Synthetic and Systems Biotechnology 14 (2026) 388–398

Contents lists available at ScienceDirect

Synthetic and Systems Biotechnology

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

Review Article

Transposon-based genome editing of industrial microorganisms: advances, challenges, and prospects

Zhiyi Zhang [a,1], Zheng Huang [a,1], Yilin Wang [a], Zheng Li [a], Zhiqiang Wen [b], Yang Gu [a,*]

[a] School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai, 200093, China
[b] Nanjing Gasgene Biochemical Technology Co. Ltd., Nanjing, 210032, China

ARTICLE INFO

Keywords:
Transposon
Industrial microorganisms
Transposon sequencing
Random integration
Functional genomics

ABSTRACT

As mobile genetic elements, transposons play a crucial role in the adaptive evolution and genome engineering of industrial microorganisms. Their applications range from high-throughput functional genomics, enabling systematic genotype-phenotype mapping via transposon sequencing, to the construction of random integration libraries for chassis development and directed evolution. Despite the emergence of precise editing tools such as CRISPR-Cas, transposon technology remains indispensable in non-model industrial strains owing to its operational simplicity and high efficiency. Recent discoveries of novel transposon systems, along with their functional enhancement, have further expanded their utility. Looking ahead, integrating transposon technology with strategies such as AI-assisted design and CRISPR-Cas-based systems will greatly advance our ability to decipher and engineer industrial microbial cell factories. This review summarizes the principles, challenges, and opportunities of transposon-associated technologies in industrial microorganisms, offering new insights into their roles in industrial biotechnology.

1. Introduction

With the rapid advancement of green biomanufacturing, industrial microorganisms have emerged as a core driving force in sustainable production. Model organisms such as *Escherichia coli*, *Bacillus subtilis*, and *Saccharomyces cerevisiae* have been successfully developed into efficient cell factories, capable of producing a wide range of value-added products [1–3]. In parallel, several non-model industrial microorganisms have demonstrated distinct and advantageous traits, underscoring their potential for practical applications. Notable examples include gas-fermenting *Clostridium* species, which can utilize one-carbon (C1) gases such as CO_2 and CO [4], as well as *Pichia pastoris* and *Vibrio natriegens* for their ability to assimilate organic C1 compounds like methanol and formic acid [5,6]. To fully harness the potential of these microorganisms and scale up their industrial use, comprehensive genetic modification and optimization are essential. Consequently, the continuous expansion and refinement of molecular genetic tools tailored to these strains remains a central research focus, attracting widespread scientific interest in the field.

Traditional homologous recombination techniques are well-established in model microorganisms like *E. coli* and *S. cerevisiae* [7–9], yet their application in most industrial strains is hindered by low efficiency, lengthy procedures, and operational complexity [10,11]. To address these limitations, recombinase-mediated site-specific systems such as Cre-loxP and FLP/FRT were subsequently developed, enabling precise DNA excision or recombination [12,13]. However, their utility remains constrained by the prerequisite of pre-engineered genomic recognition sites in the host organism. In recent years, the CRISPR-Cas system has revolutionized precise genome editing. Its targeting specificity and versatility provide unprecedented convenience for metabolic engineering. However, challenges such as delivery difficulties, cytotoxicity, and unstable editing efficiency persist in many non-model microorganisms [14]. Moreover, CRISPR-Cas-based approaches remain costly for large-scale functional genomics screening or genome-wide unbiased library construction (e.g., crRNA library construction for CRISPRi screening) [15]. These limitations underscore the ongoing need in industrial microbial engineering to develop diverse, broadly applicable tools that can be adapted to the many non-model microorganisms with industrial potential but limited genetic toolkits. Consequently, transposon technology, with its distinct mechanism independent of

Peer review under the responsibility of Editorial Board of Synthetic and Systems Biotechnology.
* Corresponding author.
E-mail address: ygu@usst.edu.cn (Y. Gu).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.synbio.2026.04.026
Received 27 January 2026; Received in revised form 7 March 2026; Accepted 26 April 2026

homologous recombination, is attracting renewed interest.

The core of transposon technology lies in the activity of transposase, which recognizes specific terminal inverted repeat sequences and catalyzes DNA fragment movement independently of the host's homologous recombination system [16]. Key advantages of this technology include: (1) broad versatility, enabling genetic manipulation across diverse microorganisms, including those resistant to homologous recombination [17]; (2) efficient random insertion, suitable for constructing genome-wide mutagenesis libraries to identify essential genes and screen functional genes associated with specific phenotypes [18]; (3) stable genomic integration, serving as an efficient one-step delivery tool for large DNA fragments [19]; and (4) programmable targeting when fused with CRISPR systems, as seen in CRISPR-associated transposases (CAST), which combine precise insertion with exploratory capability [20]. Furthermore, recent integration of transposon technology with high-throughput sequencing and microfluidic platforms is extending the utility of transposon systems in functional genomics, synthetic biology, and strain development [21].

Table 1
Applications of transposons in industrial microorganisms.

Transposons	Microorganisms	Transposition efficiency	Applications	Randomness/ targetability	Cargo sizes	References
Mini-Tn5 transposon carrying a promoter-less GFP cassette	*Escherichia coli*	1600–1900 CFU/mL	Random mutagenesis libraries	Random	3.2 kb	[43]
Tn5 with strong *rrnB* T1/T2 terminators and an aTc-inducible promoter	*Escherichia coli*	2.83×10^7 CFU/μg DNA	Genome-wide gene activation mutant libraries	Random	2.3 kb	[47]
Mini-Tn5/7-Lux	*Escherichia coli*	—	Identifying and capturing promoters	Tn5: Random Tn7: Targeted	—	[99]
Tn5	*Lactobacillus casei*	—	Random mutagenesis libraries, identifying *asnH*, an asparagine synthetase gene involved in the immune-activating capacity	Random	—	[37]
Mariner	*Bacillus subtilis*	5%; 1.26×10^{-5} per recipient cell	Random knock-in expression system for high yield production of heterologous protein	Random	—	[39,40]
		—	Random mutagenesis libraries, revealing a previously unknown link between the DnaJK chaperone system and RNase P-dependent mRNA metabolism	Random	—	[41]
Xylose-inducible *mariner*	*Clostridium acetobutylicum*	91.9%	Random mutagenesis libraries, isolating a sporulation mutant *spo0A*	Random	—	[42]
IS*Cg1*	*Corynebacterium glutamicum*	—	Random mutagenesis libraries	Random	—	[48]
Mini-Tn5	*Corynebacterium glutamicum*	3.0×10^4 CFU/μg DNA	Random mutagenesis libraries	Random	1.2 kb	[100]
IS*6100*	*Corynebacterium glutamicum*	5–10 CFU/μg DNA	Random mutagenesis libraries, combining with metabolic profiling method to analyze the genotype/phenotype correlations	Random	—	[49,50]
Tn5	*Streptomyces avermitilis*	—	Random mutagenesis libraries, identifying a previously unidentified sugar transporter TP6568 for producing high-yield avermectin B_{1a}	Random	—	[34]
	Streptomyces rimosus	—	Random mutagenesis libraries, identifying two previously uncharacterized regulators of rimocidin biosynthesis: RRT and HYP	Random	—	[35]
Mini-Tn3/7-3xHA/lacZ	*Saccharomyces cerevisiae*	—	Random mutagenesis libraries	Random	—	[44–46]
Tn5, *mariner*	*Escherichia coli*	—	Large-scale genetic screening and characterization	Random	—	[58–61]
Mariner	*Clostridium ljungdahlii*	98%	Chromosomal integration of 3-hydroxybutyrate production pathway	Random	4.8 kb	[69]
Xylose-inducible *mariner*	*Clostridium ljungdahlii*	—	Integrating 11.1 kb clusters into the genome and producing acetone and isopropanol; Integrating 17.9 kb and 7.8 kb clusters respectively into the genome and producing hexanol and butanol	Random	Up to 25.6 kb	[70,71]
Hermes	*Yarrowia lipolytica*	—	Integrating multiple copies of the 2-pyrone synthase (2-PS)-expressing cassette into the genome	Random	—	[72]
Ty (LTR sequences)	*Yarrowia lipolytica*	—	Developing multi-copy and high-copy integration tools (YaliCMulti and YaliHMulti)	Targeted	≤20 kb	[73]
Mini-Tn5	*Escherichia coli*	5.5×10^{-3} per recipient cell	Integrating multicopy of polyhydroxybutyrate (PHB) biosynthesis genes and producing more PHB than before	Random	5.7 kb	[76]
Tn7-like	*Escherichia coli*	~97%	Integrating DNA into target sites; Inhibiting byproducts and enhancing recombinant protein yield (ShCAST)	Targeted	At least 14.5 kb	[81]
Tn6677	*Escherichia coli*	≤62%	Integrating DNA sequence on the target site (INTEGRATE)	Targeted	At least 10 kb	[82]
Tn6677	*Lactococcus lactis*	4×10^{-5} per recipient cell	Integrating a 10 kb fragment at the target site with an efficiency only fivefold lower than that for a 1 kb fragment (INTEGRATE)	Targeted	10 kb	[84]
Tn6677	*Corynebacterium glutamicum*	≤55.5%	Integrating DNA fragments at the target site (INTEGRATE)	Targeted	Up to 9 kb	[85]
Tn6677	*Escherichia coli*	≤95%	Implementing multicopy chromosomal integration (MUCICAT), producing 11.59 g/L N-acetylglucosamine (GlcNAc), and inserting the *ppk* gene (No. 8) into various positions and copy numbers in the genome	Targeted	—	[87,88]

Notes: the dash indicates that the relevant information is not mentioned in this reference.

This review systematically summarizes recent research advancements and innovative applications of transposon technology in industrial microorganisms (Table 1). Key molecular mechanisms and engineering strategies of various transposon systems are discussed, supported by representative case studies illustrating their use in constructing saturated mutagenesis libraries, enabling gene knock-in/knockout, facilitating high-throughput functional screening, and driving genome rearrangements. Current technical limitations and emerging solutions are critically evaluated. Furthermore, we provide future directions, with emphasis on integrating transposon systems with automation and artificial intelligence, to inspire the development of more robust and intelligent genetic manipulation platforms for next-generation industrial biotechnology.

2. Transposon classification and mechanisms in industrial microbes

Transposons are primarily classified by their mobilization mechanism. Class I retrotransposons employ a "copy-and-paste" strategy, utilizing an RNA intermediate for genomic integration and duplication [22]. In contrast, Class II DNA transposons typically move via a "cut-and-paste" mechanism, where encoded transposases directly excise and reinsert DNA segments [16,23]. More recently, a complex "copy--out–paste-in" pathway has been characterized, involving the excision and reinsertion of a double-stranded circular DNA intermediate [16]. While these core mechanisms are established, more researches are focused on its application in gene regulation [24], strain improvement [25], and genome reintegration and evolution [26,27]. Thus, the field is moving beyond mechanistic descriptions toward a more integrated understanding of transposon functions and their biological impact.

To date, transposons have been found in nearly all sequenced

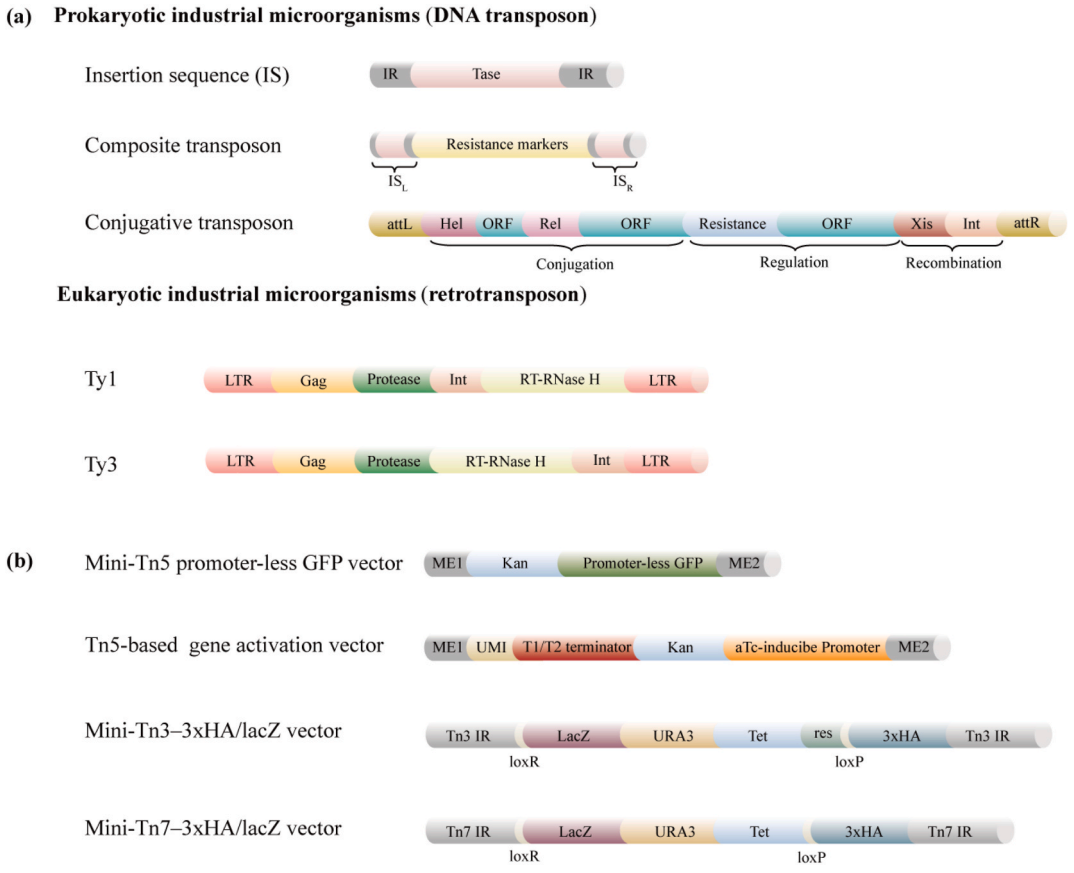

Fig. 1. The schematic diagrams of transposons. (a) The structure of transposons in prokaryotic and eukaryotic industrial microorganisms. The insertion sequence (IS) represents the simplest form, consisting of terminal inverted repeats (IR) of 4 to 10 base pairs at both ends and a single transposase gene (Tase) located centrally. Composite transposon contains an IS element flanking a gene that confers resistance to antimicrobial agents. Conjugative transposon has multiple modules. The conjugation module includes the genes encoding helicase (Hel) and relaxase (Rel) as well as other essential open reading frames (ORFs). The recombination module encompasses the genes encoding integrase (Int) and excisionase (Xis) as well as the specific attachment site (att). Eukaryotic industrial microorganisms typically harbor two types of DNA transposons: Ty1 and Ty3. The associated functional components include: long terminal repeats (LTRs), which regulate transcriptional initiation and termination; *gag*, the gene encoding the capsid protein essential for the assembly of viral-like particles; protease, responsible for the proteolytic cleavage of precursor polyproteins into mature functional subunits; reverse transcriptase in conjunction with RNase H (RT-RNase H), which mediates the synthesis of complementary DNA (cDNA) from an RNA template; *int*, the gene encoding integrase that catalyzes the integration of cDNA into the host genome. **(b)** The core structure of modified transposons. The mini-Tn5 promoter-less GFP vector contains a promoter-less GFP-kanamycin (Kan) cassette. When the transposon inserts in the correct orientation, expression of GFP and the kanamycin resistance gene is driven by the native promoter of the inserted gene, allowing real-time monitoring of gene expression. The specific inverted mosaic ends (ME1/2) are the recognition site of Tn5 transposon. The Tn5-based gene activation vector is composed of ME1/2, strong *rrnB* T1 and T2 terminators (preventing native transcriptional interference from upstream), a 6-nucleotide unique molecular identifier (UMI) (facilitating accurate quantification of the enrichment of activated genes) and an anhydrotetracycline (aTc)-inducible promoter (allowing for conditional activation of downstream genes). The mini-Tn3/7–3xHA/lacZ vector contains a lox site near each Tn3/7 end (IR), three copies of a hemagglutinin (3xHA) epitope tag, a reporter gene (*lacZ*) lacking an initiator methionine and upstream promoter sequence, and a tetracycline-resistance gene (*tet*). The recombination site res is the Tn3 recombination region.

industrial microorganisms [25]. Prokaryotic industrial microorganisms predominantly harbor DNA transposons, which exist either as simple insertion sequences (IS) or within more complex composite and conjugative transposons, such as Tn5, Tn10, Tn916 and Tn1545. In contrast, eukaryotic industrial microorganisms typically possess retrotransposons, including ty1 and ty3 (Fig. 1A). For industrial application, non-replicative transposons serve as molecular tools in functional genomics, particularly for constructing random mutant libraries, as they usually insert as single copies into the genome. This enables efficient identification of functional genes associated with key phenotypic traits. In contrast, replicative transposons amplify their copy number within the chromosome during transposition. This property allows them to be employed for multicopy integration of specific genes into the bacterial chromosome, making them suitable for investigating gene dosage effects on expression and host performance. However, the presence of multiple replicative transposon copies in the genome may also trigger genomic rearrangements, which may impact the genetic stability of industrial microbial strains.

As dynamic genomic elements, transposons exert a dual influence on industrial microorganisms, posing challenges such as genetic instability while also harboring underexplored biotechnological potential. Their influence operates through three primary mechanisms: (1) horizontal gene transfer, (2) gene expression regulation via insertion mutagenesis, and (3) genome restructuring. First, transposons act as key vectors for horizontal gene transfer within industrial microbial communities. Although this raises public health concerns, such as the spread of antibiotic resistance genes in fermentation settings [28], it also carries significant biotechnological promise. For example, transposons may facilitate the dissemination of gene clusters involved in degrading complex substrates or tolerating harsh industrial conditions, offering valuable insights into microbial adaptive evolution. Second, through insertion mutagenesis, transposons can directly modulate gene expression [24]. While this is often regarded as a source of genetic instability, this mechanism can be harnessed from a synthetic biology perspective. By engineering transposons as controllable random mutagenesis tools, they enable high-throughput screening for industrial strains with improved yields or novel traits. Third, transposon-mediated chromosomal rearrangements represent a core driver of genomic plasticity [27]. While large-scale insertions, deletions, or inversions can disrupt essential metabolic pathways, they may also create novel gene combinations or regulatory networks, equipping microorganisms with new capacities to cope with industrial stresses. Therefore, transposons play a constructive role in driving adaptive evolution and functional innovation in industrial microorganisms.

To fully explore the application potential of transposons, researchers have undertaken numerous efforts in the engineering of transposases and the optimization of the host microenvironment. As the core component of transposon systems, transposase activity is a key determinant of transposition efficiency. Enhancing transposase expression and obtaining highly active enzyme mutants through directed evolution or rational design represent primary approaches for optimizing transposase performance [29]. In addition, modulating the host microenvironment has proven to be an effective strategy for improving transposition efficiency. For example, overexpression of the DNA-PK complex subunits (Ku70, Ku80, and DNA-PKcs) in mammalian cells significantly enhanced piggyBac transposon activity [30]. Given the high evolutionary conservation of the DNA-PK complex in eukaryotes, this strategy may also hold promise for the genetic engineering of industrial microorganisms, particularly eukaryotic microbes such as yeasts. This is an area that warrants further investigation.

3. Applications of transposons in industrial microorganisms

3.1. Rapid discovery of functional genes linked to specific phenotypes

3.1.1. Transposon-based random genome mutagenesis

In industrial microorganisms, essential genes act as critical hubs for enhancing metabolic efficiency, improving environmental adaptation, and ensuring biosafety. The primary criterion for identifying such genes is that their disruption leads to microbial lethality. Consequently, systematic identification and functional analysis of essential genes constitute a fundamental step toward a deeper understanding and optimization of industrial strains. Currently, transposon-based genome-wide random mutagenesis combined with high-throughput sequencing technologies, such as Tn-seq, HITS, INSeq, and TraDIS [18], has served as a powerful tool for large-scale screening of functional gene in microorganisms [31]. These methods not only offer extensive genome coverage, but also enable systematic identification of essential genes in a single experiment, significantly accelerating functional genomics research. Moreover, these tools unlock a broader biological application, including the identification of conditionally essential genes, the discovery of key host adaptation factors, the detection of regulatory elements such as functional small RNAs, the tracking of adaptive evolutionary events, and the generation of comprehensive datasets for constructing global genetic interaction networks [32]. These expanded applications make them indispensable not only for fundamental microbiological studies, but also for applied biotechnological research.

The Tn5 and *mariner* transposon are the primary tools for constructing genome-scale mutant libraries in industrial prokaryotes due to their efficiency, randomness, and broad host applicability [33]. These systems enable both gene discovery and functional genomics across diverse bacteria. In *Streptomyces avermitilis*, Tn5 mutagenesis identified an ABC transporter and its regulator; their co-overexpression increased avermectin B_{1a} titer by over 50% [34]. The same approach applied to *Streptomyces rimosus* led to the identification of two previously uncharacterized regulators of rimocidin biosynthesis: a positive regulator (RRT) and a negative regulator (HYP) [35]. Similarly, a large Tn5 library in *Lactobacillus casei* mapped genetic bases of key industrial traits such as nutrient biosynthesis and phage sensitivity [36,37].

In contrast to Tn5, the *mariner* transposon employs a distinct, low-specificity targeting mechanism that primarily inserts at "TA" dinucleotides [38]. In *Bacillus subtilis*, it has been deployed to systematically scout for genomic integration sites that surpass the canonical *amyE* locus for heterologous protein expression [39]. A follow-up optimization confirmed the superiority of the *yjbH* site, achieving ~50% higher expression than *amyE* [40]. The mariner-based transposon was subsequently employed for insertional mutagenesis in *B. subtilis* to uncover novel regulators of gene expression [41]. Screening of approximately 12,000 insertion mutants carrying a *proB::lacZ* reporter identified the *dnaJK* and *grpE* loci as post-transcriptional regulators of *proB*, revealing a previously unknown link between the DnaJK chaperone system and RNase P-dependent mRNA metabolism. Furthermore, a xylose-inducible *mariner* system was constructed and used to successfully identify sporulation-deficient (*spo0A*) mutants in *Clostridium acetobutylicum*, an anaerobic bacterium with low transformation efficiency [42]. These findings underscore the flexibility of the *mariner* transposon in various microbial hosts.

Through deliberate design and modification, transposon systems can be further engineered to achieve functionalities beyond simple insertional mutagenesis (Fig. 1B). A key strategy involves coupling transposition with reporter systems. For example, in *E. coli*, a modified Tn5 transposon carrying a promoter-less GFP cassette enables selection and reporter assays based on the activity of the promoter upstream of the insertion site [43]. This design combines mutant selection with functional genomics, providing a direct link between genotype and phenotype. Similarly, in yeast, engineered Tn3 and Tn7 transposons, which carry selectable markers such as URA3, *lox* recombination sites, and

sequences encoding multiple copies of the hemagglutinin (HA) epitope, have been widely used for mutagenesis [44,45]. These systematically engineered platforms have proven immensely productive, collectively illuminating the function of nearly 4,000 yeast genes, showcasing the power of high-throughput transposon mutagenesis in model organisms [46]. Notably, a high-throughput strategy developed in *E. coli* employs Tn5 transposase to activate genomic gene transcription for genotype–-phenotype mapping [47]. This system incorporates a selectable marker and an inducible promoter into the cargo DNA, simplifying the selection of transposition mutants while enabling tunable gene expression. The system successfully activated a pre-integrated gene and identified three putative transporters that enhance *E. coli* resistance to glycine, thereby illustrating an innovative application of transposon mutagenesis.

Beyond engineered transposon systems, endogenous bacterial transposons represent an under-explored but valuable reservoir of genetic tools. For instance, the IS*Cg1* transposase from *Corynebacterium glutamicum* has been used to randomly insert a selectable marker into an IS-free strain, though its uncontrolled jumping activities currently limits its broader utility [48]. In another approach, an IS*6100*-based random mutagenesis system in *C. glutamicum* helped identify *cg0910*, a previously unknown gene involved in histidine biosynthesis, and enabled parallel screening of mutants at genomic and metabolic levels [49,50]. Notably, such approaches extend beyond mere gene cataloguing; they can uncover intricate functional dependencies within metabolic networks, such as those among TCA cycle metabolites or between glycine and serine biosynthesis [50], thereby providing insights that may guide the optimization of industrial bioprocesses.

3.1.2. *Transposon-associated high-throughput sequencing technologies*

The construction of random transposon insertion mutant libraries provides a foundational resource for functional genomics research. Subsequently, the rapid and comprehensive mapping of transposon insertion sites across the entire genome relies heavily on high-throughput sequencing. Various transposon insertion sequencing methods, including DLA-454, Mu-Taq, Tn-seq, IN-seq, HITs, Mu-seq, and biotin-oligonucleotide tagging, have been developed to identify transposon insertion loci on the chromosome with high sensitivity and specificity [51–57]. Nonetheless, these methods still face several limitations, including the inability to distinguish multiple insertions at the same genomic site, PCR bias during flanking sequence amplification, and insufficient detection sensitivity for rare mutants within the population.

RB-TnSeq (Random Barcoded Transposon Sequencing) represents a significant methodological advance in microbial functional genomic research. It effectively addresses previous limitations by establishing a standardized, quantitative platform for genome-wide fitness profiling [58]. This method incorporates random barcodes into transposons, which are linked to genomic integration sites via an initial insertion sequencing step (Fig. 2). This design not only improves traceability but also substantially increases throughput while reducing costs compared to earlier transposon insertion sequencing approaches. The robustness of RB-TnSeq was clearly demonstrated in its application, which conducted 387 genome-wide fitness assays across 130 bacteria, identifying 5196 genes associated with significant phenotypes in five bacterial species [58]. This large-scale systematic screening highlighted the power of RB-TnSeq for discovering functional genes. RB-TnSeq was further applied in a subsequent study that characterized 11,779 poorly annotated genes across 32 diverse bacterial species, enabling functional predictions for DNA repair proteins, enzymes, transporters, and uncharacterized protein families [59]. A key limitation of RB-TnSeq, however, is its restriction to profiling nonessential genes under given conditions. To address this, integrated approaches have been developed. For instance, combining RB-TnSeq with Dub-seq in *E. coli* allowed systematic mapping of phage-host interactions, identifying both known and novel resistance factors [60]. Similarly, coupling RB-TnSeq with CRISPRi enabled profiling of both nonessential and essential genes involved in phage infection [60]. Looking forward, RB-TnSeq continues to evolve towards more dynamic and controlled applications. Recently, an inducible, self-propagating transposon system coupled with RB-TnSeq was developed for continuous mutagenesis and dynamic rewiring of gene networks in *E. coli*, demonstrating its utility in studying carbon utilization and antibiotic resistance [61]. Collectively, these

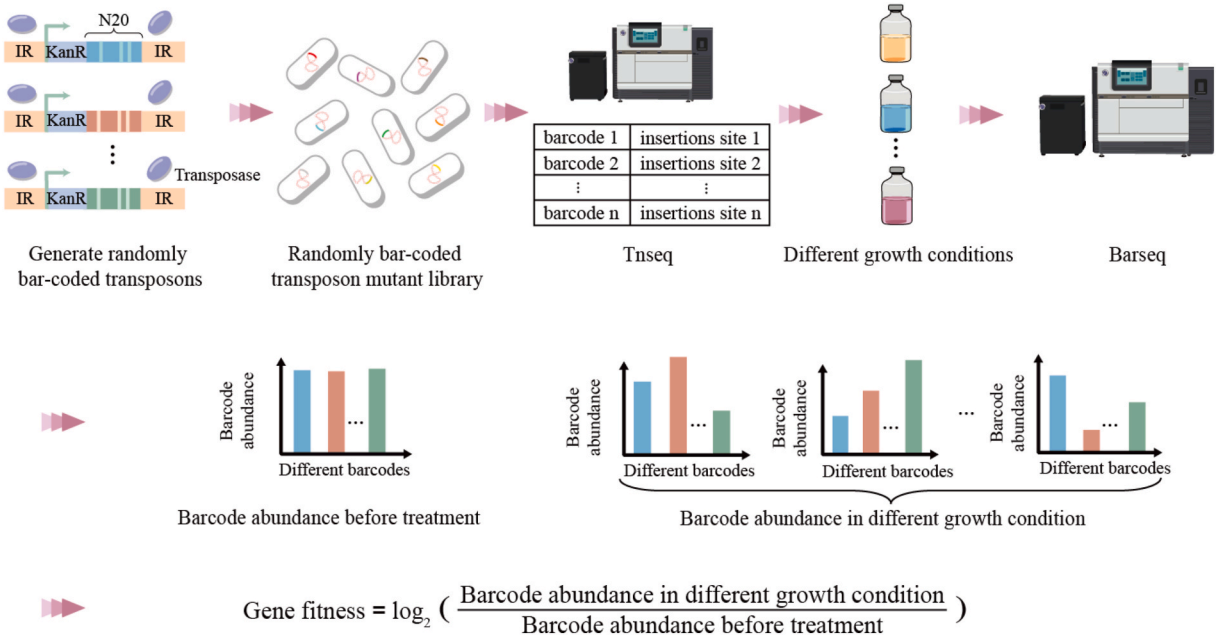

Fig. 2. Operational procedures for RB-TnSeq. A pooled mutant library is firstly generated by integrating transposons (Tn5 or mariner) that carrying unique 20 bp randomized DNA barcodes (N20) into the bacterial genome. A single comprehensive TnSeq procedure is then performed, involving genomic DNA fragmentation, library preparation, and high-throughput sequencing, to establish a precise barcode-to-insertion-site reference map. Then, phenotypic screening is carried out using streamlined BarSeq assays: the mutant pool is subjected to competitive growth under diverse selective conditions, followed by amplification of barcodes via universal primers and sequencing of samples collected before and after cultivation. Gene fitness scores are calculated based on changes in the relative abundance of barcodes.

developments underscore RB-TnSeq not merely as a gene-screening tool, but as a versatile foundational methodology upon which increasingly sophisticated functional genomic strategies can be built.

It should be noted that although transposon integration into the host microbial chromosome is randomly, the issue of "insertion hot spots" exists to varying degrees. The presence of insertion hot spots not only affects the coverage of transposon mutant libraries but also impacts the accuracy of subsequent screening for genes associated with specific phenotypes, particularly leading to the omission of genes located in "non-hot spot insertion regions" [62]. To mitigate this adverse effect, researchers have developed various optimization strategies. For instance, employing a combination of transposons with different targeting preferences, or constructing controllable transposase expression systems, can achieve more uniform genome coverage across different hosts [47,63,64]. Additionally, using long-read transposon insertion site sequencing (LoRTIS) to generate long reads that span repetitive regions can overcome the limitations of conventional short-read sequencing in precisely mapping transposon insertion sites within genomic repetitive regions [65]. Furthermore, leveraging machine learning to decipher the insertion preference mechanisms of specific transposases enables the development of computational correction tools that optimize algorithms, thereby improving the accuracy of high-throughput sequencing data analysis [66].

3.2. Transposon-mediated chromosomal integration of DNA fragments

In the engineering of industrial strains, chromosomal integration of exogenous genes or biosynthetic pathways is a pivotal strategy for constructing high-performance production strains. This approach ensures stable, long-term expression and overcomes the inherent instability often associated with plasmid-based systems [67]. Compared to methods such as homologous recombination and site-specific recombination, transposon-mediated chromosomal integration of large DNA fragments offers distinct advantages, including operational simplicity and broad host compatibility. This makes it especially suitable for genetically challenging non-model microorganisms. Furthermore, the method enables the generation of "clean" engineered strains, free of antibiotic resistance markers or residual scar sequences in the chromosome, thereby improving both biosafety and industrial applicability.

3.2.1. Random chromosomal integration

A major advantage of transposon-mediated random chromosomal integration is its ability to generate a diverse library of chassis strains with integrations at various genomic loci, facilitating the screening of optimal candidates. As shown in previous studies, the genomic integration site of exogenous DNA fragments or pathway genes may critically impact their expression efficiency [68]. This underscores the necessity for systematic evaluation and comparison of integration sites, a challenge not easily addressed by homology-directed recombination or site-specific recombination techniques, which rely on predefined loci. In contrast, transposon-mediated random integration offers a solution that precisely meets this requirement.

The *mariner* transposon system has been utilized to integrate heterologous pathways into autotrophic *C. ljungdahlii* for producing various chemicals. For example, using a rapid library screening approach, one study obtained a strain capable of producing 10.3 g/L of 3-hydroxybutyrate in continuous gas fermentation [69]. The flexibility of this technology is further demonstrated by its successful use in assembling and integrating pathways of varying complexity, i.e., from an 11.1 kb construct for acetone and isopropanol biosynthesis [70] to a more extensive 17.9 kb pathway for syngas-to-alcohols production [71]. This platform also supports iterative pathway optimization. In the latter study, the initial integrated pathway yielded butanol and hexanol, but subsequent CRISPR/Cas9-mediated marker excision and integration of an additional 7.8 kb gene cluster from *C. carboxidivorans* significantly enhanced hexanol production, reaching a titer of 251 mg/L [71].

In a parallel application with a different transposon, the Hermes transposon was employed for random genomic integration to introduce multiple copies of the 2-pyrone synthase (2-PS) expression cassette into *Yarrowia lipolytica*, generating strain yJY2039 [72]. This transposon-enabled strain exhibited enhanced triacetic acid lactone (TAL) production, contributing to a final titer of 2.6 g/L in batch fermentation. This significant improvement facilitated the subsequent one-pot synthesis of the valuable antibiotic pogostone from microbially-derived TAL, achieving a 96% yield. Harnessing native Ty retrotransposon LTR sequences in *Y. lipolytica*, two random genomic integration toolkits, YaliCMulti and YaliHMulti, were developed [73]. By exploiting the natural genomic dispersion of LTR (long terminal repeat) elements and incorporating Cre-loxP-mediated marker recycling, the toolkits facilitate iterative gene integration without selectable marker constraints.

As a further refinement and evolution of this random integration strategy, mini-transposons have been developed. These engineered, simplified DNA constructs are derived from natural transposons [74] and represent an important advancement in genetic engineering due to their reduced size and enhanced versatility. They retain the integration capability of full-length transposons while imposing a reduced genomic burden on the host, thereby improving strain stability and fitness in long-term cultivation. Furthermore, when equipped with distinct selectable markers, mini-transposons can be iteratively applied within the same host [75], greatly facilitating multi-step metabolic engineering and genome editing workflows. One of the most widely adopted systems is the Tn5 transposon. For instance, the mini-Tn5 system has been successfully used to integrate polyhydroxybutyrate (PHB) biosynthesis genes from *Ralstonia eutropha* into the genomes of *E. coli* strains, with PHB yields varying depending on the genomic insertion site [75]. Recently, a novel dual-plasmid mini-Tn5 system, in which one plasmid encodes the Tn5 transposase and the other carries the target gene cluster, was employed to stably integrate PHB synthesis genes into the *E. coli* genome. This approach resulted in enhanced PHB production compared to strains carrying the same gene cluster on a previously used high-copy-number plasmid [76]. Taken together, these studies highlight the potential of transposon-mediated chromosomal integration of large DNA fragments in developing robust industrial microbial chassis, offering a stable, scalable, and burden-minimized alternative to conventional plasmid-based expression systems.

3.2.2. CRISPR-associated transposon (CAST) systems

CRISPR-associated transposon (CAST) systems represent a novel class of genome-editing tools, created by functionally coupling CRISPR-Cas systems with Tn7-like transposons. This fusion merges the precise, RNA-guided DNA recognition of CRISPR with the inherent "cut-and-paste" integration mechanism of transposons [77]. Importantly, this design circumvents a major limitation (the inability to insert DNA directionally) of a simple transposon system, and therefore enables the programmable, directional integration of large DNA cargo into the chromosome without generating double-strand breaks (DSBs). The DSB-free nature of this process significantly reduces the likelihood of unintended genomic rearrangements and cytotoxic effects, positioning CAST systems as particularly advantageous for applications requiring high-fidelity genomic insertions.

To date, seven major subtypes have been identified or bioinformatically predicted (I–F, I–B, I–C, I-E, I-D, IV-A, and V–K), classified according to their associated Cas effector proteins [78,79]. The phylogenetic diversity of CAST systems reflects their broad natural distribution and suggests considerable potential for functional variation. However, they remain largely underexplored, with only a few systems having been adapted for biotechnological use. Specific to industrial microorganisms, CAST systems, including ShCAST, INTEGRATE, and MUCICAT, have showed great potential in genetic engineering [78]. These systems are particularly valuable for assembling complex multi-gene pathways and for metabolic engineering efforts that require

the stable genomic integration of large biosynthetic clusters.

The ShCAST system, derived from *Scytonema hofmanni*, is composed of TniQ, TnsBC, and the single-effector Cas12k (Fig. 3A) [80]. Initially, its utility in *E. coli* was limited to non-coding protospacers in specialized Pir1 strains [80]. To overcome this limitation, an optimized version termed SHOT was developed, which used a strong independent promoter to drive Cas12k expression [81]. SHOT achieves significantly higher integration efficiencies for large DNA fragments and enables targeting of previously inaccessible genomic loci, thereby expanding the utility of the ShCAST platform in industrially relevant microorganisms. Notably, SHOT has demonstrated direct applicability in bioprocess optimization: its use in metabolic intervention, specifically by reducing acidic byproduct formation, led to enhanced target protein yield, concretely illustrating its potential for industrial strain engineering [81].

INTEGRATE is derived from the *Vibrio cholerae* Tn6677 transposon and comprises TnsABC, TniQ, and the multi-subunit Cas678 complex (Fig. 3B) [82,83]. This system has demonstrated broad host compatibility, functioning efficiently in diverse microbial hosts ranging from *Lactococcus lactis* to *Corynebacterium glutamicum*. In *L. lactis*, integration of a 10 kb fragment was achieved with an efficiency only about fivefold lower than that for a 1 kb fragment [84], indicating relatively low size dependency compared to many other integration tools. Furthermore, in *C. glutamicum*, the INTEGRATE system combined with a spectinomycin-resistance marker enabled the integration of a 9 kb cargo with an efficiency of approximately 13.4% [85].

MUCICAT is an engineered CAST/INTEGRATE variant, specifically designed for multicopy chromosomal integration. It employs crRNAs targeting multicopy genomic loci or crRNA arrays for multiplexed targeting (Fig. 3C) [86]. When applied to engineer the N-acetylglucosamine (GlcNAc) biosynthesis pathway in *E. coli*, MUCICAT enabled the identification of an optimal strain harboring five genomic copies of the expression cassette. This strain produced 11.59 g/L of GlcNAc, a yield more than sixfold higher than that achieved by a plasmid-based control [87]. The system has also been used to construct ATP energy cycles and to integrate functional genes into the chromosome, with a series of evaluations, including insertion efficiency, genomic context effects, and copy number-performance relationship, conducted [88]. Thus, the development of MUCICAT represents a strategic advance in transposon-based genome editing, shifting the paradigm from single-locus editing toward scalable, combinatorial genomic writing.

3.3. Strategies for stabilizing chromosomal transposon integration

For industrial microorganisms, genetic instability arising from transposon reversion or secondary transposition has emerged as a critical factor leading to strain degeneration, yield fluctuation, and even fermentation failure [89]. To address this challenge, researchers have developed various strategies.

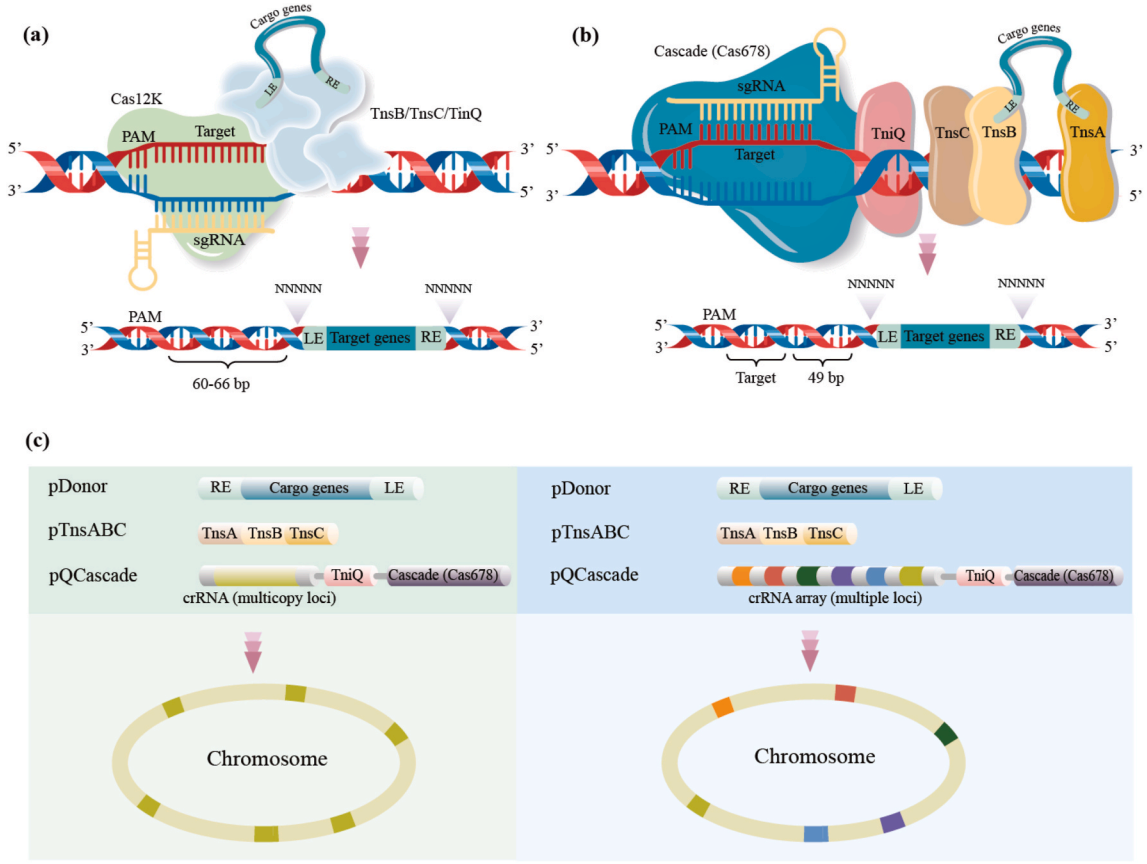

Fig. 3. Models for CRISPR-associated transposons. (a) ShCAST, a natural CRISPR-Cas/Tn7-like transposon fusion. Its catalytically inactive Cas12k binds target DNA via crRNA. The transposase (TnsB, TnsC, TniQ) then mediates precise, unidirectional insertion of donor DNA (flanked by LE/RE ends) into a genomic region approximately 60-66 downstream of the PAM, creating a 5 bp duplication (a Tn7-family hallmark). Unlike standard CRISPR editing, this cut-free mechanism requires no double-strand breaks or host repair, enabling efficient, site-specific integration of large fragments, even in non-dividing cells. **(b)** The INTEGRATION system. It operates on a principle similar to ShCAST, with specific differences in its component architecture. It employs a *Vibrio cholerae* Tn6677-derived Cascade (Cas678) complex for crRNA-guided target binding without DNA cleavage. Its adaptor protein TniQ bridges Cas678 with the transposase (TnsA, TnsB, TnsC), directing precise integration of donor DNA ~49 bp downstream of the protospacer. **(c)** MUCICAT. It is engineered based on the INTEGRATION system and consists of three modular components: pTnsABC, pDonor, and pQCascade. It is distinguished by its ability to use crRNA to simultaneously target multicopy genomic loci or multiple distinct loci.

First, for the selection of integration sites, anchoring exogenous genes in genomically highly stable regions has proven effective in minimizing the risk of genetic drift. For instance, in *Yarrowia lipolytica*, researchers have identified multiple "safe harbor" sites based on essential genes. Exogenous genes integrated at these sites exhibited better genetic stability compared to traditional sites in simulated scale-up cultures [90]. Second, the use of low-copy-number transposons (e.g., Tn5/Tn10) or suicide vectors facilitates the stable, single-copy integration of exogenous genes into the host chromosome. This method effectively circumvents the genetic instability commonly associated with multicopy insertion events. [91]. Third, leveraging inducible promoters further allows for precise control over both the timing and the intensity of transposase expression [42,47], thereby mitigating the risk of transposon reversion or secondary transposition.

Additionally, the strategies employed in other types of organisms to stabilize transposon mutations or chromosomal integration events also merit consideration for application in industrial microorganisms. For example, researchers have constructed transposon vectors with tandem repeat termini. It enabled transposase-induced deletion of one terminal sequence following genomic integration, which can permanently inactivate the transposition activity of the remaining vector, thereby achieving stable fixation of the transgene [92]. An alternative strategy involves incorporating recombinase recognition sites into the transposon vector. Following chromosomal integration, the subsequent introduction of the corresponding recombinase excises the vector backbone, thereby ensuring the stability of the integrated transposon [93].

Although transposon-related technologies have been widely applied in model industrial microorganisms such as *E. coli* and *B. subtilis*, extending these tools to non-model strains presents several challenges. For instance, high-GC content bacteria (e.g., *Corynebacterium glutamicum* and *Streptomyces* spp.) possess genomes with complex secondary structures [94], which may influence transposase insertion efficiency and insertion site distribution on the chromosomes. Furthermore, these microorganisms often possess efficient DNA repair systems [95,96], which may identify the double-strand breaks generated by transposon insertion as genomic damage. This recognition can subsequently trigger the excision of the transposon via the repair pathway, thereby rendering the insertion mutation ineffective. Similarly, for some unconventional yeasts (e.g., *Yarrowia* or *Pichia*), their compact chromatin structure and the NHEJ repair pathways may also affect transposon insertion efficiency and distribution. Consequently, the development of transposon tools for these non-model industrial microbes should consider suppressing key host repair factors, such as Ku70 or LigD, to establish an intracellular environment more conducive to stable transposon integration.

4. Conclusions and outlook

This review systematically elaborates on the principles and values of transposon-based molecular genetic tools in decoding and reconstructing industrial microorganisms. It serves not only as a vital complement to the existing toolbox but also as the preferred tool for exploring the "dark matter" of industrial microorganisms.

However, current transposon-based technologies still face challenges. Gram-positive industrial microorganisms generally exhibit low exogenous plasmid introduction efficiency, which stems from core obstacles: the physical barrier of the thick cell wall and the presence of powerful restriction-modification systems [97]. Furthermore, most transposon systems exhibit host specificity, and their efficiency drops significantly when applied to non-model industrial microorganisms, which greatly limits the general applicability of transposon-based technologies [25]. In the field of genome editing, although transposon-based CAST systems have enabled targeted integration of DNA fragments into chromosomes, their current application is hindered by significant off-target effects, necessitating further optimization [78].

Future research should prioritize the following key directions. First, expanding the application boundaries. A central task is to develop transposon systems with broader host applicability, particularly those non-model industrial microorganisms. Artificial systems such as the abovementioned CAST represent a highly promising breakthrough. By engineering and adapting their protein components, it may be possible to build a more "universal" genetic tool applicable across a wider range of industrial microorganisms. Second, deepening the transposon-based technology with artificial intelligence (AI) to revolutionize the research paradigm. Currently, machine learning is driving innovative applications of transposon technology in industrial microbiology through two primary approaches, i.e., multi-omics data integration and protein function prediction. By leveraging machine learning models, researchers can integrate multidimensional features such as the physicochemical properties of DNA sequences, chromatin accessibility, and gene transcriptional activity to accurately predict transposon insertion preferences across diverse microbial genomes [98]. Building on this, training models on large-scale Tn-Seq datasets could enable the effective extraction of key sequence features governing transposon insertion efficiency. This would provide technical support for constructing high-coverage mutant libraries or achieving efficient targeted integration, thereby facilitating the discovery of functional genes associated with specific phenotypes. Concurrently, deep learning-based protein language models are progressively supplanting traditional "blind screening" approaches in directed evolution. Through model training on vast sequence datasets, researchers can gain deeper insights into the evolutionary trajectories and structural constraints of transposase families, which will provide a theoretical foundation for the rational design of artificial transposases with enhanced activity and stability. Furthermore, in CRISPR-associated transposase (CAST) systems, AI techniques can also predict the efficiency of guide RNA-directed transposition events based on the sequence characteristics flanking the target site, thereby improving the precision and controllability of gene editing and integration. Third, exploring new biological mechanisms to empower tool innovation. Continued exploration of transposon biology remains a fertile source for catalyzing next-generation molecular tools. In particular, the discovery and mechanistic elucidation of novel transposon systems hold the potential to expand the functional boundaries and application scenarios of transposon-related technologies.

In summary, transposon-based technology has evolved from a classical random mutagenesis tool into a versatile and foundational platform in modern industrial microbiology. It is now advancing toward greater precision, intelligence, and broad applicability. The continued integration of transposon systems with CRISPR-Cas technologies, systems biology, and data science promises to enable more systematic and effective analysis, design, and reprogramming of microbial cell factories. However, ensuring genetic stability and biosafety constitutes a fundamental prerequisite for industrial application of engineered strains. Given that residual mobile elements in engineered strains pose potential risks such as genomic instability and resistance gene transfer, prioritizing the development of marker-free chassis strains with well-defined genetic backgrounds represents an inevitable pathway for mitigating biosafety risks and promoting the sustainable development of synthetic biology in the future.

CRediT authorship contribution statement

Zhiyi Zhang: Writing – original draft, Visualization, Validation, Investigation, Conceptualization. **Zheng Huang:** Validation, Investigation. **Yilin Wang:** Investigation. **Zheng Li:** Investigation. **Zhiqiang Wen:** Supervision, Project administration. **Yang Gu:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal

relationships which may be considered as potential competing interests: Zhiqiang Wen is currently employed by Nanjing Gasgene Biochemical Technology Co. Ltd.

Acknowledgement

This work was supported by the supported by the National Key R&D Program of China (No. 2025YFA0922200 to Y.G.), the Frontier Technology R&D Program of Jiangsu Province (BF2024079), the Jiangsu Carbon Peak and Carbon Neutrality Science and Technology Innovation Special Project (BT2025013), Science and Technology Commission of Shanghai Municipality (25ZR1401262, 24HC2820800), Basic Research Program of Jiangsu (2025CSJZN01400), and the National Natural Science Foundation of China (32571661).

References

[1] McElwain L, Phair K, Kealey C, Brady D. Current trends in biopharmaceuticals production in *Escherichia coli*. Biotechnol Lett 2022;44(8):917–31. https://doi.org/10.1007/s10529-022-03276-5.
[2] Chen T, Uzunovic H, Brul S, Hugenholtz J. *Bacillus subtilis* as cell factory for enhanced production of the biopolymer precursor pyridine-2,6-dicarboxylic acid. Microb Cell Fact 2025;24(1):233. https://doi.org/10.1186/s12934-025-02859-x.
[3] Nielsen J. The power of yeast. Yeast 2025;42(12):303–10. https://doi.org/10.1002/yea.70009.
[4] Fernandez-Blanco C, Robles-Iglesias R, Naveira-Pazos C, Veiga MC, Kennes C. Production of biofuels from C_1-gases with *Clostridium* and related bacteria-recent advances. Microb Biotechnol 2023;16(4):726–41. https://doi.org/10.1111/1751-7915.14220.
[5] Cai P, Wu X, Deng J, Gao L, Shen Y, Yao L, et al. Methanol biotransformation toward high-level production of fatty acid derivatives by engineering the industrial yeast *Pichia pastoris*. Proc Natl Acad Sci U S A 2022;119(29):e2201711119. https://doi.org/10.1073/pnas.2201711119.
[6] Tian J, Deng W, Zhang Z, Xu J, Yang G, Zhao G, et al. Discovery and remodeling of *Vibrio natriegens* as a microbial platform for efficient formic acid biorefinery. Nat Commun 2023;14(1):7758. https://doi.org/10.1038/s41467-023-43631-2.
[7] Thomason LC, Costantino N, Li X, Court DL. Recombineering: genetic engineering in *Escherichia coli* using homologous recombination. Curr Protoc 2023;3(2):e656. https://doi.org/10.1002/cpz1.656.
[8] Amen T, Kaganovich D. Integrative modules for efficient genome engineering in yeast. Microb Cell 2017;4(6):182–90. https://doi.org/10.15698/mic2017.06.576.
[9] 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. https://doi.org/10.1093/nar/gkn991.
[10] Meijer WJJ, Miguel-Arribas A. Genetic engineering of *Bacillus subtilis* using competence-induced homologous recombination techniques. Methods Mol Biol 2024;2819:241–60. https://doi.org/10.1007/978-1-0716-3930-6_12.
[11] Xu M, Yang N, Pan J, Hua Q, Li CX, Xu JH. Remodeling the homologous recombination mechanism of *Yarrowia lipolytica* for high-level biosynthesis of squalene. J Agric Food Chem 2024;72(17):9984–93. https://doi.org/10.1021/acs.jafc.4c01779.
[12] Che S, Zhuo Y, Yang L, Wang H, Cui Z, Fan J. Multiple genes deletion based on Cre-*loxP* marker-less gene deletion system for the strains from the genus of *Pectobacterium*. Biotechnol Lett 2024;46(6):1133–42. https://doi.org/10.1007/s10529-024-03518-8.
[13] Dan J, Deng H, Xia Y, Zhan Y, Tang N, Wang Y, et al. Application of the *FLP/LoxP-FRT* recombination system to switch the eGFP expression in a model prokaryote. Open Life Sci 2022;17(1):172–9. https://doi.org/10.1515/biol-2022-0019.
[14] Pacesa M, Pelea O, Jinek M. Past, present, and future of CRISPR genome editing technologies. Cell 2024;187(5):1076–100. https://doi.org/10.1016/j.cell.2024.01.042.
[15] Cardiff RAL, Faulkner ID, Beall JG, Carothers JM, Zalatan JG. CRISPR-Cas tools for simultaneous transcription & translation control in bacteria. Nucleic Acids Res 2024;52(9):5406–19. https://doi.org/10.1093/nar/gkae275.
[16] Hickman AB, Dyda F. DNA transposition at work. Chem Rev 2016;116(20):12758–84. https://doi.org/10.1021/acs.chemrev.6b00003.
[17] Hwang J, Ye DY, Jung GY, Jang S. Mobile genetic element-based gene editing and genome engineering: recent advances and applications. Biotechnol Adv 2024;72:108343. https://doi.org/10.1016/j.biotechadv.2024.108343.
[18] Van Opijnen T, Camilli A. Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. Nat Rev Microbiol 2013;11(7):435–42. https://doi.org/10.1038/nrmicro3033.
[19] Gelsinger DR, Vo PLH, Klompe SE, Ronda C, Wang HH, Sternberg SH. Bacterial genome engineering using CRISPR-associated transposases. Nat Protoc 2024;19(3):752–90. https://doi.org/10.1038/s41596-023-00927-3.
[20] Schmitz M, Querques I. DNA on the move: mechanisms, functions and applications of transposable elements. FEBS Open Bio 2024;14(1):13–22. https://doi.org/10.1002/2211-5463.13743.
[21] Knights HE, Jorrin B, Haskett TL, Poole PS. Deciphering bacterial mechanisms of root colonization. Environ Microbiol Rep 2021;13(4):428–44. https://doi.org/10.1111/1758-2229.12934.
[22] Finnegan DJ. Retrotransposons. Curr Biol 2012;22(11):R432–7. https://doi.org/10.1016/j.cub.2012.04.025.
[23] Hickman AB, Dyda F, Chandler M, Craig N. Mechanisms of DNA transposition. Microbiol Spectr 2015;3(2). https://doi.org/10.1128/microbiolspec.MDNA3-0034-2014.
[24] Feng G, Leem YE, Levin HL. Transposon integration enhances expression of stress response genes. Nucleic Acids Res 2013;41(2):775–89. https://doi.org/10.1093/nar/gks1185.
[25] Wu S, Tian P, Tan T. Genomic landscapes of bacterial transposons and their applications in strain improvement. Appl Microbiol Biotechnol 2022;106(19–20):6383–96. https://doi.org/10.1007/s00253-022-12170-z.
[26] Siguier P, Gourbeyre E, Chandler M. Bacterial insertion sequences: their genomic impact and diversity. FEMS Microbiol Rev 2014;38(5):865–91. https://doi.org/10.1111/1574-6976.12067.
[27] Liang Q, Kong J, Stalker J, Bradley A. Chromosomal mobilization and reintegration of *Sleeping Beauty* and *PiggyBac* transposons. Genesis 2009;47(6):404–8. https://doi.org/10.1002/dvg.20508.
[28] Mendonca AA, de Lucena BT, de Morais MM, de Morais Jr MA. First identification of Tn916-like element in industrial strains of *Lactobacillus vini* that spread the tet-*M* resistance gene. FEMS Microbiol Lett 2016;363(3). https://doi.org/10.1093/femsle/fnv240.
[29] Wang S, Zhang P, Sun Y, Fang Y, Wang P, Shao M, et al. Engineering of *BZ* transposase and transposon donor vector for enhanced efficiency and safety in gene delivery applications. Nucleic Acids Res 2025;53(18). https://doi.org/10.1093/nar/gkaf935.
[30] Jin Y, Chen Y, Zhao S, Guan KL, Zhuang Y, Zhou W, et al. DNA-PK facilitates piggyBac transposition by promoting paired-end complex formation. Proc Natl Acad Sci U S A 2017;114(28):7408–13. https://doi.org/10.1073/pnas.1612980114.
[31] Fernandez-Garcia G, Valdes-Chiara P, Villazan-Gamonal P, Alonso-Fernandez S, Manteca A. Essential genes discovery in microorganisms by transposon-directed sequencing (Tn-Seq): experimental approaches, major goals, and future perspectives. Int J Mol Sci 2024;25(20). https://doi.org/10.3390/ijms252011298.
[32] Kwon YM, Ricke SC, Mandal RK. Transposon sequencing: methods and expanding applications. Appl Microbiol Biotechnol 2016;100(1):31–43. https://doi.org/10.1007/s00253-015-7037-8.
[33] Barquist L, Boinett CJ, Cain AK. Approaches to querying bacterial genomes with transposon-insertion sequencing. RNA Biol 2013;10(7):1161–9. https://doi.org/10.4161/rna.24765.
[34] Dong Z, Li L, Du G, Zhang Y, Wang X, Li S, et al. A previously unidentified sugar transporter for engineering of high-yield *Streptomyces*. Appl Microbiol Biotechnol 2024;108(1):72. https://doi.org/10.1007/s00253-023-12964-9.
[35] Bao HY, Li HJ, Zhang YY, Bechthold A, Yu XP, Ma Z. Transposon-based identification of genes involved in the rimocidin biosynthesis in *Streptomyces rimosus* M527. World J Microbiol Biotechnol 2023;39(12):359. https://doi.org/10.1007/s11274-023-03814-x.
[36] Ito M, Kim YG, Tsuji H, Kiwaki M, Nomoto K, Tanaka R, et al. A practical random mutagenesis system for probiotic *Lactobacillus casei* using Tn5 transposition complexes. J Appl Microbiol 2010;109(2):657–66. https://doi.org/10.1111/j.1365-2672.2010.04690.x.
[37] Ito M, Kim YG, Tsuji H, Takahashi T, Kiwaki M, Nomoto K, et al. Transposon mutagenesis of probiotic *Lactobacillus casei* identifies *asnH*, an asparagine synthetase gene involved in its immune-activating capacity. PLoS One 2014;9(1):e83876. https://doi.org/10.1371/journal.pone.0083876.
[38] Foreman M, Gershoni M, Barkan D. A simplified and efficient method for Himar-1 transposon sequencing in bacteria, demonstrated by creation and analysis of a saturated transposon-mutant library in *Mycobacterium abscessus*. mSystems 2020;5(5). https://doi.org/10.1128/mSystems.00976-20.
[39] Jeong DE, So Y, Park SY, Park SH, Choi SK. Random knock-in expression system for high yield production of heterologous protein in *Bacillus subtilis*. J Biotechnol 2018;266:50–8. https://doi.org/10.1016/j.jbiotec.2017.12.007.
[40] Ye B, Tao Q, Yan X. A transposon system for random insertion of a gene expression cassette into the chromosome of *Bacillus subtilis*. J Biotechnol 2023;361:66–73. https://doi.org/10.1016/j.jbiotec.2022.11.020.
[41] Ogura M, Kanesaki Y, Yoshikawa H, Haga K. The DnaJK chaperone of *Bacillus subtilis* post-transcriptionally regulates gene expression through the YlxR(RnpM)/RNase P complex. mBio 2025;16(3). https://doi.org/10.1128/mbio.04053-24.
[42] Zhang Y, Xu S, Chai C, Yang S, Jiang W, Minton NP, et al. Development of an inducible transposon system for efficient random mutagenesis in *Clostridium acetobutylicum*. FEMS Microbiol Lett 2016;363(8). https://doi.org/10.1093/femsle/fnw065.
[43] Nazareno ES, Acharya B, Dumenyo CK. A mini-Tn5-derived transposon with reportable and selectable markers enables rapid generation and screening of insertional mutants in gram-negative bacteria. Lett Appl Microbiol 2020;72(3):283–91. https://doi.org/10.1111/lam.13423.
[44] Ross-Macdonald P, Coelho PSR, Roemer T, Agarwal S, Kumar A, Jansen R, et al. Large-scale analysis of the yeast genome by transposon tagging and gene disruption. Nature 1999;402(6760):413–8. https://doi.org/10.1038/46558.
[45] Kumar A, Seringhaus M, Biery MC, Sarnovsky RJ, Umansky L, Piccirillo S, et al. Large-scale mutagenesis of the yeast genome using a Tn7-derived multipurpose transposon. Genome Res 2004;14(10a):1975–86. https://doi.org/10.1101/gr.2875304.

[46] Kumar A. Multipurpose transposon-insertion libraries in yeast. Cold Spring Harb Protoc 2016;2016(6). https://doi.org/10.1101/pdb.top080259.
[47] Song Y, Liu Y, Li Q, Chen L, Sun Z, Zhang X. A Tn5 transposase-based system for high-efficiency genome-wide gene activation in *Escherichia coli*. ACS Synth Biol 2025;14(7):2753–63. https://doi.org/10.1021/acssynbio.5c00170.
[48] Linder M, Haak M, Botes A, Kalinowski J, Ruckert C. Construction of an IS-free *Corynebacterium glutamicum* ATCC 13032 chassis strain and random mutagenesis using the endogenous IS*Cg1* transposase. Front Bioeng Biotechnol 2021;9:751334. https://doi.org/10.3389/fbioe.2021.751334.
[49] Mormann S, Lomker A, Ruckert C, Gaigalat L, Tauch A, Puhler A, et al. Random mutagenesis in *Corynebacterium glutamicum* ATCC 13032 using an IS6100-based transposon vector identified the last unknown gene in the histidine biosynthesis pathway. BMC Genom 2006;7:205. https://doi.org/10.1186/1471-2164-7-205.
[50] Reimer LC, Spura J, Schmidt-Hohagen K, Schomburg D. High-throughput screening of a *Corynebacterium glutamicum* mutant library on genomic and metabolic level. PLoS One 2014;9(2):e86799. https://doi.org/10.1371/journal.pone.0086799.
[51] van Opijnen T, Bodi KL, Camilli A. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. Nat Methods 2009;6(10):767–72. https://doi.org/10.1038/nmeth.1377.
[52] Goodman AL, McNulty NP, Zhao Y, Leip D, Mitra RD, Lozupone CA, et al. Identifying genetic determinants needed to establish a human gut symbiont in its habitat. Cell Host Microbe 2009;6(3):279–89. https://doi.org/10.1016/j.chom.2009.08.003.
[53] Gawronski JD, Wong SM, Giannoukos G, Ward DV, Akerley BJ. Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for *Haemophilus* genes required in the lung. Proc Natl Acad Sci U S A 2009;106(38):16422–7. https://doi.org/10.1073/pnas.0906627106.
[54] McCarty DR, Latshaw S, Wu S, Suzuki M, Hunter CT, Avigne WT, et al. Mu-seq: sequence-based mapping and identification of transposon induced mutations. PLoS One 2013;8(10):e77172. https://doi.org/10.1371/journal.pone.0077172.
[55] Winterberg KM, Reznikoff WS. Screening transposon mutant libraries using full-genome oligonucleotide microarrays. Methods Enzymol 2007;421:110–25. https://doi.org/10.1016/S0076-6879(06)21011-5.
[56] Liang L, Wang Y, Han Y, Chen Y, Li M, Wu Y, et al. Expansion and improvement of ChinaMu by MuT-seq and chromosome-level assembly of the *Mu*-starter genome. J Integr Plant Biol 2024;66(4):645–59. https://doi.org/10.1111/jipb.13637.
[57] Liu S, Dietrich CR, Schnable PS. DLA-based strategies for cloning insertion mutants: cloning the *gl4* locus of maize using *Mu* transposon tagged alleles. Genetics 2009;183(4):1215–25. https://doi.org/10.1534/genetics.109.108936.
[58] Wetmore KM, Price MN, Waters RJ, Lamson JS, He J, Hoover CA, et al. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. mBio 2015;6(3):e00306–15. https://doi.org/10.1128/mBio.00306-15.
[59] Price MN, Wetmore KM, Waters RJ, Callaghan M, Ray J, Liu H, et al. Mutant phenotypes for thousands of bacterial genes of unknown function. Nature 2018;557(7706):503–9. https://doi.org/10.1038/s41586-018-0124-0.
[60] Mutalik VK, Adler BA, Rishi HS, Piya D, Zhong C, Koskella B, et al. High-throughput mapping of the phage resistance landscape in *E. coli*. PLoS Biol 2020;18(10):e3000877. https://doi.org/10.1371/journal.pbio.3000877.
[61] English MA, Alcantar MA, Collins JJ. A self-propagating, barcoded transposon system for the dynamic rewiring of genomic networks. Mol Syst Biol 2023;19(6):e11398. https://doi.org/10.15252/msb.202211398.
[62] Choe D, Kim U, Hwang S, Seo SW, Kim D, Cho S, et al. Revealing causes for false-positive and false-negative calling of gene essentiality in *Escherichia coli* using transposon insertion sequencing. mSystems 2023;8(1). https://doi.org/10.1128/msystems.00896-22.
[63] Seringhaus M, Kumar A, Hartigan J, Snyder M, Gerstein M. Genomic analysis of insertion behavior and target specificity of mini-Tn7 and Tn3 transposons in *Saccharomyces cerevisiae*. Nucleic Acids Res 2006;34(8):e57. https://doi.org/10.1093/nar/gkl184.
[64] Levitan A, Gale AN, Dallon EK, Kozan DW, Cunningham KW, Sharan R, et al. Comparing the utility of in vivo transposon mutagenesis approaches in yeast species to infer gene essentiality. Curr Genet 2020;66(6):1117–34. https://doi.org/10.1007/s00294-020-01096-6.
[65] Yasir M, Turner AK, Lott M, Rudder S, Baker D, Bastkowski S, et al. Long-read sequencing for identification of insertion sites in large transposon mutant libraries. Sci Rep 2022;12(1). https://doi.org/10.1038/s41598-022-07557-x.
[66] Zhang H, Lu T, Liu S, Yang J, Sun G, Cheng T, et al. Comprehensive understanding of Tn5 insertion preference improves transcription regulatory element identification. NAR Genom Bioinf 2021;3(4). https://doi.org/10.1093/nargab/lqab094.
[67] Wu Y, Jiang W, Gu Y. Chromosomal integration of large DNA fragments in microorganisms: a review. Sheng Wu Gong Cheng Xue Bao 2023;39(3):842–57. https://doi.org/10.13345/j.cjb.220737.
[68] Bryant JA, Sellars LE, Busby SJW, Lee DJ. Chromosome position effects on gene expression in *Escherichia coli* K-12. Nucleic Acids Res 2014;42(18):11383–92. https://doi.org/10.1093/nar/gku828.
[69] Wu Y, Ye W, Li R, Jia D, Zhu X, Jiang W, et al. Engineering *Clostridium ljungdahlii* for efficient 3-hydroxybutyrate production from one-carbon gases. Green Synth Catal 2025. https://doi.org/10.1016/j.gresc.2025.01.001.
[70] Philipps G, de Vries S, Jennewein S. Development of a metabolic pathway transfer and genomic integration system for the syngas-fermenting bacterium *Clostridium ljungdahlii*. Biotechnol Biofuels 2019;12:112. https://doi.org/10.1186/s13068-019-1448-1.
[71] Lauer I, Philipps G, Jennewein S. Metabolic engineering of *Clostridium ljungdahlii* for the production of hexanol and butanol from CO_2 and H_2. Microb Cell Fact 2022;21(1):85. https://doi.org/10.1186/s12934-022-01802-8.
[72] Yu J, Landberg J, Shavarebi F, Bilanchone V, Okerlund A, Wanninayake U, et al. Bioengineering triacetic acid lactone production in *Yarrowia lipolytica* for pogostone synthesis. Biotechnol Bioeng 2018;115(9):2383–8. https://doi.org/10.1002/bit.26733.
[73] Liu M, Wu J, Yue M, Ning Y, Guan X, Gao S, et al. YaliCMulti and YaliHMulti: stable, efficient multi-copy integration tools for engineering *Yarrowia lipolytica*. Metab Eng 2024;82:29–40. https://doi.org/10.1016/j.ymben.2024.01.003.
[74] Kahrs AF, Odenbreit S, Schmitt W, Heuermann D, Meyer TF, Haas R. An improved Tn*Max* mini-transposon system suitable for sequencing shuttle mutagenesis and gene fusions. Gene 1995;167(1–2):53–7. https://doi.org/10.1016/0378-1119(95)00671-0.
[75] Martinez-Garcia E, Aparicio T, de Lorenzo V, Nikel PI. New transposon tools tailored for metabolic engineering of gram-negative microbial cell factories. Front Bioeng Biotechnol 2014;2:46. https://doi.org/10.3389/fbioe.2014.00046.
[76] Liu M, Ge W, Zhong G, Yang Y, Xun L, Xia Y. Dual-plasmid mini-Tn5 system to stably integrate multicopy of target genes in *Escherichia coli*. ACS Synth Biol 2024;13(11):3523–38. https://doi.org/10.1021/acssynbio.4c00140.
[77] Peters JE, Makarova KS, Shmakov S, Koonin EV. Recruitment of CRISPR-Cas systems by Tn7-like transposons. Proc Natl Acad Sci U S A 2017;114(35):E7358–66. https://doi.org/10.1073/pnas.1709035114.
[78] Chen C, Li YW, Zheng YY, Li XJ, Wu N, Guo Q, et al. Expanding the frontiers of genome engineering: a comprehensive review of CRISPR-associated transposons. Biotechnol Adv 2025;78:108481. https://doi.org/10.1016/j.biotechadv.2024.108481.
[79] Rybarski JR, Hu K, Hill AM, Wilke CO, Finkelstein IJ. Metagenomic discovery of CRISPR-associated transposons. Proc Natl Acad Sci U S A 2021;118(49). https://doi.org/10.1073/pnas.2112279118.
[80] Strecker J, Ladha A, Gardner Z, Schmid-Burgk JL, Makarova KS, Koonin EV, et al. RNA-guided DNA insertion with CRISPR-associated transposases. Science 2019;365(6448):48–53. https://doi.org/10.1126/science.aax9181.
[81] Chang C-W, Huang J-W, Lu Y-H, Pham NN, Tu J, Tung Y-T, et al. Metabolic engineering of *Escherichia coli* to enhance protein production by coupling ShCAST-based optimized transposon system and CRISPR interference. J Taiwan Inst Chem Eng 2023;144. https://doi.org/10.1016/j.jtice.2023.104746.
[82] Klompe SE, Vo PLH, Halpin-Healy TS, Sternberg SH. Transposon-encoded CRISPR-Cas systems direct RNA-guided DNA integration. Nature 2019;571(7764):219–25. https://doi.org/10.1038/s41586-019-1323-z.
[83] Vo PLH, Ronda C, Klompe SE, Chen EE, Acree C, Wang HH, et al. CRISPR RNA-guided integrases for high-efficiency, multiplexed bacterial genome engineering. Nat Biotechnol 2021;39(4):480–9. https://doi.org/10.1038/s41587-020-00745-y.
[84] Pechenov PY, Garagulya DA, Stanovov DS, Letarov AV. New effective method of lactococcus genome editing using guide RNA-directed transposition. Int J Mol Sci 2022;23(22). https://doi.org/10.3390/ijms232213978.
[85] Yang S, Zhu J, Zhou X, Zhang J, Li Q, Bian F, et al. RNA-guided DNA transposition in *Corynebacterium glutamicum* and *Bacillus subtilis*. ACS Synth Biol 2023;12(7):2198–202. https://doi.org/10.1021/acssynbio.3c00193.
[86] Zhang Y, Sun X, Wang Q, Xu J, Dong F, Yang S, et al. Multicopy chromosomal integration using CRISPR-associated transposases. ACS Synth Biol 2020;9(8):1998–2008. https://doi.org/10.1021/acssynbio.0c00073.
[87] Zhang Y, Yang J, Yang S, Zhang J, Chen J, Tao R, et al. Programming cells by multicopy chromosomal integration using CRISPR-associated transposases. CRISPR J 2021;4(3):350–9. https://doi.org/10.1089/crispr.2021.0018.
[88] Ren LM, Qi YH, Cao FY, Zhou EP. Study on the framework of ATP energy cycle system in *Escherichia coli*. Appl Microbiol Biotechnol 2025;109(1):42. https://doi.org/10.1007/s00253-024-13350-9.
[89] Peng M, Liang Z. Degeneration of industrial bacteria caused by genetic instability. World J Microbiol Biotechnol 2020;36(8):119. https://doi.org/10.1007/s11274-020-02901-7.
[90] Yao ZY, Perea J, Shao ZY. Investigation of genomic loci between essential genes as stable harbors for heterologous expression in *Yarrowia Lipolytica*. In: Proceeding of 2024 AIChE Annual Meeting; 2024. 325e. https://proceedings.aiche.org/conferences/aiche-annual-meeting/2024/proceeding/paper/325e-investigation-genomic-loci-between-essential-genes-stable-harbors-heterologous-expression.
[91] Herrero M, de Lorenzo V, Timmis KN. Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. J Bacteriol 1990;172(11):6557–67. https://doi.org/10.1128/jb.172.11.6557-6567.1990.
[92] Handler AM, Zimowska GJ, Horn C. Post-integration stabilization of a transposon vector by terminal sequence deletion in *Drosophila melanogaster*. Nat Biotechnol 2004;22(9):1150–4. https://doi.org/10.1038/nbt1002.
[93] Horn C, Handler AM. Site-specific genomic targeting in *Drosophila*. Proc Natl Acad Sci U S A 2005;102(35):12483–8. https://doi.org/10.1073/pnas.0504305102.
[94] Blaha GM, Wade JT. Transcription-translation coupling in bacteria. Annu Rev Genet 2022;56:187–205. https://doi.org/10.1146/annurev-genet-072220-033342.
[95] Hoff G, Bertrand C, Piotrowski E, Thibessard A, Leblond P. Genome plasticity is governed by double strand break DNA repair in streptomyces. Sci Rep 2018;8(1):5272. https://doi.org/10.1038/s41598-018-23622-w.
[96] Ju Y, Zhang HY, Du XC, Wei JX, Liu J, Wei L, et al. DRAGON: harnessing the power of DNA repair for accelerating genome evolution in *Corynebacterium*

glutamicum. Metab Eng 2023;79:182–91. https://doi.org/10.1016/j.ymben.2023.08.002.

[97] Rozanov AS, Shaposhnikov LA, Bondarenko KD, Sazonov AE. Advances in genetic transformation of Lactic Acid bacteria: overcoming barriers and enhancing plasmid tools. Int J Mol Sci 2025;26(18). https://doi.org/10.3390/ijms26189146.

[98] Ayat M, Domaratzki M. Prediction of insertion-site preferences of transposons using support vector machines and artificial neural networks. In: Elisa B, Wen G, Bernhard S, Moti Y, editors. Lecture notes in computer science. Switzerland: Springer International Publishing; 2014. p. 183–92.

[99] Bruckbauer ST, Kvitko BH, Karkhoff-Schweizer RR, Schweizer HP. Tn*5*/7-*lux*: a versatile tool for the identification and capture of promoters in gram-negative bacteria. BMC Microbiol 2015;15(1):17. https://doi.org/10.1186/s12866-015-0354-3.

[100] Suzuki N, Inui M, Yukawa H. High-throughput transposon mutagenesis of *Corynebacterium glutamicum*. Methods Mol Biol 2011;765:409–17. https://doi.org/10.1007/978-1-61779-197-0_24.