Engineering Microbiology 4 (2024) 100115

Contents lists available at ScienceDirect

Engineering Microbiology

journal homepage: www.elsevier.com/locate/engmic

Review

Recent advances in genome-scale engineering in *Escherichia coli* and their applications

Hui Gao[a,b], Zhichao Qiu[a,c,d], Xuan Wang[a], Xiyuan Zhang[a], Yujia Zhang[a,e], Junbiao Dai[b,f,*], Zhuobin Liang[a,*]

[a] Institute of Molecular Physiology, Shenzhen Bay Laboratory, Shenzhen 518132, China
[b] CAS Key Laboratory of Quantitative Engineering Biology, Guangdong Provincial Key Laboratory of Synthetic Genomics and Shenzhen Key Laboratory of Synthetic Genomics. Shenzhen Institute of Synthetic Biology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
[c] Regenerative Medicine Program, Bellvitge Institute for Biomedical Research (IDIBELL) and Program for Clinical Translation of Regenerative Medicine in Catalonia (P-CMRC), L' Hospitalet de Llobregat, Barcelona 08908, Spain
[d] Faculty of Pharmacy and Food Science, Barcelona University, Barcelona 08028, Spain
[e] College of Life Sciences, Northwest A&F University, Shaanxi 712100, China
[f] Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Key Laboratory of Synthetic Biology, Ministry of Agriculture and Rural Affairs, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518120, China

ARTICLE INFO

Keywords:
Genome-scale engineering
E. coli
Recombineering
CRISPR-Cas

ABSTRACT

Owing to the rapid advancement of genome engineering technologies, the scale of genome engineering has expanded dramatically. Genome editing has progressed from one genomic alteration at a time that could only be employed in few species, to the simultaneous generation of multiple modifications across many genomic loci in numerous species. The development and recent advances in multiplex automated genome engineering (MAGE)-associated technologies and clustered regularly interspaced short palindromic repeats and their associated protein (CRISPR-Cas)-based approaches, together with genome-scale synthesis technologies offer unprecedented opportunities for advancing genome-scale engineering in a broader range. These approaches provide new tools to generate strains with desired phenotypes, understand the complexity of biological systems, and directly evolve a genome with novel features. Here, we review the recent major advances in genome-scale engineering tools developed for *Escherichia coli*, focusing on their applications in identifying essential genes, genome reduction, recoding, and beyond.

1. Introduction

With rapidly accumulating genome sequences of new species and breakthroughs in gene synthesis and genome editing technologies, we now have the power to unveil gene functionality in a complex genomic network, discover desired phenotypic features, and design strains with valuable industrial applications [1–3]. *E. coli* is one of the most commonly used model organisms in biological research [4–6], and its engineered strains have been established for industrial applications, such as the production of humanized monoclonal antibody fragments (Fab) [7,8] and important chemicals [9,10]. Therefore, developing powerful tools for modifying the *E. coli* genome in a desired and controllable manner is of great priority [2,11]. Over the past several decades, scientific progress in *E. coli* genome editing has evolved from random genetic modifications to targeted editing at specific positions, enabling precise genetic changes. Furthermore, advancements have allowed for multiple and more precise modifications at the genome-scale.

Homologous recombination (HR)-based methods for introducing genetic variation are widely applicable to many organisms and have the advantage of introducing specific modifications at specific genomic sites [3,12,13]. However, *E. coli*'s recombination efficiency is relatively low, hampering efficient genome engineering using HR [14,15]. Early efforts were made to increase the recombination ability of *E. coli* by overexpressing its endogenous recombinases and associated factors [16,17];

Abbreviations: MAGE, multiplex automated genome engineering; CRISPR-Cas, clustered regularly interspaced short palindromic repeats and their associated protein; HR, homologous recombination; DSB, double-strand break; MMR, mismatch repair; SSAP, single-stranded DNA-annealing protein; CAGE, conjugate assembly genome engineering; gRNA, guide RNA; PAM, protospacer adjacent motif; crRNA, CRISPR RNA; sgRNA, single guide RNA; CRISPRi, CRISPR interference; BE, base editors; Target-AID, deaminase-mediated targeted nucleotide editing; CREPE, CRISPR-Cas9-mediated genomic error-prone editing; REXER, replicon excision for enhanced genome engineering through programmed recombination; GENESIS, genome stepwise interchange synthesis.

* Corresponding authors.
E-mail addresses: daijunbiao@caas.cn (J. Dai), zbliang@szbl.ac.cn (Z. Liang).

https://doi.org/10.1016/j.engmic.2023.100115
Received 31 May 2023; Received in revised form 12 September 2023; Accepted 14 September 2023
Available online 15 September 2023

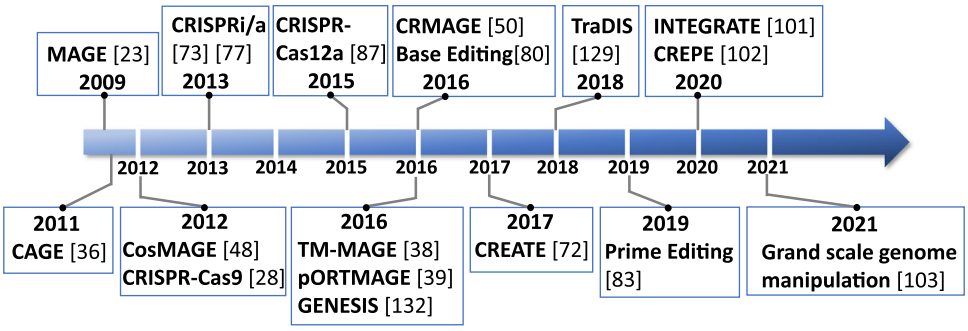

Fig. 1. Milestone development in genome engineering technology.

however, the practical application of these methods remains limited. The expression of bacteriophage-derived proteins was found to be more potent in editing a single to a few genomic sites [18,19], but is less efficient in editing multiple loci simultaneously due to low recombination frequency [19]. Although iterative implementation of the lambda red recombination system could fulfill the requirements for engineering biosynthetic pathways involving several proteins [20,21], the ability to more efficiently modify the *E. coli* genome is desperately required. The subsequent development of single-stranded DNA oligo-mediated Multiplex Automated Genome Engineering (MAGE) [22] offers a new platform for multiplex genome editing in *E. coli*. MAGE has been used to introduce small sequence changes iteratively to many sites in the *E. coli* genome, yielding strains with the desired genetic modifications and phenotypes [23]. Additionally, the CRISPR-Cas system is considered a new generation of genome engineering tools with great precision and accuracy [24–28]. In recent years, genome-scale engineering has coevolved with the development of new genome editing tools, which have greatly accelerated our understanding of microbial genomes and enabled us to reprogram biological systems on a whole-genome scale in a more controllable manner (Fig. 1). Developments include strategies to combine multiple powerful genetic editing tools and expand current genetic editing tools with new functionalities.

In the present review, we summarize the recent development of key genome engineering technologies which are applicable in genome-scale engineering of *E. coli* and discuss their applications in various areas. We begin by describing the basis of these technologies and the advances for enhancing their efficiency and reducing off-target effects. We then discuss the expansion of these technologies using diverse designs to enable genome-scale engineering (Table 1). Finally, we focus on genome-scale engineering applications in the fields of essential gene identification, genome size reduction, genome recoding, and chromosome restructuring.

2. MAGE-related genome engineering tools

MAGE introduces precise edits at multiple sites scarlessly through oligonucleotide-mediated allelic replacement in *E. coli* genome [23]. By repeatedly introducing a pool of targeting oligos into a cell and incorporating bacteriophage lambda red recombinase, MAGE enables many modifications at different chromosomal locations simultaneously and effectively without generating double-strand breaks (DSBs). Because MAGE is capable of multiplex genome engineering, it has been used to improve metabolism [23,29] and whole-genome recording [30–32] in *E. coli* and other organisms [30,33–35]. However, despite its potential in multiplex genome engineering, off-target mutations are still concerning [36]. In addition, an efficient method for target genotype screening in genome editing is lacking. To overcome these limitations, advances have been made to improve the properties of MAGE and its derivative techniques (Fig. 2) have been explored to expand its applications.

The central molecular mechanism of MAGE involves the introduction of genomic modifications via oligonucleotide annealing at the replication fork. Highly efficient genomic editing was achieved by disabling the mismatch repair (MMR) system in the original MAGE host EcNR2 strain (*mutS*$^{-}$) [23]. However, this modification results in increased undesired mutation rates, potentially affecting the strain phenotype. To overcome the limitations of high off-target mutation rates, conditional active MMR proteins are used instead of genetically disabling MMR genetically [37–39]. A plasmid-based system, called TM-MAGE, was observed with over a 5-fold reduction in off-target mutations by transiently inhibiting methyl-directed mismatch repair through inducible overexpression of Dam methylase [38]. Transient suppression of MMR was also achieved by temperature-controlled expression of the dominant-negative mutator allele of the *E. coli* MMR protein MutL [39]. As a component of the MutS-MutL complex, MutL is responsible for methyl-directed mismatch repair [40] and it allows a transient switch from the non-mutator to mutator phenotype of the host cell. The resulting pORTMAGE-expressing MutL exhibited high allelic replacement frequency and low off-target mutation rate [39]. Owing to the highly conserved nature of MutL, MMR suppression can also be extended to distant relatives of *E. coli*. Efforts have also been made to reduce the accumulation of undesired mutations by generating oligos with chemical modifications [41,42] or additional mismatches [43] to avoid mismatch detection and repair. Although higher recombination efficiency and reduced off-target effects were recorded, this method has limited applications because of the added costs and effort involved.

Furthermore, to expand the host range of MAGE applications, the required single-stranded DNA-annealing proteins (SSAPs) were screened and identified for better performance in *E. coli* and other hosts with a broader range [44,45]. Recently, a high-throughput method called “serial enrichment for efficient recombineering” (SEER) for SSAP discovery was reported [46]. A library of 122 diverse SSAP variants from seven different families was screened, and highly active variants were identified; CspRecT incorporation into pORTMAGE improved multiplex editing by 5- to 10-fold in *E. coli*. CspRecT also performed better than the lambda red system in implementing the CRISPR-Cpf1-assisted genome editing system in *E. coli* [47]. The specific interaction between SSAPs and the C-terminal peptide of the host single-stranded DNA-binding protein may contribute to the better compatibility of SSAPs between species [45].

In addition to efforts aimed at improving MAGE properties, various MAGE derivative techniques have been developed to expand its application. Similar to MAGE, conjugate assembly genome engineering (CAGE) hierarchically combines strains with the desired modifications in a pairwise manner through conjugation to construct chimeric genomes from many district strains [36]. Unlike natural conjugation, CAGE can control the direction and start site of conjugal transfer from the donor strain and precisely select the final composition of the genome with the desired modification by integrating *oriT* sequences into the donor strain genome with selectable markers. Compared with MAGE, CAGE provides the possibility of assembling large-scale chromosomal regions from many modified genomes [31,36], allowing the exploration of vast genetic modification effects in strains. By combining CAGE and MAGE, genome modification can be achieved at a higher resolution and on a larger scale than by using either technique individually.

Table 1
Summary of key genome-scale engineering tools.

Genome-scale engineering tools	Main characteristics	Applications	Limitations	Reference
MAGE	Multiplex automated genome engineering via oligonucleotide mediated allelic replacement	Facilitates the rapid and continuous generation of a diverse set of genetic changes, such as the overproduction of the industrially important isoprenoid lycopene	Limitation in insertion and deletion size	[23]
CAGE	Hierarchical conjugative assembly of many modified genomes	Genomic wide modifications, such as stop codon TAGs replacement	Potential limitations in performing genomic assembly within the same species	[36]
Cos-MAGE	Population enrichment for highly mutated genomes with Cos oligo supplements	Improves the efficiency of oligo-mediated recombineering; can be implemented without automated instrumentation	Genomic selectable markers are required	[48]
TM-MAGE	Transiently disables MMR during MAGE cycling by inducible overexpression of *E. coli* Dam methylase	Enables nearly equivalent efficiencies of allelic replacement compared strains with fully disabled MMR while maintaining an approximately 12- to 33-fold lower off-target mutation rate	Can potentially increase off-target mutation rate in subsequent cycles	[38]
pORTMAGE	Temperature-controlled expression of a dominant-negative mutator allele of the *E. coli* MMR protein MutL and related recombinase enzyme, enabling transient suppression of DNA repair during oligonucleotide integration	Applicable to a diverse set of enterobacterial species other than *E. coli* with high editing efficiency and low off-target effect	Specific enzymes and expression vectors are required	[39,46]
CRMAGE	Combines CRISPR-Cas9 and lambda red recombineering-based MAGE technology	Robustness in recombineering efficiency and compatibility with automation protocols	Additional steps are required to eliminate plasmids	[50]
CMGE	CRISPR-Cas9 assisted multiplex genome editing (CMGE) technique	Simultaneous editing of up to four loci in *E. coli* and can also be adapted for use in other prokaryotic cells	Low multi-site gene editing efficiency in *E. coli*	[66]
Target-AID	Deaminase-mediated targeted nucleotide editing	Simultaneous multiplex editing of six different genes and 41 loci in transposase genes	Editing capacity can be improved in the aspects of the various selection of PAMs and higher fidelity	[82]
CRISPR-Prime editing system	Introduces high-fidelity substitutions, deletions, insertions, and the combination thereof, both in plasmids and the chromosome of *E. coli*	Multiplexed editing in *E. coli* with the second guide RNA, providing a potential basis for developing similar toolkits for other bacteria	To achieve gene editing at more sites in *E. coli*, more plasmids need to be introduced, limiting the efficiency of gene editing	[84]
CREPE	Integrates an error-prone PCR donor libraries with the synonymous PAM mutation (SPM) on the genome, as CRISPR-Cas9- mediated genomic error-prone editing technology	Studies essential *E. coli* genes in their native genomic context	Limitation in target size and particular genomic regions	[102]
CREATE	Combines CRISPR-Cas9 and massively parallel oligomer synthesis including homologous repair cassettes that both edit loci and function as barcodes to track genotype–phenotype relationships	Achieves editing efficiencies of up to 98% in a single round of recombineering-and-CRISPR-selection; enables editing of a broad range of mutations at multiple loci in parallel	Potential for instability of a plasmid-based barcode for tracking codon level genotype–phenotype relationships	[72]
CRISPRi	Silences the gene by blocking gene transcription with dCas9:sgRNA complexes	Studies high-throughput interrogation of genome-wide gene functions and genetic interactions	Possible false positives and failures in designing sufficiency sgRNA for given locations when investigating gene functionality	[77]
INTEGRATE	Contains transposition proteins TnsA, TnsB, and TnsC, in conjunction with the RNA-guided DNA targeting complex TniQ-Cascade	Without reliance on homology arms specific to each target site, multiple simultaneous genomic insertions can be generated in the same cell using CRISPR arrays with multiple targeting spacers, achieving highly accurate and marker-free DNA integration of up to 10 kb bases at approximately 100% efficiency in bacteria	May not accomplish precise, scarless insertions or point mutations; further reductions in transposon end size or their conversion into functional parts are required	[101]
GENESIS	Iterative genome replacement by REXER, which is strictly dependent on CRISPR-Cas9 and the lambda red recombination machinery by simultaneously enhancing positive and negative selection	Enables radical and high-density changes to genomes that are not accessible through site-directed mutagenesis approaches; facilitates replacement of entire *E. coli* genome with synthetic DNA through iteration.		[103,132]

To further enhance multiplex genome engineering, Co-selection MAGE (CoS-MAGE) was applied to enrich highly mutated genomes [48] and substantially increase recombination frequencies [49] by incorporating genetic selection markers. Additionally, CRMAGE, a combination of CRISPR-Cas9 and lambda red recombineering-based MAGE technology, has shown efficient and rapid genome editing with negative selection against wild-type strains, which can be achieved through DSBs by CRISPR-Cas9 [50,51]. Computer-aided design software has also been used to develop an open-source web-based tool called Merlin, which improves MAGE performance by performing free-energy calculations and BLAST scoring on a sliding window spanning the targeted site [52]. Overall, the efficiency of MAGE-based technologies has improved since their initial development, and the continued increase in recombination gene efficiency and size will help understand complex microbial systems and developing novel strains for biotechnology applications.

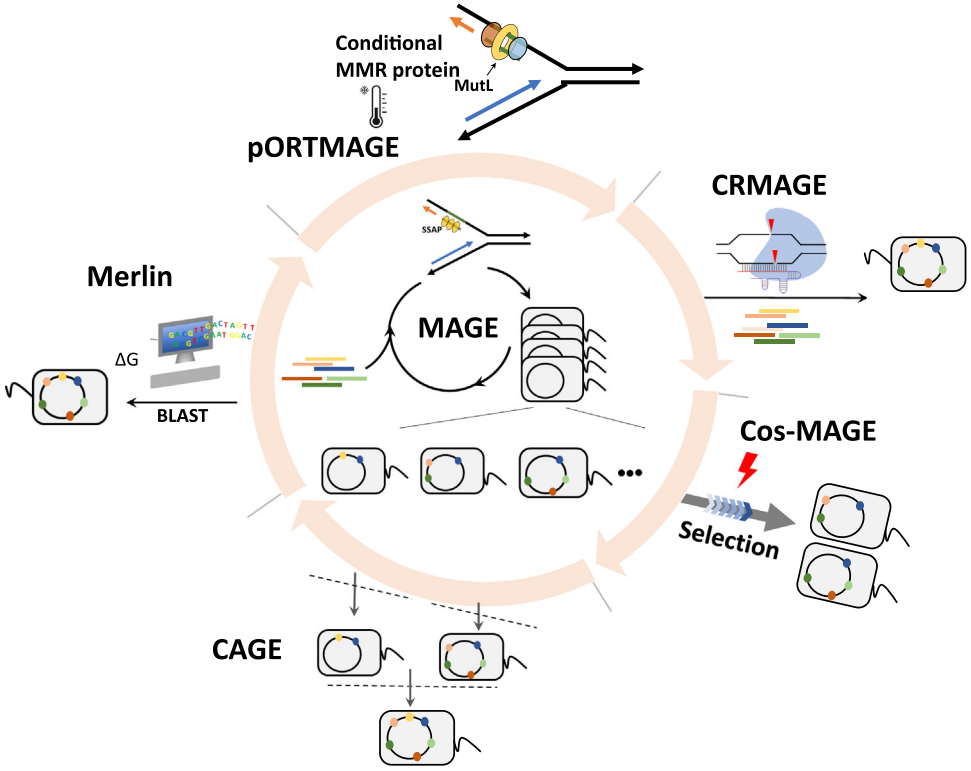

Fig. 2. MAGE-based technologies for *E.coli* genome engineering.

3. CRISPR-Cas system in genome-scale engineering

The CRISPR-Cas system, which was originally discovered in prokaryotic adaptive immunity, has been engineered as a highly valuable tool for genome editing in various organisms. Although only a decade has passed since the first report on programmable CRISPR-Cas systems [28], their remarkable genome editing capabilities have been extensively demonstrated [24,53–55]. The CRISPR-Cas system-based technology typically consists of three fundamental components: Cas nucleases, guide RNA (gRNA) derived from non-coding RNA sequences, and donor DNA [27,56,57]. Cas nucleases such as Cas9 are double-stranded (ds) DNA endonucleases that employ gRNA to precisely target and cleave DNA near a specific sequence known as the protospacer adjacent motif (PAM). The PAM site, typically a 2- to 5-base pair recognition sequence, varies across different CRISPR-Cas variants and plays a crucial role in facilitating cleavage [28]. The CRISPR-Cas system comprises a complex of CRISPR RNA (crRNA) and transactivating RNA (tracrRNA), which together facilitate the targeting. Notably, this RNA complex can be engineered into a single-guide RNA transcript, thereby effectively guiding the Cas nuclease to the desired target site for precise cleavage. Subsequently, the introduction of a double-stranded break at the target site triggers repair using an exogenous donor DNA as the repair template through the recombination machinery via HR or non-homologous end joining (NHEJ). CRISPR-Cas can achieve precise genome editing that targets single and multiple genes, introducing insertions, deletions, and nucleotide substitutions in *E. coli* [56,58]. Moreover, rational designs of these three components of the CRISPR-Cas system have enabled its applications for broader genome-scale engineering (Fig. 3).

3.1. CRISPR-Cas system for genome-scale editing based on gRNA engineering

Engineering gRNA is central to the ability of the CRISPR-Cas system to serve as a powerful tool for genome-scale editing. It plays a pivotal role in expanding the capability of multi-locus gene editing. Several strategies have been developed to optimize the gRNA design to improve efficiency [59–61] and minimize off-target effects [62], which have been reviewed elsewhere [63,64]. To achieve simultaneous multilocus gene editing, researchers have assembled multiple gRNA expression cassettes within a single plasmid. The expressed gRNAs act as guides for Cas endonucleases, directing them to multiple target sites for efficient editing. Sakuma et al. demonstrated a substantial improvement in the efficiency of multiplex gene editing by utilizing a single expression vector to deliver up to seven single guide RNA (sgRNAs) to human cells [65]. This advancement represents a major step forward in the effective delivery of multiple sgRNAs, which is crucial for efficiently employing multiplex CRISPR technology. Later, using a similar approach based on CRMAGE, CMAG was designed to simultaneously modify four loci in *E. coli* by constructing a complex plasmid containing multiple gRNA (sgRNA) cassettes [66]. The limitation of this gRNA expression approach lies in the potential challenge of the low delivery efficiency of large plasmids in the editing system. Remarkably, multiple gRNAs can be expressed by utilizing endogenous tRNA-processing machinery. This approach allows for the simultaneous editing of multiple genomic loci in various organisms, including plants [67], human cells [68], yeast [69], and other microbial species [70,71]. This system has been proven successful in increasing the efficiency of multiplex gene editing, indicating its potential to further augment the scale and speed of genetic modifications. Although attempts at creating gRNA-tRNA arrays have not been made in *E. coli*, the transformative potential of this technique in *E. coli* is promising, considering the presence of tRNA-processing systems in virtually all organisms [67]. Notable advancements have been made in genome-wide editing using gRNA libraries in the CRISPR-Cas system. One method, CRISPR-enabled trackable genome engineering (CREATE), uses episomal vectors containing a synthetic library of oligonucleotides as repair templates and a gRNA expression cassette to rapidly map genome-wide mutations [72]. Because of the potential instability of the plasmid-based barcode in this system, alternative barcoding strategies can be used to further enhance editing efficiency.

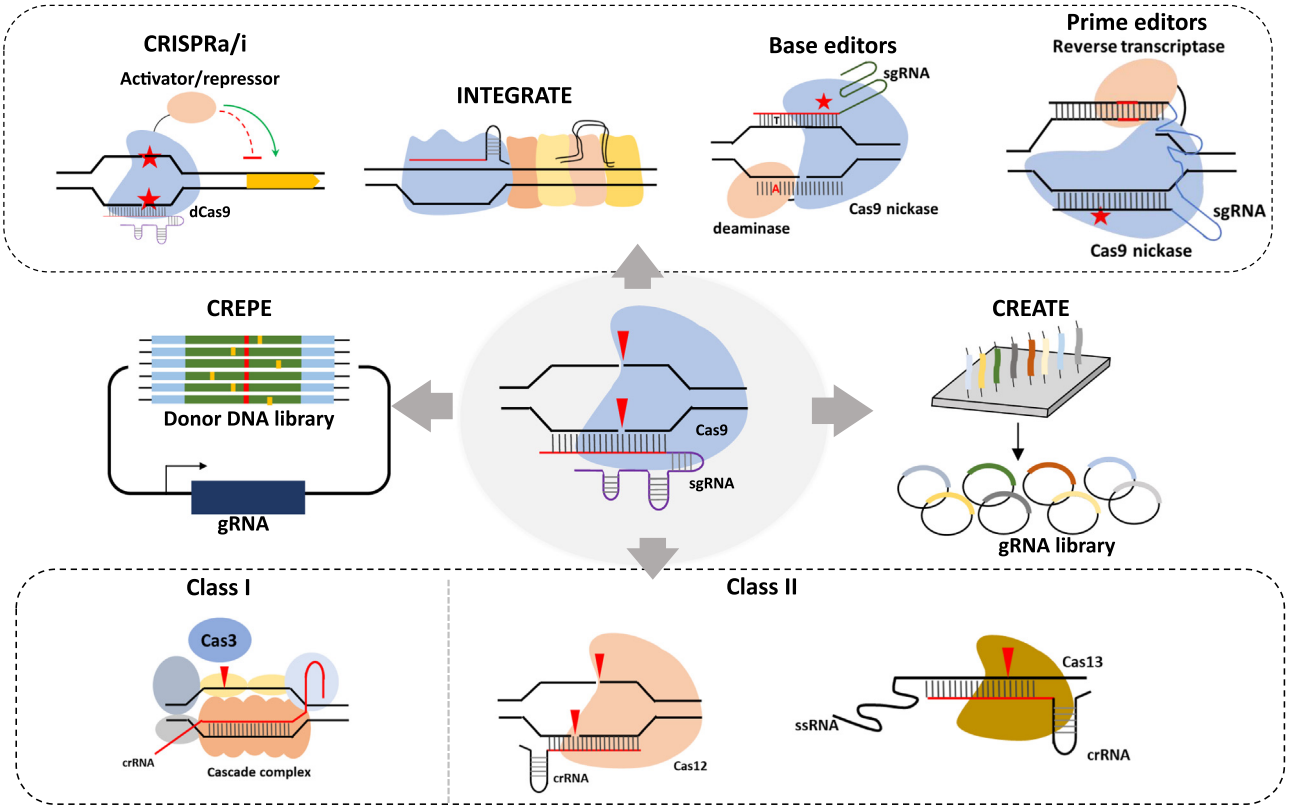

Fig. 3. CRISPR-Cas associated systems in genome-scale engineering.

3.2. CRISPR-Cas system for genome-scale editing based on Cas9 derivatives and other Cas endonucleases

While the expression of multiple gRNAs is vital for achieving multi-locus gene editing, Cas endonuclease plays a critical role in ensuring high editing efficiency and mitigating the challenges associated with off-target effects. Considerable efforts have been dedicated to expanding the application of Cas endonucleases in two primary directions. The first step involved the development of Cas9 derivatives. One such variant, dCas9, was created through two mutations (D10A and H840A in *Streptococcus pyogenes* Cas9). Although dCas9 lacks endonucleolytic activity, it retains its ability to bind to specific parts of the genome, enabling targeted binding without inducing DSBs [28,73,74]. dCas9 has been engineered as a programmable transcriptional repressor, resulting in a CRISPR interference (CRISPRi) system. CRISPRi can silence genes by obstructing gene transcription through dCas9:sgRNA complexes in *E. coli* [75,76]. It has also been used to simultaneously repress multiple target genes with high repression efficiency and site specificity [77]. Additionally, CRISPRi has been applied in genome-scale functional screening to guide future discoveries from rich genotype-phenotype maps of human cells [78]. This approach overcomes the limitations of single-cell CRISPR screening and allows the study of preselected genes [78]. Using *E. coli* as a microbial biosynthesis factory, the CRISPRi system was successfully employed to identify advantageous gene targets within related pathways. By perturbing 108 targets, researchers gained valuable insights into future gene modulation and their potential applications [79]. Moreover, chimeric dCas9 constructs with specific elements or compartments have been used to modulate gene transcription. For example, the fusion of dCas9 with the omega subunit of RNAP converts it into a transcription activator (CRISPRa) [73]. Base editors (BE) consisting of dCas9 fused with a nucleobase deaminase can replace a point mutation in the target locus with a desired base, either using cytidine deaminases [80] or adenosine deaminase [81]. Based on their ability to avoid double-stranded DNA cleavage, base editors have a low off-target rate, which makes them suitable for correcting pathogenic point mutations in human cells. Furthermore, BE has shown high potential for genome editing in *E. coli*. Deaminase-mediated targeted nucleotide editing (Target-AID), composed of dCas9 and *Petromyzon marinus* cytosine deaminase PmCDA1, allows simultaneous multiplex editing of six different genes and 41 loci located in transposase genes in *E. coli* [82]. Although BE technology reduces off-target effects compared to CRISPR-Cas9, it has limitations in editing specific sites because of the requirement for a PAM and a specific target site location. Prime editing (PE) provides more extensive editing capabilities with a broader range of edits than BE, including insertions, deletions, and point mutations [83]. The PE consists of a catalytically impaired Cas9 endonuclease fused to an engineered reverse transcriptase and a prime editing guide RNA that specifies the target site and encodes the desired edit [83]. Compared to BE, PE has demonstrated remarkable precision in human and mouse genomes, enabling the accurate insertion of genetic material up to approximately 40 base pairs and the deletion of sequences up to approximately 80 base pairs, with high ratios of desired edits to undesired byproducts [83]. Subsequently, a versatile genetic engineering toolkit based on CRISPR-prime editing with Cas9 H840A nickase (Cas9n) was reported [84]. By introducing only one nick into each DNA strand, this system can successfully modify *E. coli* genome with deletions of up to 97 bp and insertions of up to 33 bp. Although base and prime editing have substantially advanced gene editing technology in human cell lines, these techniques have not yet shown the capacity for larger genetic modifications in a genome-scale level in *E. coli* and other organisms.

CRISPR-Cas systems are highly diverse, but can be categorized into two distinct classes based on the organization of the effector module responsible for processing and adaptation [85]. Class 1 systems utilize multi-subunit effectors, whereas class 2 systems employ single large proteins as effectors. Among these, the class 2 system, characterized by its simplicity, has emerged as the most commonly used system in the field [24,86], with Cas9 being a prominent example. With the discovery of numerous CRISPR-Cas systems, novel class 2 CRISPR–Cas sys-

tems such as Cas12 [87–90] and Cas13 [91–94] have expanded substantially, which have also shown potential for genome editing [89,95,96] and RNA targeting and editing [94,97] in *E. coli*. Moreover, with the increased number of Cas endonucleases and effector complexes discovered in the CRISPR-Cas system, Class 1 systems have attracted attention and shown unique advantages and potential for advancing genome-scale engineering applications. Recently, a compact Cascade-Cas3 system has demonstrated its capacity for large targeted genome deletions without requiring a selection marker, which facilitates the manipulation of repetitive and noncoding regions of the genome [98]. A new CRISPR system, derived from the transposon-encoded CRISPR-Cas system [99], involving an RNA-guided DNA targeting complex, including TniQ, Cas6, Cas7, and Cas8, requires the transposon proteins TnsA, TnsB, and TnsC to direct the integration of large DNA fragments [100]. Later, the system with optimized insertion of transposable elements by guide RNA-assisted targeting (INTEGRATE), which contains transposition proteins TnsA, TnsB and TnsC, in conjunction with the RNA-guided DNA targeting complex TniQ-Cascade, can achieve marker-free DNA integration of up to 10 kb at approximately 100% efficiency [101]. These examples demonstrate a broad range of applications of Cas endonucleases in genome-scale engineering. Researchers have continued to refine genome-editing tools for broader and more precise applications.

3.3. CRISPR-Cas system for genome-scale editing based on donor DNA

While the system primarily utilizes guide RNA (gRNA) to direct Cas endonuclease to target sites, donor DNA, which serves as a template for introducing specific modifications, also plays a crucial role in facilitating specific genomic modifications. One research group developed the CRISPR-Cas9-mediated genomic error-prone editing (CREPE) technology, which uses error-prone PCR to generate donor libraries on the target genes in the genome [102]. Using this technology, a library of mutations can be developed across different essential and nonessential genomic loci in their native genetic context to introduce unbiased and diverse genomic mutant libraries with high editing efficiency. Additionally, donor DNA in the form of plasmids or BACs [103] has been used to insert large genes or regulatory elements into the genome, thereby enabling the study of gene function and regulation. By precisely modifying the *E. coli* genome, researchers can explore gene functions, engineer metabolic pathways, and develop customized strains for various biotechnological applications.

4. Gene synthesis-based genome-scale engineering

In addition to genome editing as a means of exploring genome functionality, genome synthesis offers a valuable approach for understanding genomic structures and functions. Recent advancements in gene synthesis technology have allowed for the construction of genome-sized fragments from oligonucleotides, enabling the generation of synthetic genomes with desired modifications to study genomic structures and functions. By synthesizing custom genomes, researchers can gain insights into the underlying principles of genome organization and function. This capability opens new avenues for investigating the relationships between genomic sequences and their phenotypic outcomes, and for designing and engineering genomes with specific properties for various applications.

An important milestone in synthetic biology was achieved through the *de novo* synthesis of infectious poliovirus cDNA, which successfully produced a synthetic virus displaying the biochemical and pathogenic characteristics of polioviruses [104]. This paved the way for subsequent groundbreaking research focused on the complete chemical synthesis and whole-genome design of *Mycoplasma*, the simplest bacterium with the smallest genome. The primary objectives of these efforts are to unravel the fundamental principles of life and decipher the intricate cellular operating systems. Initially, the synthetic genome of *Mycoplasmas* was chemically synthesized, assembled, and cloned into *Saccharomyces cerevisiae* [105]. This synthetic genome encompassed nearly all the genes found in the wild-type strain, and it featured "watermarks" to help track the modifications made during the synthesis. These remarkable achievements have contributed to our understanding of the intricacies of cellular life and offer valuable insights into genome design and engineering. With the development of genome transplantation methods, synthesized genomes have been successfully introduced into recipient strains, enabling sustained cell viability and self-replication of the cells [106]. Subsequently, by implementing three cycles of design, synthesis, and testing, the synthetic genome of *Mycoplasma mycoides* was constructed, with a genome smaller than that of any autonomously replicating cell found in nature [107]. These studies demonstrated the feasibility of rewriting genomes using large-scale fragment design, assembly, and testing. The Synthetic Yeast Genome Project (Sc2.0) opened a new era for synthetic genomics, ranging from prokaryotes to eukaryotes, with further development. The complete design and synthesis of synIII, resulting from systematic replacement of the native sequence of *S. cerevisiae* III with alternating selectable markers, was announced in 2014 [108]. Soon after this, other synthetic chromosomes were successfully built and their functionalities were confirmed [109–111], leading to the next version of the synthetic yeast genome, dubbed Sc3.0 [112]. *E. coli* has also been utilized as a model system for extensive genetic modifications at the genome-wide level, showing non-native features, such as the 57-codon genome [113]. Similar to the synthetic genome of *Mycoplasmas* construction, after computational design and de novo synthesis of the recoded fragments, segments were assembled in *S. cerevisiae* because of its high recombination efficiency. Successfully assembled segments were subsequently transformed into *E. coli* and the corresponding chromosomal sequence was deleted to test the functionality of the recoded genes. Eventually, a recoded *E. coli* genome with 57-codon was built, demonstrating the feasibility of genome-wide multiple codon replacements [113]. Furthermore, with genome-scale development, a synthetic *E. coli* genome with additional modifications for the removal of two sense codons and a stop codon was constructed, demonstrating the feasibility of synonymous codon compression [114]. Synthetic genomes have become more sophisticated, allowing us to examine biological questions that are impossible to answer using traditional approaches.

5. Application of genome-scale engineering

Extensive efforts have been made to understand genome functions and develop technologies for interpreting, designing, and building genes at the genome scale. Here, we summarize the applications of genome-scale engineering in *E. coli* over the past decade from the perspective of understanding cellular functions and reconstructing new strains with the desired phenotype (Fig. 4).

5.1. Characterization of genes' essentiality

Essential genes that are indispensable for the survival and reproduction of organisms under specific conditions [115,116], have been widely studied in prokaryotes [117,118] and eukaryotes [119,120]. Researchers have shown great interest in the essential genes of *E. coli* owing to their fundamental role in understanding bacterial evolution, the complexity of biological systems, and the modification and synthesis of genomes with desired features [121].

The traditional strategy for identifying essential genes in *E. coli* initially involved a combination of comparative genomics [122–126] and experimental approaches [127]. By exploiting these two strategies, emerging databases of essential genes in *E. coli* have been developed [127,128]. Recently, a high-throughput method called transposon-directed insertion site sequencing (TraDIS) was implemented as a benchmark against previously existing essential gene databases, eliminating the drawbacks of relying solely on single gene deletions [129]. This method can be used to reveal novel features at a fine resolution that

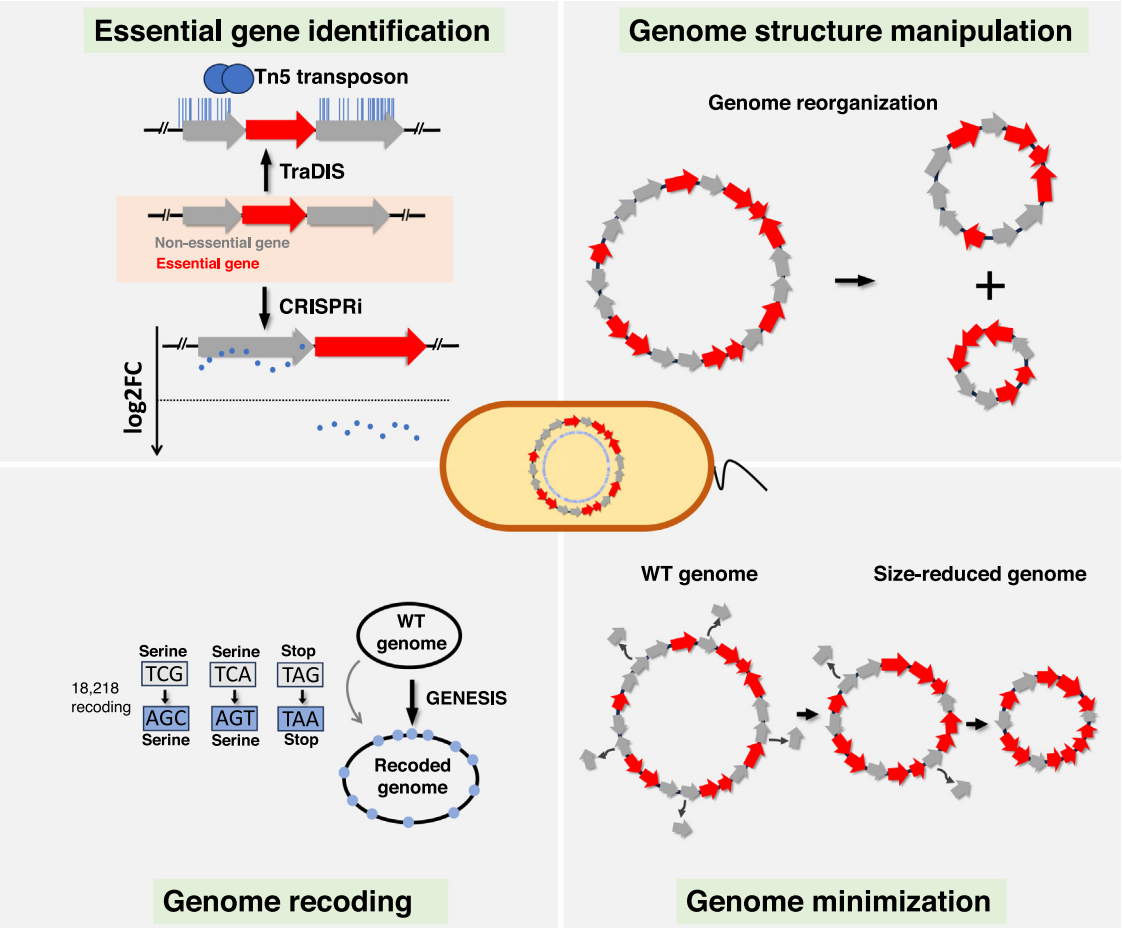

Fig. 4. Applications of genome-scale engineering.

would not have been detected using a simple conventional method; however, statistical analysis with gene-by-gene inspection of the insertion distribution within each individual gene could be biased toward genes with long coding regions.

In addition, as CRISPR-Cas technologies have emerged as versatile tools for editing genomes and modulating gene expression, several methods based on Cas9-mediated genome-wide high-throughput genetic screens have been developed to identify essential genes in *E. coli* [102,130] and other organisms [131]. A library of guide RNAs directing the dCas9 protein have been used to silence genes in the *E. coli* genome and assign gene essentiality. These pooled CRISPRi screenings can reliably identify essential genes at the genome level [130]. However, because of the intricate properties of the genome, this method may be restricted by limitations in designing sufficient sgRNAs due to the short coding length or extreme GC content in microbial genomes. Thus, an additional strategy could be used to address related problems, such as implementing an RNA-targeting CRISPR-Cas system to modulate gene expression at the RNA level and avoid direct genome editing. After identifying essential genes, CREPE can be employed to introduce the mutational library of the target essential gene with synonymous PAM (SPM) into the genome to study essential gene fitness landscapes in the native genetic context [102].

5.2. Genome recoding

Changing the genetic code has fundamental importance in understanding basic biological events and reprograming biological functions. Genome-wide stop codon replacement was achieved using MAGE and CAGE technologies. The 314 TAG stop codons were replaced with synonymous TAA codons in parallel across 32 *E. coli* strains through 18 MAGE cycles [36]. Hierarchical CAGE was used to assemble these strains with codon modifications into a single strain in vivo [36]. However, due to a deficiency in mismatch repair in the parent strain, the accumulation of off-target mutations was observed when 321 known TAGs were replaced with TAA, resulting in 355 off-target mutations [30]. Recently, Chin et al. created a synthetic *E. coli* strain that uses 61 codons for protein synthesis, with two sense codons (TCG and TCA) and a stop codon (TAG) removed [132]. This strain was built by combining two iterative approaches: replicon excision for enhanced genome engineering through programmed recombination (REXER) and genome stepwise interchange synthesis (GENESIS) [132]. A large synthetic fragment of more than 100 kb of *E. coli* genome with defined synonyms was replaced with genomic DNA through REXER, which was achieved using CRISPR-Cas9 coupled with lambda-red-mediated recombination and simultaneous positive and negative selection. Through consecutive REXER steps, known as GENESIS, the total synthesis of *E. coli* with a recoded genome was achieved, which was four times larger than that previously reported for genome replacement in any organism [114]. The same research team reported that by changing the structure of the genetic code and locking refactored codes into synthetic organisms, valuable engineered organisms could be protected from natural invaders using mobile genetic elements [133].

5.3. Manipulation of genome structure

As mentioned earlier, although the genome of *E. coli* is reduced by eliminating nonessential genes, the overall genome structure remains unchanged. Similar to eukaryotic organisms, such as *S. cerevisiae*, whose

genome structure exhibits high flexibility, the circular supercoiled chromosome of *E. coli* can also adopt different configurations. Advances in large-scale genome manipulation technologies have enabled genome manipulation in various formats. Recent studies have demonstrated that fragments as large as 700 kb can be translocated and inverted to different positions on the chromosome [134]. Additionally, the *E. coli* genome was shown to be efficiently split into two synthetic chromosomes through single-step programmed fission [134]. Conversely, they can be assembled from different strains into a single genome to form a new genomic structure [135]. Another study split the genome into three chromosomes and achieved stable maintenance in *E. coli* using the *ori-par* system [136]. These studies provide new insights into the reprogramming ability of *E. coli* genome and advance new technologies for synthetic biology and genomics. However, newly developed technologies for high-resolution three-dimensional map of the *E. coli* chromosome [137,138] have not yet been used in these engineered strains with new features. By comparing the similarities and differences in the chromosome folding mechanisms of these strains with those of the wild-type strain, we can gain insights into the general principles of DNA organization in living cells.

5.4. Genome minimization

With the increasing knowledge of essential genes and the concept of a minimal set of genes necessary to maintain a living cell, attempts have been made to reconstruct organisms with minimal gene sets to understand the fundamental interactions in cell systems through genomic engineering [102,139]. Cells with a minimized genome are believed to offer advantages, such as reduced transcriptional cost, elimination of competing functions, and limitation of unwanted regulatory interactions. Bottom-up and top-down approaches have been widely used to minimize the genomes of model bacteria [4,120,140]. The bottom-up approach creates minimal cells from synthetic fragments built with all essential genes to support cell survival. Because certain barriers remain in the bottom-up approach, such as the lack of related information on gene networks and their interactions, this approach has only been successfully applied in a limited number of cases. A notable example is the construction of JCVI-syn3.0, which contains approximately half the genome of the original strain *Mycoplasma mycoides* [107]. The top-down approach has been extensively used to reduce the genome of *E. coli* and other organisms through serial deletion of non-essential genes. Much effort has been made to reduce the *E. coli* genome using recombinase and selection markers [141] by employing methods such as I-SceI or Cre/loxp excision to eliminate the marker. To date, the second smallest genome has been that of DGF-298 (2.98 Mb), which has been reduced to 35.2% of the parent strain *E. coli* K-12 [142] without compromising the fitness. The construction of these strains relied on comparative genomics and gene deletion. Using more sophisticated bioinformatics tools in metabolic models [140] and gene interactions, we could discover potential genes that can be deleted. CRISPR-guided nickase systems have been demonstrated to enable the deletion of 133 kb of *E. coli* genome using single-stranded DNA breaks without causing cell lethality [143]. Incorporating newly developed genome editing tools into the field allows engineered *E. coli* strains with minimized genomes to be achieved.

6. Conclusions and outlook

The past decade has witnessed the development of new technologies for genome-scale engineering. Along with the systems derived from MAGE, the CRISPR-Cas system has demonstrated advancements in genetic modifications. Despite their rapid development and feasibility across diverse biological organisms, gene editing technologies that rely on the DNA DSBs mechanism have limitations such as the generation of cytotoxic effects or introduction of unwanted modifications during DSB repair. Although recent developments in base- and prime-editing methods have reduced the toxic effects by avoiding the introduction of DSB, neither of them provides a pathway-level introduction in the genome or genome-scale modification at base-pair resolution in *E. coli*. These technologies can be improved further to enable the introduction of large gene fragments into *E. coli* genome with desired phenotype. A newly developed technique called dReaMGE has shown promising results in the recombineering-assisted multiplex genome editing of kilobase-scale sequences with asymmetrically phosphonothioate modified dsDNA substrates without generating DSBs [144]. Furthermore, recently developed techniques, such as the implementation of retron-driven reverse-transcribed DNA for efficient genome modification, offer an alternative to using exogenous DNA as a template. This eliminates issues related to exogenous template delivery and availability [141–143]. With advancements in synthetic biology, genome-scale engineering can be integrated with synthetic biology principles to facilitate the construction of robust and predictable genetic circuits in *E. coli*. This includes the design and assembly of genetic modules, pathway engineering for the biosynthesis of valuable compounds, and implementation of genetic feedback control systems. These advances support the development of sustainable bioproduction processes and biosensors.

In summary, with the continued research and development of related technologies to improve editing efficiency, these emerging tools will expand the range of applications of *E. coli* in biotechnology, pharmaceuticals, and other fields, enabling the development of more efficient and sustainable processes.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Given his role as editorial board member, Dr. Junbiao Dai, had no involvement in the peer-review of this article and has no access to information regarding its peer-review. Full responsibility for the editorial process for this article was delegated to Dr. Xiaoying Bian.

CRediT authorship contribution statement

Hui Gao: Conceptualization, Writing – original draft, Writing – review & editing, Funding acquisition. **Zhichao Qiu:** Writing – original draft, Writing – review & editing. **Xuan Wang:** Writing – original draft, Writing – review & editing. **Xiyuan Zhang:** Writing – review & editing. **Yujia Zhang:** Writing – original draft, Writing – review & editing. **Junbiao Dai:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition. **Zhuobin Liang:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2018YFA0903700), the National Natural Science Foundation of China (32030004, 32150025, 31901020), Tianjin Synthetic Biotechnology Innovation Capacity Improvement Project (TSBICIP-PTJS-002), Guangdong Basic and Applied Basic Research Foundation (2023A1515030285), Shenzhen Science and Technology Program (KQTD20180413181837372) and Shenzhen Outstanding Talents Training Fund. Shenzhen Bay Laboratory startup funding.

References

[1] I. Thiele, B. Palsson, A protocol for generating a high-quality genome-scale metabolic reconstruction, Nat. Protoc. 5 (2010) 93–121, doi:10.1038/nprot.2009.203.
[2] J.D. Orth, T.M. Conrad, J. Na, J.A. Lerman, H. Nam, A.M. Feist, B. Palsson, A comprehensive genome-scale reconstruction of Escherichia coli metabolism-2011, Mol. Syst. Biol. 7 (2011) 1–9, doi:10.1038/msb.2011.65.
[3] S.K. Sharan, L.C. Thomason, S.G. Kuznetsov, D.L. Court, Recombineering: a homologous recombination-based method of genetic engineering, Nat. Protoc. 4 (2009) 206–223, doi:10.1038/nprot.2008.227.

[4] H. Kurasawa, T. Ohno, R. Arai, Y. Aizawa, A guideline and challenges toward the minimization of bacterial and eukaryotic genomes, Curr. Opin. Syst. Biol. 24 (2020) 127–134, doi:10.1016/j.coisb.2020.10.012.

[5] S. Pontrelli, T.-Y. Chiu, E.I. Lan, F.Y.-H. Chen, P. Chang, J.C. Liao, Escherichia coli as a host for metabolic engineering, Metab. Eng. 50 (2018) 16–46, doi:10.1016/j.ymben.2018.04.008.

[6] Z.D. Blount, The natural history of model organisms: the unexhausted potential of E. coli, Elife. 4 (2015) e05826, doi:10.7554/eLife.05826.

[7] S. Huleani, M.R. Roberts, L. Beales, H. Emmanouil, Escherichia coli as an antibody expression host for the production of diagnostic proteins : significance and expression, Crit. Rev. Biotechnol. 42 (2022) 756–773, doi:10.1080/07388551.2021.1967871.

[8] T.S. Castiñeiras, S.G. Williams, A.G. Hitchcock, D.C. Smith, E. coli strain engineering for the production of advanced biopharmaceutical products, FEMS Microbiol. Lett. 365 (2018) 1–10, doi:10.1093/femsle/fny162.

[9] M. Theisen, J.C. Liao, Industrial biotechnology: escherichia coli as a host, Ind. Biotechnol. (2017) 149–181, doi:10.1002/9783527807796.ch5.

[10] G.N.B. Chandresh Thakker, Irene Martínez, Ka-Yiu San, Succinate production in Escherichia coli, Biotechnol. J. 7 (2013) 213–224, doi:10.1002/biot.201100061.

[11] D. McCloskey, B. Palsson, A.M. Feist, Basic and applied uses of genome-scale metabolic network reconstructions of Escherichia coli, Mol. Syst. Biol. 9 (2013) 1–15, doi:10.1038/msb.2013.18.

[12] L.C. Thomason, J.A. Sawitzke, X. Li, N. Costantino, D.L. Court, Recombineering: genetic engineering in bacteria using homologous recombination, Curr. Protoc. Mol. Biol. 106 (2014) 1.16.1-1.16.39, doi:10.1002/0471142727.mb0116s106.

[13] D.L. Court, J.A. Sawitzke, L.C. Thomason, Genetic engineering using homologous recombination, Annu. Rev. Genet. 36 (2002) 361–388, doi:10.1146/annurev.genet.36.061102.093104.

[14] Y. Daiguan, E.H. M, L. E-Chiang, J.N. A, C.N. G, C.D. L, An efficient recombination system for chromosome engineering in Escherichia coli, Proc. Natl. Acad. Sci. 97 (2000) 5978–5983, doi:10.1073/pnas.100127597.

[15] G. Pines, E.F. Freed, J.D. Winkler, R.T. Gill, Bacterial recombineering: genome engineering via phage-based homologous recombination, ACS Synth. Biol. 4 (2015) 1176–1185, doi:10.1021/acssynbio.5b00009.

[16] Z. Youming, B. Frank, M. Joep, S. Francis, A new logic for DNA engineering using recombination in *Escherichia coli*, Nat. Genet. 20 (1998) 123–128.

[17] R. Kolodner, S.D. Hall, C. Luisi-DeLuca, Homologous pairing proteins encoded by the Escherichia coli recE and recT genes, Mol. Microbiol. 11 (1994) 23–30, doi:10.1111/j.1365-2958.1994.tb00286.x.

[18] K.C. Murphy, Use of bacteriophage λ recombination functions to promote gene replacement in Escherichia coli, 180 (1998) 2063–2071.

[19] D. Yu, H.M. Ellis, E. Lee, N.A. Jenkins, N.G. Copeland, D.L. Court, An efficient recombination system for chromosome engineering in Escherichia coli, Proc. Natl. Acad. Sci. 97 (2000), doi:10.1073/pnas.100127597.

[20] M. Juhas, J.W. Ajioka, Lambda Red recombinase-mediated integration of the high molecular weight DNA into the Escherichia coli chromosome, Microb. Cell Fact. 15 (2016) 1–10, doi:10.1186/s12934-016-0571-y.

[21] D.M. Bubnov, T.V. Yuzbashev, T.V. Vybornaya, A.I. Netrusov, S.P. Sineoky, Development of new versatile plasmid-based systems for λ Red-mediated Escherichia coli genome engineering, J. Microbiol. Methods. 151 (2018) 48–56, doi:10.1016/j.mimet.2018.06.001.

[22] T.M. Wannier, P.N. Ciaccia, A.D. Ellington, G.T. Filsinger, F.J. Isaacs, K. Javanmardi, M.A. Jones, A.M. Kunjapur, A. Nyerges, C. Pal, M.G. Schubert, G.M. Church, Recombineering and MAGE, Nat. Rev. Methods Prim. 1 (2021), doi:10.1038/s43586-020-00006-x.

[23] H.H. Wang, F.J. Isaacs, P.A. Carr, Z.Z. Sun, G. Xu, C.R. Forest, G.M. Church, Programming cells by multiplex genome engineering and accelerated evolution, Nature 460 (2009), doi:10.1038/nature08187.

[24] K.S. Makarova, Y.I. Wolf, J. Iranzo, S.A. Shmakov, O.S. Alkhnbashi, S.J.J. Brouns, E. Charpentier, D. Cheng, D.H. Haft, P. Horvath, S. Moineau, F.J.M. Mojica, D. Scott, S.A. Shah, V. Siksnys, M.P. Terns, Č. Venclovas, M.F. White, A.F. Yakunin, W. Yan, F. Zhang, R.A. Garrett, R. Backofen, J. van der Oost, R. Barrangou, E.V. Koonin, Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants, Nat. Rev. Microbiol. 18 (2020) 67–83, doi:10.1038/s41579-019-0299-x.

[25] A. Pickar-Oliver, C.A. Gersbach, The next generation of CRISPR–Cas technologies and applications, Nat. Rev. Mol. Cell Biol. 20 (2019) 490–507, doi:10.1038/s41580-019-0131-5.

[26] A.C. Komor, A.H. Badran, D.R. Liu, CRISPR-based technologies for the manipulation of eukaryotic genomes, Cell 168 (2017) 20–36, doi:10.1016/j.cell.2016.10.044.

[27] J.D. Sander, J.K. Joung, CRISPR-Cas systems for editing, regulating and targeting genomes, Nat. Biotechnol. 32 (2014) 347–355, doi:10.1038/nbt.2842.

[28] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, E. Charpentier, A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity, Science (80-) 337 (2012) 816–822.

[29] X. Zhu, D. Zhao, H. Qiu, F. Fan, S. Man, C. Bi, The CRISPR /Cas9-facilitated multiplex pathway optimization (CFPO) technique and its application to improve the Escherichia coli xylose utilization pathway, Metab. Eng. 43 (2017) 37–45, doi:10.1016/j.ymben.2017.08.003.

[30] M.J. Lajoie, A.J. Rovner, D.B. Goodman, H. Aerni, A.D. Haimovich, G. Kuznetsov, J. a Mercer, H.H. Wang, P. a Carr, J. a Mosberg, N. Rohland, P.G. Schultz, J.M. Jacobson, J. Rinehart, G.M. Church, F.J. Isaacs, Genomically recoded organisms expand biological functions, Science (80-) 342 (2013) 357–360.

[31] N.J. Ma, D.W. Moonan, F.J. Isaacs, Precise manipulation of bacterial chromosomes by conjugative assembly genome engineering, Nat. Protoc. 9 (2014) 2285–2300, doi:10.1038/nprot.2014.081.

[32] M.G. Napolitano, M. Landon, C.J. Gregg, M.J. Lajoie, L. Govindarajan, J.A. Mosberg, G. Kuznetsov, D.B. Goodman, O. Vargas-Rodriguez, F.J. Isaacs, D. Söll, G.M. Church, Emergent rules for codon choice elucidated by editing rare arginine codons in Escherichia coli, Proc. Natl. Acad. Sci. U. S. A. 113 (2016) E5588–E5597, doi:10.1073/pnas.1605856113.

[33] N.R. Sandoval, J.Y.H. Kim, T.Y. Glebes, P.J. Reeder, H.R. Aucoin, J.R. Warner, R.T. Gill, Strategy for directing combinatorial genome engineering in Escherichia coli, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 10540–10545, doi:10.1073/pnas.1206299109.

[34] S. Raman, J.K. Rogers, N.D. Taylor, G.M. Church, Evolution-guided optimization of biosynthetic pathways, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 17803–17808, doi:10.1073/pnas.1409523111.

[35] M. Amiram, A.D. Haimovich, C. Fan, Y.-S. Wang, H.-R. Aerni, I. Ntai, D.W. Moonan, N.J. Ma, A.J. Rovner, S.H. Hong, N.L. Kelleher, A.L. Goodman, M.C. Jewett, D. Söll, J. Rinehart, F.J. Isaacs, Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids, Nat. Biotechnol. 33 (2015) 1272–1279, doi:10.1038/nbt.3372.

[36] R.N.B. Isaacs FJ, P.A. Carr, H.H. Wang, M.J. Lajoie, B. Sterling, L. Kraal, A.C. Tolonen, T.A. Gianoulis, D.B. Goodman, Precise manipulation of chromosomes in vivo enables genome-wide codon replacement, Science (80-) 333 (2011) 348–353.

[37] C. Pa, Conditional DNA repair mutants enable highly precise genome engineering, Nucleic Acids Res 42 (2014), doi:10.1093/nar/gku105.

[38] R.M. Lennen, A.I.N. Wallin, M. Pedersen, M. Bonde, H. Luo, M.O.A. Sommer, M.J. Herrg, Transient overexpression of DNA adenine methylase enables efficient and mobile genome engineering with reduced off-target effects, Nucleic Acids Res 44 (2016) 1–14, doi:10.1093/nar/gkv1090.

[39] I. Nagy, B. Bálint, P. Bihari, V. Lázár, G. Apjok, A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species, Proc. Natl. Acad. Sci. 113 (2016), doi:10.1073/pnas.1520040113.

[40] A. Aronshtam, M.G. Marinus, Dominant negative mutator mutations in the mutL gene of Escherichia coli, Nucleic Acids Res. 24 (1996) 2498–2504.

[41] H.H. Wang, G. Xu, A.J. Vonner, G. Church, Modified bases enable high-efficiency oligonucleotide-mediated allelic replacement via mismatch repair evasion, Nucleic Acids Res. 39 (2011) 7336–7347, doi:10.1093/nar/gkr183.

[42] T.W. Van Ravesteyn, M. Dekker, A. Fish, T.K. Sixma, A. Wolters, R.J. Dekker, LNA modification of single-stranded DNA oligonucleotides allows subtle gene modification in mismatch-repair-proficient cells, Proc. Natl. Acad. Sci 113 (2016), doi:10.1073/pnas.1513315113.

[43] J.A. Sawitzke, N. Costantino, X.T. Li, L.C. Thomason, M. Bubunenko, C. Court, D.L. Court, Probing cellular processes with oligo-mediated recombination and using the knowledge gained to optimize recombineering, J. Mol. Biol. 407 (2011) 45–59, doi:10.1016/j.jmb.2011.01.030.

[44] T.M. Wannier, A. Nyerges, H.M. Kuchwara, M. Czikkely, D. Balogh, G.T. Filsinger, N.C. Borders, C.J. Gregg, M.J. Lajoie, X. Rios, C. Pál, G.M. Church, Improved bacterial recombineering by parallelized protein discovery, Proc. Natl. Acad. Sci. U. S. A. 117 (2020) 13689–13698, doi:10.1073/pnas.2001588117.

[45] G.T. Filsinger, T.M. Wannier, F.B. Pedersen, I.D. Lutz, J. Zhang, D.A. Stork, A. Debnath, K. Gozzi, H. Kuchwara, V. Volf, S. Wang, X. Rios, C.J. Gregg, M.J. Lajoie, S.L. Shipman, J. Aach, M.T. Laub, G.M. Church, Characterizing the portability of phage-encoded homologous recombination proteins, Nat. Chem. Biol. 17 (2021) 394–402, doi:10.1038/s41589-020-00710-5.

[46] T.M. Wannier, A. Nyerges, H.M. Kuchwara, M. Czikkely, D. Balogh, Improved bacterial recombineering by parallelized protein discovery, Proc. Natl. Acad. Sci. 117 (2020), doi:10.1073/pnas.2001588117.

[47] X. Zhu, Y. Wu, X. Lv, Y. Liu, G. Du, J. Li, L. Liu, Combining CRISPR-Cpf1 and recombineering facilitates fast and efficient genome editing in Escherichia coli, ACS Synth. Biol. (2022), doi:10.1021/acssynbio.2c00041.

[48] H.H. Wang, H. Kim, L. Cong, J. Jeong, D. Bang, G.M. Church, Genome-scale promoter engineering by coselection MAGE, Nat. Methods. 9 (2012) 591–593, doi:10.1038/nmeth.1971.

[49] P.A. Carr, H.H. Wang, B. Sterling, F.J. Isaacs, M.J. Lajoie, G. Xu, G.M. Church, J.M. Jacobson, Enhanced multiplex genome engineering through co-operative oligonucleotide co-selection, Nucleic Acids Res. 40 (2012), doi:10.1093/nar/gks455.

[50] C. Ronda, L.E. Pedersen, M.O.A. Sommer, A.T. Nielsen, CRMAGE: CRISPR optimized MAGE recombineering, Sci. Rep. 6 (2016) 1–11, doi:10.1038/srep19452.

[51] K. Umenhoffer, G. Draskovits, Á. Nyerges, I. Karcagi, B. Bogos, E. Tímár, B. Csörgő, R. Herczeg, I. Nagy, T. Fehér, C. Pál, G. Pósfai, Genome-wide abolishment of mobile genetic elements using genome shuffling and CRISPR/Cas-assisted MAGE allows the efficient stabilization of a bacterial chassis, ACS Synth. Biol. 6 (2017) 1471–1483, doi:10.1021/acssynbio.6b00378.

[52] M. Quintin, N.J. Ma, S. Ahmed, S. Bhatia, A. Lewis, F.J. Isaacs, D. Densmore, Merlin: computer-aided oligonucleotide design for large scale genome engineering with mage, ACS Synth. Biol. 5 (2016) 452–458, doi:10.1021/acssynbio.5b00219.

[53] A.C. Komor, A.H. Badran, D.R. Liu, CRISPR-based technologies for the manipulation of eukaryotic genomes, Cell 168 (2017) 20–36, doi:10.1016/j.cell.2016.10.044.

[54] A. Pickar-Oliver, C.A. Gersbach, The next generation of CRISPR–Cas technologies and applications, Nat. Rev. Mol. Cell Biol 20 (2019) 490–507, doi:10.1038/s41580-019-0131-5.

[55] C. Le, R.F. Ann, C. David, L. Shuailiang, B. Robert, H. Naomi, H.P. D, W. Xuebing, J. Wenyan, M.L. A, Z. Feng, Multiplex genome engineering using CRISPR/Cas systems, Science (80-) 339 (2013) 819–823, doi:10.1126/science.1231143.

[56] W. Jiang, D. Bikard, D. Cox, F. Zhang, L.A. Marraffini, RNA-guided editing of bacterial genomes using CRISPR-Cas systems, Nat. Biotechnol. 31 (2013) 233–239, doi:10.1038/nbt.2508.
[57] G. Liu, Q. Lin, S. Jin, C. Gao, The CRISPR-Cas toolbox and gene editing technologies, Mol. Cell. 82 (2022) 333–347, doi:10.1016/j.molcel.2021.12.002.
[58] Y. Jiang, B. Chen, C. Duan, B. Sun, J. Yang, S. Yang, Multigene editing in the Escherichia coli genome via the CRISPR-Cas9 system, Appl. Environ. Microbiol. 81 (2015) 2506–2514, doi:10.1128/AEM.04023-14.
[59] S. Kiani, A. Chavez, M. Tuttle, R.N. Hall, R. Chari, D. Ter-ovanesyan, J. Qian, B.W. Pruitt, J. Beal, S. Vora, J. Buchthal, E.J.K. Kowal, M.R. Ebrahimkhani, J.J. Collins, R. Weiss, G. Church, Cas9 gRNA engineering for genome editing, activation and repression, Nat. Methods. 12 (2015), doi:10.1038/nmeth.3580.
[60] Y. Wang, Y. Wang, D. Pan, Y. Wang, Y. Wang, D. Pan, H. Yu, Y. Zhang, W. Chen, F. Li, Z. Wu, Guide RNA engineering enables efficient CRISPR editing with a miniature Syntrophomonas palmitatica Cas12f1 nuclease, CellReports 40 (2022) 111418, doi:10.1016/j.celrep.2022.111418.
[61] J. Guo, T. Wang, C. Guan, B. Liu, C. Luo, Z. Xie, C. Zhang, X. Xing, Improved sgRNA design in bacteria via genome-wide activity profiling, Nucleic Acids Res 46 (2018) 7052–7069, doi:10.1093/nar/gky572.
[62] J. Wu, H. Yin, Engineering guide RNA to reduce the off-target effects of CRISPR, J. Genet. Genomics. 46 (2019) 523–529, doi:10.1016/j.jgg.2019.11.003.
[63] D. Allen, M. Rosenberg, A. Hendel, Using synthetically engineered guide RNAs to enhance CRISPR genome editing systems in mammalian cells, Front, Genome Ed. 2 (2021) 1–16, doi:10.3389/fgeed.2020.617910.
[64] S. Bin Moon, D.Y. Kim, J.-H. Ko, J.-S. Kim, Y.-S. Kim, Improving CRISPR genome editing by engineering guide RNAs, Trends Biotechnol. 37 (2019) 870–881, doi:10.1016/j.tibtech.2019.01.009.
[65] T. Sakuma, A. Nishikawa, S. Kume, K. Chayama, T. Yamamoto, Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system, Sci. Rep. 4 (2014) 4–9, doi:10.1038/srep05400.
[66] X. Feng, D. Zhao, X. Zhang, X. Ding, C. Bi, CRISPR/Cas9 assisted multiplex genome editing technique in Escherichia coli, Biotechnol. J. 13 (2018), doi:10.1002/biot.201700604.
[67] K. Xie, B. Minkenberg, Y. Yang, Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 3570–3575, doi:10.1073/pnas.1420294112.
[68] Q. Yuan, X. Gao, Multiplex base- and prime-editing with drive-and-process CRISPR arrays, Nat. Commun. 13 (2022) 1–13, doi:10.1038/s41467-022-30514-1.
[69] Y. Zhang, J. Wang, Z. Wang, Y. Zhang, S. Shi, J. Nielsen, Z. Liu, A gRNA-tRNA array for CRISPR-Cas9 based rapid multiplexed genome editing in Saccharomyces cerevisiae, Nat. Commun. 10 (2019) 1–10, doi:10.1038/s41467-019-09005-3.
[70] L. Song, J. Ouedraogo, M. Kolbusz, T. Truc, M. Nguyen, A. Tsang, Efficient genome editing using tRNA promoter-driven CRISPR /Cas9 gRNA in Aspergillus niger, PLoS ONE (2018) 1–17.
[71] C. Gene, G. Silencing, Development of a gRNA expression and processing platform for efficient CRISPR-Cas9-based gene editing and gene silencing in Candida tropicalis, Microbiol. Spectr. 9 (2022) 1–14.
[72] A.D. Garst, M.C. Bassalo, G. Pines, S.A. Lynch, A.L. Halweg-Edwards, R. Liu, L. Liang, Z. Wang, R. Zeitoun, W.G. Alexander, R.T. Gill, Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering, Nat. Biotechnol. 35 (2017) 48–55, doi:10.1038/nbt.3718.
[73] D. Bikard, W. Jiang, P. Samai, A. Hochschild, F. Zhang, L.A. Marraffini, Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system, Nucleic Acids Res. 41 (2013) 7429–7437, doi:10.1093/nar/gkt520.
[74] R. Heler, P. Samai, J.W. Modell, C. Weiner, G.W. Goldberg, D. Bikard, L.A. Marraffini, Cas9 specifies functional viral targets during CRISPR–Cas adaptation, Nature 519 (2015) 199–202, doi:10.1038/nature14245.
[75] M.H. Larson, L.A. Gilbert, X. Wang, W.A. Lim, J.S. Weissman, L.S. Qi, CRISPR interference (CRISPRi) for sequence-specific control of gene expression, Nat. Protoc. 8 (2013) 2180–2196, doi:10.1038/nprot.2013.132.
[76] T. Wang, C. Guan, J. Guo, B. Liu, Y. Wu, Z. Xie, C. Zhang, X.H. Xing, Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance-net, Nat. Commun. 9 (2018), doi:10.1038/s41467-018-04899-x.
[77] L.S. Qi, M.H. Larson, L.A. Gilbert, J.A. Doudna, J.S. Weissman, A.P. Arkin, W.A. Lim, Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression, Cell 152 (2013) 1173–1183, doi:10.1016/j.cell.2013.02.022.
[78] J.M. Replogle, R.A. Saunders, A.N. Pogson, J.A. Hussmann, A. Lenail, A. Guna, L. Mascibroda, E.J. Wagner, K. Adelman, G. Lithwick-Yanai, N. Iremadze, F. Oberstrass, D. Lipson, J.L. Bonnar, M. Jost, T.M. Norman, J.S. Weissman, Mapping information-rich genotype-phenotype landscapes with genome-scale Perturb-seq, Cell (2022) 1–17, doi:10.1016/j.cell.2022.05.013.
[79] L. Fang, J. Fan, S. Luo, Y. Chen, C. Wang, Y. Cao, H. Song, Genome-scale target identification in Escherichia coli for high-titer production of free fatty acids, Nat. Commun. (2021), doi:10.1038/s41467-021-25243-w.
[80] A.C. Komor, Y.B. Kim, M.S. Packer, J.A. Zuris, D.R. Liu, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage, Nature 533 (2016) 420–424, doi:10.1038/nature17946.
[81] N.M. Gaudelli, A.C. Komor, H.A. Rees, M.S. Packer, A.H. Badran, D.I. Bryson, D.R. Liu, Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage, Nature 551 (2017) 464–471, doi:10.1038/nature24644.
[82] S. Banno, K. Nishida, T. Arazoe, H. Mitsunobu, A. Kondo, Deaminase-mediated multiplex genome editing in Escherichia coli, Nat. Microbiol. 3 (2018), doi:10.1038/s41564-017-0102-6.
[83] A.V. Anzalone, P.B. Randolph, J.R. Davis, A.A. Sousa, L.W. Koblan, J.M. Levy, P.J. Chen, C. Wilson, G.A. Newby, A. Raguram, D.R. Liu, Search-and-replace genome editing without double-strand breaks or donor DNA, Nature 576 (2019) 149–157, doi:10.1038/s41586-019-1711-4.
[84] Y. Tong, T.S. Jørgensen, C.M. Whitford, T. Weber, S.Y. Lee, A versatile genetic engineering toolkit for E. coli based on CRISPR-prime editing, Nat. Commun. 12 (2021) 1–11, doi:10.1038/s41467-021-25541-3.
[85] E.V. Koonin, K.S. Makarova, E.V. Koonin, Origins and evolution of CRISPR-Cas systems, Philos. Trans. R. Soc. B. (2019).
[86] K. Chylinski, K.S. Makarova, E. Charpentier, E.V. Koonin, Classification and evolution of type II CRISPR-Cas systems, Nucleic Acids Res. 42 (2014) 6091–6105, doi:10.1093/nar/gku241.
[87] B. Zetsche, J.S. Gootenberg, O.O. Abudayyeh, I.M. Slaymaker, K.S. Makarova, P. Essletzbichler, S.E. Volz, J. Joung, J. Van Der Oost, A. Regev, E.V. Koonin, F. Zhang, Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system, Cell 163 (2015) 759–771, doi:10.1016/j.cell.2015.09.038.
[88] D.C. Swarts, M. Jinek, Mechanistic insights into the cis- and trans-acting DNase activities of Cas12a, Mol. Cell. 73 (2019) 589–600 e4, doi:10.1016/j.molcel.2018.11.021.
[89] T. Karvelis, G. Bigelyte, J.K. Young, Z. Hou, R. Zedaveinyte, K. Budre, S. Paulraj, V. Djukanovic, S. Gasior, A. Silanskas, C. Venclovas, V. Siksnys, PAM recognition by miniature CRISPR-Cas12f nucleases triggers programmable double-stranded DNA target cleavage, Nucleic Acids Res. 48 (2020) 5016–5023, doi:10.1093/nar/gkaa208.
[90] W.X. Yan, P. Hunnewell, L.E. Alfonse, J.M. Carte, E. Keston-Smith, S. Sothiselvam, A.J. Garrity, S. Chong, K.S. Makarova, E.V. Koonin, D.R. Cheng, D.A. Scott, Functionally diverse type V CRISPR-Cas systems, Science (80-) 363 (2019) 88–91, doi:10.1126/science.aav7271.
[91] W.X. Yan, S. Chong, H. Zhang, K.S. Makarova, E.V. Koonin, D.R. Cheng, D.A. Scott, Cas13d is a compact RNA-targeting type VI CRISPR effector positively modulated by a WYL-domain-containing accessory protein, Mol. Cell. 70 (2018) 327–339 e5, doi:10.1016/j.molcel.2018.02.028.
[92] S. Kannan, H. Altae-Tran, X. Jin, V.J. Madigan, R. Oshiro, K.S. Makarova, E.V. Koonin, F. Zhang, Compact RNA editors with small Cas13 proteins, Nat. Biotechnol. 40 (2022) 194–197, doi:10.1038/s41587-021-01030-2.
[93] O.O. Abudayyeh, J.S. Gootenberg, P. Essletzbichler, S. Han, J. Joung, J.J. Belanto, V. Verdine, D.B.T. Cox, M.J. Kellner, A. Regev, E.S. Lander, D.F. Voytas, A.Y. Ting, F. Zhang, RNA targeting with CRISPR-Cas13, Nature 550 (2017) 280–284, doi:10.1038/nature24049.
[94] O.O. Abudayyeh, J.S. Gootenberg, S. Konermann, J. Joung, I.M. Slaymaker, D.B.T. Cox, S. Shmakov, K.S. Makarova, E. Semenova, L. Minakhin, K. Severinov, A. Regev, E.S. Lander, E.V. Koonin, F. Zhang, C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector, Science (80-) 353 (2016), doi:10.1126/science.aaf5573.
[95] K. Okano, Y. Sato, T. Hizume, K. Honda, Genome editing by miniature CRISPR/Cas12f1 enzyme in Escherichia coli, J. Biosci. Bioeng. 132 (2021) 120–124, doi:10.1016/j.jbiosc.2021.04.009.
[96] X. Ao, Y. Yao, T. Li, T.T. Yang, X. Dong, Z.T. Zheng, G.Q. Chen, Q. Wu, Y. Guo, A multiplex genome editing method for Escherichia coli based on CRISPR-Cas12a, Front. Microbiol. 9 (2018) 1–13, doi:10.3389/fmicb.2018.02307.
[97] D.B.T. Cox, J.S. Gootenberg, O.O. Abudayyeh, B. Franklin, M.J. Kellner, J. Joung, F. Zhang, RNA editing with CRISPR-Cas13, Science (80-) 358 (2017) 1019–1027, doi:10.1126/science.aaq0180.
[98] B. Csörgő, L.M. León, I.J. Chau-Ly, A. Vasquez-Rifo, J.D. Berry, C. Mahendra, E.D. Crawford, J.D. Lewis, J. Bondy-Denomy, A compact Cascade–Cas3 system for targeted genome engineering, Nat. Methods. 17 (2020) 1183–1190, doi:10.1038/s41592-020-00980-w.
[99] J.E. Peters, K.S. Makarova, S. Shmakov, E.V. Koonin, Recruitment of CRISPR-Cas systems by Tn7-like transposons, Proc. Natl. Acad. Sci. U. S. A. 114 (2017) E7358–E7366, doi:10.1073/pnas.1709035114.
[100] S.E. Klompe, P.L.H. Vo, T.S. Halpin-Healy, S.H. Sternberg, Transposon-encoded CRISPR–Cas systems direct RNA-guided DNA integration, Nature 571 (2019) 219–225, doi:10.1038/s41586-019-1323-z.
[101] P.L.H. Vo, C. Ronda, S.E. Klompe, E.E. Chen, C. Acree, H.H. Wang, S.H. Sternberg, CRISPR RNA-guided integrases for high-efficiency, multiplexed bacterial genome engineering, Nat. Biotechnol. (2020), doi:10.1038/s41587-020-00745-y.
[102] A. Choudhury, J.A. Fenster, R.G. Fankhauser, J.L. Kaar, O. Tenaillon, R.T. Gill, CRISPR /Cas9 recombineering-mediated deep mutational scanning of essential genes in Escherichia coli, Mol. Syst. Biol. 16 (2020) 1–22, doi:10.15252/msb.20199265.
[103] W.E. Robertson, L.F.H. Funke, D. de la Torre, J. Fredens, K. Wang, J.W. Chin, Creating custom synthetic genomes in Escherichia coli with REXER and GENESIS, Nat. Protoc. 16 (2021) 2345–2380, doi:10.1038/s41596-020-00464-3.
[104] J. Cello, A.V. Paul, E. Wimmer, Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template, Science (80-) 297 (2002) 1016–1018, doi:10.1126/science.1072266.
[105] D.G. Gibson, G.A. Benders, C. Andrews-Pfannkoch, E.A. Denisova, H. Baden-Tillson, J. Zaveri, T.B. Stockwell, A. Brownley, D.W. Thomas, M.A. Algire, C. Merryman, L. Young, V.N. Noskov, J.I. Glass, J.C. Venter, C.A. Hutchison, H.O. Smith, Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome, Science (80-) 319 (2008) 1215–1220, doi:10.1126/science.1151721.
[106] D.G. Gibson, J.I. Glass, C. Lartigue, V.N. Noskov, R.Y. Chuang, M.A. Algire, G.A. Benders, M.G. Montague, L. Ma, M.M. Moodie, C. Merryman, S. Vashee, R. Krishnakumar, N. Assad-Garcia, C. Andrews-Pfannkoch, E.A. Denisova, L. Young, Z.N. Qi, T.H. Segall-Shapiro, C.H. Calvey, P.P. Parmar, C.A. Hutchison, H.O. Smith,

J.C. Venter, Creation of a bacterial cell controlled by a chemically synthesized genome, Science (80-). 329 (2010) 52–56, doi:10.1126/science.1190719.
[107] C.A. Hutchison, R.Y. Chuang, V.N. Noskov, N. Assad-Garcia, T.J. Deerinck, M.H. Ellisman, J. Gill, K. Kannan, B.J. Karas, L. Ma, J.F. Pelletier, Z.Q. Qi, R.A. Richter, E.A. Strychalski, L. Sun, Y. Suzuki, B. Tsvetanova, K.S. Wise, H.O. Smith, J.I. Glass, C. Merryman, D.G. Gibson, J.C. Venter, Design and synthesis of a minimal bacterial genome, Science (80-) 351 (2016), doi:10.1126/science.aad6253.
[108] N. Annaluru, H. Muller, L.A. Mitchell, S. Ramalingam, G. Stracquadanio, S.M. Richardson, J.S. Dymond, Z. Kuang, L.Z. Scheifele, E.M. Cooper, Y. Cai, K. Zeller, N. Agmon, J.S. Han, M. Hadjithomas, J. Tullman, K. Caravelli, K. Cirelli, Z. Guo, V. London, A. Yeluru, S. Murugan, K. Kandavelou, N. Agier, G. Fischer, K. Yang, J.A. Martin, M. Bilgel, P. Bohutski, K.M. Boulier, B.J. Capaldo, J. Chang, K. Charoen, W.J. Choi, P. Deng, J.E. DiCarlo, J. Doong, J. Dunn, J.I. Feinberg, C. Fernandez, C.E. Floria, D. Gladowski, P. Hadidi, I. Ishizuka, J. Jabbari, C.Y.L. Lau, P.A. Lee, S. Li, D. Lin, M.E. Linder, J. Ling, J. Liu, J. Liu, M. London, M. Henry, J. Mao, J.E. McDade, A. McMillan, A.M. Moore, W.C. Oh, Y. Ouyang, R. Patel, M. Paul, L.C. Paulsen, J. Qiu, A. Rhee, M.G. Rubashkin, I.Y. Soh, N.E. Sotuyo, V. Srinivas, A. Suarez, A. Wong, R. Wong, W.R. Xie, Y. Xu, A.T. Yu, R. Koszul, J.S. Bader, J.D. Boeke, S. Chandrasegaran, Total synthesis of a functional designer eukaryotic chromosome, Science (80-) 344 (2014) 55–58, doi:10.1126/science.1249252.
[109] W. Zhang, G. Zhao, Z. Luo, Y. Lin, L. Wang, Y. Guo, A. Wang, S. Jiang, Q. Jiang, J. Gong, Y. Wang, S. Hou, J. Huang, T. Li, Y. Qin, J. Dong, Q. Qin, J. Zhang, X. Zou, X. He, L. Zhao, Y. Xiao, M. Xu, E. Cheng, N. Huang, T. Zhou, Y. Shen, R. Walker, Y. Luo, Z. Kuang, L.A. Mitchell, K. Yang, S.M. Richardson, Y. Wu, B.-Z. Li, Y.-J. Yuan, H. Yang, J. Lin, G.-Q. Chen, Q. Wu, J.S. Bader, Y. Cai, J.D. Boeke, J. Dai, Engineering the ribosomal DNA in a megabase synthetic chromosome, Science (80-) 355 (2017) eaaf3981, doi:10.1126/science.aaf3981.
[110] Y. Shen, Y. Wang, T. Chen, F. Gao, J. Gong, D. Abramczyk, R. Walker, H. Zhao, S. Chen, W. Liu, Y. Luo, C.A. Müller, A. Paul-Dubois-Taine, B. Alver, G. Stracquadanio, L.A. Mitchell, Z. Luo, Y. Fan, B. Zhou, B. Wen, F. Tan, Y. Wang, J. Zi, Z. Xie, B. Li, K. Yang, S.M. Richardson, H. Jiang, C.E. French, C.A. Nieduszynski, R. Koszul, A.L. Marston, Y. Yuan, J. Wang, J.S. Bader, J. Dai, J.D. Boeke, X. Xu, Y. Cai, H. Yang, Deep functional analysis of synII, a 770-kilobase synthetic yeast chromosome, Science (80-) (2017) 355, doi:10.1126/science.aaf4791.
[111] L.A. Mitchell, A. Wang, G. Stracquadanio, Z. Kuang, X. Wang, K. Yang, S. Richardson, J.A. Martin, Y. Zhao, R. Walker, Y. Luo, H. Dai, K. Dong, Z. Tang, Y. Yang, Y. Cai, A. Heguy, B. Ueberheide, D. Fenyö, J. Dai, J.S. Bader, J.D. Boeke, Synthesis, debugging, and effects of synthetic chromosome consolidation: synVI and beyond, Science (80-) 355 (2017), doi:10.1126/science.aaf4831.
[112] J. Dai, J.D. Boeke, Z. Luo, S. Jiang, Y. Cai, Sc3.0: revamping and minimizing the yeast genome, Genome Biol. 21 (2020) 1–4, doi:10.1186/s13059-020-02130-z.
[113] N. Ostrov, M. Landon, M. Guell, G. Kuznetsov, J. Teramoto, N. Cervantes, M. Zhou, K. Singh, M.G. Napolitano, M. Moosburner, E. Shrock, B.W. Pruitt, N. Conway, D.B. Goodman, C.L. Gardner, G. Tyree, A. Gonzales, B.L. Wanner, J.E. Norville, M.J. Lajoie, G.M. Church, Design, synthesis, and testing toward a 57-codon genome, Science (80-) 353 (2016) 819 LP –822, doi:10.1126/science.aaf3639.
[114] J. Fredens, K. Wang, D. de la Torre, L.F.H. Funke, W.E. Robertson, Y. Christova, T. Chia, W.H. Schmied, D.L. Dunkelmann, V. Beránek, C. Uttamapinant, A.G. Llamazares, T.S. Elliott, J.W. Chin, Total synthesis of Escherichia coli with a recoded genome, Nature 569 (2019) 514–518, doi:10.1038/s41586-019-1192-5.
[115] C.G. Acevedo-Rocha, G. Fang, M. Schmidt, D.W. Ussery, A. Danchin, From essential to persistent genes: a functional approach to constructing synthetic life, Trends Genet 29 (2013) 273–279, doi:10.1016/j.tig.2012.11.001.
[116] Y.-C. Hwang, C.-C. Lin, J.-Y. Chang, H. Mori, H.-F. Juan, H.-C. Huang, Predicting essential genes based on network and sequence analysis, Mol. Biosyst. 5 (2009) 1672–1678, doi:10.1039/B900611G.
[117] L.A. Gallagher, J. Bailey, C. Manoil, Ranking essential bacterial processes by speed of mutant death, Proc. Natl. Acad. Sci. U. S. A. 117 (2020) 18010–18017, doi:10.1073/pnas.2001507117.
[118] J.M. Peters, A. Colavin, H. Shi, T.L. Czarny, M.H. Larson, S. Wong, J.S. Hawkins, C.H.S. Lu, B.M. Koo, E. Marta, A.L. Shiver, E.H. Whitehead, J.S. Weissman, E.D. Brown, L.S. Qi, K.C. Huang, C.A. Gross, A comprehensive, CRISPR-based functional analysis of essential genes in bacteria, Cell 165 (2016) 1493–1506, doi:10.1016/j.cell.2016.05.003.
[119] G. Giaever, A.M. Chu, L. Ni, C. Connelly, L. Riles, S. Véronneau, S. Dow, A. Lucau-Danila, K. Anderson, B. André, A.P. Arkin, A. Astromoff, M. El Bakkoury, R. Bangham, R. Benito, S. Brachat, S. Campanaro, M. Curtiss, K. Davis, A. Deutschbauer, K.-D. Entian, P. Flaherty, F. Foury, D.J. Garfinkel, M. Gerstein, D. Gotte, U. Güldener, J.H. Hegemann, S. Hempel, Z. Herman, D.F. Jaramillo, D.E. Kelly, S.L. Kelly, P. Kötter, D. LaBonte, D.C. Lamb, N. Lan, H. Liang, H. Liao, L. Liu, C. Luo, M. Lussier, R. Mao, P. Menard, S.L. Ooi, J.L. Revuelta, C.J. Roberts, M. Rose, P. Ross-Macdonald, B. Scherens, G. Schimmack, B. Shafer, D.D. Shoemaker, S. Sookhai-Mahadeo, R.K. Storms, J.N. Strathern, G. Valle, M. Voet, G. Volckaert, C. Wang, T.R. Ward, J. Wilhelmy, E.A. Winzeler, Y. Yang, G. Yen, E. Youngman, K. Yu, H. Bussey, J.D. Boeke, M. Snyder, P. Philippsen, R.W. Davis, M. Johnston, Functional profiling of the Saccharomyces cerevisiae genome, Nature 418 (2002) 387–391, doi:10.1038/nature00935.
[120] D.S. León, J. Nogales, Toward merging bottom – up and top – down model-based designing of synthetic microbial communities, Curr. Opin. Microbiol. 69 (2022) 102169, doi:10.1016/j.mib.2022.102169.
[121] R. Maddamsetti, P.J. Hatcher, A.G. Green, B.L. Williams, D.S. Marks, R.E. Lenski, Core genes evolve rapidly in the long-term evolution experiment with Escherichia coli, Genome Biol. Evol. 9 (2017) 1072–1083, doi:10.1093/gbe/evx064.
[122] O. Lukjancenko, T.M. Wassenaar, D.W. Ussery, Comparison of 61 sequenced Escherichia coli genomes, Microb. Ecol. 60 (2010) 708–720, doi:10.1007/s00248-010-9717-3.
[123] E. Martínez-Carranza, H. Barajas, L.D. Alcaraz, L. Servín-González, G.Y. Ponce-Soto, G. Soberón-Chávez, Variability of bacterial essential genes among closely related bacteria: the case of Escherichia coli, Front. Microbiol. 9 (2018) 1–7, doi:10.3389/fmicb.2018.01059.
[124] R.L. Charlebois, W.F. Doolittle, Computing prokaryotic gene ubiquity: rescuing the core from extinction, Genome Res. 14 (2004) 2469–2477, doi:10.1101/gr.3024704.
[125] C. Freiberg, B. Wieland, F. Spaltmann, K. Ehlert, H. Brötz, H. Labischinski, Identification of novel essential Escherichia coli genes conserved among pathogenic bacteria, J. Mol. Microbiol. Biotechnol. 3 (2001) 483–489.
[126] E.V. Koonin, Comparative genomics, minimal gene-sets and the last universal common ancestor, Nat. Rev. Microbiol. 1 (2003) 127–136, doi:10.1038/nrmicro751.
[127] T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K.A. Datsenko, M. Tomita, B.L. Wanner, H. Mori, Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection, Mol. Syst. Biol. 2 (2006), doi:10.1038/msb4100050.
[128] R. Zhang, Y. Lin, DEG 5.0, a database of essential genes in both prokaryotes and eukaryotes, Nucleic Acids Res. 37 (2009) 455–458, doi:10.1093/nar/gkn858.
[129] E.C.A. Goodall, A. Robinson, I.G. Johnston, S. Jabbari, K.A. Turner, A.F. Cunningham, P.A. Lund, J.A. Cole, I.R. Henderson, The essential genome of Escherichia coli K-12, MBio 9 (2017) 1–18, doi:10.1101/237842.
[130] F. Rousset, L. Cui, E. Siouve, C. Becavin, F. Depardieu, D. Bikard, Genome-wide CRISPR-dCas9 screens in E. coli identify essential genes and phage host factors, PLoS Genet 14 (2018) 1–28, doi:10.1371/journal.pgen.1007749.
[131] W. Tim, B. Kıvanç, H.N. W, K.K. M, P. Yorick, W.J. J, L.E. S, S.D. M, Identification and characterization of essential genes in the human genome, Science (80-) 350 (2015) 1096–1101, doi:10.1126/science.aac7041.
[132] K. Wang, J. Fredens, S.F. Brunner, S.H. Kim, T. Chia, J.W. Chin, Defining synonymous codon compression schemes by genome recoding, Nature 539 (2016) 59–64, doi:10.1038/nature20124.
[133] J.F. Zürcher, W.E. Robertson, T. Kappes, G. Petris, T.S. Elliott, G.P.C. Salmond, J.W. Chin, Refactored genetic codes enable bidirectional genetic isolation, Science (80-). 378 (2022) 516–523, doi:10.1126/science.add8943.
[134] T. Mukai, T. Yoneji, K. Yamada, H. Fujita, S. Nara, M. Su'etsugu, Overcoming the challenges of megabase-sized plasmid construction in Escherichia coli, ACS Synth. Biol. 9 (2020) 1315–1327, doi:10.1021/acssynbio.0c00008.
[135] K. Wang, D. De Torre, W.E. Robertson, J.W. Chin, Programmed chromosome fission and fusion enable precise large-scale genome rearrangement and assembly, Science (80-) 926 (2019) 922–926.
[136] T. Yoneji, H. Fujita, T. Mukai, M. Su'etsugu, M. Su'etsugu, Grand scale genome manipulation via chromosome swapping in Escherichia coli programmed by three one megabase chromosomes, Nucleic Acids Res. 49 (2021) 1–12, doi:10.1093/nar/gkab298.
[137] R.T. Dame, F.Z.M. Rashid, D.C. Grainger, Chromosome organization in bacteria: mechanistic insights into genome structure and function, Nat. Rev. Genet. 21 (2020) 227–242, doi:10.1038/s41576-019-0185-4.
[138] V.S. Lioy, A. Cournac, M. Marbouty, S. Duigou, J. Mozziconacci, O. Espéli, F. Boccard, R. Koszul, Multiscale structuring of the E. coli chromosome by nucleoid-associated and condensin proteins, Cell 172 (2018) 771–783 e18, doi:10.1016/j.cell.2017.12.027.
[139] R. Gil, F.J. Silva, J. Peretó, A. Moya, Determination of the core of a minimal bacterial gene set, Microbiol. Mol. Biol. Rev. 68 (2004) 518–537, doi:10.1128/mmbr.68.3.518-537.2004.
[140] L. Wang, C.D. Maranas, MinGenome: an in silico top-down approach for the synthesis of minimized genomes, ACS Synth. Biol. 7 (2018) 462–473, doi:10.1021/acssynbio.7b00296.
[141] V. Kolisnychenko, G. Plunkett, C.D. Herring, T. Fehér, J. Pósfai, F.R. Blattner, G. Pósfai, Engineering a reduced Escherichia coli genome, Genome Res 12 (2002) 640–647, doi:10.1101/gr.217202.
[142] Y. Hirokawa, H. Kawano, K. Tanaka-Masuda, N. Nakamura, A. Nakagawa, M. Ito, H. Mori, T. Oshima, N. Ogasawara, Genetic manipulations restored the growth fitness of reduced-genome Escherichia coli, J. Biosci. Bioeng. 116 (2013) 52–58, doi:10.1016/j.jbiosc.2013.01.010.
[143] K. Standage-Beier, Q. Zhang, X. Wang, Targeted large-scale deletion of bacterial genomes using CRISPR-nickases, ACS Synth. Biol. 4 (2015) 1217–1225, doi:10.1021/acssynbio.5b00132.
[144] Z. Zhang, A.E. Baxter, D. Ren, K. Qin, Z. Chen, S.M. Collins, H. Huang, C.A. Komar, P.F. Bailer, J.B. Parker, G.A. Blobel, R.M. Kohli, E.J. Wherry, S.L. Berger, J. Shi, Efficient engineering of human and mouse primary cells using peptide-assisted genome editing, (2023). https://doi.org/10.1038/s41587-023-01756-1.