BGCs, which typically span tens to hundreds of kb [37, 69, 87]—are primarily assembled in vivo through HR within host cells. In general, in vivo assembly achieves higher efficiency and greater stability than in vitro recombination methods [90], thereby accelerating the development of microbial cell factories and the optimization of biosynthetic pathways. The Gram-negative bacterium *E. coli* (Figure 3A) and the Gram-positive bacterium *S. cerevisiae* (Figure 3B) are among the most widely used hosts for the in vivo assembly of BGCs [38, 69, 87] and biochemical pathways [39, 40]. These assembled constructs enable two production routes: direct synthesis of the target metabolite in the original host or transfer of the plasmid to a specialized production chassis (Figure 3C). Moreover, in vivo recombination strategies have been extended to various other hosts, significantly advancing the fields of synthetic biology and metabolic engineering.

3.1 | In vivo DNA assembly in *E. coli* for the cloning and refactoring of BGCs

E. coli is a preferred host for DNA cloning and maintenance because its endogenous recombination can be minimized more thoroughly than in any other host, making it particularly suited for assembling secondary metabolite pathways that usually contain highly repetitive DNA regions (such as polyketide BGCs). However, endogenous RecA-mediated HR in *E. coli* is inefficient, and exogenous linear dsDNA introduced for assembly is degraded by the RecBCD complex, resulting in low assembly efficiency in vivo [91]. To overcome these limitations, the Red recombination system from λ phage (comprising Exo, Beta, and Gam proteins) and the RecET system from Rac phage (comprising RecE and RecT proteins) have been developed to significantly enhance HR efficiency in *E. coli*. Although both systems enhance recombination efficiency in *E. coli*, the Red system is more effective for recombination between linear and circular DNA. In

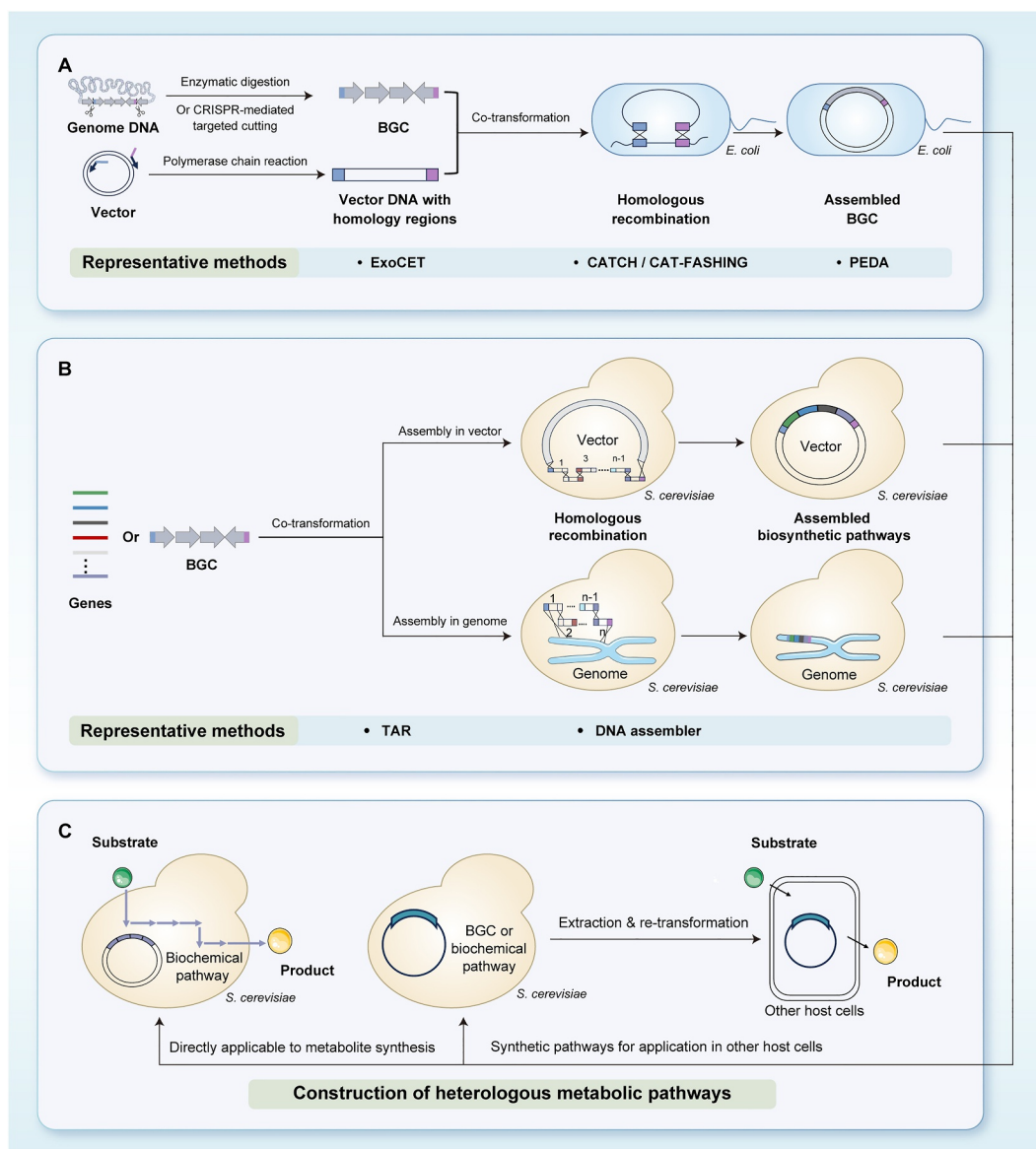


FIGURE 3 In vivo assembly strategies for the construction of heterologous metabolic pathways. In vivo DNA assembly in (A) *E. coli* and (B) *S. cerevisiae*. (C) Applications in cell factory engineering (as illustrated by the yeast system).

contrast, RecET efficiently mediates recombination between two linear DNA molecules. Fu et al. [32] first developed the linear-linear homologous recombination (LLHR) technique in 2012 using full-length RecE and RecT. Using LLHR, several BGCs have been successfully assembled, including the syringolin biosynthetic gene cluster (*syICDE*, 19 kb) from *Pseudomonas syringae* [92], the luminmycin biosynthetic pathway (18 kb) from *Photorhabdus luminescens* [68], the sevadin biosynthetic gene cluster (11.6 kb) from *Paenibacillus larvae* [69] and polyketide secondary metabolite BGCs from *Colletotrichum higginsianum* [33]. BGCs are typically obtained through enzymatic digestion of genomic DNA, but the limitation of restriction sites may hinder the release of BGCs. Ongoing improvements (Figure 3A) have incorporated

CRISPR-mediated targeted cutting systems, such as Cas9-Assisted Targeting of CHromosome segments (CATCH) [93, 94] and CRISPR/Cas12a-mediated Fast Direct Biosynthetic Gene Cluster Cloning Platform (CAT-FASHING) [95], which enable rapid and efficient capture of larger BGCs. Beyond pathway assembly, combining RecET with Red recombination or CRISPR/Cas systems has enabled effective refactoring of targeted BGCs via promoter engineering and RBS optimization, significantly enhancing the yield of natural products [96, 97]. Using the recently developed CRISPR/Cas9 and RecET-mediated Refactoring (CRISPR/Cas9 and RecET-mediated Refactoring) strategy, He et al. [97] refactored the 74-kb daptomycin BGC by introducing diverse combinations of promoters with varying strengths, which resulted in a 20.4-fold yield improvement of daptomycin.

The efficiency of direct cloning is limited by the chance that both the linear cloning vector and the target BGC simultaneously enter an *E. coli* cell during electroporation. In 2018, researchers developed the Exonuclease Combined with RecET recombination (ExoCET) method, which employs an initial *in vitro* step that exploits the 3'→5' exonuclease activity of T4 polymerase to preanneal linear vectors with genomic fragments [34]. This approach enabled the successful construction of large BGCs that had previously been challenging to clone using RecET recombination alone. Notably, whereas a large 106 kb fragment previously required cloning in three separate steps, ExoCET enables single-step cloning, substantially improving the efficiency of BGC construction.

3.2 | *In vivo* DNA assembly in *S. cerevisiae* for the construction of multigene metabolic pathways

S. cerevisiae exhibits a high HR rate, where DNA fragments with 40 bp of overlapping sequence achieve assembly efficiencies exceeding 90% [39, 98, 99]. Leveraging its ability to recombine small homologous and divergent DNA sequences during transformation, Larionov et al. [35] developed the transformation-associated recombination (TAR) method in 1996, which enables selective and accurate cloning of chromosomal regions from complex genomes (Figure 3B). Their findings indicated the highest yield of yeast artificial chromosomes (YACs) when using vectors lacking autonomously replicating sequence (ARS) elements. However, for high G + C DNA constructs with a total size of over 200 kb, stable maintenance proved unattainable without integrating ARS sequences [36]. The updated TAR vector incorporates a yeast selectable marker, ARS, and a centromere sequence to ensure proper propagation, segregation, and selection in *S. cerevisiae*. The efficiency of TAR remains low (0.1%–2%) due to nonhomologous end joining (NHEJ), which often results in misassembly [100]. Several strategies have been developed to enhance TAR efficiency, including: (1) separating essential elements (yeast episome and selection marker) [101]; (2) introducing the *URA3* gene as a counterselectable marker [102]; (3) using the α subunit of K toxin as a counterselectable marker [37]; and (4) employing CRISPR/Cas9 to pretreat genomic DNA, which increased region-positive cloning efficiency by 35-fold [103]. Recently, to overcome the challenge of low efficiency in capturing BGCs and difficulty in stably maintaining large BGCs (e.g., glycopeptide antibiotic BGCs, typically >70 kb) using traditional TAR methods, Xu et al. [38] developed an optimized system by optimizing the expression of counter selection marker *URA3*, reasonably designing hook sequences, and incorporating a copy-number control replicon

(oriV-ori2-repE-sopABC). This optimized system achieved a 70% capture efficiency for the 76-kb corbomycin BGC, a substantial improvement over the traditional TAR efficiency (below 10%). Notably, the captured BGCs demonstrated stable integration and efficient expression in other hosts, which achieved a 19-fold increase in the final corbomycin yield on the GPAHex platform. TAR eliminates the need for *in vitro* manipulation and enables the assembly of extremely long DNA fragments, pioneering HR-mediated assembly in yeast cells. This method has been widely applied to assemble individual or entire genomes [104–107], assemble BGCs from complex genomes [37, 108, 109], and reconstruct gene clusters for synthetic biology applications [110–112]. More recently, TAR has been integrated into an automated biofoundry platform, significantly improving the scalability and precision of large-scale DNA assembly [113]. This platform enables the construction of expression-tunable multigene libraries, such as the β -carotene metabolic pathway.

Similarly, Shao et al. [39] developed the DNA assembler method in *S. cerevisiae* in 2009, enabling the one-step assembly of multiple gene expression cassettes into biosynthetic pathways via *in vivo* HR. Functional biosynthetic pathways ranging from 9 to 19 kb—including a three-gene D-xylose utilization pathway, a five-gene zeaxanthin biosynthetic pathway, and a combined eight-gene D-xylose utilization and zeaxanthin biosynthetic pathway—can be rapidly assembled on either plasmids or chromosomes [39, 40]. Beyond pathway assembly, the DNA assembler has also been employed to optimize metabolic pathways using combinatorial transcriptional engineering (COMPACTER), enabling rapid tuning of gene expression in xylose and cellobiose utilization pathways [114]. Currently, DNA assembler-assembled plasmids are applied in targeted genome editing [115], metabolite synthesis [116], and unconventional carbon source utilization [117].

3.3 | *In vivo* DNA assembly in other microorganisms

Advances in sequencing and HR technologies have facilitated the adaptation of *in vivo* DNA assembly for use in a diverse range of microorganisms. *Bacillus subtilis* (*B. subtilis*) has been one of the most widely used hosts for DNA assembly. Owing to its natural high extracellular DNA transformability, extracellular ssDNA can readily enter the cytoplasm of *B. subtilis* and integrate into the genome via RecA-mediated HR. Several *in vivo* assembly methods have been developed in *B. subtilis*, including inchworm elongation (IWe), domino assembly, and inducible RecA expression BGM vector (iREX) [118–120]. Comprehensive reviews of DNA assembly techniques in *B. subtilis* have been provided by Li et al.

[121]. Beyond *E. coli*, *S. cerevisiae*, and *B. subtilis*, in vivo DNA assembly methods have also been established in various other microorganisms [42–44, 122]. In NHEJ-deficient strains of *Aspergillus nidulans*, efficient DNA fragment fusion requires only 25 bp sequence overlaps, facilitating the construction and optimization of metabolic pathways [42]. Shi et al. developed an efficient CRISPR/Cas9-based genome editing and one-step DNA assembly system in *Starmerella bombicola*, enabling the efficient assembly of multiple DNA fragments for the production of single acid-type sophorolipids with high yield and low cost [43]. More recently, a method termed simultaneous in vivo assembly and targeted genome integration of multiple DNA fragments (SATIMD) was developed in *Trichoderma reesei* (*T. reesei*). SATIMD employs CRISPR/Cas9 to introduce specific DSBs in the genome, thereby enhancing HR efficiency and enabling the assembly of a 32.7-kb sorbicillinoid BGC via the one-step integration of 10 DNA fragments at a specific locus [44]. This suggests that SATIMD is a powerful tool for heterologous gene cluster assembly, thus speeding up cell factory construction. Notably, neither the number of fragments nor the length of flanking homology arms affected assembly efficiency in *T. reesei*, a finding that warrants further investigation.

Direct assembly of large DNA fragments in hosts with limited HR capacity remains challenging. Increasing HR efficiency within these hosts offers the most direct and effective solution. Although recombination systems such as Red and RecET—developed initially in *E. coli*—have been widely used to improve HR efficiency in many Gram-negative bacteria, their application is often restricted to specific species [123]. To overcome this limitation, Huang et al. [122] identified an operon (XBJ1_operon 0213) from *Xenorhabdus bovienii* that encodes RecET-like recombinases. Expression of this operon significantly enhanced recombineering efficiency in both *E. coli* and *X. stockiae*, enabling the development of a robust recombineering system in *X. stockiae*. This technology holds significant application value in the identification and modification of BGCs in natural products. In a separate study, researchers developed a phage enzyme-assisted in vivo DNA assembly (PEDA) method [124] by combinatorially expressing T5 exonuclease and T4 DNA ligase (Figure 3). In this system, T5 exonuclease generates single-stranded overhangs from linear DNA ends, whereas T4 DNA ligase seals nicks after annealing of homologous sequences. Remarkably, PEDA enables the assembly of fragments with homologous overlaps as short as 5 bp in *E. coli* and is applicable to a broad range of species, including *Ralstonia eutropha*, *Pseudomonas putida* (*P. putida*), *Lactobacillus plantarum*, and *Yarrowia lipolytica*. Using this method, a plasmid library containing different RBS variants was constructed through direct in vivo DNA assembly in *P. putida* KT2440, confirming the applicability of PEDA for metabolic flux regulation.

4 | SYNTHETIC GENOME ASSEMBLY ENABLES GENOME-SCALE REPROGRAMMING OF HOST METABOLISM

The primary significance of synthetic genome assembly lies in its capacity for host metabolic reprogramming at the genome scale (Figure 4). As an extension of in vivo assembly, it uses the wild-type genome as a reference, employing a hierarchical strategy to construct synthetic segments and stepwise replace native genomic sequences, driving key innovations in large DNA fragment assembly [48, 55, 58]. Now, advances in synthetic genome assembly have progressed from viral genomes to prokaryotic genomes and then to multiple chromosomes in eukaryotic genomes [125–127]. De novo genome synthesis facilitates genome-wide engineering, endowing synthetic strains with expanded or enhanced industrial capabilities, such as global codon recoding for noncanonical amino acid (ncAA) incorporation [128, 129], antiviral immunity [130, 131], and insertion of global flexibility elements for accelerated evolution [132].

4.1 | Development of de novo synthetic genome assembly technologies

Synthetic genome assembly technology has evolved from viral to prokaryotic and then to eukaryotic genomes, leading to significant enhancements in length, accuracy, speed, and stability of genome assembly. For smaller viral genomes, the complete synthetic genomes of poliovirus (about 7 kb) [125] and ΦX174 bacteriophage (6 kb) [133] could be assembled in vitro with polymerase cycling assembly (PCA)-derived methods. For large genomic constructs ranging from 10 kb to Mb, where DNA fragments exhibit instability, the yeast-based TAR approach is employed. This method facilitates the combination of small DNA fragments with a YAC vector to form genome-YAC hybrids in yeast cells, enabling the assembly of large virus genomes such as SARS-CoV-2 [134] and mycoplasma genomes such as *Mycoplasma genitalium* (583 kb) [126] and *Mycoplasma mycoides* (1.08 Mb) [135].

For bacterial genome synthesis at the Mb scale, stepwise assembly of large synthetic fragments (>100 kb) is the principal strategy. Synthetic genome assembly enables the genome-wide removal of target codons, which can be reassigned for producing novel materials. To assemble a recoded *E. coli* genome with complete elimination of UAG stop codons, Isaacs et al. [7] used Multiplex automated genome engineering (MAGE) to modify 314 TAG-containing genes in the genome and created conjugative assembly genome engineering for codon mutation fragment transfer and assembly (Figure 4A). However, MAGE introduces

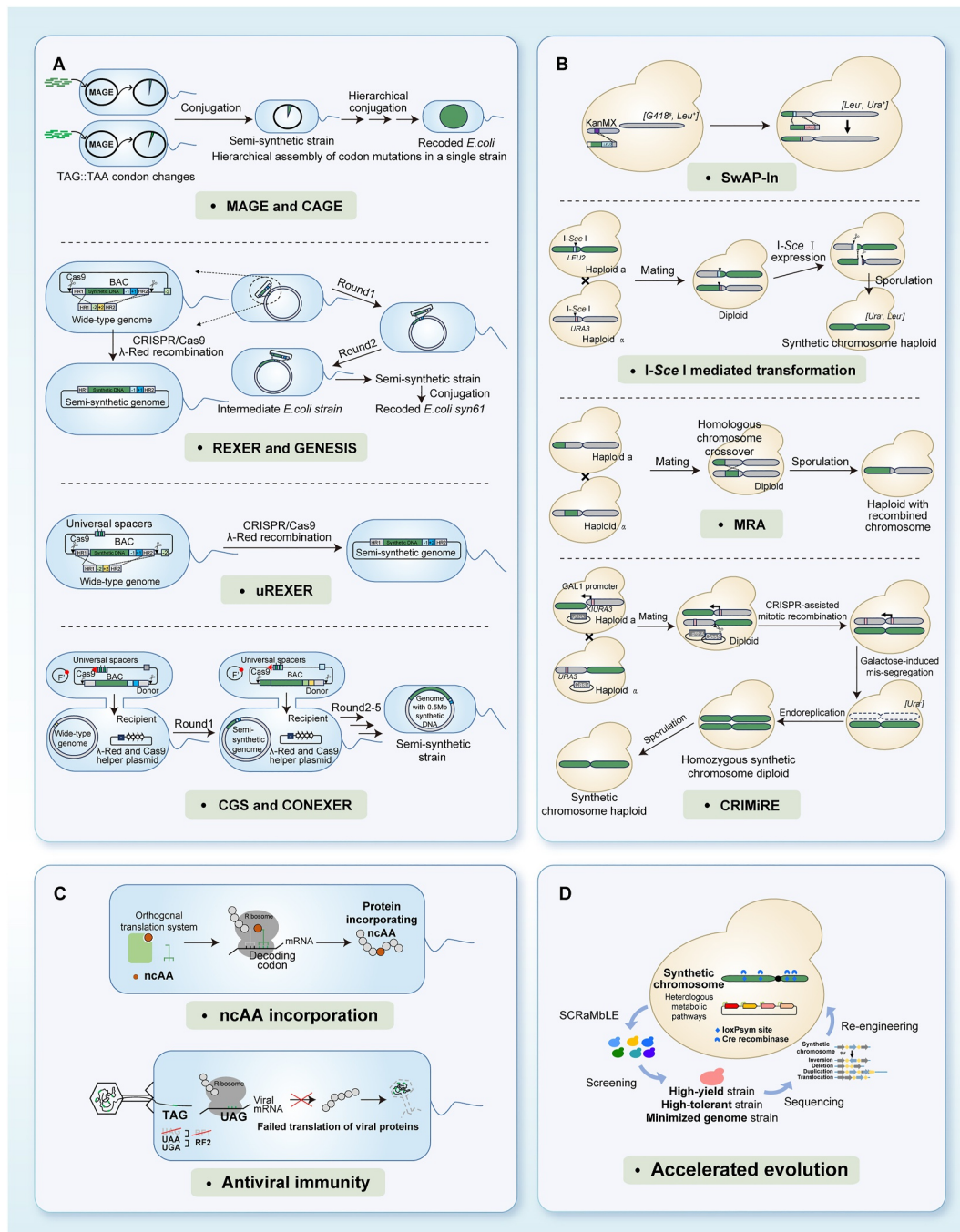


FIGURE 4 Synthetic genome assembly technologies drive applications of genomically recoded *E. coli* and synthetic yeast. The representative synthetic genome assembly methods in (A) *E. coli* and (B) *S. cerevisiae*. (C) Genome-wide codon removal in *E. coli* facilitates application of ncAA incorporation and antiviral immunity. (D) Genome-wide loxPysm insertion in a synthetic yeast chromosome allows accelerated evolution. The synthetic chromosomal fragments are represented in green, and the wild-type chromosomal fragments are represented in gray.

off-target mutations and is unsuitable for removing the much more abundant sense codons [7, 136]. To further compress to 61 codons, the replicon excision enhanced recombination (REXER), a dual-selection marker genome integration technique, was used [45]. REXER utilizes CRISPR/Cas9 and λ -Red recombination systems to drive efficient and precise replacement of

wild-type genomic segments (Figure 4A). The iterative marker-swapping process, termed genome stepwise interchange synthesis (GENESIS), combined with conjugation, enabled the construction of the final synthetic genome strain Syn61 [137]. To deeply compress the genetic code, George Church's team worked toward constructing a fully 57-codon recoded synthetic strain

(*rE.coli*-57 or *Ec_Syn57*). Their strategy employed the yeast TAR method to assemble about 50 kb of recoded fragments for stepwise genome replacement [138] and a multiomics-based iterative construction and troubleshooting (SynOMICS) approach to identify and correct issues in the synthetic genome assembly [139]. In contrast, Jason Chin's team developed the uREXER method based on REXER, which employs universal CRISPR/Cas9 spacers to streamline the process (Figure 4A). Through iterative uREXER (GENESIS assembly) and conjugation assembly, these recoded fragments were compiled into a single chassis, Syn57 [46]. Moreover, the team also upgraded the continuous genome synthesis method—conjugation coupled with programmed excision for enhanced recombination (CONEXER). Compared to REXER, CONEXER integrates bacterial conjugation and omits intermediate sequencing (Figure 4A), facilitating single integration from 4 days to 1 day [47].

In eukaryotic genome synthesis, Sc2.0 is a milestone project aimed at the complete synthesis of yeast chromosomes. A core principle involves genome-wide introduction of inducible genetic flexibility elements to enable genome rearrangement for phenotypic evolution of yeast strains [48]. To achieve these fully synthetic yeast chromosomes, the standardized switching auxotrophies progressively for integration (SwAP-In) method was developed [127] (Figure 4B). This method progressively replaced native chromosomal regions with synthetic megachunks (30–50 kb) by employing alternating selection with auxotrophic markers (*URA3* and *LEU2*), successfully generating several synthetic chromosomes, including synI [49], synIII [127], synV [50], synVI [51], synIX [52], synX [53], and synXIV [54]. Although the SwAP-In method minimizes off-target effects by utilizing yeast HR, multiple rounds of replacement make the assembly process time-consuming, and efficiency decreases as the number of rounds extends. To accelerate this process, I-Sce I-mediated transformation enabled parallel replacement from both chromosome ends in haploids, followed by diploid hybridization and cleavage by I-Sce I to successfully assemble chromosomes synII [55], synVII [56], and synXIII [57] (Figure 4B). This method reduces the total assembly time by almost 50%, significantly boosting synthetic chromosome assembly efficiency. The meiotic recombination-mediated assembly method further streamlines and accelerates the assembly process by combining segments from different haploid strains through meiosis (Figure 4B), ultimately assembling chromosomes synIV [59], synVIII [60], synXII [58], and synXVI [61]. However, the recombination process during assembly is random, and cross-mixing with wild-type sequences becomes a major bottleneck, especially for long synthetic chromosomes. CRISPR/Cas9-mediated mitotic recombination with endoreduplication (CRIMiRE) leverages the unique cleavage

sites of CRISPR/Cas9 to ensure chromosome recombination at precise locations and generates a homozygous diploid configuration to prevent cross-mixing with wild-type sequences, enabling efficient assembly of synXI [62] and synXV [63] (Figure 4B). After the assembly of individual synthetic chromosomes, multichromosome consolidation strategies such as “endoreduplication intercross” and “KAR1 mutation” have successfully combined up to 6.5 synthetic chromosomes in a single strain, bringing the goal of a fully synthetic yeast genome within the immediate future [140, 141].

4.2 | Applications of synthetic genome-engineered *E. coli* in ncAA incorporation and antiviral immunity

The assembly of the *E. coli* synthetic genome has created genomically recoded organisms, such as synthetic *E. coli* with stop codons swapped from TAG to TAA, Syn61, *rE.coli*-57 (or *Ec_Syn57*), and Syn57, which free up 1–7 codons. These strains accelerate genetic code expansion (GCE) technology development by employing orthogonal translation systems (aminoacyl-tRNA synthetase/tRNA pairs). By harnessing this system, vacant codons can be reassigned to incorporate ncAAs into target proteins, producing novel peptide drugs, biopolymers, and industrial enzymes (Figure 4C). For instance, the C321.ΔA *E. coli* strain, featuring genome-wide UAG to UAA codon replacement and RF1 knockout, converts UAG into a functional sense codon for ncAA insertion. Lajoie et al. [142] employed this system to achieve high efficiency and specific insertion of pAcF with undetectable read-through. Cheng et al. [143] achieved efficient Sec incorporation at UAG codons, enabling production of mammalian selenoprotein thioredoxin reductases with unsurpassed purity and yield. To improve the yield of ncAA-containing proteins, Yi et al. [144] integrated the T7 RNA polymerase system into C321.ΔA, achieving a 2.5-fold increase in target protein production. Furthermore, Grome et al. [145] compressed stop codons into a single codon, generating the C321.ΔA ΔTGA variant. Both UAG and UGA codons were reassigned to enable multisite incorporation of two ncAAs into a protein with >99% accuracy. In the Syn61 strain, genes decoding vacant codons were deleted (Syn61Δ3), and three mutually orthogonal ncAA incorporation systems were introduced. By reassigning TCG, TCA, and TAG codons, Robertson et al. achieved simultaneous incorporation of three ncAAs into proteins, producing diverse noncanonical heteropolymers and macrocycles [128]. Spinck et al. utilized the Syn61Δ3 platform to encode five distinct ncAAs (AllocK, AlkynK, pIF, CbzK, and pAzF) at TCG and TAG codons in the macrocycle-encoding genes, generating 25 macrocyclic peptides,

each incorporating two distinct ncAAs [129]. Through targeted hydroxy acid incorporation, they also achieved the synthesis of 12 diverse non-natural depsipeptide macrocycles. With the synthesis of Syn57, which has created a deeply recoded strain to genetically encode up to seven distinct ncAAs into proteins, and the synthesis of noncanonical macrocycles and polymers comprising up to seven novel monomers [46]. These achievements demonstrate the immense potential of synthetic genome-assembled strains for functional protein biomanufacturing.

On the other hand, recoding the *E. coli* genome to reassign the UAG stop codon effectively creates a virus-resistant host, providing high biosafety for industrial fermentation. As shown in Figure 4C, in the UAG codon- and RF1-deficient host C321.ΔA, it disrupts proper termination of T7 phage (genome contains 6 UAG among 60 stop codons) gene expression, preventing viral amplification and propagation [142]. Ma et al. also discovered that the C321.ΔA strain exhibits multiviral (λ, M13, P1, and MS2) resistance at titers up to 10¹¹ PFU/mL and impaired conjugative plasmids (F and RK2) up to 10⁵-fold [146]. The Syn61Δ3 strain, with more deletions of codons and translation components, exhibits broad-spectrum phage resistance, with negligible growth impact from the phage cocktail (λ, P1vir, T4, T6, and T7) [128]. In addition, the Syn61Δ3 strain, with its refactored genetic codes, not only resists viruses written in the canonical code but also blocks genetic information leakage from synthetic organisms to natural organisms [130]. These synthetic genome-assembled strains provide a genetic firewall against the natural ecosystem; they thus represent a secure production chassis for pharmaceuticals and bioactive compounds.

4.3 | Applications of the synthetic chromosome yeast for accelerated evolution

In the Sc2.0 project, to introduce inducible genetic flexibility into synthetic chromosomes, thousands of symmetric Cre-loxP recombination (loxPsym) sites were inserted into the 3' UTR of each nonessential gene, enabling a Cre recombinase-mediated rearrangement of synthetic chromosomes called synthetic chromosome rearrangement and modification by loxP-mediated evolution (SCRaMbLE). These modifications, including deletions, inversions, duplications, and translocations of chromosome regions, generate combinatorial stochastic diversity of synthetic chromosomes [147]. This has been shown to accelerate genome-wide metabolic regulation and evolution in yeast, enabling applications in conditional tolerant strain engineering, natural compound production enhancement, and chassis cell simplification (Figure 4D). In terms of enhancing strain tolerance

under environmental stress conditions, Luo et al. [148] used SCRaMbLE and efficient selection based on a loxP-mediated switch of two auxotrophic markers (ReSCuES) in the synXII strain, generating three yeast strains capable of growing at 39.5°C, improving the OD600 value by 21-fold under 0.6% acetic acid conditions, and generating a tolerant strain capable of growing in 8% ethanol. Ma et al. [149] employed SCRaMbLE in yeast strains harboring either one (synV) or two (synV and synX) synthetic chromosomes, obtaining alkali-tolerant strains capable of withstanding pH 8.0. Whole-genome sequencing analysis confirmed that the deletion of *SPT2* was responsible for the enhanced alkali tolerance. Additionally, researchers also screened strains with improved tolerance to antibiotics such as rapamycin [150] and hygromycin [151], as well as growth-affecting factors such as caffeine [152] and high salt [153] via SCRaMbLE.

In terms of sugar utilization, Blount et al. [154] introduced the xylose metabolic pathway into the synV strain. By SCRaMbLE, they successfully obtained a strain with a 5.85-fold improvement in growth capability on a medium where xylose was the sole carbon source. Structure variant analysis revealed that the deletion of *MXR1* was critical for xylose utilization, offering potential for environmentally friendly application of yeast. In enhancing the production of natural compounds, Zhang et al. [155] achieved a 15.8-fold increase in lycopene yield by applying SCRaMbLE in a diploid strain harboring synII, attributing part of this improvement to the deletion of the *PEX32* gene. Jia et al. [156] employed a combinatorial approach using an AND gate to enable precise control of SCRaMbLE and multiplex SCRaMbLE iterative cycling (MuSIC) in a diploid yeast carrying synIII and synV. After five cycles, they achieved a 38.8-fold increase in carotenoid production. Additionally, the SCRaMbLE of synthetic chromosome yeast has demonstrated potential in boosting the biosynthesis of more compounds, including violacein [157], astaxanthin [158], betulinic acid [159], and geraniol [160], as well as phenylethyl acetate and phenylethylamine [161]. Furthermore, SCRaMbLE-based systems can also be employed to build metabolically simplified chassis cells by deleting nonessential chromosomal genes. This strategic reduction will create a cleaner background for heterologous gene expression, helping improve metabolic resource partitioning and enhance metabolic pathway efficiency [162, 163].

5 | SYNTHESIZING NeoChrs FOR EXPANDING COMPLEX BIOSYNTHETIC NETWORKS

Constructing efficient *S. cerevisiae* cell factories typically requires the integration of multiple exogenous genes into native chromosomes. For example, the

implementation of the complete QS-21 biosynthetic pathway necessitated the introduction of 38 heterologous genes from seven enzyme families [164]. However, integrating such large genetic constructs might disrupt chromosomal architecture, potentially affecting the expression of neighboring genes and compromising chromosome replication. Given that *S. cerevisiae* exhibits tolerance to variations in chromosome number [165, 166], synthetic NeoChrs—which are modular, supernumerary artificial chromosomes providing an orthogonal platform for stable, high-capacity endogenous or/and heterologous gene expression without disrupting native chromosomal function—have emerged as platforms for reprogramming native cellular processes and introducing novel functionalities [5, 64, 66]. In contrast to the construction of synthetic chromosomes in Sc2.0, which relies on existing genomic frameworks [48, 55], there are no native templates for the design of NeoChrs, and thus, NeoChrs need to be de novo designed from individual accessory genes and regulatory elements. To assemble NeoChrs, structural elements (centromere, ARS, telomeres, and auxotrophic markers), along with PCR-amplified essential gene sequences, were cotransformed into *S. cerevisiae* for in vivo assembly. The overall construction workflow of NeoChr is illustrated in Figure 5A.

5.1 | Constructing NeoChrs via de novo design and assembly

Several synthetic NeoChrs have been constructed to expand the capabilities of microbial cell factories. In 2021, Postma et al. [64] exploited yeast HR to assemble up to 44 DNA parts in a single step into modular circular NeoChrs (Figure 5B). Specifically, they first assembled TUs—each consisting of a promoter, gene, and terminator—via Golden Gate cloning. Subsequently, they de novo assembled these fragments, which featured 60 bp nongenic overhangs, via the highly efficient HR machinery of *S. cerevisiae*. The stability of NeoChrs and their effect on host physiology are key considerations for their application in synthetic biology. Interestingly, the introduction of four differently sized empty NeoChrs resulted in reduced specific growth rates and the loss of NeoChrs in approximately 20% of the population [64]. To address this, the authors hypothesized that a linear configuration mimicking native yeast chromosomes could improve performance. They constructed a 100-kb linear NeoChr using short telomere seed sequences for telomere formation (Figure 5C) [5]. However, results showed that linear NeoChrs exhibited lower assembly efficiency and fidelity compared to their circular counterparts, although both types were comparable in growth rate, stability, and phenotypic effects.

Because of the limitations of one-step assembly—namely, (1) the reduced efficiency of in vivo assembly as the number of fragments increases, and (2) the decreased probability that all transformed fragments will successfully enter the nucleus—large-size NeoChrs are typically constructed through iterative genome assembly involving multiple rounds of HR with a small number of fragments. Mitchell et al. developed the eSwAP-In strategy (extrachromosomal switching auxotrophies progressively for integration) [167]. This method relies on iterative use of two different auxotrophic markers (*URA3* and *LEU2*) embedded near the rightmost ends of DNA fragments that are designed for stepwise assembly. Applying the eSwAP-In strategy, Kutyna et al. [66] designed and constructed a NeoChr incorporating diverse elements from the *S. cerevisiae* pan-genome, enabling the Sc2.0 parent strain to acquire a wide range of traits found in industrial and environmental isolates. Specifically, 17 unique genomic segments were identified from over 200 *S. cerevisiae* strains and integrated in silico into a single 211 kb DNA molecule, termed the pan-genomic NeoChr (PGNC). The design followed Sc2.0 principles, including substitution of all TAA stop codons with TAG, insertion of watermarks in 36 ORFs, and placement of 63 loxP sites. The authors compared circular (PGNC^{circ}) and linear (PGNC^{lin}) chromosomal variants, observing that the growth curve of PGNC^{circ} was comparable to that of the wild-type strain, whereas PGNC^{lin} showed a slightly extended lag phase and reduced cell density and maximum specific growth rate [66]. These findings imply that intracellular mechanisms governing NeoChr stability and function remain to be fully understood. Future improvements may rely on screening additional combinations of structural and regulatory elements.

The iterative processes involve the extraction of the assembled DNA from yeast, enrichment in vitro or in *E. coli*, and retransformation into yeast for subsequent rounds of assembly. To improve the import and export efficiency of large DNA across host cells and thereby increase overall assembly efficiency, He et al. developed a yeast life cycle-mediated iterative assembly method (YLC-assembly) [168]. In this approach, mating between haploid yeast strains brings two DNA fragments into a diploid cell for recombination, after which meiosis generates spores that can be used in the next assembly cycle. Although YLC-assembly achieves a high success rate for megabase-scale DNA assembly, each round takes approximately 11 days due to the sporulation process required [168]. To overcome this limitation, a haploidization-based DNA assembly and delivery strategy in yeast (HAnDy) was recently developed, enabling the de novo assembly of a 1.024-Mb NeoChr containing 542 exogenous genes [6]. Three key design principles guided the construction of the NeoChr: (1) clustering of functionally related genes

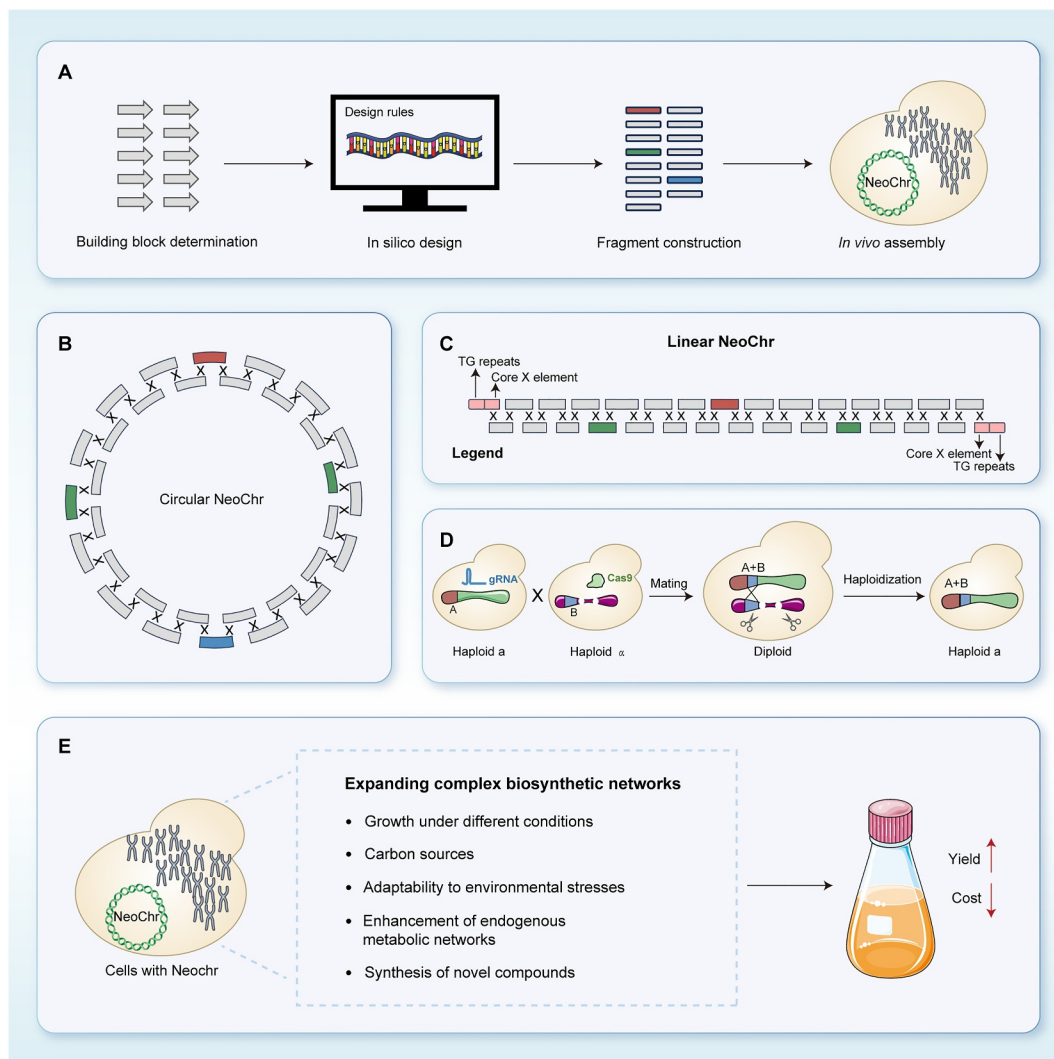


FIGURE 5 De novo construction of NeoChrs. (A) Schematic representation of the design and construction of NeoChrs. The in vivo assembly design for (B) circular NeoChrs and (C) linear NeoChrs. (D) Schematic illustration of CRISPR/Cas9-mediated haploidization. The red boxes represent the centromere for segregation. The green boxes represent ARS for replication. The blue boxes represent the telomerator for circular NeoChr. The pink box represents Tess for linear NeoChr. (E) NeoChr expands the capabilities of microbial cell factories (as illustrated by the yeast system). ARS, autonomously replicating sequence.

identified from 1011 *S. cerevisiae* isolates that are variable across the population but absent in S288C; (2) establishment of an orthogonal rearrangement system utilizing *loxP* and *vox* sites; and (3) incorporation of a precise genome editing system using “GGCC” sequences as specific gRNA target sites. The HANdy method embeds the DNA assembly process within a CRISPR/Cas9-mediated programmed haploidization cycle (Figure 5D). Briefly, two sets of episomal DNAs, preconstructed in haploid strains of opposite mating types, are brought together through mating. The DNAs are then assembled into a NeoChr via HR, and haploidization is completed through targeted genome elimination using CRISPR/Cas9. The resulting haploid strain can subsequently mate with another to initiate the next round of assembly.

5.2 | NeoChrs confer phenotypic plasticity to the host strain

The widely used laboratory strain *S. cerevisiae* S288C—and its derivatives typically used for yeast cell factory construction—lack many genes that contribute to the phenotypic diversity observed in industrial and environmental isolates. De novo-synthesized NeoChr function as mobile genetic modules that enhance host adaptation to various environmental stresses and enable the utilization of diverse carbon sources [6, 66]. For instance, the introduction of a PGNC into an Sc2.0-derived strain conferred distinct phenotypes compared to the parental strain. The PGNC strain exhibited at least a 2-fold increase in relative growth under nine different conditions, including D-melibiose, palatinose,

butyric acid, 5% sodium formate, 4% sodium lactate, neomycin, carbonyl cyanide 4-(trifluoromethoxy) phenylhydrazone (FCCP), deoxy-D-glucose, and ibuprofen [66]. Notably, only the PGNC-containing strain could metabolize alternative carbon sources such as palatinose (a disaccharide of glucose and fructose), melibiose (a disaccharide of glucose and galactose), and butyric acid. The ORF(s) responsible for these phenotypes can potentially be identified by analyzing intermediate PGNC strains obtained during stepwise assembly; consequently, their direct introduction into chassis strains can enhance tolerance to complex fermentation environments and increase product yield. Likewise, the introduction of NeoChrs assembled via HAnDy into six phylogenetically distinct yeast strains significantly improved host adaptability to various environmental stresses, including temperature, osmotic pressure, heavy metals, and the ability to use glutaric acid as a sole carbon source [6]. The NeoChr equips microbial cell factories with enhanced robustness for complex fermentation processes by enabling adaptation to environmental stresses, and the utilization of novel carbon sources further creates economic advantages by permitting the use of cheaper, renewable feedstocks (Figure 5E).

5.3 | NeoChrs expand the metabolic network of the host strain

NeoChrs not only enhance global metabolite production but also enable the synthesis of novel compounds in *S. cerevisiae*. Quantitative metabolite analysis of the host strain after the introduction of HAnDy-assembled NeoChrs confirmed the enrichment of eight endogenous metabolites, including pantothenic acid, myristic acid, and pipercolic acid. Furthermore, nine putative novel metabolites and the aromatic compound methyl anthranilate were detected [6]. Enhancing the production of key endogenous metabolites provides cell factories with abundant precursors, thereby increasing their overall biosynthetic capacity. Furthermore, enabling the synthesis of novel compounds allows for the development of new microbial production platforms and opens avenues for identifying previously unknown biosynthetic pathways (Figure 5E). In addition, NeoChrs offer significant advantages in pathway swapping [5, 64] and microbial cell factory construction [5]. The concept of modular pathway swapping involves the genetic reduction and clustering of genes from a specific metabolic pathway, allowing facile replacement with alternative configurations. This approach is particularly valuable for the metabolic reprogramming of industrial cell factories. For instance, genes encoding glycolysis have been relocated either to a single chromosomal locus [169] or to a defined landing pad on NeoChrs [5, 64]. Postma et al. [5] combined pathway swapping with NeoChrs to

rewire endogenous metabolic networks and introduce new heterologous pathways in *S. cerevisiae*. Specifically, 20 native genes involved in glycolysis, ethanol fermentation, and the pentose phosphate pathway were relocated from native chromosomes to NeoChrs. Additionally, the whole pathway for aromatic amino acid biosynthesis—including feedback-resistant alleles (10 genes)—along with 11 plant genes required for peltargonidin 3-O-glucoside (P3G) synthesis from tyrosine and phenylalanine, was assembled onto a single NeoChr. Expression of the de novo anthocyanin pathway from this NeoChr enabled anthocyanin production, whereas increasing the copy number of *coAtCHS3* and repairing *coAtANS* further enhanced biosynthetic efficiency.

6 | SUMMARY AND OUTLOOK

This review illustrates how large DNA fragment assembly has revolutionized the construction, modification, and evolution of microbial cell factories. In vitro assembly allows flexible assembly of standardized parts or synthetic pathways without host limitation, whereas in vivo assembly offers higher efficiency and stability for larger biosynthetic pathways (>10 kb). Together, they significantly accelerate strain construction and optimization. Synthetic genome assembly elevates microbial engineering beyond local optimization to constructing purpose-built ideal chassis. Assembly of additional NeoChrs further provides cell factories with a flexible and scalable platform, thereby enabling versatile network expansion. Ultimately, these large DNA fragment assembly techniques lay a solid foundation for industrial applications of advanced microbial cell factories in biomanufacturing.

However, large DNA fragment assembly strategies are usually applied to model organisms such as *E. coli* and *S. cerevisiae*, and their limited host diversity restricts broader application in metabolic engineering. On the one hand, future efforts should focus on developing efficient in vivo assembly systems for more hosts (e.g., *T. reesei*, *Vibrio natriegens*, and *Y. lipolytica*), expanding applications of these species in metabolic engineering and synthetic biology [44, 170, 171]. On the other hand, delivery technologies are needed to transfer assembled large DNA fragments across cells. Current methods, including bacterial conjugation [137], yeast mating [6], and induced cell fusion [172], have enabled large DNA fragment transfer between *E. coli* strains, yeast strains, and even across *E. coli* and yeast. Nevertheless, interspecies transfer efficiency remains low. Future work should therefore prioritize optimizing transfer conditions and elucidating the underlying fusion mechanisms to improve efficiency. Additionally, developing novel delivery systems to expand the host range, as well as leveraging nuclear

isolation-transplantation technologies such as synNICE for precise large DNA fragment transfer [173], are critical.

The engineering of microbial cell factories continues to be constrained by the time-consuming and costly process of large DNA fragment assembly. Several promising avenues are emerging to address these challenges. Advances in high-throughput DNA synthesis technologies such as microchip-based massive parallel synthesis are poised to provide higher-quality starting materials at a significantly reduced cost [10]. Automated high-throughput platforms, integrated with established methods such as Biopart Assembly Standard for Idempotent Cloning (BASIC) [174], Robot-assisted MoClo (RoboMoClo) [24], and Uracil-Specific Excision Reagent (USER) [175], enable robust and parallel construction of complex genetic constructs. Looking forward, the integration of these modular, automated DNA synthesis and sequencing technologies with automated assembly platforms is paving the way for fully automated biological system engineering. When further combined with AI-driven design, large DNA fragment assembly will power a high-throughput, high-precision Design-Build-Test-Learn (DBTL) cycle. Together, these advances are expected to slash the cost and time by orders of magnitude, thereby dramatically accelerating the development cycle of microbial cell factories.

In addition, despite advances in NeoChr and genome assembly, their application in engineering microbial cell factories remains nascent. Looking ahead, NeoChr assembly supports large functional modules “plug and play.” Combining the heterologous pathway module with interconnected metabolic modules (e.g., glycolysis, precursor synthesis, and cofactor regeneration) on NeoChrs allows for facile manipulation and testing of ultralarge metabolic pathways. The NeoChr or synthetic genome assembly also enables building pan-genomes or hybrid genomes to aggregate superior traits across hundreds or thousands of strains for expanding adaptability to complex industrial conditions [6, 66]. Moreover, the utilization of genome-recoded chassis (e.g., GCE in *E. coli*) provides for secure and efficient production of novel biochemicals. This paradigm is now extending to yeast for more precise strain regulation and biocontainment [176, 177]. In the near term, we are likely to see these technologies being applied for systematic, accurate, and in-depth engineering of industrial strains with hyper tolerance and productivity.

AUTHOR CONTRIBUTIONS

Yu Zhang: Conceptualization; investigation; project administration; visualization; writing—original draft; writing—review and editing. **Wenqian Liu:** Conceptualization; investigation; visualization; writing—original draft; writing—review and editing. **Dongyuan Cheng:** Conceptualization; investigation; visualization; writing—

original draft; writing—review and editing. **Jiazhang Lian:** Supervision; writing—review and editing. **Xian Fu:** Conceptualization; Funding acquisition; supervision; writing—review and editing. **Yue Shen:** Conceptualization; Funding acquisition; supervision; writing—review and editing.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (No. 2024YFA0916500), the National Natural Science Foundation of China (Nos. 32322047 and 32471490), and Jiangsu Provincial Department of Science and Technology (No. BM2023009). We thank Xiandi Zhang for her suggestions on the DNA synthesis technology and Zhang Zhang for her assistance on the illustration improvements.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article.

ETHICS STATEMENT

This article does not contain any studies with human or animal materials performed by any of the authors.

REFERENCES

- [1] Cho JS, Kim GB, Eun H, Moon CW, Lee SY. Designing microbial cell factories for the production of chemicals. *JACS Au*. 2022;2(8):1781–99.
- [2] Srinivasan P, Smolke CD. Biosynthesis of medicinal tropane alkaloids in yeast. *Nature*. 2020;585(7826):614–9.
- [3] Zhang J, Hansen LG, Gudich O, Viehriq K, Lassen LMM, Schrübbers L, et al. A microbial supply chain for production of the anti-cancer drug vinblastine. *Nature*. 2022;609(7926):341–7.
- [4] Liang F, Xie Y, Zhang C, Zhao Y, Motawia MS, Kampranis SC. Elucidation of the final steps in taxol biosynthesis and its biotechnological production. *Nat Synth*. 2025;4:1212–22.
- [5] Postma ED, Hassing EJ, Mangkusaputra V, Geelhoed J, de la Torre P, van den Broek M, et al. Modular, synthetic chromosomes as new tools for large scale engineering of metabolism. *Metab Eng*. 2022;72:1–13.
- [6] Ma Y, Su S, Fu Z, Zhou C, Qiao B, Wu Y, et al. Convenient synthesis and delivery of a megabase-scale designer accessory chromosome empower biosynthetic capacity. *Cell Res*. 2024;34(4):309–22.
- [7] Isaacs FJ, Carr PA, Wang HH, Lajoie MJ, Sterling B, Kraal L, et al. Precise manipulation of chromosomes *in vivo* enables genome-wide codon replacement. *Science*. 2011;333(6040):348–53.
- [8] Kosuri S, Church GM. Large-scale *de novo* DNA synthesis: technologies and applications. *Nat Methods*. 2014;11(5):499–507.
- [9] Yu M, Tang X, Li Z, Wang W, Wang S, Li M, et al. High-throughput DNA synthesis for data storage. *Chem Soc Rev*. 2024;53(9):4463–89.
- [10] Zhang X, Jiang X, Wang Y, Chen Q, Jiang H, Zhang H, et al. Scaling DNA synthesis with a microchip-based massively parallel synthesis system. *Nat Biotechnol*. 2025.

- [11] Hoose A, Vellacott R, Storch M, Freemont PS, Ryadnov MG. DNA synthesis technologies to close the gene writing gap. *Nat Rev Chem*. 2023;7(3):144–61.
- [12] Forget SM, Krawczyk MJ, Knight AM, Ching C, Copeland RA, Mahmoodi N, et al. Evolving a terminal deoxynucleotidyl transferase for commercial enzymatic DNA synthesis. *Nucleic Acids Res*. 2025;53(4):gkaf115.
- [13] Gao N, Yu A, Yang W, Zhang X, Shen Y, Fu X. Enzymatic *de novo* oligonucleotide synthesis: emerging techniques and advancements. *Biotechnol Adv*. 2025;82:108604.
- [14] Shetty RP, Endy D, Knight TF. Engineering BioBrick vectors from BioBrick parts. *J Biol Eng*. 2008;2:5.
- [15] Ray KA, Lutgens JD, Bista R, Zhang J, Desai RR, Hirsch M, et al. Assessing and harnessing updated polyketide synthase modules through combinatorial engineering. *Nat Commun*. 2024;15(1):6485.
- [16] Anderson JC, Dueber JE, Leguia M, Wu GC, Goler JA, Arkin AP, et al. BglBricks: a flexible standard for biological part assembly. *J Biol Eng*. 2010;4:1.
- [17] Bird JE, Marles Wright J, Giachino A. A user's guide to Golden Gate cloning methods and standards. *ACS Synth Biol*. 2022;11(11):3551–63.
- [18] Taylor GM, Mordaka PM, Heap JT. Start-Stop assembly: a functionally scarless DNA assembly system optimized for metabolic engineering. *Nucleic Acids Res*. 2019;47(3):e17.
- [19] Engler C, Youles M, Gruetzner R, Ehnert TM, Werner S, Jones JDG, et al. A Golden Gate modular cloning toolbox for plants. *ACS Synth Biol*. 2014;3(11):839–43.
- [20] Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. A modular cloning system for standardized assembly of multi-gene constructs. *PLoS One*. 2011;6(2):e16765.
- [21] Guo Y, Dong J, Zhou T, Auxillos J, Li T, Zhang W, et al. YeastFab: the design and construction of standard biological parts for metabolic engineering in *Saccharomyces cerevisiae*. *Nucleic Acids Res*. 2015;43(13):e88.
- [22] Zhou Y, Li G, Dong J, Xing XH, Dai J, Zhang C. MiYA, an efficient machine-learning workflow in conjunction with the YeastFab assembly strategy for combinatorial optimization of heterologous metabolic pathways in *Saccharomyces cerevisiae*. *Metab Eng*. 2018;47:294–302.
- [23] Blazquez B, León DS, Torres-Bacete J, Gómez-Luengo Á, Kniewel R, Martínez I, et al. Golden Standard: a complete Standard, portable, and interoperative MoClo tool for model and non-model proteobacteria. *Nucleic Acids Res*. 2023;51(19):e98.
- [24] Kang DH, Ko SC, Heo YB, Lee HJ, Woo HM. RoboMoClo: a robotics-assisted modular cloning framework for multiple gene assembly in biofoundry. *ACS Synth Biol*. 2022;11(3):1336–48.
- [25] Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods*. 2009;6(5):343–5.
- [26] Yan P, Zeng Y, Shen W, Tuo D, Li X, Zhou P. Nimble Cloning: a simple, versatile, and efficient system for standardized molecular cloning. *Front Bioeng Biotechnol*. 2019;7:460.
- [27] Li MZ, Elledge SJ. Harnessing homologous recombination *in vitro* to generate recombinant DNA via SLIC. *Nat Methods*. 2007;4(3):251–6.
- [28] Quan J, Tian J. Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS One*. 2009;4(7):e6441.
- [29] Li SY, Zhao GP, Wang J. C-Brick: a new standard for assembly of biological parts using Cpfl. *ACS Synth Biol*. 2016;5(12):1383–8.
- [30] Lei C, Li SY, Liu JK, Zheng X, Zhao GP, Wang J. The CCTL (Cpfl-assisted cutting and Taq DNA ligase-assisted ligation) method for efficient editing of large DNA constructs *in vitro*. *Nucleic Acids Res*. 2017;45(9):e74.
- [31] Enghiad B, Xue P, Singh N, Boob AG, Shi C, Petrov VA, et al. PlasmidMaker is a versatile, automated, and high throughput end-to-end platform for plasmid construction. *Nat Commun*. 2022;13(1):2697.
- [32] Fu J, Bian X, Hu S, Wang H, Huang F, Seibert PM, et al. Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bio prospecting. *Nat Biotechnol*. 2012;30(5):440–6.
- [33] Geistdoth-Kiener A, Totozafy JC, Le Goff G, Vergne J, Sakai K, Ouazzani J, et al. Yeast-based heterologous production of the coltochlorin family of fungal secondary metabolites. *Metab Eng*. 2023;80:216–31.
- [34] Wang H, Li Z, Jia R, Yin J, Li A, Xia L, et al. ExoCET: exonuclease *in vitro* assembly combined with RecET recombination for highly efficient direct DNA cloning from complex genomes. *Nucleic Acids Res*. 2018;46(5):e28.
- [35] Larionov V, Kouprina N, Graves J, Chen XN, Korenberg JR, Resnick MA. Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination. *Proc Natl Acad Sci U S A*. 1996;93(1):491–6.
- [36] Noskov VN, Karas BJ, Young L, Chuang RY, Gibson DG, Lin YC, et al. Assembly of large, high G plus C bacterial DNA fragments in yeast. *ACS Synth Biol*. 2012;1(7):267–73.
- [37] Kurylenko O, Paluszczak A, Luzhetskyy A, Rebets Y. An improved transformation-associated recombination cloning approach for direct capturing of natural product biosynthetic gene clusters. *Microb Biotechnol*. 2024;17(12):e70067.
- [38] Xu M, Wang W, Waglechner N, Culp EJ, Guitor AK, Wright GD. GPAHex-A synthetic biology platform for type IV-V glycopeptide antibiotic production and discovery. *Nat Commun*. 2020;11(1):5232.
- [39] Shao Z, Zhao H, Zhao H. DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res*. 2009;37(2):e16.
- [40] Shao Z, Luo Y, Zhao H. Rapid characterization and engineering of natural product biosynthetic pathways via DNA assembler. *Mol Biosyst*. 2011;7(4):1056–9.
- [41] Zhao Y, Che Y, Zhang F, Wang J, Gao W, Zhang T, et al. Development of an efficient pathway construction strategy for rapid evolution of the biodegradation capacity of *Pseudomonas putida* KT2440 and its application in bioremediation. *Sci Total Environ*. 2021;761:143239.
- [42] Jarczyńska ZD, Garcia Vanegas K, Deichmann M, Nørskov Jensen C, Scheeper MJ, Futyma ME, et al. A versatile *in vivo* DNA assembly toolbox for fungal strain engineering. *ACS Synth Biol*. 2022;11(10):3251–63.
- [43] Shi Y, Zhang L, Zhang M, Chu J, Xia Y, Yang H, et al. A CRISPR-Cas9 system-mediated genetic disruption and multi-fragment assembly in *Starmerella bombicola*. *ACS Synth Biol*. 2022;11(4):1497–509.
- [44] Fang Y, Meng X, Liu L, Li Z, Jia K, Liu W. Simultaneous *in vivo* assembly and targeted genome integration of gene clusters in *Trichoderma reesei*. *ACS Synth Biol*. 2025;14(2):575–84.
- [45] Wang K, Fredens J, Brunner SF, Kim SH, Chia T, Chin JW. Defining synonymous codon compression schemes by genome recoding. *Nature*. 2016;539(7627):59–64.
- [46] Robertson WE, Rehm FBH, Spinck M, Schumann RL, Tian R, Liu W, et al. *Escherichia coli* with a 57-codon genetic code. *Science*. 2025;390(6771):eady4368.
- [47] Zürcher JF, Kleefeldt AA, Funke LFH, Birnbaum J, Fredens J, Grazioli S, et al. Continuous synthesis of *E. coli* genome sections and Mb-scale human DNA assembly. *Nature*. 2023;619(7970):555–62.
- [48] Richardson SM, Mitchell LA, Stracquandano G, Yang K, Dymond JS, DiCarlo JE, et al. Design of a synthetic yeast genome. *Science*. 2017;355(6329):1040–4.
- [49] Luo J, Vale-Silva LA, Raghavan AR, Mercy G, Heldrich J, Sun X, et al. Synthetic chromosome fusion: effects on mitotic

- and meiotic genome structure and function. *Cell Genom.* 2023; 3(11):100439.
- [50] Xie ZX, Li BZ, Mitchell LA, Wu Y, Qi X, Jin Z, et al. "Perfect" designer chromosome V and behavior of a ring derivative. *Science.* 2017;355(6329):eaaf4704.
 - [51] Mitchell LA, Wang A, Stracquadanio G, Kuang Z, Wang X, Yang K, et al. Synthesis, debugging, and effects of synthetic chromosome consolidation: synvi and beyond. *Science.* 2017; 355(6329):eaaf4831.
 - [52] McCulloch LH, Sambasivam V, Hughes AL, Annaluru N, Ramalingam S, Fanfani V, et al. Consequences of a telomerase-related fitness defect and chromosome substitution technology in yeast synIX strains. *Cell Genom.* 2023; 3(11):100419.
 - [53] Wu Y, Li BZ, Zhao M, Mitchell LA, Xie ZX, Lin QH, et al. Bug mapping and fitness testing of chemically synthesized chromosome X. *Science.* 2017;355(6329):eaaf4706.
 - [54] Williams TC, Kroukamp H, Xu X, Wightman ELI, Llorente B, Borneman AR, et al. Parallel laboratory evolution and rational debugging reveal genomic plasticity to *S. cerevisiae* synthetic chromosome XIV defects. *Cell Genom.* 2023;3(11): 100379.
 - [55] Shen Y, Wang Y, Chen T, Gao F, Gong J, Abramczyk D, et al. Deep functional analysis of synII, a 770-kilobase synthetic yeast chromosome. *Science.* 2017;355(6329):eaaf4791.
 - [56] Shen Y, Gao F, Wang Y, Wang Y, Zheng J, Gong J, et al. Dissecting aneuploidy phenotypes by constructing Sc2.0 chromosome VII and SCRaMBLEing synthetic disomic yeast. *Cell Genom.* 2023;3(11):100364.
 - [57] Zhou C, Wang Y, Huang Y, An Y, Fu X, Yang D, et al. The *de novo* design and synthesis of yeast chromosome XIII facilitates investigations on aging. *Nat Commun.* 2024;15(1): 10139.
 - [58] Zhang W, Zhao G, Luo Z, Lin Y, Wang L, Guo Y, et al. Engineering the ribosomal DNA in a megabase synthetic chromosome. *Science.* 2017;355(6329):eaaf3981.
 - [59] Zhang W, Lazar-Stefanita L, Yamashita H, Shen MJ, Mitchell LA, Kurasawa H, et al. Manipulating the 3D organization of the largest synthetic yeast chromosome. *Mol Cell.* 2023; 83(23):4424–37.e5.
 - [60] Lauer S, Luo J, Lazar-Stefanita L, Zhang W, McCulloch LH, Fanfani V, et al. Context-dependent neocentromere activity in synthetic yeast chromosome VIII. *Cell Genom.* 2023;3(11): 100437.
 - [61] Goold HD, Kroukamp H, Erpf PE, Zhao Y, Kelso P, Calame J, et al. Construction and iterative redesign of synXVI a 903 kb synthetic *Saccharomyces cerevisiae* chromosome. *Nat Commun.* 2025;16(1):841.
 - [62] Blount BA, Lu X, Driessen MRM, Jovicevic D, Sanchez MI, Ciurkot K, et al. Synthetic yeast chromosome XI design provides a testbed for the study of extrachromosomal circular DNA dynamics. *Cell Genom.* 2023;3(11):100418.
 - [63] Foo JL, Kitano S, Susanto AV, Jin Z, Lin Y, Luo Z, et al. Establishing chromosomal design-build-test-learn through a synthetic chromosome and its combinatorial reconfiguration. *Cell Genom.* 2023;3(11):100435.
 - [64] Postma ED, Dashko S, van Breemen L, Taylor Parkins SK, van den Broek M, Daran JM, et al. A supernumerary designer chromosome for modular *in vivo* pathway assembly in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 2021;49(3): 1769–83.
 - [65] Jiang S, Luo Z, Wu J, Yu K, Zhao S, Cai Z, et al. Building a eukaryotic chromosome arm by *de novo* design and synthesis. *Nat Commun.* 2023;14(1):7886.
 - [66] Kutyna DR, Onetto CA, Williams TC, Goold HD, Paulsen IT, Pretorius IS, et al. Construction of a synthetic *Saccharomyces cerevisiae* pan-genome neo-chromosome. *Nat Commun.* 2022;13(1):3628.
 - [67] Xu P, Vansiri A, Bhan N, Koffas MAG. ePathBrick: a synthetic biology platform for engineering metabolic pathways in *E. coli*. *ACS Synth Biol.* 2012;1(7):256–66.
 - [68] Bian X, Plaza A, Zhang Y, Müller R. Luminmycins a-c, cryptic natural products from *Phototrhax luminescens* identified by heterologous expression in *Escherichia coli*. *J Nat Prod.* 2012; 75(9):1652–5.
 - [69] Tang Y, Frewert S, Harmrolfs K, Herrmann J, Karmann L, Kazmaier U, et al. Heterologous expression of an orphan NRPS gene cluster from *Paenibacillus larvae* in *Escherichia coli* revealed production of sevadicin. *J Biotechnol.* 2015;194: 112–4.
 - [70] Young R, Haines M, Storch M, Freemont PS. Combinatorial metabolic pathway assembly approaches and toolkits for modular assembly. *Metab Eng.* 2021;63:81–101.
 - [71] Pingoud A, Wilson GG, Wende W. Type II restriction endonucleases--a historical perspective and more. *Nucleic Acids Res.* 2014;42(12):7489–527.
 - [72] Naem M, Hao S, Chu M, Zhang X, Huang X, Wang J, et al. Efficient biosynthesis of D/L-alanine in the recombinant *Escherichia coli* BL21(DE3) by biobrick approach. *Front Bioeng Biotechnol.* 2024;12:1421167.
 - [73] Cheng F, Wang CJ, Gong XX, Sun KX, Liang XH, Xue YP, et al. Assembly and engineering of BioBricks to develop an efficient NADH regeneration system. *Appl Environ Microbiol.* 2025;91(1):e0104124.
 - [74] Sarrion-Perdigones A, Vazquez-Vilar M, Palací J, Castelijn B, Forment J, Ziarsolo P, et al. GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. *Plant Physiol.* 2013;162(3):1618–31.
 - [75] Zhang J, Li T, Hong Z, Ma C, Fang X, Zheng F, et al. Biosynthesis of hybrid neutral lipids with archaeal and eukaryotic characteristics in engineered *Saccharomyces cerevisiae*. *Angew Chem Int Ed Engl.* 2023;62(4):e202214344.
 - [76] Frey M, Bathe U, Meink L, Balcke GU, Schmidt J, Frolov A, et al. Combinatorial biosynthesis in yeast leads to over 200 diterpenoids. *Metab Eng.* 2024;82:193–200.
 - [77] Hamedirad M, Weisberg S, Chao R, Lian J, Zhao H. Highly efficient single-pot scarless Golden Gate assembly. *ACS Synth Biol.* 2019;8(5):1047–54.
 - [78] Sorida M, Bonasio R. An efficient cloning method to expand vector and restriction site compatibility of Golden Gate assembly. *Cell Rep Methods.* 2023;3(8):100564.
 - [79] Sarrion-Perdigones A, Falconi EE, Zandalinas SI, Juárez P, Fernández-del-Carmen A, Granell A, et al. GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules. *PLoS One.* 2011;6(7):e21622.
 - [80] Avilan L. Assembling multiple fragments: the Gibson assembly. *Methods Mol Biol.* 2023;2633:45–53.
 - [81] Li H, Xu S, Liu Y, Lu Y, Ning Y. Efficient *de novo* assembly of 100 kb-scale human functional immunoglobulin heavy variable (IGHV) gene fragments *in vitro*. *ACS Synth Biol.* 2025; 14(7):2559–71.
 - [82] Santos-Moreno J, Schaeferli Y. A framework for the modular and combinatorial assembly of synthetic gene circuits. *ACS Synth Biol.* 2019;8(7):1691–7.
 - [83] Xia Y, Li K, Li J, Wang T, Gu L, Xun L. T5 exonuclease-dependent assembly offers a low-cost method for efficient cloning and site-directed mutagenesis. *Nucleic Acids Res.* 2019;47(3):e15.
 - [84] Wang JW, Wang A, Li K, Wang B, Jin S, Reiser M, et al. CRISPR/Cas9 nuclease cleavage combined with Gibson assembly for seamless cloning. *Biotechniques.* 2015;58(4): 161–70.
 - [85] Kudo K, Hashimoto T, Hashimoto J, Kozono I, Kagaya N, Ueoka R, et al. *In vitro* Cas9-assisted editing of modular polypeptide synthase genes to produce desired natural product derivatives. *Nat Commun.* 2020;11(1):4022.

- [86] Wei P, Xue W, Zhao Y, Ning G, Wang J. CRISPR-based modular assembly of a UAS-cDNA/ORF plasmid library for more than 5500 *Drosophila* genes conserved in humans. *Genome Res.* 2020;30(1):95–106.
- [87] Guo Y, Cai G, Li H, Lin Z, Shi S, Jin J, et al. A CRISPR-Cas9-mediated large-fragment assembly method for cloning genomes and biosynthetic gene cluster. *Microorganisms.* 2024;12(7):1462.
- [88] Enghiad B, Zhao H. Programmable DNA-guided artificial restriction enzymes. *ACS Synth Biol.* 2017;6(5):752–7.
- [89] Liu J, Liu F, Luo X, Chen M, Wang C, Wang L, et al. Gibson assembly interposition improves amplification efficiency of long DNA and multifragment overlap extension PCR. *Bio-techniques.* 2023;74(6):286–92.
- [90] Qi DD, Jin J, Liu D, Jia B, Yuan YJ. *In vitro* and *in vivo* recombination of heterologous modules for improving biosynthesis of astaxanthin in yeast. *Microb Cell Fact.* 2020;19(1):103.
- [91] Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A.* 2000;97(12):6640–5.
- [92] Bian X, Huang F, Stewart FA, Xia L, Zhang Y, Müller R. Direct cloning, genetic engineering, and heterologous expression of the syringolin biosynthetic gene cluster in *E. coli* through Red/ET recombineering. *Chembiochem.* 2012;13(13):1946–52.
- [93] Jiang W, Zhao X, Gabrieli T, Lou C, Ebenstein Y, Zhu TF. Cas9-Assisted targeting of chromosome segments CATCH enables one-step targeted cloning of large gene clusters. *Nat Commun.* 2015;6:8101.
- [94] Jiang W, Zhu TF. Targeted isolation and cloning of 100-kb microbial genomic sequences by Cas9-assisted targeting of chromosome segments. *Nat Protoc.* 2016;11(5):960–75.
- [95] Liang M, Liu L, Xu F, Zeng X, Wang R, Yang J, et al. Activating cryptic biosynthetic gene cluster through a CRISPR-Cas12a-mediated direct cloning approach. *Nucleic Acids Res.* 2022;50(6):3581–92.
- [96] Wang H, Li Z, Jia R, Hou Y, Yin J, Bian X, et al. RecET direct cloning and redalpha-beta recombineering of biosynthetic gene clusters, large operons or single genes for heterologous expression. *Nat Protoc.* 2016;11(7):1175–90.
- [97] He F, Liu X, Tang M, Wang H, Wu Y, Liang S. CRISPR: an efficient technology for multiplexed refactoring of biosynthetic gene clusters. *Nucleic Acids Res.* 2024;52(18):11378–93.
- [98] Juhas M, Ajioka JW. High molecular weight DNA assembly *in vivo* for synthetic biology applications. *Crit Rev Biotechnol.* 2017;37(3):277–86.
- [99] Noskov VN, Koriabine M, Solomon G, Randolph M, Barrett JC, Leem SH, et al. Defining the minimal length of sequence homology required for selective gene isolation by TAR cloning. *Nucleic Acids Res.* 2001;29(6):E32.
- [100] Wang W, Zheng G, Lu Y. Recent advances in strategies for the cloning of natural product biosynthetic gene clusters. *Front Bioeng Biotechnol.* 2021;9:100085.
- [101] Kuijpers NG, Solis-Escalante D, Bosman L, van den Broek M, Pronk JT, Daran JM, et al. A versatile, efficient strategy for assembly of multi-fragment expression vectors in *Saccharomyces cerevisiae* using 60 bp synthetic recombination sequences. *Microb Cell Fact.* 2013;12:47.
- [102] Tang X, Li J, Millán-Aguinaga N, Zhang JJ, O'Neill EC, Ugalde JA, et al. Identification of thiotetronic acid antibiotic biosynthetic pathways by target-directed genome mining. *ACS Chem Biol.* 2015;10(12):2841–9.
- [103] Kouprina N, Kim J-H, Larionov V. Highly selective, CRISPR/Cas9-mediated isolation of genes and genomic loci from complex genomes by TARcloning in yeast. *Curr Protoc.* 2021;1(8):e207.
- [104] Karas BJ, Tagwerker C, Yonemoto IT, Hutchison CA 3rd, Smith HO. Cloning the *Acholeplasma laidlawii* PG-8A genome in *Saccharomyces cerevisiae* as a yeast centromeric plasmid. *ACS Synth Biol.* 2012;1(1):22–8.
- [105] Tagwerker C, Dupont CL, Karas BJ, Ma L, Chuang RY, Benders GA, et al. Sequence analysis of a complete 1.66 Mb *Prochlorococcus marinus* MED4 genome cloned in yeast. *Nucleic Acids Res.* 2012;40(20):10375–83.
- [106] Karas BJ, Molparia B, Jablanovic J, Hermann WJ, Lin YC, Dupont CL, et al. Assembly of eukaryotic algal chromosomes in yeast. *J Biol Eng.* 2013;7(1):30.
- [107] Walker EJJ, Pampuch M, Chang N, Cochrane RR, Karas BJ. Design and assembly of the 117-kb *Phaeodactylum tricornutum* chloroplast genome. *Plant Physiol.* 2024;194(4):2217–28.
- [108] Back D, O'Donnell TJ, Axt KK, Gurr JR, Vanegas JM, Williams PG, et al. Identification, heterologous expression, and characterization of the tolypodiol biosynthetic gene cluster through an integrated approach. *ACS Chem Biol.* 2023;18(8):1797–807.
- [109] Campos-Magana MA, dos Santos VAPM, Garcia-Morales L. Enabling access to novel bacterial biosynthetic potential from ONT draft genomic data. *Microb Biotechnol.* 2025;18(3):e70104.
- [110] Kim S-H, Lu W, Ahmadi MK, Montiel D, Ternei MA, Brady SF. Atolypenes, tricyclic bacterial sesterterpenes discovered using a multiplexed *in vitro* Cas9-TAR gene cluster refactoring approach. *ACS Synth Biol.* 2019;8(1):109–18.
- [111] Kim H, Ji CH, Je HW, Kim JP, Kang HS. mpCRISTAR: multiple plasmid approach for CRISPR/Cas9 and TAR-mediated multiplexed refactoring of natural product biosynthetic gene clusters. *ACS Synth Biol.* 2020;9(1):175–80.
- [112] Naseri G, Behrend J, Rieper L, Mueller-Roeber B. COMPASS for rapid combinatorial optimization of biochemical pathways based on artificial transcription factors. *Nat Commun.* 2019;10(1):2615.
- [113] Seong MJ, Yoon YR, Kwon KK, Kim H, Lee SG, Shin J, et al. Automated construction of a yeast-based multigene library via homologous recombination in a biofoundry workflow. *ACS Synth Biol.* 2025;14(5):1549–56.
- [114] Du J, Yuan Y, Si T, Lian J, Zhao H. Customized optimization of metabolic pathways by combinatorial transcriptional engineering. *Nucleic Acids Res.* 2012;40(18):e142.
- [115] Xu H, Shi L, Boob AG, Park W, Tan SI, Tran VG, et al. Discovery, characterization, and application of chromosomal integration sites for stable heterologous gene expression in *Rhodotorula toruloides*. *Metab Eng.* 2025;89:22–32.
- [116] Boob AG, Tan SI, Zaidi A, Singh N, Xue X, Zhou S, et al. Design of diverse, functional mitochondrial targeting sequences across eukaryotic organisms using variational autoencoder. *Nat Commun.* 2025;16(1):4151.
- [117] Guo F, Liu K, Qiao Y, Zheng Y, Liu C, Wu Y, et al. Evolutionary engineering of *saccharomyces cerevisiae*: crafting a synthetic methylotroph via self-reprogramming. *Sci Adv.* 2024;10(51):eadq3484.
- [118] Ogawa T, Iwata T, Kaneko S, Itaya M, Hirota J. An inducible RecA expression *Bacillus subtilis* genome vector for stable manipulation of large DNA fragments. *BMC Genom.* 2015;16(1):209.
- [119] Itaya M, Fujita K, Kuroki A, Tsuge K. Bottom-up genome assembly using the *Bacillus subtilis* genome vector. *Nat Methods.* 2008;5(1):41–3.
- [120] Itaya M, Tsuge K, Koizumi M, Fujita K. Combining two genomes in one cell: stable cloning of the synechocystis PCC6803 genome in the *Bacillus subtilis* 168 genome. *Proc Natl Acad Sci U S A.* 2005;102(44):15971–6.
- [121] Li J, Cheng L, Ma Y, Cai Y, Dai J. Recent technology advancements in large-scale DNA assembly. *Quant Biol.* 2024;13(1):e79.
- [122] Huang X, Sun Y, Liu S, Li Y, Li C, Sun Y, et al. Recombineering using RecET-like recombinases from *xenorhabdus* and

- its application in mining of natural products. *Appl Microbiol Biotechnol.* 2022;106(23):7857–66.
- [123] Hu S, Fu J, Huang F, Ding X, Stewart AF, Xia L, et al. Genome engineering of *Agrobacterium tumefaciens* using the lambda red recombination system. *Appl Microbiol Biotechnol.* 2014; 98(5):2165–72.
- [124] Pang Q, Ma S, Han H, Jin X, Liu X, Su T, et al. Phage enzyme-assisted direct *in vivo* DNA assembly in multiple microorganisms. *ACS Synth Biol.* 2022;11(4):1477–87.
- [125] Cello J, Paul AV, Wimmer E. Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. *Science.* 2002;297(5583):1016–8.
- [126] Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, et al. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science.* 2008;319(5867):1215–20.
- [127] Annaluru N, Muller H, Mitchell LA, Ramalingam S, Stracquadanio G, Richardson SM, et al. Total synthesis of a functional designer eukaryotic chromosome. *Science.* 2014; 344(6179):55–8.
- [128] Robertson WE, Funke LFH, de la Torre D, Fredens J, Elliott TS, Spinck M, et al. Sense codon reassignment enables viral resistance and encoded polymer synthesis. *Science.* 2021;372(6546):1057–62.
- [129] Spinck M, Piedrafita C, Robertson WE, Elliott TS, Cervettini D, de la Torre D, et al. Genetically programmed cell-based synthesis of non-natural peptide and depsipeptide macrocycles. *Nat Chem.* 2023;15(1):61–9.
- [130] Zürcher JF, Robertson WE, Kappes T, Petris G, Elliott TS, Salmond GPC, et al. Refactored genetic codes enable bidirectional genetic isolation. *Science.* 2022;378(6619):516–23.
- [131] Nyerges A, Vinke S, Flynn R, Owen SV, Rand EA, Budnik B, et al. A swapped genetic code prevents viral infections and gene transfer. *Nature.* 2023;615(7953):720–7.
- [132] Zhou S, Wu Y, Xie ZX, Jia B, Yuan YJ. Directed genome evolution driven by structural rearrangement techniques. *Chem Soc Rev.* 2021;50(22):12788–807.
- [133] Smith HO, Hutchison CA, 3rd, Pfannkoch C, Venter JC. Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides. *Proc Natl Acad Sci U S A.* 2003;100(26):15440–5.
- [134] Thi Nhu Thao T, Labrousseau F, Ebert N, V'kovski P, Stalder H, Portmann J, et al. Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. *Nature.* 2020;582(7813):561–5.
- [135] Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, et al. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science.* 2010; 329(5987):52–6.
- [136] Blattner FR, Plunkett G, 3rd, Bloch CA, Perna NT, Burland V, Riley M, et al. The complete genome sequence of *Escherichia coli* K-12. *Science.* 1997;277(5331):1453–62.
- [137] Fredens J, Wang K, de la Torre D, Funke LFH, Robertson WE, Christova Y, et al. Total synthesis of *Escherichia coli* with a recoded genome. *Nature.* 2019;569(7757):514–8.
- [138] Ostrov N, Landon M, Guell M, Kuznetsov G, Teramoto J, Cervantes N, et al. Design, synthesis, and testing toward a 57-codon genome. *Science.* 2016;353(6301):819–22.
- [139] Nyerges A, Chiappino-Pepe A, Budnik B, Baas-Thomas M, Flynn R, Yan S, et al. Synthetic genomes unveil the effects of synonymous recoding. *bioRxiv.* 2024(16):599206.
- [140] Zhao Y, Coelho C, Hughes AL, Lazar-Stefanita L, Yang S, Brooks AN, et al. Debugging and consolidating multiple synthetic chromosomes reveals combinatorial genetic interactions. *Cell.* 2023;186(24):5220–36.e16.
- [141] Guo Z, Yin H, Ma L, Li J, Ma J, Wu Y, et al. Direct transfer and consolidation of synthetic yeast chromosomes by abortive mating and chromosome elimination. *ACS Synth Biol.* 2022; 11(10):3264–72.
- [142] Lajoie MJ, Rovner AJ, Goodman DB, Aerni HR, Haimovich AD, Kuznetsov G, et al. Genomically recoded organisms expand biological functions. *Science.* 2013;342(6156):357–60.
- [143] Cheng Q, Arnér ES. Selenocysteine insertion at a predefined UAG codon in a release factor 1 (RF1)-depleted *Escherichia coli* host strain bypasses species barriers in recombinant selenoprotein translation. *J Biol Chem.* 2017;292(13): 5476–87.
- [144] Yi H, Zhang J, Ke F, Guo X, Yang J, Xie P, et al. Comparative analyses of the transcriptome and proteome of *Escherichia coli* C321.ΔA and further improving its noncanonical amino acids containing protein expression ability by integration of T7 RNA polymerase. *Front Microbiol.* 2021;12:744284.
- [145] Grome MW, Nguyen MTA, Moonan DW, Mohler K, Gurara K, Wang S, et al. Engineering a genomically recoded organism with one stop codon. *Nature.* 2025;639(8054):512–21.
- [146] Ma NJ, Isaacs FJ. Genomic recoding broadly obstructs the propagation of horizontally transferred genetic elements. *Cell Syst.* 2016;3(2):199–207.
- [147] Shen Y, Stracquadanio G, Wang Y, Yang K, Mitchell LA, Xue Y, et al. SCRaMbLE generates designed combinatorial stochastic diversity in synthetic chromosomes. *Genome Res.* 2016;26(1):36–49.
- [148] Luo Z, Wang L, Wang Y, Zhang W, Guo Y, Shen Y, et al. Identifying and characterizing SCRaMbLEd synthetic yeast using ReSCuES. *Nat Commun.* 2018;9(1):1930.
- [149] Ma L, Li Y, Chen X, Ding M, Wu Y, Yuan YJ. SCRaMbLE generates evolved yeasts with increased alkali tolerance. *Microb Cell Fact.* 2019;18(1):52.
- [150] Li Y, Wu Y, Ma L, Guo Z, Xiao W, Yuan Y. Loss of heterozygosity by SCRaMbLEing. *Sci China Life Sci.* 2019;62(3): 381–93.
- [151] Ong JY, Swidah R, Monti M, Schindler D, Dai J, Cai Y. SCRaMbLE: a study of its robustness and challenges through enhancement of hygromycin B resistance in a semi-synthetic yeast. *Bioengineering (Basel).* 2021;8(3):42.
- [152] Shen MJ, Wu Y, Yang K, Li Y, Xu H, Zhang H, et al. Heterozygous diploid and interspecies SCRaMbLEing. *Nat Commun.* 2018;9(1):1934.
- [153] Kang J, Li J, Guo Z, Zhou S, Su S, Xiao W, et al. Enhancement and mapping of tolerance to salt stress and 5-fluorocytosine in synthetic yeast strains via SCRaMbLE. *Synth Syst Biotechnol.* 2022;7(3):869–77.
- [154] Blount BA, Gowers GOF, Ho JCH, Ledesma-Amaro R, Jovicevic D, McKiernan RM, et al. Rapid host strain improvement by *in vivo* rearrangement of a synthetic yeast chromosome. *Nat Commun.* 2018;9(1):1932.
- [155] Zhang Y, Chiu TY, Zhang JT, Wang SJ, Wang SW, Liu LY, et al. Systematical engineering of synthetic yeast for enhanced production of lycopene. *Bioengineering (Basel).* 2021;8(1):14.
- [156] Jia B, Wu Y, Li BZ, Mitchell LA, Liu H, Pan S, et al. Precise control of SCRaMbLE in synthetic haploid and diploid yeast. *Nat Commun.* 2018;9(1):1933.
- [157] Liu W, Luo Z, Wang Y, Pham NT, Tuck L, Pérez-Pi I, et al. Rapid pathway prototyping and engineering using *in vitro* and *in vivo* synthetic genome SCRaMbLE-in methods. *Nat Commun.* 2018;9(1):1936.
- [158] Jia B, Jin J, Han M, Li B, Yuan Y. Directed yeast genome evolution by controlled introduction of trans-chromosomal structural variations. *Sci China Life Sci.* 2022;65(9):1703–17.
- [159] Gowers GF, Chee SM, Bell D, Suckling L, Kern M, Tew D, et al. Improved betulinic acid biosynthesis using synthetic yeast chromosome recombination and semi-automated rapid LC-MS screening. *Nat Commun.* 2020;11(1):868.

- [160] Zhang Y, Yuan M, Wu X, Zhang Q, Wang Y, Zheng L, et al. The construction and optimization of engineered yeast chassis for efficient biosynthesis of 8-hydroxygeraniol. *mLife*. 2023;2(4):438–49.
- [161] Jia ZC, Liu D, Ma HD, Cui YH, Li HM, Li X, et al. Yeast metabolic engineering for biosynthesis of caffeic acid-derived phenethyl ester and phenethyl amide. *ACS Synth Biol*. 2023;12(12):3635–45.
- [162] Luo Z, Yu K, Xie S, Monti M, Schindler D, Fang Y, et al. Compacting a synthetic yeast chromosome arm. *Genome Biol*. 2021;22(1):5.
- [163] Wang P, Xu H, Li H, Chen H, Zhou S, Tian F, et al. SCRaMbLEing of a synthetic yeast chromosome with clustered essential genes reveals synthetic lethal interactions. *ACS Synth Biol*. 2020;9(5):1181–9.
- [164] Liu Y, Zhao X, Gan F, Chen X, Deng K, Crowe SA, et al. Complete biosynthesis of QS-21 in engineered yeast. *Nature*. 2024;629(8013):937–44.
- [165] Shao Y, Lu N, Wu Z, Cai C, Wang S, Zhang LL, et al. Creating a functional single-chromosome yeast. *Nature*. 2018;560(7718):331–5.
- [166] de Vries ARG, Pronk JT, Daran J-MG. Industrial relevance of chromosomal copy number variation in *saccharomyces* yeasts. *Appl Environ Microbiol*. 2017;83(11):e03206–16.
- [167] Mitchell LA, McCulloch LH, Pinglay S, Berger H, Bosco N, Brosh R, et al. *De novo* assembly and delivery to mouse cells of a 101 kb functional human gene. *Genetics*. 2021;218(1):iyab038.
- [168] He B, Ma Y, Tian F, Zhao GR, Wu Y, Yuan YJ. YLC-assembly: large DNA assembly via yeast life cycle. *Nucleic Acids Res*. 2023;51(15):8283–92.
- [169] Kuijpers NGA, Solis-Escalante D, Luttik MAH, Bisschops MMM, Boonekamp FJ, van den Broek M, et al. Pathway swapping: toward modular engineering of essential cellular processes. *Proc Natl Acad Sci U S A*. 2016;113(52):15060–5.
- [170] Su C, Cui H, Wang W, Liu Y, Cheng Z, Wang C, et al. Bioremediation of complex organic pollutants by engineered *Vibrio natriegens*. *Nature*. 2025;642(8069):1024–33.
- [171] Zhang C, Liu M, Wang X, Cheng J, Xiang J, Yue M, et al. *De novo* synthesis of reticuline and taxifolin using re-engineered homologous recombination in *Yarrowia lipolytica*. *ACS Synth Biol*. 2025;14(2):585–97.
- [172] Karas BJ, Jablanovic J, Sun L, Ma L, Goldgof GM, Stam J, et al. Direct transfer of whole genomes from bacteria to yeast. *Nat Methods*. 2013;10(5):410–2.
- [173] Liu Y, Zhou J, Liu D, Hu X, Yang L, Song XR, et al. *De novo* assembly and delivery of synthetic megabase-scale human DNA into mouse early embryos. *Nat Methods*. 2025;22(8):1686–97.
- [174] Storch M, Haines MC, Baldwin GS. DNA-BOT: a low-cost, automated DNA assembly platform for synthetic biology. *Synth Biol (Oxf)*. 2020;5(1):ysaa010.
- [175] Cavaleiro AM, Kim SH, Seppälä S, Nielsen MT, Nørholm MHH. Accurate DNA assembly and genome engineering with optimized uracil excision cloning. *ACS Synth Biol*. 2015;4(9):1042–6.
- [176] Zhang H, Fu X, Gong X, Wang Y, Zhang H, Zhao Y, et al. Systematic dissection of key factors governing recombination outcomes by GCE-SCRaMbLE. *Nat Commun*. 2022;13(1):5836.
- [177] Chang T, Ding W, Yan S, Wang Y, Zhang H, Zhang Y, et al. A robust yeast biocontainment system with two-layered regulation switch dependent on unnatural amino acid. *Nat Commun*. 2023;14(1):6487.

How to cite this article: Zhang Y, Liu W, Cheng D, Lian J, Fu X, Shen Y. Advances in large DNA fragment assembly for microbial cell factory engineering. *Quantitative Biology*. 2026;e70039. <https://doi.org/10.1002/qub2.70039>