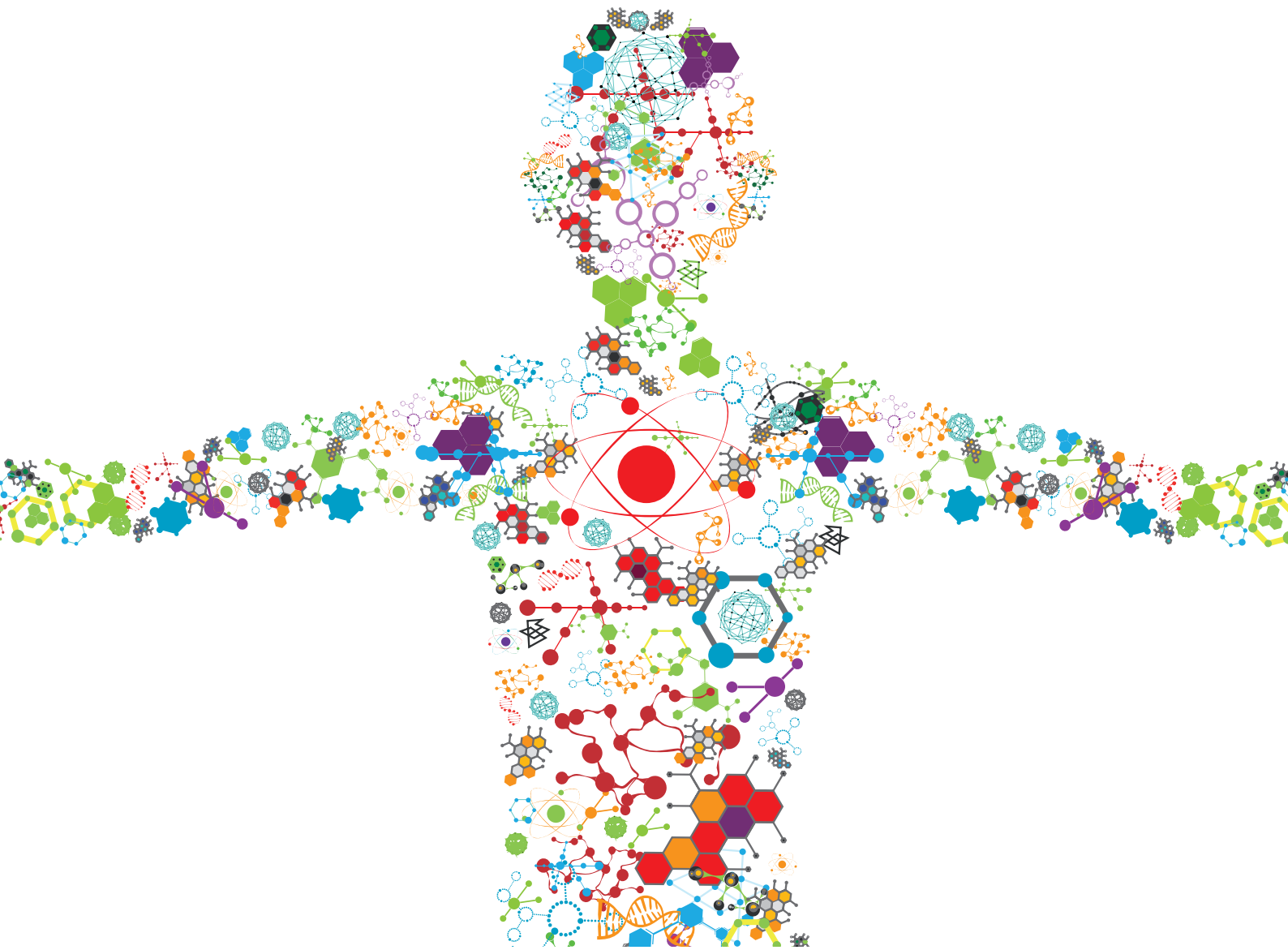
DEVELOPMENT AND APPLICATION OF NOVEL GENOME ENGINEERING TOOLS IN MICROBIAL BIOTECHNOLOGY

EDITED BY: Jiazhang Lian, Yi Wang, Yunzi Luo and Chun Li

PUBLISHED IN: Frontiers in Bioengineering and Biotechnology



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88966-405-4
DOI 10.3389/978-2-88966-405-4

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

DEVELOPMENT AND APPLICATION OF NOVEL GENOME ENGINEERING TOOLS IN MICROBIAL BIOTECHNOLOGY

Topic Editors:
Jiazhang Lian, Zhejiang University, China
Yi Wang, Auburn University, United States
Yunzi Luo, Tianjin University, China
Chun Li, Tsinghua University, China

Citation: Lian, J., Wang, Y., Luo, Y., Li, C., eds. (2021). Development and Application of Novel Genome Engineering Tools in Microbial Biotechnology. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-405-4

Table of Contents

frontiers in Bioengineering and Biotechnology

EDITORIAL
published: 07 December 2020
doi: 10.3389/fbioe.2020.621851

Editorial: Development and Application of Novel Genome Engineering Tools in Microbial Biotechnology

Jiazhang Lian[1]*, *Yi Wang*[2]*, *Yunzi Luo*[3]* and *Chun Li*[4]*

[1] Key Laboratory of Biomass Chemical Engineering (Education Ministry), College of Chemical and Biological Engineering, Zhejiang University, Hangzhou, China, [2] Biosystems Engineering Department, Auburn University, Auburn, AL, United States, [3] Frontier Science Center for Synthetic Biology and Key Laboratory of Systems Bioengineering (Ministry of Education), Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), School of Chemical Engineering and Technology, Tianjin University, Tianjin, China, [4] Department of Chemical Engineering, Tsinghua University, Beijing, China

Keywords: CRISPR-Cas, genome engineering, microbial cell factories, microbial biotechnology, metabolic engineering, synthetic biology

Editorial on the Research Topic

Development and Application of Novel Genome Engineering Tools in Microbial Biotechnology

Over the past few years, novel genome engineering tools, especially the CRISPR-Cas system, have emerged and revolutionized our ability to modify microorganisms for both fundamental studies and biotechnological applications. More specifically, recent advances in genome engineering tools have enabled the assembly of multiple and/or large DNA fragments with high efficiency and fidelity, the optimization of biosynthetic pathways in combination and a high throughput manner, and the evolution of microbial cells at an unprecedented speed and scale. Consequently, microorganisms have been extensively engineered to enable us to elucidate microbial cellular machinery and produce renewable fuels and chemicals at high titer, yield, and productivity, or to discover novel natural products as potential drugs. This Research Topic aims to provide an overview of the current status, recent progress, challenges, and future perspectives in the field of microbial genome engineering. It brings together a collection of contributions including original research articles, short communications, reviews, and mini-reviews, from communities involved with genome engineering, metabolic engineering, and microbial biotechnology.

As a revolutionary genome engineering tool, the CRISPR-Cas system has been widely adopted in nearly all kingdoms of life. CRISPR-based multiplex genome engineering tools have been well-established for an increasing list of microorganisms for biotechnological applications. Ding et al. from Beijing University of Chemical Technology provided a comprehensive overview of various CRISPR-Cas tools for microbial genome engineering, including gene activation, gene interference, orthogonal CRISPR systems, and precise single base editing, followed by recent applications in metabolic engineering toward the production of chemicals and natural compounds. Wang K. et al. from Zhejiang University constructed a genetic cassette with triple controls of Cas9 activities at transcriptional, translational, and protein levels, together with the over-expression of the ATP synthase β-subunit AtpD, to achieve efficient genome editing in *Streptomyces*. Zhang J. et al. from Sichuan University developed three efficient CRISPR-FnCas12a systems for multiplex genome editing in several *Streptomyces* strains, among which the CRISPR-FnCas12a1 system was used to efficiently edit the industrial strain *Streptomyces hygroscopicus*, the CRISPR-FnCas12a2 system enabled the deletion of large fragments ranging from 21.4 to 128 kb, and the

OPEN ACCESS

Edited and reviewed by:
Jean Marie François, Institut Biotechnologique de Toulouse (INSA), France

***Correspondence:**
Jiazhang Lian
jzlian@zju.edu.cn
Yi Wang
yiwang3@auburn.edu
Yunzi Luo
yunzi.luo@tju.edu.cn
Chun Li
lichun@tsinghua.edu.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *27 October 2020*
Accepted: *18 November 2020*
Published: *07 December 2020*

Citation:
Lian J, Wang Y, Luo Y and Li C (2020) Editorial: Development and Application of Novel Genome Engineering Tools in Microbial Biotechnology. Front. Bioeng. Biotechnol. 8:621851. doi: 10.3389/fbioe.2020.621851

CRISPR-FnCas12a3 system was employed to successfully recognize a broad spectrum of PAM sequences. Hao et al. from Jiangnan University designed a CRISPR-Cpf1-based toolkit for efficient genome editing in *Bacillus subtilis*, including the flexible deletion of a single gene, multiple genes, a large gene cluster, or gene knock-in. Hu et al. from China Agricultural University report the first application of CRISPR-Cas9 in genome editing of the apicomplexan parasite *Eimeria tenella* for the elucidation of gene functions at the single-gene level as well as for the systematic functional analysis of an entire gene family. Zhou et al. from Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences reported the establishment of a CRISPR-Cas12a-assisted genome editing system in *Amycolatopsis mediterranei*, an industrial producer of rifamycin SV. Kwon et al. from Gyeongsang National University review recent advancements in toolbox development for genome and transcriptome engineering in solventogenic *Clostridium*. Cen et al. from Zhejiang University of Technology reviewed tools in genetic engineering for *Fusarium fujikuroi*, a host for the industrial production of gibberellins. Sun J. et al. from Zhejiang University reported an efficient cytosine base editing system in *Pseudomonas putida*, as well as several other *Pseudomonas* species, without the need for DNA strand breakage and donor DNA templates. Besides genome editing, Zhang K. et al. from Tianjin Institute of Industrial Biotechnology at the Chinese Academy of Sciences established a CRISPR/Cas13d-mediated RNA knockdown platform for *Escherichia coli* and *Corynebacterium glutamicum*.

Of equal importance to the development of genome engineering tools are their applications in the establishment of microbial chassis for the over-production of high value-added compounds, the high level expression of recombinant proteins and enzymes, and the discovery of novel natural products. Zeng et al. from Tianjin University and Liu et al. from Zhejiang University engineered *Saccharomyces cerevisiae* for enhanced production of dihydroartemisinic acid and crocetin, respectively. Xin et al. from Dalian University of Technology employed CRISPR-Cas9 genome engineering tools to delete histidine kinase genes in *Clostridium beijerinckii* for enhanced butanol production. Lo et al. from National Renewable Energy Laboratory utilized the CRISPR-Cas9 system to switch *C. ljungdahlii* from acetogenic to ethanologenic metabolism. Sun Y. et al. from Jiangnan University improved the production of prodigiosin by relieving CpxR temperature-sensitive inhibition in *Serratia marcescens*. Yang et al. from Beijing Institute of Technology reviewed the strategies of optimizing the location and adaptation of pathways on the whole-genome scale for plant natural product biosynthesis in microbial cell factories. In addition to value-added small molecules, Wang Q. et al. from Shanghai Jiaotong University and Zhang T. et al. from Beijing Institute of Technology review the genetic engineering tools developed for the engineering of filamentous fungi in the production of recombinant proteins. Köbbing et al. from RWTH Aachen University systematically characterized the context-dependent effects on synthetic promoters. Lin et al. from Beijing University of Chemical Technology reviewed recent achievements in synthetic biology tools for the cloning of intact natural product biosynthetic gene clusters (BGCs) from complex genome sequences for natural product discovery.

Although the CRISPR-Cas system is now a fashionable technology, other genome engineering tools have applications in microbial biotechnology. Chen et al. from Shenzhen Institute of Advanced Technology in the Chinese Academy of Sciences reviewed recent advances in RNAi-assisted strain engineering in *S. cerevisiae*.

This Research Topic presents a valuable collection of articles focusing on recent developments and applications of novel genome engineering tools for both model (i.e., *E. coli, B. subtilis,* and *S. cerevisiae*) and non-model (i.e., *P. putida*, *Streptomyces, Clostridium*, and filamentous fungi) microorganisms. As a revolutionary technology, genome engineering and particularly the CRISPR-Cas system, is expected to play increasingly more essential roles in microbial biotechnology as we address the challenges faced by human society.

AUTHOR CONTRIBUTIONS

JL drafted the manuscript. YW, YL, and CL read and revised the manuscript. All authors approved the manuscript.

FUNDING

This work was supported by the National Key Research and Development Program of China (2018YFA0901800 and 2018YFA0903300), the Natural Science Foundation of China (21736002), and the Alabama Agricultural Experiment Station (USDA-NIFA Hatch project ALA014-1017025).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

BRIEF RESEARCH REPORT
published: 01 November 2019
doi: 10.3389/fbioe.2019.00304

Multi-Layer Controls of Cas9 Activity Coupled With ATP Synthase Over-Expression for Efficient Genome Editing in *Streptomyces*

Kai Wang[1,2†], *Qing-Wei Zhao*[1,3†], *Yi-Fan Liu*[1,2†], *Chen-Fan Sun*[1,2], *Xin-Ai Chen*[1,2], *Richard Burchmore*[4], *Karl Burgess*[5], *Yong-Quan Li*[1,2*] *and Xu-Ming Mao*[1,2*]

[1] *Institute of Pharmaceutical Biotechnology & Research Center for Clinical Pharmacy, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China,* [2] *Zhejiang Provincial Key Laboratory for Microbial Biochemistry and Metabolic Engineering, Hangzhou, China,* [3] *Zhejiang Provincial Key Laboratory for Drug Evaluation and Clinical Research, Hangzhou, China,* [4] *Wolfson Wohl Cancer Research Centre, Institute of Infection, Immunity and Inflammation, University of Glasgow, Glasgow, United Kingdom,* [5] *School of Biological Sciences, Institute of Quantitative Biology, Biochemistry and Biotechnology, University of Edinburgh, Edinburgh, United Kingdom*

Efficient genome editing is a prerequisite of genetic engineering in synthetic biology, which has been recently achieved by the powerful CRISPR/Cas9 system. However, the toxicity of Cas9, due to its abundant intracellular expression, has impeded its extensive applications. Here we constructed a genetic cassette with triple controls of Cas9 activities at transcriptional, translational and protein levels, together with over-expression of the ATP synthase β-subunit AtpD, for the efficient genome editing in *Streptomyces*. By deletion of *actII-ORF4* in *Streptomyces coelicolor* as a model, we found that constitutive expression of *cas9* had about 90% editing efficiency but dramatically reduced transformation efficiency by 900-fold. However, triple controls of Cas9 under non-induction conditions to reduce its activity increased transformation efficiency over 250-fold, and had about 10% editing efficiency if combined with *atpD* overexpression. Overall, our strategy accounts for about 30-fold increased possibility for successful genome editing under the non-induction condition. In addition, about 80% editing efficiency was observed at the *actII-ORF4* locus after simultaneous induction with thiostrepton, theophylline and blue light for Cas9 activity reconstitution. This improved straightforward efficient genome editing was also confirmed in another locus *redD*. Thus, we developed a new strategy for efficient genome editing, and it could be readily and widely adaptable to other *Streptomyces* species to improve genetic manipulation for rapid strain engineering in *Streptomyces* synthetic biology, due to the highly conserved genetic cassettes in this genus.

Keywords: pleiotropic controls, inducible Cas9 activity, homologous recombination, genome editing, *Streptomyces*

OPEN ACCESS

Edited by:
Yunzi Luo,
Sichuan University, China

Reviewed by:
Yinhua Lu,
Shanghai Normal University, China
Cameron Ross MacPherson,
Milieu Intérieur Project,
Institut Pasteur, France

***Correspondence:**
Yong-Quan Li
lyq@zju.edu.cn
Xu-Ming Mao
xmmao@zju.edu.cn

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *17 August 2019*
Accepted: *17 October 2019*
Published: *01 November 2019*

Citation:
Wang K, Zhao Q-W, Liu Y-F, Sun C-F, Chen X-A, Burchmore R, Burgess K, Li Y-Q and Mao X-M (2019) Multi-Layer Controls of Cas9 Activity Coupled With ATP Synthase Over-Expression for Efficient Genome Editing in Streptomyces. Front. Bioeng. Biotechnol. 7:304. doi: 10.3389/fbioe.2019.00304

INTRODUCTION

Given the high efficiency and precise targeting, CRISPR/Cas9 is regarded as the most powerful genome editing toolkit for dissection of molecular mechanisms and control of gene expression (Xu and Qi, 2018). Particularly, it has exhibited great potential in synthetic biology for cell reprogramming, protein engineering and circuitry design (Fellmann et al., 2017; Ho and Chen, 2017), and also for microbial metabolic engineering and manipulation of genetically intractable microorganisms (Cho et al., 2018b; Shapiro et al., 2018).

However, the off-target effects and toxic activity of Cas9 have been observed, due to excessive production of Cas9, along with non-specific guiding by sgRNA, for non-specific binding and cleavage of non-target DNA (Reisch and Prather, 2015; Cui and Bikard, 2016; Cho et al., 2018a). Meanwhile, many efforts have been taken to reduce Cas9 toxicity by controlling its endogenous nuclease activity (Richter et al., 2017). These include transient delivery of Cas9/sgRNA ribonucleoprotein to avoid over-accumulation of intracellular Cas9 during cell propagation (DeWitt et al., 2017), and chemically-inducible expression of Cas9 to reduce off-target cleavages and facilitate precise genome editing (Cao et al., 2016). Moreover, post-translational control of Cas9 by fusion of Cas9 to ERT2 (Liu et al., 2016) or Cas9 splits to FKBP and FRB (Zetsche et al., 2015), respectively, can sequester Cas9 in the cytoplasm, and trans-localization of Cas9 into nuclei for genome editing is efficiently induced by 4-hydroxytamoxifen (Liu et al., 2016) or rapamycin (Zetsche et al., 2015).

The visible consequence of Cas9 toxicity is the drastic reduction of transformant number during introduction of genetic elements into cells, as already observed in *Escherichia coli* (Cui and Bikard, 2016), yeast (DiCarlo et al., 2013), and *Streptomyces* (Tong et al., 2015; Zeng et al., 2015; Cao et al., 2016), and most often results in failure of genetic manipulations. *Streptomyces* is a genus of soil-dwelling filamentous bacteria rich in gene clusters and production of clinically used pharmaceuticals (Kieser et al., 2000), and synthetic biology has been highly developed for their precious natural products (Palazzotto et al., 2019). However, this formidable obstacle will severely impede efficient genetic development of the rich reserves of natural products, particularly in some newly isolated species. Though CRISPR/Cas9 toolkits with *cas9* expressed under a thiostrepton-inducible promoter *tipAp* (Huang et al., 2015; Cao et al., 2016) or constitutive promoters (Cobb et al., 2015; Zeng et al., 2015) have been reported in this genus, the genome editing system should be improved for more efficiency.

Recently in *Streptomyces*, a theophylline-dependent riboswitch has been developed (Horbal and Luzhetskyy, 2016). An 85-bp riboswitch can reduce downstream protein expression to an extremely low level, but several hundred-fold induction was observed when the inducer is added, suggesting that it is an ideal tool for the translational control. Moreover, another version of split Cas9 system—blue light inducible Cas9 reconstitution—has been developed. The fungal photo-receptor Vivid is engineered into two proteins, positive Magnet (pMag) and negative Magnet (nMag), which hetero-dimerize upon exposure to blue light and have switch-off kinetics by four orders of magnitude (Kawano et al., 2015). This photo-switchable system allows development of an photo-activatable split Cas9 (N713-pMag and nMag-C714) for the precise controls of Cas9 activity for gene editing (Nihongaki et al., 2015).

DSB (double-strand breakage) caused by Cas9 should be precisely repaired through homology directed recombination (HDR), which is exerted by the ATP-dependent DNA repair system. ATP plays important and essential roles in HDR, as it is associated with recombinase (RecA/Rad51) filaments on DNA, and its hydrolysis is essential for the dynamic interaction between RecA-ssDNA (Reymer et al., 2015), RecA sliding along DNA (Kim et al., 2017), and conformational change during Rad51 filament disassembly (Brouwer et al., 2018).

Here we constructed a genetic cassette for efficient genome editing in *Streptomyces* by pleiotropic controls of Cas9 activity at transcription, translation and protein levels, combined with overexpression of ATP synthase β-subunit AtpD. Using *Streptomyces coelicolor* as a model, this strategy has dramatically enhanced the transformation efficiency by about 250-fold and increased nearly 30-fold probability to obtain mutants over the traditional CRISPR/Cas9 system under the non-induction condition. Moreover, about 80% deletion rate was observed at the *actII-ORF4* locus in the unedited transformants when Cas9 activity was induced and reconstituted. In addition, this strategy by uncoupling DNA introduction and Cas9-mediated DNA cleavage to improve the chance of genome editing was also verified at the *redD* locus in *S. coelicolor*. Thus, here we provide an efficient genetic tool for rapid genome editing in *S. coelicolor*, and it could be also widely and readily applied in other *Streptomyces* species.

MATERIALS AND METHODS

Strains and Media

Streptomyces coelicolor M145 (Kieser et al., 2000) and a daptomycin producer *Streptomyces roseosporus* SW0702 (Mao et al., 2015) were used for genome editing. *E. coli* DH5α (Merck) was used as a general plasmid cloning host. *E. coli* ET12567/pUZ8002 was used to introduce plasmids into *Streptomyces* by inter-species conjugation. *E. coli* strains were cultured in LB medium. Liquid 3% TSB plus 5 % PEG6000 was used for vegetative growth of *Streptomyces* for conjugation after inoculum of spores. Solid R2YE was used as the medium for *S. coelicolor* to produce actinorhodin (Act) and undecylprodigiosin (Red), antibiotic sensitivity assays (with 50 μg/ml apramycin) and also for colony PCR, while MSF medium was for spore preparation and conjugation (Kieser et al., 2000).

Plasmid Construction

All the primers and plasmids used in this work are listed in **Tables S1**, **S2**, respectively, and the detailed procedures for plasmid construction were described in **Supplementary Materials and Methods**. Chemically synthesized DNA sequences of pMag and nMag were shown in **Figures S28**, **S29**, respectively.

Colony Counting

For plates with <100 transformants, the total colonies were counted directly. For conjugations expected with over 100 transformants according to the preliminary experiments, the bacterial mixtures were plated in 1: 15 or 1: 3 dilution in triplicate. Each plate was divided into eight equal parts, and only one representative part was counted. The total transformants were estimate based on counting from each plate, and shown as the mean ± SD (standard deviation) from three independent transformations.

Genotype Validation by Colony PCR

The transformants on the MSF medium were patched on the R5 medium containing 50 μg/ml apramycin, and further cultured at 30°C for 3 days. The genotypes of transformants were verified by colony PCR with 2 × Taq Plus Master Mix (Dye Plus) (Vazyme), as very little mycelia were transferred to the PCR reactions, and lysed by heating at 95°C for 5 min before PCR. Primers 46 + 47 and 48 + 49 were used for confirmative PCR of *actII-ORF4* deletion, while primers 50 + 51 and 52 + 53 for confirmation of *redD* deletion.

Induction of the Reconstituted Cas9

The spores of *S. coelicolor* strains containing the inducible Cas9 cassette and overexpressed *atpD* were inoculated in TSB + PEG6000, and cultured in a cell-tissue bottle for more transmissibility and proximity of blue light. Cas9 activity was induced with the combination of inducers, including 5 μg/ml thiostrepton, 4 mM theophylline and blue light irradiation (470 ± 20 nm LED light source) with a power of 6.4 W and distance of 10 cm. After induction for 5 days, the mycelia were streaked on R2YE medium for genotype verification by colony PCR.

RESULTS

Design of an Inducible Split-Cas9 Genetic Cassette

Constitutively over-expressed *cas9* is proposed to be toxic to many microbial species in genome editing, as shown by the drastic reduction in the number of transformants (DiCarlo et al., 2013; Zeng et al., 2015; Cao et al., 2016; Cui and Bikard, 2016). To overcome this problem, we set out to control intracellular Cas9 activity at transcriptional, translational and protein levels. Our hypothesis is that the reduced Cas9 activity should be beneficial for cell viability during transformation, and subsequently efficient genome editing would occur after Cas9 activity is induced for DSB coupled with HDR.

For this purpose, we developed an inducible genetic cassette to control the Cas9 activity in *Streptomyces*. The codon optimized Cas9 (from *Streptococcus pyogenes*) in the plasmid pWHU2653 was put under the control of an inducible promoter *tipAp* instead of the constitutive *aac(3)IV* promoter [*aac(3)IVp*] (Huang et al., 2015; Zeng et al., 2015). Subsequently, a theophylline-inducible riboswitch was inserted after *tipAp* for the translational control (Rudolph et al., 2015; Horbal and Luzhetskyy, 2016). In addition, Cas9 was split expressed as N713-pMag and nMag-C714 fusions to control Cas9 activity at the protein level (Nihongaki et al., 2015). Moreover, since HDR is highly dependent on energy, so *atpD*, encoding the β-subunit of ATP synthase from *S. coelicolor* (SCO5373), which is the catalytic subunit for ATP synthesis, was expressed under a strong constitutive promoter *ermEp** (**Figure 1**).

Genome Editing Under Non-induction Conditions

To test our above hypothesis, we attempted to delete *actII-ORF4* in *S. coelicolor* M145 as a proof of concept, since this gene has been shown to have 100% editing efficiency if *cas9* is expressed from *tipAp*. Deletion of *actII-ORF4* disrupted the production of the blue pigment actinorhodin (Act) (Huang et al., 2015). The spacer sequence and homologous arms were the same as described previously (**Figure 2A**; Huang et al., 2015). High transformation efficiency was observed without Cas9 (>21,000 transformants), but only about 24 or 30 transformants appeared if *cas9* was expressed under *aac(3)IVp* or *tipAp* (**Figure 2B**, **Figure S1**). However, further introduction of riboswitch (*tipAp-ribo-cas9*) and split Cas9 (*TM-cas9*) under *tipAp* significantly enhanced transformants to about 530 and 5000, which was about 18 (528/30) and 168 (5032/30) -fold higher than the control *tipAp-cas9*, respectively. Moreover, about 6400 transformants were obtained if riboswitch and split fusion expression were combined to control Cas9 activity (*TRM-cas9*) (**Figure 2B**, **Figure S1**). These results suggested that reduction of Cas9 activity at translational and protein levels could significantly improve transformation.

Diagnostic PCR data showed that no editing at *actII-ORF4* was observed if Cas9 was not included (**Figure S2**). High efficiency (about 90%) of *actII-ORF4* deletion was found when *cas9* was expressed from *aac(3)IVp* or *tipAp* (**Figure 2C**, **Figure S3**), though very few transformants were observed (**Figure 2B**). However, if *cas9* was controlled under both riboswitch and split proteins (*TRM-cas9*), no successful deletion was observed. Instead, most transformants showed single cross-over integration of the plasmid into the genome (**Figure S4**), though high transformation efficiency was observed.

Our above data suggested that reduced Cas9 toxicity by pleiotropic controls of Cas9 activity significantly improved cell viability during transformation, but no genome editing was observed. However, leakage effects have been reported both from riboswitch and split Cas9 under non-induction conditions (Kawano et al., 2015; Nihongaki et al., 2015; Rudolph et al., 2015; Horbal and Luzhetskyy, 2016), suggesting that Cas9 could have residual activity *in vivo*, thus causing DSB. We speculated that genome editing to some extents might occur if we could enhance the efficiency of HDR after Cas9-caused DSB.

The bacterial recombinase RecA for HDR is highly conserved among *Streptomyces* species (**Figures S5**, **S6**). It is anticipated that increased RecA concentration would promote recombination after DSB to support microbial survival (Prentiss et al., 2015). However, our data showed that over-expression of *recA* from *S. coelicolor* (SCO5769) dramatically reduced transformation efficiency in about 60% (6,400–2,700 transformants) (**Figure S7**), and very few plasmid integrations (only 1/19 cross-over) were observed (**Figure S8**). These data suggested that RecA has a negative role in DSB repair after Cas9 cleavage in *Streptomyces*. However, over-expression of

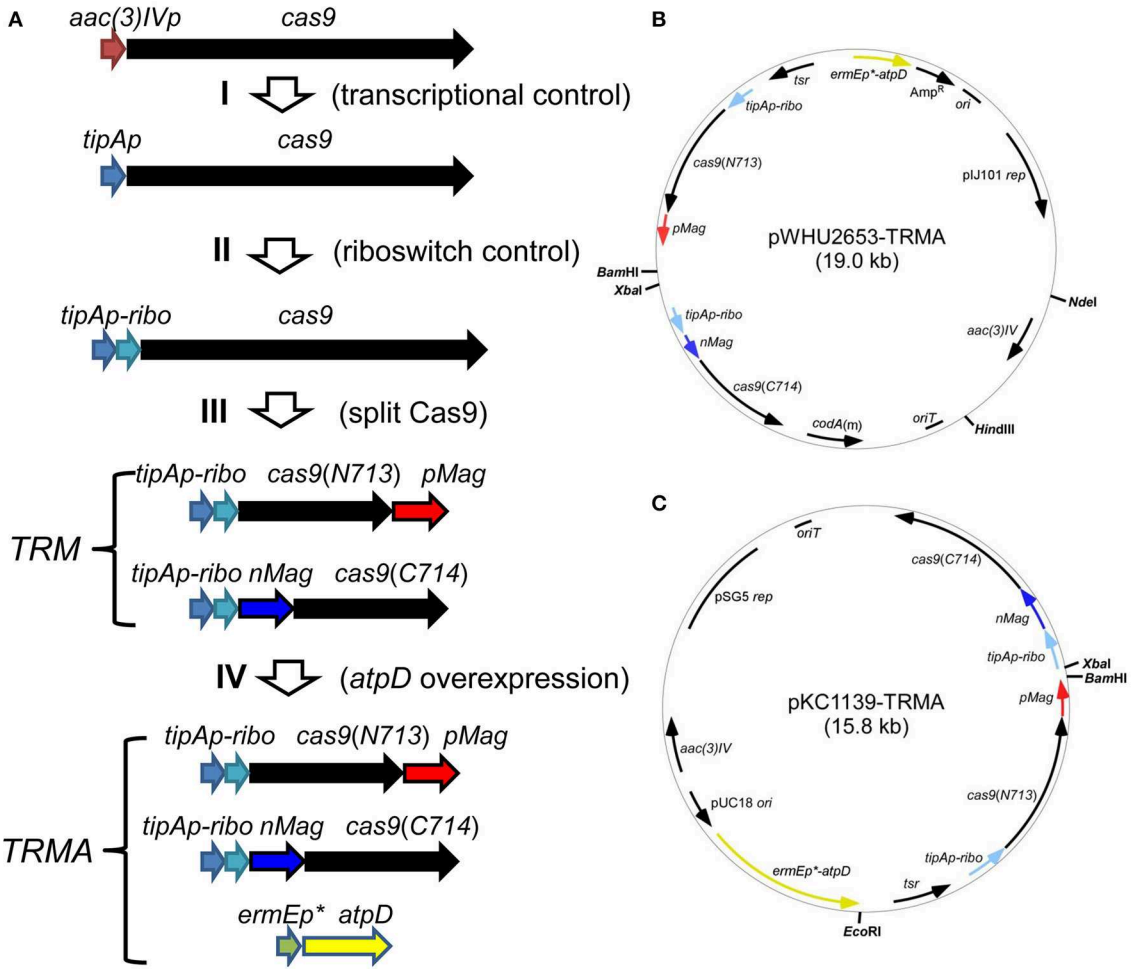


FIGURE 1 | Engineered controls of Cas9 activities and ATP supply for genome editing. **(A)** Strategy for the stepwise controls of Cas9 under the expression of *tipAp*, riboswitch and blue light for the pleiotropic controls at transcriptional (I), translational (II), and protein (III) levels. The ATP synthase gene *atpD* from *S. coelicolor* was additionally overexpressed under *ermEp** (IV). **(B,C)** Two vectors containing genetic cassettes based on pIJ101 *ori* (from pWHU2653) and pSG5 *ori* (pKC1139) were shown as pWHU2653-TRMA **(B)** and pKC1139-TRMA **(C)**, respectively. All the genetic cassettes were labeled as in **(A)** and multiple-cloning sites are shown in bold.

AtpD, the β-subunit of *Streptomyces* F1F0-type ATP synthase (Capaldi and Aggeler, 2002), did not significantly influence the transformation efficiency (6,400–7,200 transformants) (*TRMA-cas9*) (**Figure 2B**), and interestingly, about 10% genome editing efficiency was observed (2/18) (**Figure 2D**). These data suggested that elevation of *aptD* expression promotes HDR on Cas9-caused DSB. Again, further introduction of RecA reduced over half transformation (**Figure S7**) and diminished genome editing even with over-expressed AtpD (**Figure S9**). Thus, based on our editing percentage from *TRMA-cas9* transformation, it was estimated that over 700 mutants would be obtained, which was over 30 (700/24*90%)-fold higher in chance than the constitutively expressed *cas9* [*aac(3)IVp-cas9*].

To further confirm our hypothesis, another gene *redD* was chosen for editing with this engineered CRISP/Cas9 system (**Figure S10**), together with the same homologous arms and spacer used in a previous study (Huang et al., 2015). Consistent with the results from *actII-ORF4* editing, over 500-fold reduction of transformation efficiency was observed if Cas9 was constitutively expressed from *aac(3)IVp* (48 colonies) or *tipAp* (36 colonies) compared to the control without Cas9 (about 27,000 transformants), while introduction of riboswitch, Mag-based split proteins or both, for Cas9 activity control increased transformation efficiency with about 17-fold (840/48), 100-fold (4860/48), or 115-fold (5520/48) compared to Cas9 expressed under *aac(3)IVp* (**Figure S11**). Diagnostic PCR showed that no genome editing was observed without Cas9 (**Figure S12**), and Cas9 expressed from *aac(3)IVp* conferred <10% (1/12) editing efficiency, though most strains showed integrative forms of the plasmid (**Figure S13**), consistent with the previous report of <30% *redD* editing efficiency (Huang et al., 2015). After Cas9 was controlled under riboswitch and the split form (*TRM-cas9*), most transformants displayed replicative forms of the plasmid (**Figure S14**). However, further over-expression of AtpD

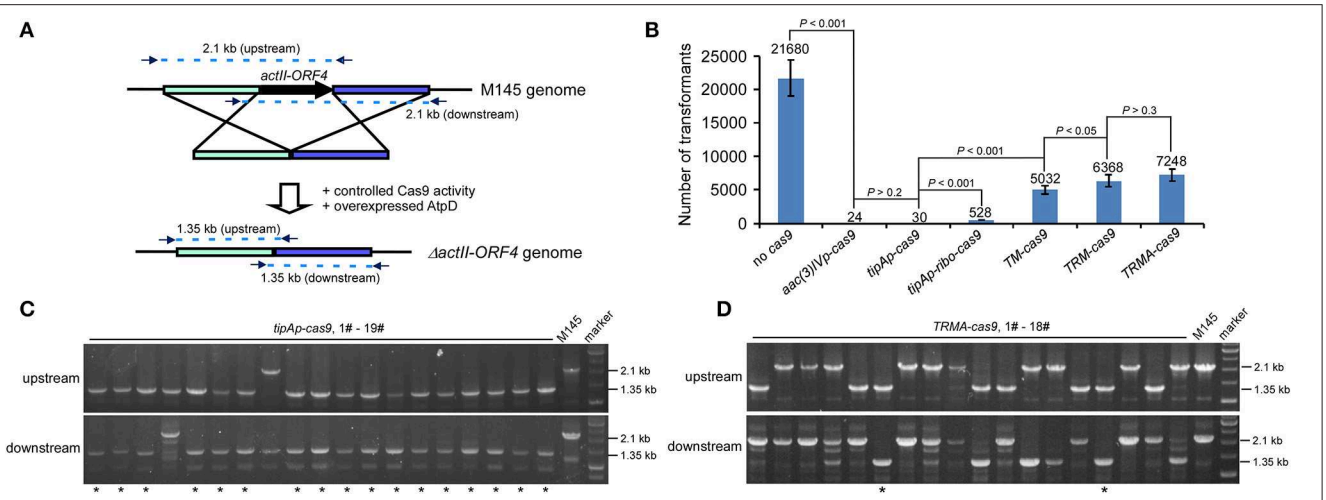


FIGURE 2 | *actII-ORF4* editing in *S. coelicolor* with fine-tuned controls of Cas9 and overexpression of *atpD* on the pIJ101 *ori* plasmid under non-induction conditions. **(A)** Diagram of *actII-ORF4* deletion in *S. coelicolor* based on engineered controls of Cas9. The primers for diagnostic PCR were shown as arrows, and the predicted sizes of PCR fragments in wild type (M145) and the Δ*actII-ORF4* mutant were also shown. The two 2.1 kb fragments are the predicted upstream and downstream PCR products at *actII-ORF4* locus of wild type, while the two 1.35 kb fragments are the corresponding PCR products after *actII-ORF4* is deleted in the mutant. Desired genome editing is verified by simultaneous genotype investigation of the upstream and downstream regions by PCR. **(B)** Transformation efficiency during *actII-ORF4* deletion in *S. coelicolor* with various combinatory controls of Cas9 activities based on the pIJ101 *ori* plasmid. The estimated transformants were shown as mean ± SD (standard deviation) from three independent transformations, and the average transformant numbers were shown on the corresponding columns. *ribo*, theophylline-inducible riboswitch; *TM*, both Cas9 N713-pMag and nMagHigh1-Cas9 C714 split fusions expressed under *tipAp*; *TRM*, both Cas9 N713-pMag and nMagHigh1-Cas9 C714 splits expressed under *tipAp* and riboswitch; *TRMA*, overexpressed *atpD* along with *TRM*. Statistical significance ($P < 0.001$ or $P < 0.05$) was determined by two-tailed unpaired Student's test. **(C,D)** Diagnostic PCR for CRISPR/Cas9-based *actII-ORF4* deletion with Cas9 expressed from *tipAp* **(C)** or with engineered control of Cas9 **(D)** as shown in **Figure 1A**. The successfully edited mutants were indicated with asterisks.

promoted integration of the plasmid and 8.6% (2/23) editing efficiency was obtained (**Figure S15**), which could be used for estimation that about 475 mutants would be obtained. The data from genome editing at the *redD* locus suggested that our strategy had over 100-fold [475/48*(1/12)] higher possibility than the traditional method with constitutively expressed *cas9* to obtain the genetic modified strains (**Figures S11, S15**).

All our above results suggested that reduced intracellular Cas9 activity by pleiotropic controls is significantly beneficial for cell viability during transformation, and resulted in a higher chance to obtain the desired genome-edited mutants if combined with over-expressed AtpD, while the recombinase RecA is detrimental to HDR after Cas9-mediated DSB possibly by prevention of efficient plasmid integration, since the toxicity from over-expressed RecA has been observed in a closely related species *S. lividans* (Vierling et al., 2000).

Engineered CRISPR/Cas9 System on a Temperature-Sensitive Backbone

Here we engineered the CRISPR/Cas9 system on the plasmid pWHU2653, which contains a counter-selection marker *codA* for plasmid loss in the presence of 5-FC (Zeng et al., 2015). However, this selection system is not always suitable in some *Streptomyces* species with high resistance (Dubeau et al., 2009), and we observed that the loss rate of this plasmid was low (about 4%) without the selection pressure of 5-FC (**Figure S16**). Thus, we developed the same CRISPR/Cas9 system in a *Streptomyces* temperature-sensitive but multiple-copy plasmid pKC1139 (Bierman et al., 1992), which will be readily lost over 34°C and therefore advantageous for multiple cycles of genome editing (**Figure 1C**).

We found that this engineered CRISPR/Cas9 system on pKC1139 conferred comparable transformation efficiency as on pWHU2653, with or without over-expressed AtpD (**Figure 3A**, **Figure S17**). Consistent with results shown above, although residual Cas9 activity was still observed, no genome editing was detected even on this multiple-copy plasmid backbone without over-expressed AtpD, as exemplified by deletion of *actII-ORF4* (**Figure S18**) or *redD* (**Figure S19**) in *S. coelicolor*. Estimated 10% (2/20) and 8.3% (1/12) editing efficiency was found on *actII-ORF4* and *redD* loci (**Figure 3B**, **Figure S20**), respectively, when *atpD* was further over-expressed. As expected, the plasmid was lost at a rate of 100% (30/30) at the restricted temperature (37°C) (**Figure S21**).

Efficient Genome Editing Under Induction Conditions

Most transformants remained genetically intact with triple-layer controlled Cas9 (**Figures 2**, **3**). Then we investigated whether the genome could be further edited upon induction of Cas9 activity. We found that the transformants with pKC1139-TRMA-cas9, either in single cross-over or replicative form, were not genetically edited at the *actII-ORF4* locus if only induced with chemicals (Tsr and Theo) or the blue light (**Figures S22**, **S23**). However, when all inducers were included, efficient genome editing (80% = 16/20)

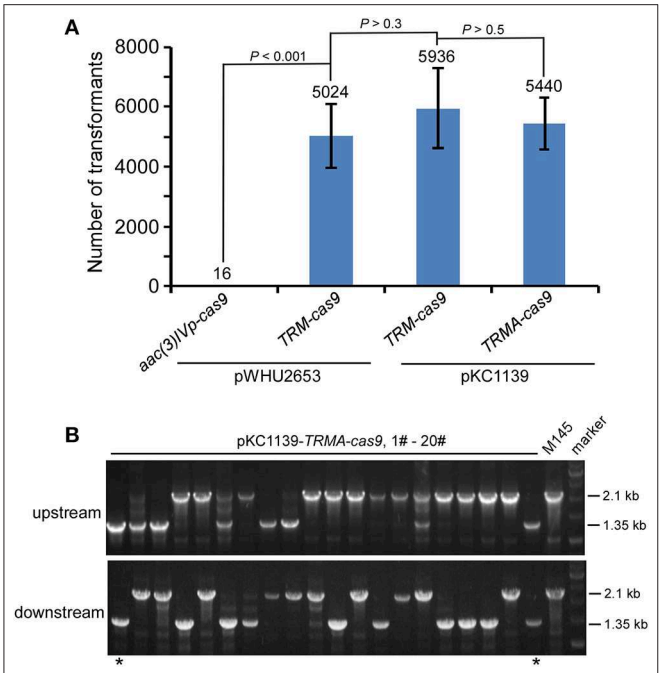


FIGURE 3 | *actII-ORF4* deletion in *S. coelicolor* with engineered Cas9 and *atpD* overexpression from the pSG5 *ori* plasmid under non-induction conditions. **(A)** Transformation efficiency for *actII-ORF4* deletion in *S. coelicolor* with engineered controls of Cas9 activities as in **Figure 2B**, but based on the pSG5 *ori* plasmid. The estimated transformants were shown as mean ± SD (standard deviation) from three independent transformations, and the average transformant numbers were shown on the corresponding columns. Statistical significance ($P < 0.001$) was determined by two-tailed unpaired Student's test. **(B)** Diagnostic PCR for *actII-ORF4* deletion based on the same strategy as **Figure 2D**, but with pSG5 *ori* on the pKC1139 plasmid. The successfully deleted strains were indicated with asterisks.

was observed with both forms (**Figure 4**). Phenotypic analysis of actinhordin (Act) biosynthesis in *S. coelicolor* also confirmed that all the *ΔactII-ORF4* mutants failed to produce this blue pigment, regardless of plasmid backbones and induction/non-induction conditions (**Figure S24**), suggesting that they were all successfully edited. The similar results were obtained on the *redD* locus, as genome editing was only observed after induction with both chemicals and blue light (**Figures S25–S27**), but 35% (7/20) and 40% (8/20) efficiency were only detected from single cross-over and replicative plasmids, respectively (**Figure S27**), which might result from the low efficiency of genome editing at this locus (Huang et al., 2015).

DISCUSSION

Toxicity from the constitutively active Cas9 will dramatically reduce the transformation efficiency, and most often, no transformant at all will be obtained, thus obstructing efficient genetic manipulations for desired genome editing (DiCarlo et al., 2013; Zeng et al., 2015; Cao et al., 2016; Cui and Bikard, 2016). In these cases, introduction of genetic cassettes into cells and subsequent genome editing occur simultaneously at one step. Here we demonstrated an alternative strategy to uncouple transformation and genome editing in two steps, and showed that it can be applied straightforwardly for efficient genome editing in *Streptomyces*.

In this work, a genetic cassette was developed for multi-layer controls of Cas9 activity at transcriptional, translational and protein levels. These include thiostrepton-inducible *tipAp*, the theophylline-inducible riboswitch (Rudolph et al., 2015) and the blue light-inducible Cas9 reconstitution system (Nihongaki et al., 2015) to reduce intracellular Cas9 toxicity. Combination of these two latter cassettes has the best effects with over 250-fold increased transformation. Further over-expression of ATP synthase resulted in about 10% editing efficiency. Though not as high in genome editing as the traditional method with constitutively expressed *cas9* (Huang et al., 2015; Zeng et al., 2015), our new strategy is estimated to enhance about 30-fold of chance to obtain the genetically edited mutants. Moreover, it is encouraging that subsequent reconstitution of Cas9 activity in the unedited transformants showed 40–80% editing efficiency, probably depending on the genomic environments or the 20-bp spacer sequences.

Here in the well-studied model *Streptomyces* species *S. coelicolor*, the transformation efficiency dropped about 900-fold (21680/24 = 903) with constitutively expressed *cas9* (**Figure 2B**), and far <100 transformants could be obtained, though the genome editing efficiency was almost 100%. This similar observation was also reported previously (Zeng et al., 2015). However, in some genetically intractable industrial or newly identified *Streptomyces* strains, the efficient conjugation is not well-established. The low transformation efficiency in these strains will most often result in null transformant with overexpressed Cas9 and significantly delay the study of microorganisms themselves, as well as strain improvement or development. Our strategy by uncoupling of transformation and genome editing can potentially enhance the chance to obtain the transformants, and make it sure to successfully edit the strains subsequently after induction. Though more time might be needed than the traditional Cas9 expression cassette if transformants could be obtained, our strategy will be more efficient than the traditional gene knock-out methods in *Streptomyces* solely depending on homologous recombination, such as pKC1139-based gene deletion (Bierman et al., 1992) and PCR-targeting (Gust et al., 2003). In addition, the theophylline-inducible riboswitch and the blue light-inducible pMag/nMag cassettes could have wide applicability in *Streptomyces*, suggesting that the fine-tuned controls of Cas9 activity will be readily adaptable in other *Streptomyces* species. This idea was further confirmed by our data on a daptomycin producer *Streptomyces roseosporus*, in which a putative daptomycin-producing related gene *dptP* was used as a model (Miao et al., 2005). Constitutive expression of *cas9* again reduced transformation efficiency for over 1000-fold, while introduction of stepwise combination of the regulatory genetic elements (*tipAp*, riboswitch, and Mag proteins) to reduce Cas9 activity could increase

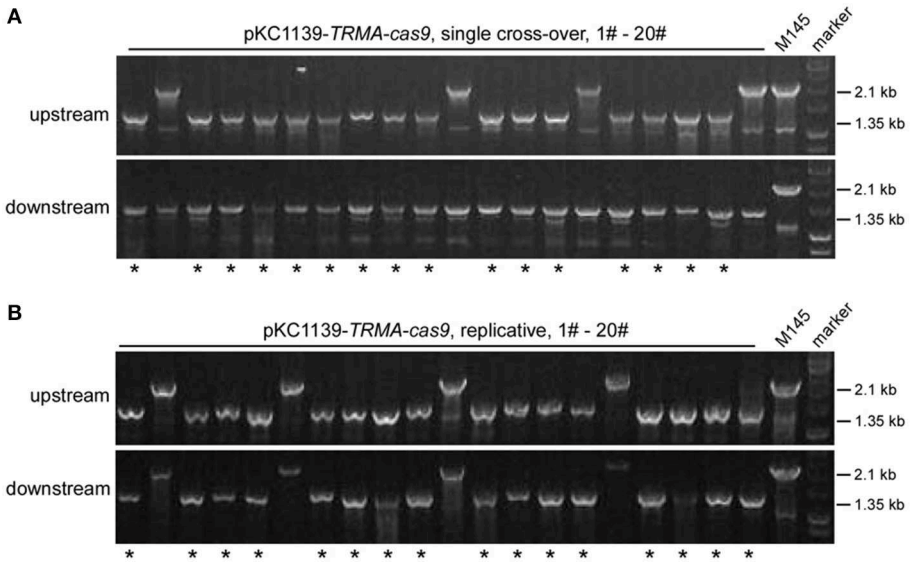


FIGURE 4 | *actII-ORF4* deletion in *S. coelicolor* with engineered Cas9 and *atpD* overexpression from the pSG5 *ori* plasmid under induction conditions. The unedited transformants containing pKC1139-TRMA-cas9 in the single cross-over **(A)** or replicative form **(B)** from **Figure 3B** were simultaneously induced with thiostrepton, theophylline and blue light, and the genotype of twenty single colonies was investigated with diagnostic PCR as in **Figure 2A**. The successfully deleted strains were indicated with asterisks.

transformation efficiency (**Figure S30**). This wide feasibility of this genetic system holds a highly promising application in synthetic biology of *Streptomyces*, and will benefit the microbiologists for efficient genome editing and industry for rapid strain improvement.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

KW, Y-QL, and X-MM have conceived the project. KW, Q-WZ, Y-FL, C-FS, X-AC, Y-QL, and X-MM have designed experiments. KW, Q-WZ, Y-FL, and C-FS have demonstrated the experiments. All authors have analyzed the data, prepared the manuscript, and approved the final version.

FUNDING

This work was financially supported by National Key Research and Development Program (2016YFD0400805), Natural Science Foundation of China (31571284, 11811530638, 3173002, and 31520103901).

ACKNOWLEDGMENTS

We gratefully thank Yuhui Sun in Wuhan University, China, for his kind gift of plasmid pWHU2653, and Mark J. Buttner in John Innes Centre, UK, for the plasmid pIJ8600. These plasmids are critically helpful for experiment controls and cassette construction.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2019.00304/full#supplementary-material

REFERENCES

Bierman, M., Logan, R., Obrien, K., Seno, E. T., Rao, R. N., and Schoner, B. E. (1992). Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* 116, 43–49. doi: 10.1016/0378-1119(92)90627-2

Brouwer, I., Moschetti, T., Candelli, A., Garcin, E. B., Modesti, M., Pellegrini, L., et al. (2018). Two distinct conformational states define the interaction of human RAD51-ATP with single-stranded DNA. *EMBO J.* 37:e98162. doi: 10.15252/embj.201798162

Cao, J., Wu, L., Zhang, S. M., Lu, M., Cheung, W. K., Cai, W., et al. (2016). An easy and efficient inducible CRISPR/Cas9 platform with improved specificity for multiple gene targeting. *Nucleic Acids Res.* 44:e149. doi: 10.1093/nar/gkw660

Capaldi, R. A., and Aggeler, R. (2002). Mechanism of the F(1)F(0)-type ATP synthase, a biological rotary motor. *Trends Biochem. Sci.* 27, 154–160. doi: 10.1016/S0968-0004(01)02051-5

Cho, S., Choe, D., Lee, E., Kim, S. C., Palsson, B., and Cho, B. K. (2018a). High-level dCas9 expression induces abnormal cell morphology in *Escherichia coli*. *ACS Synth. Biol.* 7, 1085–1094. doi: 10.1021/acssynbio.7b00462

Cho, S., Shin, J., and Cho, B. K. (2018b). Applications of CRISPR/Cas system to bacterial metabolic engineering. *Int. J. Mol. Sci.* 19:E1089. doi: 10.3390/ijms19041089

Cobb, R. E., Wang, Y., and Zhao, H. (2015). High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS Synth. Biol.* 4, 723–728. doi: 10.1021/sb500351f

Cui, L., and Bikard, D. (2016). Consequences of Cas9 cleavage in the chromosome of *Escherichia coli*. *Nucleic Acids Res.* 44, 4243–4251. doi: 10.1093/nar/gkw223

DeWitt, M. A., Corn, J. E., and Carroll, D. (2017). Genome editing via delivery of Cas9 ribonucleoprotein. *Methods* 121–122, 9–15. doi: 10.1016/j.ymeth.2017.04.003

DiCarlo, J. E., Norville, J. E., Mali, P., Rios, X., Aach, J., and Church, G. M. (2013). Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* 41, 4336–4343. doi: 10.1093/nar/gkt135

Dubeau, M. P., Ghinet, M. G., Jacques, P. E., Clermont, N., Beaulieu, C., and Brzezinski, R. (2009). Cytosine deaminase as a negative selection marker for gene disruption and replacement in the genus *Streptomyces* and other actinobacteria. *Appl. Environ. Microbiol.* 75, 1211–1214. doi: 10.1128/AEM.02139-08

Fellmann, C., Gowen, B. G., Lin, P. C., Doudna, J. A., and Corn, J. E. (2017). Cornerstones of CRISPR-Cas in drug discovery and therapy. *Nat. Rev. Drug Discov.* 16, 89–100. doi: 10.1038/nrd.2016.238

Gust, B., Challis, G. L., Fowler, K., Kieser, T., and Chater, K. F. (2003). PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1541–1546. doi: 10.1073/pnas.0337542100

Ho, P., and Chen, Y. Y. (2017). Mammalian synthetic biology in the age of genome editing and personalized medicine. *Curr. Opin. Chem. Biol.* 40, 57–64. doi: 10.1016/j.cbpa.2017.06.003

Horbal, L., and Luzhetskyy, A. (2016). Dual control system - a novel scaffolding architecture of an inducible regulatory device for the precise regulation of gene expression. *Metab. Eng.* 37, 11–23. doi: 10.1016/j.ymben.2016.03.008

Huang, H., Zheng, G., Jiang, W., Hu, H., and Lu, Y. (2015). One-step high-efficiency CRISPR/Cas9-mediated genome editing in *Streptomyces*. *Acta Biochim. Biophys. Sin.* 47, 231–243. doi: 10.1093/abbs/gmv007

Kawano, F., Suzuki, H., Furuya, A., and Sato, M. (2015). Engineered pairs of distinct photoswitches for optogenetic control of cellular proteins. *Nat. Commun.* 6:6256. doi: 10.1038/ncomms7256

Kieser, T., Bibb, M. J., Butter, M. J., Chater, K. F., and Hopwood, D. A. (2000). *Practical Streptomyces Genetics*. Norwich: The John Innes Foundation.

Kim, S. H., Ahn, T., Cui, T. J., Chauhan, S., Sung, J., Joo, C., et al. (2017). RecA filament maintains structural integrity using ATP-driven internal dynamics. *Sci. Adv.* 3:e1700676. doi: 10.1126/sciadv.1700676

Liu, K. I., Ramli, M. N., Woo, C. W., Wang, Y., Zhao, T., Zhang, X., et al. (2016). A chemical-inducible CRISPR-Cas9 system for rapid control of genome editing. *Nat. Chem. Biol.* 12, 980–987. doi: 10.1038/nchembio.2179

Mao, X. M., Luo, S., Zhou, R. C., Wang, F., Yu, P., Sun, N., et al. (2015). Transcriptional regulation of the daptomycin gene cluster in *Streptomyces roseosporus* by an autoregulator, AtrA. *J. Biol. Chem.* 290, 7992–8001. doi: 10.1074/jbc.M114.608273

Miao, V., Coeffet-Legal, M. F., Brian, P., Brost, R., Penn, J., Whiting, A., et al. (2005). Daptomycin biosynthesis in *Streptomyces roseosporus*: cloning and analysis of the gene cluster and revision of peptide stereochemistry. *Microbiology* 151(Pt 5), 1507–1523. doi: 10.1099/mic.0.27757-0

Nihongaki, Y., Kawano, F., Nakajima, T., and Sato, M. (2015). Photoactivatable CRISPR-Cas9 for optogenetic genome editing. *Nat. Biotechnol.* 33, 755–760. doi: 10.1038/nbt.3245

Palazzotto, E., Tong, Y., Lee, S. Y., and Weber, T. (2019). Synthetic biology and metabolic engineering of actinomycetes for natural product discovery. *Biotechnol. Adv.* 37:107366. doi: 10.1016/j.biotechadv.2019.03.005

Prentiss, M., Prevost, C., and Danilowicz, C. (2015). Structure/function relationships in RecA protein-mediated homology recognition and strand exchange. *Crit. Rev. Biochem. Mol. Biol.* 50, 453–476. doi: 10.3109/10409238.2015.1092943

Reisch, C. R., and Prather, K. L. J. (2015). The no-SCAR (scarless Cas9 assisted recombineering) system for genome editing in *Escherichia coli*. *Sci. Rep.* 5:15096. doi: 10.1038/srep15096

Reymer, A., Babik, S., Takahashi, M., Norden, B., and Beke-Somfai, T. (2015). ATP hydrolysis in the RecA-DNA filament promotes structural changes at the protein-DNA interface. *Biochemistry* 54, 4579–4582. doi: 10.1021/acs.biochem.5b00614

Richter, F., Fonfara, I., Gelfert, R., Nack, J., Charpentier, E., and Moglich, A. (2017). Switchable Cas9. *Curr. Opin. Biotechnol.* 48, 119–126. doi: 10.1016/j.copbio.2017.03.025

Rudolph, M. M., Vockenhuber, M. P., and Suess, B. (2015). Conditional control of gene expression by synthetic riboswitches in *Streptomyces coelicolor*. *Meth. Enzymol.* 550, 283–299. doi: 10.1016/bs.mie.2014.10.036

Shapiro, R. S., Chavez, A., and Collins, J. J. (2018). CRISPR-based genomic tools for the manipulation of genetically intractable microorganisms. *Nat. Rev. Microbiol.* 16, 333–339. doi: 10.1038/s41579-018-0002-7

Tong, Y., Charusanti, P., Zhang, L., Weber, T., and Lee, S. Y. (2015). CRISPR-Cas9 based engineering of Actinomycetal genomes. *ACS Synth. Biol.* 4, 1020–1029. doi: 10.1021/acssynbio.5b00038

Vierling, S., Weber, T., Wohlleben, W., and Muth, G. (2000). Transcriptional and mutational analyses of the *Streptomyces lividans recX* gene and its interference with RecA activity. *J. Bacteriol.* 182, 4005–4011. doi: 10.1128/JB.182.14.4005-4011.2000

Xu, X., and Qi, L. S. (2018). A CRISPR-dCas toolbox for genetic engineering and synthetic biology. *J. Mol. Biol.* 431, 34–47. doi: 10.1016/j.jmb.2018.06.037

Zeng, H., Wen, S., Xu, W., He, Z., Zhai, G., Liu, Y., et al. (2015). Highly efficient editing of the actinorhodin polyketide chain length factor gene in *Streptomyces coelicolor* M145 using CRISPR/Cas9-CodA(sm) combined system. *Appl. Microbiol. Biotechnol.* 99, 10575–10585. doi: 10.1007/s00253-015-6931-4

Zetsche, B., Volz, S. E., and Zhang, F. (2015). A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nat. Biotechnol.* 33, 139–142. doi: 10.1038/nbt.3149

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 25 February 2020
doi: 10.3389/fbioe.2020.00128

Efficient Single-Gene and Gene Family Editing in the Apicomplexan Parasite *Eimeria tenella* Using CRISPR-Cas9

*Dandan Hu[1], Xinming Tang[1], Choukri Ben Mamoun[2], Chaoyue Wang[1], Si Wang[1], Xiaolong Gu[1], Chunhui Duan[1], Sixin Zhang[1], Jinxia Suo[1], Miner Deng[1], Yonglan Yu[3], Xun Suo[1] and Xianyong Liu[1]**

[1] Key Laboratory of Animal Epidemiology and Zoonosis, Ministry of Agriculture, National Animal Protozoa Laboratory, College of Veterinary Medicine, China Agricultural University, Beijing, China, [2] Department of Internal Medicine and Microbial Pathogenesis, School of Medicine, Yale University, New Haven, CT, United States, [3] Department of Clinical Veterinary Medicine, College of Veterinary Medicine, China Agricultural University, Beijing, China

OPEN ACCESS

Edited by:
Yi Wang,
Auburn University, United States

Reviewed by:
Hamid Dolatshad,
University of Oxford, United Kingdom
Hongyu Han,
Shanghai Veterinary Research Institute (CAS), China

***Correspondence:**
Xianyong Liu
liuxianyong@cau.edu.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *13 November 2019*
Accepted: *10 February 2020*
Published: *25 February 2020*

Citation:
Hu D, Tang X, Ben Mamoun C, Wang C, Wang S, Gu X, Duan C, Zhang S, Suo J, Deng M, Yu Y, Suo X and Liu X (2020) Efficient Single-Gene and Gene Family Editing in the Apicomplexan Parasite Eimeria tenella Using CRISPR-Cas9. Front. Bioeng. Biotechnol. 8:128. doi: 10.3389/fbioe.2020.00128

Eimeria species are pathogenic protozoa with a wide range of hosts and the cause of poultry coccidiosis, which results in huge economic losses to the poultry industry. These parasites encode a genome of ∼8000 genes that control a highly coordinated life cycle of asexual replication and sexual differentiation, transmission, and virulence. However, the function and physiological importance of the large majority of these genes remain unknown mostly due to the lack of tools for systematic analysis of gene functions. Here, we report the first application of CRISPR-Cas9 gene editing technology in *Eimeria tenella* for analysis of gene function at a single gene level as well as for systematic functional analysis of an entire gene family. Using a transgenic line constitutively expressing Cas9, we demonstrated successful and efficient loss of function through non-homologous end joining as well as guided homologous recombination. Application of this approach to the study of the localization of EtGRA9 revealed that the gene encodes a secreted protein whose cellular distribution varied during the life cycle. Systematic disruption of the ApiAp2 transcription factor gene family using this approach revealed that 23 of the 33 factors expressed by this parasite are essential for development and survival in the host. Our data thus establish CRISPR-Cas9 as a powerful technology for gene editing in *Eimeria* and will set the stage for systematic functional analysis of its genome to understand its biology and pathogenesis, and will make it possible to identify and validate new targets for coccidiosis therapy.

Keywords: *Eimeria tenella*, CRISPR-Cas9, genetic engineering, apicomplexa, ApiAp2

INTRODUCTION

Eimeria tenella, the causative agent of the intestinal disease coccidiosis, is an apicomplexan protozoan parasite of critical importance in veterinary medicine, and most notably in poultry. The parasite disrupts the intestinal epithelial cells of young poultry causing hemorrhagic cecal disease and even death (Chapman et al., 2013), which results in major economic losses worldwide

(Blake and Tomley, 2014). The disease is thus considered a major risk to animal health of poultry. *E. tenella* undergoes a complex life cycle in the cecal epithelium of chicken following ingestion of sporulated oocysts. Three cycles of schizogony and a subsequent gametogony occur before unsporulated oocysts are shed with feces and undergo sporulation (Chapman et al., 2013).

Genome sequencing revealed that the life cycle of *E. tenella* is controlled by more than 8000 genes (Reid et al., 2014). While significant progress has been achieved over the past several years to understand the function of some of these genes in *E. tenella* development, differentiation, virulence, and susceptibility to therapy, the large majority of its genes remain inaccessible and their function unknown (Reid et al., 2014; Blake, 2015). Functional analysis in *Eimeria* genes has so far been limited to expression of recombinant proteins and localization of these proteins by indirect fluorescence microscopy. Although the first transfection of *Eimeria* was reported 20 years ago (Kelleher and Tomley, 1998), and several exogenous antigens have been successfully expressed in the parasite for the purpose of generating transgenic lines that could be used as vaccine strains (Pastor-Fernández et al., 2018; Tang et al., 2018a,b), no genetic editing tools that could be used in large-scale and systematic functional analysis have been developed heretofore.

Clustered regularly interspaced short palindromic repeats (CRISPR) is a system for DNA recognition used as a defense mechanism in bacteria and archaea (Rodolphe et al., 2007). In recent years, the type II CRISPR system from *Streptococcus* has been successfully used to introduce double-stranded breaks (DSBs) in the genomic DNA of several protozoan parasites including *Plasmodium* (Mehdi et al., 2014; Wagner et al., 2014), *Toxoplasma* (Shen et al., 2014; Sidik et al., 2014), *Leishmania* (Sollelis et al., 2015; Zhang and Matlashewski, 2015), *Trypanosoma* (Peng et al., 2015; Medeiros et al., 2017), and *Cryptosporidium* (Vinayak et al., 2015). In this system, the targeted genome cleavage is achieved by an RNA–Protein complex consisting of a customizable 20 nucleotide sequence that directs a Cas9 endonuclease to the genome location of interest through RNA–DNA hybridization (Jinek et al., 2012). DSBs introduced by Cas9 are then repaired by the cellular machinery through either non-homologous end-joining pathway (NHEJ), creating insertions or deletions (indels), or homology-directed repair (HDR) if an appropriate DNA template is provided (Haber, 2000). Due to the lack of a continuous *in vitro* culture system (Bussière et al., 2018) for *Eimeria* and a low transfection efficiency in this parasite (Clark et al., 2008; Liu et al., 2008), alternative strategies to co-transfection of two plasmids containing Cas9-gRNA cassettes and donor DNA or to the use of a single but large plasmid constructs are thus critical to a successful implementation of the CRISPR-based gene editing approach in this parasite. Interestingly, studies in mammalian cells and other parasites have demonstrated successful stable expression of the Cas9 protein (Koike-Yusa et al., 2014; Peng et al., 2015; Sidik et al., 2016) as a strategy to enable genome-wide gene loss-of-function screening and to circumvent the difficulty of gene delivery in species with low transfection efficiency.

To develop an efficient gene disruption tool for *Eimeria*, we first constructed a stable *E. tenella* parasite line that expresses the *S. pyogenes* Cas9(SpCas9) endonuclease. Using this line, we achieved efficient target gene disruption by introducing a single vector harboring the U6-gRNA and donor DNA. As a proof of principle, we implemented this approach to tag the putative *E. tenella* dense granule protein 9 (EtGRA9), and monitored its expression throughout the life-cycle of the parasite. To further assess the application of this strategy for large-scale and systematic analysis of the *E. tenella* genome, we applied the CRISPR-Cas9-based technology to disrupt each of the 33 members of the AP2 transcription factors family. To the best of our knowledge, this is the first successful targeted gene disruption in *Eimeria* species. This data will set the stage for large-scale and systematic functional analysis of the *E. tenella* genome to help unravel the physiological function of the ∼8000 genes encoded by *E. tenella* in parasite's development, differentiation, virulence, and susceptibility to antimicrobials.

RESULTS

Generation of a Stable Cas9-Expressing Transgenic *E. tenella* Line

Although the humanized SpCas9 sequence has been used in other apicomplexan parasites including *T. gondii*, our attempts to use it in *E. tenella* was not successful, likely due to weak expression of the transgene (data not shown). Therefore, we optimized the sequence of SpCas9 using the *E. tenella* codon usage preference. The 4203 bp optimized Cas9 (eCas9) nucleotide sequence was synthesized and cloned downstream of the Histone 4 gene promoter to generate the eCas9-NLS-2A-YFP transfection vector (**Figure 1A**). In this construct, eCas9 harbors a C-terminal nuclear targeting sequence and two Flag tags and is fused to the reporter molecule YFP with the two sequences separated by the "self-cleavage" 2A peptide. Immunoblot analysis using a monoclonal anti-Flag antibody showed successful expression of eCas9 in *E. tenella* as a ∼156 kDa protein (**Figure 1B**). To further localize eCas9 in transgenic parasites, *E. tenella* sporozoites were stained with the anti-Flag monoclonal antibody, and examined by confocal microscopy. As shown in **Figure 1C**, the fluorescence signal for eCas9 expression co-localized with the nuclear stain DAPI, whereas the YFP signal was confined to the large and small refractile bodies of the parasite. As expected, the expression of eCas9-YFP was constitutive throughout the life cycle of the parasite (**Figure 2**).

To generate a genetically stable eCas9-expressing *E. tenella* line, single sporocysts were isolated by FACS and inoculated into chicken. Two clones were selected following infection of 35 chicken with single sporocysts. Analysis of these two transgenic clones showed little to no differences in total oocyst output compared to the isogenic wild type strain and overall similar oocyst output curve, and intestine lesion score as the isogenic parental line (**Figures 1E–G**).

To identify the insertion site of the Cas9 cassette in the *E. tenella* genome, total genomic DNA from the Et-HCYA

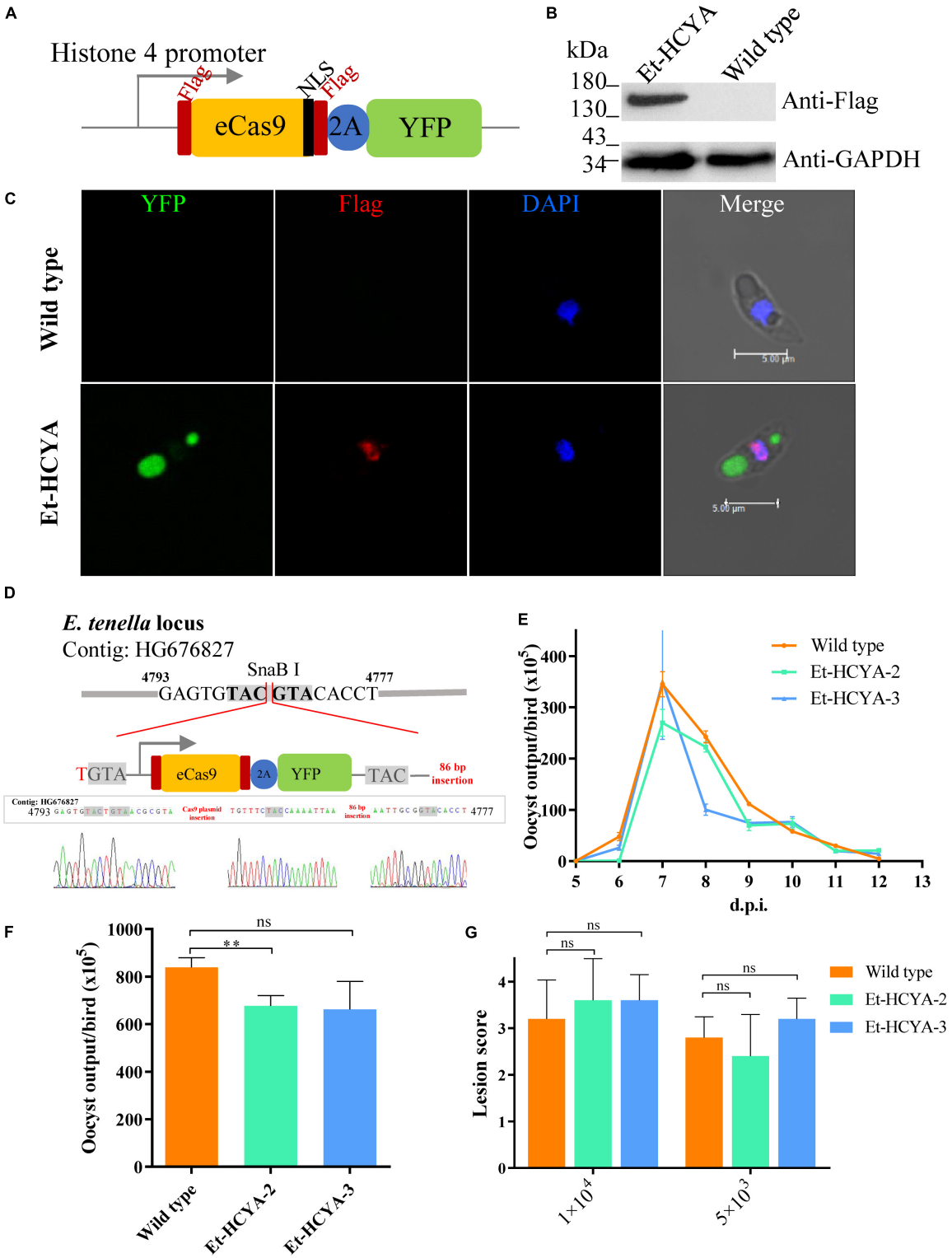


FIGURE 1 | Construction and identification of Cas9-expressing Et-HCYA lines. **(A)** Schematic illustration of plasmid expressing eCas9 and YFP separated by a "self-cleavage" 2A peptide. Stably expression of Cas9 protein in Et-HCYA was identified by Western blot **(B)** and immunofluorescent assay **(C)** using mouse anti-Flag mAb. **(D)** Schematics of the insertion site of Cas9 plasmid. The insertion site was determined by genome-walking and identified by PCR and sequencing. Oocyst output curves **(E)** and total oocyst outputs **(F)** of eCas9-expressing lines. Chicken (n = 4) were infected with 1000 oocysts, and oocyst outputs were detected daily in triplicates. **(G)** Intestine lesion scores after chickens were infected with eCas9-expressing clones. Chickens (n = 5) were infected with 5000/10,000 oocyst/bird, and the lesion scores were scored 7 d.p.i. ns, not significant; $^{**}p < 0.01$.

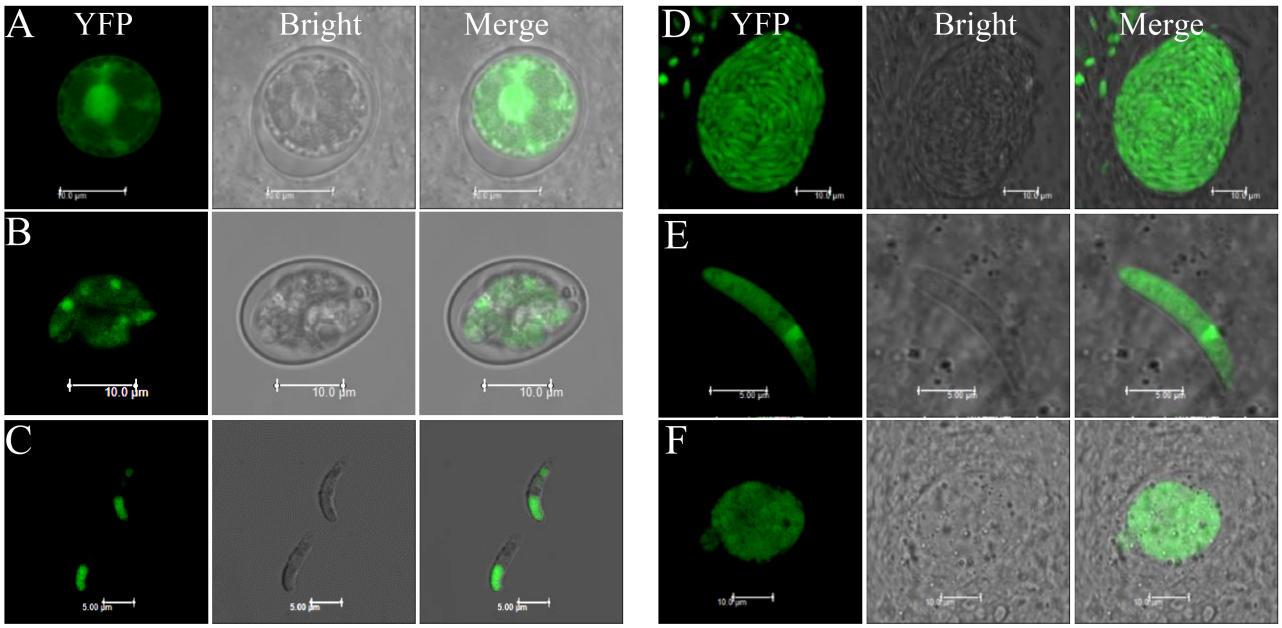


FIGURE 2 | Cas9 protein expression in all stages of EtHCYA. Chicken were infected with 10,000 Et-HCYA clone #3, then the cecal smears were prepared at 96, 120, and 156 h.p.i. for the detection of merozoites, gametocytes, and unsporulated oocysts, respectively. Oocysts were collected and purified from feces. Sporozoites were purified from sporulated oocysts. **(A)** Unsporulated oocyst; **(B)** sporulated oocyst; **(C)** sporozoites; **(D)** schizont; **(E)** merozoite; and **(F)** gametocyte.

transgenic line (clone #3) was extracted and used for genome walking using specific primers (**Supplementary Table S1**). Sequence analysis identified the site of insertion to be in an intergenic region between nucleotides 4777 and 4793 of the *E. tenella* genomic contig HG676827 with one base pair insertion in the N-terminal and 86 bp insertion on the C-terminal region of the inserted plasmid, respectively (**Figure 1D**).

Site-Specific Cas9-Mediated Mutagenesis in *E. tenella*

To examine the efficiency of double strand DNA cleavage activity of eCas9 in eCas9-expressing *Eimeria*, the Et-HCYA line was transfected with gRNA-expressing plasmid targeting the endogenous *yfp* locus from eCas9-NLS-2A-YFP plasmid (**Figure 3A**). The vector harbors the DHFR-TS selectable marker (*T. gondii* bifunctional dihydrofolate reductase-thymidylate synthase mutant which confers resistance to pyrimethamine), the mCherry reporter under the Mic2 promoter, and the YFP guide RNA under the control of the U6 promoter. Following transfection, ∼7% of oocysts analyzed by flow cytometry were found to express mCherry. Of these, ∼2% were negative for YFP fluorescence (**Figures 3B,C**), suggesting a cleavage efficiency of ∼29%. Analysis of the genomic DNA of the transfectants by PCR, to detect the targeted sites repaired by NHEJ, identified multiple bands of different lengths. As expected, the control Et-HCYA line showed a single band (**Figure 3D**). Subsequent cloning and sequencing of all the bands identified in the transfectants revealed four distinct deletions of 4, 8, 522, and 699 bp (**Figures 3E,F**), all occurring near the cleavage target site and the protospacer adjacent motif (PAM, **Figure 3E**).

Cas9-Mediated Gene Disruption in *E. tenella*

To evaluate whether eCas9 can mediate double-stranded gene break and homologous recombination in *E. tenella*, the eCas9-expressing sporozoites were transfected with a plasmid containing the YFP gRNA as well as 5′ and 3′ fragments of homology to sequences in the eCas9-Flag-2A-YFP vector (**Figure 4A**). Because the targeting construct was designed so that the DHFR-mCherry cassette lacks a promoter region, only successful gene replacement by homologous recombination would result in expression of the selectable marker under the control of the Histone 4 promoter that otherwise drives the expression of the eCas9 cassette in the Et-HCYA line (**Figure 4A**). Flow cytometry analysis of the resulting transfectants showed a 67% loss of the YFP signal in oocysts compared to the parental strain and identified ∼4% of oocysts that express mCherry fluorescence only (**Figures 4C,D**). These findings suggest that the eCas9-mediated cleavage occurs with a high efficiency in *E. tenella*, but that the recombination rate is low. Cell sorting of the $mCherry^{+}$ YFP^{-} sporocysts and subsequent passaging produced a homogenous *E. tenella* population of $mCherry^{+}$ YFP^{-} transfectants (Cas9-YFP-KO line) by the third generation as demonstrated by flow cytometry, PCR, and sequencing (**Figure 4B**).

Thirty-two potential off-targets were found in the Et-HCYA genome with maximum mismatch number of five nucleotides using the Cas-OFFinder software (**Supplementary Table S2**). Sequencing of the genome of Et-HCYA and Cas9-YFP-KO line to detect variants identified 138 single nucleotide variations (SNVs) in the Cas9-YFP-KO line. However, none of these mutations were

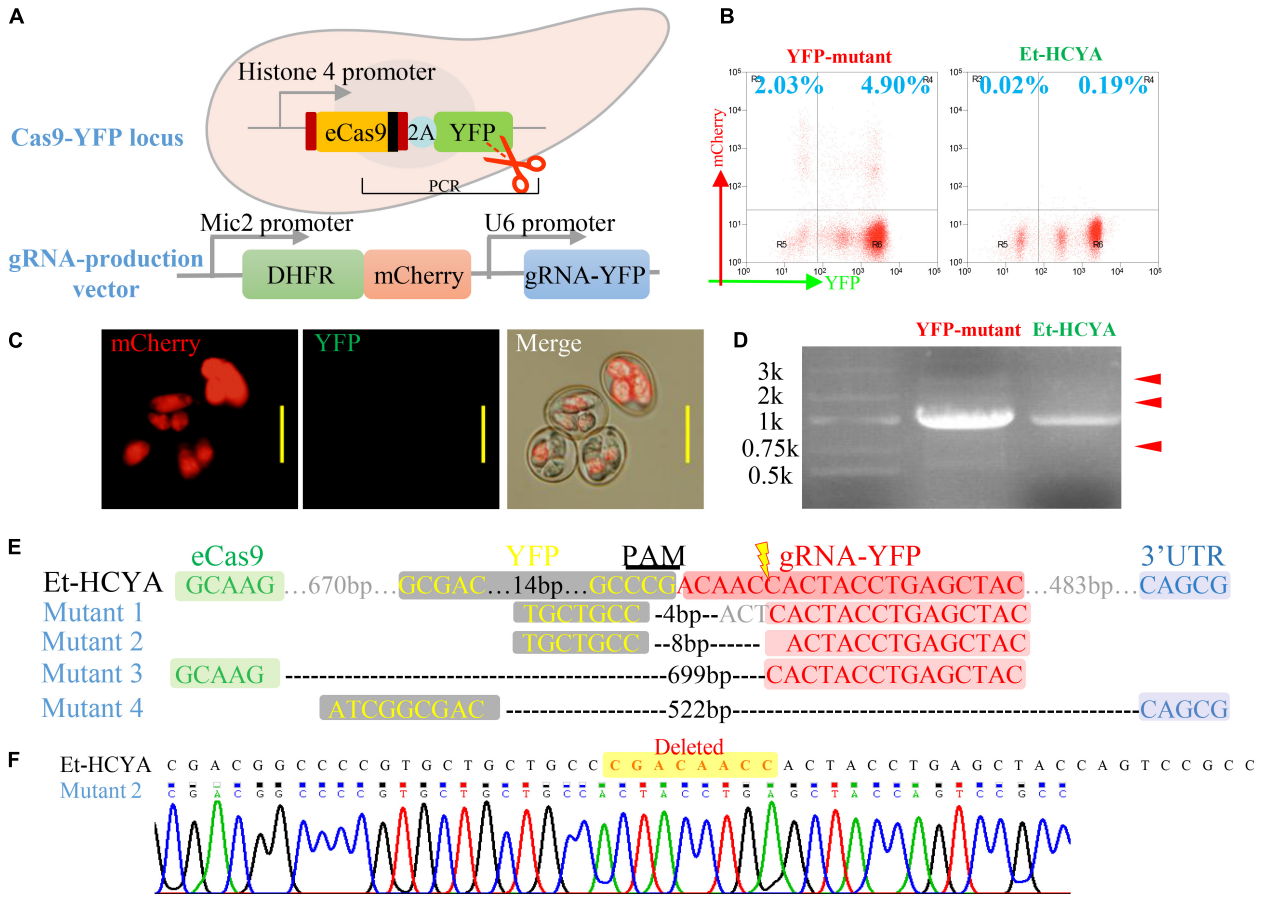


FIGURE 3 | CRISPR/Cas9-mediated YFP mutagenesis in transgenic EtHCYA. **(A)** Schematics showing the mutagenesis strategy by transfecting a vector encoding a gRNA targeting endogenous *yfp* locus in Et-HCYA. **(B)** The mutation rate of the transfected oocysts was analyzed by flow cytometry. Left: Et-HCYA transfected with gRNA-producing vector; Right: Et-HCYA control. **(C)** Fluorescence observation of mutant oocysts under microscope. Bar = 20 μm. **(D)** Cleaved *yfp* loci in the transfected population were identified by PCRs, and all the mutated amplicons (red arrowhead) were cloned and sequenced. **(E)** All detected mutations in the *yfp* loci induced by Cas9 cleavage. **(F)** Chromatogram of mutant 2.

found within the 2 kb region surrounding the predicted off-target sites. Altogether, these data suggest that these SNVs are not the result of Cas9 cleavage but could have emerged from random mutations during the three consecutive passages in chicken.

CRISPR/Cas9-Mediated Gene Tagging of *E. tenella* EtGRA9 Secretory Protein

Studies in *T. gondii* have shown that the dense granule protein TgGRA9 associates with the vacuolar network membranes (Adjogble et al., 2004). To evaluate the expression pattern of the putative *Etgra9* gene in *E. tenella* development, we applied the Cas9-mediated approach. The EtGRA9 protein (ToxoDB ID: ETH_00028350) shares 25% similarity with the *T. gondii* GRA9, contains a 23 aa signal peptide and a putative *N*-palmitoyl cysteine. To monitor the expression profile of EtGRA9 during *E. tenella* life cycle, the C-terminal region of the protein was tagged with mCherry by Cas9-mediated homology recombination in the Et-HCYA line (**Figure 5A**). Cell sorting of sporocysts after subsequent passages produced a GRA9:mCherry line, which was further validated for correct expression of the gene under the endogenous promoter by PCR and sequencing (**Figure 5B**). Fluorescence microscopy showed EtGRA9 to be expressed in all *E. tenella* stages, and to be uniformly distributed in unsporulated oocysts, merozoites, and gametocytes *in vivo* (**Figure 5C**). In sporocysts, however, EtGRA9 was primarily targeted to the Stieda bodies and dotted distributed (**Figures 5D,E**). Following sporozoite excystation, two localization patterns were found. Sporocysts containing the residue body showed localization of EtGRA9 to this organelle, whereas those lacking the residue body showed localization of EtGRA9 to the Stieda body (**Figure 5D**). GRA9:mCherry sporozoites were also cultured in Madin-Darby Bovine Kidney (MDBK) cells to monitor protein localization following sporozoite invasion. EtGRA9 was observed as a single dot in the extra cellular sporozoites, while the protein was dotted distributed in a few more number after 3 h post-invasion. By 24 h post-invasion, the protein could be found in a greater number of dots and the parasitophorous vacuole (PV). EtGRA9 was subsequently distributed in the immature multi-nucleated

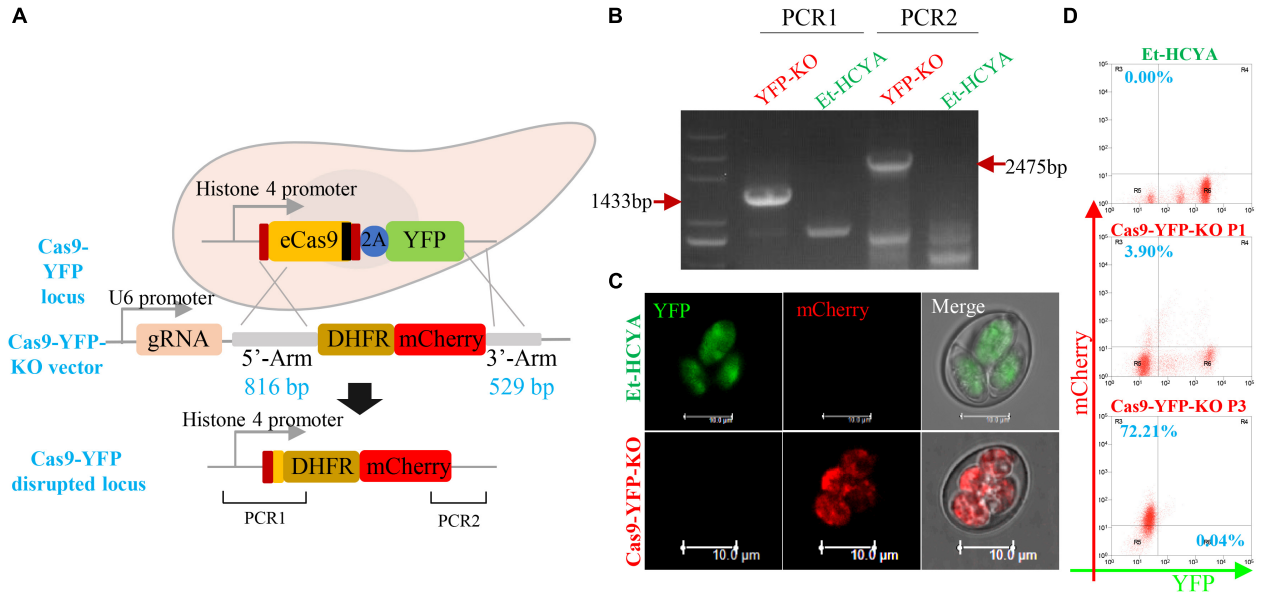


FIGURE 4 | CRISPR/Cas9-directed YFP gene disruption in transgenic EtHCYA. **(A)** Schematics showing CRISPR/Cas9-mediated disruption of endogenous *cas9-yfp* locus by transfecting Cas9-YFP-KO vector. This vector contains a gRNA targeting *yfp* driven by EtU6 promoter, and DHFR:mCherry coding sequence (without promoter) flanked by 5′ and 3′ homologous regions. Endogenous *cas9-yfp* locus disruption was determined by PCR **(B)** and fluorescent microscopy **(C)**. The specific bands in Cas9-YFP-KO parasites were marked with red arrows. **(D)** Cas9 editing efficiency detected by flow cytometry. P1: first generation of progeny oocysts post-transfection; P3: third generation post-transfection.

schizont, and then localized to each merozoite (**Figure 5F**). Following merozoite released and invasion of new cells, EtGRA9 was then redirected to the PV. Immuno-electron microscopy was used to identify the details of these EtGRA9 containing dots in the merozoites. To our surprise, the EtGRA9 labeling signal was not found in the dense granule but in the vesicles (**Figure 5G**). These results suggest that the EtGRA9 is a secreted protein that may help *E. tenella* sporozoites release from sporocysts, and may play an important role during invasion by sporozoites and merozoites. Interestingly, whereas tagging of the endogenous *Etgra9* gene could be easily achieved, no knock-out lines could be isolated using the eCas9-mediated gene disruption, suggesting that the gene is essential for parasite's viability.

Cas9-Mediated Gene Disruption for Functional Analysis of 33 *E. tenella* AP2 Family Genes

Studies in other protozoan parasites showed an important role for AP2 transcription factors in virulence, sexual commitment, and development (Painter et al., 2011; Radke et al., 2013; Kafsack et al., 2014; Sinha et al., 2014). Our analysis of the *Eimeria* genome identified a family of 33 AP2 genes encoding proteins with the AP2 domain. To establish the systematic analysis of this gene family in *E. tenella*, we designed guide RNAs targeting each of the 33 AP2 genes (see **Supplementary Table S3**). These were then cloned under the regulatory control of the U6 promoter in the eCas9 targeting vector and the resulting constructs used to conduct six independent transfections of the eCas9-expressing *E. tenella* line Et-HCYA (**Figure 6A**). Following transfection, viable mCherry^{+} parasites were selected by FACS and analyzed by PCR, high-throughput sequencing, and read mapping to identify the gRNAs present in viable transfectants (**Figures 6B–G**). Interestingly, of the 33 AP2 genes targeted for disruption, only 10 could be successfully targeted as determined by the presence of the gRNAs (**Figure 6H**). Of these, nine gRNAs were readily detectable after eCas9 cleavage. These gRNAs target the following AP2 genes: ETH_00016895, ETH_00031980, ETH_00009105, ETH_00002450, ETH_00009540, ETH_00014800, ETH_00033770, ETH_00028930, and ETH_00042195. The 10th gRNA targeted ETH00036635 and was associated with a low read count, suggesting that disruption of this gene results in a defect in parasite's growth (**Figure 6H**). Together, these data suggest that of the 33 AP2 transcription factors of *E. tenella*, 23 are essential to the parasites, and 10 are dispensable. These findings will set the stage for future detailed phenotypic analyses of viable mutants as well as genetic analyses by conditional disruption of essential genes to assess which stages of *E. tenella* life cycle they control.

DISCUSSION

The lack of genetic tools for functional analysis of *Eimeria* parasites is a major obstacle to the advancement of the biology of these important protozoa and to efforts aimed to control coccidiosis. In this study, we report the first successful application of CRISPR-Cas9-mediated gene editing in *Eimeria* and provide evidence for its use for loss-of-function analysis of the AP2

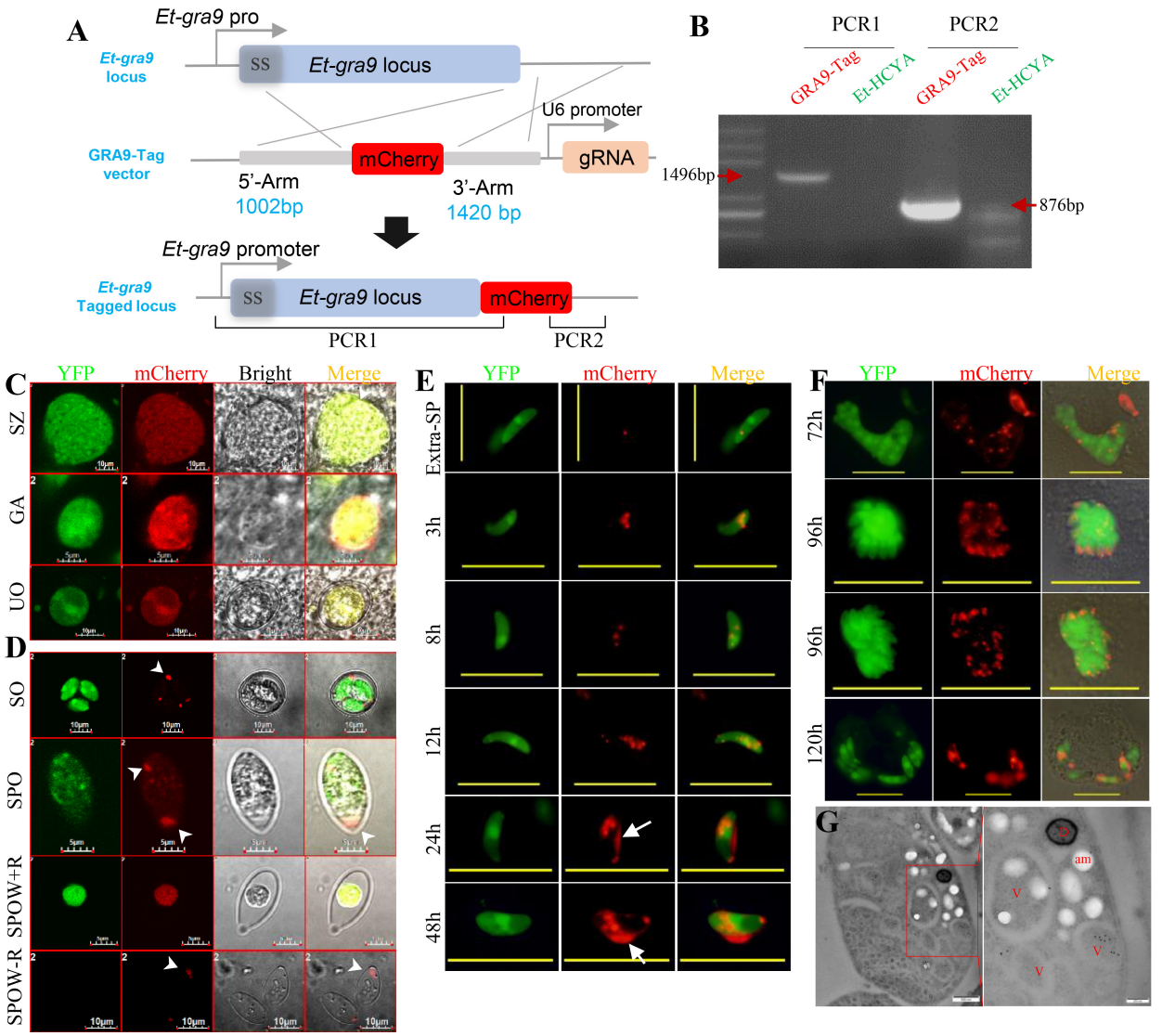


FIGURE 5 | C-terminal tagging of *Etgra9* mediated by CRISPR/Cas9 and the observation of expression pattern of tagged *Etgra9* in parasites. **(A)** Schematics showing CRISPR/Cas9-mediated targeting of *Etgra9* gene. Et-HCYA were transfected with GRA9-Tag vector, which contains a U6-gRNA cassette and a homologous recombination cassette with mCherry fused to the C-terminal of *Etgra9* gene. **(B)** Oocysts with yellow and mCherry fluorescence were sorted by FACS, then genomic DNA was extracted for PCR identification. **(C)** Localization of mCherry-tagged EtGRA9 in different developmental stages *in vivo*. The cecal smears were prepared at 96, 120, and 156 h.p.i. for the detection of merozoites, gametocytes, and unsporulated oocysts, respectively. UO, unsporulated oocyst; SZ, schizont; and GA, gametocyte. **(D)** Localization of mCherry-tagged EtGRA9 during excystation. Sporocyst (SPO) were released from sporulated oocyst (SO) by glass beads grinding, and sporozoites (Extra-SP) were excysted after bile-trypsin digestion, resulting empty sporocyst wall with (SPOW + R)/without (SPOW–R) residue body. Sporozoites were observed at different time points after being inoculated onto MDBK cells **(E,F)**. White arrowheads show EtGRA9 in Stieda bodies and dense granule; white arrows show EtGRA9 secreted into the parasitophorous vacuole. The localization of EtGRA9 was also observed by immune-electron microscopy **(G)**. The ultrathin sections of purified second generation merozoites were labeled with anti-mCherry antibody and followed by immunogold labeling. D, dense granule; V, vesicle; am, amylopectin. Bars in panels **C**, **D**, and **G** are indicated in graph, and bars are 20 μm in panels **E** and **F**.

transcription factor family. We constructed a stable Cas9-expressing *E. tenella* line that undergoes a life cycle that is indistinguishable from that of the wild-type parasite. Using this transgenic line, we found that the cleavage activity of Cas9 occurs with high efficiency. We also used this line to demonstrate successful deletion of the *yfp* gene inserted into the chromosome as well as successful tagging of the EtGRA9 secreted protein. Moreover, we provide evidence that the technology can be applied to target the 33 ApiAp2 transcription factors encoded by *E. tenella* to determine which ones are essential for development and survival of the parasite.

Although our data represent a major advance in *E. tenella* genetic manipulation, further optimizations are needed to streamline the use of this technology for gene editing in this parasite. First, although some progress has been made in the *in vitro* culture of *Eimeria* (Shi et al., 2008; Bussière et al., 2018),

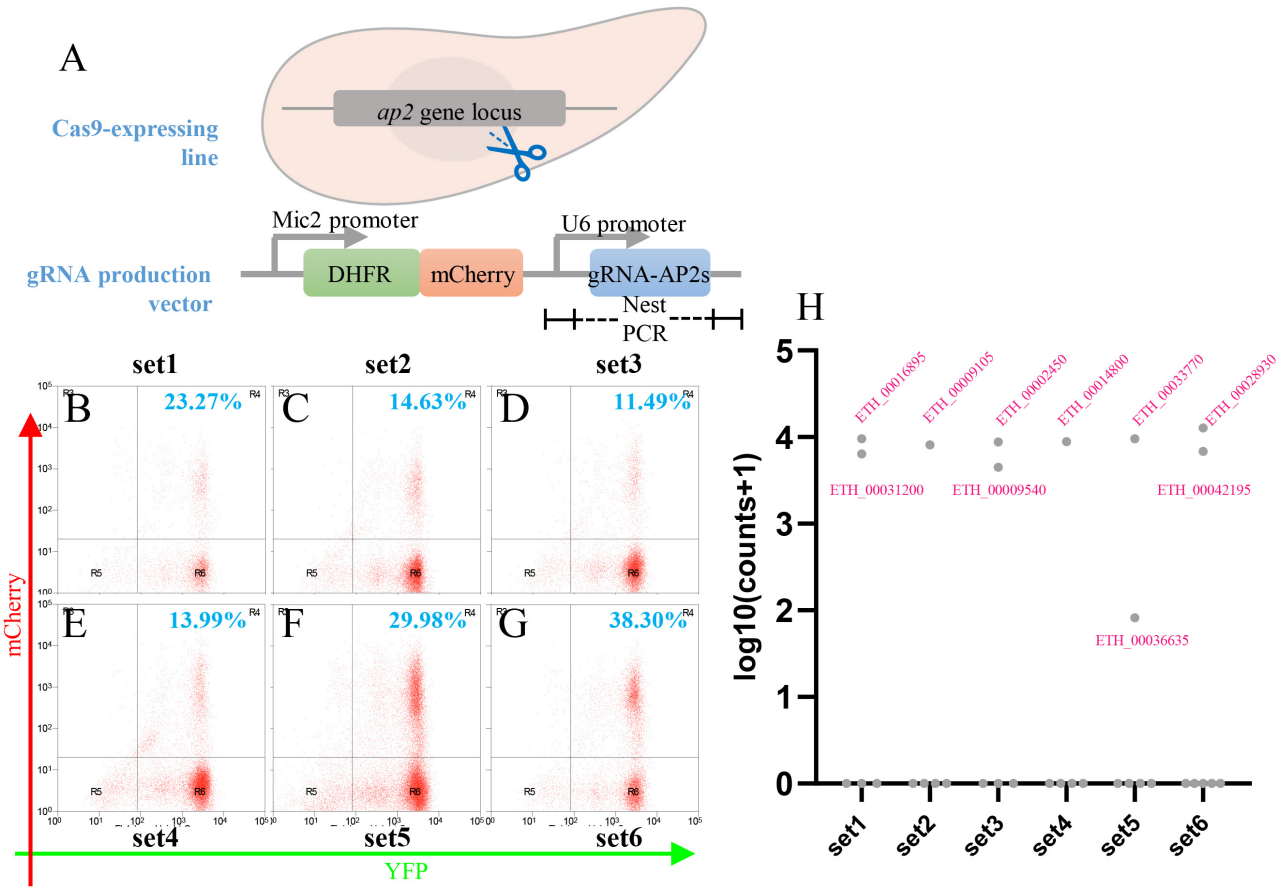


FIGURE 6 | Cas9-mediated *Etap2* family gene disruption. **(A)** Schematics showing the *Etap2* family genes disrupted by transfection with gRNA-producing vectors. **(B–G)** Thirty-three *Etap2* genes were screened by six sets of transfections, and the efficiency of each transfection was informed by flow cytometry analysis. **(H)** Log10 abundances of gRNAs remained in the positive oocysts determined by pooled sequencing of nest-PCR amplicons. Target gene IDs were indicated for their gRNAs presented in the sequencing reads.

the majority of cultured parasites do not complete their life-cycle. Thus, cloning of single parasites requires propagation of transgenic parasites in animals. This strategy, however, is not very efficient as only ∼10% of the clones could be isolated by *in vivo* propagation. Second, unlike trypanosomatids that lack an NHEJ pathway (Burton et al., 2007; Glover et al., 2010), *Eimeria* species repair DSB DNAs mainly by NHEJ pathway because of the conserved KU70-/KU80-dependent mechanism. This problem, however, has recently been solved in *T. gondii* by deleting the NHEJ repair pathway through disruption of the KU80 coding gene (Fox et al., 2009) or by using Cas9-mediated DSB followed by cloning single tachyzoites (Shen et al., 2014; Sidik et al., 2014). Finally, the low transfection efficiency in *Eimeria* (∼0.2%) (Clark et al., 2008; Liu et al., 2008) significantly limits the use of strategies that employ co-transfection of multiple plasmids, such as Cas9-gRNA plasmid and donor plasmid.

Interestingly, unlike our results in *E. tenella*, stable expression of Cas9 protein in *T. cruzi* (Peng et al., 2015) and *T. gondii* (Sidik et al., 2016) reduced the fitness of these parasites. The Cas9 protein altered the growth kinetics of *T. cruzi*, and this alteration could be rescued by deleting the Cas9 gene integrated in the *T. cruzi* genome (Peng et al., 2015). Constitutively expressed Cas9 protein is also lethal to *T. gondii*, and a "decoy" gRNA was necessary to reduce partially the toxicity of Cas9-expression in *T. gondii* (Sidik et al., 2016). According to Markus et al. (2019), viral 2A peptides can greatly improve the stability of Cas9 expressed in *T. gondii* tachyzoites. Thus, in our study, we fused a 2A peptide to the C-terminal end of Cas9, which made it possible to achieve a stable and constitutive Cas9-expression in *E. tenella* (in all life stages) without an impact on parasite development.

The CRISPR-Cas9 strategy developed in this study has many advantages. For example, it can reduce the time needed to construct Cas9 plasmids, and eliminates the need for co-transfection of a Cas9 plasmid and a donor vector, which may greatly reduce the editing efficiency. The strategy also enables possibilities in multi-gene family disruption by transfecting with gRNAs targeting conserved regions (Peng et al., 2015) or by multiple gRNA delivery (Sidik et al., 2016). Our analysis of the transgenic line identified limitations in the *Eimeria* system that need to be addressed in order to create a strategy amenable to large-scale functional analysis. *Eimeria* species repair their genome DSBs mainly by NHEJ, and that the ratio of random

insertions in this parasite is higher than recombination when transfecting with linearized plasmid DNA. Another challenge is the lack of a reliable continuous *in vitro* culture system, which makes it difficult to clone individual parasites. Therefore, in order to avoid random insertions, we employed targeting constructs that harbor the DHFR-mCherry cassette without a promoter. Thus, only insertions resulting from homologous recombination result in expression of the maker–reporter fusion. However, this strategy might have limited application for genes whose expression is weak. Improvement of the *in vitro* culture system or disruption of the activity of the NHEJ pathway in *E. tenella* may help overcome this challenge.

Systematic functional analysis of *Eimeria* parasites could also benefit from strategies that employ a gene targeting strategy using PCR products rather than cloning. So far attempts to knock out the endogenous YFP in the Et-HCYA line by using DHFR:mCherry PCR products with 39 bp homology recombinant regions have not been successful. Further optimization of this approach is desirable and could pave the way for knockout of genes at a large scale. Recently, the CRISPR-Cas9 gene editing system in *Trypanosome* was optimized by transfecting parasites with ribonucleoprotein complexes consisting of the *S. aureus* Cas9 and gRNAs (Medeiros et al., 2017). Using this transfection strategy, a very high (near 100%) editing efficiency in *T. cruzi* and other kinetoplastids was achieved (Medeiros et al., 2017). Application of this strategy in *Eimeria* is also warranted.

In summary, we successfully adapted the CRISPR-Cas9 gene editing system for the genetic manipulation of *E. tenella* and created a transgenic line that constitutively express Cas9 without affecting the life cycle or fitness of the parasite. Using this line, we were able to manipulate both individual genes (*yfp* and *Etgra9*) as well as multiple genes (gene family of transcription factors in this study). It is anticipated that this toolbox will usher a new era in the functional genomic of *Eimeria* species, advance the biology of the parasite and help develop more effective strategies to control coccidiosis.

MATERIALS AND METHODS

Ethics Statement

Use of animals in this study was approved by the Beijing Administration Committee of Laboratory Animals and performed in accordance with the China Agricultural University Institutional Animal Care and Use Committee guidelines [approval number: AW05(7)069102-2].

Animals and Parasites

One- to six-week-old AA broilers used for *E. tenella* passages and transfection experiments were purchased from Beijing Arbor Acres Poultry Breeding (Beijing, China). Ten-day-old SPF chickens used in pathogenicity studies were purchased from Merial Animal Health Co., Ltd. (Beijing, China). All birds were fed with a coccidia-free diet and water *ad libitum*. The wild-type *E. tenella* Beijing strain used was maintained in the lab. The procedures for collection, purification, and sporulation of the parasite were carried out as previously described (Eckert, 1995).

MDBK Cell Cultural and Parasite *in vitro* Infection

Madin-Darby Bovine Kidney cells were cultured as previously reported (Tierney and Mulcahy, 2003). *In vitro E. tenella* sporozoites infections were performed as described by Bussière et al. (2018). To remove the extra-cellular parasites, three times of wash with PBS were performed before sampling.

Codon Optimization of SpCas9 Using *E. tenella* Optimal Codon Usage

For optimal expression of the SpCas9 protein in *E. tenella*, the gene was synthesized using *E. tenella* optimal codon usage, which was downloaded from Codon Usage Database[1]. Codon optimization was achieved using the online software OPTIMIZER[2]. The optimized Cas9 sequence could be found in **Supplementary Data S1**.

Plasmid Construction

To construct the eCas9/YFP plasmid, the *eCas9* gene and the YFP gene encoding enhanced yellow fluorescent protein were expressed in a single cassette under the regulatory control of the Histone 4 promoter. The sequence was designed to also include a self-cleavage 2A peptide sequence between eCas9 and YFP.

To determine the cleavage activity of Cas9 in Et-HCYA line, a gRNA-producing plasmid (pgRNA) targeting YFP locus was constructed. In this plasmid, the mCherry sequence was fused to the pyrimethamine resistant gene TgDHFR-ts-m2m3 (DHFR) positive selectable marker and expressed under the control of the EtMic2 promoter. The plasmid also harbors the U6-gRNA cassette. The small guide RNA was designed using EuPaGDT[3]. The YFP-KO vector consisted of two cassettes, the gRNA was regulated by EtU6 promoter, and the DHFR positive selectable marker and mCherry. These sequences were flanked by 5′ (816 bp) and 3′ (529 bp) homology sequences. For the GRA9-tag vector, the coding region of mCherry was flanked by 5′ (gene sequence of *Etgra9*, 1002 bp) and 3′ (3′downstream of *Etgra9*, 1420 bp) homology sequences, followed by a gRNA cassette derived by EtU6.

For the genome-wide *Etap2* gene family disruption screen, *E. tenella ap2* genes were identified by searching against ToxoDB v42 using PF00847 and PF14733 with 1.0E−05 as threshold. Thirty-three *ap2* vectors were constructed based on pgRNA with substitution of gRNAs.

All PCR amplifications were performed using Q5® High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, United States) and primers were listed in **Supplementary Table S1**. Elements in each plasmid were linked by seamless assembly strategy (ClonExpress MultiS One Step Cloning Kit, Vazyme Biotech Co., Ltd., Nanjing, China) based on pEASY-Blunt-Simple cloning vector (TransGen Biotech, Beijing, China)

[1]http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=5802

[2]http://genomes.urv.es/OPTIMIZER/

[3]http://grna.ctegd.uga.edu/

backbone with double SnaBI sites (for linearization) on the edge of inserts.

Transfection and Establishment of Cas9-Expressing *E. tenella* Line

Eimeria tenella sporozoites were harvested and purified from freshly purified oocysts by Percoll centrifugation and bile-trypsin digestion (Dulski and Turner, 1988). The *Sna*BI-mediated transfection procedures were performed as previously reported (Qin et al., 2014). For *Et-ap2* gene family disruptions, five to six gRNA coffering plasmids were mixed as a group (six group in total), and 20 μg of each linearized plasmid was used for one transfection with $\sim 1 \times 10^7$ sporozoites. These sporozoites were subsequently inoculated to six 2-week-old chicken through cloaca. Pyrimethamine (150 mg/L) was added to the drinking water with solubilize formula (kindly provided by Dr. Tuanyuan Shi from Zhejiang Academy of Agricultural Science, China).

For enrichment of positive transgenic parasites, sporocysts of the first-generation oocysts after transfection were extracted and purified as introduced previously (Dulski and Turner, 1988), and then sorted by flow cytometry. Briefly, for sporocyst isolation, purified sporulated oocyst were broken by vortexing with equal volume of 1 mm glass beads for 30 s (3000 rpm), then the broken oocyst pellets were resuspended with 50% percoll, remove the oocyst wall fragments in the up-layer after centrifuge for 1 min at 10,000 × *g*, then the sporocyst pellets were washed with PBS for three times. Single sporocyst was sorted and isolated to generate genetically stable Cas9-expressing *E. tenella* lines (Et-HCYA). To locate the insertion site of eCas9 plasmid in Et-HCYA #3 clone, genome walking experiments were performed according to the manufacture's instruction (Takara Biomedical Technology, Beijing, China).

Indirect Immunofluorescence Assay and Immunoblotting

Immunofluorescence assay (IFA) and immunoblotting were performed to confirm the expression of eCas9 in *E. tenella* as previously described (Tang et al., 2018a). Mouse anti-Flag monoclonal antibody (1:200) and Cy3-conjugated goat anti-mouse IgG (1:200) were used for the IFA and western blot analysis. Extracellular sporozoites were fixed with 4% formaldehyde for 10 min, and permeabilized with 0.25% Triton X-100 in PBS for 8 min and stained with anti-Flag antibody, and the nuclei were stained with DAPI.

Total oocyst proteins of Et-HCYA and wild strain were extracted for immunoblotting. Anti-Flag mAb (1:1000) and horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG (1:2000) were used to detect Flag-tagged eCas9 protein. Anti-GAPDH mouse mAb (1:1000) was used as a control.

Immuno-Electron Microscopy

Briefly, samples were fixed in 2% PFA and 0.2% GA overnight. After rinsing with phosphate buffer (PB) (0.1 M, pH 7.4), samples were dehydrated through a graded ethanol series (30, 50, 70, 80, 90, and 100%, 10 min each). Samples were infiltrated in a graded mixtures (3:1, 1:1, 1:3) of ethanol and LR White resin (Ted Pella) then changed two to three times pure resin. Finally, samples were embedded in pure resin and polymerized for 48 h at 50°C. The ultrathin sections (70 nm thick) were sectioned with microtome (Leica EM UC6). After rinsing with 0.1 M PB, sections were blocked in 5% goat serum in 1% BSA buffer (with 0.15% Glycine in PB) for 30 min, then incubated with 1:10 diluted Rabbit Anti-mCherry antibody (ab167453, Abcam, Cambridge, United Kingdom) for 2 h. After rinsing with 0.1 M PB six times again, sections were followed by immunogold labeled with 10-nm protein A-gold (1: 50; Cell Microscopy Center, University Medical Center Utrecht, Utrecht, Netherlands) for 1 h. Following rinses with PB, sections were stained with 2% uranyl acetate for 5 min, and imaged by a transmission electron microscope (FEI Tecnai Spirit 120kV).

Oocyst Output Curves and Virulence Determination

The effect of eCas9 expression on the life cycle of *E. tenella* was determined by comparing the oocyst output and intestine damages between the Et-HCYA clones and the parental wild-type parent strain. One-week-old AA broilers (n = 4) were infected with 1000 fresh oocysts/bird/strain, oocyst outputs were measured daily during 5–12 days post infection (d.p.i.). Total oocyst outputs were also calculated by using a modified McMaster chamber (Haug et al., 2006). Ten-day-old SPF chicken were inoculated with 5000/10,000 fresh oocysts/bird (n = 5) and were then sacrificed at 7 d.p.i. for lesion scoring of the ceca (Johnson and Reid, 1970).

High-Throughput gRNA Detection in EtAp2 Family Screen

2–3 × 10^5 sporocysts of each group were sorted by FACS and mixed before infecting chicken for second generation merozoites. Genomic DNA was extracted from these merozoites, and the integrated gRNA sequences were amplified by nested PCR. Subsequently, the amplicons were used for illumina library construction using NEBNext® Multiplex Small RNA Library Prep Set for Illumina® kit (New England Biolabs, Ipswich, MA, United States) following the manufacturer's instruction. The library was then used for Miseq using PE250 platform. The output reads were filtered and mapped to the gRNA library, and then the mapped reads were counted for each gRNA.

Whole Genome Sequencing and Off-Target Evaluation

Second generation merozoites of Et-HCYA clone #3 and Cas9-YFP-KO line were isolated for gDNA extraction. Genome DNA were sheared into ∼300 bp fragments and were then processed and sequenced using illumina Novaseq 6000 platform according to the manufacturer's protocol. Paired-end reads were filtered and then mapped to *E. tenella* reference genome (an unpublished new version) by bwa-mem, and the SAM files were sorted and converted to BAM files, and PCR duplications were also marked by GATK v4.0.11. The small nucleotide variations (SNVs) were called by freebayes v1.02 with parameters of "-C 5 –min-mapping-quality 30 –min-base-quality

20 –min-coverage 10" and samtools mpileup pipeline with default settings. The concordance SNVs called from the two methods were kept and used for filtration with "QUAL < 60.0, QD < 20.0, FS > 13.0, MQ < 30.0, MQRankSum < −1.65, ReadPosRankSum < −1.65." In house python script was used to identify different SNVs between Et-HCYA and Cas9-YFP-KO lines. Potential off-target sites were predicted by Cas-OFFinder (Bae et al., 2014) with maximum five nucleotides mismatch, and the 2000 bp region around the potential off-target sites were checked for its mutation manually.

Statistical Analysis

Unpaired multiple *t*-tests were used for the analysis in total oocyst output and lesion scores. All bar plots depict the mean with standard deviations shown as error bars.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the NCBI via SRA accession: PRJNA587588: https://www.ncbi.nlm.nih.gov/sra/PRJNA587588.

ETHICS STATEMENT

The animal study was reviewed and approved by the China Agricultural University Institutional Animal Care and Use Committee.

AUTHOR CONTRIBUTIONS

XL, XS, and CM conceived and designed the study. DH performed the experiments, analyzed the data, and drafted the manuscript. XT, CW, SW, XG, and CD helped in plasmid construction, data analysis, and manuscript writing. XT, SZ, JS, YY, and MD helped in animal experiments and oocyst collection. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (31772728 and 31873007), the earmarked fund for China Agriculture Research System (CARS-43), and the National Key Research and Development Program of China (2016YFD0501300 and 2018YFD0500300).

ACKNOWLEDGMENTS

We would like to show our appreciation to Prof. Bang Shen in Huazhong Agricultural University for his help in identifying *Eimeria* U6 promoter. We are also grateful to Dr. Tuanyuan Shi from Zhejiang Academy of Agricultural Science for the improvement of experimental method on preparing solubilized pyrimethamine. We thank the Flow Cytometry Core at National Center for Protein Sciences at Peking University, particularly Hongxia Lv and Liying Du, for technical help. Finally, we would like to thank Xixia Li, Li Wang, and Xueke Tan for help with immuno-electron microscopy at the Center for Biological Imaging (CBI), Institute of Biophysics, Chinese Academy of Sciences.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00128/full#supplementary-material

TABLE S1 | Primers used for this study (primers for plasmid construction are not listed).

TABLE S2 | Potential off-target sites of Cas9-YFP-KO.

TABLE S3 | Detailed read counts and other information for guide RNAs involved in this study. gRNAs were designed in EuPaGDT (http://grna.ctegd.uga.edu), and the top ranked gRNAs for each gene were selected for the transfection studies.

DATA S1 | *E. tenella*-codon usage optimized nucleotide sequence of SpCas9.

REFERENCES

Adjogble, K. D., Mercier, C., Dubremetz, J.-F., Hucke, C., MacKenzie, C. R., Cesbron-Delauw, M.-F., et al. (2004). GRA9, a new *Toxoplasma gondii* dense granule protein associated with the intravacuolar network of tubular membranes. *Int. J. Parasitol.* 34, 1255–1264. doi: 10.1016/j.ijpara.2004.07.011

Bae, S., Park, J., and Kim, J.-S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 30, 1473–1475. doi: 10.1093/bioinformatics/btu048

Blake, D. P. (2015). *Eimeria* genomics: where are we now and where are we going? *Vet. Parasitol.* 212, 68–74. doi: 10.1016/j.vetpar.2015.05.007

Blake, D. P., and Tomley, F. M. (2014). Securing poultry production from the ever-present *Eimeria* challenge. *Trends Parasitol.* 30, 12–19. doi: 10.1016/j.pt.2013.10.003

Burton, P., McBride, D. J., Wilkes, J. M., Barry, J. D., and McCulloch, R. (2007). Ku heterodimer-independent end joining in *Trypanosoma brucei* cell extracts relies upon sequence microhomology. *Eukaryot Cell* 6, 1773–1781. doi: 10.1128/ec.00212-07

Bussière, F. I., Niepceron, A., Sausset, A., Esnault, E., Silvestre, A., Walker, R. A., et al. (2018). Establishment of an *in vitro* chicken epithelial cell line model to investigate *Eimeria tenella* gamete development. *Parasit Vectors* 11, 44. doi: 10.1186/s13071-018-2622-1

Chapman, H. D., Barta, J. R., Blake, D., Gruber, A., Jenkins, M., Smith, N. C., et al. (2013). Chapter two – A selective review of advances in coccidiosis research. *Adv. Parasitol.* 83, 93–171. doi: 10.1016/B978-0-12-407705-8.00002-1

Clark, J. D., Billington, K., Bumstead, J. M., Oakes, R. D., Soon, P. E., Sopp, P., et al. (2008). A toolbox facilitating stable transfection of *Eimeria* species. *Mol. Biochem. Parasitol.* 162, 77–86. doi: 10.1016/j.molbiopara.2008.07.006

Dulski, P., and Turner, M. (1988). The purification of sporocysts and sporozoites from *Eimeria tenella* oocysts using Percoll density gradients. *Avian Dis.* 32, 235–239.

Eckert, J. (1995). *Guidelines on Techniques in Coccidiosis Research*. ECSC-EC-EAEC . Brussels: Office for Official Publications of the European Communities.

Fox, B. A., Ristuccia, J. G., Gigley, J. P., and Bzik, D. J. (2009). Efficient gene replacements in *Toxoplasma gondii* strains deficient for nonhomologous end joining. *Eukaryot Cell* 8, 520–529. doi: 10.1128/EC.00357-08

Glover, L., Jun, J., and Horn, D. (2010). Microhomology-mediated deletion and gene conversion in African trypanosomes. *Nucleic Acids Res.* 39, 1372–1380. doi: 10.1093/nar/gkq981

Haber, J. E. (2000). Partners and pathways: repairing a double-strand break. *Trends Genet.* 16, 259–264.

Haug, A., Williams, R. B., and Larsen, S. (2006). Counting coccidial oocysts in chicken faeces: a comparative study of a standard McMaster technique and a new rapid method. *Vet. Parasitol.* 136, 233–242. doi: 10.1016/j.vetpar.2005.11.024

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10.1126/science.1225829

Johnson, J., and Reid, W. M. (1970). Anticoccidial drugs: lesion scoring techniques in battery and floor-pen experiments with chickens. *Exp. Parasitol.* 28, 30–36. doi: 10.1016/0014-4894(70)90063-9

Kafsack, B. F., Rovira-Graells, N., Clark, T. G., Bancells, C., Crowley, V. M., Campino, S. G., et al. (2014). A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature* 507, 248. doi: 10.1038/nature12920

Kelleher, M., and Tomley, F. (1998). Transient expression of β -galactosidase in differentiating sporozoites of *Eimeria tenella*. *Mol. Biochem. Parasitol.* 97, 21–31. doi: 10.1016/s0166-6851(98)00128-5

Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M. D. C., and Yusa, K. (2014). Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat. Biotechnol.* 32, 267. doi: 10.1038/nbt.2800

Liu, X., Shi, T., Ren, H., Su, H., Yan, W., and Suo, X. (2008). Restriction enzyme-mediated transfection improved transfection efficiency in vitro in Apicomplexan parasite *Eimeria tenella*. *Mol. Biochem. Parasitol.* 161, 72–75. doi: 10.1016/j.molbiopara.2008.06.006

Markus, B. M., Bell, G. W., Lorenzi, H. A., and Lourido, S. (2019). Optimizing systems for Cas9 expression in *Toxoplasma gondii*. *mSphere* 4:e00386-19. doi: 10.1128/mSphere.00386-19

Medeiros, L. C. S., South, L., Peng, D., Bustamante, J. M., Wang, W., Bunkofske, M., et al. (2017). Rapid, selection-free, high-efficiency genome editing in protozoan parasites using CRISPR-Cas9 ribonucleoproteins. *mBio* 8:e01788-17. doi: 10.1128/mBio.01788-17

Mehdi, G., Molly, G., Cameron Ross, M., Rafael Miyazawa, M., Artur, S., and Jose-Juan, L. R. (2014). Genome editing in the human malaria parasite *Plasmodium falciparum* using the CRISPR-Cas9 system. *Nat. Biotechnol.* 32, 819–821. doi: 10.1038/nbt.2925

Painter, H. J., Campbell, T. L., and Llinás, M. (2011). The Apicomplexan AP2 family: integral factors regulating *Plasmodium* development. *Mol. Biochem. Parasitol.* 176, 1–7. doi: 10.1016/j.molbiopara.2010.11.014

Pastor-Fernández, I., Kim, S., Billington, K., Bumstead, J., Marugán-Hernández, V., Küster, T., et al. (2018). Development of cross-protective *Eimeria*-vectored vaccines based on apical membrane antigens. *Int. J. Parasitol.* 48, 505–518. doi: 10.1016/j.ijpara.2018.01.003

Peng, D., Kurup, S. P., Yao, P. Y., Minning, T. A., and Tarleton, R. L. (2015). CRISPR-Cas9-mediated single-gene and gene family disruption in *Trypanosoma cruzi*. *mBio* 6:e02097-1. doi: 10.1128/mBio.02097-14

Qin, M., Liu, X. Y., Tang, X. M., Suo, J. X., Tao, G. R., and Suo, X. (2014). Transfection of *Eimeria mitis* with yellow fluorescent protein as reporter and the endogenous development of the transgenic parasite. *Plos One* 9:e114188. doi: 10.1371/journal.pone.0114188

Radke, J. B., Lucas, O., De Silva, E. K., Ma, Y., Sullivan, W. J., Weiss, L. M., et al. (2013). ApiAP2 transcription factor restricts development of the *Toxoplasma* tissue cyst. *Proc. Natl. Acad. Sci. U.S.A.* 110, 6871–6876. doi: 10.1073/pnas.1300059110

Reid, A. J., Blake, D. P., Ansari, H. R., Billington, K., Browne, H. P., Bryant, J., et al. (2014). Genomic analysis of the causative agents of coccidiosis in domestic chickens. *Genome Res.* 24, 1676–1685. doi: 10.1101/gr.168955.113

Rodolphe, B., Christophe, F., Hélène, D., Melissa, R., Patrick, B., Sylvain, M., et al. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712. doi: 10.1126/science.1138140

Shen, B., Kevin, M. B., Tobie, D. L., and Sibley, L. D. (2014). Efficient gene disruption in diverse strains of *Toxoplasma gondii* using CRISPR/CAS9. *mBio* 5, 01114–01114.

Shi, T., Liu, X., Hao, L., Li, J., Gh, A. N., Abdille, M., et al. (2008). Transfected *Eimeria tenella* could complete its endogenous development in vitro. *J. Parasitol.* 94, 978–981. doi: 10.1645/GE-1412.1

Sidik, S. M., Hackett, C. G., Tran, F., Westwood, N. J., and Lourido, S. (2014). Efficient genome engineering of *Toxoplasma gondii* using CRISPR/Cas9. *PLoS One* 9:e100450. doi: 10.1371/journal.pone.0100450

Sidik, S. M., Huet, D., Ganesan, S. M., Huynh, M.-H., Wang, T., Nasamu, A. S., et al. (2016). A genome-wide CRISPR screen in *Toxoplasma* identifies essential Apicomplexan genes. *Cell* 166, 1423.–1435. doi: 10.1016/j.cell.2016.08.019

Sinha, A., Hughes, K. R., Modrzynska, K. K., Otto, T. D., Pfander, C., Dickens, N. J., et al. (2014). A cascade of DNA-binding proteins for sexual commitment and development in *Plasmodium*. *Nature* 507, 253. doi: 10.1038/nature12970

Sollelis, L., Ghorbal, M., MacPherson, C. R., Martins, R. M., Kuk, N., Crobu, L., et al. (2015). First efficient CRISPR—Cas9—mediated genome editing in Leishmania parasites. *Cell Microbiol.* 17, 1405–1412. doi: 10.1111/cmi.12456

Tang, X., Liu, X., Yin, G., Suo, J., Tao, G., Zhang, S., et al. (2018a). A novel vaccine delivery model of the Apicomplexan *Eimeria tenella* expressing *Eimeria maxima* antigen protects chickens against infection of the two parasites. *Front. Immunol.* 8:1982. doi: 10.3389/fimmu.2017.01982

Tang, X., Suo, J., Li, C., Du, M., Wang, C., Hu, D., et al. (2018b). Transgenic *Eimeria tenella* expressing profilin of *Eimeria maxima* elicits enhanced protective immunity and alters gut microbiome of chickens. *Infect. Immun.* 86:e00888-17.. doi: 10.1128/IAI.00888-17

Tierney, J., and Mulcahy, G. (2003). Comparative development of *Eimeria tenella* (Apicomplexa) in host cells in vitro. *Parasitol. Res.* 90, 301–304. doi: 10.1007/s00436-003-0846-1

Vinayak, S., Pawlowic, M. C., Sateriale, A., Brooks, C. F., Studstill, C. J., Bar-Peled, Y., et al. (2015). Genetic modification of the diarrhoeal pathogen *Cryptosporidium parvum*. *Nature* 523, 477. doi: 10.1038/nature14651

Wagner, J. C., Platt, R. J., Goldfless, S. J., Zhang, F., and Niles, J. C. (2014). Efficient CRISPR-Cas9–mediated genome editing in *Plasmodium falciparum*. *Nat Methods* 11, 915. doi: 10.1038/nmeth.3063

Zhang, W. W., and Matlashewski, G. (2015). CRISPR-Cas9-Mediated Genome Editing in *Leishmania donovani*. *mBio* 6:e00861. doi: 10.1128/mBio.00861-15

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 17 March 2020
doi: 10.3389/fbioe.2020.00152

Metabolic Engineering of *Saccharomyces cerevisiae* for Enhanced Dihydroartemisinic Acid Production

Bo-Xuan Zeng[1,2], *Ming-Dong Yao*[1,2], *Ying Wang*[1,2], *Wen-Hai Xiao*[1,2]* *and Ying-Jin Yuan*[1,2]

[1] *Frontier Science Center for Synthetic Biology and Key Laboratory of Systems Bioengineering (Ministry of Education), School of Chemical Engineering and Technology, Tianjin University, Tianjin, China,* [2] *Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), Tianjin University, Tianjin, China*

OPEN ACCESS

Edited by:
Jiazhang Lian,
Zhejiang University, China

Reviewed by:
Chris J. Paddon,
Amyris (United States), United States
Xixian Chen,
Independent Researcher, Singapore, Singapore
Yong-Su Jin,
University of Illinois at Urbana–Champaign, United States

***Correspondence:**
Wen-Hai Xiao
wenhai.xiao@tju.edu.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *08 November 2019*
Accepted: *14 February 2020*
Published: *17 March 2020*

Citation:
Zeng B-X, Yao M-D, Wang Y, Xiao W-H and Yuan Y-J (2020) Metabolic Engineering of Saccharomyces cerevisiae for Enhanced Dihydroartemisinic Acid Production. Front. Bioeng. Biotechnol. 8:152. doi: 10.3389/fbioe.2020.00152

Direct bioproduction of DHAA (dihydroartemisinic acid) rather than AA (artemisinic acid), as suggested by previous work would decrease the cost of semi-biosynthesis artemisinin by eliminating the step of initial hydrogenation of AA. The major challenge in microbial production of DHAA is how to efficiently manipulate consecutive key enzymes ADH1 (artemisinic alcohol dehydrogenase), DBR2 [artemisinic aldehyde Δ11(13) reductase] and ALDH1 (aldehyde dehydrogenase) to redirect metabolic flux and elevate the ratio of DHAA to AA (artemisinic acid). Herein, DHAA biosynthesis was achieved in *Saccharomyces cerevisiae* by introducing a series of heterologous enzymes: ADS (amorpha-4,11-diene synthase), CYP71AV1 (amorphadiene oxidase), ADH1, DBR2 and ALDH1, obtaining initial DHAA/AA ratio at 2.53. The flux toward DHAA was enhanced by pairing fusion proteins DBR2-ADH1 and DBR2-ALDH1, leading to 1.75-fold increase in DHAA/AA ratio (to 6.97). Moreover, to promote the substrate preference of ALDH1 to dihydroartemisinic aldehyde (the intermediate for DHAA synthesis) over artemisinic aldehyde (the intermediate for AA synthesis), two rational engineering strategies, including downsizing the active pocket and enhancing the stability of enzyme/cofactor complex, were proposed to engineer ALDH1. It was found that the mutant H194R, which showed better stability of the enzyme/NAD^+ complex, obtained the highest DHAA to AA ratio at 3.73 among all the mutations. Then the mutant H194R was incorporated into above rebuilt fusion proteins, resulting in the highest ratio of DHAA to AA (10.05). Subsequently, the highest DHAA reported titer of 1.70 g/L (DHAA/AA ratio of 9.84) was achieved through 5 L bioreactor fermentation. The study highlights the synergy of metabolic engineering and protein engineering in metabolic flux redirection to get the most efficient product to the chemical process, and simplified downstream conversion process.

Keywords: dihydroartemisinic acid, artemisinic acid, protein engineering, synthetic biology, *Saccharomyces cerevisiae*

INTRODUCTION

Artemisinin, a well-known sesquiterpene lactone isolated from extracts of *Artemisia annua* with excellent anti-malaria properties, had been designated as first-line antimalarial drugs by WHO in 2002 (Paddon and Keasling, 2014). Artemisinin production and prices vary greatly as they depend on plant extraction, which relies on the supply of plant materials (Chunyin and Cook, 2012; Peplow, 2013). Therefore, a stable and sustainable supply of artemisinin is highly desirable. This breakthrough was achieved by the Amyris, Inc., in which a semi-synthetic process of artemisinin production was developed (Westfall et al., 2012; Paddon et al., 2013). The semi-synthesis of artemisinin consists of two parts: (1) *de novo* biosynthesis of AA (artemisinic acid) with a very high titer (25 g/L) by an engineered *Saccharomyces cerevisiae* and (2) extract AA from yeast culture and transform it to artemisinin by chemical process. To be noted, the first step of the chemical process was reduction of AA to dihydroartemisinic acid (DHAA). Although the biosynthesis of DHAA has been demonstrated (Zhang et al., 2008), the productivity (about 15.7 mg/L) has not yet to be optimized in a heterologous host. Although hydrogenation conversion of AA to DHAA was efficient (conversion efficiency 99%), the production of unexpected isoform (R, S)-DHAA (about 6%) was not avoided (Paddon et al., 2013; Turconi et al., 2014; Pieber et al., 2015). Instead, biosynthesis of DHAA produced very little (R, S) isomer in yeast (Zhang et al., 2008). In order to execute a more efficient way of synthesizing artemisinin, engineering DHAA biosynthesis in microbes would open up a promising alternative route.

The biosynthesis pathways of DHAA and AA both start from amorpha-4,11-diene that is oxidized to AO (artemisinic aldehyde) through AOH (artemisinic alcohol) by CYP17AV1 (amorphadiene oxidase) and ADH1 (artemisinic alcohol dehydrogenase) (**Figure 1**). AO is a joint intermediate followed by two branch biosynthesis pathways: (1) directly oxidized to AA by ALDH1 (aldehyde dehydrogenase) or (2) reduced to dihydroartemisinic aldehyde (DHAO) by DBR2 (artemisinic aldehyde Δ11(13) reductase) and then oxidized to DHAA by ALDH1 (Chen et al., 2017). The enzymatic parameters of DBR2 and ALDH1 to their substrates AO and DHAO were summarized in **Supplementary Table S1** (Zhang et al., 2008; Teoh et al., 2009). Because of the slightly higher reported affinity of ALDH1 [For substrate AO, Km(DBR2) = 19 μM, Km(ALDH1) = 2.58 μM], there is no advantage for DBR2 to bind with AO. Furthermore, no enzymes (even DBR2) have been reported to directly catalyze AA to DHAA. Thus, it was inevitable to produce a large amount of AA as by-product of DHAA production. So far, the highest ratio of DHAA/AA reaches just 1.67 achieved by enzymatic reaction *in vitro* (Chen et al., 2017). Therefore, the major challenge in production of DHAA by recombinant yeasts is efficient redirection of carbon flux to DHAA rather than AA.

Branch-point regulatory mechanisms are involved in many natural metabolic pathways (such as TCA cycle) (Chen et al., 2017). In these pathways, the enzymes can work together spatiotemporally due to protein–protein interactions and channel the metabolites between sequential enzymes without equilibration in the aqueous phase inside cells (Sweetlove and Fernie, 2018). The assemblies of consecutive enzymes are formed either from large clusters of multiple copies of enzymes, or by pairwise interactions of enzymes from single complexes which are beneficial for enzymes to reach substrate saturation (Zhang, 2011) so that the reaction flux is regulated at a branch point. Pham et al. (2015) once co-localized enzymes responsible for GABA (Gamma-aminobutyric acid) biosynthesis together to switch the metabolic flux toward GABA from TCA cycle and finally increased the production of GABA by 2.7 fold.

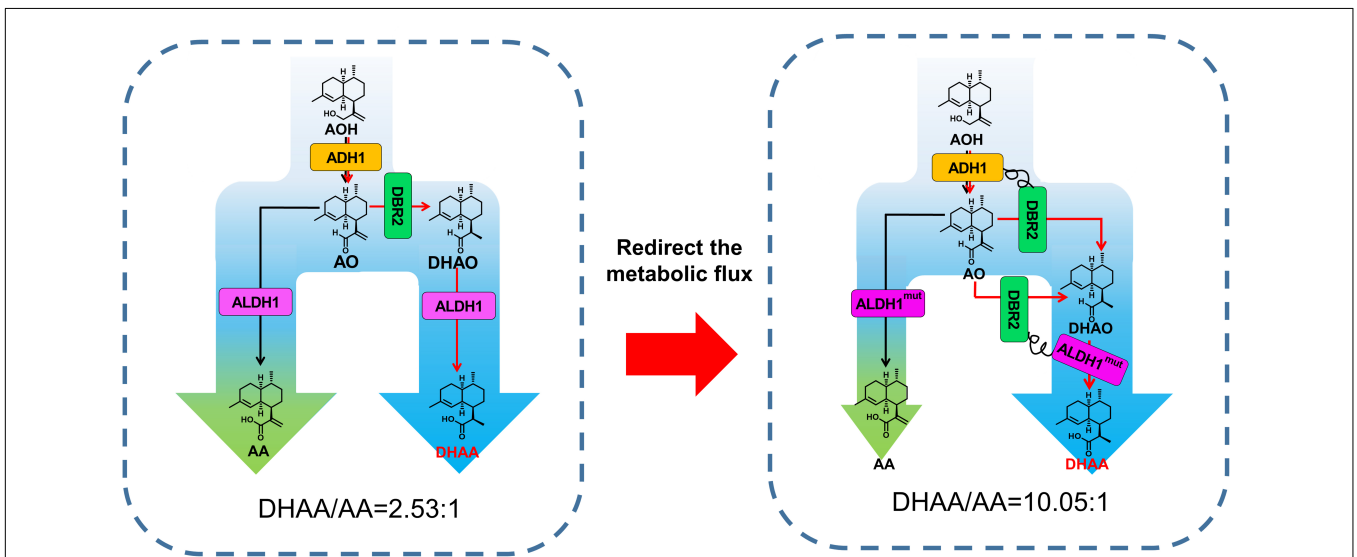


FIGURE 1 | Schematic diagram of tailoring the ratio of DHAA to AA by protein fusion and modifying ALDH1 catalysis preference. The broader blue arrow indicates enhanced flux toward DHAA. AOH, artemisinic alcohol; AO, artemisinic aldehyde; DHAO, dihydroartemisinic aldehyde; ADH1, artemisinic alcohol dehydrogenase of *A. annua*; DBR2, artemisinic aldehyde Δ11(13) reductase of *Artemisia annua*; ALDH1, artemisinic aldehyde dehydrogenase of *A. annua*.

Referring to DHAA biosynthetic pathway itself and key enzymes, two obstacles should be overcome to achieve a high ratio of DHAA/AA and a high DHAA titer: (1) redirecting the metabolic flux from AA toward DHAA *via* the expression of DBR2 requires the assembly of the pathway enzyme in a desired order and promotes the reactions of metabolites along a specified pathway (Miles et al., 1999); (2) rationally engineering ALDH1 to shift the substrate specificity from AO to DHAO.

Herein, the biosynthesis pathway of DHAA was successfully rebuilt in *S. cerevisiae* with high FPP supply. In order to improve the ratio of DHAA/AA by increasing the substrate accessibility of AO by DBR2 as well as that of DHAO by ALDH1, fusion proteins of paired enzymes (ADH1-DBR2 or DBR2-ALDH1) were adopted to reorganize the biosynthetic pathway of DHAA (**Figure 1**). Meanwhile, ALDH1, as the joint enzyme for biosynthesis of DHAA and AA, was also rationally engineered to shift the substrate specificity from AO to DHAO. Correspondingly, the ratio of DHAA/AA was enhanced by 3.34 fold (from 2.53 to 10.05) through integrating these above two strategies, without the compromise of DHAA production. Eventually, the highest DHAA titer of 1.70 g/L (DHAA/AA ratio of 9.84) was achieved in a 5 L bioreactor through high density fermentation. The study highlights the importance of redirecting metabolic flux toward a desired target *via* consecutive enzyme-enabled reorganization.

MATERIALS AND METHODS

Strains and Medium

All *E. coli* used for plasmid construction were cultured at 37°C in Luria-Bertani (LB) medium (1% tryptone, 0.5% yeast extract, and 1% NaCl) with 100 μg/ml ampicillin or 34 μg/ml chloramphenicol if necessary.

All engineered yeast strains were derived from *S. cerevisiae* CEN.PK2-1C (Entian and Kötter, 2007) obtained from EUROSCARF (Frankfurt, Germany) and were listed in **Table 1**. *S. cerevisiae* strains were cultured in YPD medium (2% tryptone, 1% yeast extract, and 2% glucose) or in synthetic complete (SC) drop-out medium at 30°C. All the medium formulations for yeast culture are available in our previous work (Su et al., 2015).

Plasmid Construction

Plasmids used in this study are listed in **Table 2**. The genes *ADS* (ACCESSION Q9AR04), *CYP71AV1* (ACCESSION Q1PS23), *DBR2* (ACCESSION KC505370.1), *ALDH1* (ACCESSION JQ609276.1), *ADH1* (ACCESSION JF910157.1), *CYB5* (ACCESSION JQ582841.1), *CPR1* (ACCESSION DQ318192.1) from *Artemisia annua* were codon-optimized for expression in yeast and synthesized by GenScript, Inc. (China). To overexpress three key genes for DHAA biosynthesis (including ADS, CYP71AV1, and DBR2), plasmid pZBX040 were first constructed based on the multi-copy plasmid pRS425. Then the plasmid with fusion protein DBR2-ADH1 or ADH1-DBR2 (including pZBX059, pZBX060, pZBX067, pZBX069) were constructed based on pZBX040. To substitute ALDH1 integrated in the genome with fusion protein DBR2-ALDH1, the cassette *HIS3*-T_{HIS3}-P_{GAL7}-*DBR2-Linker-ALDH1*-T_{TDH1} were assembled in pSB1C3 (obtained from the Registry of Standard Biological Parts[1]) to form plasmid pZBX100, pZBX101. The cassettes with mutated ALDH1 were constructed based on pZBX101. All the primers used in this work are listed in **Supplementary Table S2**.

TABLE 1 | Yeast strains used in this study.

Yeast strains	Description	Source
CEN.PK2.1C	*MAT a; ura3-52, trp1-289, leu2-3,112, his3Δ1, MAL2-8C, SUC2*	Invitrogen
Sc027	CEN.PK2-1C derivative; *leu2-3,112:G418^R_P_{GAL7}-CYB5_T_{ERG19}(RC)-ERG19(RC)-P_{GAL1}(RC)_P_{GAL10}-ERG8-T_{ERG8}; his3Δ1:HIS3_P_{GAL7}-ALDH1-T_{TDH1}_T_{ERG12}(RC)-ERG12(RC)-P_{GAL1}(RC)_P_{GAL10}-ERG10-T_{ERG10}; ade1Δ:T_{HMG1}(RC)-tHMG1(RC)-P_{GAL1}(RC)_P_{GAL10}-IDI1-T_{IDI1}_ADE1; ura3-52:T_{HMG1}(RC)-tHMG1(RC)-P_{GAL1}(RC)_P_{GAL10}-ERG13-T_{ERG13}; trp1-289:T_{HMG1}(RC)-tHMG1(RC)-P_{GAL1}(RC)_P_{GAL10}-ERG20-T_{ERG20}_TRP1; gal1/10/7Δ:natA_P_{GAL3}-CPR1-T_{CYC1};*	This study
Sc057	Sc027 derivative; *GAL80Δ:URA3_P_{GAL7}-AaADH1-T_{TDH1}*	This study
Sc077	Sc027 derivative; *GAL80Δ:URA3*	This study
Sc085	Sc057 derivative; pZBX040	This study
Sc113	Sc077 derivative; pZBX059	This study
Sc115	Sc077 derivative; pZBX067	This study
Sc146	Sc077 derivative; pZBX060	This study
Sc147	Sc077 derivative; pZBX069	This study
Sc352	Sc057 derivative; *his3Δaldh1Δ:hphA*	This study
Sc361	Sc077 derivative; pZBX067; *ΔALDH1:P_{GAL7}-DBR2-Linker1-ALDH1-T_{TDH1}*	This study
Sc467	Sc057 derivative; pZBX040; *ΔALDH1:P_{GAL7}-DBR2-Linker2-ALDH1-T_{TDH1}*	This study
Sc468	Sc057 derivative; pZBX040; *ΔALDH1:P_{GAL7}-DBR2-Linker1-ALDH1-T_{TDH1}*	This study
Sc470	Sc085 derivative; *ura3down:hphMX6-P_{GAL7}-DBR2-T_{CYC1}*	This study
Sc429	Sc057 derivative; pZBX040 *ΔALDH1:P_{GAL7}-ALDH1^{H194R}-T_{TDH1}*	This study
Sc457	Sc057 derivative; pZBX067; *ΔALDH1:P_{GAL7}-DBR2-Linker1-ALDH1^{H194R}-T_{TDH1}*	This study

*The DNA fragment followed by '(RC)' represents that the orientation of the DNA fragment is reversed. Sequences of linker1 and linker2 showed in **Table 3**; In strain Sc470, the cassette hphMX6-P_{GAL7}-DBR2-T_{CYC1} was inserted into the downstream of cassette T_{HMG1}(RC)-tHMG1(RC)-P_{GAL1}(RC)_P_{GAL10}-ERG13-T_{ERG13} which had been integrated in the locus in ura3.*

Construction of pZBX040

The DNA fragments *ADS*, *CYP71AV1*, $P_{GAL1,10}$, T_{PGK1}, and T_{ADH1} were amplified by PCR and joined together by overlap extension PCR(OE-PCR) to obtain cassette T_{ADH1}(RC)-*CYP71AV1(RC)*-P_{GAL10}(RC)_P_{GAL1}-*ADS*-T_{PGK1}. The cassette was digested with *Bam*HI and *Xho*I and inserted into pRS425 to obtain plasmid pZBX020. Fragments *DBR2*, P_{GAL7}, T_{CYC1} were amplified by PCR and joined together by OE-PCR to obtain

[1]http://parts.igem.org/Part:pSB1C3

TABLE 2 | Plasmids used in this study.

Plasmid	Description	Source
pZBX020	*pRS425_T_{ADH1}(RC)-CYP71AV1(RC)-P_{GAL10}(RC)_P_{GAL1}-ADS-T_{PGK1}*	This study
pZBX040	*pRS425_P_{GAL7}-DBR2-T_{CYC1}_T_{ADH1}(RC)-CYP71AV1(RC)-P_{GAL10}(RC)_P_{GAL1}-ADS-T_{PGK1}*	This study
pZBX059	*pRS425_P_{GAL7}-ADH1-linker1-DBR2-T_{CYC1}_T_{ADH1}(RC)-CYP71AV1(RC)-P_{GAL10}(RC)_P_{GAL1}-ADS-T_{PGK1}*	This study
pZBX060	*pRS425_P_{GAL7}-ADH1-linker2-DBR2-T_{CYC1}_T_{ADH1}(RC)-CYP71AV1(RC)-P_{GAL10}(RC)_P_{GAL1}-ADS-T_{PGK1}*	This study
pZBX067	*pRS425_P_{GAL7}-DBR2-linker1-ADH1-T_{CYC1}_T_{ADH1}(RC)-CYP71AV1(RC)-P_{GAL10}(RC)_P_{GAL1}-ADS-T_{PGK1}*	This study
pZBX069	*pRS425_P_{GAL7}-DBR2-linker2-ADH1-T_{CYC1}_T_{ADH1}(RC)-CYP71AV1(RC)-P_{GAL10}(RC)_P_{GAL1}-ADS-T_{PGK1}*	This study
pZBX101	*pSB1C3_HIS3-T_{HIS3}_P_{GAL7}-DBR2-Linker1-ALDH1-T_{TDH1}*	This study
pZBX100	*pSB1C3_HIS3-T_{HIS3}_P_{GAL7}-DBR2-Linker2-ALDH1-T_{TDH1}*	This study
pZBX199	*pSB1C3_HIS3-T_{HIS3}_P_{GAL7}-ALDH1^{H194R}-T_{TDH1}*	This study
pZBX196	*pSB1C3_HIS3-T_{HIS3}_P_{GAL7}-ALDH1^{G227V}-T_{TDH1}*	This study
pZBX197	*pSB1C3_HIS3-T_{HIS3}_P_{GAL7}-ALDH1^{G227F}-T_{TDH1}*	This study
pZBX194	*pSB1C3_HIS3-T_{HIS3}_P_{GAL7}-ALDH1^{G223V}-T_{TDH1}*	This study
pZBX195	*pSB1C3_HIS3-T_{HIS3}_P_{GAL7}-ALDH1^{G223F}-T_{TDH1}*	This study
pZBX218	*pSB1C3_HIS3-T_{HIS3}_P_{GAL7}-DBR2-Linker1-ALDH1^{H194R}-T_{TDH1}*	This study

cassette P_{GAL7}-*DBR2*-T_{CYC1}. The cassette was digested with *Xho*I and *Pst*I and inserted into pZBX020 to obtain plasmid pZBX040.

Construction of Plasmids for Fusion Protein ADH1-DBR2 and DBR2- ADH1

Fragments P_{GAL7}-*ADH1, DBR2*-T_{CYC1} were amplified by PCR and then assembled together to obtain cassettes of different types of P_{GAL7}-*ADH1-linker-DBR2*-T_{CYC1} by OE-PCR. These cassettes were digested with *Xho*I and *Pst*I and inserted into pZBX020 to obtain the plasmids with the fusion protein *ADH1-DBR2* (including pZBX059, pZBX060). Fragments P_{GAL7}-*DBR2, ADH1,* T_{CYC1} were amplified by PCR and then linked together to form cassettes of different types of P_{GAL7}-*DBR2-linker-ADH1*-T_{CYC1} by OE-PCR. These cassettes were digested with *Xho*I and *Pst*I and inserted into pZBX020 to obtain the plasmids with the fusion protein *DBR2-ADH1* (including pZBX067, pZBX069). The sequences of two linkers in the fusion proteins used in this work are listed in **Table 3**.

TABLE 3 | Linkers used in this study.

Linker	Sequence	Type
Linker1	EAAAKEAAAKA	Rigid
Linker2	GGGGSGGGGSGGGGS	Flexible

Construction of Plasmid for Fusion Protein DBR2-ALDH1

Fragment *HIS3*-P_{GAL7} was amplified by PCR from the genome of Sc085 followed by digestion of *Eco*RI and *Bsa*I. The fragments of *DBR2*, and *ALDH1*-T_{TDH1} were amplified by PCR from Sc085. Fragment *DBR2* was digested with *Bsa*I, and fragment *ALDH1*-T_{TDH1} was digested with *Bsa*I and *Pst*I. The vector pSB1C3 was digested with *Eco*RI, and *Bsa*I. All above four fragments were ligated together by T4 ligase to construct all the plasmid with the fusion protein *DBR2-ALDH1* (including pZBX099, pZBX100, pZBX101). The cassettes of the plasmids were released from plasmid followed by digestion of *Pme*I before transformation to the strain Sc352 for integration.

Construction of Plasmid for ALDH1 Mutants

Each DNA fragment of an ALDH1 mutant was divided into two parts P_{GAL7}-*ALDH1a* and *ALDH1b* with 40 bp overlapping according to the mutated site, and was amplified by PCR. Each pair of P_{GAL7}-*ALDH1a* and *ALDH1b* was ligated to form intact ALDH1 mutant by OE-PCR followed by digestion of *Spe*I and *Kpn*I. The vector pZBX100 was digested with *Spe*I, and *Kpn*I for insertion of the fragment of P_{GAL7}-ALDH1 mutant, obtaining pZBX194, pZBX195, pZBX196, pZBX197, and pZBX199.

The process of construction of pZBX218 was the same as that of pZXB101 except that the fragment *ALDH1*-T_{TDH1} was amplified from plasmid pZBX199. All the cassettes of the plasmid described here were released from plasmid by digestion of *Pme*I and transformed to the strain Sc352.

Fermentation Condition

Medium for Fermentation

All the fermentation medium (FM) recipe used in this work was prepared similarly to that used in Westfall's work with some modifications (Westfall et al., 2012). The medium was composed of 8 g/L KH_2PO_4, 15 g/L $(NH4)_2SO_4$, 6.2 g/L $MgSO_4 \cdot 7H_2O$, 40 g/L glucose, 12 ml/L vitamin solution, and 10 ml/L trace metal solution. The vitamin solution included 0.05 g/L biotin, 1 g/L calcium pantothenate,1 g/L nicotinic acid, 25 g/L myo-inositol,1 g/L thiamine HCl, 1 g/L pyridoxal HCl, 0.2 g/L *p*-aminobenzoic acid and 2 g/L adenine sulfate. The trace metal solution was composed of 5.75 g/L $ZnSO_4 \cdot 7H_2O$, 0.32 g/L $MnCl_2 \cdot 4H_2O$, 0.32 g/L Anhydrous $CuSO_4$, 0.47 g/L $CoCl_2 \cdot 6H_2O$, 0.48 g/L $Na_2MoO_4 \cdot 2H_2O$, 2.9 g/L $CaCl_2 \cdot 2H_2O$, 2.8 g/L $FeSO_4 \cdot 7H_2O$, and 80 ml/L EDTA solution (containing 0.5 mol/L Na_2EDTA pH = 8.0). The amino acid solution (10 ml/L), including 2 g/L methionine, 6 g/L tryptophan, 8 g/L isoleucine, 5 g/L phenylalanine, 10 g/L sodium glutamate, 20 g/L threonine, 10 g/L aspartate, 15 g/L valine, 40 g/L serine and 2 g/L arginine, was supplemented to the medium in our study for obtaining better cell growth use. The FM bases (8 g/L KH_2PO_4, 15 g/L $(NH_4)_2SO_4$, and 6.2 g/L $MgSO_4 \cdot 7H_2O$) and glucose stock solution (667 g/L) were sterilized using an autoclave. Vitamin solution, trace metal solution, and amino acid solution were sterilized by filtration. All the components of the FM were mixed together after sterilization. Glucose stock solution (667 g/L), filtrated 95% (v/v) ethanol solution and filtrated feed stock solution (80 g/L KH_2PO_4, 150 g/L $(NH4)_2SO_4$,

62 g/L $MgSO_4 \cdot 7H_2O$) also needed to be prepared for 5 L bioreactor fermentation. pH was adjusted to 5 by 10 mol/L NaOH prior to use.

Flask Fermentation

To prepare seed vials, single isolates of each strain from agar plates were grown for 18 h in FM medium at 30°C and 200 rpm. And then the cultures were inoculated into fresh FM medium at initial OD_{600} of 0.05 and grown for another 16 h cultivation at 30°C. The seed culture was transferred into 250 ml flask containing 25 ml FM medium at initial OD_{600} of 0.2. The cells were grown at 30°C with shaking at 200 rpm. After 24 h, 5 ml IPM (Isopropyl myristate) and 20 g/L ethanol was added to each flask. The whole fermentation process continued for 120 h until harvest.

5 L Bioreactor Fermentation

Seed culture preparation was the same as in-flask fermentation. The seed culture (200 mL) was inoculated into a 5 L bioreactor containing a 2 L batch FM medium with 20 g/L glucose. The pH was controlled at 5 by adding 5 mol/L NaOH. The gas flow and the temperature were maintained at 1.5 vvm and 30°C. During the fermentation, glucose solution was fed into the bioreactor at a speed of 0.3 mL/min after the glucose was depleted at the running time of about 6 h. During the glucose feeding stage, the dissolved oxygen level (DO) was kept at 40% by cascading agitation from 400 to 600 rpm. When OD_{600} increased to 50 (about 30–36 h), the feeding of glucose was switched to ethanol feeding. The concentration of ethanol was maintained at 10 g/L to guarantee no starvation and no excess carbon accumulation. During the ethanol feeding process, DO was automatically maintained at about 30% by cascading stirring.

The Measurement of DHAA and AA and Other Intermediates

After harvest, the fermentation broth was centrifuged at 12,000 *g* for 2 min and the IPM phase was collected. And then 50 μL organic phase was mixed with 950 μL methanol. After filtrated with 0.22 μm Nylon66 filter, the sample was ready for HPLC analysis.

A 10 μL aliquot was injected into waters e2695 HPLC with ultraviolet detection at 194 nm. A thermoHypersil BDS C18 column (4.6 mm × 150 mm × 5 μm) was used for separation, with the following gradient (channel A: acetonitrile, channel B: water plus 0.1% formic acid): 0–3 min 65% A, gradually increased to 100% A from 3 to 10 min, held at 100% A from 10 to 13 min, decreased to 65% A from 13 to 15 min, kept 65% A from 15 to 18 min. The column was held at 25°C during the separation. Under this condition, DHAA and AA were found to elute at 6.67 and 7.43 min, respectively (see **Figure 2B**). The concentrations of both products in the sample were calculated using the calibration curves of standards (HPLC ≥ 98%) which was purchased from Chengdu Pufei De Biotech, Co., Ltd. Standards of DHAO, AO, AOH, and AD were purchased from TRC (Toronto Research Chemicals).

Homology Modeling, Molecular Docking, and Structural Analysis

Three-dimensional structure models of ALDH1 were constructed using the program Swiss-Model[2] to get the structure information accordingly. The structures were modeled using the high sequence homology (>41% identified) and high resolution crystal structure of the aldehyde dehydrogenase family protein (indole-3-acetaldehyde dehydrogenase) combined with NAD^+ from *tomato* as the template (pdb id:5iuw-A). The structure models were subjected to energy minimization using the Swiss-Pdb Viewer. Afterward the docking of enzyme and ligand were performed using the AutoDockVina program (Oleg and Olson, 2010). The docking studies were run with DHAO or AO as ligands and the structure model of ALDH1 with NAD^+. The DHAO and AO structure files were retrieved from ZINC site (Irwin et al., 2012). The docking cluster analysis was performed in the AutoDockVina program environment, and clusters were characterized by binding energy (in kilocalories per mole). Establishment of dative bonds between ligand, NAD^+ and the corresponding amino acids were followed by energy minimization. The built complex structural analysis was done using Pymol software (Delano, 2010). The mutation at the specific amino acid site was also introduced using this software, which allowed exploration of the spatial and molecular interactions among amino acids.

RESULTS AND DISCUSSION

Strain Construction for DHAA Production

Referring to previous work (Westfall et al., 2012), CEN.PK2.1C was selected as the host strain for DHAA production. In order to decouple cell growth and product production, inducible promotor GAL (P_{GAL1}, P_{GAL7}, P_{GAL10}) was explored to control all the overexpressed genes (Peng et al., 2017). *GAL80* was deleted to eliminate the demand of galactose for de-repressing the promoter GAL. To increase the metabolic flux toward FPP, all the genes of MVA pathway (Westfall et al., 2012) were overexpressed in the genome. Among these genes, tHMG1 was integrated into the genome with three copies (**Supplementary Figure S1**). Three key heterologous genes for DHAA biosynthesis including *ADS, CYP71AV1* and *DBR2* were constructed in a multi-copy plasmid pRS425 for higher level expression. Meanwhile, four other heterologous genes, including *ADH1*, *ALDH1*, *CYB5* and *CPR1*, were integrated in the genome at one copy to enhance the metabolic flux to DHAA (**Figure 2A**). The strain for DHAA production (Sc085) was characterized in shake flask fermentation and the products were detected by HPLC at a wavelength of 194 nm. The results showed the presence of DHAA and AA, indicated by the observation that the retention time of DHAA and AA peaks was same as the mix standard (6.67 min for DHAA, 7.43 min for AA) (**Figure 2B**). And 327 mg/L of DHAA and 129 mg/L of AA were successfully detected. The DHAA/AA ratio of 2.53 was much higher than the reported highest ratio (DHAA/AA = 1.67) (**Figure 2C**) (Chen et al., 2017).

[2]http://swissmodel.expasy.org/

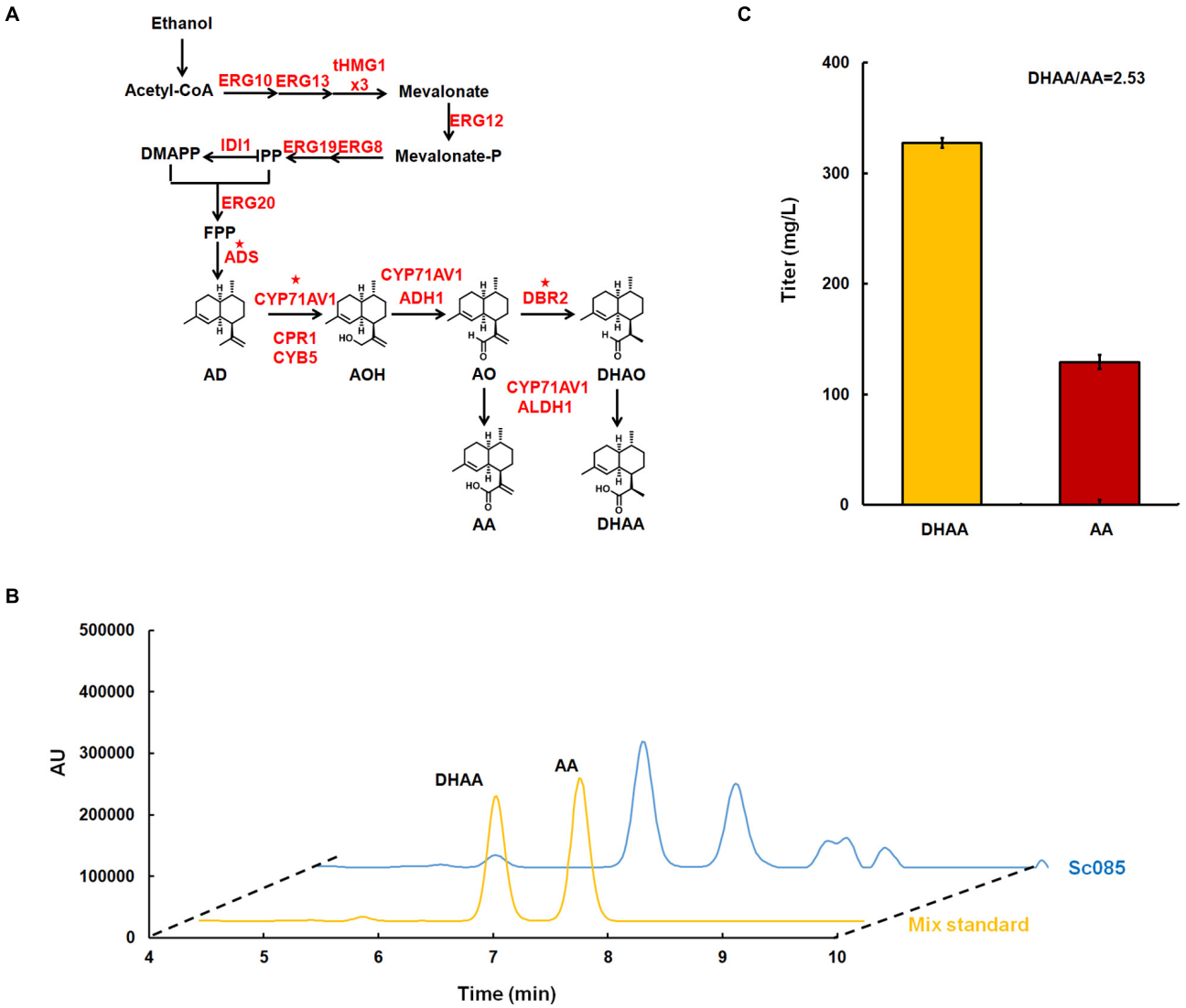
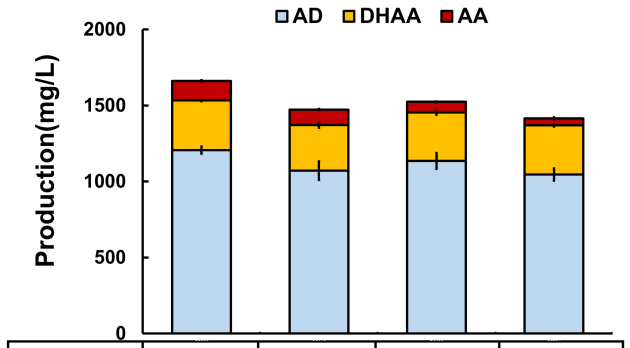


FIGURE 2 | Microbial production of dihydroartemisinic aldehyde in *Saccharomyces. cerevisiae*. **(A)** The biosynthesis pathway of DHAA derived from ethanol constructed in *S. cerevisiae*. All over-expressed genes are marked in red. The genes with star superscripted are over-expressed in multi-copy 2 μ plasmid pRS425 and others are over-expressed in the genome with strong promoter. For gene tHMG1, it is over-expressed with three copies in the genome, which is labeled as 'x3.' **(B)** Liquid chromatograms results for the mix standard of DHAA and AA (line in yellow), and the initial strain for DHAA production (line in blue). The DHAA eluted at about 6.67 min and AA eluted at 7.43 min. **(C)** The production of DHAA and AA in strain Sc085. And the ratio of DHAA/AA were shown below.

Nevertheless, there was still about 30% of the metabolic flux flowing toward AA (129 mg/L), probably indicating increasing expression level of DBR2 was needed to enhance the transformation from AO to DHAO. However, because of the reportedly slightly higher affinity of ALDH1 to the joint intermediate AO [Km(ALDH1) = 2.58 μM vs. Km(DBR2) = 19 μM] (Zhang et al., 2008; Teoh et al., 2009), individual over-expression of *DBR2* might not be sufficient to tackle this issue. When Sc470 held another copy of *DBR2* within the chromosome, the DHAA/AA ratio was reduced as well as the DHAA titer and no AOH, AO or DHAO was accumulated (**Figure 3** and **Supplementary Figure S4**). To summarize, it was difficult to directly control the reaction route toward DHAA rather than AA by simply increasing the expression of these genes. Considering that ALDH1 is a promiscuous enzyme that can catalyze AO and DHAO simultaneously, we need to reconfigure the enzymes (ADH1, DBR2, and ALDH1) in a desired order and switch the catalysis preference of ALDH1 to tailor the ratio of DHAA/AA.

Switch of the Biosynthesis Pathway to DHAA by Fusion Proteins

To increase the substrate accessibility of AO by DBR2, the biosynthetic route from AOH to DHAO was re-built *via* protein fusion of ADH1 and DBR2 (**Figure 1**). As known, linker type



Strain	Sc085	Sc115	Sc468	Sc361
DBR2-ADH1	-	+	-	+
DBR2-ALDH1	-	-	+	+
DHAA/AA	2.53	2.95	4.58	6.97
DHAA%	19.6	20.4	20.9	22.8

FIGURE 3 | The effects of protein fusions on production of DHAA. DBR2 is individually over-expressed both in pRS425 and genome in strain Sc470, which can be taken as the control to exclude the effect of overexpression of DBR2 on the production of DHAA. Proportion of DHAA (DHAA%) was calculated as the production of DHAA versus the sum production of DHAA, AA, AD, AO, and DHAO.

(rigid or flexible) (Lu and Feng, 2008; Chen et al., 2012) and fusion orientation (forward or reverse) (Zhang et al., 2017) could significantly affect the performance of fusion proteins. Borneman and colleagues (Lee et al., 2016) once pointed out that fusing a coumarate-CoA ligase (4CL) with benzalacetone synthase from *Rheum palmatum* (RpBAS) in the 4CL-RpBAS orientation gave rise to significant improvement on final raspberry ketone levels, but the reverse version (RpBAS-4CL) did not work. Therefore, forward and reverse fusions of ADH1 and DBR2 with rigid and flexible linkers (**Supplementary Table S2**) were constructed and the combined effects on the DHAA output was investigated. As shown in **Supplementary Figure S2**, it was observed that reverse fusion of ADH1 and DBR2 could significantly increase the ratio of DHAA to AA. To be noted, reverse fusion of ADH1 and DBR2 with rigid linker achieved the highest ratio of DHAA to AA at 2.95, obtaining strain Sc115 (**Figure 3** and **Supplementary Figure S2**). While forward fused protein ADH1-DBR2 tended to accumulate more AA than that of reverse fused protein. To further adjust the ratio of DHAA to AA, increasing the substrate accessibility of DHAO by ALDH1 should also be considered. Based on the result from the fusion of ADH1 and DBR2, the N-terminal of DBR2 should be exposed (**Supplementary Figure S2**). Thus, the fusion direction here was set as DBR2-ALDH1. Rigid and flexible linkers were also chosen to construct the fusion protein. Strains harboring the fusion protein DBR2-ALDH1 obtained much higher ratio of DHAA to AA compared that in control strain (Sc085) (**Supplementary Figure S3**). It was also shown that the fusion protein with a rigid linker got the highest ratio of DHAA to AA at 4.58 (**Figure 3** and **Supplementary Figure S3**). Consequently, in order to promote flux from AOH to DHAA, the fusion proteins DBR2-ADH1 and DBR2-ALDH1 were co-expressed to form the strain Sc361 which resulted in a lower production of AA with no concomitant increase in DHAA (323.2 mg/L) and raised the DHAA/AA to 6.76 (**Figure 3**).

Since there is a high accumulation of AD (about 1.05–1.20 g/L) and no accumulation of any other intermediates (including AO, DHAO, AOH) in all strains described above (**Figure 3** and **Supplementary Figures S2, S3**), the oxidation of AD was also the final step of biosynthesis pathway for all strains, and AOH could be converted to the end product AA or DHAA in time. Additionally, if comparing the effect of the fusion proteins DBR2-ADH1 and DBR2-ALDH1 on the DHAA/AA ratio, DBR2-ALDH1 exhibited a more important role in tailoring DHAA/AA ratio (from 2.53 to 4.58), while the other only increased DHAA/AA ratio to 2.95. Although the sum of the product was reduced, combining two fusion proteins would slightly increase the conversion yield of DHAA transformed from AD [proportion of DHAA (DHAA%) increased from 19.7 to 22.8%]. All of the above data demonstrated that engineering minimal conversion of AO to AA is crucial to increase the DHAA/AA ratio and slightly increase the DHAA%. Thus, in order to get optimal ratio of DHAA to AA, reducing the catalysis preference of the shared enzyme ALDH1 to AO or enhancing the catalysis preference to that of DHAO to smooth flux toward DHAA seems to be especially important, in addition to protein fusion.

Improving the Ratio of DHAA/AA by Changing ALDH1 Specificity

ALDH1 can employ AO or DHAO as the substrate and may have a slightly stronger binding affinity to AO, which will potentially restrain the production of DHAA (**Supplementary Table S1**). Herein, we attempted to change the substrate specificity of ALDH1 by rationally modifying its structure to improve the preference to DHAO.

Resorting to the molecular structure, DHAO and AO structures are very similar. They both have the aldehyde group which will then be oxidized to the carboxyl group. Nevertheless, the AO molecule has an additional vinyl group adjacent to the aldehyde group (**Supplementary Figure S5**) that forms a conjugated structure. Such conjugated structure in AO will affect the electronic cloud distribution of the catalyzed aldehyde group. Therefore, the structure difference of the catalyzed site of substrates would cause distinct binding force of the shared ALDH1 to DHAO or to AO. By analyzing the complex structure of ALDH1 with its cofactor NAD^+ and substrates, it was found that the nicotinamide group of NAD^+ just bound to the catalytic site of substrates in the active pocket of ALDH1 (**Supplementary Figure S5**). Consequently, a hypothesis was proposed that the binding characteristic of NAD^+ to DHAO or to AO potentially made the crucial contributions to the differences in ALDH1's affinities to its substrates. With this regard, the structural modifications of ALDH1 were carried out in two aspects. Firstly, the binding pocket, which for both NAD^+ and substrate, was downsized to shorten the distance between the nicotinamide group of NAD^+ and the catalyzed aldehyde group of DHAO. Secondly, the binding force of ALDH1 and NAD^+

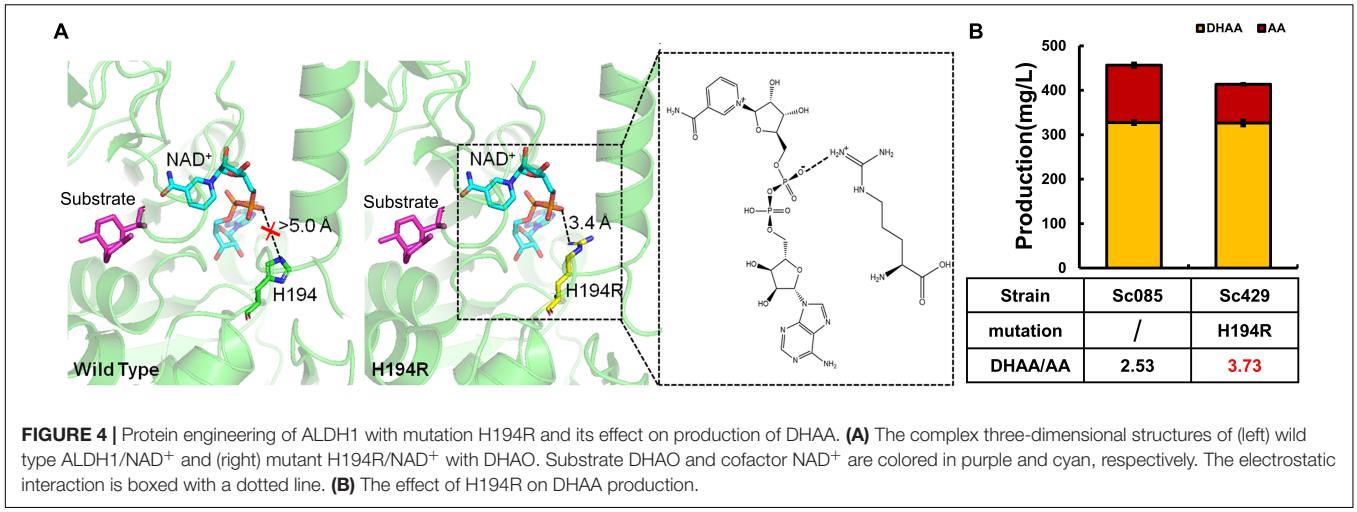


FIGURE 4 | Protein engineering of ALDH1 with mutation H194R and its effect on production of DHAA. **(A)** The complex three-dimensional structures of (left) wild type ALDH1/NAD^+ and (right) mutant H194R/NAD^+ with DHAO. Substrate DHAO and cofactor NAD^+ are colored in purple and cyan, respectively. The electrostatic interaction is boxed with a dotted line. **(B)** The effect of H194R on DHAA production.

was enhanced to improve the complex structural stability and enhance the enzymatic catalytic activity to DHAO, since NAD^+ is loosely bound to ALDH1 (Butterworth, 2010).

Accordingly, an attempt was made to replace the residues G223 and G227 with branch chains of a larger size (like V or F) by squeezing them through the space within the binding pocket. However, by further comparing the complex structures between the wild type ALDH1/NAD^+ and mutant (G223V/F, G227V/F)/NAD^+, it was found that although the binding pocket structure of the mutant complex was downsized, the adenine group of NAD^+ was extruded out of the original position, due to the steric effect of large branch chains of the mutated residues (**Supplementary Figure S6**). This structure weakened the binding force between NAD^+ and ALDH1 mutants (**Supplementary Figure S6**), probably damaging or even destroying the enzymatic activity of ALDH1 mutants. As expected, either the ratio of DHAA/AA or the DHAA titer were dramatically decreased by mutagenesis of G223F and G227V/F (**Supplementary Figure S6F**), compared to that of the wild type ALDH1. Although G223V increased the ratio of DHAA/AA, the actual production of DHAA was reduced significantly (**Supplementary Figure S6F**). The significant accumulation of DHAO was found in all mutants at G223 and G227, showing that the activity of these ALDH1 mutants were reduced. Although CYP71AV1 was reported to transform AO to AA (Teoh et al., 2009), its activity might not be strong enough to produce more AA or DHAA.

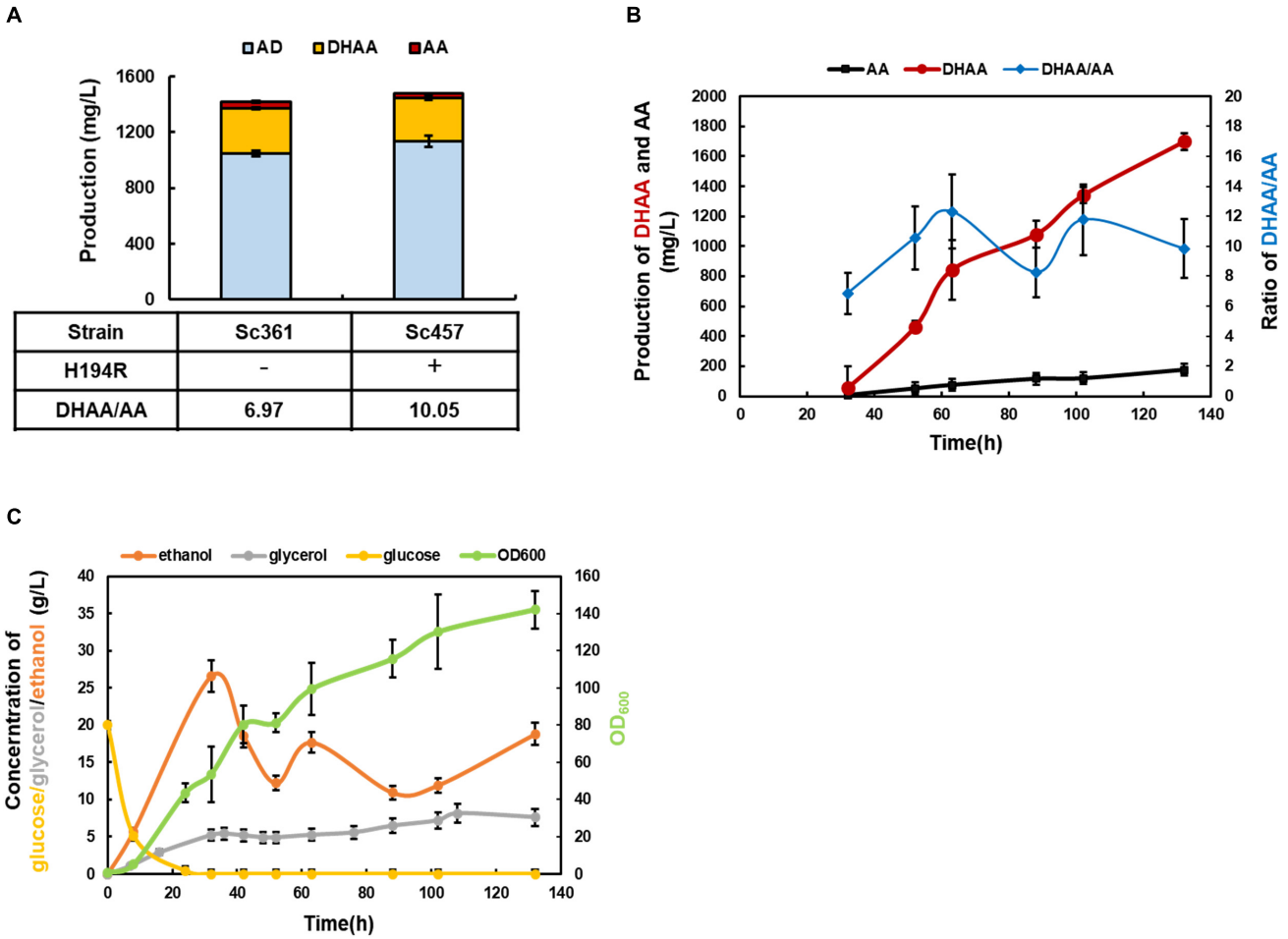
Alternatively, we sought to enhance the binding between ALDH1, NAD^+ and the substrate. The site-directed mutant H194R was constructed to enhance the binding of enzyme to NAD^+. Comparing with the complex structures of wild type ALDH1/NAD^+, the mutant H194R formed an additional electrostatic interaction with the phosphate group of NAD^+ (**Figure 4A**). Such electrostatic interactions would improve the structural stability of the mutant complex, which would significantly regulate the enzymatic catalysis activity to substrate (AO and DHAO).

Consequently, H194R significantly increased the ratio of DHAA/AA to 3.73 without reducing the production of DHAA (**Figure 4**). Different from other mutants in G223 or G227, the strain with mutant H194R (Sc429) didn't accumulate DHAO and any other intermediates (AO, AOH) in the product. H194R seemed to affect the activity of ALDH1 with AO and reduced the production of AA (see **Supplementary Figure S6F**). In summary, our experimental results were well-consistent with our modeling based on structural analysis (**Figure 4**).

DHAA Enrichment by 5 L Fed-Batch Fermentation

Mutated ALDH1(H194R) was introduced into strain Sc361 to replace the wild type of ALDH1 existing in fusion protein DBR2-ALDH1, obtaining strain Sc457. As shown in **Figure 5A**, by further reducing the production of AA, the highest ratio of DHAA/AA at 10.05 was achieved at shake flask level. The DHAA titer was maintained at 311 mg/L without accumulation of AOH, AO, or DHAO (**Figure 5** and **Supplementary Figure S4**). Taking the DHAA/AA ratio and the DHAA titer together into considerations, strain Sc457 was chosen for the following fermentation process.

In order to promote DHAA production and investigate the performance of strain Sc457 in fermenters, 5 L bioreactor fermentation was conducted using the carbon restriction strategy. As shown in **Figure 5B**, OD_{600} increased sharply before 32 h when glucose was utilized as a carbon source. When the carbon source was switched to ethanol, the OD_{600} was gradually promoted to 142 until harvest. Under above control process, the highest reported DHAA titer at 1.70 g/L in microbes was accomplished after 132 h fermentation (**Figure 5B**). A slight accumulation of AA was also obtained. It was also observed that the average ratio of DHAA to AA during the fermentation process was quite steady, demonstrating that the genetic stability of strain Sc457 was quite stable and the process was scalable to mimic the strain performance in shake flask to some extent. As the current DHAA production is not as high as that of AA in Amyris, Inc. (Paddon et al., 2013), further optimization of the fermentation efficiency via integrating medium and feed strategy (Westfall et al., 2012) as well as off-gas analysis feedback control could be tried.



Strain	Sc361	Sc457
H194R	-	+
DHAA/AA	6.97	10.05

FIGURE 5 | The characterization of the final strain Sc457 and related bioprocess development in 5 L bioreactor. **(A)** The DHAA production in strain Sc457 in which mutated ALDH1(H194R) was introduced into strain Sc361 to replace the wild type of ALDH1 existing in fusion protein DBR2-ALDH1. **(B)** The DHAA production and DHAA/AA ratio of Sc457 in 5 L bioreactor fermentation during the time course. **(C)** The concentrations of major metabolites (glucose, ethanol, glycerol, acetate) during the fermentation.

During the fermentation process, a significant accumulation of glycerol was also found at about 5 g/L in glucose consumption stage. Ho and colleagues reported both UBR2 and GUT1 with single point mutation could be regarded as targets for establishing glycerol utilization in strains of the CEN.PK family (Ho et al., 2017). It was also found that regulatory and metabolic trade-offs of glycerol utilization in *S. cerevisiae* were revealed by laboratory evolution, which could guide us to enhance glycerol consumption rationally in our study (Strucko et al., 2018). Furthermore, no acetate was detected during the whole process due to carbon restriction strategy (data not shown).

CONCLUSION

The combination of metabolic engineering and protein engineering has shown great performances in constituting microbial cell factories for the production of natural products with complex structures. In this study, fusion proteins and modifying ALDH1 catalysis preference strategies have been adopted and conducted to successfully tailor the ratio of DHAA/AA in *S. cerevisiae*. Promoted flux toward DHAA *via* ADH1, DBR2 and ALDH1 was firstly reconstituted by pairing fusion proteins DBR2-ADH1 and DBR2-ALDH1. The theoretical model of enhancing the stability of the enzyme/cofactor complex was assumed and executed to switch the catalysis preference of ALDH1 toward DHAA. Consequently, the ratio of DHAA/AA was elevated from 2.53 to 10.05 with the highest DHAA titer reaching 1.70 g/L (DHAA/AA ratio of 9.84) in 5 L bioreactor fermentation. This study shows the potential of oriented arrangement of consecutive heterologous enzymes in the reconstitution of microbial cell factories. On the other hand, the reduced sum flux of the biosynthesis pathway might be enhanced by other strategies of metabolic engineering which will be studied in future work.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the ADS (ACCESSION Q9AR04), CYP71AV1 (ACCESSION Q1PS23), DBR2 (ACCESSION KC505370.1), ALDH1 (ACCESSION JQ609276.1), ADH1 (ACCESSION JF910157.1), CYB5 (ACCESSION JQ582841.1), and CPR1 (ACCESSION DQ318192.1).

AUTHOR CONTRIBUTIONS

B-XZ and W-HX conceived of the study and participated in fed-batch fermentation. B-XZ and YW participated in strain construction and carried out the molecular genetic studies. M-DY carried out the protein analysis. M-DY, YW, and W-HX participated in design and coordination of the study as well as helped to draft the manuscript. W-HX supervised the whole research and revised the manuscript. All the authors read and approved the final manuscript.

FUNDING

This work was financially supported by National Natural Science Foundation of China (21621004, 21676192, and 31600052) and The National Key Research and Development Program of China (2018YFA0900702).

ACKNOWLEDGMENTS

The authors are grateful to thank Dr. Jian zha and Dr. Xia Wu from Shaanxi University of Science & Technology for their kind discussion.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00152/full#supplementary-material

REFERENCES

Butterworth, P. J. (2010). *Lehninger: Principles of Biochemistry*, 4th edn. New York, NY: Freeman & Co, 293–294.

Chen, X., Zaro, J. L., and Shen, W. C. (2012). Fusion protein linkers: property, design and functionality. *Adv .Drug Deliv .Rev.* 65, 1357–1369. doi: 10.1016/j.addr.2012.09.039

Chen, X., Zhang, C., and Too, H. P. (2017). Multienzyme biosynthesis of dihydroartemisinic acid. *Molecules* 22:1422. doi: 10.3390/molecules22091422

Chunyin, Z., and Cook, S. P. (2012). A concise synthesis of (+)-artemisinin. *J. Am. Chem. Soci.* 134, 13577–13579. doi: 10.1021/ja3061479

Delano, W. L. (2010). *The PyMol Molecular Graphics System*. New York:Schrodinger, LLC.

Entian, K. D., and Kötter, P. (2007). 25 yeast genetic strain and plasmid collections. *Methods Microbiol.* 36, 629–666. doi: 10.1016/s0580-9517(06)36025-4

Ho, P. W., Swinnen, S., Duitama, J., and Nevoigt, E. (2017). The sole introduction of two single-point mutations establishes glycerol utilization in Saccharomyces cerevisiae CEN.P*K* derivatives. *Biotechnol. Biofuels* 10:10. doi: 10.1186/s13068-016-0696-6

Irwin, J. J., Teague, S., Mysinger, M. M., Bolstad, E. S., and Coleman, R. G. (2012). ZINC: a free tool to discover chemistry for biology. *J. Chem. Inform. Model.* 52, 1757–1768. doi: 10.1021/ci3001277

Lee, D., Lloyd, N. D. R., Pretorius, I. S., and Borneman, A. R. (2016). Heterologous production of raspberry ketone in the wine yeast Saccharomyces cerevisiae via pathway engineering and synthetic enzyme fusion. *Microbial Cell Factories* 15:49. doi: 10.1186/s12934-016-0446-2

Lu, P., and Feng, M. G. (2008). Bifunctional enhancement of a β-glucanase-xylanase fusion enzyme by optimization of peptide linkers. *Appl. Microbiol. Biotechnol.* 79, 579–587. doi: 10.1007/s00253-008-1468-4

Miles, E. W., Rhee, S., and Davies, D. R. (1999). The molecular basis of substrate channeling. *J. Biol. Chem.* 274, 12193–12196. doi: 10.1074/jbc.274.18.12193

Oleg, T., and Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Computa. Chem.* 31, 455–461. doi: 10.1002/jcc.21334

Paddon, C. J., and Keasling, J. D. (2014). Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat. Rev. Microbiol.* 12:355. doi: 10.1038/nrmicro3240

Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., et al. (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496:528. doi: 10.1038/nature12051

Peng, B., Plan, M. R., Carpenter, A., Nielsen, L. K., and Vickers, C. E. (2017). Coupling gene regulatory patterns to bioprocess conditions to optimize synthetic metabolic modules for improved sesquiterpene production in yeast. *Biotechno. Biofuels* 10, 43. doi: 10.1186/s13068-017-0728-x

Peplow, M. (2013). Malaria drug made in yeast causes market ferment. *Nature* 494:160. doi: 10.1038/494160a

Pham, V. D., Lee, S. H., Si, J. P., and Hong, S. H. (2015). Production of gamma-aminobutyric acid from glucose by introduction of synthetic scaffolds between isocitrate dehydrogenase, glutamate synthase and glutamate decarboxylase in recombinant *Escherichia coli*. *J. Biotechnol.* 207, 52–57. doi: 10.1016/j.jbiotec.2015.04.028

Pieber, B. U., Glasnov, T., and Kappe, C. O. (2015). Continuous flow reduction of artemisinic acid utilizing multi-injection strategies-closing the gap towards a fully continuous synthesis of antimalarial drugs. *Chemistry* 21, 4368–4376. doi: 10.1002/chem.201406439

Strucko, T., Zirngibl, K., Pereira, F., Kafkia, E., Mohamed, E. T., Rettel, M., et al. (2018). Laboratory evolution reveals regulatory and metabolic trade-offs of glycerol utilization in Saccharomyces cerevisiae. *Metab. Eng.* 47:S1096717617304032. doi: 10.1016/j.ymben.2018.03.006

Su, W., Xiao, W. H., Wang, Y., Liu, D., Zhou, X., and Yuan, Y. J. (2015). Alleviating redox imbalance enhances 7-dehydrocholesterol production in engineered *Saccharomyces cerevisiae*. *Plos One* 10:e0130840. doi: 10.1371/journal.pone.0130840

Sweetlove, L. J., and Fernie, A. R. (2018). The role of dynamic enzyme assemblies and substrate channelling in metabolic regulation. *Nat. Commun.* 9:2136. doi: 10.1038/s41467-018-04543-8

Teoh, K. H., Polichuk, D. R., Reed, D. W., and Covello, P. S. (2009). Molecular cloning of an aldehyde dehydrogenase implicated in artemisinin biosynthesis in *Artemisia annua*. *Bot. Bot.* 87, 635–642. doi: 10.1139/b09-032

Turconi, J., Griolet, F., Guevel, R., Oddon, G., Villa, R., Geatti, A., et al. (2014). Correction to semisynthetic artemisinin, the chemical path to industrial production. *Organ. Process Res. Dev.* 18, 831–831. doi: 10.1021/op5001397

Westfall, P. J., Pitera, D. J., Lenihan, J. R., Diana, E., Woolard, F. X., Rika, R., et al. (2012). Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proc. Natl. Acad. Sci. U.S.A.*109, 655–656. doi: 10.1073/pnas.1110740109

Zhang, Y., Teoh, K. H., Reed, D. W., Maes, L., Goossens, A., Olson, D. J., et al. (2008). The molecular cloning of artemisinic aldehyde Delta11(13) reductase and its role in glandular trichome-dependent biosynthesis of artemisinin

in *Artemisia annua*. *J. Biol. Chem.* 283, 21501–21508. doi: 10.1074/jbc.M803090200

Zhang, Y., Wang, Y., Wang, S., and Fang, B. (2017). Engineering bi-functional enzyme complex of formate dehydrogenase and leucine dehydrogenase by peptide linker mediated fusion for accelerating cofactor regeneration. *Engi. Life Sci.* 17, 989–996. doi: 10.1002/elsc.201600232

Zhang, Y. H. P. (2011). Substrate channeling and enzyme complexes for biotechnological applications. *Biotechnol. Adv.* 29, 715–725. doi: 10.1016/j.biotechadv.2011.05.020

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

BRIEF RESEARCH REPORT
published: 20 March 2020
doi: 10.3389/fbioe.2020.00214

Metabolic Engineering of Histidine Kinases in *Clostridium beijerinckii* for Enhanced Butanol Production

*Xin Xin, Chi Cheng, Guangqing Du, Lijie Chen and Chuang Xue**

School of Bioengineering, Dalian University of Technology, Dalian, China

Clostridium beijerinckii, a promising industrial microorganism for butanol production, suffers from low butanol titer and lack of high-efficiency genetical engineering toolkit. A few histidine kinases (HKs) responsible for Spo0A phosphorylation have been demonstrated as functionally important components in regulating butanol biosynthesis in solventogenic clostridia such as *C. acetobutylicum*, but no study about HKs has been conducted in *C. beijerinckii*. In this study, six annotated but uncharacterized candidate HK genes sharing partial homologies (no less than 30%) with those in *C. acetobutylicum* were selected based on sequence alignment. The encoding region of these HK genes were deleted with CRISPR-Cas9n-based genome editing technology. The deletion of *cbei2073* and *cbei4484* resulted in significant change in butanol biosynthesis, with butanol production increased by 40.8 and 17.3% (13.8 g/L and 11.5 g/L vs. 9.8 g/L), respectively, compared to the wild-type. Faster butanol production rates were observed, with butanol productivity greatly increased by 40.0 and 20.0%, respectively, indicating these two HKs are important in regulating cellular metabolism in *C. beijerinckii*. In addition, the sporulation frequencies of two HKs inactivated strains decreased by 96.9 and 77.4%, respectively. The other four HK-deletion (including *cbei2087, cbei2435, cbei4925,* and *cbei1553*) mutant strains showed few phenotypic changes compared with the wild-type. This study demonstrated the role of HKs on sporulation and solventogenesis in *C. beijerinckii*, and provided a novel engineering strategy of HKs for improving metabolite production. The hyper-butanol-producing strains generated in this study have great potentials in industrial biobutanol production.

Keywords: *Clostridium beijerinckii*, CRISPR-Cas9n, histidine kinases, butanol, sporulation

OPEN ACCESS

Edited by:
Yi Wang,
Auburn University, United States

Reviewed by:
Petra Patakova,
University of Chemistry and Technology Prague, Czechia
Zhiqiang Wen,
Nanjing University of Science and Technology, China
Fengxue Xin,
Nanjing Tech University, China

****Correspondence:***
Chuang Xue
xue.1@dlut.edu.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *14 January 2020*
Accepted: *03 March 2020*
Published: *20 March 2020*

Citation:
Xin X, Cheng C, Du G, Chen L and Xue C (2020) Metabolic Engineering of Histidine Kinases in Clostridium beijerinckii for Enhanced Butanol Production. Front. Bioeng. Biotechnol. 8:214. doi: 10.3389/fbioe.2020.00214

INTRODUCTION

During the past decades, with concerns about diminishing petroleum reserves and fluctuations in oil prices, renewable biofuels have gained intensive attentions. As a substitute for gasoline, butanol has promising physical features such as high energy density, low volatility, and less corrosivity, but clostridia-based biobutanol production is economically unfavorable due to low product titer and productivity (Xue et al., 2013; Yang et al., 2018). *Clostridium beijerinckii* is an important industrial microorganism which produces butanol, but is notoriously difficult to metabolically engineer and hard to break through the limitation of low product concentration (Xue et al., 2017a). Since CRISPR-Cas9 has been explored as a powerful and effective tool for genome editing of lots

of organisms including *C. beijerinckii* (Li et al., 2016; Wang et al., 2016a), we expected to focus on editing the genome of *C. beijerinckii* based on this system.

Clostridium beijerinckii is an organism historically used for ABE (acetone, butanol, ethanol) fermentation. Batch ABE fermentation is characterized by two distinctive phases, acidogenesis and solventogenesis (Herman et al., 2017). During exponential growth, short-chain fatty acids, including acetic acid and butyric acid, are produced and accumulated in the system, which causes a drop in the culture pH. As the culture reaches a low enough pH, the formed acids can be re-assimilated and the culture pH rises. At the same time, solvent production is initiated (Li et al., 2011). The metabolic transition from acidogenesis to solventogenesis is consistent with the initiation of the complex sporulation process. During sporulation, a starch-like carbohydrate called granulose accumulates in the form of a swollen, bright-phase *Clostridium* bacterium, in which the endospore begins to develop. Further morphological development produces free-spores, heat- and chemical-resistant cell types that do not contribute to solvent production (Al-Hinai et al., 2015; Herman et al., 2017).

Histidine kinases (HKs) are involved in perceiving and transducing environmental signals to trigger multiple cellular responses, such as cell division (Jacobs et al., 1999), nitrogen metabolism (Loomis et al., 1997), antibiotic resistance (Wilke et al., 2015) and sporulation (Xue and Cheng, 2019). Orphan HKs are HKs lacking an adjoining response regulator, as is typical for most HKs in prokaryotes (Steiner et al., 2011). Several orphan HKs, regulating spore development in *C. perfringens* (Hiscox et al., 2013), *C. difficile* (Underwood et al., 2009), *C. botulinum* (Wörner et al., 2006) and *C. thermocellum* (Mearls and Lynd, 2014), etc., have been identified. Especially, orphan HKs directly phosphorylate Spo0A, which is the master regulator of both sporulation and solventogenesis in *C. acetobutylicum* (Steiner et al., 2011). Spo0A directly controls the genes with the presence of one or more "0A boxes" (TGNCGAA) in their 5′ regulatory regions (Wilkinson et al., 1995). The "0A boxes" are present in the upstream regions of many genes involving sporulation such as *sigE* and *spoIIE*, and central metabolism such as *adc*, *sol* operon, *bdhA* and *bdhB* (Wilkinson et al., 1995; Kolek et al., 2017). To regulate the transcription of these genes, the phosphorylated Spo0A will bind to the "0A boxes" to function. In *C. acetobutylicum,* there are two pathways for Spo0A activation: one depends on a HK Cac0323, and the other involves two HKs Cac0903 and Cac3319, respectively (Steiner et al., 2011). Individual mutants of the three HKs inactivation showed that the sporulation frequency was decreased by 95–99% compared to the wild-type strain (Steiner et al., 2011), indicating these HKs have similar characteristics and are essential for spore development. More importantly, inactivation of Cac3319 in *C. acetobutylicum* was found to give positive phenotypic changes, including increased butanol tolerance and production (Xu et al., 2015). However, no relevant study has been performed in *C. beijerinckii*, another important butanol-producing strain that could more preferably utilize lignocellulosic hydrolysate (Xiao et al., 2012). There are 85 HKs in the genome of *C. beijerinckii* and only one of them has been investigated to explore its regulatory function. The three-component system consists of HK LytS and response regulator YesN, which directly regulates the transcription of *xylFGH* genes and is responsible for xylose transportation (Sun et al., 2015). However, whether orphan HKs could regulate sporulation and solventogenesis in *C. beijerinckii*, as they did in *C. acetobutylicum*, was still unknown. The genome sequence data indicated that clostridia largely shared the key HK genes for phosphorylation of Spo0A (Al-Hinai et al., 2015), and thus we hypothesized that analogous HKs existed and were functional in *C. beijerinckii*.

In this study, we reported the identification and verification of orphan HKs in *C. beijerinckii* for their regulatory function in solventogenesis and sporulation. Six HK candidates were identified by sequence alignment with *cac3319*, *cac0323,* and *cac0903* in *C. acetobutylicum*. Using CRISPR-Cas9n system, we successfully deleted these HK genes, and fermentation analysis showed that two of them were important for regulating butanol biosynthesis and spore development. This study demonstrated the role of HKs on sporulation and solventogenesis in *C. beijerinckii* for the first time. The HK-engineered strains can be used as robust workhorses for enhanced butanol production, and also can be further engineered for production of valuable metabolites.

MATERIALS AND METHODS

Bacterial Strains, Culture Conditions, and Plasmids

All bacterial strains and plasmids used in this study were listed in **Supplementary Table S1**. The *Escherichia coli* DH5α was used for plasmid cloning. The *E. coli* transformants were grown aerobically at 37°C in Luria-Bertani (LB) medium or on solid LB agar (2% w/v) plate supplemented with ampicillin (100 μg/mL) when necessary. *C. beijerinckii* CC101 was an adaptive mutant of *C. beijerinckii* NCIMB 8052 (ATCC 51743) from ST Yang's lab (Lu et al., 2013), and its mutant strains were grown anaerobically at 37°C in Clostridium Growth Medium (CGM) (Li et al., 2016) supplemented with erythromycin (40 μg/mL) as necessary.

Reagents and Enzymes

All restriction enzymes used in this study were purchased from New England Biolabs (Beverly, MA, United States). The DNA polymerase KOD FX (Toyobo, Osaka, Japan) was used for DNA amplification and colony extension PCR to screen the positive transformants. Recombinant plasmids were assembled through the ClonExpress One Step Cloing Kit (Vazyme Biotech, Nanjing, China).

Plasmid Construction

All the DNA oligonucleotides used in this study were synthesized in Sangon Biotech Co., Ltd. (Shanghai, China) and listed in **Supplementary Table S2**. The schematic diagram of the plasmid construction and gene-editing process was shown in **Supplementary Figure S1**. CRISPR-based genome editing plasmid, pNICKclos 2.0-*xylR* (Li et al., 2016), was used as the vector to delete HK genes. As the selected target region (N20-NGG) was crucial for sgRNA targeting, alignment research (NCBI BLAST) was performed to ensure that the 23-bp

target sequence had no sequence similarity elsewhere in the genome and thus could avoid multiple off-target candidates. The sgRNA-encoding region targeting the *cbei2073* was placed following the Pj23119 synthetic promoter through the steps described below. The primers pNICKclos-2073-1-1/pNICKclos-2073-3 were used to generate a Pj23119-sgRNA-2073-1 cassette from pNICKclos 2.0-*xylR* and then the cassette was used as the template to PCR amplify the Pj23119-sgRNA-2073-2 cassette with primers pNICKclos-2073-1-2/pNICKclos-2073-3. The 1.2-kb upstream and 1.2-kb downstream homology arms (HAs) of the selected target region (N20-NGG) in *cbei2073* were amplified from the genome of *C. beijerinckii* CC101 using primers pairs pNICKclos-2073-2/pNICKclos-2073-5 and pNICKclos-2073-4/pNICKclos-2073-6, respectively. Then the Pj23119-sgRNA-2073-2 cassette and the two HAs were joined by overlap extension PCR using primers pNICKclos-2073-1-2/pNICKclos-2073-6, generating a fragment in which the two HAs were separated by a *Pst*I restriction site. Finally, the fragment was fused with the *Spe*I/*Xho*I linearized pNICKclos 2.0-*xylR* plasmid through the One Step Cloning Kit, yielding pNICKclos 2.0-*cbei2073*. All the other plasmids (pNICKclos 2.0-*cbei2087*, pNICKclos 2.0-*cbei2435*, pNICKclos 2.0-*cbei1553*, pNICKclos 2.0-*cbei4925*, and pNICKclos 2.0-*cbei4484*) used for gene deletion based on CRISPR-Cas9 nickase in *C. beijerinckii* CC101 were also derived from pNICKclos 2.0-*xylR* using similar methods as described above.

Transformation of *C. beijerinckii*

Plasmids for genome editing were transformed into *C. beijerinckii* via electroporation (Mermelstein et al., 1992) according to the following protocol. To prepare electrocompetent cells, 2 mL stock culture in 20% glycerol was inoculated into 100 mL CGM medium until the optical density at 600 nm (OD_{600}) reached 0.4–0.6. All of the following steps were done at low temperatures (4°C) and the buffers used were pre-cooled. The culture was centrifuged at 4500 r/min for 10 min and then the cells were resuspended in 30 mL ETM buffer (270 mM sucrose, 4.4 mM NaH_2PO_4, 0.6 mM Na_2HPO_4, 10 mM $MgCl_2$, pH 7.4) for 10 min. The resuspended cells were centrifuged and resuspended in 2.5 ml ET buffer (270 mM sucrose, 4.4 mM NaH_2PO_4, 0.6 mM Na_2HPO_4, pH 7.4). The mixture of 190 μL competent cells and 10 μL plasmid DNA (1–2 μg) was transferred to an electroporation cuvette with a 0.2 cm gap width. Electroporation was performed using a Bio-Rad Micropulser (Bio-Rad Laboratories, Hercules, CA, United States) at 1.8 kV. After electroporation, samples were mixed with 400 μL CGM immediately and transferred to another 400 μL CGM in a 2-mL tube. After incubation for 4–6 h at 37°C, 200 μL recovered cultures were plated on CGM agar containing 40 μg/mL erythromycin and grown anaerobically at 37°C for 48 h. Transformant colonies were picked from the agar plates and subjected to PCR-based verification.

FIGURE 1 | Comparison of HK phosphodonor active site domains. The DHp/HisKA domains of several putative orphan HKs in *C. beijerinckii* were compared to known orphan kinases Cac3319, Cac0903, and Cac0323 (*C. acetobutylicum*), respectively. Identical residues conserved with Cac3319 or Cac0903 and Cac0323 are shaded black, identical residues conserved with Cac0903 are shaded dark gray, and identical residues conserved with Cac0323 are shaded light gray. The phosphorylated histidine residue is denoted by an asterisk.

Mutant Screening

Individual transformant colonies were all picked and analyzed using colony-PCR after electroporation. Then the PCR productions were digested by restriction endonuclease and sequenced to confirm the positive transformant colonies. The primers used were selected from sequence located about 100 bp up- and down-stream of the HAs on the genome named *cbei2073*-For/*cbei2073*-Rev, etc. The restriction site (*Pst*I) between the two HAs in the editing template was confirmed by cleavage of the fragment amplified by colony extension PCR with *Pst*I restriction enzyme. Parent strain *C. beijerinckii* CC101 genome was used as a negative control.

Plasmid Curing

To remove the pNICKclos 2.0 plasmids from mutants, the positive transformants were cultured in 5 mL of CGM liquid medium (M1) without antibiotic pressure at first. After growing for 12 h, 50 μL of M1 was inoculated into 5 mL CGM medium without antibiotic (M2). After another 12 h, the OD_{600} of M2 increased to ∼0.8. Then 100 μL of the M2 broth was spread on a CGM medium ager plate without antibiotic. After 12 h, the individual colonies were spread onto CGM ager with erythromycin (50 μg/mL) and CGM ager without antibiotic. The cells that grew on the later but could not grow on the former were considered to have lost the plasmid.

Batch Fermentation

Batch fermentations with various engineered *C. beijerinckii* strains were performed anaerobically in P2 medium in a 3-L bioreactor with 1.5-L working volume. The P2 medium was prepared as previously described (Xue et al., 2016), which contained glucose ∼90 g/L; yeast extract 1 g/L; CH_3COONH_4 2.2 g/L; K_2HPO_4 0.5 g/L; KH_2PO_4 0.5 g/L; $MgSO_4{\cdot}7H_2O$ 0.2 g/L; $MnSO_4{\cdot}H_2O$ 0.01 g/L; $FeSO_4{\cdot}7H_2O$ 0.01 g/L; NaCl 0.01 g/L; para-amino-benzoic acid 1 mg/L; thiamin 1 mg/L; and biotin 1 mg/L. The *C. beijerinckii* strains were first incubated anaerobically in 100 mL of CGM at 37°C until OD_{600} reached ∼1.0 and then the seed culture was transferred into the bioreactor containing 1.4 L P2 medium. All the media used in this study were sterilized at 121°C and 15 psig for 15 min and purged with N_2 for 20 min through a sterile 0.2 μm filter to remove dissolved oxygen.

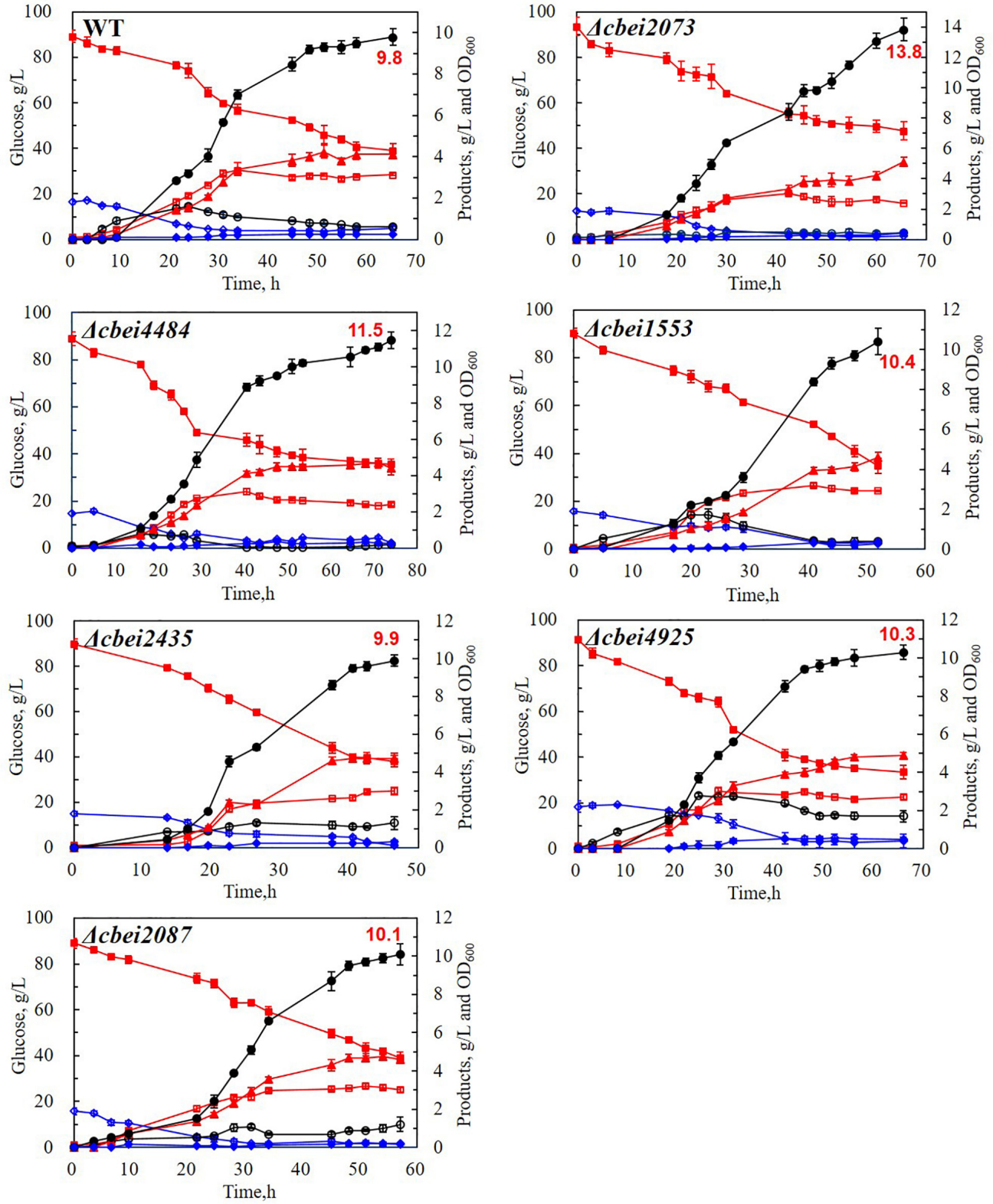


FIGURE 2 | Fermentation results of the wild-type (WT) and mutant strains in batch fermentation. Symbols: glucose (filled squares), OD_{600} (open squares), butanol (filled circles), acetone (filled triangles), ethanol (filled diamonds); butyrate (open circles), acetate (open diamonds).

TABLE 1 | Fermentation parameters of wild-type and mutant strains in batch fermentation.

Strains	$OD_{600(max)}$	Glucose consumption (g/L)	Acids (g/L)		Solvents (g/L)				Yield (g/g)		Productivity (g/L/h)	
			Acetic acid	Butyric acid	Acetone	Butanol	Ethanol	ABE	Butanol	ABE	Butanol	ABE
CC101	3.1 ± 0.1	50.1 ± 3.1	0.5 ± 0.1	0.6 ± 0.1	4.1 ± 0.2	9.8 ± 0.4	0.2 ± 0.0	14.1 ± 0.7	0.20 ± 0.02	0.28 ± 0.03	0.15 ± 0.02	0.22 ± 0.03
Δ*cbei2073*	3.0 ± 0.1	58.4 ± 4.2	0.2 ± 0.0	0.4 ± 0.0	5.1 ± 0.3	13.8 ± 0.8	0.4 ± 0.1	19.4 ± 0.7	0.24 ± 0.02	0.33 ± 0.02	0.21 ± 0.02	0.30 ± 0.03
Δ*cbei4484*	3.1 ± 0.2	53.5 ± 2.7	0.3 ± 0.0	0.2 ± 0.0	4.4 ± 0.4	11.5 ± 0.5	0.3 ± 0.0	16.1 ± 0.5	0.21 ± 0.01	0.30 ± 0.01	0.18 ± 0.01	0.25 ± 0.02
Δ*cbei2087*	3.2 ± 0.1	49.9 ± 3.3	0.2 ± 0.0	1.2 ± 0.4	4.6 ± 0.1	10.1 ± 0.6	0.2 ± 0.0	14.9 ± 0.6	0.20 ± 0.00	0.30 ± 0.00	0.18 ± 0.00	0.26 ± 0.00
Δ*cbei2435*	3.0 ± 0.2	51.7 ± 4.1	0.2 ± 0.0	1.3 ± 0.3	4.8 ± 0.3	9.9 ± 0.3	0.3 ± 0.1	15.0 ± 0.5	0.19 ± 0.01	0.29 ± 0.02	0.21 ± 0.02	0.32 ± 0.03
Δ*cbei1553*	3.2 ± 0.1	55.3 ± 4.4	0.3 ± 0.0	0.4 ± 0.0	4.6 ± 0.3	10.4 ± 0.7	0.3 ± 0.0	15.3 ± 0.4	0.19 ± 0.01	0.28 ± 0.02	0.20 ± 0.02	0.29 ± 0.02
Δ*cbei4925*	3.0 ± 0.2	57.7 ± 3.8	0.5 ± 0.1	1.7 ± 0.3	4.9 ± 0.1	10.3 ± 0.4	0.4 ± 0.1	15.6 ± 0.9	0.18 ± 0.01	0.27 ± 0.02	0.16 ± 0.00	0.24 ± 0.01

The results obtained were presented as the mean ± SD for triplicates of cultures.

Sporulation Frequency Assay

The strains were cultured in liquid CGM medium for 5 days. Then the cell cultures were treated in 80°C for 10 min, after which 100 μL of them were plated on CGM agar directly or following diluting for 10 times. The sporulation frequency of strains was characterized by the numbers of heat-resistant colonies.

Analytical Methods

The fermentation samples were periodically taken from the bioreactor and the cell growth was analyzed by measuring OD_{600} using a spectrophotometer (Thermo Spectronic, United States). Samples for glucose and product analyses were pelleted by centrifugation at 8000 × *g* for no less than 5 min. The solvents of the fermentation broth (acetone, butanol, and ethanol) were analyzed using a gas chromatograph (Agilent 6890A GC, United States) as previously described (Xue et al., 2017b; Yang et al., 2019). Glucose, acetate and butyrate were analyzed by HPLC (Waters 1525, United States) equipped with an Aminex HPX-87H column (300 mm × 7.8 mm) maintained at 50°C. Dilute H_2SO_4 (10 mM, 0.5 mL/min) was used as the eluent (Xiao et al., 2019).

RESULTS

Screening of Target Genes and Construction of Engineered Strains

To explore the uncharacterized HKs in *C. beijerinckii*, sequence alignments between the nucleotide sequence of three orphan HK genes in *C. acetobutylicum* (*cac3319*, *cac0323,* and *cac0903*) and the whole genome of *C. beijerinckii* were performed. Six candidates, including *cbei2073*, *cbei2087*, *cbei4925*, *cbei2435*, *cbei1553*, and *cbei4484*, were identified. Among them, *cbei2073* showed 46% homology with *cac3319*, and the others showed 31–36% homology with *cac0323* or *cac0903* gene. By further amino acid sequence alignment of HisKA domain, Cbei2073 showed 78% sequence similarity with Cac3319 and the others showed no less than 70% with Cac0903 or Cac0323. Also, Cbei2073 was special among the six in that its phosphorylation sites and their adjacent regions were highly conserved and similar to Cac3319, which has been proved to be the most functional HK in regulating butanol biosynthesis in *C. acetobutylicum* (**Figure 1**; Xu et al., 2015). Based on these analyses, these six HKs were finally selected as candidate homologous proteins.

After selecting these targeting genes, recombinant plasmids (pNICKclos 2.0-*cbei2073*, pNICKclos 2.0-*cbei2087*, pNICKclos 2.0-*cbei2435*, pNICKclos 2.0-*cbei1553*, pNICKclos 2.0-*cbei4925*, and pNICKclos 2.0-*cbei4484*) were first constructed to delete the corresponding gene, respectively. Then they were separately transferred into *C. beijerinckii* via electroporation. A low number of transformants could been identified after electroporation, but the verification results showed that the accuracy rates (number of correctly edited transformants/total number of transformants screened) were no less than 60% (**Supplementary Table S3**).

Impact of HK Genes on ABE Fermentation

To investigate the effect of HKs on solventogenesis and cell growth, batch fermentation performances of wild-type *C. beijerinckii* and HK-inactivated strains were compared (**Figure 2** and **Table 1**). Cbei2073 inactivated strain showed enhanced butanol production with 40.8 and 40.0% increases in butanol titer and butanol productivity, respectively, compared to wild-type *C. beijerinckii*. In addition, increases in ethanol and acetone were also observed in Cbei2073 inactivated strain. According to the fermentation kinetics study, it could be concluded that Cbei2073 played an important role in regulating solvent production, which was similar to HK Cac3319 of *C. acetobutylicum*. Cbei4484 inactivated strain showed enhanced butanol production with 17.3 and 20.0% increases in butanol titer and butanol productivity, respectively. Fermentation of Cbei1553 mutant strain showed that this HK had minor effects on cell growth and solvent production, but its inactivation resulted in 40.0% increase in butanol productivity. The ABE titers of Cbei2435, Cbei2087, and Cbei4925 inactivated strains were similar to that of the wild strain. However, Cbei2435 and Cbei2087 mutant strains showed slight increases in butyrate titer and significant increases in butanol productivity.

Effect of HKs on Sporulation Frequency

As spore-forming cells were metabolically inactive, it is hypothesized that the enhanced butanol production was resulted from reduced sporulation frequency caused by inactivation of HK. The heat-resistant colonies of all the six HK mutants and wild-type strain were tested. It was found that the heat-resistant colonies of four HK genes (*cbei2435*, *cbei1553*, *cbei2087,* and *Cbei4925*) deleted strains were in the same level with wild-type strain, which were much more than the other two mutants (*Δcbei2073* and *Δcbei4484*). Since *cbei2073* and *cbei4484* had obvious effect on butanol production, we compared the sporulation frequency of wild-type strain, *cbei2073* and *cbei4484* deleted strains by counting the number of heat-resistant colonies following incubation in liquid CGM medium for 5 days. As excepted, significant decreases of sporulation frequency were observed for two HKs inactivated strains, with spore formation reduced by 96.9 and 77.4% of the wild-type strain, respectively (**Figure 3**). It was indicated that these two HKs might have effects on sporulation and its related genes.

DISCUSSION

CRISPR-Cas9n based gene editing technology is a method which is not only effective but also widely used in bacteria including *C. acetobutylicum*, *C. beijerinckii*, *Bacillus licheniformis*, etc. (Wang et al., 2015, 2016a,b; Li et al., 2016, 2018; Zhang et al., 2018). Despite the wide application of this system, it remains reliant on homologous recombination (HR) (Li et al., 2016). In clostridia, HR is notoriously inefficient, which cannot be solved by CRISPR-Cas9 technologies alone (Bruder et al., 2016; Charubin et al., 2018). For this reason, the numbers of individual transformant colonies obtained after the transformation of plasmids into *C. beijerinckii* and the edited cells were both small (**Supplementary Table S3**). However, the verification results showed that the accuracy rates (60.0–100.0%) were relative higher than those (6.7–100.0%) obtained in a previous study (Li et al., 2016).

Orphan HKs in *Clostridium* species are reported to directly phosphorylate Spo0A, which is the master regulator for initiation of sporulation and solventogenesis (Steiner et al., 2011; Dürre, 2014). To explore the regulatory function of HKs on biobutanol synthesis, the 85 HK genes in *C. beijerinckii* were analyzed and compared with three HKs (Cac3319, Cac0903, and Cac0323) in *C. acetobutylicum* with elucidated functions. Cac3319 inactivated strains could produce 44.4% more butanol (Xu et al., 2015), indicating its vital role in regulating solvent production in *C. acetobutylicum*. In this study, Cac3319 and Cbei2073 were

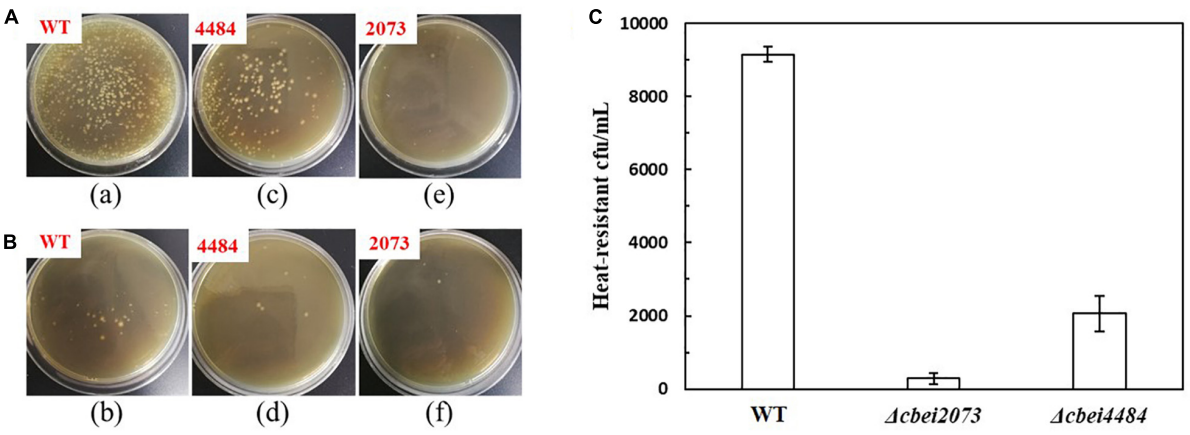


FIGURE 3 | Sporulation frequencies of the wild-type (WT) and mutant strains. The heat-resistant colony-forming units (CFU) were evaluated by treating the cell cultures (5 days) in 80°C for 10 min, after which 100 μL of them were plated on CGM agar directly **(A)** or following diluting for 10 times **(B)**. **(C)** The numbers of heat-resistant colonies.

similar based on the sequence alignment results. The deletion of *cbei2073* gene led to significant change in butanol biosynthesis with butanol production increased by 40.8%, indicating that Cbei2073 was also most effective in regulating biobutanol synthesis among the selected six HKs. However, some HKs including Cbei1553, Cbei4925, Cbei2435, and Cbei2087 seemed to have no apparent effect on ABE synthesis, and the growth of all the strains with inactivated HKs was not affected.

In *C. acetobutylicum*, sporulation is closely associated with solventogenesis, and both of them are controlled by Spo0A (Steiner et al., 2011). Cac3319, Cac0903, and Cac0323 are essential for Spo0A activation, which controls the initiation of sporulation and solventogenesis (Steiner et al., 2011; Xu et al., 2015). The "0A box" that can be bound with Spo0A∼P has been proved present in the upstream regions of lots of genes whose expressions are either activated or repressed at the initiation of solventogenesis in both *C. acetobutylicum* and *C. beijerinckii* (Wilkinson et al., 1995). Generally, the sporulating phenotype is considered to be profitable for solvent formation, but solvent production ceases when mature spores form (Kolek et al., 2016). In addition, overexpression of *spo0A* in *C. beijerinckii* NRRL B-598, sharing high genome homology with *C. beijerinckii* NCIMB 8052, led to cessation of production at a low ABE concentration, indicating that Spo0A is a master regulator of solventogenesis in *C. beijerinckii* (Kolek et al., 2017). Furthermore, without spore formation, *C. beijerinckii* NRRL B-598 could produce more butanol in RCM medium, compared with the sporulating phenotype (Branska et al., 2018). Therefore, we speculated that Cbei2073 and Cbei4484 played a role on phosphorylation of Spo0A due to their high similarity in sequence with Cac3319, Cac0903, and Cac0323 and further regulated both butanol synthesis and sporulation frequency. As shown in the results, the inactivation of the two HKs (Cbei2073 and Cbei4484) may lower Spo0A∼P level and then decrease sporulation frequency, finally leading to the improved butanol production. Selective inhibition of sporulation process was beneficial for improving the economics of ABE fermentation since solvent can be produced during a longer time frame (Cheng et al., 2019).

In conclusion, the results of the present study demonstrated that by deleting HK genes, the sporulation frequency might be decreased, with different degrees of improvements in butanol titers, productivities and yields (**Table 1**). Therefore, this study provided a novel strategy for promoting production of metabolites applicable for a broad of bacteria.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

CX developed the research scheme. XX performed the experiments and drafted the manuscript. XX, CC, GD, and CX were involved in the data interpretation and result discussion. XX, CC, GD, LC, and CX were involved in the manuscript revision. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Key R&D Program of China (2018YFB1501703), the National Natural Science Foundation of China (NSFC) with Grant number (21878035 and 21576045), the Liaoning Revitalization Talents Program (XLYC1807269), the Dalian Science and Technology Innovation Project (2018J12SN074), the Liaoning Innovative Talent Support Program (LR2017005), the Youth Science and Technology Star Project of Dalian (2017RQ003), the Talent Cultivation Plan of "Xinghai Scholar" from Dalian University of Technology, and the Fundamental Research Funds for the Central Universities (DUT19ZD213).

ACKNOWLEDGMENTS

We would like to thank Prof. Shang-Tian Yang (The Ohio State University, United States) for donating the strain *C. beijerinckii* CC101 and Prof. Sheng Yang and Prof. Weihong Jiang (Shanghai Institutes for Biological Sciences, China) for providing the vector pNICKclos 2.0-*xylR* used for genome editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00214/full#supplementary-material

REFERENCES

Al-Hinai, M. A., Jones, S. W., and Papoutsakis, E. T. (2015). The *Clostridium* sporulation programs: diversity and preservation of endospore differentiation. *Microbiol. Mol. Biol. Rev.* 79, 19–37. doi: 10.1128/MMBR.00025-14

Branska, B., Pechacova, Z., Kolek, J., Vasylkivska, M., and Patakova, P. (2018). Flow cytometry analysis of *Clostridium beijerinckii* NRRL B-598 populations exhibiting different phenotypes induced by changes in cultivation conditions. *Biotechnol. Biofuels* 11:99. doi: 10.1186/s13068-018-1096-x

Bruder, M. R., Pyne, M. E., Moo-Young, M., Chung, D. A., and Chou, C. P. (2016). Extending CRISPR-Cas9 technology from genome editing to transcriptional engineering in the genus *Clostridium*. *Appl. Environ. Microbiol.* 82, 6109–6119. doi: 10.1128/AEM.02128-16

Charubin, K., Bennett, R. K., Fast, A. G., and Papoutsakis, E. T. (2018). Engineering *Clostridium* organisms as microbial cell-factories: challenges & opportunities. *Metab. Eng.* 50, 173–191. doi: 10.1016/j.ymben.2018.07.012

Cheng, C., Bao, T., and Yang, S. T. (2019). Engineering *Clostridium* for improved solvent production: recent progress and perspective. *Appl. Environ. Microbiol.* 103, 5549–5566. doi: 10.1007/s00253-019-09916-7

Dürre, P. (2014). Physiology and sporulation in *Clostridium*. *Microbiol. Spectr.* 2:4. doi: 10.1128/microbiolspec.TBS-0010-2012

Herman, N. A., Kim, S. J., Li, J. S., Cai, W., Koshino, H., and Zhang, W. (2017). The industrial anaerobe *Clostridium acetobutylicum* uses polyketides to regulate

cellular differentiation. *Nat. Commun.* 8:1514. doi: 10.1038/s41467-017-01809-5

Hiscox, T. J., Harrison, P. F., Chakravorty, A., Choo, J. M., Ohtani, K., Shimizu, T., et al. (2013). Regulation of sialidase production in *Clostridium perfringens* by the orphan sensor histidine kinase ReeS. *PLoS One* 8:e73525. doi: 10.1371/journal.pone.0073525

Jacobs, C., Domian, I. J., Maddock, J. R., and Shapiro, L. (1999). Cell cycle-dependent polar localization of an essential bacterial histidine kinase that controls DNA replication and cell division. *Cell* 97, 111–120. doi: 10.1016/S0092-8674(00)80719-9

Kolek, J., Branska, B., Drahokoupil, M., Patakova, P., and Melzoch, K. (2016). Evaluation of viability, metabolic activity and spore quantity in clostridial cultures during ABE fermentation. *FEMS Microbiol. Lett* 363:fnw031. doi: 10.1093/femsle/fnw031

Kolek, J., Diallo, M., Vasylkivska, M., Branska, B., Sedlar, K., López-Contreras, A. M., et al. (2017). Comparison of expression of key sporulation, solventogenic and acetogenic genes in *C. beijerinckii NRRL B*-598 and its mutant strain overexpressing spo0A. *Appl. Microbiol. Biotechnol.* 101, 8279–8291. doi: 10.1007/s00253-017-8555-3

Li, K., Cai, D., Wang, Z., He, Z., and Chen, S. (2018). Development of an efficient genome editing tool in *Bacillus licheniformis* using CRISPR-Cas9 nickase. *Appl. Environ. Microbiol.* 84, e2608–e2617. doi: 10.1128/AEM.02608-17

Li, Q., Chen, J., Minton, N. P., Zhang, Y., Wen, Z., Liu, J., et al. (2016). CRISPR−based genome editing and expression control systems in *Clostridium acetobutylicum* and *Clostridium beijerinckii*. *Biotechnol. J.* 11, 961–972. doi: 10.1002/biot.201600053

Li, S. Y., Srivastava, R., Suib, S. L., Li, Y., and Parnas, R. S. (2011). Performance of batch, fed-batch, and continuous A–B–E fermentation with pH-control. *Bioresour. Technol.* 102, 4241–4250. doi: 10.1016/j.biortech.2010.12.078

Loomis, W. F., Shaulsky, G., and Wang, N. (1997). Histidine kinases in signal transduction pathways of eukaryotes. *J. Cell. Sci.* 110, 1141–1145.

Lu, C., Dong, J., and Yang, S. T. (2013). Butanol production from wood pulping hydrolysate in an integrated fermentation–gas stripping process. *Bioresour. Technol.* 143, 467–475. doi: 10.1016/j.biortech.2013.06.012

Mearls, E. B., and Lynd, L. R. (2014). The identification of four histidine kinases that influence sporulation in *Clostridium thermocellum*. *Anaerobe* 28, 109–119. doi: 10.1016/j.anaerobe.2014.06.004

Mermelstein, L. D., Welker, N. E., Bennett, G. N., and Papoutsakis, E. T. (1992). Expression of cloned homologous fermentative genes in *Clostridium acetobutylicum* ATCC 824. *Nat. Biotechnol.* 10, 190–195. doi: 10.1038/nbt0292-190

Steiner, E., Dago, A. E., Young, D. I., Heap, J. T., Minton, N. P., and Young, M. (2011). Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in *Clostridium acetobutylicum*. *Mol. Microbiol.* 80, 641–654. doi: 10.1111/j.1365-2958.2011.07608.x

Sun, Z., Chen, Y., Yang, C., Yang, S., Gu, Y., and Jiang, W. (2015). A novel three−component system−based regulatory model for d−xylose sensing and transport in *Clostridium beijerinckii*. *Mol. Microbiol.* 95, 576–589. doi: 10.1111/mmi.12894

Underwood, S., Guan, S., Vijayasubhash, V., Baines, S. D., Graham, L., Lewis, R. J., et al. (2009). Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. *J. Bacteriol.* 191, 7296–7305. doi: 10.1128/JB.00882-09

Wang, Y., Zhang, Z. T., Seo, S. O., Choi, K., Lu, T., Jin, Y. S., et al. (2015). Markerless chromosomal gene deletion in *Clostridium beijerinckii* using CRISPR/Cas9 system. *J. Biotechnol.* 200, 1–5. doi: 10.1016/j.jbiotec.2015.02.005

Wang, Y., Zhang, Z. T., Seo, S. O., Lynn, P., Lu, T., Jin, Y. S., et al. (2016a). Gene transcription repression in *Clostridium beijerinckii* using CRISPR−dCas9. *Biotechnol. Bioeng.* 113, 2739–2743. doi: 10.1002/bit.26020

Wang, Y., Zhang, Z. T., Seo, S. O., Lynn, P., Lu, T., Jin, Y. S., et al. (2016b). Bacterial genome editing with CRISPR-Cas9: deletion, integration, single nucleotide modification, and desirable "clean" mutant selection in *Clostridium beijerinckii* as an example. *ACS Synth. Biol.* 5, 721–732. doi: 10.1021/acssynbio.6b00060

Wilke, K. E., Francis, S., and Carlson, E. E. (2015). Inactivation of multiple bacterial histidine kinases by targeting the ATP-binding domain. *ACS Chem. Biol.* 10, 328–335. doi: 10.1021/cb5008019

Wilkinson, S. R., Young, D. I., Gareth Morris, J., and Young, M. (1995). Molecular genetics and the initiation of solventogenesis in *Clostridium beijerinckii* (formerly *Clostridium acetobutylicum*) NCIMB 8052. *FEMS Microbiol. Rev.* 17, 275–285. doi: 10.1111/j.1574-6976.1995.tb00211.x

Wörner, K., Szurmant, H., Chiang, C., and Hoch, J. A. (2006). Phosphorylation and functional analysis of the sporulation initiation factor Spo0A from *Clostridium botulinum*. *Mol. Microbiol.* 59, 1000–1012. doi: 10.1111/j.1365-2958.2005.04988.x

Xiao, H., Li, Z., Jiang, Y., Yang, Y., Jiang, W., and Gu, Y. (2012). Metabolic engineering of D-xylose pathway in *Clostridium beijerinckii* to optimize solvent production from xylose mother liquid. *Metab. Eng.* 14, 569–578. doi: 10.1016/j.ymben.2012.05.003

Xiao, M., Wang, L., Wu, Y., Cheng, C., Chen, L., and Chen, H. (2019). Hybrid dilute sulfuric acid and aqueous ammonia pretreatment for improving butanol production from corn stover with reduced wastewater generation. *Bioresour. Technol.* 278, 460–463. doi: 10.1016/j.biortech.2019.01.079

Xu, M., Zhao, J., Yu, L., Tang, I. C., Xue, C., and Yang, S. T. (2015). Engineering *Clostridium acetobutylicum* with a histidine kinase knockout for enhanced n-butanol tolerance and production. *Appl. Microbiol. Biotechnol.* 99, 1011–1022. doi: 10.1007/s00253-014-6249-7

Xue, C., and Cheng, C. (2019). "Butanol production by *Clostridium*," in *Advances in Bioenergy*, ed. A. Jackman (San Diego, CA: Academic Press), 35–77. doi: 10.1016/bs.aibe.2018.12.001

Xue, C., Liu, F., Xu, M., Zhao, J., Chen, L., Ren, J., et al. (2016). A novel *in situ* gas stripping-pervaporation process integrated with acetone-butanol-ethanol fermentation for hyper n-butanol production. *Biotechnol. Bioeng.* 113, 120–129. doi: 10.1002/bit.25666

Xue, C., Zhao, J., Chen, L., Yang, S. T., and Bai, F. (2017a). Recent advances and state-of-the-art strategies in strain and process engineering for biobutanol production by *Clostridium acetobutylicum*. *Biotechnol. Adv.* 35, 310–322. doi: 10.1016/j.biotechadv.2017.01.007

Xue, C., Liu, M., Guo, X., Hudson, E. P., Chen, L., Bai, F., et al. (2017b). Bridging chemical-and bio-catalysis: high-value liquid transportation fuel production from renewable agricultural residues. *Green Chem.* 19, 660–669. doi: 10.1039/C6GC02546C

Xue, C., Zhao, X. Q., Liu, C. G., Chen, L. J., and Bai, F. W. (2013). Prospective and development of butanol as an advanced biofuel. *Biotechnol. Adv.* 31, 1575–1584. doi: 10.1016/j.biotechadv.2013.08.004

Yang, D., Cheng, C., Bao, M., Chen, L., Bao, Y., and Xue, C. (2019). The pervaporative membrane with vertically aligned carbon nanotube nanochannel for enhancing butanol recovery. *J. Membr. Sci.* 577, 51–59. doi: 10.1016/j.memsci.2019.01.032

Yang, D., Tian, D., Xue, C., Gao, F., Liu, Y., Li, H., et al. (2018). Tuned fabrication of the aligned and opened CNT membrane with exceptionally high permeability and selectivity for bioalcohol recovery. *Nano Lett.* 18, 6150–6156. doi: 10.1021/acs.nanolett.8b01831

Zhang, Z. T., Jiménez-Bonilla, P., Seo, S. O., Lu, T., Jin, Y. S., Blaschek, H. P., et al. (2018). "Bacterial genome editing with CRISPR-Cas9: taking *Clostridium beijerinckii* as an example," in *Synthetic Biology*, ed. J. Braman (New York, NY: Humana Press), 297–325. doi: 10.1007/978-1-4939-7795-6_17

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

REVIEW
published: 27 March 2020
doi: 10.3389/fbioe.2020.00232

The Gibberellin Producer *Fusarium fujikuroi*: Methods and Technologies in the Current Toolkit

Yu-Ke Cen[1,2], Jian-Guang Lin[1,2], You-Liang Wang[1,2], Jun-You Wang[1,2], Zhi-Qiang Liu[1,2] and Yu-Guo Zheng[1,2]*

[1] *Key Laboratory of Bioorganic Synthesis of Zhejiang Province, College of Biotechnology and Bioengineering, Zhejiang University of Technology, Hangzhou, China,* [2] *Engineering Research Center of Bioconversion and Biopurification of Ministry of Education, Zhejiang University of Technology, Hangzhou, China*

In recent years, there has been a noticeable increase in research interests on the *Fusarium* species, which includes prevalent plant pathogens and human pathogens, common microbial food contaminants and industrial microbes. Taken the advantage of gibberellin synthesis, *Fusarium fujikuroi* succeed in being a prevalent plant pathogen. At the meanwhile, *F. fujikuroi* was utilized for industrial production of gibberellins, a group of extensively applied phytohormone. *F. fujikuroi* has been known for its outstanding performance in gibberellin production for almost 100 years. Research activities relate to this species has lasted for a very long period. The slow development in biological investigation of *F. fujikuroi* is largely due to the lack of efficient research technologies and molecular tools. During the past decade, technologies to analyze the molecular basis of host-pathogen interactions and metabolic regulations have been developed rapidly, especially on the aspects of genetic manipulation. At the meanwhile, the industrial fermentation technologies kept sustained development. In this article, we reviewed the currently available research tools/methods for *F. fujikuroi* research, focusing on the topics about genetic engineering and gibberellin production.

Keywords: *Fusarium fujikuroi*, genetic engineering, fermentation, tools, gibberellic acid, CRISPR-cas

OPEN ACCESS

Edited by:
Yi Wang,
Auburn University, United States

Reviewed by:
Raimund Nagel,
University of Leipzig, Germany
Masatoshi Nakajima,
The University of Tokyo, Japan
Javier Avalos,
University of Seville, Spain

****Correspondence:***
Zhi-Qiang Liu
microliu@zjut.edu.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *21 September 2019*
Accepted: *06 March 2020*
Published: *27 March 2020*

Citation:
Cen Y-K, Lin J-G, Wang Y-L, Wang J-Y, Liu Z-Q and Zheng Y-G (2020) The Gibberellin Producer Fusarium fujikuroi: Methods and Technologies in the Current Toolkit. Front. Bioeng. Biotechnol. 8:232. doi: 10.3389/fbioe.2020.00232

INTRODUCTION TO *F. fujikuroi*

F. fujikuroi is a prevalent plant pathogen, which causes the bakanae disease of the rice plant. The sick plants grew inordinately long, and eventually felled off and died. This phytopathogen was latterly found causing devastating disease in many other economically important plants, including maize, sugarcane, wheat, asparagus etc. In the early 20th century, scientists from Japan, United Kingdom, and United States isolated the active compounds, gibberellic acids (GAs), which was also isolated later from the higher plants (Mitchell et al., 1951). Since then, differentially structured GAs were isolated, and GAs became a large family of structurally identical diterpenoids with 136 known isoforms, of which some are active plant hormones, including GA_1, GA_3, GA_4, and GA_7 (Blake et al., 2000; Bomke and Tudzynski, 2009; Rodrigues et al., 2012). GAs are now classified as one of the five major types of phytohormones, namely the auxins, cytokinins, gibberellins, abscisic acid, and ethylene. The use of ppm (parts per million, mg/l) levels of GAs may result in physiological effects such as elimination of dormancy in seeds, acceleration of seed germination, improvement in crop yield, promotion of fruit setting, overcoming of dwarfism etc. GAs have been widely applied to

improve the quality and quantity of fruit, crop and omamental plants. Although GAs are present extensively in plants, fungi and bacteria, *F. fujikuroi* is the only organism being applied for industrial production of GAs, as it shows excellent productivity. Taking the product GA_3, a representative GA product of *F. fujikuroi*, as an example, the yield in industrial submerged fermentation (SMF) has reached more than 2g/L after 7 days fermentation, while in solid state fermentations (SSF), its yield reached 7 g/kg support or even higher after 9 days fermentation. These values are much higher than the reported yield of other microbes. Besides, many other valuable secondary metabolites were also discovered to be produced by *F. fujikuroi*, indicating the potential of *F. fujikuroi* to be apply for production of other chemicals (Janevska and Tudzynski, 2018).

GA induced signal transduction is very complicated in plants, while overdose of GAs may result in plant death (Eckardt, 2002). As a phytopathogen, fusaria can be loosely classified as hemibiotrophs. Upon infection with *F. fujikuroi*, the plant becomes sick/weak, and gets subsequently more easy to be invaded, which could be largely due to the contribution of GAs secretion. Research from Wiemann et al. showed that the infected rice plant experienced dramatically increased invasive fungal growth of a GA-secreting wild type *F. fujikuroi* when compared to its GA-deficient mutant (Wiemann et al., 2013). At the meanwhile, the enlarged plant body by the abnormal elongation might also provide the pathogen additional space and nutrients. Eventually, the infection turns to the stage of killing and consuming the host body, while the fusaria become necrotrophic in this stage (Ma et al., 2013).

It would be interesting to reveal the underlying mechanism of the virulence factors of *F. fujikuroi*, which may help to discover the potential antifungal targets or to develop a strain that are non-pathogenic and safe for the agricultural environment. Currently, we still lack the systemic knowledge to control the pathogenesis of *F. fujikuroi*. The virulence/pathogenicity genes of some *Fusarium* species has been characterized and summarized in some review articles (Michielse and Rep, 2009; Walter et al., 2010; Kazan et al., 2012). The virulence linked host-pathogen interaction is a very complicated process with a massive amount of genes and regulators involved. Based on the infection strategies, Ma et al. classified the virulence genes into two types. The genes of the first class were named as the basic pathogenicity genes, which is universal in the *Fusarium* genus and shared with many other pathogens. Genes of mitogen-activated protein kinase (MAPK) signaling pathways, Ras proteins (small GTPases), G-protein signaling components and cAMP pathways etc. are involved in this class. These genes usually correlate also globally with the cell fitness. The genes in the second class were named as the specialized pathogenicity genes, which is usually specific to a *Fusarium* species on specific hosts (Ma et al., 2013). GAs production is apparently a key virulence factor of *F. fujikuroi* and requires a set of specialized pathogenicity genes. However, GAs production is not essential to the virulence. Deletion of the entire GAs gene cluster could neither impair the host-cell colonization nor abolish the invasive growth completely in a rice-root infection experiment (Wiemann et al., 2013). Besides GAs, *F. fujikuroi* synthesizes a large amount of other metabolites, of which many are toxic compounds. In addition, many secreted enzymes may also help the fusaria to penetrate the cell wall and ultimately invade the plant. Bashyal et al. analyzed the *F. fujikuroi* genome and predicted that there were 1194 secretory proteins, of which 38% proteins might relate to the virulence. Moreover, out of secretory proteins, 5% were polysaccharide lyases, 7% were glycosyl transferases, 20% were carbohydrate esterases, and 41% were glycosyl hydrolases (Bashyal et al., 2017). It is interesting to exploit further experimentally the specialized pathogenicity genes, especially some secreted cell wall degrading enzymes and mycotoxins in *F. fujikuroi* (Desmond et al., 2008).

Beside the gibberellin-producing fusaria, the helminthosporol-producing *Helminthosporium sativum* was also focused. Helminthosporol is a natural sesquiterpenoid that is able to induce GAs like bioactivity (Miyazaki et al., 2017) and cause seedling blight and root rot in some plants (Pringle, 1976). However, far less is known about the biosynthesis of helminthosporol and the biology of *H. sativum*, when compared to GAs and *F. fujikuroi*. Compared to helminthosporol, GAs are prevalently present in the high plants in nature, thus have a broader application, whereas helminthosporol is a plant growth regulator that synthesized by the microorganism. Besides, although helminthosporol and its analogs helminthosporal and helminthosporic acid, have GA-like activity in some plants, they act less efficient or differentially in many experiments. For instance, helminthosporol and helminthosporic acid work less efficiently than GAs in reversing 2-chloroethyl-trimethylammonium chloride induced dwarfing on the hypocotyls of lettuce seedlings and in stimulating sugar release from de-embryonated barley (Briggs, 1966). The GA biosynthetic inhibitor prohexadione did not inhibit the shoot elongation caused helminthosporic acid. *H. sativum* infected wheat was not elongated, because helminthosporol has no GA activity in wheat. *H. sativum* did not infect the rice plant as a host, although helminthosporol may promote rice seedlings (Miyazaki et al., 2017).

A prerequisite to characterize the molecular biology is having efficient analysis tools. After a number of *F. fujikuroi* genome sequences are available (Jeong et al., 2013; Wiemann et al., 2013; Bashyal et al., 2017; Niehaus et al., 2017), it is more eager than ever to develop molecular tools. However, unlike many other fungi, such as the baker's yeast and some phylogenetically close *Aspergillus* spp., *F. fujikuroi* is critically short of molecular tools. Lacking of handy molecular tools has become the major obstacle for the development of the *F. fujikuroi* research, especially in the aspect of genetic modification of this fungus. Genetic manipulation is difficult in this microorganism, which is due to poor protoplast formation, inefficient transformation, low homologous recombination (HR) rate etc. In this review, currently applying methods and tools, including the methods to identify *Fusarium* species, plant infection assays, sexual cross method, promoters for gene expression, plasmid toolbox, protocol of protoplast preparation, transformation technologies, genome editing strategies, RNA-mediated gene silencing assay, protein fluorescent tags, methods of biomass quantification, gibberellin fermentation technologies and strategies of strain improvement, have been reviewed. We summarized the currently

using materials and techniques for *F. fujikuroi* research, providing a perspective in the development of molecular tools for this industrial and agricultural important fungus.

IDENTIFICATION OF *FUSARIUM* SPECIES

The *Fusarium* species are ubiquitous in nature, and are extensively distributed in soil, plants and various organic substances. Identification of the *Fusarium* species becomes crucial for agricultural application, healthcare purpose and scientific investigation. To date, hundreds of *Fusarium* genome sequences have been deposited in the database. These genome sequences can be used as efficient and essential tools for identification of *Fusarium* species, gene/enzyme mining, evolutionary, and phylogenetic analysis etc. (Ward et al., 2002; O'Donnell et al., 2013; Bashyal et al., 2017). There are currently 68 *Fusarium* species with their genome sequences available in the NCBI (National Center of Biotechnology Information, United States) database. Among them, *Fusarium oxysporum*, *F. fujikuroi*, *Fusarium proliferatum* and *Fusarium graminearum* are the best focused four species. The numbers of their genome assemblies in the database are 222, 18, 13, and 11 respectively. These *Fusarium* species are all prevalent phytopathogens and economically very important, thus are also better studied and with more molecular tools available when compared with the other species in the *Fusarium* genus. The *Fusarium fujikuroi* species complex (FFCS), previously known as *Gibberella fujikuroi* species complex, contains about 50–100 phylogenetically close *Fusarium* species, of which *F. fujikuroi*, *Fusarium proliferatum,* and *Fusarium verticillioides* are best studied. The taxonomy of FFSC was based on the evolutionary, biological and morphological species concepts (Kvas et al., 2009; Summerell et al., 2010), whereas the modern biology also employs sequencing data.

Conventionally, the identification of microorganisms is mainly based on the morphology. The morphological identification is generally based on the macroscopic and microscopic characteristics. The macroscopic characteristics include the colony appearance, pigmentations and growth rates. The FFSC cells present usually as white to dark purple cottony aerial mycelium. The microscopic characteristics include the microscopic observation of the macroconidia, microconidia, chlamydospores, the mode of microconidial formations etc. At the moment, the most recent and systemically documented guide for morphological characterization of the *Fusarium* species was contributed by Leslie and Summerell (Leslie and Summerell, 2006). However, the morphological identification is time consuming and could easily result in misidentifications, especially for the phylogenetically close species (Hsuan et al., 2011; Raja et al., 2017). Although it might be problematic to use morphology alone, this method is still helpful in practice and is frequently used now in combination with other molecular means.

The MALDI-TOF MS (Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry) assay is an advanced tool for rapid and accurate identification of microorganisms. This technique has been widely applied to identify bacteria, yeasts and other fungi (Pinto et al., 2011; Mesureur et al., 2018; Pauker et al., 2018; Quero et al., 2019; Kim et al., 2019), especially for the identification of the human pathogens, whereas might be relatively less popular in the plant pathogen research at this moment. However, a broader application can be foresee in the near future based on its excellent accuracy and efficiency, and fast development of analyzing equipment. Briefly, this assay is carried out based on the mass spectral readout of the molecular mass from the ionized protein mixture. Thus, each cell culture may result in a very specific mass spectral pattern, which can be taken as the unique fingerprint to identify a microorganism from the very closely related species (Huschek and Witzel, 2019). To be taken as an identification tool, a database of such mass spectral patterns has to be established beforehand. A MALDI-TOF MS database has been established with 24 reference strains for identification of mainly the clinical isolates belonging to the FFSC. It was reported that 93.6% of the isolates can be correctly identified to the species level (Al-Hatmi et al., 2015). Recently, Wigmann et al. expanded the database (Wigmann et al., 2019). In their work, MALDI-TOF MS was carried out for 49 species from the species complex, taking the sequencing data of the translation elongation factor 1 α (*TEF1*α) gene as the reference. The MALDI-TOF MS fingerprints were then taken as a database to screen over 80 isolates from the FFSC, and resulted in a high correct-identification-rate of 94.61%.

PCR based cell identification is another type of rapid, accurate and cost effective method to identify the microorganisms. Unlike the MALDI-TOF MS method, the PCR based methods require only the routine facilities in a molecular lab. The PCR based methods have been developed for the *Fusarium* species identification since many years ago, whereas without a standardized protocol. Usually, different genomic loci were targeted, and ended up with diverse forms of results. The galactose oxidase gene *gaoA* was taken as the PCR target to identify the *Fusarium* species, as the gene region has very low homology among the fungi (Niessen and Vogel, 1997; de Biazio et al., 2008). The internal transcribed spacer (ITS) regions have been successfully used to identify some closely related fungi. The ITS regions of the conserved rDNA have been successfully used to identify some *Fusarium* species (Abd-Elsalam et al., 2003; Lacmanova et al., 2009). The *TEF1*α gene is usually a single copy gene in the *Fusarium* genus, and is frequently employed for species identification, as it also presents a high level of sequence polymorphism in different species. Other genes such as the β-tubulin, RNA polymerase II (*RPB2*), nitrate reductase, phosphate permase, and the mitochondrial small subunit were also targeted for PCR identification. However, for a better resolution, a multi-locus sequence typing (MLST) method should be used by targeting multiple genes. Usually, at least three gene loci were taken for such identifications (Baayen et al., 2000; O'Donnell et al., 2000; Skovgaard et al., 2001). As an example, Ke et al. (2016) identified the *Fusarium* species by PCR of ITS, *RPB2* and *TEF1*α. Faria et al. (2012) developed a multiplex PCR method after testing 6 pairs of primers targeting different genes/genomic DNA of different *Fusarium*

species. The failure/success of PCR amplification, using different pairs of primers, was counted to determine the belonging of a specific *Fusarium* species (Faria et al., 2012). Recently, a *TEF1α* LAMP (Loop-Mediated Isothermal Amplification) based identification method has been developed for detection of the seedborne *F. fujikuroi* and *Magnaporthe oryzae* in rice seeds. Four independent *F. fujikuroi* isolates were tested taking their serially diluted DNA samples as the amplification templates. Based on the time-to-positive of the LAMP assay, the authors claimed that this assay showed a detection sensitivity/limit of 100–999 fg (vary among different isolates) of *F. fujikuroi* DNA (Ortega et al., 2018).

PLANT INFECTION ASSAYS

Although *F. fujikuroi* may invade many plants, the rice plant is a preferred host. The ability to cause rice bakanae disease has become the hallmark of the microorganism *F. fujikuroi*. Thus, the rice plant was frequently chosen as the host to investigate the virulence of *F. fujikuroi*. Wiemann et al. investigated *F. fujikuroi* virulence that linked to a velvet-like protein complex using a rice plant infection assay. In their experiment, the husks removed rice seeds were incubated for 3 days in agar gel for germination and co-incubated subsequently with 5 mm diameter *F. fujikuroi* mycelial plugs in Vermiculite filled test tubes. The infected plants were grown for another 7 days, supplying with water and nutrients. The germination period and growth period were both implemented at 28°C under a 12 h light – 12 h dark cycle. Finally, the bakanae symptoms such as chlorotic stems and leaves were observed and documented (Wiemann et al., 2010). Adam et al. infected the rice plants using conidia samples of different *F. fujikuroi* strains. In their experiment, the rice seeds were germinated for 2 days for seedlings with developed shoots/roots length of 1–2 mm. A fixed amount of conidia were co-inoculated then with the prepared seedlings for infection. The infected plants were grown for another 10 days with the programmed lighting and nutrient supply. Finally, the plant length and internodal distances were recorded, while the paler pigmentation of the bakanae disease was characterized and verified by measuring their content ratio of chlorophylls/carotenoids (Adam et al., 2018). Similar to the previously described experiment, the whole assay took around 2 weeks to evaluate the systemic *F. fujikuroi* infection of the rice plant *in vivo* (see **Figure 1**). A rice/maize root infection assay was carried out by Wiemann et al. to evaluate the pathogenicity of GAs production. The pathogens were inoculated by co-cultivation with the geminated rice and maize seeds. Fixed temperature and humidity, and programmed lighting cycles (differentially for rice and maize) were supplied to the infected seedlings in agar gel support. After 10 days of growth the root samples were collected for visualization of invasive growth of the corresponding pathogen by fluorescence microscopy, and the penetration events were quantified. At the meanwhile, pathogen spores (10^4/ml) were collected for measurement of relevant mRNA by RT-PCR (Wiemann et al., 2013). On the basis of the described infection assays, mostly the rice seedlings were preferred to be chosen as the host plant. Different fungal samples, such as mycelia and conidia, can be used to infect the seedlings, while different methods are available to characterize the pathogenicity/virulence (see **Figure 1**).

SEXUAL CROSS

Crossing is always a powerful method to combine genotypes, exchange genetic materials and obtain large-scale mutations. Studt et al. implemented a sexual cross between a pair of mating partners, C1995 and IMI 58289, two well-studied laboratory strains with opposite mating types, Mat-2 and Mat-1, respectively. With the experiment, they pinpointed that a newly found gene *FSR1* was involved in perithecial pigmentation in *F. fujikuroi* (Studt et al., 2012). The crossing experiment was carried out using a carrot agar medium, which has been reported by Klittich et al. previously (Klittich and Leslie, 1988). The crossing protocol has been well documented by Zakaria et al. (2011). Briefly, in their experiment the female parent and male parent were inoculated in the carrot medium and complete medium respectively. After 7 days growth at 25°C, the mycelium from the male parent was harvest and suspended in Tween 60 for spore suspension, which was subsequently spread into the mycelium of the female parent on a carrot agar plate. The carrot agar plate was then incubated at 27°C for a few weeks until the perithecia were produced. The *F. fujikuroi* species complex was divided into many biological species, designated as mating populations A to J. *F. fujikuroi* is the mating population C. Generally, *F. fujikuroi* is heterothallic, and should be readily crossed in the laboratory. However, sexual fertility varies from strain to strain, making the sexual crosses not always successful (Zakaria et al., 2011).

PROMOTERS

Selection of a proper promoter is crucial in genetic engineering. Usually, based on the research purpose and the gene to be expressed, a native promoter, a constitutively expressing promoter or an inducible promoter can be used. In *F. fujikuroi*, the most frequently used promoters, such as the *gpdA* (Michielse et al., 2014) *oliC* and *trpC* (Rosler et al., 2016) promoters, are originated from the *Aspergillus* spp. and provide very strong expression. The native strong *glnA* promoter can be induced under the nitrogen starvation condition, while can be repressed under the addition of NH_4NO_3 or glutamine (Teichert et al., 2004). The transcriptional regulation of *glnA* is on the basis of the transcription factor AreA, which is extensively involved in regulation of a wide range of metabolism pathways. Thus, nitrogen starvation/induction is closely linked to the synthesis of many important secondary metabolites, suggesting that potential conflict between the *glnA* expression and the research purpose has to be taken care of before the *glnA* promoter is chosen. The *glnA* promoter has been used for conditional expression of a gene in *F. fujikuroi* (Teichert et al., 2006). The *alcA* promoter is another strong inducible promoter that has been successfully applied in *F. fujikuroi* research (Teichert et al., 2006). The *alcA* promoter

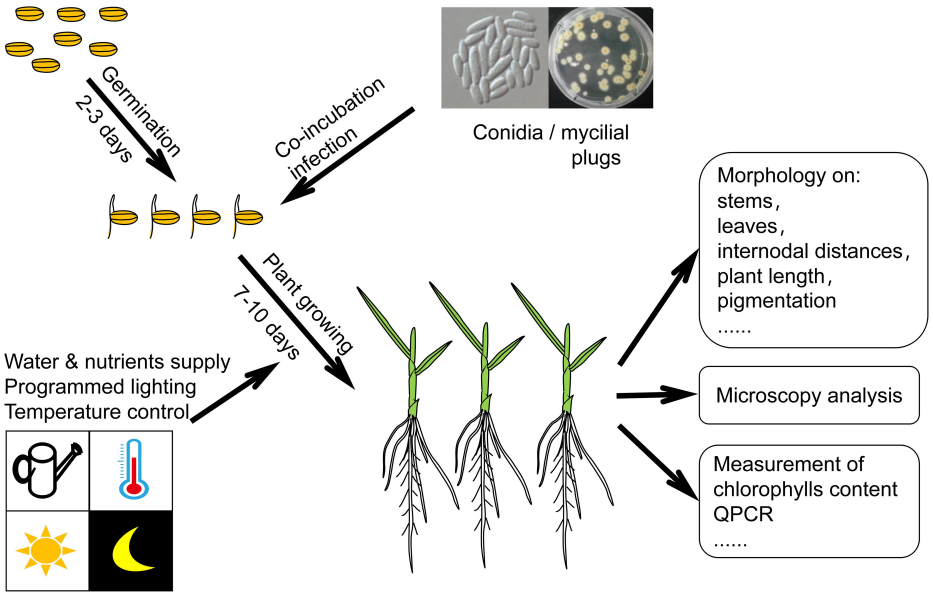


FIGURE 1 | Schematic diagram of a *F. fujikuroi*-rice plant infection assay.

driven gene expression can be well induced by 1% (V/V) ethanol and repressed by 2% (W/V) glucose in *F. fujikuroi*. The *alcA* promoter is originated from *Aspergillus nidulans*. *alcA* and two other genes, *alcR* and *aldA*, are the genes of ethanol regulon, and are all transcriptional regulated by CREA and ALCR proteins (Mathieu and Felenbok, 1994). The *glnA* and *alcA* promoters are both strong promoters. Besides these, a Tet-on system has been developed for *F. fujikuroi* (Janevska et al., 2017). This Tet-on system was established based on the adaptation of a Tet-on system of *Aspergillus niger* (Meyer et al., 2011). This expression construct was composed by an *oliC* promoter, a tetracycline-dependent transactivator rtTA2^{S}-M2, an *A. fumigatus* terminator *TcgrA*, and an rtTA2S-M2-dependent promoter *tetO7*:Pmin (see **Figure 2**). The constructed Tet-on promoter has been shown to successfully activate a silent gene cluster in *F. fujikuroi* by adding 50 μl/ml doxycycline.

PLASMID TOOLBOX

Plasmids have formed an essential part in molecular biology and genetic manipulation. However, compared to the model organisms, such as *Saccharomyces cerevisiae*, which has the most diverse plasmids, *F. fujikuroi* has almost no specific plasmid for use. In *F. fujikuroi*, the currently working plasmids are generally integrative plasmids that originated from the plasmid toolbox of the *Aspergillus* species. An *A. nidulans* DNA fragment AMA1, which enables autonomous replication (AR) of plasmids in some *Aspergillus* species, such as *A. nidulans*, *A. niger,* and *Aspergullus oryzae* (Gems et al., 1991), has been tested in *F. fujikuroi*. The transformation efficiency increased over two times by using the AMA1 integrated AR plasmid in comparison with the backbone plasmid. However, this plasmid showed to be very instable inside the cells. After 10–19 days incubation in non-selective medium, only 8–44% of the cells still contain the selection marker (Bruckner et al., 1992). In addition, the southern blot test of the cells transformed with this plasmid gave very weak bands when compared to the cells transformed with the backbone plasmid, although the total amount of DNA used for this assay is equal for both conditions, suggesting that transformation of AMA1 plasmid resulted in a low copy number of the plasmid maintaining in the cells.

It would be interesting to test a centromeric plasmid based on this AR construct. In yeast, a centromeric plasmid usually works as a small chromosome, with one or two copies in each cell, which provides a stable expression profile. Some years ago, yeast centromere *CEN11* had been tested with an plasmid in *A. nidulans* (Boylan et al., 1986). However, this yeast centromere seems to be unfunctional in *A. nidulans*, since it has no effect in plasmid stability, and does not prevent chromosomal integration of the vector. Thus, to construct a centromeric plasmid, it might be necessary to test a native centromere of *F. fujikuroi*. In fact, several *Fusarium* species have been known to contain dispensable mini-chromosomes. These mini-chromosomes stay independently from the other chromosomes, can somehow communicate between neighborhood cells and contribute to the pathogenicity (Nagy et al., 1995; Ma et al., 2010; Ma and Xu, 2019; Peng et al., 2019). However, little is known about the entity of these mini-chromosomes. It would be very interesting to exploit how these chromosomes are utilized, manage to replicate, and are selectively present in the host fungi.

The selective markers are essential for the screening of plasmid transformations. In *F. fujikuroi*, the choices of selection markers are very limited. Drug resistance markers, as the representatives the nourseothricin resistance marker *nat1* (Teichert et al., 2006; Bomke et al., 2008; Janevska et al., 2017), hygromycin

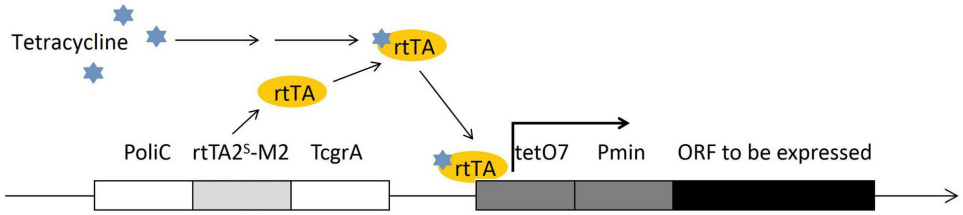


FIGURE 2 | The construct of Tet-on promoter for conditional expression in *F. fujikuroi*. The promoter region is composed of a tetracycline-dependent transactivator rtTA2^S-M2 (on the left of the construct, encodes rtTA protein) and an rtTA protein driven operator tetO7. The tetracycline activated rtTA protein is capable to bind the tetO7 operator and induce the targeted gene expression.

resistance marker *hph* (Studt et al., 2013a; Wagner et al., 2013) and geneticin (G418) resistance marker *nptII* (Castrillo et al., 2013), were the most frequently used selective markers in *F. fujikuroi*. Nutrition selection markers were hardly seen to be used in this microorganism, for instance the use of auxotrophic complementary marker genes, which could be due to the lack of constructed auxotrophic strains. Sanchez-Fernandez et al. (1991) mutated the nitrate reductase gene *niaD* in *F. fujikuroi*, and developed a selection system employing a complementary *niaD* gene of *A. niger*. With this system, the transformants were screened for the ability to utilize nitrate as the sole nitrogen source. The *A. niger niaD* gene was subsequently replaced by a native *niaD* gene of *F. fujikuroi* for future applications (Tudzynski et al., 1996; Prado et al., 2004). **Table 1** has listed the currently using transformants-screening markers for *F. fujikuroi*. Wiemann et al. (2012) knocked out the Sfp-Type 4′-Phosphopantetheinyl Transferase Ppt1. The resulted mutant strain became lysine auxotrophic and dramatically increased in GAs yield. It might be interesting to apply this strain for future lysine auxotrophic screening. Twaruschek et al. developed a plasmid that is able to recycle markers for continuous genetic engineering in *F. graminearum*, as such to overcome the shortage of selection markers in this species. In their strategy, the recombinase system Cre-loxP was activated upon induced expression to remove the marker genes after genetic engineering, while *URA3*/*pyrG* was involved in the system to counterselect marker removal isolates with 5-fluoroorotic acid (Twaruschek et al., 2018). Similarly, the yeast FLP recombinase has been applied in *Candida albicans* to recycle the nourseothricin resistance marker (Reuss et al., 2004), while a Cre-loxP marker recycling system has also been tested in *A. oryzae* (Mizutani et al., 2012). Induced expression of recombinases has been widely employed to recycle the selective makers, and in the meanwhile to remove the redundant DNA fragment after genetic engineering. This is a feasible strategy to get applied also in *F. fujikuroi*, especially when multiple genes need to be disrupted.

TABLE 1 | Summary of the currently using markers for transformants screening in *F. fujikuroi*.

Type of selection		Marker gene	Example reference
Drug resistance markers	Nourseothricin resistance	*nat1*	Teichert et al., 2006; Janevska et al., 2017; Bomke et al., 2008
	Hygromycin resistance	*hph*	Studt et al., 2013a; Wagner et al., 2013
	Geneticin resistance	*nptII*	Castrillo et al., 2013
Nitrogen source	Nitrate dependent	*niaD*	Sanchez-Fernandez et al., 1991; Tudzynski et al., 1996; Prado et al., 2004

PROTOPLAST PREPARATION AND TRANSFORMATION TECHNOLOGIES

In many filamentous fungi, successful protoplast generation is the prerequisite of single cell isolation, efficient transformation, successful genetic engineering etc. *F. fujikuroi* is a polynuclear mycelial fungus. Some important industrial applying strains do not produce conidia, making it obliged to prepare protoplast for single cell isolation. Efficient cell wall degradation enzymes are of substantial importance to generate protoplasts. Some degradation enzymes, such as the snailase (Mink et al., 1990) and chitinase (Patil et al., 2013; Halder et al., 2014), have been frequently used to deconstruct the hyphae in some fungi. Shi et al. (2019) performed a series of optimizations on the preparation of protoplasts. This is the latest update about the optimized protocol for protoplast production in *F. fujikuroi*. They have tested five enzymes, including the lysozyme, snailase, cellulase, lysing enzyme and driselase. Only the lysing enzyme (Sigma-Aldrich, United States) and driselase (Sigma-Aldrich, United States) treated cells gave a reasonable amount of living protoplasts. The lysing enzyme and driselase were then tested in combination and the optimum ratio was obtained at 3:2 with a total concentration of 15 mg/L. Finally, the hydrolyzing time was optimized based on the amount of produced protoplast and cell regeneration efficiency, and the optimal hydrolysis time was ultimately chosen at 3.5 h.

There are many transformation methods available for *F. fujikuroi*. Electroporation, PEG (polyethylene glycol)-mediated transformation and *Agrobacterium* transformation are the three most popular transformation methods in filamentous fungi. The PEG-mediated transformation is an easy-to-operate method that usually combines a heat-shock process. This method has been frequently used to transform *F. fujikuroi*

(Bruckner et al., 1992; Linnemannstons et al., 1999; Fernandez-Martin et al., 2000). Besides *F. fujikuroi*, PEG-mediated transformation method has also been frequently used in many other filamentous fungi, such as *Aspergillus fumigatus* (Fuller et al., 2015), Stagonospora nodorum (Liu and Friesen, 2012), and *Pseudogymnoascus verrucosus* (Diaz et al., 2019). Electroporation is another frequently used method for DNA transformation in fungi. This method is usually known to achieve high transformation efficiency. However, the applied voltage needs to be carefully adjusted, especially when the transformation is applied to the protoplast, a very week cell form. Garcia-Martinez et al. (2015) successfully implemented electroporation in *F. fujikuroi* at the voltage amplitude of 600 V/mm, with one pulse duration of 200 μs, in a cuvette with 1 mm electrode distance. The *Agrobacterium* transformation method has been successfully developed for the filamentous fungi for many years. It was claimed to be much more efficient than the conventional techniques (de Groot et al., 1998). In the *Agrobacterium* transformation protocol, both conidia and protoplast can be used as the host. This method has been succeed in transforming DNA in many different fungi, including *A. niger*, *Aspergillus awamori*, *Trichoderma rees*ei, *Colletotrichum gloeosporioides*, *Neurospora crassa* etc. The *Agrobacterium* transformation method has also been successfully applied in some *Fusarium* species, such as *F. oxysporum* (Islam et al., 2012), *Fusarium venenatum* (de Groot et al., 1998), and *F. proliferatum* (Bernardi-Wenzel et al., 2016). Zhu et al. (2009) tested the PEG and Agrobacterium-mediated transformations of a plasmid in *F. fujikuroi* using respectively protoplast and conidia as the competent cells. With their protocol, 15 transformants per μg of DNA and 37 transformants per 1×10^6 conidia were harvest respectively by the PEG-mediated and Agrobacterium-mediated transformations.

GENOME EDITING

Homologous Recombination (HR)

Unlike the baker's yeast, genome editing is usually inefficient in most of fungi. This is mainly due to the bad transformation efficiency and low HR rate. In *S. cerevisiae*, efficienct gene targeting can be carried out using 30–40 bp homologous flanking sequence on each side of the donor DNA, while in *F. fujikuroi* or many other fungi, we usually use 500 bp or longer homologous flanking sequences, even though the correct integration rate is still very low. We harvested 1 correct mutant after screening over 100 transformants when deleting a gene in *F. fujikuroi*, although a donor DNA construct harboring 700 bp homologous flanking sequence on each side was used (data not shown). In many other organisms, the "correct transformant rate" can be significantly improved by blocking the Non-Homologous End Joining (NHEJ) system. Usually, either gene *ku70/80* or gene *lig4* was knocked out to eliminate NHEJ in different fungi, including *N. crassa* (Ninomiya et al., 2004), *Kluyveromyces lactis* (Kooistra et al., 2004), *Cryptococcus neoformans* (Goins et al., 2006), *Aspergillus* spp. (da Silva et al., 2006; Takahashi et al., 2006; Meyer et al., 2007), *Pichia ciferrii* (Schorsch et al., 2009), and *Candida glabrata* (Ueno et al., 2007; Cen et al., 2015). In comparison to *KU70/80*, deletion of the *LIG4* gene showed to have less side effect, except for losing the NHEJ function (Daley et al., 2005; Schorsch et al., 2009; Cen et al., 2015). It will be interesting if we can delete the *lig4* gene in *F. fujikuroi*, as such to enhance the gene targeting efficiency in this species. However, according to our previous work, deficient NHEJ could not always make noticeably increase in HR efficiency, but reduce dramatically the ectopic integration rate (Cen et al., 2015). The CRISPR-cas system was frequently reported to promote the HR efficiency when compared to the classical approaches (Lin et al., 2014; Zhang Y. et al., 2016; Cen et al., 2017; Chung et al., 2017). Thus, applying the CRISPR-cas technology in an NHEJ deficient *F. fujikuroi* strain can be expected to ideally improve the genetic engineering efficiency. We suggest to apply the CRISPR-cas system in combination of *lig4* disruption (Cen et al., 2017).

CRISPR-Cas

The Clustered Regulatory Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated (Cas) system has brought a remarkable development in genome engineering efficiency in different organisms during the past few years (Doudna and Charpentier, 2014; Albadri et al., 2017; Ermert et al., 2019; Pu et al., 2019; Song et al., 2019). Briefly, the CRISPR-cas system employs a guide RNA (gRNA) and an endonuclease, mostly a single nuclease Cas9 (Makarova et al., 2011; Chylinski et al., 2014), as the two working elements for a site directed DNA cutting. The cut DNA can be then repaired by either NHEJ or HR (see **Figure 3**). The NHEJ repair frequently caused unpredictable mutagenesis, while the homology directed repair may result in a seamless genetic editing. The first fungal CRISPR-Cas system was established in *S. cerevisiae* (DiCarlo et al., 2013). After that, the CRISPR-Cas system was developed very rapidly in different fungi, including *C. albicans* (Vyas et al., 2015), *C. glabrata* (Enkler et al., 2016; Cen et al., 2017), *C. neoformans* (Arras et al., 2016; Wang, 2018), *A. niger* (Zheng et al., 2018, 2019), *A. fumigatus* (Fuller et al., 2015), *A. oryzae* (Katayama et al., 2016) etc.

Although the CRISPR-cas technology has been used for a few years in filamentous fungi (Nodvig et al., 2015), this system was only tested successfully in *F. fujikuroi* very recently. Shi et al. (2019) developed a CRISPR-cas system for genome editing in *F. fujikuroi* as the first time at the beginning of 2019. In their work, three nuclear localization signal (NLS) peptides, the classical SV40 NLS, an endogenous histone NLS and an endogenous *Velvet* NLS, was tested to import the Cas9 protein into the nucleus. Finally, the NLS of native histone H2B was chosen to fuse with the Cas9 protein, as it gave the best mutagenesis rate. The promoter selection for gRNA transcription was the second challenge. Shi et al. (2019) evaluated three promoters, the polymerase II promoter, the endogenous polymerase III *U6* promoter and the endogenous 5SrRNA promoter, of which the 5SrRNA promoter gave the best editing efficiency. The resulted CRISPR-Cas system showed a genome editing efficiency of approximately 60–80%.

The advantage of this *F. fujikuroi* CRISPR-Cas system is that the expression of Cas9 protein did not show any effect to the cell growth, while the non-specific toxicity resulting from

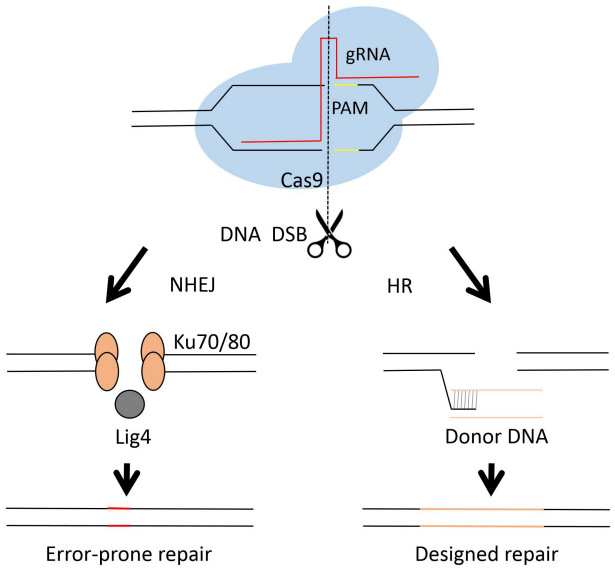


FIGURE 3 | Diagram of Cas9 complex and DNA repair in genome engineering. The CRISPR-Cas system employs a short guide RNA to direct the Cas9 protein, an endonuclease, to a specific cutting site in the genome and generates DNA double strands break (DBS). The lethal DBS can be repaired by either NHEJ or HR. In the NHEJ process, proteins such as Ku70, Ku80 and Lig4 are involved. The NHEJ process results in error-prone repairs. In the HR process, usually a donor DNA is employed for precise genetic engineering.

Cas9 expression or nuclease activity has been widely noticed in many other organisms (Morgens et al., 2016, 2017; Munoz et al., 2016; Cen et al., 2017). In some other microorganisms, such as *E. coli* and baker's yeast, usually the CRISPR-Cas elements can be eliminated after the genetic manipulation, which not only avoided the possible toxicity of the Cas9 nuclease, but also enabled continuous genome editing. In *E. coli*, a dual-plasmid-based system was employed. The gRNA sequence was included in one plasmid (named as pTarget). The *cas9* gene, a temperature-sensitive replicon *repA101*(Ts) and an arabinose induced gRNA expression cassette that targets the pTarget plasmid were edited in another plasmid (named as pCas). The pTarget can be eliminated by adding arabinose, while the pCas can be removed by rising the temperature (Jiang et al., 2015). In *S. cerevisiae*, another dual-plasmid-based system was employed using a centromeric plasmid and a 2μ plasmid. In the centromeric plasmid, Cas9 can be expressed upon galactose induction. The gRNA containing 2μ plasmid can be eliminated by culturing the cells in non-selective medium (DiCarlo et al., 2013) for continuous genome editing. However, the current *F. fujikuroi* CRISPR-Cas system is a one-off system. For multi-gene editing, multiple gRNA expression cassettes have to be cloned into one plasmid to execute genetic engineering one time for all gene targets. The multi-gene disruption efficiency was also tested by Shi et al. The disruption efficiencies were 79.2, 10.8, and 4.2% for single, double and triple gene disruption, respectively. This result indicates that genetic manipulation for 3 genes or more would be very difficult in *F. fujikuroi* by this system.

Another type of CRISPR-Cas system has been developed for several non-*fujikuroi* fusaria, using the *in vitro* prepared Cas9 protein/gRNA ribonucleoproteins (RNPs). Transformation of *in vitro* prepared Cas9/gRNA RNPs has been claimed to be able to reduce the specific integration of the donor DNA, while the CRISPR elements can be degraded naturally after genetic manipulation (Wang et al., 2018). However, compared to the plasmid-transformation method, this RNPs-transformation method is more complicated in handling, as the Cas9 protein needs to be additionally purified and concentrated before the transformation. Ferrara et al. (2019) developed a CRISPR-Cas9 system for *F. proliferatum* by transforming *in vitro*-assembled dual Cas9 RNPs. Using this method, the genomic DNA was cut twice at a specific locus and the donor DNA can target the DSB (DNA double strand break) efficiently using a short homologous flanking sequence of 50 bp on each side (Ferrara et al., 2019). Such efficient HR (efficient DNA integration using 35-60 bp flanking arms) has also been achieved in some other filamentous fungi, including *Penicillium chrysogenum* (Pohl et al., 2016) and *A. fumigatus* (Zhang C. et al., 2016). The efficient homology directed repair may simplify the construction of donor DNAs and reduces off target rate. It is very interesting to also achieve it in *F. fujikuroi*.

RNA-MEDIATED GENE SILENCING

An alternative method to gene deletion is to silence gene expression at the post-transcriptional level, mostly known as RNAi (RNA interference). The RNAi technology has become an excellent tool to exploit gene function in microorganisms, plants and animals. Briefly, RNAi employs a specific double strand RNA and the homologous based mechanism to attack and degrade the targeted mRNA, and finally knockdown the gene expression. This technology has been successfully established in different filamentous fungi for many years (Liu et al., 2002; Goldoni et al., 2004; Mouyna et al., 2004; Ullan et al., 2008). McDonald et al. applied the RNAi technology in three filamentous fungal phytopathogens, two aspergilli and *F. graminearum*, over a decade ago (McDonald et al., 2005). More recently, Nino-Sanchez et al. developed an RNAi system for *F. oxysporum* on the basis of the dsRNA expression cassette used for *P. chrysogenum* and *Acremonium chrysogenum* (Nino-Sanchez et al., 2016). Compared to gene deletion, RNAi is apparently a better choice to target the essential genes. Unfortunately, to date there's no report about RNAi application in *F. fujikuroi*. RNAi is highly conserved among many eukaryotes, thus has a great potential to be used also in *F. fujikuroi*.

FLUORESCENT PROTEIN TAGS

Bimolecular Fluorescence Complementation

Investigation of protein-protein interaction is essential for understanding the signal transduction and regulation of metabolism, and helps to reveal many intrinsic biological

mechanisms. Bimolecular Fluorescence Complementation (BiFC) assay is a decent tool for *in vivo* observation of protein-protein interaction. Michielse et al. (2014) have applied the BiFC assay in *F. fujikuroi* several years ago. They tagged two transcription factors with two splitted gene fragments of an enhanced yellow fluorescent protein (EYFP). The two gene fragments encode the N terminal amino acids 1–154 and C terminal amino acids 155–238. Two transcription factors were then tagged with these two gene fragments and co-transformed into *F. fujikuroi*. The resulted transformant showed a co-localized fluorescence signal in the nucleus. This method was previously tested by Hoff et al. in another filamentous fungus (Hoff and Kuck, 2005), and was directly applied in *F. fujikuroi*. It would be more confident if we can test the BiFC assay systemically with additional controls to eliminate errors such as EYFP self-aggregation before we can start to apply it extensively in *F. fujikuroi*.

Fluorescent Proteins

Fluorescent-protein tags are common tools to monitor the protein localization. It has been extensively practiced in the research of *F. fujikuroi* (Studt et al., 2013b; Pfannmuller et al., 2015). Michielse et al. (2014) tagged the GATA transcription factors AreA and AreB with GFP and RFP respectively to track their cytosolic or nucleic localization. Wiemann et al. (2010) tagged two proteins of a *velvet* complex, FfVel1 and FfLae1, by enhanced GFP and YFP, and visualized a nuclear co-localization signal. Garcia-Martinez et al. tagged the light-sensitive gene *carO* with enhanced YFP and visualized the membrane localization signal. To successfully express a functional fusion protein, they used an 18 bp DNA fragment to bridge the gene and tag (Garcia-Martinez et al., 2015).

QUANTIFICATION OF BIOMASS

Biomass quantification techniques are very basic but important to monitor the cell proliferation. Usually, the microorganisms can be quantified simply by counting the cells using a hemocytometer, measuring the optical density, or weight the cell dry weight. Due to the filamentous nature of many fungi, dry weight measurement become the only feasible way to efficiently determine the biomass. However, *F. fujikuroi* is an industrial production microbe. Those traditional measurements cannot satisfy the complicate fermentation medium, which usually contains insoluble components, such as the corn/rice flour, soybean pulp and arachis flour. These components may form sticky paste like medium or present as insoluble particles. The tetrazolium salt (XTT) method is another efficient way to quantify the biomass. XTT can rapidly penetrate into the living cells and being catalyzed by the active dehydrogenase. Thus, the XTT method can also discriminate the active biomass from the cell debris and bio-inactive particles. The XTT reaction uses a color change for readout and has been applied in many filamentous fungi (Meletiadis et al., 2001; Antachopoulos et al., 2007; Moss et al., 2008). We developed an XTT assay to measure the active biomass of *F. fujikuroi* (Cen et al., 2018). The established method was then tested and approved using the industrial fermentation conditions. Using this method, the cell growth can be well monitored during the fermentation.

There are other methods available for biomass quantification, for instance, the ELISA method and the quantitative PCR (QPCR). These methods have a very high resolution, and are able to distinguish a trace amount of difference in biomass. Besides, they can also distinguish the targeted sample in a complicated mixture. As an example, both ELISA and QPCR were used to quantify the biomass of filamentous fungi in infected plants (Brunner et al., 2012; Song et al., 2014; Feckler et al., 2017). However, when the cell sample is in a large quantity, such as the cells in fermentation, then the big dilution factor could confer significant error to these assays.

FERMENTATION TECHNOLOGIES

Medium Composition

Gibberellin fermentation has a very long history since 1950s (Darken et al., 1959; Rodrigues et al., 2012). Development of fermentation technologies has been sustained for decades for the development of the gibberellin industry. The optimization of the fermentation conditions started since the early 1950s (Borrow et al., 1955). Darken et al. (1959) tested different carbon sources as the first time, and concluded that addition of slowly utilized carbon source may result in increased GAs production, while slow-feed of glucose in a fermentation also positively affected the GAs production. The carbon-catabolite-repression has been known for a long time, while addition of a large quantity of glucose inhibits GAs production. However, little is known about the molecular basis of this phenomenon. Based on the cost efficiency, the currently used industrial fermentation media are mostly composed of a large quantity of starch. Plant oils were also successfully used for gibberellin fermentations (Gancheva et al., 1984; Gokdere and Ates, 2014). The addition of plant oils was interpreted to inert the carbon-catabolite-repression (Tudzynski, 1999). However, it has not been experimentally verified. Addition of plant oils might also functions to balance the nutritional needs of the fungi and release the metabolic burden of alternative biosynthesis pathways. Besides, as a phytopathogen, it is reasonable that *F. fujikuroi* secretes GAs to infect plants in the case that nutrients are poorly present in nature. More complex carbon sources would possibly resemble the natural system better. Nitrogen inhibition has also been known for many years in gibberellin production (Borrow et al., 1964). Kuhr et al. (1961) investigated the influence of different nitrogen sources on the production of GAs. The complex organic nitrogen sources, such as the peanut meal, soybean meal and yeast extract, are favorable for the production of GAs. The molecular basis of nitrogen inhibition has been well studied during the past 10 years (Mihlan et al., 2003; Wagner et al., 2010, 2013; Michielse et al., 2014; Tudzynski, 2014). Other fermentation parameters, including the growth temperature, kinetics of nutrient metabolisms and impact of some other nutrients were also studied many years ago (Borrow et al., 1964; Bruckner et al., 1991). The plant extract seems to be in favor of, as they were frequently reported to promote the GA synthesis. Sucrose, corn steep liquor, glycerol, soybean pulp, arachis flour etc. were frequently chosen as

TABLE 2 | Summary of various SSFs experiments for GA_3 production.

Microorganism	Substrate	Productivity	Fermentation time	References
Fusarium fujikuroi	Wheat bran with soluble Starch	1.22 mg g^{-1}	7 days	Kumar and Lonsane, 1986
Fusarium moniliforme	Maize flour mixed with wheat bran	19.3 mg g^{-1}	18 days	Qian et al., 1994
Fusarium fujikuroi	Maize cob	4.8 mg g^{-1}	22.9 days	Pastrana et al., 1995
Fusarium fujikuroi	Wheat bran	6.8 mg g^{-1}	7.9 days	Agosin et al., 1997
Fusarium fujikuroi	Wheat bran	3 mg g^{-1}	11 days	Bandelier et al., 1997
Fusarium fujikuroi	Coffee husk with cassava bagasse	0.4925 mg g^{-1}	7 days	Machado et al., 2002
Fusarium fujikuroi	Coffee husk with cassava bagasse	0.925 mg g^{-1}	6 days	Machado et al., 2004
Fusarium fujikuroi	Wheat bran and soluble starch	4.7 mg g^{-1}	6 days	Corona et al., 2005
Fusarium fujikuroi	Cassava bagasse	1.58 mg g^{-1}	8 days	Otálvaro et al., 2008
Fusarium moniliforme	Citric pulp with sucrose	5.9 mg g^{-1}	3 days	Rodrigues et al., 2009
Fusarium proliferatum	Pigeon pea pod	7.8 mg g^{-1}	8−10 days	Satpute et al., 2010
	Pea pods	6.4 mg g^{-1}		
	Corncobs	6.1 mg g^{-1}		
	Sorghum straw	5.5 mg g^{-1}		
Fusarium moniliforme	Jatropha seed cake	225 mg g^{-1}	6 days	Rangaswamy and Balu, 2010
Fusarium moniliforme	Jatropha seed cake	105 mg g^{-1}	4 days	Rangaswamy, 2012
Fusarium moniliforme	Wheat bran with soluble starch	1.16 mg g^{-1}	7 days	Desai and Panchal, 2016
Fusarium fujikuroi	Citric pulp	7.34 mg g^{-1}	9 days	de Oliveira et al., 2017

the key components of fermentation media to promote GAs production, especially in the modern gibberellin industry (Cen et al., 2018). These fermentation factors, such the plant extract and temperature (25–28°C), are more like simulations of the nature environment, in which *F. fujikuroi* secretes GAs to invade the plant to survive and propagate. To date, modifications of fermentation conditions are still on going with the purpose to increase the productivity, reduce the cost and make it compatible to the following processings.

Fermentation Types

Different types of fermentation have been established for GAs production. Currently, the most widely used industrial fermentation is the SMF. However, the SMF usually requires high energy consumption, is deficient in aeration, gets frequently contaminated and ends up with a large amount of waste water. The SSF is the most frequently tested fermentation other than SMF. The GAs yield is much higher in SSF in comparison to that of the other types of fermentation. The reported GAs yield reached more than 5 mg/g support after 7 days fermentation (Corona et al., 2005; Rodrigues et al., 2009; Satpute et al., 2010). The reason could be that SSF mimics best the growth conditions of this microorganism in nature, and is able to overcome all the previously mentioned shortcomings of SMF. In addition, the increased GAs yield might be largely due to the increased aeration. The supplied oxygen might be consumed during the GAs anabolism, as the monooxygenases play very important roles during the synthesis of GAs (Tudzynski, 1999, 2005; Albermann et al., 2013). Or probably the enhanced mitochondrial respiration benefited the cell growth in general and indirectly improved GAs synthesis. In an SSF system, the selection of support/substrate materials is crucial, as it may provide nutrients, serve as a support material, and induce product synthesis. Different types of substrate materials have been tested for GAs production, including wheat bran (Agosin et al., 1997; Bandelier et al., 1997), Coffee husk (Machado et al., 2002, 2004), citric pulp (Rodrigues et al., 2009; de Oliveira et al., 2017), Pigeon pea pods, Corncobs, Sorghum straw (Satpute et al., 2010) etc. The use of these plant derived materials/wastes reduced the production cost whereas all resulted in a high productivity of GAs. A variety of tested SSF experiments for GA_3 production have been listed in **Table 2**. Water activity reflects the active part of moisture content, and is usually considered to correlate with microbial growth. It has been found that the water activity has a significant impact on GAs fermentation in a SSF. Usually, the water activity needs to reach 0.99 or higher for optimal cell growth and efficient GA production (Gelmi et al., 2002; Corona et al., 2005). However, unlike SMF, these SSF technologies are all non-conventional setup, and are currently only tested in the laboratory scale. The fermentation scale-up tests, further processings, such as the extraction and purification technologies, stay to be investigated before it can be applied in the industry. Oliveira et al. developed a semisolid state fermentation (SSSF) system using submerged citric pulp particles. They compared the SMF, SSF, and SSSF systems using similar citric pulp based media in both bubble columns (BCR) and Erlenmeyer flask reactors. In general, the GA productivity in BCRs is lower than that in the Erlenmeyer flasks, while the SSSF system is better than the SMF system in yield, indicating that an SSSF system could be possibly used to replace the SMF system while keep the current industrial facilities and production processes.

STRAIN IMPROVEMENT FOR GA PRODUCTION

F. fujikuroi is able to produce a variety of valuable secondary metabolites. There are at least over 10 different types

of metabolites that are known to be produced by this microorganism. Forty seven putative gene clusters of different secondary metabolites have later been revealed and well documented (Wiemann et al., 2013; Janevska and Tudzynski, 2018). Among all these secondary metabolites, GAs are the iconic products, as they were discovered first, widely produced in industry, extensively applied in farming and with its genetic basis of metabolism best studied.

The excellent GAs yield is the trademark of *F. fujikuroi*. Besides optimization of the fermentation processes, strain improvement is also very crucial for industrial application. In industry, the researchers traditionally engaged in random mutagenesis and screening for improved production strains. These methods usually employ extreme physical conditions or toxic chemical reagents, and require an efficient screening protocol and a large amount of labor. An advantage over knock-out mutations is that these methods may generate unpredictable mutations on genes in question rather than just simply turn them off. This is probably the main reason why these methods are favored and widely practiced.

The GAs synthesis pathway (see **Figure 4**) has been completely uncovered for about two decades (MacMillan, 1997; Tudzynski and Holter, 1998; Tudzynski, 1999; Hedden and Thomas, 2012). Many efforts have been contributed to improve the GA productivity by altering the metabolic pathway. Wiemann et al. deleted the *ppt1* gene, which is involved in the post-translational modification of some key enzymes of the non-gibberellin metabolic pathways. The mutated strain in the C-1995 strain background showed to have over twofold increased GAs yield, indicating that elimination of other metabolites synthesis might somehow reduce the metabolic burden and enhance the metabolic flow to GAs. However, this phenotype could not be reproduced in another *ppt1* mutant in the IMI58289 strain background (Wiemann et al., 2012). Albermann et al. (2013) overexpressed *ggs2*, the first gene of the gibberellin-specific pathway (see **Figure 4**), and resulted in an increased productivity of 50%. Interestingly, the increased productivity was due to the increased product of $GA_4/_7$, while the GA_3 kept its original yield. Beside this, they overexpressed a truncated HmgR protein with the N terminal deleted. The N terminal of HmgR corresponds to the regulatory domain that involved in the inhibitory feedback to the protein. The constructed strain resulted in a 2.5-fold increase in GAs yield (Albermann et al., 2013). Compared to GA3, GA_4, and GA_7 may perform more excellently when treating different plants (Kim and Miller, 2009; Curry, 2012; Qian et al., 2018). Tudzynski et al. deleted the *P450-3* gene and the resulted mutant could not produce GA_1 and GA_3 anymore, while the production of GA_4 and GA_7 was significantly increased (Tudzynski et al., 2003). Shi et al. (2019) deleted the *P450-3* gene and harvested a strain with 4.6 times increase in GA_4 and GA_7 production. Based on this strain, they additionally overexpressed the */ks* genes and the truncated *hmgR* gene. The resulted strain showed to have increased production of GA_4/GA_7 mixtures of approximately eightfold (Shi et al., 2019).

To date, the genetic engineering works are mostly base on the GAs metabolic pathway. To engineer a gibberellin specific producer, genes that relate to the pathogenicity and other secondary metabolites might be redundant and will increase metabolic burden. Beside this, there is a great potential to reveal the upstream regulation network of gibberellin metabolism. Some unknown global regulators, such as the nutrient sensing and signaling pathways, might also be critical to GAs synthesis. Some of those regulators can certainly be the future targets to get genetically engineered for a supreme GA producer.

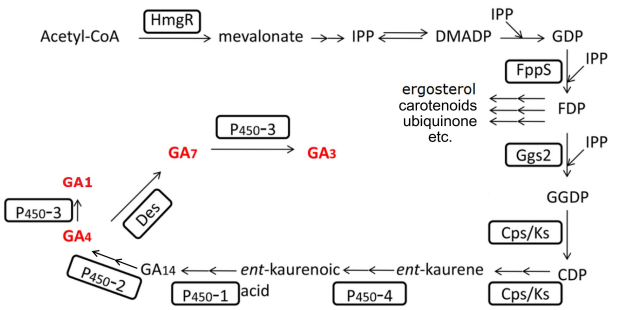


FIGURE 4 | Gibberellin biosynthesis pathway in *F. fujikuroi*. The involved enzymes are highlighted by boxes; arrows indicate sequential biosynthesis steps (double arrows indicate multiple reactions); the rest names are either the main intermediates or end products. The bioactive GAs are marked in bold red letters.

CONCLUSION

F. fujikuroi has been designated as a prevalent plant pathogen for around one century, while as an industrial GAs producer for over 50 years. The research activities concerning its virulence, drug resistance, host-pathogen interaction, metabolic pathways, signal transductions, fermentation and strain improvement have been carried out for many years. However, there are still a lot unresolved puzzles on the biology of *F. fujikuroi*. The slow progress of biological research is largely due to the sluggish development of efficient molecular tools/technologies. In the recent years, with the increased research interests in this species, the toolkit for *F. fujikuroi* is expanding rapidly.

The optimized protoplast preparation and DNA transformation protocols, the established CRISPR-Cas system, the increased choices of promoters and selection markers etc. can satisfy now the basic needs for genetic manipulations. Challenges are also emerging. Firstly, it might be interesting to knock out the NHEJ machinery to increase the efficiency of transformants screening. The HR efficiency is relatively low and varies among different isolates of this species. In classic genome engineering, usually homologous flanking arms of 500–1500 bp are required for homology directed DNA integration. Even though, the editing efficiency was not ideal. Microhomology-mediated DNA integration using 35–60 bp homologous flanking arms has been achieved in many filamentous fungi with a CRISPR-Cas

system. Optimization of the current CRISPR-Cas system is then necessary to enhance the HR efficiency for more efficient genome editing. Besides, it will be also interesting to construct a CRISPR-Cas system that is able to pop out the CRISPR-cas elements for continuous genetic manipulations. Secondly, investigation of large DNA fragment deletion is missing. A large amount of gene clusters of different secondary metabolites, secreted proteins and virulence genes are harbored in the *F. fujikuroi* genome. It would be interesting to develop a genome engineering tool to efficiently alter a large DNA fragment. Thirdly, RNAi should be tested in this microorganism to complement with the current genetic engineering methods, while development of episomal plasmids could be a valuable attempt in the research field.

There are many methods available for identification of the *Fusarium* species, such as the morphological identification, PCR methods and MALDI-TOF MS method. Among them, the morphological identification is less precise, while the MALDI-TOF MS is the most accurate and efficient with the fingerprints database of the FFSC established. However, MALDI-TOF MS requires a large facility, thus is not very popular at the current stage. The PCR methods were frequently used and also precise. This method took several conserved genes as the PCR targets. However, the amplified DNA fragments should always be necessary to be sequenced for confirmation. The *in vivo* plant infection assays have been tested extensively with *F. fujikuroi*, taking mostly the rice seedlings as the hosts. This assay can be efficiently implemented within 2 weeks and end up with different virulence parameters. Different fluorescent proteins (GFP, RFP, YFP) have been used to tag *F. fujikuroi* proteins. The BIFC assay was once tested to analyze protein-protein interaction. Further assessment of this technology is necessary, as some essential controls were missing. The FRET (fluorescence resonance energy transfer) technology is another important tool to analyze protein-protein interaction, whereas is presently missing for *F. fujikuroi* research.

Due to the economic importance of gibberellins, the fermentation technologies and strain improvement studies were better developed. However, the literatures of medium optimization are mostly from many years ago. The current industrial production conditions of SMF need to be updated. Other technologies, such as the SSF systems, were mostly carried out at the laboratory level. Further studies about the scale-up testes, system optimization and adaptation to post-fermentational processings remain to be carried out. The reported strain improvement works were mostly based on engineering of the GAs metabolic pathway. We lack knowledge about the upstream regulation/signal-transduction network. It will be also interesting to engineer a GAs production strain with the redundant pathogenesis genes and gene clusters of other secondary metabolites deleted.

AUTHOR CONTRIBUTIONS

Y-KC created the idea and wrote the first draft. J-GL participated in writing the section of CRISPR-Cas. Y-LW helped to correct **Table 2**. Y-LW and J-YW participated in writing the chapter of Fermentation technologies. Z-QL and Y-GZ finalized the draft.

FUNDING

This research was supported by Zhejiang Provincial Natural Science Foundation of China under Grant No. Q19C010006, and Educational Commission of Zhejiang Province, China under Grant No. Y201738062.

REFERENCES

Abd-Elsalam, K. A., Aly, I. N., Abdel-Satar, M. A., Khalil, M. S., and Verreet, J. A. (2003). PCR identification of *Fusarium* genus based on nuclear ribosomal-DNA sequence data. *Afri. J. Biotechnol.* 2:4.

Adam, S., Deimel, J., Pardo-Medina, J., Garcia-Martinez, T., Konte, M. C., Limon, J., et al. (2018). Protein activity of the *Fusarium fujikuroi* rhodopsins CarO and OpsA and their relation to fungus-plant interaction. *Int. J. Mol. Sci.* 19:215. doi: 10.3390/ijms19010215

Agosin, E., Maureira, M., Biffani, V., and Perez, F. (1997). "Production of gibberellins by solid substrate cultivation of *Gibberella fujikuroi*," in *Advances in Solid State Fermentation*, eds S. Roussos, B. K. Lonsane, M. Raimbault, and G. Viniegra-Gonzalez (Dordrecht: Springer), 355–366. doi: 10.1007/978-94-017-0661-2_29

Albadri, S., De Santis, F., Di Donato, V., and Del Bene, F. (2017). "CRISPR/Cas9-mediated knockin and knockout in zebrafish," in *Genome Editing in Neurosciences*, eds R. Jaenisch, F. Zhang, and F. Gage (Cham: Springer), 41–49. doi: 10.1007/978-3-319-60192-2_4

Albermann, S., Linnemannstons, P., and Tudzynski, B. (2013). Strategies for strain improvement in *Fusarium fujikuroi*: overexpression and localization of key enzymes of the isoprenoid pathway and their impact on gibberellin biosynthesis. *Appl. Microbiol. Biotechnol.* 97, 2979–2995. doi: 10.1007/s00253-012-4377-5

Al-Hatmi, A. M., Normand, A. C., van Diepeningen, A. D., Hendrickx, M., de Hoog, G. S., and Piarroux, R. (2015). Rapid identification of clinical members of *Fusarium fujikuroi* complex using MALDI-TOF MS. *Future Microbiol.* 10, 1939–1952. doi: 10.2217/fmb.15.108

Antachopoulos, J., Meletiadis, T. S., Roilides, E., and Walsh, T. J. (2007). Use of high inoculum for early metabolic signalling and rapid susceptibility testing of *Aspergillus* species. *J. Antimicrob. Chemother.* 59, 230–237. doi: 10.1093/jac/dkl488

Arras, S. D., Chua, S. M., Wizrah, M. S., Faint, J. A., Yap, A. S., and Fraser, J. A. (2016). Targeted genome editing via CRISPR in the pathogen *Cryptococcus neoformans*. *PLoS One* 11:e0164322. doi: 10.1371/journal.pone.0164322

Baayen, R. P., O'Donnell, K., Bonants, P. J. M., Cigelnik, E., Kroon, L. P. N. M., Roebroeck, E. J. A., et al. (2000). Gene genealogies and AFLP analyses in the *Fusarium oxysporum* complex identify monophyletic and nonmonophyletic formae speciales causing wilt and rot disease. *Phytopathology* 90, 891–900. doi: 10.1094/PHYTO.2000.90.8.891

Bandelier, S., Renaud, R., and Durand, A. (1997). Production of gibberellic acid by fed-batch solid state fermentation in an aseptic pilot-scale reactor. *Process Biochem.* 32, 141–145. doi: 10.1016/s0032-9592(96)00063-5

Bashyal, B. M., Rawat, K., Sharma, S., Kulshreshtha, D., Gopala Krishnan, S., Singh, A. K., et al. (2017). Whole genome sequencing of *Fusarium fujikuroi* provides insight into the role of secretory proteins and cell wall degrading enzymes in causing bakanae disease of rice. *Front. Plant Sci.* 8:2013. doi: 10.3389/fpls.2017.02013

Bernardi-Wenzel, J., Quecine, M. C., Azevedo, J. L., and Pamphile, J. A. (2016). Agrobacterium-mediated transformation of *Fusarium proliferatum*. *Genet. Mol. Res.* 15:15027944. doi: 10.4238/gmr.15027944

Blake, P. S., Taylor, D. R., Crisp, C. M., Mander, L. N., and Owen, D. J. (2000). of endogenous gibberellins in strawberry, including the novel gibberellins GA_{123}, GA_{124} and GA_{125}. *Phytochemistry* 55, 887–890. doi: 10.1016/s0031-9422(00)00237-5

Bomke, C., and Tudzynski, B. (2009). Diversity, regulation, and evolution of the gibberellin biosynthetic pathway in fungi compared to plants and bacteria. *Phytochemistry* 70, 1876–1893. doi: 10.1016/j.phytochem.2009.05.020

Bomke, C., Rojas, M. C., Hedden, P., and Tudzynski, B. (2008). Loss of Gibberellin production in *Fusarium verticillioides* (*Gibberella fujikuroi* MP-A) Is due to a deletion in the gibberellic acid gene cluster. *Appl. Environ. Microb.* 74, 7790–7801. doi: 10.1128/AEM.01819-08

Borrow, P. W., Brian, V. E. C., Curtis, P. J., Hemming, H. G., Henehan, C., Jeffreys, E. G., et al. (1955). Gibberellic acid, a metabolic product of the fungus *Gibberella fujikuroi* : some observations on its production and isolation. *J. Sci. Food Agric.* 6, 340–348. doi: 10.1002/jsfa.2740060609

Borrow, S., Brown, E.G., Kessell, R. H., Lloyd, E. C., Lloyd, P. B., Rothwell, A., et al. (1964). The effect of varied temperature on the kinetics of metabolism of *Gibberella fujikuroi* in stirred culture. *Can. J. Microbiol.* 10, 445–466. doi: 10.1139/m64-055

Boylan, M. T., Holland, M. J., and Timberlake, W. E. (1986). Saccharomyces cerevisiae centromere *CEN1*1 does not induce chromosome instability when integrated into the *Aspergillus nidulans* genome. *Mol. Cell. Biol.* 6, 3621–3625. doi: 10.1128/mcb.6.11.3621

Briggs, D. E. (1966). Gibberellin-like activity of helminthosporol and helminthosporic acid. *Nature* 210:2.

Bruckner, B., Blechschmidt, D., and Recknagel, R. D. (1991). Optimization of nutrient medium for biosynthesis of gibberellic-acid. *J. Basic Microb.* 31, 243–250.

Bruckner, B., Unkles, S. E., Weltring, K., and Kinghorn, J. R. (1992). Transformation of *Gibberella fujikuroi*: effect of the *Aspergillus nidulans* AMA1 sequence on frequency and integration. *Curr. Genet.* 22, 313–316. doi: 10.1007/bf00317927

Brunner, K., Farnleitner, A., and Mach, R. L. (2012). *Novel Methods for the Quantification of Pathogenic Fungi in Crop Plants: Quantitative PCR and ELISA Accurately Determine Fusarium Biomass*. London: IntechOpen.

Castrillo, M., Garcia-Martinez, J., and Avalos, J. (2013). Light-dependent functions of the *Fusarium fujikuroi* CryD DASH cryptochrome in development and secondary metabolism. *Appl. Environ. Microbiol.* 79, 2777–2788. doi: 10.1128/aem.03110-12

Cen, Y. K., Fiori, A., and Van Dijck, P. (2015). Deletion of the DNA ligase IV gene in *Candida glabrata* significantly increases gene-targeting efficiency. *Eukaryotic Cell* 14, 783–791. doi: 10.1128/EC.00281-14

Cen, Y. K., Lin, J. G., Wang, J. Y., Liu, Z. Q., and Zheng, Y. G. (2018). Colorimetric assay for active biomass quantification of *Fusarium fujikuroi*. *J. Microbiol. Methods* 155, 37–41. doi: 10.1016/j.mimet.2018.11.009

Cen, Y., Timmermans, B., Souffriau, B., Thevelein, J. M., and Van Dijck, P. (2017). Comparison of genome engineering using the CRISPR-Cas9 system in *C. glabrata* wild-type and *lig4* strains. *Fungal Genet. Biol.* 107, 44–50. doi: 10.1016/j.fgb.2017.08.004

Chung, M. E., Yeh, I. H., Sung, L. Y., Wu, M. Y., Chao, Y. P., Ng, I. S., et al. (2017). Enhanced integration of large DNA into *E. coli* chromosome by CRISPR/Cas9. *Biotechnol. Bioeng.* 114, 172–183. doi: 10.1002/bit.26056

Chylinski, K., Makarova, K. S., Charpentier, E., and Koonin, E. V. (2014). Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Res.* 42, 6091–6105. doi: 10.1093/nar/gku241

Corona, A., Sáez, D., and Agosin, E. (2005). Effect of water activity on gibberellic acid production by *Gibberella fujikuroi* under solid-state fermentation conditions. *Process Biochem.* 40, 2655–2658. doi: 10.1016/j.procbio.2004.11.008

Curry, E. (2012). Increase in epidermal planar cell density accompanies decreased russeting of 'golden delicious' apples treated with gibberellin A(4+7). *Hortscience* 47, 232–237. doi: 10.21273/hortsci.47.2.232

da Silva, M. E., Ferreira, M. R. K., Savoldi, M., Goldman, M. H., Hartl, A., Heinekamp, T., et al. (2006). The *akuB*(*KU80*) mutant deficient for nonhomologous end joining is a powerful tool for analyzing pathogenicity in *Aspergillus fumigatus*. *Eukaryot Cell* 5, 207–211.

Daley, J. M., Palmbos, P. L., Wu, D. L., and Wilson, T. E. (2005). Nonhomologous end joining in yeast. *Annu. Rev. Genet.* 39, 431–451.

Darken, M. A., Jensen, A. L., and Shu, P. (1959). Production of gibberellic acid by fermentation. *Appl. Microbiol.* 7, 301–303. doi: 10.1128/aem.7.5.301-303.1959

de Biazio, G. R., Leite, G. G., Tessmann, D. J., and Barbosa-Tessmann, I. P. (2008). A new PCR approach for the identification of *Fusarium graminearum*. *Braz. J. Microbiol.* 39, 554–560. doi: 10.1590/S1517-838220080003000028

de Groot, M. J., Bundock, P., Hooykaas, P. J., and Beijersbergen, A. G. (1998). *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nat. Biotechnol.* 16, 839–842. doi: 10.1038/nbt0998-839

de Oliveira, J., Rodrigues, C., Vandenberghe, L. P. S., Camara, M. C., Libardi, N., and Soccol, C. R. (2017). Gibberellic acid production by different fermentation systems using citric pulp as substrate/support. *Biomed. Res. Int.* 2017:5191046. doi: 10.1155/2017/5191046

Desai, P. V., and Panchal, R. R. (2016). Study of gibberellic acid production by solid state fermentation using *Fusarium moniliforme* sheldon. *Int. J. Appl. Sci. Biotechnol.* 4, 402–407. doi: 10.3126/ijasbt.v4i3.15588

Desmond, O. J., Manners, J. M., Stephens, A. E., Maclean, D. J., Schenk, P. M., Gardiner, D. M., et al. (2008). The *Fusarium* mycotoxin deoxynivalenol elicits hydrogen peroxide production, programmed cell death and defence responses in wheat. *Mol. Plant Pathol.* 9, 435–445. doi: 10.1111/j.1364-3703.2008.00475.x

Diaz, P., Villanueva, V. O., Gil-Duran, C., Fierro, F., Chavez, R., and Vaca, I. (2019). Genetic transformation of the filamentous fungus *Pseudogymnoascus verrucosus* of antarctic origin. *Front. Microbiol.* 10:2675. doi: 10.3389/fmicb.2019.02675

DiCarlo, J. E., Norville, J. E., Mali, P., Rios, X., Aach, J., and Church, G. M. (2013). Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* 41, 4336–4343. doi: 10.1093/nar/gkt135

Doudna, J. A., and Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science* 346:1077.

Eckardt, N. A. (2002). Foolish seedlings and DELLA regulators: the functions of rice *SLR1* and *Arabidopsis RGL1* in GA signal transduction. *Plant Cell* 14, 1–5. doi: 10.1105/tpc.140110

Enkler, L., Richer, D., Marchand, A. L., Ferrandon, D., and Jossinet, F. (2016). Genome engineering in the yeast pathogen *Candida glabrata* using the CRISPR-Cas9 system. *Sci. Rep.* 6:35766. doi: 10.1038/srep35766

Ermert, A. L., Nogue, F., Stahl, F., Gans, T., and Hughes, J. (2019). /Cas9-mediated knockout of *Physcomitrella patens* phytochromes. *Methods Mol. Biol.* 2026, 237–263. doi: 10.1007/978-1-4939-9612-4_20

Faria, C. B., Abe, C. A., da Silva, C. N., Tessmann, D. J., and Barbosa-Tessmann, I. P. (2012). New PCR assays for the identification of *Fusarium verticillioides*, *Fusarium subglutinans*, and other species of the *Gibberella fujikuroi* complex. *Intern. J. Mol. Sci.* 13, 115–132. doi: 10.3390/ijms13010115

Feckler, A., Schrimpf, M. B., Barlocher, F., Baudy, P., Cornut, J., and Schulz, R. (2017). Quantitative real-time PCR as a promising tool for the detection and quantification of leaf-associated fungal species - A proof-of-concept using *Alatospora pulchella*. *PLoS One* 12:e0174634. doi: 10.1371/journal.pone.0174634

Fernandez-Martin, R., Cerda-Olmedo, E., and Avalos, J. (2000). Homologous recombination and allele replacement in transformants of *Fusarium fujikuroi*. *Mol. Gen. Genet.* 263, 838–845. doi: 10.1007/s004380000249

Ferrara, M., Haidukowski, M., Logrieco, A. F., Leslie, J. F., and Mule, G. (2019). Cas9 system for genome editing of *Fusarium proliferatum*. *Sci. Rep.* 9:19836. doi: 10.1038/s41598-019-56270-9

Fuller, K. K., Chen, S., Loros, J. J., and Dunlap, J. C. (2015). Development of the CRISPR/Cas9 system for targeted gene disruption in *Aspergillus fumigatus*. *Eukaryot Cell* 14, 1073–1080. doi: 10.1128/EC.00107-15

Gancheva, V., Dimova, T., Kamenov, K., and Futekova, M. (1984). Biosynthesis of gibberellins. III. Optimization of the nutrient medium for gibberellin biosynthesis by using mathematical methods to plan experiments. *Acta Microbiol. Bulg.* 14, 80–84.

Garcia-Martinez, J., Brunk, M., Avalos, J., and Terpitz, U. (2015). The CarO rhodopsin of the fungus *Fusarium fujikuroi* is a light-driven proton pump that retards spore germination. *Sci. Rep.* 5:7798.

Gelmi, C., Érez-Correa, R. P., and Agosin, E. (2002). Modelling *Gibberella fujikuroi* growth and GA_3 production in solid-state fermentation. *Process Biochem.* 37, 1033–1040. doi: 10.1016/s0032-9592(01)00314-4

Gems, D., Johnstone, I. L., and Clutterbuck, A. J. (1991). An autonomously replicating plasmid transforms *Aspergillus nidulans* at high frequency. *Gene* 98, 61–67. doi: 10.1016/0378-1119(91)90104-j

Goins, C. L., Gerik, K. J., and Lodge, J. K. (2006). Improvements to gene deletion in the fungal pathogen *Cryptococcus neoformans*: absence of Ku proteins increases homologous recombination, and co-transformation of independent DNA molecules allows rapid complementation of deletion phenotypes. *Fungal Genet. Biol.* 43, 531–544. doi: 10.1016/j.fgb.2006.02.007

Gokdere, M., and Ates, S. (2014). Extractive fermentation of gibberellic acid with free and immobilized *Gibberella fujikuroi*. *Prep. Biochem. Biotech.* 44, 80–89. doi: 10.1080/10826068.2013.792275

Goldoni, M., Azzalin, G., Macino, G., and Cogoni, C. (2004). Efficient gene silencing by expression of double stranded RNA in *Neurospora crassa*. *Fungal Genet. Biol.* 41, 1016–1024. doi: 10.1016/j.fgb.2004.08.002

Halder, S. K., Maity, C., Jana, A., Ghosh, K., Das, A., Paul, T., et al. (2014). Chitinases biosynthesis by immobilized *Aeromonas hydrophila* SBK1 by prawn shells valorization and application of enzyme cocktail for fungal protoplast preparation. *J. Biosci. Bioeng.* 117, 170–177. doi: 10.1016/j.jbiosc.2013.07.011

Hedden, P., and Thomas, S. G. (2012). Gibberellin biosynthesis and its regulation. *Biochem. J.* 444, 11–25. doi: 10.1042/BJ20120245

Hoff, B., and Kuck, U. (2005). Use of bimolecular fluorescence complementation to demonstrate transcription factor interaction in nuclei of living cells from the filamentous fungus *Acremonium chrysogenum*. *Curr. Genet.* 47, 132–138. doi: 10.1007/s00294-004-0546-0

Hsuan, H. M., Salleh, B., and Zakaria, L. (2011). Molecular identification of *Fusarium* species in *Gibberella fujikuroi* species complex from rice, sugarcane and maize from peninsular Malaysia. *Int. J. Mol. Sci.* 12, 6722–6732. doi: 10.3390/ijms12106722

Huschek, D., and Witzel, K. (2019). Rapid dereplication of microbial isolates using matrix-assisted laser desorption ionization time-of-flight mass spectrometry: a mini-review. *J. Adv. Res.* 19, 99–104. doi: 10.1016/j.jare.2019.03.007

Islam, M. N., Nizam, S., and Verma, P. K. (2012). A highly efficient *Agrobacterium* mediated transformation system for chickpea wilt pathogen *Fusarium oxysporum* f. sp. ciceri using DsRed-Express to follow root colonisation. *Microbiol. Res.* 167, 332–338. doi: 10.1016/j.micres.2012.02.001

Janevska, S., and Tudzynski, B. (2018). Secondary metabolism in *Fusarium fujikuroi*: strategies to unravel the function of biosynthetic pathways. *Appl. Microbiol. Biotechnol.* 102, 615–630. doi: 10.1007/s00253-017-8679-5

Janevska, S., Arndt, B., Baumann, L., Apken, L. H., Mauriz Marques, L. M., Humpf, H. U., et al. (2017). Establishment of the inducible tet-on system for the activation of the silent trichosetin gene cluster in *Fusarium fujikuroi*. *Toxins* 9:126. doi: 10.3390/toxins9040126

Jeong, H., Lee, S., Choi, G. J., Lee, T., and Yun, S. H. (2013). Draft genome sequence of *Fusarium fujikuroi* B14, the causal agent of the bakanae disease of rice. *Genome Announc.* 1:e0035-13.

Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., and Yang, S. (2015). Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl. Environ. Microbiol.* 81, 2506–2514. doi: 10.1128/AEM.04023-14

Katayama, T., Tanaka, Y., Okabe, T., Nakamura, H., Fujii, W., Kitamoto, K., et al. (2016). Development of a genome editing technique using the CRISPR/Cas9 system in the industrial filamentous fungus *Aspergillus oryzae*. *Biotechnol. Lett.* 38, 637–642. doi: 10.1007/s10529-015-2015-x

Kazan, K., Gardiner, D. M., and Manners, J. M. (2012). On the trail of a cereal killer: recent advances in *Fusarium graminearum* pathogenomics and host resistance. *Mol. Plant Pathol.* 13, 399–413. doi: 10.1111/j.1364-3703.2011.00762.x

Ke, X. L., Lu, M. X., and Wang, J. G. (2016). of *Fusarium solani* species complex from infected zebrafish (Danio rerio). *J. Vet. Diagn. Invest.* 28, 688–692. doi: 10.1177/1040638716669539

Kim, E., Kim, H. J., Yang, S. M., Kim, C. G., Choo, D. W., and Kim, H. Y. (2019). Rapid identification of *Staphylococcus* species isolated from food samples by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J. Microbiol. Biotechnol.* 29, 548–557. doi: 10.4014/jmb.1901.01046

Kim, H. J., and Miller, W. B. (2009). GA(4+7) plus BA enhances postproduction quality in pot tulips. *Postharvest. Biol. Technol.* 51, 272–277. doi: 10.1016/j.postharvbio.2008.07.002

Klittich, C., and Leslie, J. F. (1988). Nitrate reduction mutants of *fusarium moniliforme* (*gibberella fujikuroi*). *Genetics* 118, 417–423.

Kooistra, R., Hooykaas, P. J., and Steensma, H. Y. (2004). Efficient gene targeting in *Kluyveromyces lactis*. *Yeast* 21, 781–792. doi: 10.1002/yea.1131

Kuhr, I., Podojil, M., and Sevcik, V. (1961). The influence of the nitrogen source on the production of gibberellic acid in submerse cultivation of *Gibberella fujikuroi*. *Folia Microbiol.* 6, 18–21. doi: 10.1007/bf02869590

Kumar, P. K. R., and Lonsane, B. K. (1986). Gibberellic acid by solid state fermentation : consistent and improved yields. *Biotechnol. Bioeng.* 30, 267–271. doi: 10.1002/bit.260300217

Kvas, M., Marasas, W. F. O., Wingfield, B. D., Wingfield, M. J., and Steenkamp, E. T. (2009). Diversity and evolution of *Fusarium* species in the *Gibberella fujikuroi* complex. *Fungal Divers.* 34, 1–21.

Lacmanova, J., Pazlarova, M., Kostelanska, L., and Hajslova, J. (2009). PCR-based identification of toxinogenic *Fusarium* Species. *Czech. J. Food Sci.* 27, 90–94. doi: 10.17221/634-cjfs

Leslie, J. F., and Summerell, B. A. (2006). *The Fusarium Laboratory Manual*. Ames: Blackwell Publishing.

Lin, S., Staahl, B., Alla, R. K., and Doudna, J. A. (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife* 3:e04766. doi: 10.7554/eLife.04766

Linnemannstons, P., Voss, T., Hedden, P., Gaskin, P., and Tudzynski, B. (1999). Deletions in the gibberellin biosynthesis gene cluster of *Gibberella fujikuroi* by restriction enzyme-mediated integration and conventional transformation-mediated mutagenesis. *Appl. Environ. Microbiol.* 65, 2558–2564. doi: 10.1128/aem.65.6.2558-2564.1999

Liu, H., Cottrell, T. R., Pierini, L. M., Goldman, W. E., and Doering, T. L. (2002). interference in the pathogenic fungus *Cryptococcus neoformans*. *Genetics* 160, 463–470.

Liu, Z., and Friesen, T. L. (2012). Polyethylene glycol (PEG)-mediated transformation in filamentous fungal pathogens. *Methods Mol. Biol.* 835, 365–375. doi: 10.1007/978-1-61779-501-5_21

Ma, L. J., and Xu, J. R. (2019). Shuffling effector genes through mini-chromosomes. *PLoS Genet.* 15:e1008345. doi: 10.1371/journal.pgen.1008345

Ma, L. J., Geiser, D. M., Proctor, R. H., Rooney, A. P., O'Donnell, K., Trail, F., et al. (2013). *Fusarium* pathogenomics. *Annu. Rev. Microbiol.* 67, 399–416. doi: 10.1146/annurev-micro-092412-155650

Ma, L. J., van der Does, H. C., Borkovich, K. A., Coleman, J. J., Daboussi, M. J., Di Pietro, A., et al. (2010). Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* 464, 367–373.

Machado, C. M., Oishi, B. O., Pandey, A., and Soccol, C. R. (2004). Kinetics of *Gibberella fujikuroi* growth and gibberellic acid production by solid-state fermentation in a packed-bed column bioreactor. *Biotechnol. Prog.* 20, 1449–1453. doi: 10.1021/bp049819x

Machado, C. M., Soccol, C. R., de Oliveira, B. H., and Pandey, A. (2002). Gibberellic acid production by solid-state fermentation in coffee husk. *Appl. Biochem. Biotechnol.* 103, 179–191.

MacMillan, J. (1997). Biosynthesis of the gibberellin plant hormones. *Nat. Prod. Rep.* 14, 221–243.

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., et al. (2011). and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9, 467–477. doi: 10.1038/nrmicro2577

Mathieu, M., and Felenbok, B. (1994). The *Aspergillus nidulans* CREA protein mediates glucose repression of the ethanol regulon at various levels through competition with the ALCR-specific transactivator. *EMBO J.* 13, 4022–4027. doi: 10.1002/j.1460-2075.1994.tb06718.x

McDonald, T., Brown, D., Keller, N. P., and Hammond, T. M. (2005). RNA silencing of mycotoxin production in *Aspergillus* and *Fusarium* species. *Mol. Plant Microbe. Int.* 18, 539–545.

Meletiadis, J., Mouton, J. W., Meis, J. F., Bouman, B. A., Donnelly, J. P., and Verweij, P. E. (2001). Colorimetric assay for antifungal susceptibility testing of *Aspergillus* species. *J. Clin. Microbiol.* 39, 3402–3408. doi: 10.1128/jcm.39.9.3402-3408.2001

Mesureur, J., Arend, S., Celliere, B., Courault, P., Cotte-Pattat, P. J., Totty, H., et al. (2018). A MALDI-TOF MS database with broad genus coverage for species-level identification of *Brucella*. *PLoS Neglect. Trop. Dis.* 12:e0006874. doi: 10.1371/journal.pntd.0006874

Meyer, V., Arentshorst, M., El-Ghezal, A., Drews, A. C., Kooistra, R., van den Hondel, C. A., et al. (2007). Highly efficient gene targeting in the *Aspergillus niger* kusA mutant. *J. Biotechnol.* 128, 770–775. doi: 10.1016/j.jbiotec.2006.12.021

Meyer, V., Wanka, F., van Gent, J., Arentshorst, M., van den Hondel, C. A., and Ram, A. F. (2011). Fungal gene expression on demand: an inducible, tunable, and metabolism-independent expression system for *Aspergillus niger*. *Appl. Environ. Microbiol.* 77, 2975–2983. doi: 10.1128/AEM.02740-10

Michielse, C. B., and Rep, M. (2009). Pathogen profile update: *Fusarium oxysporum*. *Mol. Plant Pathol.* 10, 311–324. doi: 10.1111/j.1364-3703.2009.00538.x

Michielse, C. B., Pfannmuller, A., Macios, M., Rengers, P., Dzikowska, A., and Tudzynski, B. (2014). The interplay between the GATA transcription factors AreA, the global nitrogen regulator and AreB in *Fusarium fujikuroi*. *Mol. Microbiol.* 91, 472–493. doi: 10.1111/mmi.12472

Mihlan, M., Homann, V., Liu, T. W. D., and Tudzynski, B. (2003). AreA directly mediates nitrogen regulation of gibberellin biosynthesis in *Gibberella fujikuroi*, but its activity is not affected by NMR. *Mol. Microbiol.* 47, 975–991. doi: 10.1046/j.1365-2958.2003.03326.x

Mink, M., Holtke, H. J., Kessler, C., and Ferenczy, L. (1990). Endonuclease-free, protoplast-forming enzyme preparation and its application in fungal transformation. *Enzy. Microb. Technol.* 12, 612–615. doi: 10.1016/0141-0229(90)90135-d

Mitchell, J. W., Skaggs, D. P., and Anderson, W. P. (1951). Plant growth-stimulating hormones in immature bean seeds. *Science* 114, 159–161. doi: 10.1126/science.114.2954.159

Miyazaki, S., Jiang, K., Kobayashi, M., Asami, T., and Nakajima, M. (2017). Helminthosporic acid functions as an agonist for gibberellin receptor. *Biosci. Biotech. Biochem.* 81, 2152–2159. doi: 10.1080/09168451.2017.1381018

Mizutani, O., Masaki, K., Gomi, K., and Iefuji, H. (2012). Modified Cre-loxP recombination in *Aspergillus oryzae* by direct introduction of Cre recombinase for marker gene rescue. *Appl. Environ. Microb.* 78, 4126–4133. doi: 10.1128/AEM.00080-12

Morgens, D. W., Deans, R. M., Li, A., and Bassik, M. C. (2016). Systematic comparison of CRISPR/Cas9 and RNAi screens for essential genes. *Nat. Biotechnol.* 34, 634–636. doi: 10.1038/nbt.3567

Morgens, D. W., Wainberg, M., Boyle, E. A., Ursu, O., Araya, C. L., Tsui, C. K., et al. (2017). Genome-scale measurement of off-target activity using Cas9 toxicity in high-throughput screens. *Nat. Commun.* 8:15178. doi: 10.1038/ncomms15178

Moss, B. J., Kim, Y., Nandakumar, M. P., and Marten, M. R. (2008). Quantifying metabolic activity of filamentous fungi using a colorimetric XTT assay. *Biotechnol. Prog.* 24, 780–783. doi: 10.1021/bp070334t

Mouyna, C., Henry, T. L., Doering, L., and Latge, J. P. (2004). Gene silencing with RNA interference in the human pathogenic fungus *Aspergillus fumigatus*. *FEMS Microbiol. Lett.* 237, 317–324. doi: 10.1111/j.1574-6968.2004.tb09713.x

Munoz, D. M., Cassiani, P. J., Li, L., Billy, E., Korn, J. M., Jones, M. D., et al. (2016). CRISPR screens provide a comprehensive assessment of cancer vulnerabilities but generate false-positive hits for highly amplified genomic regions. *Cancer Discov.* 6, 900–913. doi: 10.1158/2159-8290.CD-16-0178

Nagy, R., Taborhegyi, E., Wittner, A., and Hornok, L. (1995). Minichromosomes in *Fusarium sporotrichioides* are mosaics of dispersed repeats and unique sequences. *Microbiology* 141, 713–719. doi: 10.1099/13500872-141-3-713

Niehaus, E. M., Kim, H. K., Munsterkotter, M., Janevska, S., Arndt, B., Kalinina, S. A., et al. (2017). Comparative genomics of geographically distant *Fusarium fujikuroi* isolates revealed two distinct pathotypes correlating with secondary metabolite profiles. *PLoS Pathog.* 13:e1006670. doi: 10.1371/journal.ppat.1006670

Niessen, M. L., and Vogel, R. F. (1997). Specific identification of *Fusarium graminearum* by PCR with gaoA targeted primers. *Syst. Appl. Microbiol.* 20, 111–123.

Ninomiya, Y., Suzuki, K., Ishii, C., and Inoue, H. (2004). Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining. *Proc. Natl. Acad. Sci. U.S.A.* 101, 12248–12253. doi: 10.1073/pnas.0402780101

Nino-Sanchez, J., Casado-Del Castillo, V., Tello, V., De Vega-Bartol, J. J., Ramos, B., Sukno, S. A., et al. (2016). The *FTF* gene family regulates virulence and expression of SIX effectors in *Fusarium oxysporum*. *Mol. Plant Pathol.* 17, 1124–1139. doi: 10.1111/mpp.12373

Nodvig, C. S., Nielsen, J. B., Kogle, M. E., and Mortensen, U. H. (2015). Cas9 system for genetic engineering of filamentous fungi. *PLoS One* 10:e0133085. doi: 10.1371/journal.pone.0133085

O'Donnell, K., Kistler, H. C., Tacke, B. K., and Casper, H. H. (2000). Gene genealogies reveal global phylogeographic structure and reproductive isolation among lineages of *Fusarium graminearum*, the fungus causing wheat scab. *Proc. Natl. Acad. Sci. U.S.A.* 97, 7905–7910. doi: 10.1073/pnas.130193297

O'Donnell, K., Rooney, A. P., Proctor, R. H., Brown, D. W., McCormick, S. P., Ward, T. J., et al. (2013). Phylogenetic analyses of *RPB1* and *RPB2* support a middle Cretaceous origin for a clade comprising all agriculturally and medically important fusaria. *Fungal Genet. Biol.* 52, 20–31. doi: 10.1016/j.fgb.2012.12.004

Ortega, S. F., Tomlinson, J., Hodgetts, J., Spadaro, D., Gullino, M. L., and Boonham, N. (2018). Development of loop-mediated isothermal aamplification assays for the detection of seedborne fungal pathogens *Fusarium fujikuroi* and *Magnaporthe oryzae* in rice seed. *Plant Dis.* 102, 1549–1558. doi: 10.1094/PDIS-08-17-1307-RE

Otálvaro, ÁM., Gutiérrez, G. D., Pierotty, D. A., Parada, F. A., and Algecira, N. A. (2008). Gibberellic acid production by *Gibberella fujikuroi* under solid-state fermentation of cassava bagasse and rice hull. *J. Biotechnol.* 136:S371. doi: 10.1016/j.jbiotec.2008.07.853

Pastrana, L. M., Gonzalez, M. P., Pintado, J., and Murado, M. A. (1995). Interactions affecting gibberellic-acid production in solid-state culture - a factorial study. *Enzy. Microb. Technol.* 17, 784–790. doi: 10.1016/0141-0229(94)00024-l

Patil, N. S., Waghmare, S. R., and Jadhav, J. P. (2013). Purification and characterization of an extracellular antifungal chitinase from *Penicillium ochrochloron* MTCC 517 and its application in protoplast formation. *Process Biochem.* 48, 176–183. doi: 10.1016/j.procbio.2012.11.017

Pauker, V. I., Thoma, B. R., Grass, G., Bleichert, P., Hanczaruk, M., Zoller, L., et al. (2018). Improved discrimination of *Bacillus anthracis* from closely related species in the *Bacillus cereus* sensu lato group based on matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J. Clin. Microbiol.* 56:e01900-17. doi: 10.1128/JCM.01900-17

Peng, Z., Oliveira-Garcia, E., Lin, G. F., Hu, Y., Dalby, M., Migeon, P., et al. (2019). Effector gene reshuffling involves dispensable mini-chromosomes in the wheat blast fungus. *PLoS Genet.* 15:e1008272. doi: 10.1371/journal.pgen.1008272

Pfannmuller, D., Wagner, C. S., Schonig, B., Boeckstaens, M., Marini, A. M., and Tudzynski, B. (2015). The general amino acid permease FfGap1 of *Fusarium fujikuroi* is sorted to the vacuole in a nitrogen-dependent, but Npr1 kinase-independent manner. *PLoS One* 10:e0125487. doi: /10.1371/journal.pone.0125487

Pinto, C. H., Zahra, M., van Hal, S., Olma, T., Maszewska, K., Iredell, J. R., et al. (2011). Matrix-assisted laser desorption ionization-time of flight mass spectrometry identification of yeasts is contingent on robust reference spectra. *PLoS One* 6:e25712. doi: 10.1371/journal.pone.0025712

Pohl, C., Kiel, J. A., Driessen, A. J., Bovenberg, R. A., and Nygard, Y. (2016). /Cas9 based genome editing of *Penicillium chrysogenum*. *ACS Synth. Biol.* 5, 754–764. doi: 10.1021/acssynbio.6b00082

Prado, M. M., Prado-Cabrero, A., Fernandez-Martin, R., and Avalos, J. (2004). A gene of the opsin family in the carotenoid gene cluster of *Fusarium fujikuroi*. *Curr. Genet.* 46, 47–58.

Pringle, R. B. (1976). Comparative biochemistry of the phytopathogenic fungus Helminthosporium. XVI. The production of victoxinine by *H. sativum* and *H. victoriae*. *Can. J. Biochem.* 54, 783–787. doi: 10.1139/o76-112

Pu, X., Liu, L., Li, P., Huo, H., Dong, X., Xie, K., et al. (2019). A CRISPR/LbCas12a-based method for highly efficient multiplex gene editing in *Physcomitrella patens*. *Plant J.* 100, 863–872. doi: 10.1111/tpj.14478

Qian, C., Ren, N., Wang, J., Xu, Q., Chen, X., and Qi, X. (2018). Effects of exogenous application of CPPU, NAA and GA4+7 on parthenocarpy and fruit quality in cucumber (*Cucumis sativus L.*). *Food Chem.* 243, 410–413. doi: 10.1016/j.foodchem.2017.09.150

Qian, X. M., Dupreez, J. C., and Kilian, S. G. (1994). Factors affecting gibberellic-acid production by *Fusarium Moniliforme* in solid-state cultivation on starch. *World J. Microb. Biot.* 10, 93–99. doi: 10.1007/BF00357571

Quero, L., Girard, V., Pawtowski, A., Treguer, S., Weill, A., Arend, S., et al. (2019). Development and application of MALDI-TOF MS for identification of food spoilage fungi. *Food Microbiol.* 81, 76–88. doi: 10.1016/j.fm.2018.05.001

Raja, H. A., Miller, A. N., Pearce, C. J., and Oberlies, N. H. (2017). Fungal identification using molecular tools: a primer for the natural products research community. *J. Nat. Prod.* 80, 756–770. doi: 10.1021/acs.jnatprod.6b01085

Rangaswamy, V. (2012). Improved production of gibberellic acid by *Fusarium moniliforme*. *J. Microbiol. Res.* 2, 51–55. doi: 10.5923/j.microbiology.20120203.02

Rangaswamy, V., and Balu, G. (2010). *Process for Gibberellic Acid Production with Fusarium Moniliforme Strains*. Rabale: Reliance Life Sciences Pvt. Ltd.

Reuss, O., Vik, A., Kolter, R., and Morschhauser, J. (2004). The *SAT1* flipper, an optimized tool for gene disruption in *Candida albicans*. *Gene* 341, 119–127. doi: 10.1016/j.gene.2004.06.021

Rodrigues, C., Vandenberghe, L. P. D., Teodoro, J., Oss, J. F., Pandey, A., and Soccol, C. R. (2009). A new alternative to produce gibberellic acid by solid state fermentation. *Braz. Arch. Biol. Technol.* 52, 181–188. doi: 10.1590/s1516-89132009000700023

Rodrigues, C., Vandenberghe, L. P., de Oliveira, J., and Soccol, C. R. (2012). New perspectives of gibberellic acid production: a review. *Crit. Rev. Biotechnol.* 32, 263–273. doi: 10.3109/07388551.2011.615297

Rosler, S. M., Sieber, C. M., Humpf, H. U., and Tudzynski, B. (2016). Interplay between pathway-specific and global regulation of the fumonisin gene cluster in the rice pathogen *Fusarium fujikuroi*. *Appl. Microbiol. Biotechnol.* 100, 5869–5882. doi: 10.1007/s00253-016-7426-7

Sanchez-Fernandez, R., Unkles, S. E., Campbell, E. I., Macro, J. A., Cerda-Olmedo, E., and Kinghorn, J. R. (1991). Transformation of the filamentous fungus *Gibberella fujikuroi* using the *Aspergillus niger niaD* gene encoding nitrate reductase. *Mol. Gen. Genet.* 225, 231–233. doi: 10.1007/bf00269853

Satpute, D., Sharma, V., Murarkar, K., Bhotmange, M., and Dharmadhikari, D. (2010). Solid-state fermentation for production of gibberellic acid using agricultural residues. *Int. J. Environ. Pollut.* 43, 201–213.

Schorsch, C., Kohler, T., and Boles, E. (2009). Knockout of the DNA ligase IV homolog gene in the sphingoid base producing yeast *Pichia ciferrii* significantly increases gene targeting efficiency. *Curr. Genet.* 55, 381–389. doi: 10.1007/s00294-009-0252-z

Shi, T.-Q., Gao, J., Wang, W.-J., Wang, K.-F., Xu, G.-Q., Huang, H., et al. (2019). CRISPR/Cas9-Based genome editing in the filamentous fungus *Fusarium fujikuroi* and its application in strain engineering for gibberellic acid production. *ACS Synth. Biol.* 8, 445–454. doi: 10.1021/acssynbio.8b00478

Skovgaard, K., Nirenberg, H. I., O'Donnell, K., and Rosendahl, S. (2001). of *Fusarium oxysporum* f. sp vasinfectum races inferred from multigene genealogies. *Phytopathology* 91, 1231–1237. doi: 10.1094/PHYTO.2001.91.12.1231

Song, R., Zhai, Q., Sun, L., Huang, E., Zhang, Y., Zhu, Y., et al. (2019). CRISPR/Cas9 genome editing technology in filamentous fungi: progress and perspective. *Appl. Microbiol. Biotechnol.* 103, 6919–6932. doi: 10.1007/s00253-019-10007-w

Song, Z. W., Vail, A., Sadowsky, M. J., and Schilling, J. S. (2014). Quantitative PCR for measuring biomass of decomposer fungi in planta. *Fungal Ecol.* 7, 39–46. doi: 10.1016/j.funeco.2013.12.004

Studt, L., Humpf, H. U., and Tudzynski, B. (2013a). Signaling governed by G proteins and cAMP is crucial for growth, secondary metabolism and sexual development in *Fusarium fujikuroi*. *PLoS One* 8:e58185. doi: 10.1371/journal.pone.0058185

Studt, L., Schmidt, F. J., Jahn, L., Sieber, C. M., Connolly, L. R., Niehaus, E. M., et al. (2013b). Two histone deacetylases, FfHda1 and FfHda2, are important for *Fusarium fujikuroi* secondary metabolism and virulence. *Appl. Environ. Microbiol.* 79, 7719–7734. doi: 10.1128/AEM.01557-13

Studt, L., Wiemann, P., Kleigrewe, K., Humpf, H. U., and Tudzynski, B. (2012). Biosynthesis of fusarubins accounts for pigmentation of *Fusarium fujikuroi* perithecia. *Appl. Environ. Microb.* 78, 4468–4480. doi: 10.1128/AEM.00823-12

Summerell, B. A., Laurence, M. H., Liew, E. C. Y., and Leslie, J. F. (2010). Biogeography and phylogeography of *Fusarium*: a review. *Fungal Divers.* 44, 3–13. doi: 10.1080/19440049.2014.984244

Takahashi, T., Masuda, T., and Koyama, Y. (2006). Enhanced gene targeting frequency in *ku70* and *ku80* disruption mutants of *Aspergillus sojae* and *Aspergillus oryzae*. *Mol. Genet. Genom.* 275, 460–470. doi: 10.1007/s00438-006-0104-1

Teichert, S., Schonig, B., Richter, S., and Tudzynski, B. (2004). Deletion of the *Gibberella fujikuroi* glutamine synthetase gene has significant impact on transcriptional control of primary and secondary metabolism. *Mol. Microbiol.* 53, 1661–1675. doi: 10.1111/j.1365-2958.2004.04243.x

Teichert, S., Wottawa, M., Schonig, B., and Tudzynski, B. (2006). Role of the *Fusarium fujikuroi* TOR kinase in nitrogen regulation and secondary metabolism. *Eukaryot. Cell* 5, 1807–1819. doi: 10.1128/ec.00039-06

Tudzynski, B. (1999). Biosynthesis of gibberellins in *Gibberella fujikuroi*: biomolecular aspects. *Appl. Microbiol. Biot.* 52, 298–310. doi: 10.1007/s002530051524

Tudzynski, B. (2005). Gibberellin biosynthesis in fungi: genes, enzymes, evolution, and impact on biotechnology. *Appl. Microbiol. Biotechnol.* 66, 597–611. doi: 10.1007/s00253-004-1805-1

Tudzynski, B. (2014). Nitrogen regulation of fungal secondary metabolism in fungi. *Front. Microbiol.* 5:656. doi: 10.3389/fmicb.2014.00656

Tudzynski, B., and Holter, K. (1998). Gibberellin biosynthetic pathway in *Gibberella fujikuroi*: evidence for a gene cluster. *Fungal Genet. Biol.* 25, 157–170.

Tudzynski, B., Mende, K., Weltring, K. M., Kinghorn, J. R., and Unkles, S. E. (1996). The *Gibberella fujikuroi* niaD gene encoding nitrate reductase: isolation, sequence, homologous transformation and electrophoretic karyotype location. *Microbiology* 142(Pt 3), 533–539. doi: 10.1099/13500872-142-3-533

Tudzynski, B., Mihlan, M., Rojas, M. C., Linnemannstons, P., Gaskin, P., and Hedden, P. (2003). Characterization of the final two genes of the gibberellin biosynthesis gene cluster of *Gibberella fujikuroi* - des and P450-3 encode GA(4) desaturase and the 13-hydroxylase, respectively. *J. Biol. Chem.* 278, 28635–28643. doi: 10.1074/jbc.m301927200

Twaruschek, K., Sporhase, P., Michlmayr, H., Wiesenberger, G., and Adam, G. (2018). New plasmids for *Fusarium* transformation allowing positive-negative selection and eficient Cre-loxP mediated marker recycling. *Front. Microbiol.* 9:1954. doi: 10.3389/fmicb.2018.01954

Ueno, K., Uno, J., Nakayama, H., Sasamoto, K., Mikami, Y., and Chibana, H. (2007). Development of a highly efficient gene targeting system induced by transient repression of *YKU80* expression in *Candida glabrata*. *Eukaryot Cell* 6, 1239–1247. doi: 10.1128/ec.00414-06

Ullan, R. V., Godio, R. P., Vaca, I., Garcia-Estrada, C., Feltrer, R., and Kosalkova, K. (2008). RNA-silencing in *Penicillium chrysogenum* and *Acremonium chrysogenum*: validation studies using beta-lactam genes expression. *J. Microbiol. Methods* 75, 209–218. doi: 10.1016/j.mimet.2008.06.001

Vyas, V. K., Barrasa, M. I., Fink, G. R., and Candida, A. (2015). A Candida *albicans* CRISPR system permits genetic engineering of essential genes and gene families. *Science Adv.* 1:e1500248. doi: 10.1126/sciadv.1500248

Wagner, D., Schmeinck, A., Morozov, I. Y., Caddick, M. X., and Tudzynski, B. (2010). The bZIP ranscription factor MeaB mediates nitrogen metabolite repression at specific loci. *Eukaryot Cell* 9, 1588–1601.

Wagner, D., Wiemann, P., Huss, K., Brandt, U., Fleissner, A., and Tudzynski, B. (2013). A sensing role of the glutamine synthetase in the nitrogen regulation network in *Fusarium fujikuroi*. *PLoS One* 8:e80740. doi: 10.1371/journal.pone.0080740

Walter, S., Nicholson, P., and Doohan, F. M. (2010). Action and reaction of host and pathogen during *Fusarium* head blight disease. *New Phytol.* 185, 54–66. doi: 10.1111/j.1469-8137.2009.03041.x

Wang, P. (2018). Two distinct approaches for CRISPR-Cas9-mediated gene editing in *Cryptococcus neoformans* and related species. *mSphere* 3:e00208-18. doi: 10.1128/mSphereDirect.00208-18

Wang, Q., Cobine, P., and Coleman, J. (2018). Efficient genome editing in *Fusarium oxysporum* based on CRISPR/Cas9 ribonucleoprotein (RNP) complexes. *Phytopathology* 108, 21–29. doi: 10.1016/j.fgb.2018.05.003

Ward, T. J., Bielawski, J. P., Kistler, H. C., Sullivan, E., and O'Donnell, K. (2002). Ancestral polymorphism and adaptive evolution in the trichothecene mycotoxin gene cluster of phytopathogenic *Fusarium*. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9278–9283. doi: 10.1073/pnas.142307199

Wiemann, P., Albermann, S., Niehaus, E. M., Studt, L., von Bargen, K. W., Brock, N. L., et al. (2012). The Sfp-type 4′-phosphopantetheinyl transferase Ppt1 of *Fusarium fujikuroi* controls development, secondary metabolism and pathogenicity. *PLoS One* 7:e37519. doi: 10.1371/journal.pone.0037519

Wiemann, P., Brown, D. W., Kleigrewe, K., Bok, J. W., Keller, N. P., Humpf, H. U., et al. (2010). FfVel1 and FfLae1, components of a velvet-like complex in *Fusarium fujikuroi*, affect differentiation, secondary metabolism and virulence. *Mol. Microbiol.* 77, 972–994. doi: 10.1111/j.1365-2958.2010.07263.x

Wiemann, P., Sieber, C. M., von Bargen, K. W., Studt, L., Niehaus, E. M., Espino, J. J., et al. (2013). Deciphering the cryptic genome: genome-wide analyses of the rice pathogen *Fusarium fujikuroi* reveal complex regulation of secondary metabolism and novel metabolites. *PLoS Pathog.* 9:e1003475. doi: 10.1371/journal.ppat.1003475

Wigmann, E. F., Behr, J., Vogel, R. F., Niessen, L., and Ms, M. A. L. D. I.-T. O. F. (2019). fingerprinting for identification and differentiation of species

within the *Fusarium fujikuroi* species complex. *Appl. Microbiol. Biotechnol.* 103, 5323–5337. doi: 10.1007/s00253-019-09794-z

Zakaria, L., Hsuan, H. M., and Salleh, B. (2011). Mating populations of *fusarium* section liseola from rice, sugarcane and maize. *Trop. Life Sci. Res.* 22, 93–101.

Zhang, C., Meng, X., Wei, X., and Lu, L. (2016). Highly efficient CRISPR mutagenesis by microhomology-mediated end joining in *Aspergillus fumigatus*. *Fungal Genet. Biol.* 86, 47–57. doi: 10.1016/j.fgb.2015.12.007

Zhang, Y., Huang, H., Zhang, B., and Lin, S. (2016). CRISPR- enhanced DNA homologous recombination for gene editing in zebrafish. *Zebrafish* 135, 107–120. doi: 10.1016/bs.mcb.2016.03.005

Zheng, X., Zheng, P., Sun, J., Kun, Z., and Ma, Y. (2018). Heterologous and endogenous U6 snRNA promoters enable CRISPR/Cas9 mediated genome editing in *Aspergillus niger*. *Fungal Biol. Biotechnol.* 5:2. doi: 10.1186/s40694-018-0047-4

Zheng, X., Zheng, P., Zhang, K., Cairns, T. C., Meyer, V., Sun, J., et al. (2019). 5S rRNA promoter for guide RNA expression enabled highly efficient CRISPR/Cas9 genome editing in *Aspergillus niger*. *ACS Synth. Biol.* 8, 1568–1574. doi: 10.1021/acssynbio.7b00456

Zhu, T. H., Wang, W. X., Yang, X., Wang, K., and Cui, Z. F. (2009). Construction of two Gateway vectors for gene expression in fungi. *Plasmid* 62, 128–133. doi: 10.1016/j.plasmid.2009.06.005

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

MINI REVIEW
published: 08 March 2020
doi: 10.3389/fbioe.2020.00293

Genetic Engineering of Filamentous Fungi for Efficient Protein Expression and Secretion

Qin Wang[1], Chao Zhong[2,3] and Han Xiao[1]**

[1] State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic & Developmental Sciences, and Laboratory of Molecular Biochemical Engineering, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China, [2] Materials and Physical Biology Division, School of Physical Science and Technology, ShanghaiTech University, Shanghai, China, [3] Materials Synthetic Biology Center, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

Filamentous fungi are considered as unique cell factories for protein production due to the high efficiency of protein secretion and superior capability of post-translational modifications. In this review, we firstly introduce the secretory pathway in filamentous fungi. We next summarize the current state-of-the-art works regarding how various genetic engineering strategies are applied for enhancing protein expression and secretion in filamentous fungi. Finally, in a future perspective, we discuss the great potential of genome engineering for further improving protein expression and secretion in filamentous fungi.

Keywords: filamentous fungi, genetic engineering, protein secretion, expression, genome engineering

OPEN ACCESS

Edited by:
Jiazhang Lian,
Zhejiang University, China

Reviewed by:
Xiao-Jun Ji,
Nanjing Tech University, China
Xiaoyun Su,
Chinese Academy of Agricultural Sciences, China

****Correspondence:***
Chao Zhong
zhongchao@shanghaitech.edu.cn
Han Xiao
smallhan@sjtu.edu.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *15 January 2020*
Accepted: *19 March 2020*
Published: *08 March 2020*

Citation:
Wang Q, Zhong C and Xiao H (2020) Genetic Engineering of Filamentous Fungi for Efficient Protein Expression and Secretion. Front. Bioeng. Biotechnol. 8:293. doi: 10.3389/fbioe.2020.00293

INTRODUCTION

Protein production has a broad application in life sciences, biotechnology, medicine and material sciences. Filamentous fungi are powerful and efficient cell factories for protein production at the industrial scale, and over half of the commercially available proteins were produced by filamentous fungi[1]. Many species of filamentous fungi are generally regarded as safe (GRAS), and exhibit superior protein secretory capability. For example, 25–30 g/L of glucoamylase was obtained from fermentation medium of *Aspergillus niger*, while *Trichoderma reesei* was able to secret 100 g/L of cellulose (Ward, 2012). Compared to prokaryotes, filamentous fungi own the mature systems for post-translational processing (e.g., glycosylation, protease cleavage, and disulfide bond formation) (Karnaukhova et al., 2007), which are indispensable for protein function and activity. Although yeasts are able to perform post-translational modification, they tend to produce proteins in the form of high mannose-type glycosylation. In contrast, filamentous fungi have less extensive hyper-mannosylation of glycoproteins, which could be directly converted to mammalian type of glycoproteins with pharmaceutical potential (Punt et al., 2002; Deshpande et al., 2008). In addition, due to the metabolic diversity, filamentous fungi can efficiently utilize many types of monosaccharides including xylose, arabinose, and galactose, while yeasts can only metabolize glucose and mannose (Cavka and Jönsson, 2014).

To further improve the production of various proteins by filamentous fungi, traditional strategies including optimization of fermentation process and obtaining beneficial mutants via random mutagenesis, were widely adopted in the past. Here, rather than providing a comprehensive view of achieving efficient protein production, we focus on summarizing the

[1] https://amfep.org/about-enzymes/

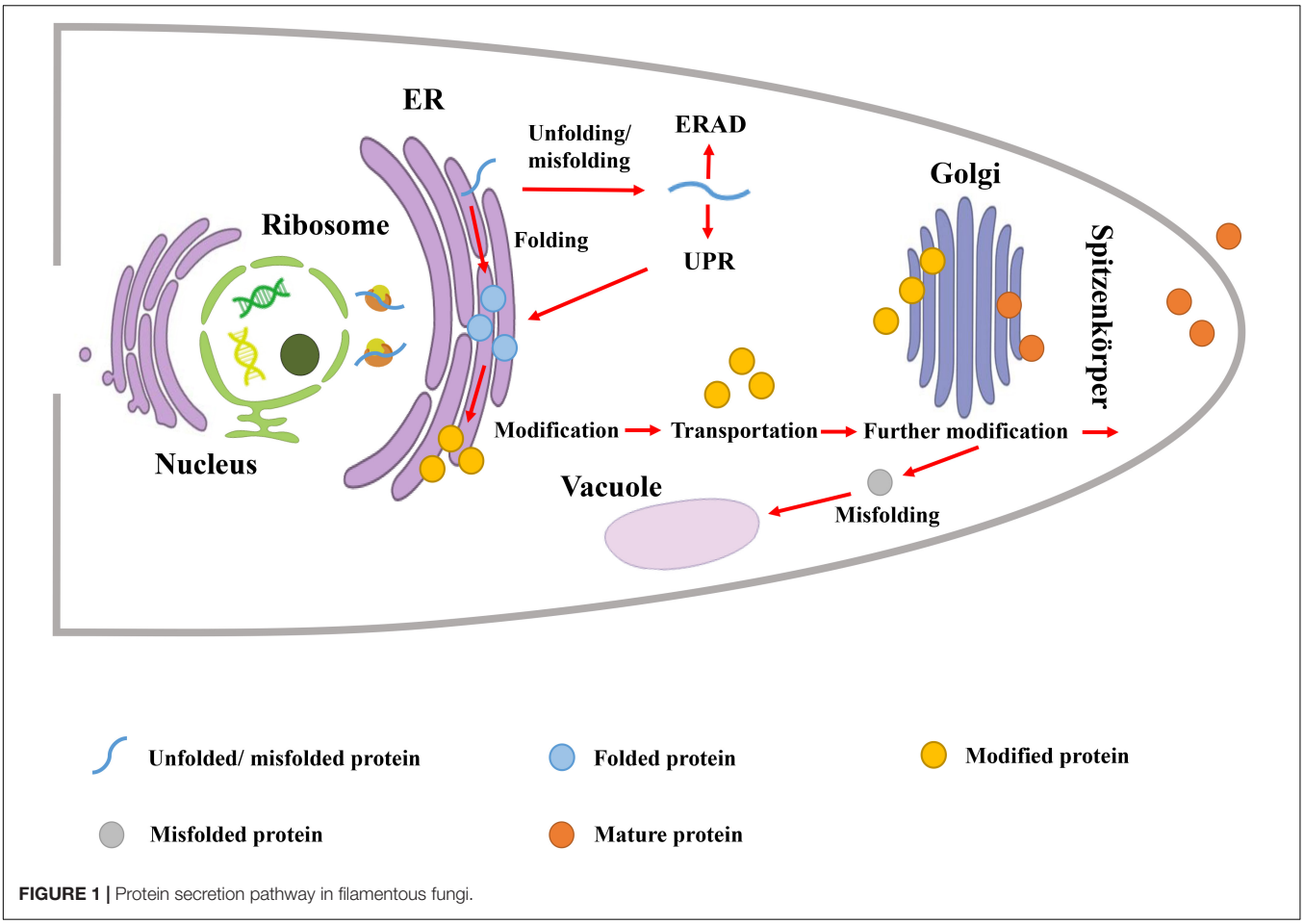


FIGURE 1 | Protein secretion pathway in filamentous fungi.

strategies based on genetic engineering of this particular cell factory to enhance protein expression and secretion. We also provide new ideas in terms of cell factory engineering.

PROTEIN SECRETION PATHWAY IN FILAMENTOUS FUNGI

Protein secretion pathway in filamentous fungi involves three major steps including: polypeptide transfer from ribosome to endoplasmic reticulum (ER), protein folding and modification in ER, transportation of the folded protein vesicles to the Golgi apparatus and extracellular environment (**Figure 1**). In the first step, the co- or post-translational transport pathway is responsible for the polypeptide transfer from the ribosome to ER. In the co-translational transport pathway, the signal peptide recognition particle (SRP) first binds to the signal peptide sequence to block translation (Halic et al., 2006). Then, SRP directs the ribosome-mRNA-nascent peptide complex to target the ER membrane and binds to the SRP receptor. Subsequently, SRP is released from the complex, translation resumes, and the nascent polypeptide enters ER lumen through the Sec61p transport complex (Conesa et al., 2001). In the post-translational transport pathway, the nascent polypeptide is translated in the cytosol, and kept unfolded by interacting with Hsp70 chaperone and co-chaperones (Conesa et al., 2001). This complex is able to target ER through interaction with the membrane receptor Sec62p-Sec72p-Sec73p subcomplex (Conesa et al., 2001). The ER luminal chaperone binding immunoglobulin protein (BiP) and the membrane protein Sec63p assist the aforementioned complex to enter ER (Haßdenteufel et al., 2018).

The second step is protein folding and modification in ER, which requires the assistance of a series of molecular chaperones and folding enzymes, including calnexin (ClxA), BiP, and protein disulfide isomerase (PDI) (Saloheimo and Pakula, 2012). For nascent peptides with correct folding, they are subjected to modifications such as glycosylation. As one of the most common and important post-translational modifications, glycosylation can significantly affect protein stability, localization, and secretion (Mitra et al., 2006). After proper folding and glycosylation, secreted proteins are transported extracellularly. On the other hand, the unfolded protein response (UPR) and ER-associated protein degradation (ERAD) are responsible for dealing with nascent peptides with incorrect folding (Bernasconi and Molinari, 2011; Wang et al., 2014). The UPR detects the presence of unfolded proteins in ER and induces the biosynthesis of chaperones and folding enzymes, while the ERAD degrades the misfolded proteins.

The third step is to transport the folded protein vesicles to the Golgi apparatus by fusion with target membrane, and secrete it to the extracellular environment (Spang, 2008). In filamentous fungi, Golgi-derived secretory vesicles are transmitted to the apical plasma membrane through apical vesicle clusters in Spitzenkörper (Virag and Harris, 2006). The formation, transportation and fusion of vesicles are mediated by a large number of proteins, including GTP-binding proteins (e.g., Sar, ARF) for vesicle budding, and Rab GTPases for fusion with Golgi (Hutagalung and Novick, 2011), etc. Specific fusion of vesicles with the target membrane is the critical process, which is mediated by soluble N-ethylmaleimide-sensitive factor-associated protein receptor (SNARE). Based on the localization, SNARE is divided into two categories: the vesicle SNARE (v-SNARE) and the target membrane SNARE (t-SNARE) (Söllner et al., 1993). In filamentous fungi, v-SNARE protein SNC1, and t-SNARE proteins SSO1 and SSO2, are involved in bubble fusion (Valkonen et al., 2007).

DIVERSE STRATEGIES FOR ENHANCED PROTEIN EXPRESSION AND SECRETION VIA GENETIC ENGINEERING

To enhance the protein expression and secretion in filamentous fungi, enhancing the intracellular protein production by optimization of the transcription and/or the codon of the target protein, is an effective strategy, as summarized in a few of reviews (Saunders et al., 1989; Jeenes et al., 1991; Nevalainen et al., 2005; Su et al., 2012). In order to bring new insights, we will discuss other genetic engineering strategies, including replacing original signal peptide with a more efficient one, fusion of heterologous protein to a naturally secreted one, regulation of UPR and ERAD, optimization of the intracellular transport process, construction of a protease-deficient strain, regulation of mycelium morphology, and optimization of the sterol regulatory element binding protein (SREBP) in this section (**Table 1**).

REPLACING ORIGINAL SIGNAL PEPTIDE WITH A MORE EFFICIENT ONE

The signal peptide sequence plays vital role in protein secretion. Replacing with a more efficient peptide in target protein tends to increase its secretion efficiency. Xu et al. replaced the original signal peptide AglB of α-galactosidase with a glucoamylase (GlaA) signal peptide in *A. niger*, and the activity of extracellular α-galactosidase increased nearly ninefold (Xu et al., 2018). Wang et al. used green fluorescent protein as a reporter gene in *P. oxalicum* to test the secretion efficiency of three signal peptides, PoxGA15A, PoxAmy13A, and PoxCbhCel7A-2. Then they selected the optimal signal peptide PoxGA15A to drive the secretion of endogenous raw starch-degrading enzymes, which was 3.4 times higher than the parental strain (Wang et al., 2018).

FUSION OF HETEROLOGOUS PROTEIN TO A NATURALLY SECRETED ONE

Fusion of heterologous protein to a naturally secreted one can enhance protein stability, promote translocation, and prevent protein from degradation. The in-frame fusion of human protein granulocyte colony stimulating factor (G-CSF) with an endogenous highly secreted glucoamylase allowed secretion of 5–10 mg/L of G-CSF by *A. niger* (Kraševec et al., 2014). When bovine chymosin (CHY) was fused with alpha-amylase (AmyB), the engineered *A. oryzae* was able to produce two times higher amount of CHY than that with none fused CHY, while multiple genes involved in ER folding and protein secretion pathway increased significantly in the fused CHY producing strain (Ohno et al., 2011). It should be noted that the fusion carrier protein could greatly affect the secretion. In order to secrete *Escherichia coli* β-glucuronidase (GUS) protein in *Penicillium funiculosum*, researchers attempted to use xylanase as a carrier. The modular structure, a catalytic domain separated from the cellulose-binding domain by a linker with serine and threonine rich sequence, enables some xylanases as a group of unique protein carrier (Alcocer et al., 2003). It was reported that xylanase A (XYNA) is an effective carrier protein, while XYNB and XYNC are ineffective (Alcocer et al., 2003).

REGULATION OF UPR AND ERAD TO PROMOTE PROTEIN SECRETION

Correct protein folding is one of the many prerequisites to protein secretion. Abnormal folding proteins could form toxic aggregates exerting pressure on the ER, and trigger the feedback regulation called repression under secretion stress (RESS) to affect protein secretion (Pakula et al., 2003). UPR and ERAD are considered as two important ways to regulate protein folding, and enhanced protein secretion could be achieved via regulation of UPR and ERAD. For example, overexpression of the transcription factor *hac1* in *Aspergillus awamori* led to 7- and 2.8-fold increases in laccase and bovine prechymotrypsin production, respectively (Valkonen et al., 2003). Overexpression of *bip1* and *hac1* in *T. reesei* exhibited 1.5- and 1.8-fold improvement on secretion of an *A. niger* glucose oxidase (Wu et al., 2017).

To avoid degradation of some heterologous proteins or semi-folded proteins, deleting key genes involved in ERAD is a solution. Deletion of the ERAD factor *doaA* and overexpression of the oligosaccharyltransferase *sttC* responsible for glycosylation of secretory proteins (Yan and Lennarz, 2002) in *A. niger* caused an increase in β-glucuronidase yield (Jacobs et al., 2009). In addition, autophagy is considered as another way to degrade the misfolded proteins (Kario et al., 2011). Disruption of autophagy-related gene *aoatg15* in *A. oryzae* caused a threefold increase in secretion of bovine chymosin (Yoon et al., 2013).

Of particular note, manipulation of certain gene may cause quite different effects in different strains. For example, overexpression of *bip1* promoted protein secretion in *T. reesei* (Wu et al., 2017) and *A. awamori* (Lombraña et al., 2004), while reduced protein secretion was observed in *A. niger* by

TABLE 1 | Typical examples for genetic engineering of filamentous fungi for enhanced protein secretion.

Protein of interest and its origin	Host	Strategy	Fold-change of protein secretion	References
α-Galactosidase from *A. niger*	*A. niger*	Replacing the original signal peptide with a glucoamylase (GlaA) signal peptide in *A. niger*	Approximately 9-fold increase	Xu et al., 2018
Erythropoietin from human	*T. reesei*	Adopting the cellobiohydrolase I (CBH) signal peptide and optimizing *cbh1* promoter	Not applicable	Zhong et al., 2011
Chymosin from bovine	*A. oryzae*	Fusing target protein with a naturally secreted protein α-amylase	2-fold increase	Ohno et al., 2011
β-Glucuronidase from *A. niger*	*A. niger*	Regulating the UPR and ERAD by overexpression of *sttC* and deletion of *dorA*	Not quantified	Jacobs et al., 2009
Glucose oxidase from *A. niger*	*T. reesei*	Regulating the UPR and ERAD by overexpression of *bip1* or *hac1*	1.5–1.8-fold increase	Wu et al., 2017
Glucose oxidase from *T. reesei*	*T. reesei*	Optimizing the intracellular transport process by overexpression of *snc1*	2.2-fold increase	Wu et al., 2017
Prochymosin from bovine	*A. niger*	Optimizing the intracellular transport process by deletion of *Aovip36* or *Aoemp47*, and fusing the target protein with α-amylase	Approximately 2-fold increase	Hoang et al., 2015
Cellulase from *T. reesei*	*T. reesei*	Constructing a protease-deficient strain by deletion of *res-1*, *cre-1*, *gh1-1,* and *alp-1*	5-fold increase	Liu et al., 2017
Laccase from *Trametes versicolor*	*A. niger*	Constructing a protease-deficient strain by deletion of *pepAa*, *pepAb,* or *pepAd*	1.21–1.42-fold increase	Wang et al., 2008
Glucoamylase from *A. niger*	*A. niger*	Regulating mycelium morphology by deletion of *racA*	4-fold increase	Fiedler et al., 2018
Cellulase from *N. crassa*	*N. crassa*	Regulating SREBP by deletion of *dsc-2*, *tul-1*, *sah-2*, *dsc-4*, *scp-1,* or *rbd-2*	Not quantified	Reilly et al., 2015; Qin et al., 2017

adopting the same strategy (Conesa et al., 2002). These effects could be attributed to the multifunction of BiP. BiP is able to promote protein translocation and folding, as well as to promote ER-associated protein degradation. Similarly, overexpression of *hac1* could promote protein secretion, which may also affect cell growth in certain strains (Valkonen et al., 2003; Carvalho et al., 2012). In addition, deletion of *derA* in *A. niger* can promote protein production (Carvalho et al., 2011), while deletion of the same gene affected the cell growth of *Aspergillus fumigatus* (Richie et al., 2011). It's not difficult to see that the effect of protein secretion by regulating UPR and ERAD is host-dependent. Thus, a deep understanding of the complexity and specificity of the interactions between the components of the secretory pathway in a particular host is required prior to the manipulation.

OPTIMIZATION OF THE INTRACELLULAR TRANSPORT PROCESS

Before being secreted outside, proteins are transported between ER and Golgi tendencies via vesicles. In this process, the ER-Golgi cargo receptor recruits the secreted proteins into the vesicles, thereby facilitating their transport (Dancourt and Barlowe, 2010). Optimization of the intracellular protein transport process allows enhanced protein secretion. In *A. oryzae*, the cargo receptor AoVip36 is localized in the ER and AoEmp47 is localized in the Golgi compartment. Deletion of AoVip36, responsible for anterograde transport, caused a 30% reduction of the endogenous α-amylase activity, and overexpression of this gene led to the increased secretion of EGFP (Hoang et al., 2015). In addition, deletion of *Aovip36* or *Aoemp47* increased the secretion of bovine prochymosin by approximately twofold (Hoang et al., 2015). In *Aspergillus nidulans*, gene *podB* is predicted to encode the subunit of the Golgi-conserved oligomeric complex (Gremillion et al., 2014), which is involved in Golgi retrograde vesicle transport, and affects cell polar growth, germination, and protein glycosylation (Wuestehube et al., 1996; Harris et al., 1999; Suvorova et al., 2002; Gremillion et al., 2014). A G-to-T mutation at nucleotide #751 in *podB1* led to significant increase in cellulase and xylanase activities (Boppidi et al., 2018). Wu et al. overexpressed *snc1* gene, which is involved in fusion of vesicles and plasma membrane, and observed a 2.2-fold increase in secretion of an *A. niger* glucose oxidase in *T. reesei* (Wu et al., 2017).

In addition to being successfully secreted outside, some heterologous proteins may be transported to the vacuole for degradation (Masai et al., 2003). Disruption of the vacuolar sorting receptor encoding gene *Aovps10* resulted in three and twofold increases in the production yields of bovine chymosin and human lysozyme in *A. oryzae*, respectively (Yoon et al., 2010).

CONSTRUCTION OF A PROTEASE-DEFICIENT STRAIN

The efficient production of certain endogenous protein in filamentous fungi disturbs the secretion of the protein of interest, and construction of a protease-deficient strain can strongly support the modification and secretion of the target protein. Disruption of alkaline serine protease SPW in *T. reesei* reduced the extracellular total protease activity by about 50%, and improved the production and stability of the heterologous alkaline endoglucanase EGV from *Humicola insolens* (Zhang et al., 2014). To construct a cellulase hyper-producing strain, β-glucosidase encoding gene *gh1-1*, alkaline protease encoding gene *alp-1*, and cellulase production related genes *cre-1* and *res-1* were simultaneously deleted in *Myceliophthora thermophile*. The secreted cellulase of the resulted strain was five times higher than that of the original strain (Liu et al., 2017).

REGULATION OF MYCELIUM MORPHOLOGY

Proteins are mainly secreted at vigorously growing mycelial tips in filamentous fungi (Wessels, 1993), and the mycelium morphology is especially important to protein secretion. The increased branching of the mycelium tip usually facilitates endogenous protein secretion. Lin et al. screened 90 morphological mutants of *Neurospora crassa* and found that disruption of *gul-1* led to a marked decrease in viscosity of the culture medium, while overexpression of *gul-1* led to a sharp increase in viscosity. In the *gul-1* disrupted strain, 25% and 56% increases were observed in the total extracellular protein concentration and β-glucosidase activity, respectively (Lin et al., 2018), suggesting that cell wall integrity has a significant effect on protein secretion. In *A. niger*, the Rho GTPase RacA regulates the polymerization and depolymerization of actin at the tip of mycelium (Kwon et al., 2011). When *racA* was deleted, the mycelial tip increased by about 20%, the number of secreted vesicles increased, and the secretion of glucoamylase increased 4 times as compared to the wild type strain (Fiedler et al., 2018). Similarly, deletion of *racA* resulted in a hyperbranched phenotype and three folds increase of cellulase activity in *T. reesei* (Fitz et al., 2019).

REGULATION OF SREBP

In filamentous fungi, SREBP, responsible for regulating sterol homeostasis under challenging environments, is strongly associated with protein secretion, including linkages to the UPR (Qin et al., 2017) and formation of hyphae branches (Willger et al., 2008). After analysis the phenotype of a 567 single-gene deletion collection of *N. crassa*, researchers found that deletion of *dsc-2* and *tul-1 (dsc-1)* significantly increased the secretion of proteins (Reilly et al., 2015). In *Schizosaccharomyces pombe* and *A. fumigatus*, homologs of Dsc-2 and Tul-1 are part of the Golgi E3 ligase complex (Dsc complex), which can activate SREBP orthologs Sre1 and SreA through proteolytic cleavage (Lloyd et al., 2013). In addition, deletion of the unit of Dsc complex Dsc-4 and the Sre1/SreA homolog SAH-2 also showed a high secretion phenotype of cellulases (Reilly et al., 2015). Homologs of SAH-2 and TUL-1 from *N. crassa* are discovered in *T. reesei*, and their deletions enhanced the capability of protein secretion (Reilly et al., 2015). In a follow-up study, deletion of gene *scp-1* and *rbd-2*, encoding SREBP cleavage activating protein and rhomboid protease respectively, also led to the high producing phenotype of cellulose (Qin et al., 2017).

CONCLUSION AND PERSPECTIVES

Owing to the powerful protein secretion pathway, filamentous fungi are attractive cell factories for protein expression and secretion. For all the discussed strategies, replacing original signal peptide with a more efficient one, regulation of UPR and ERAD, optimization of the intracellular transport process, and construction of a protease-deficient strain have been successfully applied to improve the production of endogenous and heterologous proteins by filamentous fungi, while fusion of heterologous protein to a naturally secreted one is extremely effective for production of heterologous protein (**Table 1**). As for regulation of mycelium morphology and optimization of SREBP, although they were mainly adopted for production of endogenous protein, we believe that they are also applicable for production of heterologous protein. However, most efforts in genetic engineering of filamentous fungi for enhanced protein expression and secretion were solely based on the protein of interest, the secretory pathway or the host. Although these engineering strategies significantly improved target protein production, they were mainly related to single gene or pathway.

With the aid of multiple gene editing technologies (e.g., DNA recombination, RNAi, CRISPR-Cas), genome engineering strategies introduce deletion, insertion and/or point mutations across the genome via a trackable manner to accelerate strain evolution (Si et al., 2015). Compared with traditional metabolic engineering strategies, genome engineering allows rapid tracking and discovery of novel determinants (Xiao and Zhao, 2014; Si et al., 2017), editing of key determinant with single-nucleotide precision (Garst et al., 2017; Bao et al., 2018), or simultaneous manipulating multiple pathways (Barbieri et al., 2017; Liang et al., 2017). Apart from the unicellular model organisms (e.g., *Saccharomyces cerevisiae*), many filamentous fungi, particularly the mushroom-forming fungi, contain two different nuclei with different genetic contents (Gehrmann et al., 2018). In addition to the heterogeneity, many important medicinal mushrooms also exhibit low efficiency on gene transformation and homologous recombination (HR), which pose a great challenge to establish gene editing tools for genome engineering (Wang et al., 2020). To circumvent these difficulties, developing effective technologies for single spore isolation, gene delivery and/or improving HR efficiency are highly required in these filamentous fungi. It is notable that the target performances of the engineered strains, which are greatly improved by the aforementioned genome

engineering strategies, can usually be screened out via cell growth or color. Thus, high-throughput screening methods are highly required to ensure the success of genome engineering. In fact, the fluorescence-activated cell sorting (FACS) assisted the intracellular protein production has been extensively adopted in filamentous fungi, but such strategy is difficult to screen out the beneficial mutants with enhanced protein secretion capacity (Throndset et al., 2010). To solve this problem, displaying the fluorescence protein on the cell surface, coupled by FACS, allows screening of the cellulose hypersecretors from *T. reesei* (Gao et al., 2018). As a promising alternative, the droplet-based microfluidic high-throughput screening platform has been established in *T. reesei* and *A. niger* (Beneyton et al., 2016; He et al., 2019). In future, we believe that harnessing the great potential of genome engineering will further increase protein expression and secretion by filamentous fungi.

AUTHOR CONTRIBUTIONS

CZ and HX designed this manuscript. QW and HX wrote this manuscript. QW, CZ, and HX revised this manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 31971344) and Municipal Natural Science Foundation of Shanghai (Nos. 17ZR1448900 and 18ZR1420300).

REFERENCES

Alcocer, M. J. C., Furniss, C. S. M., Kroon, P. A., Campbell, M., and Archer, D. B. (2003). Comparison of modular and non-modular xylanases as carrier proteins for the efficient secretion of heterologous proteins from *Penicillium funiculosum*. *Appl. Microbiol. Biotechnol.* 60, 726–732. doi: 10.1007/s00253-002-1184-4

Bao, Z., HamediRad, M., Xue, P., Xiao, H., Tasan, I., Chao, R., et al. (2018). Genome-scale engineering of *Saccharomyces cerevisiae* with single-nucleotide precision. *Nat. Biotechnol.* 36, 505–508. doi: 10.1038/nbt.4132

Barbieri, E. M., Muir, P., Akhuetie-Oni, B. O., Yellman, C. M., and Isaacs, F. J. (2017). Precise editing at DNA replication forks enables multiplex genome engineering in eukaryotes. *Cell* 171, 1453–1467.e1413. doi: 10.1016/j.cell.2017.10.034

Beneyton, T., Wijaya, I. P. M., Postros, P., Najah, M., Leblond, P., Couvent, A., et al. (2016). High-throughput screening of filamentous fungi using nanoliter-range droplet-based microfluidics. *Sci. Rep.* 6:27223. doi: 10.1038/srep27223

Bernasconi, R., and Molinari, M. (2011). ERAD and ERAD tuning: disposal of cargo and of ERAD regulators from the mammalian ER. *Curr. Opin. Cell Biol.* 23, 176–183. doi: 10.1016/j.ceb.2010.10.002

Boppidi, K. R., Ribeiro, L. F. C., Iambamrung, S., Nelson, S. M., Wang, Y., Momany, M., et al. (2018). Altered secretion patterns and cell wall organization caused by loss of PodB function in the filamentous fungus *Aspergillus nidulans*. *Sci. Rep.* 8:11433. doi: 10.1038/s41598-018-29615-z

Carvalho, N. D. S. P., Arentshorst, M., Kooistra, R., Stam, H., Sagt, C. M., van den Hondel, C. A., et al. (2011). Effects of a defective ERAD pathway on growth and heterologous protein production in *Aspergillus niger*. *Appl. Microbiol. Biotechnol.* 89, 357–373. doi: 10.1007/s00253-010-2916-5

Carvalho, N. D. S. P., Jørgensen, T. R., Arentshorst, M., Nitsche, B. M., van den Hondel, C. A., Archer, D. B., et al. (2012). Genome-wide expression analysis upon constitutive activation of the HacA bZIP transcription factor in *Aspergillus niger* reveals a coordinated cellular response to counteract ER stress. *BMC Genomics* 13:350. doi: 10.1186/1471-2164-13-350

Cavka, A., and Jönsson, L. J. (2014). Comparison of the growth of filamentous fungi and yeasts in lignocellulose-derived media. *Biocatal. Agric. Biotechnol.* 3, 197–204. doi: 10.1016/j.bcab.2014.04.003

Conesa, A., Jeenes, D., Archer, D. B., van den Hondel, C. A. M., and Punt, P. J. (2002). Calnexin overexpression increases manganese peroxidase production in *Aspergillus niger*. *Appl. Environ. Microbiol.* 68, 846–851. doi: 10.1128/AEM.68.2.846-851.2002

Conesa, A., Punt, P. J., van Luijk, N., and van den Hondel, C. A. (2001). The secretion pathway in filamentous fungi: a biotechnological view. *Fungal Genet. Biol.* 33, 155–171. doi: 10.1006/fgbi.2001.1276

Dancourt, J., and Barlowe, C. (2010). Protein sorting receptors in the early secretory pathway. *Annu. Rev. Biochem.* 79, 777–802. doi: 10.1146/annurev-biochem-061608-091319

Deshpande, N., Wilkins, M. R., Packer, N., and Nevalainen, H. (2008). Protein glycosylation pathways in filamentous fungi. *Glycobiology* 18, 626–637. doi: 10.1093/glycob/cwn044

Fiedler, M. R. M., Barthel, L., Kubisch, C., Nai, C., and Meyer, V. (2018). Construction of an improved *Aspergillus niger* platform for enhanced glucoamylase secretion. *Microb. Cell Fact.* 17, 95–95. doi: 10.1186/s12934-018-0941-8

Fitz, E., Gamauf, C., Seiboth, B., and Wanka, F. (2019). Deletion of the small GTPase *rac1* in *Trichoderma reesei* provokes hyperbranching and impacts growth and cellulase production. *Fungal Biol. Biotechnol.* 6:16. doi: 10.1186/s40694-019-0078-5

Gao, F., Hao, Z., Sun, X., Qin, L., Zhao, T., Liu, W., et al. (2018). A versatile system for fast screening and isolation of *Trichoderma reesei* cellulase hyperproducers based on DsRed and fluorescence-assisted cell sorting. *Biotechnol. Biofuels* 11:261. doi: 10.1186/s13068-018-1264-z

Garst, A. D., Bassalo, M. C., Pines, G., Lynch, S. A., Halweg-Edwards, A. L., Liu, R., et al. (2017). Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nat. Biotechnol.* 35, 48–55. doi: 10.1038/nbt.3718

Gehrmann, T., Pelkmans, J. F., Ohm, R. A., Vos, A. M., Sonnenberg, A. S. M., Baars, J. J. P., et al. (2018). Nucleus-specific expression in the multinuclear mushroom-forming fungus *Agaricus bisporus* reveals different nuclear regulatory programs. *Proc. Natl. Acad. Sci. U.S.A.* 115, 4429–4434. doi: 10.1073/pnas.1721381115

Gremillion, S. K., Harris, S. D., Jackson-Hayes, L., Kaminskyj, S. G. W., Loprete, D. M., Gauthier, A. C., et al. (2014). Mutations in proteins of the conserved oligomeric Golgi complex affect polarity, cell wall structure, and glycosylation in the filamentous fungus *Aspergillus nidulans*. *Fungal Genet. Biol.* 73, 69–82. doi: 10.1016/j.fgb.2014.10.005

Halic, M., Blau, M., Becker, T., Mielke, T., Pool, M. R., Wild, K., et al. (2006). Following the signal sequence from ribosomal tunnel exit to signal recognition particle. *Nature* 444, 507–511. doi: 10.1038/nature05326

Harris, S. D., Hofmann, A. F., Tedford, H. W., and Lee, M. P. (1999). Identification and characterization of genes required for hyphal morphogenesis in the filamentous fungus *Aspergillus nidulans*. *Genetics* 151, 1015–1025.

Haßdenteufel, S., Johnson, N., Paton, A. W., Paton, J. C., High, S., and Zimmermann, R. (2018). Chaperone-mediated Sec61 channel gating during ER import of small precursor proteins overcomes Sec61 inhibitor-reinforced energy barrier. *Cell Rep.* 23, 1373–1386. doi: 10.1016/j.celrep.2018.03.122

He, R., Ding, R., Heyman, J. A., Zhang, D., and Tu, R. (2019). Ultra-high-throughput picoliter-droplet microfluidics screening of the industrial cellulase-producing filamentous fungus *Trichoderma reesei*. *J. Ind. Microbiol. Biotechnol.* 46, 1603–1610. doi: 10.1007/s10295-019-02221-2

Hoang, H. D., Maruyama, J. I., and Kitamoto, K. (2015). Modulating endoplasmic reticulum-Golgi cargo receptors for improving secretion of carrier-fused heterologous proteins in the filamentous fungus *Aspergillus oryzae*. *Appl. Environ. Microbiol.* 81, 533–543. doi: 10.1128/AEM.02133-14

Hutagalung, A. H., and Novick, P. J. (2011). Role of Rab GTPases in membrane traffic and cell physiology. *Physiol. Rev.* 91, 119–149. doi: 10.1152/physrev.00059.2009

Jacobs, D. I., Olsthoorn, M. M. A., Maillet, I., Akeroyd, M., Breestraat, S., Donkers, S., et al. (2009). Effective lead selection for improved protein production in *Aspergillus niger* based on integrated genomics. *Fungal Genet. Biol.* 46(1, Suppl.), S141–S152. doi: 10.1016/j.fgb.2008.08.012

Jeenes, D. J., Mackenzie, D. A., Roberts, I. N., and Archer, D. B. (1991). Heterologous protein production by filamentous fungi. *Biotechnol. Genet. Eng. Rev.* 9, 327–367. doi: 10.1080/02648725.1991.10750006

Kario, E., Amar, N., Elazar, Z., and Navon, A. (2011). A new autophagy-related checkpoint in the degradation of an ERAD-M target. *J. Biol. Chem.* 286, 11479–11491. doi: 10.1074/jbc.M110.177618

Karnaukhova, E., Ophir, Y., Trinh, L., Dalal, N., Punt, P. J., Golding, B., et al. (2007). Expression of human α1-proteinase inhibitor in *Aspergillus niger*. *Microb. Cell Fact.* 6:34. doi: 10.1186/1475-2859-6-34

Kraševec, N., Milunović, T., Lasnik, M. A., Lukančič, I., Komel, R., and Porekar, V. G. (2014). Human granulocyte colony stimulating factor (G-CSF) produced in the filamentous fungus *Aspergillus niger*. *Acta Chim. Slov.* 61, 709–717.

Kwon, M. J., Arentshorst, M., Roos, E. D., van den Hondel, C. A., Meyer, V., and Ram, A. F. J. (2011). Functional characterization of Rho GTPases in *Aspergillus niger* uncovers conserved and diverged roles of Rho proteins within filamentous fungi. *Mol. Microbiol.* 79, 1151–1167. doi: 10.1111/j.1365-2958.2010.07524.x

Liang, L., Liu, R., Garst, A. D., Lee, T., Nogué, V. S. I., Beckham, G. T., et al. (2017). CRISPR EnAbled trackable genome engineering for isopropanol production in *Escherichia coli*. *Metab. Eng.* 41, 1–10. doi: 10.1016/j.ymben.2017.02.009

Lin, L., Sun, Z., Li, J., Chen, Y., Liu, Q., Sun, W., et al. (2018). Disruption of *gul-1* decreased the culture viscosity and improved protein secretion in the filamentous fungus *Neurospora crassa*. *Microb. Cell Fact.* 17:96. doi: 10.1186/s12934-018-0944-5

Liu, Q., Gao, R., Li, J., Lin, L., Zhao, J., Sun, W., et al. (2017). Development of a genome-editing CRISPR/Cas9 system in thermophilic fungal *Myceliophthora* species and its application to hyper-cellulase production strain engineering. *Biotechnol. Biofuels* 10:1. doi: 10.1186/s13068-016-0693-9

Lloyd, S. J.-A., Raychaudhuri, S., and Espenshade, P. J. (2013). Subunit architecture of the Golgi Dsc E3 ligase required for sterol regulatory element-binding protein (SREBP) cleavage in fission Yeast. *J. Biol. Chem.* 288, 21043–21054. doi: 10.1074/jbc.M113.468215

Lombraña, M., Moralejo, F. J., Pinto, R., and Martín, J. F. (2004). Modulation of *Aspergillus awamori* thaumatin secretion by modification of bipA gene expression. *Appl. Environ. Microbiol.* 70, 5145–5152. doi: 10.1128/AEM.70.9.5145-5152.2004

Masai, K., Maruyama, J.-I., Nakajima, H., and Kitamoto, K. (2003). *In vivo* visualization of the distribution of a secretory protein in *Aspergillus oryzae* hyphae using the RntA-EGFP fusion protein. *Biosci. Biotechnol. Biochem.* 67, 455–459. doi: 10.1271/bbb.67.455

Mitra, N., Sinha, S., Ramya, T. N. C., and Surolia, A. (2006). N-linked oligosaccharides as outfitters for glycoprotein folding, form and function. *Trends Biochem. Sci.* 31, 156–163. doi: 10.1016/j.tibs.2006.01.003

Nevalainen, K. M. H., Te'o, V. S. J., and Bergquist, P. L. (2005). Heterologous protein expression in filamentous fungi. *Trends Biotechnol.* 23, 468–474. doi: 10.1016/j.tibtech.2005.06.002

Ohno, A., Maruyama, J.-I., Nemoto, T., Arioka, M., and Kitamoto, K. (2011). A carrier fusion significantly induces unfolded protein response in heterologous protein production by *Aspergillus oryzae*. *Appl. Microbiol. Biotechnol.* 92, 1197–1206. doi: 10.1007/s00253-011-3487-9

Pakula, T. M., Laxell, M., Huuskonen, A., Uusitalo, J., Saloheimo, M., and Penttilä, M. (2003). The effects of drugs inhibiting protein secretion in the filamentous fungus *Trichoderma reesei*: evidence for down-regulation of genes that encode secreted proteins in the stressed cells. *J. Biol. Chem.* 278, 45011–45020. doi: 10.1074/jbc.m302372200

Punt, P. J., van Biezen, N., Conesa, A., Albers, A., Mangnus, J., and van den Hondel, C. (2002). Filamentous fungi as cell factories for heterologous protein production. *Trends Biotechnol.* 20, 200–206. doi: 10.1016/S0167-7799(02)01933-9

Qin, L., Wu, V. W., and Glass, N. L. (2017). Deciphering the regulatory network between the SREBP pathway and protein secretion in *Neurospora crassa*. *mBio* 8:e00233-17. doi: 10.1128/mBio.00233-17

Reilly, M. C., Qin, L., Craig, J. P., Starr, T. L., and Glass, N. L. (2015). Deletion of homologs of the SREBP pathway results in hyper-production of cellulases in *Neurospora crassa* and *Trichoderma reesei*. *Biotechnol. Biofuels* 8:121. doi: 10.1186/s13068-015-0297-9

Richie, D. L., Feng, X., Hartl, L., Aimanianda, V., Krishnan, K., Powers-Fletcher, M. V., et al. (2011). The virulence of the opportunistic fungal pathogen *Aspergillus fumigatus* requires cooperation between the endoplasmic reticulum-associated degradation pathway (ERAD) and the unfolded protein response (UPR). *Virulence* 2, 12–21. doi: 10.4161/viru.2.1.13345

Saloheimo, M., and Pakula, T. M. (2012). The cargo and the transport system: secreted proteins and protein secretion in *Trichoderma reesei* (*Hypocrea jecorina*). *Microbiology* 158(Pt 1), 46–57. doi: 10.1099/mic.0.053132-0

Saunders, G., Picknett, T. M., Tuite, M. F., and Ward, M. (1989). Heterologous gene expression in filamentous fungi. *Trends Biotechnol.* 7, 283–287. doi: 10.1016/0167-7799(89)90048-6

Si, T., Chao, R., Min, Y., Wu, Y., Ren, W., and Zhao, H. (2017). Automated multiplex genome-scale engineering in yeast. *Nat. Commun.* 8:15187. doi: 10.1038/ncomms15187

Si, T., Xiao, H., and Zhao, H. (2015). Rapid prototyping of microbial cell factories via genome-scale engineering. *Biotechnol. Adv.* 33, 1420–1432. doi: 10.1016/j.biotechadv.2014.11.007

Söllner, T., Whiteheart, S. W., Brunner, M., Erdjument-Bromage, H., Geromanos, S., Tempst, P., et al. (1993). SNAP receptors implicated in vesicle targeting and fusion. *Nature* 362, 318–324. doi: 10.1038/362318a0

Spang, A. (2008). Membrane traffic in the secretory pathway. *Cell. Mol. Life Sci.* 65, 2781–2789. doi: 10.1007/s00018-008-8349-y

Su, X., Schmitz, G., Zhang, M., Mackie, R. I., and Cann, I. K. O. (2012). "Chapter One - heterologous gene expression in filamentous fungi," in *Advances in Applied Microbiology*, eds G. M. Gadd and S. Sariaslani (Cambridge, MA: Academic Press), 1–61.

Suvorova, E. S., Duden, R., and Lupashin, V. V. (2002). The Sec34/Sec35p complex, a Ypt1p effector required for retrograde intra-Golgi trafficking, interacts with Golgi SNAREs and COPI vesicle coat proteins. *J. Cell Biol.* 157, 631–643. doi: 10.1083/jcb.200111081

Throndset, W., Kim, S., Bower, B., Lantz, S., Kelemen, B., Pepsin, M., et al. (2010). Flow cytometric sorting of the filamentous fungus *Trichoderma reesei* for improved strains. *Enzyme Microb. Technol.* 47, 335–341. doi: 10.1016/j.enzmictec.2010.09.003

Valkonen, M., Kalkman, E. R., Saloheimo, M., Penttilä, M., Read, N. D., and Duncan, R. R. (2007). Spatially segregated SNARE protein interactions in living fungal cells. *J. Biol. Chem.* 282, 22775–22785. doi: 10.1074/jbc.m700916200

Valkonen, M., Ward, M., Wang, H., Penttilä, M., and Saloheimo, M. (2003). Improvement of foreign-protein production in *Aspergillus niger* var. *awamori* by constitutive induction of the unfolded-protein response. *Appl. Environ. Microbiol.* 69, 6979–6986. doi: 10.1128/AEM.69.12.6979-6986.2003

Virag, A., and Harris, S. D. (2006). The Spitzenkörper: a molecular perspective. *Mycol. Res.* 110, 4–13. doi: 10.1016/j.mycres.2005.09.005

Wang, G., Zhang, D., and Chen, S. (2014). Effect of earlier unfolded protein response and efficient protein disposal system on cellulase production in Rut C30. *World J. Microbiol. Biotechnol.* 30, 2587–2595. doi: 10.1007/s11274-014-1682-4

Wang, L., Zhao, S., Chen, X., Deng, Q., Li, C., and Feng, J. (2018). Secretory overproduction of a raw starch-degrading glucoamylase in *Penicillium oxalicum* using strong promoter and signal peptide. *Appl. Microbiol. Biotechnol.* 102, 9291–9301. doi: 10.1007/s00253-018-9307-8

Wang, P., Xiao, H., and Zhong, J. (2020). CRISPR-Cas9 assisted functional gene editing in the mushroom *Ganoderma lucidum*. *Appl. Microbiol. Biotechnol.* 104, 1661–1671. doi: 10.1007/s00253-019-10298-z

Wang, Y., Xue, W., Sims, A. H., Zhao, C., Wang, A., Tang, G., et al. (2008). Isolation of four pepsin-like protease genes from *Aspergillus niger* and analysis of the effect of disruptions on heterologous laccase expression. *Fungal Genet. Biol.* 45, 17–27. doi: 10.1016/j.fgb.2007.09.012

Ward, O. P. (2012). Production of recombinant proteins by filamentous fungi. *Biotechnol. Adv.* 30, 1119–1139. doi: 10.1016/j.biotechadv.2011.09.012

Wessels, J. G. H. (1993). Tansley Review No. 45 Wall growth, protein excretion and morphogenesis in fungi. *New Phytol.* 123, 397–413. doi: 10.1111/j.1469-8137.1993.tb03751.x

Willger, S. D., Puttikamonkul, S., Kim, K. H., Burritt, J. B., Grahl, N., Metzler, L. J., et al. (2008). A sterol-regulatory element binding protein is required for cell polarity, hypoxia adaptation, azole drug resistance, and virulence in *Aspergillus fumigatus*. *PLoS Pathog.* 4:e1000200. doi: 10.1371/journal.ppat.1000200

Wu, Y., Sun, X., Xue, X., Luo, H., Yao, B., Xie, X., et al. (2017). Overexpressing key component genes of the secretion pathway for enhanced secretion of an *Aspergillus niger* glucose oxidase in *Trichoderma reesei*. *Enzyme Microb. Technol.* 106, 83–87. doi: 10.1016/j.enzmictec.2017.07.007

Wuestehube, L. J., Duden, R., Eun, A., Hamamoto, S., Korn, P., Ram, R., et al. (1996). New mutants of *Saccharomyces cerevisiae* affected in the transport of proteins from the endoplasmic reticulum to the Golgi complex. *Genetics* 142, 393–406.

Xiao, H., and Zhao, H. (2014). Genome-wide RNAi screen reveals the E3 SUMO-protein ligase gene *SIZ1* as a novel determinant of furfural tolerance in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* 7:78. doi: 10.1186/1754-6834-7-78

Xu, Y., Wang, Y., Liu, T., Zhang, H., Zhang, H., and Li, J. (2018). The GlaA signal peptide substantially increases the expression and secretion of α-galactosidase in *Aspergillus niger*. *Biotechnol. Lett.* 40, 949–955. doi: 10.1007/s10529-018-2540-5

Yan, Q., and Lennarz, W. J. (2002). Studies on the function of oligosaccharyl transferase subunits: Stt3p is directly involved in the glycosylation process. *J. Biol. Chem.* 277, 47692–47700. doi: 10.1074/jbc.m208136200

Yoon, J., Aishan, T., Maruyama, J., and Kitamoto, K. (2010). Enhanced production and secretion of heterologous proteins by the filamentous fungus *Aspergillus oryzae* via disruption of vacuolar protein sorting receptor gene *Aovps10*. *Appl. Environ. Microbiol.* 76, 5718–5727. doi: 10.1128/AEM.03087-09

Yoon, J., Kikuma, T., Maruyama, J., and Kitamoto, K. (2013). Enhanced production of bovine chymosin by autophagy deficiency in the filamentous fungus *Aspergillus oryzae*. *PLoS One* 8:e62512. doi: 10.1371/journal.pone.0062512

Zhang, G., Zhu, Y., Wei, D., and Wang, W. (2014). Enhanced production of heterologous proteins by the filamentous fungus *Trichoderma reesei* via disruption of the alkaline serine protease SPW combined with a pH control strategy. *Plasmid* 71, 16–22. doi: 10.1016/j.plasmid.2014.01.001

Zhong, Y., Liu, X., Xiao, P., Wei, S., and Wang, T. (2011). Expression and secretion of the human erythropoietin using an optimized *cbh1* promoter and the native CBH I signal sequence in the industrial fungus *Trichoderma reesei*. *Appl. Biochem. Biotechnol.* 165, 1169–1177. doi: 10.1007/s12010-011-9334-8

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

MINI REVIEW
published: 16 April 2020
doi: 10.3389/fbioe.2020.00282

Synthetic Biology Tools for Genome and Transcriptome Engineering of Solventogenic *Clostridium*

Seong Woo Kwon[1†], Kuppusamy Alagesan Paari[2†], Alok Malaviya[3] and Yu-Sin Jang[1]**

[1] Department of Agricultural Chemistry and Food Science Technology, Division of Applied Life Science (BK21 Plus Program), Institute of Agriculture & Life Science (IALS), Gyeongsang National University, Jinju, South Korea, [2] Department of Life Sciences, CHRIST (Deemed to be University), Bengaluru, India, [3] Applied and Industrial Biotechnology Laboratory (AIBL), Department of Life Sciences, CHRIST (Deemed to be University), Bengaluru, India

Strains of *Clostridium* genus are used for production of various value-added products including fuels and chemicals. Development of any commercially viable production process requires a combination of both strain and fermentation process development strategies. The strain development in *Clostridium* sp. could be achieved by random mutagenesis, and targeted gene alteration methods. However, strain improvement in *Clostridium* sp. by targeted gene alteration method was challenging due to the lack of efficient tools for genome and transcriptome engineering in this organism. Recently, various synthetic biology tools have been developed to facilitate the strain engineering of solventogenic *Clostridium*. In this review, we consolidated the recent advancements in toolbox development for genome and transcriptome engineering in solventogenic *Clostridium*. Here we reviewed the genome-engineering tools employing mobile group II intron, *pyrE* alleles exchange, and CRISPR/Cas9 with their application for strain development of *Clostridium* sp. Next, transcriptome engineering tools such as untranslated region (UTR) engineering and synthetic sRNA techniques were also discussed in context of *Clostridium* strain engineering. Application of any of these discussed techniques will facilitate the metabolic engineering of clostridia for development of improved strains with respect to requisite functional attributes. This might lead to the development of an economically viable butanol production process with improved titer, yield and productivity.

Keywords: *Clostridium*, synthetic biology, mobile intron, CRISPR, Cas, synthetic sRNA, UTR

OPEN ACCESS

Edited by:
Yi Wang,
Auburn University, United States

Reviewed by:
Yang Gu,
Shanghai Institutes for Biological Sciences (CAS), China
Zhiqiang Wen,
Nanjing University of Science and Technology, China

****Correspondence:***
Alok Malaviya
alokkumar.malaviya@christuniversity.in
Yu-Sin Jang
jangys@gnu.ac.kr

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *19 January 2020*
Accepted: *17 March 2020*
Published: *16 April 2020*

Citation:
Kwon SW, Paari KA, Malaviya A and Jang Y-S (2020) Synthetic Biology Tools for Genome and Transcriptome Engineering of Solventogenic Clostridium. Front. Bioeng. Biotechnol. 8:282. doi: 10.3389/fbioe.2020.00282

INTRODUCTION

Strain improvement for production of fuels or any biobased industrial product could be achieved by employing any of the following two strategies: (i) heterologous expression of metabolic pathway genes in a non-native producers, and (ii) improvement of native producers (Arora et al., 2019; Banerjee et al., 2019; Choi et al., 2019). However, achieving titer values in heterologous host matching to those being produced by native organisms, it requires a significant effort with high chances of failure. Therefore, the strategy of improving native strains with necessary genes of the desired pathway and cofactor regeneration capability is preferred (Park et al., 2018; Rhie et al., 2019).

However, this strategy of strain improvement in *Clostridium* sp. has been limited by the availability of appropriate genome engineering tools.

Clostridium genus comprises many industrially important strains for biorefinery applications such as cellulosic and hemicellulosic biomass degradation, carbon fixation, advanced biofuel and platform chemical production and as anti-cancer therapeutics (Jang et al., 2012; Malaviya et al., 2012; Liu J. et al., 2015; Jones et al., 2016; Staedtke et al., 2016; Noh et al., 2018; Woo et al., 2018; Xin et al., 2018; Strecker et al., 2019a). The full potential of *Clostridium* genus for biorefinery applications could only be realized by advancement in the synthetic biology toolkits for strain improvement. During the last decade, tremendous progresses have been made in the development of genome engineering toolkit for strain engineering of *Clostridium* species. Development of genetic tools in *Clostridium* have been well reviewed by various research groups (Pyne et al., 2014; Liu Y. J. et al., 2015; Minton et al., 2016; Moon et al., 2016; Joseph et al., 2018; Kuehne et al., 2019; McAllister and Sorg, 2019; Wen et al., 2019b,c). Most of these reports are focused on couple of tools with an explanation in depth.

In this work, we have reviewed overall recent toolbox for genome and transcriptome engineering in solventogenic *Clostridium*, which could be used to develop improved clostridia strains, for production of sustainable and commercially viable industrial scale products. Brief features of the synthetic toolbox are summarized in **Table 1**. Consolidated information in this review dealing with strain improvement tools for *Clostridium* will aid the scientific and industrial sector to select the appropriate tools for strain improvement.

MOBILE GROUP II INTRON BASED GENE-KNOCKOUT

Mobile group II intron technology is also known as "ClosTron" when applied in context of *Clostridium* genus. In this method a gene is disrupted by inserting the mobile intron into a target locus in the chromosome by a process termed as retrohoming, making this technology a convenient, efficient and specific method of gene disruption (Heap et al., 2007, 2010; Shao et al., 2007; Jang et al., 2012, 2014; Mohr et al., 2013; Liu Y. J. et al., 2015). Among various mobile group II introns, *Ll.LtrB* and *TeI3c/4c* have been extensively used for gene knockout in the solventogenic *Clostridium*. *Ll.LtrB* intron includes intron RNA domain and open reading frame (ORF) domain. Intron RNA domain contains splicing sites consisting of exon binding sites (EBS) 1, EBS 2, and δ (**Figure 1A**). The ORF domain contains genes encoding reverse transcriptase (RTase), maturase, and endonuclease (**Figure 1A**). *TeI3c/4c* intron has been employed to develop genome engineering tool for thermophilic *Clostridium thermocellum*, since the intron could be melted down at high temperatures (Mohr et al., 2013).

Moreover, *Ll.LtrB* intron has further been modified to include a retrotransposition-activated selection marker (RAM) (Zhong et al., 2003). RAM consists of a selection marker and is inserted into the intron. A group I intron is inserted into the marker to inactivate the marker itself. Inserted group I intron is self catalytically spliced out of mRNA in an orientation dependent manner, so that a functional marker gene can only be expressed after successful chromosomal insertion occurs (Joseph et al., 2018).

At the first stage of the clostridia gene knockout using *Ll.LtrB* intron, single gene knockouts mutant, such as *spo0A, pta, ack, ptb, buk, hbd, hydA* and *argA* variants have been constructed across the *Clostridium* genus, including *C. acetobutylicum*, *C. beijerinckii*, *C. botulinum*, and *C. difficile* (Heap et al., 2010; Dingle et al., 2011; Jang et al., 2012; Baban et al., 2013; Honicke et al., 2014; Lawson and Rainey, 2016; Liu et al., 2016). In 2012, a new method for second gene deletion was reported which could overcome the necessity of removing the plasmid used for the first gene deletion and resulted in the construction of various *C. acetobutylicum* strains, including *pta*/*buk*, *pta*/*ctfB*, *ptb*/*buk*, and triple mutant *pta*/*buk*/*ctfB* strains (Jang et al., 2012). In this technique, two genes encoding the erythromycin and chloramphenicol resistance enzymes were used as mutant selection marker and the concept of plasmid incompatibility was employed (Jang et al., 2012). In 2014, the same group reported the fourth and fifth gene deletion process for the construction of mutants *pta*/*buk*/*ctfB*/*adhE1* and *pta*/*buk*/*ctfB*/*adhE1*/*hydA* of *C. acetobutylicum* (Jang et al., 2014).

Curing and off-target manipulation remained one of the major limitations of mobile group II intron technology (Wen et al., 2019c). Curing efficiency of the plasmid containing mobile intron was enhanced by cloning *pyrF* (orotidine 5-phosphate decarboxylase) to ClosTron plasmid. The *pyrF* encodes essential enzyme of pyrimidine biosynthesis which can use 5-fluoroorotic acid (FOA) as a substrate and converts it to toxic compound and is widely used as counter selection marker (Sato et al., 2005; Tripathi et al., 2010; Heap et al., 2012). Once FOA gets converted to toxic compound by *pyrF* in the ClosTron plasmid, only cured strain could survive in the FOA added media. The cured strain can be rapidly selected by *pyrF*-based screening system, even on one plate (Cui et al., 2014).

Another problem with ClosTron is that it accidently affects and manipulates the off-target genome and cause unexpected genotypes and phenotypes (Heap et al., 2012). To overcome this, a highly regulated ClosTron system has been developed by inducing L-arabinose inducer (ARAi) to reduce off-target possibility (Zhang J. et al., 2015). To verify the impact of inducible ClosTron using ARAi system, pSY6-mspI (Cui et al., 2012) and pGZ-pyrF-cipC (Cui et al., 2014) were modified by introducing ARAi system in *C. cellulolyticum* H10 Δ*pyrF* strain. Surprisingly, it was found that the off-target manipulation frequency was decreased to 0 by inducible ClosTron ARAi system (Zhang J. et al., 2015).

GENOME EDITING USING *PYRE* ALLELES

Recently, allele coupled exchange (ACE) method has been developed which facilitates the insertion of complex heterologous DNA of varying size into the host genome (Ng et al., 2013;

TABLE 1 | Summary of synthetic biology tools and strategies applied for genome and transcriptome engineering of solventogenic *Clostridium*.

Categories	Tools and strategy	Brief description	Selection guide	References
Genome engineering	Mobile group II Intron	• Site-directed disruptions based on retrohoming of mobile group II introns • Insertion of intron into target site • Plasmid based • Ribonucleoprotein complex formation • Retrotransposition-activated selection marker (RAM) to help in selection	• Knockout • Knockdown	Chen et al., 2005, 2007; Shao et al., 2007; Heap et al., 2007; Baban et al., 2013; Jang et al., 2012, 2014; Pyne et al., 2014; Liu Y. J. et al., 2015; Liu et al., 2016; Xu et al., 2015; Meaney et al., 2015, 2016; Lawson et al., 2016
	pyrE allele exchange	• Works on the principle of deactivating an easily screenable gene (*pyrE*) • Complementing the mutant strain with a heterologous version of *pyrE* gene as a counter selective marker	• Knockout • Insertion • Exchange	Tripathi et al., 2010; Heap et al., 2012; Ng et al., 2013; Bankar et al., 2015; Zhang N. et al., 2015; Croux et al., 2016; Ehsaan et al., 2016a;
	CRISPR/Cas	• RNA-guided target specific DNA cleavage system • Originated from bacterial adaptive immune system • Needs single guide RNA (sgRNA), Cas endonuclease, and homologous arms for recombination	• Knockout • Knockdown	Xu et al., 2015, 2017; Nagaraju et al., 2016; Bruder et al., 2016; Li et al., 2016; Pyne et al., 2016; Wang Y. et al., 2017, Wang et al., 2018
	Phage serine integrase-mediated genome engineering	• Use two heterologous phage attachment/integration systems • Dual Integrase Cassette Exchange (DICE) strategy • Needs CRISPR/Cas9 assistance	• Knockout • Insertion	Huang et al., 2019
Transcriptome engineering	Synthetic regulatory RNA (sRNA)	• Knockdown tool based on synthetically designed sRNA • Complementarily binds to target mRNAs and block translation	• Knockdown • Overexpression (by repressor knockdown)	Cho and Lee, 2017
	Untranslated regions (UTR) engineering	• UTR modulation • Better mRNA stability by addition of small stem loop structure in the 5'-UTR	• Knockdown • Overexpression (by repressor knockdown)	Lee et al., 2016
	CRISPRi	• Knockdown tool using catalytically inactivated effector dCas9 proteins	• Knockdown • Overexpression (by repressor knockdown)	Bruder et al., 2016; Li et al., 2016; Wang et al., 2016b; Wen et al., 2017; Woolston et al., 2018; Muh et al., 2019

Zhang N. et al., 2015; Ehsaan et al., 2016a; Minton et al., 2016). In ACE, a counter selection marker is coupled to a desired double crossover event (**Figure 1B**). The counter selection marker entitles the isolation of double cross over through homologous recombination. The *pyrE* and *codA* genes are the most frequently used selectable marker in ACE Technology. The gene *codA* encodes for the enzyme cytosine deaminase while, *pyrE* encodes orotate phosphoribosyl transferase, which is a key enzyme required in the *de novo* pathway for pyrimidine biosynthesis.

In clostridia genome editing, *pyrE* allele has been primarily employed. Mutant and wild type *pyrE* allele confers resistance and sensitivity to FOA, respectively. The advantages of *pyrE* allele based recombination includes: (i) rapid insertion of heterologous DNA, (ii) double crossovers which forms the stable integration, (iii) allows large insert size, and (iv) has higher efficiency as compared to simple ClosTron and random mutagenesis (Ng et al., 2013; Ehsaan et al., 2016b; Minton et al., 2016).

The *pyrE* cassettes consists of two arms, i.e., right homology arm (RHA) and left homology arm (LHA) with the internal region comprising of *pyrE* gene (**Figure 1B**). A plasmid is constructed with a selectable marker (antibiotic resistance gene), origin of replication and a sequence containing ∼300-bp homologous to *pyrE* gene and a longer sequence of ∼1,200-bp homologous to the adjacent region of 3′ end of *pyrE*. Once the *pyrE* based pseudo-suicide plasmid is delivered into *Clostridium* cells, single crossover is formed through homologous recombination. Subsequently, the single crossover mutant is inoculated into the media containing FOA and uracil (Heap et al., 2012). Metabolization of FOA kills the single crossover cells carrying the active *pyrE* gene. Inactivation of *pyrE* happens only if double recombination had occurred on both 1200-bp long sequence and 300-bp short sequence and the FOA does not affect the cells obtained by such double crossovers (Ng et al., 2013). The final double crossovers are formed by ACE of shorter left homology arm of 300-bp by the second single crossover, which

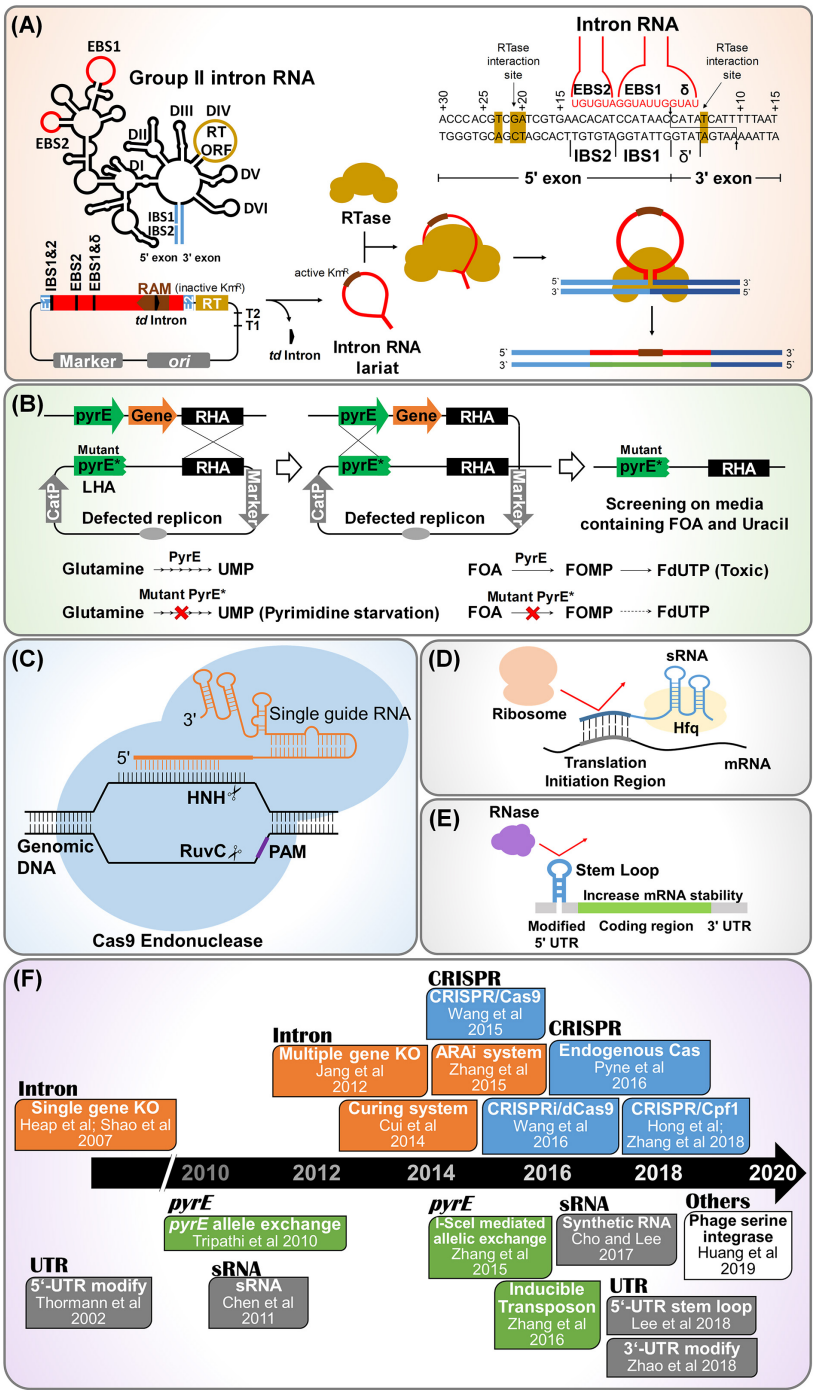


FIGURE 1 | Synthetic biology tools developed for genome and transcriptome engineering of solventogenic *Clostridium*. **(A)** Mobile group II intron-based genome engineering. Also known as ClosTron in context of *Clostridium* sp. In this technology, site directed gene disruption is achieved by insertion of the mobile group II intron into the target locus of chromosome. Abbreviations: RAM, retrotranscription-activated marker (typically kanamycin resistant marker containing self-splicing group I intron, phage T4 *td* intron); RTase, reverse transcriptase; EBS, exon binding site; IBS, intron binding site. **(B)** *pyrE* based allele exchange technology for genome engineering. Here, *pyr*E encoding orotate phosphoribosyl transferase is used as counter selection marker to ensure double crossover event. The *pyrE*-mutant (PyrE*) and wild type (PyrE) are resistant and sensitive to 5-fluoroorotic acid (FOA), respectively. Abbreviation: RHA, right homology arm. **(C)** CRISPR/Cas system for genome engineering. This needs single guide RNA containing crRNA and tracrRNA, Cas endonuclease, and homologous arm for recombination. Abbreviation: PAM, protospacer-adjacent motif. **(D)** Synthetic regulatory RNA (sRNA) based knockdown strategy. sRNA are having regulatory role in gene expression, mediated by chaperon Hfq. sRNA binds to complimentary mRNA sequences, prohibiting ribosome clamping at ribosome binding site located in translation initiation region. **(E)** 5′-UTR engineering for regulation of gene expression. The insertion of a small stem loop structure in the 5′-UTR increases the mRNA stability by blocking RNase, resulting in a high gene expression. **(F)** Timeline of notable events in the development of synthetic biology tools for genome and transcriptome engineering of solventogenic *Clostridium*.

also leads to the excision of the plasmid (Minton et al., 2016). This technology has been found to be applicable for many species of *Clostridium* genus (Heap et al., 2012).

Butanol yield in *C. pasteurianum* has been reported to be improved by application of *pyrE* based genome editing toolkit. For this, deletion mutations were created in three genes of *C. pasteurianum*: hydrogenase (*hydA*), redox response regulator (*rex*), and glycerol dehydratase (*dhaBCE*), using plasmid pMTL-KS01. This resulted in increased availability of NADPH in cell due to depletion of 1,3-propanediol synthesis, which eventually contributed to improved butanol production (Schwarz et al., 2017). Similarly, successful expression of cellulosomal subunits in *C. acetobutylicum* has also been achieved using this method (Kovacs et al., 2013). Few other *Clostridium* species modified using ACE technology includes *C. acetobutylicum*, *C. sporogenes*, and *C. difficile* (Heap et al., 2012; Ng et al., 2013; Bankar et al., 2015; Zhang J. et al., 2015; Ehsaan et al., 2016b; Minton et al., 2016; Willson et al., 2016).

CRISPR/CAS BASED CLOSTRIDIA GENOME ENGINEERING

Clustered regulatory interspaced short palindromic repeats (CRISPR) have been developed as one of the most advanced genetic engineering tools along with CRISPR-associated (Cas) protein (Doudna and Charpentier, 2014). As bacterial genome manipulation tool, CRISPR/Cas system needs single guide RNA (sgRNA), Cas endonuclease, and homologous arms for recombination (Jiang et al., 2013). The *Streptococcus pyogenes* type II CRISPR was the first CRISPR system which was exploited for genome engineering applications. Cas9 endonuclease is the basis of CRISPR based genome editing system. Cas9 recognize the protospacer adjacent motif (PAM) site (5′-NGG-3′ in *S. pyogenes*) and cleave at the 3′ end of the target gene (Mojica et al., 2009; Garneau et al., 2010; Jinek et al., 2012) (**Figure 1C**).

Various strains of *Clostridium* genus have been manipulated using the CRISPR/Cas9 system including *C. acetobutyricum* (Bruder et al., 2016; Li et al., 2016; Wasels et al., 2017), *C. beijerinkii* (Wang et al., 2016b), *C. autoethanogenum* (Nagaraju et al., 2016), *C. difficile* (McAllister et al., 2017; Wang et al., 2018), *C. cellulolyticum* (Xu et al., 2015, 2017), *C. pasteurianum* (Pyne et al., 2016), *C. ljungdahlii* (Huang et al., 2016), and *C. saccharoperbutylacetonicum* (Wang S. et al., 2017). However, the expression of Cas9 becomes detrimental for bacteria, including clostridia, in terms of the toxicity it causes. The mutation of 10^{th} amino acid (aspartic acid to alanine) in Cas9 inactivates it's RuvC-like nuclease domain resulting in formation of Cas9 nickase (Cas9n), which can cleave only single-strand of DNA (Jinek et al., 2012; Nishimasu et al., 2014; Swarts and Jinek, 2018; Li et al., 2019). Cas9n have advantage in terms of overcoming the toxicity caused by expression of Cas9. Introduction of highly regulated inducible promoter for Cas9 expression is another strategy to circumvent the associated toxicity (Wang et al., 2016a; McAllister et al., 2017; Wasels et al., 2017). Nevertheless, the CRISPR/Cas9n system is still being used for clostridia genome editing (Wang et al., 2016b; Wang S. et al., 2017; Wang Y. et al., 2017; Wang et al., 2018; McAllister et al., 2017; Wasels et al., 2017).

Moreover, modified CRISPR systems like CRISPR interference (CRISPRi) and dCas9 has also been developed to knockdown of the essential genes required for host survival (Jinek et al., 2012; Qi et al., 2013; Peters et al., 2016; Zheng et al., 2019). The dCas9 has two silenced catalytic domains (D10A and H840A; RuvC-like and HNH domains, respectively) which remains bound and block the target DNA instead of cleavage. CRISPRi/dCas9 system has also been applied to develop several mutant strains of *Clostridium* sp. (Bruder et al., 2016; Wang et al., 2016a,b; Wen et al., 2017; Woolston et al., 2018; Muh et al., 2019).

Similar to Cas9, the Cpf1 from *Acidaminococcus* sp. is another protein that is used for PAM recognition in CRISPR based system. While Cas9 recognizes G-rich PAM site, the PAM recognition site for Cpf1 is T-rich (5′-TTTN-3′) (Swarts and Jinek, 2018) making it best suited for application in AT-rich organisms like *Clostridium* sp. (Zetsche et al., 2015; Yamano et al., 2016). Single CRISPR/Cpf1 system plasmid can make multiple mutants in a single application (Zetsche et al., 2017; Hong et al., 2018; Zhang et al., 2018a). CRISPR/Cpf1 system has been applied in *C. ljungdahlii*, *C. difficile*, and *C. beijerinckii* (Hong et al., 2018; Zhang et al., 2018a; Zhao et al., 2019).

Additionally, endogenous CRISPR systems have been developed in *C. pasteurianum* and *C. tyrobutyricum* to overcome the toxic effect associated with Cas9 and Cpf1 endonucleases (Pyne et al., 2016; Zhang et al., 2018b). The endogenous CRISPR system uses endonuclease encoded by the genome and can contain multiple pre-crRNAs under one promoter, facilitating multiple genome modification using a single plasmid (Luo et al., 2014; Makarova et al., 2015; Pyne et al., 2016; Zhang et al., 2018b).

SYNTHETIC SRNA AND UNTRANSLATED REGION ENGINEERING AS POTENTIAL DOMAINS FOR *CLOSTRIDIUM* STRAIN IMPROVEMENT

Prokaryotic small RNAs (sRNA) are short strands of ribonucleotides (about 50–500 nucleotides) which have a regulatory role in maintaining the cellular processes (Gottesman, 2004). Based on the existence of natural sRNA, synthetic small RNAs are produced to alter the gene expression of the organisms. Many such naturally occurring sRNAs have been detected and analyzed in *Clostridium* sp. (Chen et al., 2011), which leads to the development of synthetic sRNA (Na et al., 2013).

The sRNA mediated gene expression usually results in repression of the gene which complements the sRNA nucleotide sequence, mediated by a protein called Hfq (De Lay et al., 2013). Hfq is the chaperone mediated protein which stabilizes the sRNA-mRNA binding. The translation process is prevented by sRNA binding to ribosome binding site (RBS) or by masking the access to the start codon (Na et al., 2013; Yoo et al., 2013). Recently, Cho and Lee (2017) have reported the development of synthetic small

regulatory RNA (sRNA) system for controlled gene expression in *C. acetobutylicum*, consisting of a target recognition site, MicC scaffold, and an RNA chaperon Hfq (**Figure 1D**). In this study, *C. acetobutylicum* Hfq was found to be ineffective in binding with *Escherichia coli* MicC scaffold-based synthetic sRNA, however Hfq from *E. coli* itself resulted in much enhanced knockdown efficiency. This *E. coli* MicC-Hfq sRNA system was used to knockdown *adhE1* gene expression resulting in 40% reduction in butanol production. Further, this synthetic sRNA system was used to knockdown the *pta* gene expression in PJC4BK strain, resulting in PJC4BK (pPta-Hfq^{Eco}) strain with improvement of butanol titer from 14.9 to 16.9 g/l (Cho and Lee, 2017).

Untranslated regions (UTRs) are non-coding regions in the mRNA helps to regulate the gene expression. UTRs are present on both the ends of the mRNA (5′-UTR and 3′-UTR) (**Figure 1E**). There are sufficient reports to confirm that the 5′-UTR in *C. acetobutylicum* has the regulatory effect on the secondary structure of enzyme *adhE1,* which is involved in solvent production (Thormann et al., 2002; Scotcher et al., 2003). Lee et al. (2016) has recently found that the presence of a single stranded short 5′-UTR in the solventogenic *C. acetobutylicum* leads to decreased gene expression (**Figure 1E**). The insertion of a small stem loop structure in the 5′-UTR was found to increase the mRNA stability and gene expression by 4.6 folds, without any modification in the promoter or RBS (Lee et al., 2016). On the other hand, the 3′-UTR mostly harbors the terminator sequence for transcription process in mRNA (Richard and Manley, 2009). sRNA sequence containing the codons that regulates the post transcriptional and translation machinery is also attached to 3′-UTR. Most importantly 3′-UTR confer stability to the mRNA (Zhao et al., 2018). Although, there are very limited studies related to 3′-UTR regions in *Clostridium*, the presence of transcripts with long 3′-UTR is confirmed in *Clostridium* (Ralston and Papoutsakis, 2018). Although several RNAseq studies were reported in the *Clostridium*, only few studies show the data related to regulation of mRNA based on 5′- and 3′-UTRs, leaderless transcripts and non-coding RNA (Soutourina et al., 2013; Wilson et al., 2013; Sedlar et al., 2018). Further research in RNAseq and proteomics will explore the complex regulations that control mRNA stability and degradation, which will be more useful to construction synthetic toolkit.

In conclusion, many *Clostridium* sp. have potential to be utilized at industrial scale to produce value added chemicals, including butanol as fossil fuel substitute. Up to date, their true potential was underexploited due to challenges in strain improvement and unavailability of genome and transcriptome editing tools for this genus. Nevertheless, during the last decade, synthetic biology toolkits for *Clostridium* sp. have been expanded rapidly (**Figure 1F**). Furthermore, a recent advancement, such as phage serine integrase mediated site-specific genome engineering technique for *C. ljungdahlii* could be extended to other *Clostridium* species (Huang et al., 2019). The synthetic biology techniques that have been applied in other microorganisms may also be adopted to solventogenic clostridia in the near future: CRISPR associated site-specific insertion of transposons and base editing techniques (Ronda et al., 2015; Zhang et al., 2016; Lim and Choi, 2019; Strecker et al., 2019b). Utilization of improved clostridia strains could be a starting point for development of an industrial scale, commercially viable bio-based fuel and chemical production using *Clostridium* sp. using a consolidated bioprocessing concept (Wen et al., 2019a). Furthermore, these synthetic biology tools could be applied to another biotechnology fields such as degradation of plastics, such as polyethylene terephthalate and polyethylene.

AUTHOR CONTRIBUTIONS

Y-SJ and AM conceived the project. All authors analyzed the literature, compiled data, planned content, wrote the manuscript, read, and approved the final manuscript.

FUNDING

This work was supported by a grant from the Ministry of Science and ICT (MSIT) through the National Research Foundation (NRF) of Korea (NRF-2019R1A4A1029125). This work was further supported by Cooperative Research Program for Agricultural Science and Technology Development (Project No. PJ01492601), Rural Development Administration, Republic of Korea. AM was supported by a grant from Centre for Research Projects, CHRIST (Deemed to be University), Bangalore (MRPDSC-1829).

REFERENCES

Arora, R., Behera, S., Sharma, N. K., and Kumar, S. (2019). Evaluating the pathway for co-fermentation of glucose and xylose for enhanced bioethanol production using flux balance analysis. *Biotechnol. Bioproc. Eng.* 24, 924–933. doi: 10.1007/s12257-019-0026-5

Baban, S. T., Kuehne, S. A., Barketi-Klai, A., Cartman, S. T., Kelly, M. L., Hardie, K. R., et al. (2013). The role of flagella in *Clostridium difficile* pathogenesis: comparison between a non-epidemic and an epidemic strain. *PLoS One* 8:e73026. doi: 10.1371/journal.pone.0073026

Banerjee, S., Mishra, G., and Roy, A. (2019). Metabolic engineering of bacteria for renewable bioethanol production from cellulosic biomass. *Biotechnol. Bioproc. Eng.* 24, 713–733. doi: 10.1007/s12257-019-0134-2

Bankar, S. B., Jurgens, G., Survase, S. A., Ojamo, H., and Granström, T. (2015). Genetic engineering of *Clostridium acetobutylicum* to enhance isopropanol-butanol-ethanol production with an integrated DNA-technology approach. *Renew. Energy* 83, 1076–1083. doi: 10.1016/j.renene.2015.05.052

Bruder, M. R., Pyne, M. E., Moo-Young, M., Chung, D. A., and Chou, C. P. (2016). Extending CRISPR-Cas9 technology from genome editing to transcriptional engineering in the genus *Clostridium*. *Appl. Environ. Microbiol.* 82, 6109–6119. doi: 10.1128/aem.02128-16

Chen, Y., Caruso, L., McClane, B., Fisher, D., and Gupta, P. (2007). Disruption of a toxin gene by introduction of a foreign gene into the chromosome of *Clostridium perfringens* using targetron-induced mutagenesis. *Plasmid* 58, 182–189. doi: 10.1016/j.plasmid.2007.04.002

Chen, Y., Indurthi, D. C., Jones, S. W., and Papoutsakis, E. T. (2011). Small RNAs in the genus *Clostridium*. *mBio* 2:e340-10.

Chen, Y., McClane, B. A., Fisher, D. J., Rood, J. I., and Gupta, P. (2005). Construction of an alpha toxin gene knockout mutant of *Clostridium perfringens* type A by use of a mobile group II intron. *Appl. Environ. Microbiol.* 71, 7542–7547. doi: 10.1128/AEM.71.11.7542-7547.2005

Cho, C., and Lee, S. Y. (2017). Efficient gene knockdown in *Clostridium acetobutylicum* by synthetic small regulatory RNAs. *Biotechnol. Bioeng.* 114, 374–383. doi: 10.1002/bit.26077

Choi, Y. Y., Hong, M.-E., Chang, W. S., and Sim, S. J. (2019). Autotrophic biodiesel production from the thermotolerant microalga *Chlorella sorokiniana* by enhancing the carbon availability with temperature adjustment. *Biotechnol. Bioproc. Eng.* 24, 223–231. doi: 10.1007/s12257-018-0375-5

Croux, C., Nguyen, N. P., Lee, J., Raynaud, C., Saint-Prix, F., Gonzalez-Pajuelo, M., et al. (2016). Construction of a restriction-less, marker-less mutant useful for functional genomic and metabolic engineering of the biofuel producer *Clostridium acetobutylicum*. *Biotechnol. Biofuels* 9:23. doi: 10.1186/s13068-016-0432-2

Cui, G. Z., Hong, W., Zhang, J., Li, W. L., Feng, Y., Liu, Y. J., et al. (2012). Targeted gene engineering in *Clostridium cellulolyticum* H10 without methylation. *J. Microbiol. Methods* 89, 201–208. doi: 10.1016/j.mimet.2012.02.015

Cui, G. Z., Zhang, J., Hong, W., Xu, C., Feng, Y., Cui, Q., et al. (2014). Improvement of ClosTron for successive gene disruption in *Clostridium cellulolyticum* using a *pyrF*-based screening system. *Appl. Microbiol. Biotechnol.* 98, 313–323. doi: 10.1007/s00253-013-5330-y

De Lay, N., Schu, D. J., and Gottesman, S. (2013). Bacterial small RNA-based negative regulation: Hfq and its accomplices. *J. Biol. Chem.* 288, 7996–8003. doi: 10.1074/jbc.r112.441386

Dingle, T. C., Mulvey, G. L., and Armstrong, G. D. (2011). Mutagenic analysis of the *Clostridium difficile* flagellar proteins, FliC and FliD, and their contribution to virulence in hamsters. *Infect. Immun.* 79, 4061–4067. doi: 10.1128/iai.05305-11

Doudna, J. A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346:1258096.

Ehsaan, M., Kuehne, S. A., and Minton, N. P. (2016a). "*Clostridium difficile* genome editing using pyrE alleles," in *Clostridium Difficile*, eds A. P. Roberts, and P. Mullany, (Berlin: Springer), 35–52. doi: 10.1007/978-1-4939-6361-4_4

Ehsaan, M., Kuit, W., Zhang, Y., Cartman, S. T., Heap, J. T., Winzer, K., et al. (2016b). Mutant generation by allelic exchange and genome resequencing of the biobutanol organism *Clostridium acetobutylicum* ATCC 824. *Biotechnol. Biofuels* 9:4.

Garneau, J. E., Dupuis, M. E., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., et al. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71. doi: 10.1038/nature09523

Gottesman, S. (2004). The small RNA regulators of *Escherichia coli*: roles and mechanisms. *Annu. Rev. Microbiol.* 58, 303–328. doi: 10.1146/annurev.micro.58.030603.123841

Heap, J. T., Ehsaan, M., Cooksley, C. M., Ng, Y. K., Cartman, S. T., Winzer, K., et al. (2012). Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. *Nucleic Acids Res.* 40:e59. doi: 10.1093/nar/gkr1321

Heap, J. T., Kuehne, S. A., Ehsaan, M., Cartman, S. T., Cooksley, C. M., Scott, J. C., et al. (2010). The ClosTron: mutagenesis in *Clostridium* refined and streamlined. *J. Microbiol. Methods* 80, 49–55. doi: 10.1016/j.mimet.2009.10.018

Heap, J. T., Pennington, O. J., Cartman, S. T., Carter, G. P., and Minton, N. P. (2007). The ClosTron: a universal gene knock-out system for the genus *Clostridium*. *J. Microbiol. Methods* 70, 452–464. doi: 10.1016/j.mimet.2007.05.021

Hong, W., Zhang, J., Cui, G., Wang, L., and Wang, Y. (2018). Multiplexed CRISPR-Cpf1-mediated genome editing in *Clostridium difficile* toward the understanding of pathogenesis of *C. difficile* infection. *ACS Synth. Biol.* 7, 1588–1600. doi: 10.1021/acssynbio.8b00087

Honicke, D., Lutke-Eversloh, T., Liu, Z., Lehmann, D., Liebl, W., and Ehrenreich, A. (2014). Chemostat cultivation and transcriptional analyses of *Clostridium acetobutylicum* mutants with defects in the acid and acetone biosynthetic pathways. *Appl. Microbiol. Biotechnol.* 98, 9777–9794. doi: 10.1007/s00253-014-6040-9

Huang, H., Chai, C., Li, N., Rowe, P., Minton, N. P., Yang, S., et al. (2016). CRISPR/Cas9-Based efficient genome editing in *Clostridium ljungdahlii*, an autotrophic gas-fermenting bacterium. *ACS Synth. Biol.* 5, 1355–1361. doi: 10.1021/acssynbio.6b00044

Huang, H., Chai, C., Yang, S., Jiang, W., and Gu, Y. (2019). Phage serine integrase-mediated genome engineering for efficient expression of chemical biosynthetic pathway in gas-fermenting *Clostridium ljungdahlii*. *Metab. Eng* 52, 293–302. doi: 10.1016/j.ymben.2019.01.005

Jang, Y.-S., Im, J. A., Choi, S. Y., Lee, J. I., and Lee, S. Y. (2014). Metabolic engineering of *Clostridium acetobutylicum* for butyric acid production with high butyric acid selectivity. *Metab. Eng.* 23, 165–174. doi: 10.1016/j.ymben.2014.03.004

Jang, Y. S., Lee, J. Y., Lee, J., Park, J. H., Im, J. A., Eom, M. H., et al. (2012). Enhanced butanol production obtained by reinforcing the direct butanol-forming route in *Clostridium acetobutylicum*. *mBio* 3:e00314-12.

Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L. A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* 31, 233–239. doi: 10.1038/nbt.2508

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10.1126/science.1225829

Jones, A. J., Venkataramanan, K. P., and Papoutsakis, T. (2016). Overexpression of two stress-responsive, small, non-coding RNAs, 6S and tmRNA, imparts butanol tolerance in *Clostridium acetobutylicum*. *FEMS Microbiol. Lett.* 363:fnw063. doi: 10.1093/femsle/fnw063

Joseph, R. C., Kim, N. M., and Sandoval, N. R. (2018). Recent developments of the synthetic biology toolkit for *Clostridium*. *Front. Microbiol.* 9:154. doi: 10.3389/fmicb.2018.00154

Kovacs, K., Willson, B. J., Schwarz, K., Heap, J. T., Jackson, A., Bolam, D. N., et al. (2013). Secretion and assembly of functional mini-cellulosomes from synthetic chromosomal operons in *Clostridium acetobutylicum* ATCC 824. *Biotechnol. Biofuels* 6:117. doi: 10.1186/1754-6834-6-117

Kuehne, S. A., Rood, J. I., and Lyras, D. (2019). Clostridial genetics: genetic manipulation of the pathogenic clostridia. *Microbiol Spectr* 7:GPP3-0040-2018.

Lawson, P. A., and Rainey, F. A. (2016). Proposal to restrict the genus *Clostridium prazmowski* to *Clostridium butyricum* and related species. *Int. J. Syst. Evol. Microbiol.* 66, 1009–1016. doi: 10.1099/ijsem.0.000824

Lawson, P. A., Citron, D. M., Tyrrell, K. L., and Finegold, S. M. (2016). Reclassification of *Clostridium difficile* as *Clostridioides difficile* (Hall and O'Toole 1935) Prevot 1938. *Anaerobe* 40, 95–99. doi: 10.1016/j.anaerobe.2016.06.008

Lee, J., Jang, Y. S., Papoutsakis, E. T., and Lee, S. Y. (2016). Stable and enhanced gene expression in *Clostridium acetobutylicum* using synthetic untranslated regions with a stem-loop. *J. Biotechnol.* 230, 40–43. doi: 10.1016/j.jbiotec.2016.05.020

Li, Q., Chen, J., Minton, N. P., Zhang, Y., Wen, Z., Liu, J., et al. (2016). CRISPR-based genome editing and expression control systems in *Clostridium acetobutylicum* and *Clostridium beijerinckii*. *Biotechnol. J.* 11, 961–972.

Li, Q., Seys, F. M., Minton, N. P., Yang, J., Jiang, Y., Jiang, W., et al. (2019). CRISPR-Cas9(D10A) nickase-assisted base editing in the solvent producer *Clostridium beijerinckii*. *Biotechnol. Bioeng.* 116, 1475–1483. doi: 10.1002/bit.26949

Lim, H., and Choi, S. K. (2019). Programmed gRNA removal system for CRISPR-Cas9-mediated multi-round genome editing in *Bacillus subtilis*. *Front. Microbiol.* 10:1140. doi: 10.3389/fmicb.2019.01140

Liu, J., Guo, T., Wang, D., Shen, X., Liu, D., Niu, H., et al. (2016). Enhanced butanol production by increasing NADH and ATP levels in *Clostridium beijerinckii* NCIMB 8052 by insertional inactivation of Cbei_4110. *Appl. Microbiol. Biotechnol.* 100, 4985–4996. doi: 10.1007/s00253-016-7299-9

Liu, J., Sun, J., Wang, F., Yu, X., Ling, Z., Li, H., et al. (2015). Neuroprotective effects of *Clostridium butyricum* against vascular dementia in mice via metabolic butyrate. *Biomed. Res. Int.* 2015:412946.

Liu, Y. J., Zhang, J., Cui, G. Z., and Cui, Q. (2015). Current progress of targetron technology: development, improvement and application in metabolic engineering. *Biotechnol. J.* 10, 855–865. doi: 10.1002/biot.201400716

Luo, M. L., Mullis, A. S., Leenay, R. T., and Beisel, C. L. (2014). Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression. *Nucleic Acids Res.* 43, 674–681. doi: 10.1093/nar/gku971

Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13, 722–736.

Malaviya, A., Jang, Y. S., and Lee, S. Y. (2012). Continuous butanol production with reduced byproducts formation from glycerol by a hyper producing mutant

of *Clostridium pasteurianum*. *Appl. Microbiol. Biotechnol.* 93, 1485–1494. doi: 10.1007/s00253-011-3629-0

McAllister, K. N., Bouillaut, L., Kahn, J. N., Self, W. T., and Sorg, J. A. (2017). Using CRISPR-Cas9-mediated genome editing to generate *C. difficile* mutants defective in selenoproteins synthesis. *Sci. Rep.* 7:14672.

McAllister, K. N., and Sorg, J. A. (2019). CRISPR genome editing systems in the genus *Clostridium*: a timely advancement. *J. Bacteriol.* 201:e00219-19.

Meaney, C. A., Cartman, S. T., McClure, P. J., and Minton, N. P. (2015). Optimal spore germination in *Clostridium botulinum* ATCC 3502 requires the presence of functional copies of SleB and YpeB, but not CwlJ. *Anaerobe* 34, 86–93.

Meaney, C. A., Cartman, S. T., McClure, P.J., and Minton, N. P. (2016). The role of small acid-soluble proteins (SASPs) in protection of spores of *Clostridium botulinum* against nitrous acid. *Int. J. Food Microbiol.* 216, 25–30. doi: 10.1016/j.ijfoodmicro.2015.08.024

Minton, N. P., Ehsaan, M., Humphreys, C. M., Little, G. T., Baker, J., Henstra, A. M., et al. (2016). A roadmap for gene system development in *Clostridium*. *Anaerobe* 41, 104–112. doi: 10.1016/j.anaerobe.2016.05.011

Mohr, G., Hong, W., Zhang, J., Cui, G.-Z., Yang, Y., Cui, Q., et al. (2013). A targetron system for gene targeting in thermophiles and its application in *Clostridium thermocellum*. *PLoS ONE* 8:e69032. doi: 10.1371/journal.pone.0069032

Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155, 733–740. doi: 10.1099/mic.0.023960-0

Moon, H. G., Jang, Y. S., Cho, C., Lee, J., Binkley, R., and Lee, S. Y. (2016). One hundred years of clostridial butanol fermentation. *FEMS Microbiol Lett* 363:fnw001. doi: 10.1093/femsle/fnw001

Muh, U., Pannullo, A. G., Weiss, D. S., and Ellermeier, C. D. (2019). A xylose-inducible expression system and a CRISPR interference plasmid for targeted knockdown of gene expression in *Clostridioides difficile*. *J. Bacteriol.* 201, e711–e718.

Na, D., Yoo, S. M., Chung, H., Park, H., Park, J. H., and Lee, S. Y. (2013). Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. *Nat. Biotechnol.* 31, 170–174. doi: 10.1038/nbt.2461

Nagaraju, S., Davies, N. K., Walker, D. J., Kopke, M., and Simpson, S. D. (2016). Genome editing of *Clostridium autoethanogenum* using CRISPR/Cas9. *Biotechnol. Biofuels* 9:219.

Ng, Y. K., Ehsaan, M., Philip, S., Collery, M. M., Janoir, C., Collignon, A., et al. (2013). Expanding the repertoire of gene tools for precise manipulation of the *Clostridium difficile* genome: allelic exchange using *pyrE* alleles. *PLoS ONE* 8:e56051. doi: 10.1371/journal.pone.0056051

Nishimasu, H., Ran, F. A., Hsu, P. D., Konermann, S., Shehata, S. I., Dohmae, N., et al. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* 156, 935–949. doi: 10.1016/j.cell.2014.02.001

Noh, H. J., Woo, J. E., Lee, S. Y., and Jang, Y.-S. (2018). Metabolic engineering of *Clostridium acetobutylicum* for the production of butyl butyrate. *Appl. Microbiol. Biotechnol.* 102, 8319–8327. doi: 10.1007/s00253-018-9267-z

Park, M.-R., Kim, S.-K., and Jeong, G.-T. (2018). Biosugar production from *Gracilaria verrucosa* with sulfamic acid pretreatment and subsequent enzymatic hydrolysis. *Biotechnol. Bioproc. Eng.* 23, 302–310. doi: 10.1007/s12257-018-0090-2

Peters, J. M., Colavin, A., Shi, H., Czarny, T. L., Larson, M. H., Wong, S., et al. (2016). A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell* 165, 1493–1506. doi: 10.1016/j.cell.2016.05.003

Pyne, M. E., Bruder, M., Moo-Young, M., Chung, D. A., and Chou, C. P. (2014). Technical guide for genetic advancement of underdeveloped and intractable *Clostridium*. *Biotechnol. Adv.* 32, 623–641. doi: 10.1016/j.biotechadv.2014.04.003

Pyne, M. E., Bruder, M. R., Moo-Young, M., Chung, D. A., and Chou, C. P. (2016). Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in *Clostridium*. *Sci. Rep.* 6:25666.

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., et al. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173–1183. doi: 10.1016/j.cell.2013.02.022

Ralston, M. T., and Papoutsakis, E. T. (2018). RNAseq-based transcriptome assembly of *Clostridium acetobutylicum* for functional genome annotation and discovery. *AIChE J.* 64, 4271–4280. doi: 10.1002/aic.16396

Rhie, M. N., Kim, H. T., Jo, S. Y., Chu, L. L., Baritugo, K.-A., Baylon, M. G., et al. (2019). Recent advances in the metabolic engineering of *Klebsiella pneumoniae*: a potential platform microorganism for biorefineries. *Biotechnol. Bioproc. Eng.* 24, 48–64. doi: 10.1007/s12257-018-0346-x

Richard, P., and Manley, J. L. (2009). Transcription termination by nuclear RNA polymerases. *Genes Dev.* 23, 1247–1269. doi: 10.1101/gad.1792809

Ronda, C., Maury, J., Jakociunas, T., Jacobsen, S. A., Germann, S. M., Harrison, S. J., et al. (2015). CrEdit: CRISPR mediated multi-loci gene integration in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* 14:97.

Sato, T., Fukui, T., Atomi, H., and Imanaka, T. (2005). Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon *Thermococcus kodakaraensis*. *Appl. Environ. Microbiol.* 71, 3889–3899. doi: 10.1128/aem.71.7.3889-3899.2005

Schwarz, K. M., Grosse-Honebrink, A., Derecka, K., Rotta, C., Zhang, Y., and Minton, N. P. (2017). Towards improved butanol production through targeted genetic modification of *Clostridium pasteurianum*. *Metab. Eng.* 40, 124–137. doi: 10.1016/j.ymben.2017.01.009

Scotcher, M. C., Huang, K. X., Harrison, M. L., Rudolph, F. B., and Bennett, G. N. (2003). Sequences affecting the regulation of solvent production in *Clostridium acetobutylicum*. *J. Ind. Microbiol. Biotechnol.* 30, 414–420. doi: 10.1007/s10295-003-0057-x

Sedlar, K., Koscova, P., Vasylkivska, M., Branska, B., Kolek, J., Kupkova, K., et al. (2018). Transcription profiling of butanol producer *Clostridium beijerinckii* NRRL B-598 using RNA-Seq. *BMC Genomics* 19:415. doi: 10.1186/s12864-018-4805-8

Shao, L., Hu, S., Yang, Y., Gu, Y., Chen, J., Yang, Y., et al. (2007). Targeted gene disruption by use of a group II intron (targetron) vector in *Clostridium acetobutylicum*. *Cell Res.* 17, 963–965. doi: 10.1038/cr.2007.91

Soutourina, O. A., Monot, M., Boudry, P., Saujet, L., Pichon, C., Sismeiro, O., et al. (2013). Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*. *PLoS Genet.* 9:e1003493. doi: 10.1371/journal.pgen.1003493

Staedtke, V., Roberts, N. J., Bai, R. Y., and Zhou, S. (2016). *Clostridium novyi*-NT in cancer therapy. *Genes Dis.* 3, 144–152. doi: 10.1016/j.gendis.2016.01.003

Strecker, J., Jones, S., Koopal, B., Schmid-Burgk, J., Zetsche, B., Gao, L., et al. (2019a). Engineering of CRISPR-Cas12b for human genome editing. *Nat. Commun.* 10:212. doi: 10.1017/cbo9781316771440.007

Strecker, J., Ladha, A., Gardner, Z., Schmid-Burgk, J. L., Makarova, K. S., Koonin, E. V., et al. (2019b). RNA-guided DNA insertion with CRISPR-associated transposases. *Science* 365, 48–53. doi: 10.1126/science.aax9181

Swarts, D. C., and Jinek, M. (2018). *Cas9 versus Cas12a/Cpf1: Structure-Function Comparisons and Implications for Genome Editing*. Hoboken, NJ: Wiley Interdiscip. Rev. RNA, e1481.

Thormann, K., Feustel, L., Lorenz, K., Nakotte, S., and Durre, P. (2002). Control of butanol formation in *Clostridium acetobutylicum* by transcriptional activation. *J. Bacteriol.* 184, 1966–1973. doi: 10.1128/jb.184.7.1966-1973.2002

Tripathi, S. A., Olson, D. G., Argyros, D. A., Miller, B. B., Barrett, T. F., Murphy, D. M., et al. (2010). Development of *pyrF*-based genetic system for targeted gene deletion in *Clostridium thermocellum* and creation of a *pta* mutant. *Appl. Environ. Microbiol.* 76, 6591–6599. doi: 10.1128/aem.01484-10

Wang, S., Dong, S., Wang, P., Tao, Y., and Wang, Y. (2017). Genome editing in *Clostridium saccharoperbutylacetonicum* N1-4 with the CRISPR-Cas9 system. *Appl. Environ. Microbiol.* 83:e233-17.

Wang, Y., Zhang, G., Zhao, X., and Ling, J. (2017). Genome shuffling improved the nucleosides production in *Cordyceps kyushuensis*. *J. Biotechnol.* 260, 42–47. doi: 10.1016/j.jbiotec.2017.08.021

Wang, S., Hong, W., Dong, S., Zhang, Z. T., Zhang, J., Wang, L., et al. (2018). Genome engineering of *Clostridium difficile* using the CRISPR-Cas9 system. *Clin. Microbiol. Infect.* 24, 1095–1099.

Wang, Y., Zhang, Z. T., Seo, S. O., Lynn, P., Lu, T., Jin, Y. S., et al. (2016a). Bacterial genome editing with CRISPR-Cas9: deletion, integration, single nucleotide modification, and desirable "clean" mutant selection in *Clostridium beijerinckii* as an example. *ACS Synth. Biol.* 5, 721–732. doi: 10.1021/acssynbio.6b00060

Wang, Y., Zhang, Z. T., Seo, S. O., Lynn, P., Lu, T., Jin, Y. S., et al. (2016b). Gene transcription repression in *Clostridium beijerinckii* using CRISPR-dCas9. *Biotechnol. Bioeng.* 113, 2739–2743. doi: 10.1002/bit.26020

Wasels, F., Jean-Marie, J., Collas, F., Lopez-Contreras, A. M., and Lopes Ferreira, N. (2017). A two-plasmid inducible CRISPR/Cas9 genome editing tool for

Clostridium acetobutylicum. *J. Microbiol. Methods* 140, 5–11. doi: 10.1016/j.mimet.2017.06.010

Wen, Z., Ledesma-Amaro, R., Lin, J., Jiang, Y., and Yang, S. (2019a). Improved n-butanol production from *Clostridium cellulovorans* by integrated metabolic and evolutionary engineering. *Appl. Environ. Microbiol* 85, e2560–e2518.

Wen, Z., Li, Q., Liu, J., Jin, M., and Yang, S. (2019b). Consolidated bioprocessing for butanol production of cellulolytic clostridia: development and optimization. *Microb. Biotechnol.* 13, 410–422. doi: 10.1111/1751-7915.13478

Wen, Z., Lu, M., Ledesma-Amaro, R., Li, Q., Jin, M., and Yang, S. (2019c). Targetron technology applicable in solventogenic clostridia: revisiting 12 years' advances. *Biotechnol. J.* 15:e1900284.

Wen, Z., Minton, N. P., Zhang, Y., Li, Q., Liu, J., Jiang, Y., et al. (2017). Enhanced solvent production by metabolic engineering of a twin-clostridial consortium. *Metab. Eng.* 39, 38–48. doi: 10.1016/j.ymben.2016.10.013

Willson, B. J., Kovacs, K., Wilding-Steele, T., Markus, R., Winzer, K., and Minton, N. P. (2016). Production of a functional cell wall-anchored minicellulosome by recombinant *Clostridium acetobutylicum* ATCC 824. *Biotechnol. Biofuels* 9:109.

Wilson, C. M., Rodriguez, M. Jr., Johnson, C. M., Martin, S. L., Chu, T. M., Wolfinger, R. D., et al. (2013). Global transcriptome analysis of *Clostridium thermocellum* ATCC 27405 during growth on dilute acid pretreated *Populus* and switchgrass. *Biotechnol. Biofuels* 6:179.

Woo, J. E., Lee, S. Y., and Jang, Y.-S. (2018). Effects of nutritional enrichment on acid production from degenerated (non-solventogenic) *Clostridium acetobutylicum* strain M5. *Appl. Biol. Chem.* 61, 469–472. doi: 10.1007/s13765-018-0372-6

Woolston, B. M., Emerson, D. F., Currie, D. H., and Stephanopoulos, G. (2018). Rediverting carbon flux in *Clostridium ljungdahlii* using CRISPR interference (CRISPRi). *Metab. Eng.* 48, 243–253. doi: 10.1016/j.ymben.2018.06.006

Xin, F., Yan, W., Zhou, J., Wu, H., Dong, W., Ma, J., et al. (2018). Exploitation of novel wild type solventogenic strains for butanol production. *Biotechnol. Biofuels* 11:252.

Xu, T., Li, Y., He, Z., Van Nostrand, J. D., and Zhou, J. (2017). Cas9 nickase-assisted RNA repression enables stable and efficient manipulation of essential metabolic genes in *Clostridium cellulolyticum*. *Front. Microbiol.* 8:1744. doi: 10.3389/fmicb.2017.01744

Xu, T., Li, Y., Shi, Z., Hemme, C. L., Li, Y., Zhu, Y., et al. (2015). Efficient genome editing in *Clostridium cellulolyticum* via CRISPR-Cas9 nickase. *Appl. Environ. Microbiol.* 81, 4423–4431. doi: 10.1128/aem.00873-15

Yamano, T., Nishimasu, H., Zetsche, B., Hirano, H., Slaymaker, I. M., Li, Y., et al. (2016). Crystal structure of Cpf1 in complex with guide RNA and target DNA. *Cell* 165, 949–962.

Yoo, S. M., Na, D., and Lee, S. Y. (2013). Design and use of synthetic regulatory small RNAs to control gene expression in *Escherichia coli*. *Nat. Protoc.* 8, 1694–1707. doi: 10.1038/nprot.2013.105

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163, 759–771. doi: 10.1016/j.cell.2015.09.038

Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E. M., et al. (2017). Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat. Biotechnol.* 35, 31–34. doi: 10.1038/nbt.3737

Zhang, J., Hong, W., Zong, W., Wang, P., and Wang, Y. (2018a). Markerless genome editing in *Clostridium beijerinckii* using the CRISPR-Cpf1 system. *J. Biotechnol.* 284, 27–30. doi: 10.1016/j.jbiotec.2018.07.040

Zhang, J., Zong, W., Hong, W., Zhang, Z. T., and Wang, Y. (2018b). Exploiting endogenous CRISPR-Cas system for multiplex genome editing in *Clostridium tyrobutyricum* and engineer the strain for high-level butanol production. *Metab. Eng.* 47, 49–59. doi: 10.1016/j.ymben.2018.03.007

Zhang, J., Liu, Y. J., Cui, G. Z., and Cui, Q. (2015). A novel arabinose-inducible genetic operation system developed for *Clostridium cellulolyticum*. *Biotechnol. Biofuels* 8:36. doi: 10.1186/s13068-015-0214-2

Zhang, N., Shao, L., Jiang, Y., Gu, Y., Li, Q., Liu, J., et al. (2015). I-SceI-mediated scarless gene modification via allelic exchange in *Clostridium*. *J. Microbiol. Methods* 108, 49–60. doi: 10.1016/j.mimet.2014.11.004

Zhang, Y., Xu, S., Chai, C., Yang, S., Jiang, W., Minton, N. P., et al. (2016). Development of an inducible transposon system for efficient random mutagenesis in *Clostridium acetobutylicum*. *FEMS Microbiol. Lett.* 363:fnw065. doi: 10.1093/femsle/fnw065

Zhao, J. P., Zhu, H., Guo, X. P., and Sun, Y. C. (2018). AU-rich long 3' untranslated region regulates gene expression in bacteria. *Front. Microbiol.* 9:3080. doi: 10.3389/fmicb.2018.03080

Zhao, R., Liu, Y., Zhang, H., Chai, C., Wang, J., Jiang, W., et al. (2019). CRISPR-cas12a-mediated gene deletion and regulation in *Clostridium ljungdahlii* and its application in carbon flux redirection in synthesis gas fermentation. *ACS Synth. Biol.* 8, 2270–2279. doi: 10.1021/acssynbio.9b00033

Zheng, Y., Su, T., and Qi, Q. (2019). Microbial CRISPRi and CRISPRa systems for metabolic engineering. *Biotechnol. Bioproc. Eng.* 24, 579–591.

Zhong, J., Karberg, M., and Lambowitz, A. M. (2003). Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. *Nucleic Acids Res.* 31, 1656–1664. doi: 10.1093/nar/gkg248

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 03 June 2020
doi: 10.3389/fbioe.2020.00344

Improved Prodigiosin Production by Relieving CpxR Temperature-Sensitive Inhibition

Yang Sun[1], Lijun Wang[1], Xuewei Pan[1], Tolbert Osire[1], Haitian Fang[2,3], Huiling Zhang[2,3], Shang-Tian Yang[4], Taowei Yang[1] and Zhiming Rao[1]*†*

[1] The Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi, China, [2] Ningxia Key Laboratory for Food Microbial-Applications Technology and Safety Control, Yinchuan, China, [3] School of Agriculture, Ningxia University, Yinchuan, China, [4] Department of Chemical and Biomolecular Engineering, The Ohio State University, Columbus, OH, United States

OPEN ACCESS

Edited by:
Jiazhang Lian,
Zhejiang University, China

Reviewed by:
Dae-Hee Lee,
Korea Research Institute of Bioscience and Biotechnology (KRIBB), South Korea
Jiabin Li,
First Affiliated Hospital, Anhui Medical University, China

***Correspondence:**
Taowei Yang
ytw1228@163.com
Zhiming Rao
raozhm@jiangnan.edu.cn

†Present address:
Zhiming Rao,
School of Biotechnology, Jiangnan University, Wuxi, China

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *25 November 2019*
Accepted: *27 March 2020*
Published: *03 June 2020*

Citation:
Sun Y, Wang L, Pan X, Osire T, Fang H, Zhang H, Yang S-T, Yang T and Rao Z (2020) Improved Prodigiosin Production by Relieving CpxR Temperature-Sensitive Inhibition. Front. Bioeng. Biotechnol. 8:344. doi: 10.3389/fbioe.2020.00344

Prodigiosin (PG) is a typical secondary metabolite mainly produced by *Serratia marcescens*. CpxR protein is an OmpR family transcriptional regulator in Gram-negative bacteria. Firstly, it was found that insertion mutation of *cpxR* in *S. marcescens* JNB 5-1 by a transposon Tn5G increased the production of PG. Results from the electrophoretic mobility shift assay (EMSA) indicated that CpxR could bind to the promoter of the *pig* gene cluster and repress the transcription levels of genes involved in PG biosynthesis in *S. marcescens* JNB 5-1. In the Δ*cpxR* mutant strain, the transcription levels of the *pig* gene cluster and the genes involved in the pathways of PG precursors, such as proline, pyruvate, serine, methionine, and S-adenosyl methionine, were significantly increased, hence promoting the production of PG. Subsequently, a fusion segment composed of the genes *proC*, *serC*, and *metH*, responsible for proline, serine, and methionine, was inserted into the *cpxR* gene in *S. marcescens* JNB 5-1. On fermentation by the resultant engineered *S. marcescens*, the highest PG titer reached 5.83 g/L and increased by 41.9%, relative to the parental strain. In this study, we revealed the role of CpxR in PG biosynthesis and provided an alternative strategy for the engineering of *S. marcescens* to enhance PG production.

Keywords: CpxR, *Serratia marcescens*, prodigiosin, temperature-sensitive, metabolic engineering

INTRODUCTION

Prodigiosin (PG), a red-colored tripyrrole, has gained increased interest because of its immunosuppressive, anticancer activity. It also has antifungal, antibacterial, antiprotozoal, and antimalarial properties; hence, it has a high potential for pharmaceutical application (Gastmeier, 2014; Gupta et al., 2014). Moreover, it is a good alternative to synthetic colorants and a promising source of food colorants. Prodigiosin is a typical secondary metabolite mainly produced by *Serratia marcescens* (Williamson et al., 2006; Papireddy et al., 2011), which usually appears in the later stages of bacterial growth (Williams, 1973). Incubation temperature plays an important role in PG synthesis in *S. marcescens*. PG could be efficiently produced at 28°C and sharply reduced at 37°C or higher (Williams et al., 1971; Tanaka et al., 2004).

The gene cluster for the biosynthesis of PG has been identified in *Serratia*, and most genes in the cluster have been functionally assigned. The bifurcated pathway culminates into two enzymatically condensed terminal products, 2-methyl-3-n-amyl-pyrrole (MAP) and 4-methoxy-2,2′-bipyrrole-5-carbaldehyde (MBC), and the *pig* clusters *pigD*, *pigE*, and *pigB* encode proteins that are involved in the production of MAP, while *pigA* and *pigF–pigN* encode proteins responsible for the synthesis of MBC (Grimont and Grimont, 1978; Von Graevenitz, 1980; Harris, 2004; Williamson et al., 2006). Proline is converted to 4-hydroxy-2,2′-bipyrrole-5-carbaldehyde (HBC) through a six-step reaction prior to methylation of the HBC hydroxyl group by S-adenosyl methionine-dependent PigF, resulting in the formation of MBC.

Previous studies have largely focused on quorum sensing (Matilla et al., 2015), signal transduction (Fineran et al., 2005), and two-component regulation systems (Horng et al., 2010) to determine the regulatory mechanisms which play an important role in the production of PG. But the thermoregulated mechanism of the production of this secondary metabolite is yet to be elucidated. Moreover, bacteria have evolved a wide variety of signal transduction systems in order to cope with changes in the external environment to ensure their survival and reproduction. This regulatory system has been divided into three categories according to the number of components involved: one-component regulatory system, two-component regulatory system (TCS), and three-component regulatory system (Marijuán et al., 2010). The *cpx* system is a canonical TCS, broadly conserved among Gram-negative bacteria, which consists of the sensor histidine kinase protein CpxA and the regulatory protein CpxR. Studies have suggested the role of CpxA in detecting changes in the external environment, including pH and the overexpression of envelope proteins (such as NlpE or pili subunits), as well as toxic concentrations of metal ions (Nakayama and Watanabe, 1995; Hunke et al., 2012; May et al., 2019). The detection of these stress signals activates CpxA, leading to autophosphorylation at the conserved 151 histidine residue. The histidine residue then carries a phosphate group and transfers it to the aspartic acid residue at position 51 of the response regulatory protein CpxR, hence activating CpxR (Yamamoto and Ishihama, 2006; MacRitchie et al., 2008). The phosphorylated CpxR finally binds to the specific sequence of the target gene promoter and initiates transcription. In this study, we firstly found that there was an increase in the proline, pyruvate, serine, and methionine biosynthetic pathway in the Δ*cpxR* mutant strain, which was beneficial to PG production. So, in this study, we discuss the effects of CpxR on prodigiosin production and further details of its mechanism function.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

S. marcescens JNB5-1 mutants derived from *S. marcescens* JNB 5-1 were grown in Luria–Bertani (LB) medium or fermentation medium. *Escherichia coli* BL21 (DE3) and *E. coli* DH5α, selected for the expression host, and the *E. coli S17-1 λpir* that pUT-Km replicated in were cultured in LB medium. The antibiotics used for selection in *E. coli* or *S. marcescens* JNB 5-1 were ampicillin (Amp), Kan, Chl, apramycin (Apm), clindamycin (Cli), and gentamicin (GM) at concentrations of 100, 50, 25, 50, 100, and 25 μg/ml, respectively. All incubations were performed either at 30 or 37°C. The strains in this work are listed in **Supplementary Table S1**.

Tn5G Transposon Mutagenesis

The method used for mutagenesis was the same as that in our previous research (Pan et al., 2019).

Measuring the Promoter of the *pig* Gene Cluster

Promoters of different lengths were amplified by PCR, and the plasmid pKK 232-8 (or pKKG) was linearized by double-enzyme digestion, followed by recombination of the linear DNA fragment obtained above by T4 ligase (Takara). Then, the recombinant plasmid pKK 232-8 (or pKKG) was transformed into *E. coli DH5α* competent cells, positive clones were selected, and successful construction was verified.

Measuring the Production of PG

The fermentation broth was dissolved in acidic ethanol (pH 3.0), moderately diluted (50-, 250-, and 1,000-fold), and the diluted sample was sealed and stored for 8 h (fully dissolved), centrifuged at 3,276 × g for 10 min, and the supernatant was taken. OD_{535} was determined. The blank control is acidic ethanol (Kalivoda et al., 2010).

Assay of PigF

A general procedure for measuring PigF activity was adopted from previous studies (Arnow, 1937; Dhar and Rosazza, 2000). The PigF assay buffer was made of 0.05 M Na_2HPO_4-KH_2PO_4 (pH 7.0). The assay mixture was composed of the assay buffer containing 1 mM dithiothreitol (DTT), 10 mM $MgCl_2$, 2 mM S-adenosyl methionine, 1 mM 4-hydroxy-2,20-bipyrrole-5-carbaldehyde, and 2.5 mg of the enzyme in a final volume of 1 ml. The enzyme reaction mixtures were incubated at 30°C for 25 min before being terminated by the addition of 1 ml of 0.5 M HCl, followed by the addition of 1 ml each of 10% $NaNO_2$, 10% $NaMoO_4$, and 1 M NaOH. The absorbance at 510 nm of this solution was immediately measured to determine the amount of substrates consumed in the reaction.

CAT and GFP Fluorescence Experiments

The activity of the chloramphenicol acetyltransferase (CAT) enzyme was measured using Shaw's method, with slight modifications (Shaw, 1975). The reaction system was made of 250 mmol L^{-1} Tris-HCl (pH 7.8), 0.1 mmol L^{-1} acetyl-CoA, 0.4 mg mL^{-1} DTNB, and an appropriate amount of crude enzyme solution. The above reaction mixture was incubated for 2 min in 37°C water, and chloramphenicol was added to a final concentration of 0.1 mmol L^{-1}. The light absorption value of A412 was measured immediately after mixing. The reaction solution without chloramphenicol was used as a control.

For the green fluorescent protein (GFP) fluorescence experiments, the BL21 (DE3)-containing plasmid with GFP was cultured in a 24-well plate at 37°C, 180 rpm, for 12 h. The

fluorescence intensity (the excitation/emission wavelengths were 480/510) was measured and the BL21 (DE3)-containing plasmid without GFP was used as a control.

Construction of Mutants

Allelic replacement of the gene open reading frame (ORF) was performed by plasmid pUT-Km, which is a suicide plasmid with a π-protein-dependent origin of replication from the R6K plasmid, so the plasmid can only be replicated in strains expressing the π protein, such as λpir-lysogenic *E. coli S17-1 λpir*, SM10(λpir) (de Lorenzo et al., 1990).

For gene knockout, the upstream (containing 9 bp of gene to be knocked out) and downstream (containing 9 bp of gene to be knocked out) homology arms of the gene to be knocked out and the resistance marker gene were amplified by PCR (**Supplementary Figure S1**). The plasmid pUT-Km is linearized by double-enzyme digestion and recombination of the linear DNA fragment obtained above using the ClonExpress II One Step Cloning Kit (Vazyme, Nanjing, China). Then, the recombinant plasmid pUT-Km was transformed into *E. coli S17-1 λpir* competent cells, positive clones were selected, and successful construction was verified.

For gene knock-in, the knockout method was slightly modified, with harvest and amplification of the upstream homology arms, downstream homology arms, and resistance genes, but should amplify a gene to be replaced in the knock-in method and then homologous recombination of the upstream homology arm, replacement gene, resistance marker gene, and the downstream homology arm. The recombinant plasmid pUT-Km was mobilized into *S. marcescens* JNB 5-1 by conjugation. Transconjugants grown on the LB plates contain both Apm (50 μg/ml) and Cli (100 μg/ml) (Maseda et al., 2009). Mutant candidates were screened by colony PCR using primer pairs, which are listed in **Supplementary Table S3** (Zhang et al., 2010). The construction of mutants that replaced the *cpxR* loci in *S. marcescens* JNB5-1 was done as described above, and the primers used are listed in **Supplementary Table S3**. The recombinant strain was obtained by screening on the agar plates containing Apm and Chl, with the results showing that the gene was successfully inserted into the gene of *S. marcescens* JNB 5-1.

RNA Extraction and Quantitative Real-Time PCR

RNA was prepared from cultures with stationary-phase cultures in LB medium at 30 or 37°C, respectively. Briefly, total RNA was treated with RNase-free pipette tips and RNA was purified using a FastPure Cell/Tissue Total RNA Isolation Kit (Vazyme, Nanjing, China). Of the total RNA, 1 μg was subjected to reverse transcription with 4 μl 5× HiScript II qRT SuperMix II(Vazyme, Nanjing, China), and the cDNA was subjected to real-time quantitative PCR analysis using the AceQ qPCR SYBR Green Master Mix (Vazyme, Nanjing, China), with the indicated primer pairs listed in **Supplementary Table S1**. 16S rRNA was used as the internal reference gene and monitored in real time with the StepOnePlus PCR system (Applied Biosystems).

Electrophoretic Mobility Shift Assay

You's method was slightly modified (You et al., 2018). The *cpxR* gene was amplified with the indicated primers (listed in **Supplementary Table S3**), cloned into plasmid pET-28a for expression with 6× His-tag, and then purified. The *pig* gene promoter and target gene ATG front sequences (possible promoter regions) were amplified with the indicated primers (listed in **Supplementary Table S3**) and ligated to the plasmid pMD19T (Simple) (listed in **Supplementary Table S2**). Subsequently, the promoter was amplified with the fluorescent-labeled primers M13-47-cy3 and RV-M-Cy3. The purified protein and the Cy3-labeled promoter were incubated in 2× binding buffer (40 mM Tris-HCl, pH 7.5), 4 mM $MgCl_2$, 100 mM NaCl, 10% glycerol, 2 mM DTT, 0.2 mg/ml bovine serum albumin (BSA), 0.02 mg/ml poly(dI-dC), and 1 mM EDTA for 30 min at 25°C. The incubated mixture was electrophoresed on a 5% native polyacrylamide gel electrophoresis (PAGE) for about 1 or 1.5 h. The gel was visualized with ImageQuant LAS 4000 (GE Healthcare Life Sciences, USA).

Shake Flask Fermentation Assay

The method with our previous research (Pan et al., 2019), with slight modifications. Briefly, JNB5-1 and SMCR were grown at 30°C and 37°C in a rotary shaker with fermentation medium (Sucrose 2%, Beef extract 1.5%, $CaCl_2$ 1%, L-proline 0.75%, $MgSO_4{\cdot}7H_2O$ 0.02%, and $FeSO_4{\cdot}7H_2O$ 0.006%) at 180 rpm for 120 h. The production of PG were determined at different fermentation time intervals (0, 12, 24, 36, 48, 60, 72, 84, 96, and 120 h) in triplicate.

Statistical Analysis

Experiments were done at least twice, with a minimum of three biological replicates. Origin software and TB tools were used for the drawing of the figures, with Student's *t*-tests, Mann–Whitney *U*-test, Fisher's exact test, and one-way ANOVA with Tukey's posttest. Significance was set at a *P*-value of <0.05

RESULTS

PG Production Could Be Enhanced by Disruption of CpxR in *S. marcescens* JNB 5-1

A Tn5G transposon was employed to identify genes, particularly those that influence PG production in *S.marcescens* JNB 5-1. The results from inverse PCR and sequencing of a mutant strain with a significantly increased PG production revealed that transposon insertion into the 38-bp position of a gene in JNB 5-1 could enhance PG production, and the selected gene was later mapped to the BVG90_20710 gene encoding a DNA-binding response regulator on the *S. marcescens* strain UMH8 genome (**Figure 1A**). Furthermore, bioinformatics analysis predicted that the selected gene from *S. marcescens* JNB 5-1 encodes the two-component regulatory system DNA-binding transcriptional regulator CpxR with a homology of 100% (699/699) to *cpxR* of the *S. marcescens* strain UMH8. At the protein level, the selected gene had a 91.38, 90.95, and 91.38% similarity to the *cpxR* of *E. coli* K12, *Salmonella enterica*, and *Yersinia pestis*, respectively

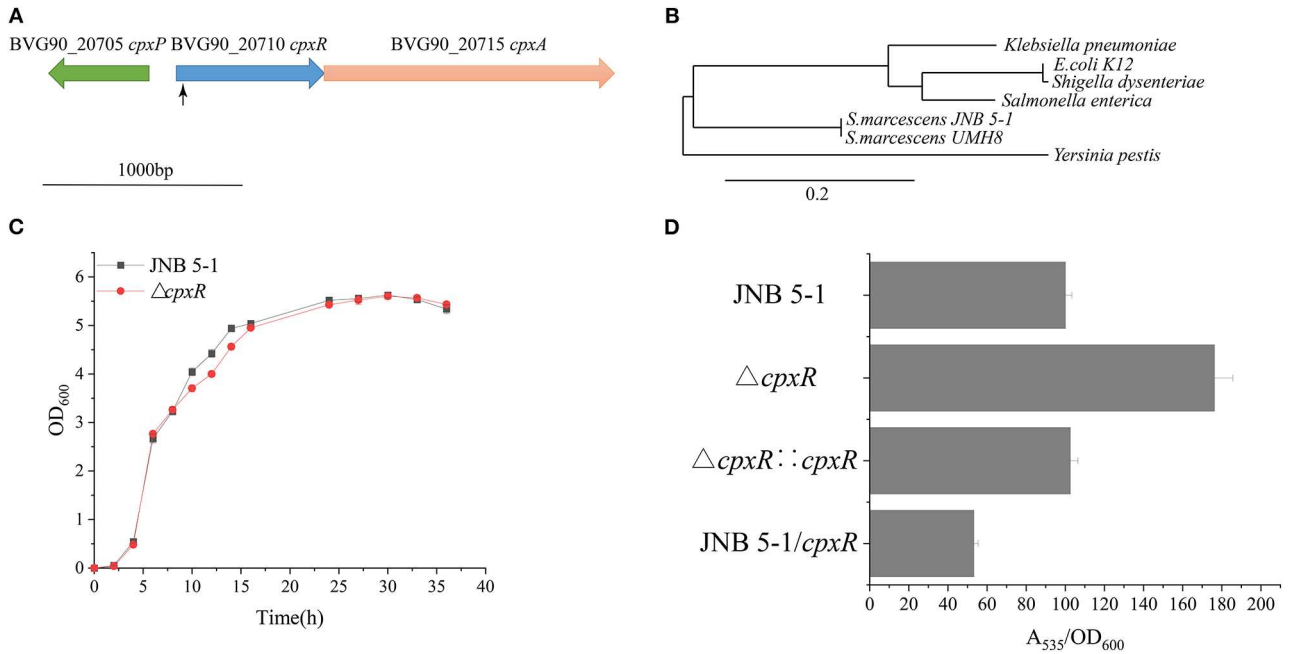


FIGURE 1 | Isolation of increased prodigiosin (PG) production in the Δ*cpxR* mutant. **(A)** Relative positions of *cpxA*, *cpxR*, and *cpxP* in the CpxARP system in the *Serratia marcescens* UMH8 strain. The locus inserted by the transposon is marked with a *black arrow*. **(B)** Evolutionary relationship with several known proteins among the Enterobacteriaceae. The Phylogram (Dereeper et al., 2008) was drawn from www.Phylogeny.fr in the mode of "one-click" using default setting. The number of substitutions per site is proportional to the length of the branch. **(C)** Growth pattern of JNB 5-1 and the Δ*cpxR* mutant in Luria–Bertani (LB) medium. **(D)** PG yield in JNB 5-1, Δ*cpxR* mutant, complementary strain, and overexpression strain. All the yield displayed by A_{535}/OD_{600} normalized to JNB 5-1 cultivated at 30°C; the mean ± SD from three independent experiments are shown.

(**Supplementary Figure S2**). Phylogenetic analysis suggests that the BVG90_20710-encoded protein is structurally intermediate in relatedness between *Y. pestis* and the highly similar *Klebsiella pneumoniae* (**Figure 1B**).

The Δ*cpxR* mutant displayed a similar growth pattern to that of JNB 5-1 when cultivated in LB medium, indicating that the deletion of *cpxR* had no significant effect on the growth of JNB 5-1 (**Figure 1C**). Additionally, there was a 76% increase in PG production observed in the Δ*cpxR* mutant compared to *S. marcescens* JNB 5-1, and the complementary strain had restored the level of PG similar to that of JNB 5-1 (**Figure 1D**). The Δ*cpxR* mutant complementary analysis using the *cpxR* on the JNB 5-1 chromosome further supported the observed phenotypic changes that increased PG production. To confirm that the phenotypic change of JNB 5-1 was due to the loss of *cpxR* or effect on the adjacent genes, the in-frame knockout of *cpxR* was introduced into the JNB 5-1 chromosome by plasmid pUT mini-Km, and the resulting mutant had a similar phenotype to the transposon insertion mutant, resulting in increased PG production. These results indicated that the increase in PG production was due to the mutation of *cpxR* rather than to another mutation or effects on adjacent genes. The Δ*cpxR* mutant phenotypes were also complemented by *cpxR* from *E. coli* K12, *S. enterica*, and *Y. pestis*, with the resulting phenotype showing consistent PG production and growth to that complemented by *cpxR* from JNB 5-1. These results suggest that *cpxR* is highly conserved among the Enterobacteriaceae. Additionally, plasmid-borne overexpression of *cpxR* in the JNB 5-1 strain (with empty plasmid) showed a 46.8% decrease in PG production, indicating the inhibitory effect of *cpxR* (**Figure 1D**).

CpxR Binds to the *pig* Gene Cluster Promoter

To further investigate the effect of CpxR on PG production in JNB 5-1, we analyzed the transcription levels of the *pig* gene cluster genes from JNB 5-1 and the Δ*cpxR* mutant by quantitative real-time PCR (qRT-PCR). The results showed that the transcription levels of the *pig* gene cluster genes in the Δ*cpxR* mutant were upregulated (**Figure 4**). This was consistent with the previously discovered phenomenon, which further confirmed that the yield change of PG was influenced by CpxR. We hypothesized that CpxR affected the genes' expressions on the *pig* gene cluster by possibly binding to the *pig* gene cluster promoter.

A promoter plays a critical role in providing transcriptional capacity for nearly every natural and synthetic circuit or pathway (Redden and Alper, 2015). To gain more understanding of the mechanism that CpxR regulates prodigiosin biosynthesis, we constructed a series of recombinant plasmids with different promoter lengths in the putative promoter region of *pigA*. Based on the analysis of the CAT and GFP fluorescence experiments and the Virtual Footprint online software (Munch et al., 2005), the sequence between −263 and the start codon before *pigA* was identified as the promoter of the *pig*

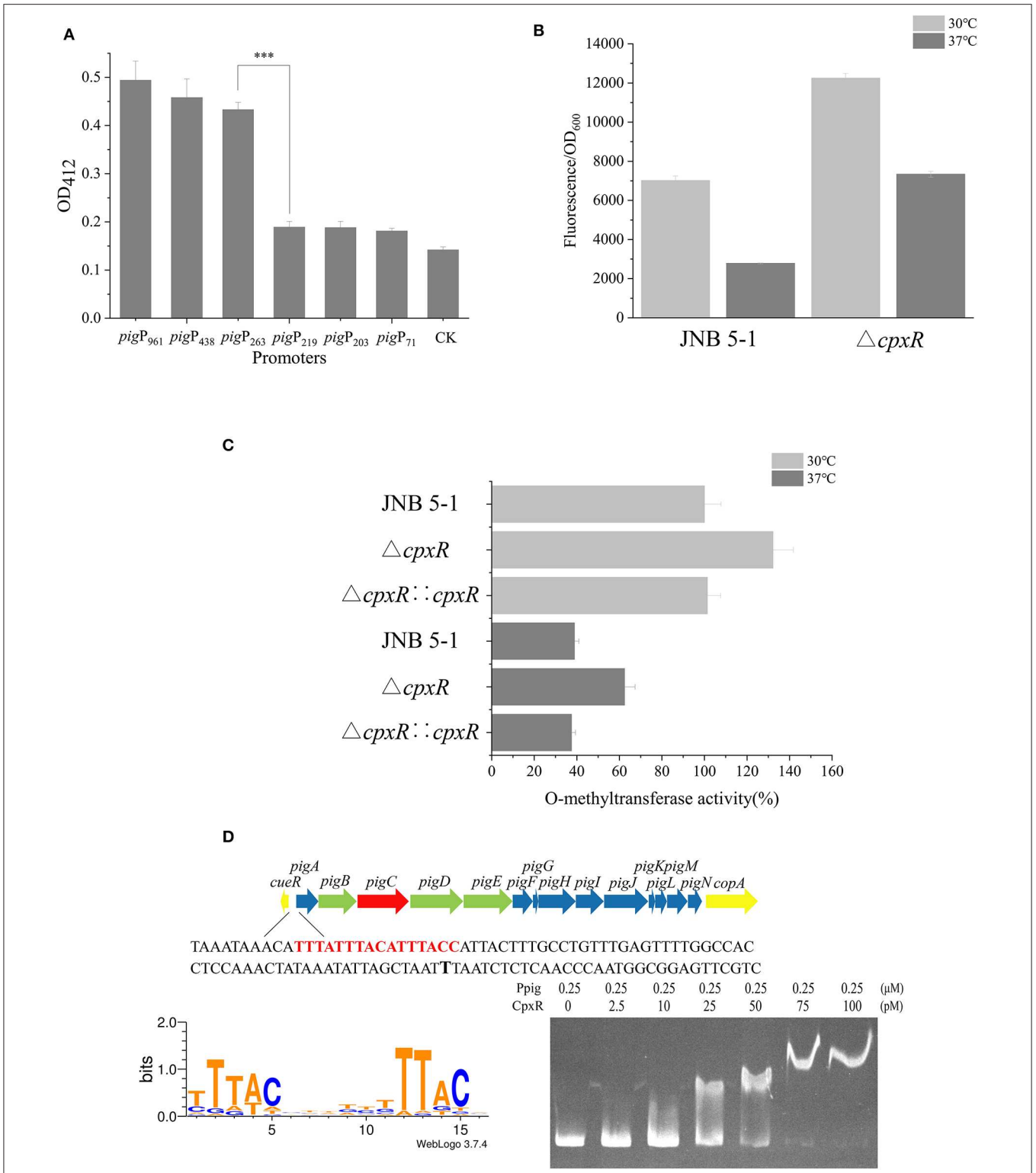


FIGURE 2 | CpxR interaction with the *pig* gene cluster promoter region in JNB 5-1. **(A)** Characterization of promoter activity at different lengths in chloramphenicol acetyltransferase (CAT) activity. **(B)** Green fluorescent protein (GFP) fluorescence intensity of Ppig-*gfp* in JNB 5-1 and the Δ*cpxR* mutant at 30 and 37°C. All the results displayed as fluorescence intensity/OD_{600}. **(C)** Whole *O*-methyltransferase activity from the crude cell extract derived from wild-type (with empty plasmid), *cpxR* mutant, and *cpxR* complement strain at 30 and 37°C. All the results displayed as percentage normalized to JNB 5-1 cultivated at 30°C. **(D)** CpxR bound to the *pig* promoter. The *top picture* is the *pig* gene cluster of the 16 genes in *S. marcescens* involved in the prodigiosin biosynthetic pathway. The *yellow block arrow on both*

(Continued)

FIGURE 2 | *ends* indicate the genes that are involved in copper ion transport. The 14 genes in the *middle* are involved in the biosynthesis of prodigiosin, in which the *green block arrow* indicates the genes encoding the protein that synthesizes MAP. *Blue block arrow* indicates the gene encoding the protein that synthesizes MBC, and *red block arrow* shows genes encoding the condensing enzyme in the condensation of MAP and MBC to form prodigiosin. The shown sequence is the front sequence of *pigA* (pig promoter), the *red region* is the CpxR binding region online, and the *black bold T* is the −10 region. The *bottom left picture* is the motif of CpxR harvested online, and *bottom right picture* is the result of the electrophoretic mobility shift assay (EMSA) wherein CpxR bound to the *pig* promoter. The mean ± SD from three independent experiments are shown. Significant differences in activity are calculated by one-way ANOVA. ***$P < 0.001$.

gene cluster used in subsequent experiments (**Figure 2A** and **Supplementary Figure S3**).

To investigate the effect of CpxR on the transcriptional level of the *pig* gene cluster in JNB 5-1, we constructed a recombinant plasmid carrying the fusion reporter gene (P*pig-gfp*, a *pig* gene cluster promoter with *gfp*), as shown in **Figure 2B**. Evidently, the Δ*cpxR* mutant possessed a higher fluorescence intensity relative to JNB 5-1 at 30 and 37°C, respectively, while the fluorescence intensity of GFP at 30°C was higher than that at 37°C in both the Δ*cpxR* mutant and JNB 5-1, which is consistent with our hypothesis. Specifically, we found that the Δ*cpxR* mutant was 74.6 and 163.5% higher than JNB 5-1 at 30 and 37°C, respectively. However, in the JNB 5-1 and Δ*cpxR* mutants, the fluorescence intensity of GFP at 30°C was 151.9 and 66.9% higher than that at 37°C, respectively. Taken together, these results suggested that temperature had an effect on *cpxR* and modulated the *pig* gene cluster promoter (**Figure 2B**).

Furthermore, a key *pigF* gene from the *pig* gene cluster encoding *O*-methyltransferase (PigF) was carried out to verify the effect of CpxR on PG biosynthesis in JNB 5-1 at 30 and 37°C. The results showed that *O*-methyltransferase activity was the highest in the Δ*cpxR* mutant compared with JNB 5-1 and the *cpxR* complementary strain at 30 and 37°C (**Figure 2C**), implying the significant effect of CpxR on *pigF* or the *pig* gene cluster.

EMSA was performed to further gain insight into the regulation mechanism of CpxR on the *pig* cluster promoter. As shown in **Figure 2D**, the binding affinity increased with an increase in the amount of CpxR protein, suggesting that CpxR bound to the *pig* gene cluster promoter. Interestingly, there was an oblique upward gradient in the binding affinity of CpxR even when the amount of *pig* promoter added is maintained at 0.25 μM. Online analysis on Regulon DB (Santos-Zavaleta et al., 2019) predicted that the CpxR motif is "TTTACNNNNNTTTACN" and binding site was upstream of the *pigA* "TTTATTTACATTTACC" in JNB 5-1 (**Figure 2D**). The CpxR motif was highly conserved in the *pig* gene cluster promoter when mapping it to the annotated *S. marcescens* (**Supplementary Figure S4**).

Inhibition of PG Production by the *cpx* System Corresponds to Temperature in *S. marcescens* JNB 5-1

Considering the impact of temperature on PG biosynthesis, we investigated the relationship between the *cpx* system and temperature. A qRT-PCR assay was performed on *cpx* for further analysis. As shown in **Figure 3A**, *cpxR* and *cpxA* were upregulated, with 3.72- and 5.44-fold increases at the transcription level. Subsequently, dynamic transcription-level analysis of the *cpx* system between 30 and 37°C in JNB 5-1 showed a significantly increased transcription level of *cpxA* at 37°C (**Figures 3A,B**). Initially, there was no change in the transcription level of *cpxR* (0–6 h), but then an increase in the transcription level was observed at the exponential phase (6–12 h), as shown in **Figure 3B**. This may be attributed to the activation of *cpxR* by CpxA in response to environmental stress. Both *cpxA* and *cpxR* represented a downtrend when the temperature was lowered to 30°C. Subsequently, after 12 h of JNB 5-1 cultivation, the temperatures were altered from 30 to 37°C and *vice versa*, and then dynamic transcription-level tests on *cpxA* and *cpxR* were performed. The results indicated that the transcription level of *cpxA* increased 3.95- and 5.91-fold when cultivated at 37°C after 12 and 24 h, respectively, and decreased by 56.2 and 35.8% when cultivated at 30°C, respectively. A similar trend was observed in *cpxR*, where the transcription level increased by 2.69- and 4.46-fold when turned to 37°C, respectively, and decreased by 46.9 and 34.8% when lowered to 30°C, respectively (**Figure 3B**). These results suggested that the *cpx* system might affect PG production in the *S. marcescens* JNB 5-1 strain in response to temperature.

Subsequently, we examined the transcript level of the *cpx* system regulatory factor CpxP, which could be activated by CpxR transcription and also acts as an accessory protein to regulate the *cpx* system in the negative feedback loop, to further confirm the relationship between the *cpx* system and temperature. A significant difference in the transcription level of *cpxP* from JNB 5-1 was observed between 30 and 37°C, with the transcription level at 37°C 15-fold higher than that at 30°C. However, the transcription level of *cpxP* in the Δ*cpxR* mutant decreased to 11% when compared to JNB 5-1 cultivated at 30°C, and the transcription level of *cpxP* in the Δ*cpxA* mutant increased by 16.47-fold (**Figure 3A**). Furthermore, a similar trend to *cpxR* where the transcription level changed in response to temperature was observed, shown in **Figure 3A**. Taken together, these results suggested that CpxAR-dependent auxiliary protein CpxP expression is temperature-dependent, indicating that there is a substantial impact on CpxAR since the *cpx* system could be activated at 37°C (**Figure 3A**).

Δ*cpxR* Mutant Promotes the Supply of PG Synthesis Precursors

To gain insight on the effect of CpxR in JNB 5-1 and the changes in the transcription levels of genes on the metabolic pathways involved in PG synthesis in the Δ*cpxR* mutant, we performed qPCR analysis on the genes related to the known precursors for PG synthesis (proline, serine, and L-methionine synthesis) in *S. marcescens*. We found that the transcription levels of the precursors involved in PG synthesis and energy genes

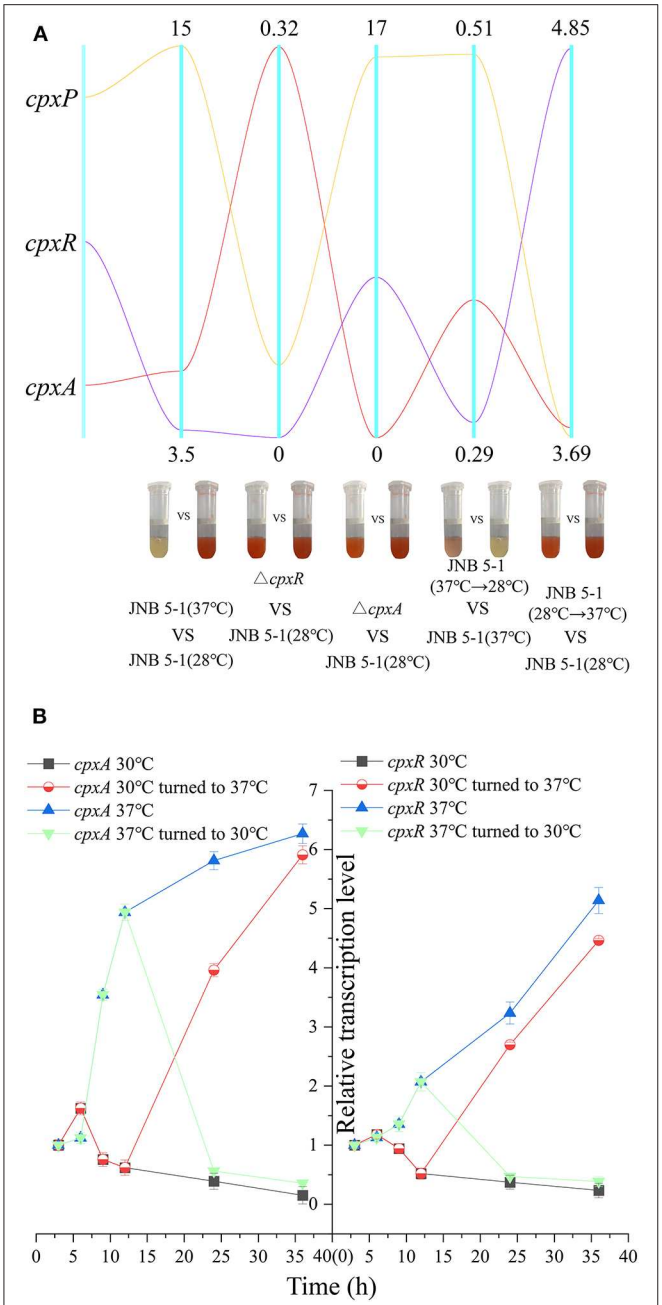


FIGURE 3 | Dynamic analysis of the *cpx* system in response to temperature. **(A)** Dynamic analysis of the *pig* gene cluster and the *cpx* system. The five *vertical line data* represent the comparisons of JNB 5-1 between 30 and 37°C, of the *cpxR* mutant and JNB 5-1 at 30°C, between the *cpxA* mutant and JNB 5-1 at 30°C, between the JNB 5-1 transferred from 37 to 30°C and JNB 5-1 cultivated at 37°C, and between the JNB 5-1 transferred from 30 to 37°C and JNB 5-1 cultivated at 30°C. *Yellow line* represents gene *cpxP*, *purple line* represents gene *cpxR*, and *red line* represents gene *cpxA*. The *tubes* are the color change of the each strain (meaning change in PG yield). **(B)** Dynamic transcription level analysis of *cpxA* and *cpxR*. qPCR is performed for the characterization of dynamic change of transcription in *cpxA* and *cpxR* in JNB 5-1. The temperatures were altered from 30 to 37°C after 12 h of JNB 5-1 cultivation and *vice versa*. All the values of qRT-PCR were normalized to the first sample (3 h).

were improved. The results showed that the transcript levels of glutamate-5-semialdehyde dehydrogenase (*proA*), glutamate 5-kinase (*proB*), pyrroline-5-carboxylate reductase (*proC*), and proline iminopeptidase (*pip*) were upregulated 4.97-, 5.43-, 6.57-, and 3.23-fold, respectively, in the Δ*cpxR* mutant compared to JNB 5-1, indicating that proline biosynthesis was enhanced in the Δ*cpxR* mutant (**Figure 4**). Furthermore, the transcription levels of phosphoserine aminotransferase (*serC*) and phosphoserine phosphatase (*serB*), related to serine synthesis, were also upregulated 4.55- and 3.62-fold, respectively, indicating that serine biosynthesis was promoted in the Δ*cpxR* mutant. qRT-PCR analysis also showed that methionine synthase (*metH*) and S-adenosyl methionine decarboxylase (*speD*) were upregulated 4.81- and 2.74-fold, respectively, showing that S-adenosyl methionine, required for the transformation of HBC to MBC, increased in the Δ*cpxR* mutant (**Figure 4**).

Moreover, the transcription levels of 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase (*gpmA*), glyceraldehyde 3-phosphate dehydrogenase (*gapA*), and pyruvate kinase (*pykF*) related to pyruvate synthesis were upregulated 4.21-, 5.14-, and 4.67-fold, respectively. The transcript level of aldehyde dehydrogenase gene (*aldB*), related to acetaldehyde synthesis, was also upregulated 5.83-fold. Acetaldehyde metabolism is essential for the supply of acetyl-CoA, which could be catalyzed to malonyl-CoA that is beneficial for PG synthesis (**Figure 4**).

Reconstruction of *S. marcescens* JNB 5-1 Strain for Efficient Production of Prodigiosin

Proline, serine, and methionine are important factors that promote PG production, considering thiamine and sodium acetate are also beneficial for PG biosynthesis. Therefore, proline, serine, methionine, thiamine, sodium acetate, and a proline–methionine combination were added into the medium for JNB 5-1 growth. The combination of proline–methionine showed the highest PG production relative to the control, and the use of proline or methionine alone also increased PG production (**Figure 5B**). Subsequently, we performed the plasmid-borne *proA*, *proB*, *proC*, *serB*, *serC*, and *metH* overexpression on strain JNB 5-1. As anticipated, plasmid-borne *proC*, *serC*, and *metH* overexpression in strain JNB 5-1 increased by 52.3%, 33.7 and 37.5%, respectively, compared with JNB 5-1 in LB medium harboring empty plasmid (**Figure 5C**).

Further, in order to overcome the low level of single-copy gene expression by homologous recombination, we integrated the *proC*, *serC*, and *metH* into the *cpxR* loci (**Figure 5A**). The titer of PG increased 79% in the recombinant strains with *proC*, *serC*, and *metH* integrated into the *cpxR* loci (the obtained recombinant strain was renamed SMCH) compared with JNB 5-1 in LB medium (**Figure 5C**). To fully substantiate the roles of *proC*, *serC*, and *metH* and *cpxR*, SMCH was subjected to shake flask fermentation for the production of PG in the fermentation medium. Overall, the PG production in SMCH was significantly higher than that of JNB 5-1 at both 30 and 37°C. Specifically, at 30°C, SMCH grew slowly compared to JNB 5-1 at the beginning

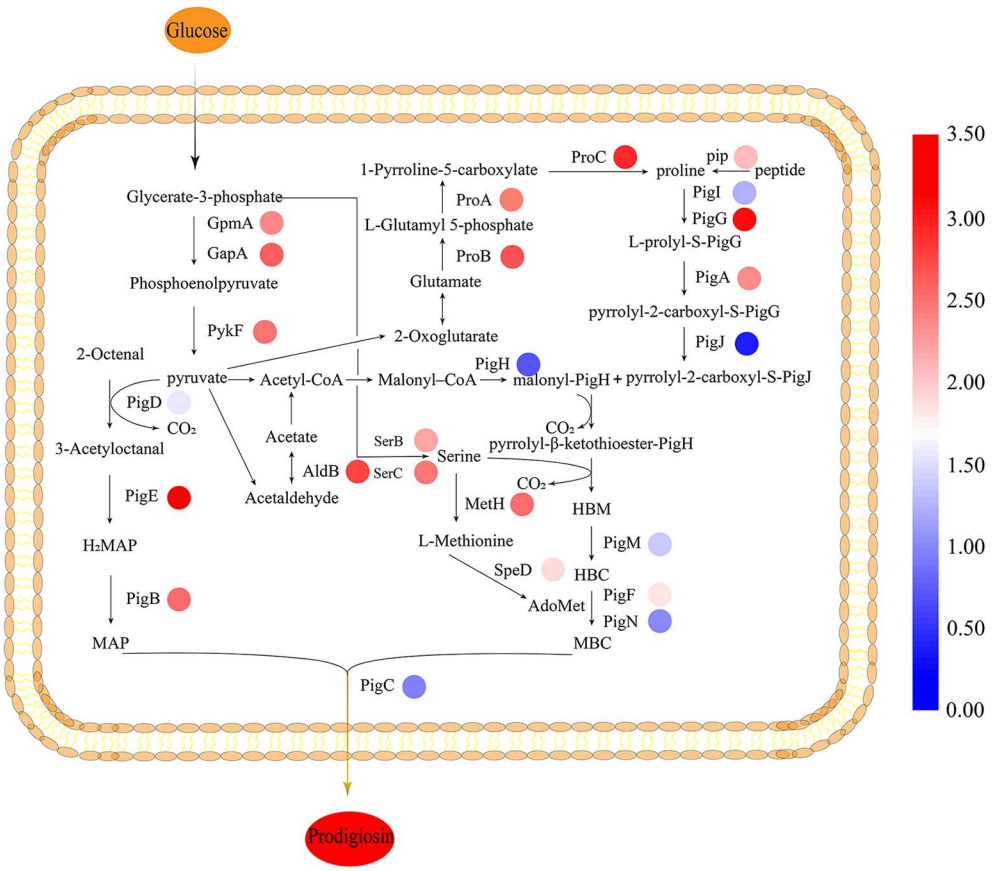


FIGURE 4 | Biosynthesis pathway of prodigiosin in JNB 5-1. All the relative transcription-level values were $\log_2$-transformed. The biosynthesis of prodigiosin in *Serratia marcescens* is synthesized by a bifurcated pathway of MAP and MBC condensation to form prodigiosin by the condensing enzyme PigC. The *colored rings beside the genes* represent the transcription level, shown as $\log_2$ (fold change). *MAP*, 2-methyl-3-n-amyl-pyrrole; *MBC*, 4-methoxy-2,2′-bipyrrole-5-carbaldehyde; *AdoMet*, S-adenosyl methionine; *HBC*, 4-hydroxy-2,2′-bipyrrole-5-carbaldehyde; *HBM*, 4-hydroxy-2,2′-bipyrrole-5-methanol. JNB 5-1 and the *cpxR* mutant were cultivated at 30°C.

of growth, and there was no difference between the two strains at the logarithmic phase, as shown in the **Figures 5D,E**. JNB 5-1 entered the death phase after 84 h, with the stationary period commencing in SMCH and the death phase later beginning at 120 h. Significantly, the highest yield of PG produced in SMCH was 5.83 g/L, which was 41.9% higher than that in JNB 5-1. When cultured at 37°C, there was no significant difference in the growth curve of SMCH and JNB 5-1 from 0 to 84 h, when JNB 5-1 entered the stationary phase. Specific to PG production, SMCH reached the highest yield of 590 mg/L, which was 64.1% higher than JNB 5-1. The results showed that the integration of *proC*, *serC*, and *metH* into the *cpxR* gene resulted in an increase in the expressions of *proC*, *serC*, and *metH*, thus improving the production of prodigiosin in *S. marcescens* JNB 5-1.

DISCUSSION

We herein show a novel insight into the biosynthesis of PG with regulator response to temperature (RRT), in which a two-component regulatory system was thermoregulated and the regulatory protein can bind to the *pig* gene cluster promoter. The Cpx two-component regulatory system that inhibits PG production was screened by transposon insertion, which was thermoregulated and transferred this effect to PG biosynthesis. Subsequently, we reconstructed the key gene of the proline, serine, and methionine metabolic pathway by knock-in and knockout, in which the copy number between *proC* and the Cpx two-component regulatory system changed. This strategy can be used in metabolically engineered strains to improve the copy numbers of *proC*, *serC*, and *metH*. Lastly, we achieved a hyperproducing *S. marcescens* through the entire genome-modified metabolic strain.

There have been many reports on the effects of the production of PG. Kim et al. improved the prodigiosin production from 0.658 to 1.628 g/L by an antibiotic mutagenesis using chloramphenicol (Kim et al., 2008). Haddix and colleagues harvested the highest production of 1.314 g/L by modifying various fermentation parameters (Haddix and Shanks, 2020); however, the regulatory network may play an important role. The two-component regulation system is an expression system established by bacteria to cope with changes in the external environment (Cheung and Hendrickson, 2010). Previous reports on the Cpx

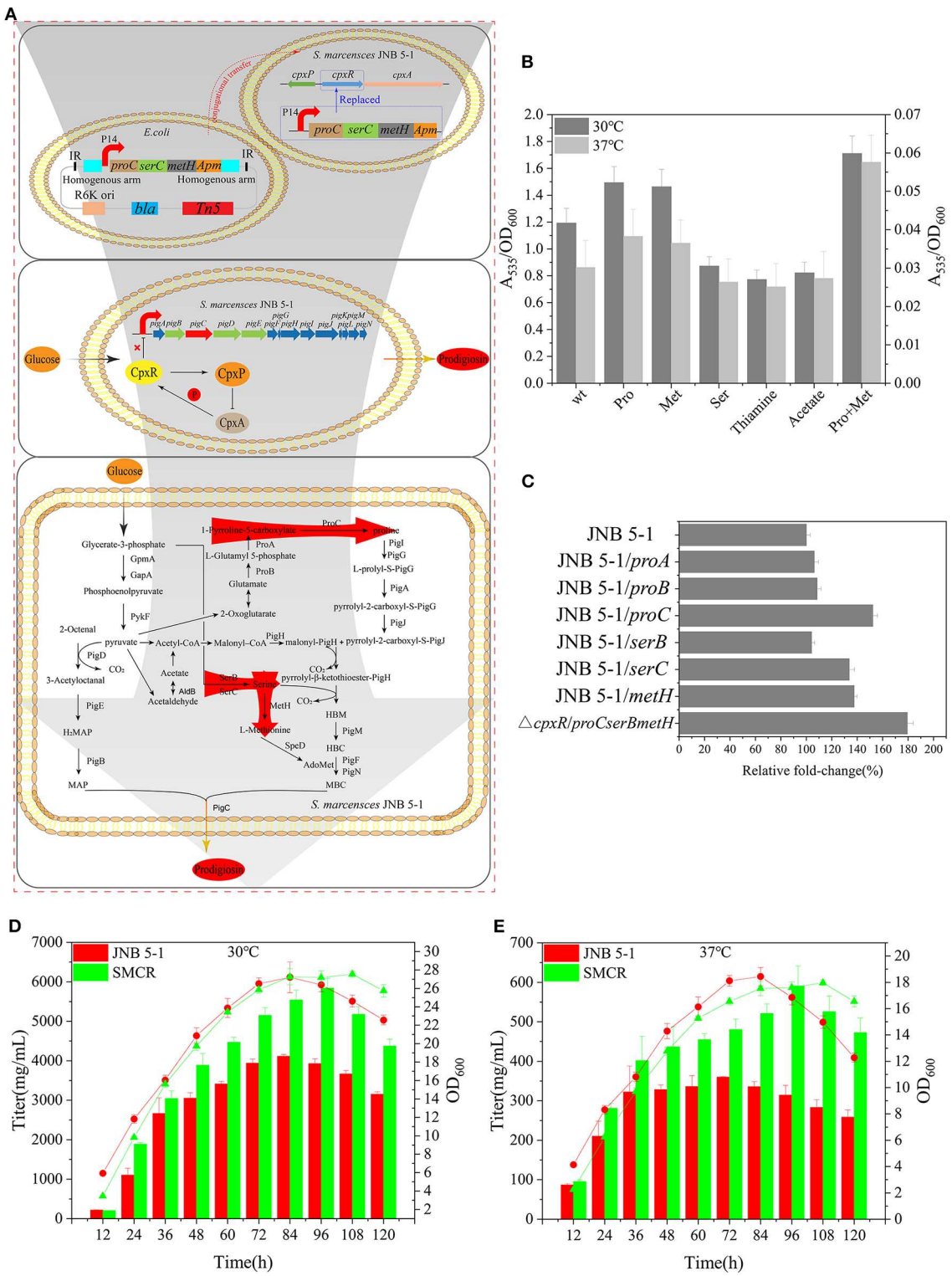


FIGURE 5 | Reconstruction of *Serratia marcescens* JNB 5-1 and fermentation. **(A)** Workflow of the reconstruction of *S. marcescens* JNB 5-1. The *top picture* displays the reconstruction of *S. marcescens* JNB 5-1, replacing *cpxR* with *proC*, *serC*, and *metH*. The pUT plasmid was used as a knock-in plasmid to send the genes to the JNB 5-1 chromosome by conjugative transfer into JNB 5-1. Subsequently, homologous recombination after entering the cell (*blue box*). The *middle picture* represents the relationship between CpxA, CpxR, and CpxP and the regulation of CpxR on the *pig* gene cluster. The picture displays the biosynthesis of prodigiosin in the hyperproducing *S. marcescens*, in which *cpxR* was replaced by *proC*, *serC*, and *metH*. *Red arrow region* is the enhanced pathway by

(Continued)

FIGURE 5 | overexpression of *proC*, *serC*, and *metH*. **(B)** Addition of different precursors cultivated in Luria–Bertani (LB) medium at 30 and 37°C. All the yield displayed as A_{535}/OD_{600}. *Left y*-axils and *right y*-axils represent A_{535}/OD_{600} from strains cultivated at 30 and 37°C, respectively. **(C)** Plasmid-borne overexpression of *proA*, *proB*, *proC*, *serB*, *serC*, and *metH* in JNB 5-1 in LB cultivated at 30°C. The last result displayed the promotion of mutant replacing the *cpxR* loci by *proC*, *serC*, and *metH* compared with JNB 5-1. All the yield displayed as A_{535}/OD_{600} normalized to JNB 5-1 cultivated at 30°C. **(D,E)** Fermentation of mutants and JNB 5-1 cultivated at 30 and 37°C, respectively. *Lines* and *columns* represent OD_{600} and titer, respectively. The means ± SD from three independent experiments are shown.

two-component regulatory system have shown that CpxR could regulate the *degP*, *dsbA*, and *ppiA* genes that encode heat shock protease, disulfide oxidoreductase, and peptide prolyl isomerase, respectively (De Wulf et al., 1999). Our study correlated the Cpx two-component regulatory system to temperature regulation and PG production. The deletion of CpxR resulted in an increase of PG production in JNB 5-1, but did not affect JNB 5-1 growth. However, the overexpression of *cpxR* not only reduced the ability to synthesize PG in JNB 5-1 but also affected the growth of JNB 5-1. Spinola et al. (2010) found that the deletion of *cpxA* led to the accumulation of CpxR in *Haemophilus ducreyi* and weakened its survival (viability) *in vivo*. A similar result was observed, as shown in **Figure 3A**, where the transcription level of *cpxR* was upregulated in the Δ*cpxA* mutant. Thus, the overexpression of *cpxR* and the deletion of *cpxA* in JNB 5-1 showed that the growth of JNB 5-1 was influenced. García Véscovi Eleonora et al. also found that CpxR expression is temperature-dependent by changes in the *cpxP* transcriptional levels (Bruna et al., 2018). However, we observed that the *cpx* system changed with temperature change (**Figure 3A**). The transcription levels of the genes in the *pig* gene cluster also changed in *S. marcescens* JNB5-1 and in the Δ*cpxR* and Δ*cpxA* mutants. Furthermore, it was confirmed by EMSA that CpxR bound to the *pig* gene cluster promoter. However, we also noticed that PG production increased in the Δ*cpxR* mutant, but decreased in the Δ*cpxA* mutant, and this could be attributed to the influence of *cpxA* on the growth of *S. marcescens* JNB 5-1 (Delhaye et al., 2016). Moreover, since CpxA is a protein that responds to changes in the surrounding cell membrane environment, such as pH and salt ion concentration, deletion of *cpxA* probably resulted in the inability of cells to adapt to changes in the external environment. We found that *S. marcescens* JNB 5-1 had a highest titer (4.11 g/L) of PG after 84 h, while SMCH produced 5.83 g/L of PG after 96 h. We hypothesized that this was related to the relieved product feedback inhibition due to *cpxR* knockout, which was in agreement with previous reports that CpxR is related to cell membrane biosynthesis (May et al., 2019; Simpson and Trent, 2019).

Proline, synthesized *via* the glutamate pathway by a series of enzymes (ProA, ProB, and ProC), is involved in the first step of MBC synthesis catalyzed by PigG and PigI. It activates L-proline using ATP, resulting in the interaction of L-proline with the thiol group in PigG to form a complex L-prolyl-S-PigG under the action of Pig I, and then PigA oxidizes the L-prolyl-S-PCP to pyrrolyl-2-carboxyl-S-PCP (Williamson et al., 2006). Our research showed that the transcription levels of *proA*, *proB*, and *proC* involved in proline biosynthesis increased to varying degrees in the Δ*cpxR* mutant, resulting in increased proline synthesis, which was beneficial to PG biosynthesis. This result is in line with a study by Siva et al. that reported enhanced PG production in *Serratia rubidaea* after the addition of proline alone or in combination with methionine (Siva et al., 2012). Moreover, an increased PG production was achieved by plasmid overexpression of *proC* and by adding proline to the fermentation medium in our study. Subsequently, the expression of *proC* was increased by inserting *proC* into the *cpxR* loci, hence a reconstructed *S. marcescens* JNB 5-1 strain with efficient production of prodigiosin.

The transcription levels of the genes involved in pyruvate, serine, and methionine metabolism were also upregulated. Pyruvate is involved in the first step of the MAP branch in the PG biosynthesis bifurcated pathway, which forms 3-acetyloctanal with 2-octenal under the catalysis of PigD and then generates MAP *via* the catalysis of PigE and PigB (Williamson et al., 2006). Wasserman et al. also confirmed the derivation of C2 of MAP, and its attached methyl group is from pyruvate by ^{14}C and ^{13}C labeling (Wasserman et al., 1973). Initially, we added pyruvate and thiamine to the medium; however, there was no significant increase in PG production. We therefore suspected that pyruvate could be involved in many reactions in the bacteria and was diverted by other metabolic pathways. On the other hand, the maximum rate-limiting factor for PG production could be in the MBC branch. Also, it was predicted that serine involved in the biosynthesis of MBC was condensed with the pyrrolyl-β-ketothioester of PigH, forming 4-hydroxy-2,2′-bipyrrole-5-methanol (HBM) (Williamson et al., 2006), which is an intermediate of HBC required for the synthesis of MBC *via* the action of S-adenosyl methionine-dependent PigF and PigN (Williamson et al., 2005). The replication of the plasmid containing individual genes (*proC*, *serC*, and *metH*) was carried out by means of circular replication to achieve higher copy numbers. However, the replication generated a large amount of single DNAs, making the plasmid unstable (Aoki et al., 1987; Zhang et al., 2019), thus limiting its industrial application. Therefore, we used integrated expression strategy to insert the gene of the key synthetic enzyme for PG into a certain position on the chromosome of *S. marcescens* JNB5-1, allowing the inserted gene to replicate as the host chromosome replicates. Although the integrated expression level was lower than the plasmid expression level, by increasing the copy number of the integrated *proC*, we were able to achieve high levels of target protein production. In addition, we knocked out a suppressor gene of *S. marcescens* JNB 5-1, resulting in a decreased inhibitory effect on prodigiosin synthesis in *S. marcescens* JNB 5-1 (Zhang et al., 2019). As anticipated, the integration of *proC*, *serC*, and *metH* into the *cpxR* loci enhanced the transcription and expression of *proC*, *serC*, and *metH* in *S. marcescens* JNB5-1.

In conclusion, our study demonstrated that the two-component regulatory system *cpx* in *S. marcescens* JNB 5-1 is temperature-regulated and inhibited PG biosynthesis. The proposed RRT system hinted that temperature drives the expression of the *cpx* system, further regulating PG synthesis. The reconstructed high copy numbers of *proC*, *serC*, and *metH* in *S. marcescens* JNB 5-1 by homologous recombination increased the PG yield.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

ZR and TY conceived and designed the study and critically revised the manuscript. TO critically revised the manuscript. YS carried out the experiments, analyzed the data, and drafted the manuscript. LW carried out the experiments. XP, HF, HZ, and S-TY contributed to the revision of the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Key Research and Development Program of China (2018YFA0900300), the National Natural Science Foundation of China (31870066, 21778024, 31570085), National First-Class Discipline Program of Light Industry Technology and Engineering (LITE2018-06), the Program of Introducing Talents of Discipline to Universities (111-2-06), key research and development program of Ningxia hui autonomous region (2017BY069), the science and technology innovation team foundation of Ningxia hui autonomous region (KJT2017001), Top-notch Academic Programs Project of Jiangsu Higher Education Institutions and the Priority Academic Program Development of Jiangsu Higher Education Institution.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00344/full#supplementary-material

REFERENCES

Aoki, T., Noguchi, N., Sasatsu, M., and Kono, M. (1987). Complete nucleotide sequence of pTZ12, a chloramphenicol-resistance plasmid of *Bacillus subtilis*. *Gene* 51, 107–111. doi: 10.1016/0378-1119(87)90481-1

Arnow, L. E. (1937). colorimetric determination of the components of 3,4-dihydroxyphenylalaninetyrosine mixtures. *J. Biol. Chem.* 118, 531–537.

Bruna, R. E., Molino, M. V., Lazzaro, M., Mariscotti, J. F., and García Véscovi, E. (2018). CpxR-dependent thermoregulation of *Serratia marcescens* PrtA metalloprotease expression and its contribution to bacterial biofilm formation. *J. Bacteriol.* 6:18. doi: 10.1128/JB.00006-18

Cheung, J., and Hendrickson, W. A. (2010). Sensor domains of two-component regulatory systems. *Curr. Opin. Microbiol.* 13, 116–123. doi: 10.1016/j.mib.2010.01.016

de Lorenzo, V., Herrero, M., Jakubzik, U., and Timmis, K. N. (1990). Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. *J. Bacteriol.* 172, 6568–6572. doi: 10.1128/JB.172.11.6568-6572.1990

De Wulf, P., Kwon, O., and Lin, E. C. (1999). The CpxRA signal transduction system of *Escherichia coli*: growth-related autoactivation and control of unanticipated target operons. *J. Bacteriol.* 181, 6772–6778. doi: 10.1128/JB.181.21.6772-6778.1999

Delhaye, A., Collet, J., and Laloux, G. (2016). Fine-tuning of the cpx envelope stress response is required for cell wall homeostasis in *Escherichia coli*. *mBio* 7, e16–e47. doi: 10.1128/mBio.00047-16

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., et al. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 36, W465–W469. doi: 10.1093/nar/gkn180

Dhar, K., and Rosazza, J. P. (2000). Purification and characterization of *Streptomyces griseus* catechol O-methyltransferase. *Appl. Environ. Microbiol.* 66, 4877–4882. doi: 10.1128/AEM.66.11.4877-4882.2000

Fineran, P. C., Slater, H., Everson, L., Hughes, K., and Salmond, G. P. C. (2005). Biosynthesis of tripyrrole and β-lactam secondary metabolites in *Serratia*: integration of quorum sensing with multiple new regulatory components in the control of prodigiosin and carbapenem antibiotic production. *Mol. Microbiol.* 56, 1495–1517. doi: 10.1111/j.1365-2958.2005.04660.x

Gastmeier, P. (2014). *Serratia marcescens*: an outbreak experience. *Front. Microbiol.* 5:81. doi: 10.3389/fmicb.2014.00081

Grimont, P. A., and Grimont, F. (1978). The genus *Serratia*. *Ann. Rev. Microbiol.* 32:221. doi: 10.1146/annurev.mi.32.100178.001253

Gupta, N., Hocevar, S. N., Moultonmeissner, H. A., Stevens, K. M., Mcintyre, M. G., Jensen, B., et al. (2014). Outbreak of *Serratia marcescens* bloodstream infections in patients receiving parenteral nutrition prepared by a compounding pharmacy. *Clin. Infect. Dis.* 59, 1–8. doi: 10.1093/cid/ciu218

Haddix, P. L., and Shanks, R. (2020). Production of prodigiosin pigment by Serratia marcescens is negatively associated with cellular ATP levels during high-rate, low-cell-density growth. *Can. J. Microbiol.* 66, 243–255. doi: 10.1139/cjm-2019-0548

Harris, A. K. P. (2004). The *Serratia* gene cluster encoding biosynthesis of the red antibiotic, prodigiosin, shows species- and strain-dependent genome context variation. *Microbiology* 150, 3547–3560. doi: 10.1099/mic.0.27222-0

Horng, Y., Chang, K., Liu, Y., Lai, H., and Soo, P. (2010). The RssB/RssA two-component system regulates biosynthesis of the tripyrrole antibiotic, prodigiosin, in *Serratia marcescens*. *Int. J. Med. Microbiol.* 300, 304–312. doi: 10.1016/j.ijmm.2010.01.003

Hunke, S., Keller, R., and Müller, V. S. (2012). Signal integration by the Cpx-envelope stress system. *FEMS Microbiol. Lett.* 326, 12–22. doi: 10.1111/j.1574-6968.2011.02436.x

Kalivoda, E. J., Stella, N. A., Aston, M. A., Fender, J. E., Thompson, P. P., Kowalski, R. P., et al. (2010). Cyclic AMP negatively regulates prodigiosin production by *Serratia marcescens*. *Res. Microbiol.* 161, 158–167. doi: 10.1016/j.resmic.2009.12.004

Kim, S. J., Lee, H. K., Lee, Y. K., and Yim, J. H. (2008). Mutant selection of *Hahella chejuensis* KCTC 2396 and statistical optimization of medium components for prodigiosin yield-up. *J. Microbiol.* 46, 183–188. doi: 10.1007/s12275-008-0037-y

MacRitchie, D. M., Buelow, D. R., Price, N. L., and Raivio, T. L. (2008). Two-component signaling and gram negative envelope stress response systems. *Adv. Exp. Med. Biol.* 631, 80–110. doi: 10.1007/978-0-387-78885-2_6

Marijuán, P. C., Navarro, J., and Del Moral, R. (2010). On prokaryotic intelligence: strategies for sensing the environment. *Biosystems* 99, 94–103. doi: 10.1016/j.biosystems.2009.09.004

Maseda, H., Hashida, Y., Konaka, R., Shirai, A., and Kourai, H. (2009). Mutational upregulation of a resistance-nodulation-cell division-type multidrug efflux pump, SdeAB, upon exposure to a biocide, cetylpyridinium chloride, and antibiotic resistance in *Serratia marcescens*. *Antimicrob. Agents Chemother.* 53, 5230–5235. doi: 10.1128/AAC.00631-09

Matilla, M. A., Leeper, F. J., and Salmond, G. P. C. (2015). Biosynthesis of the antifungal haterumalide, oocydin A, in *Serratia*, and its regulation by quorum sensing, RpoS and Hfq. *Environ. Microbiol.* 17, 2993–3008. doi: 10.1111/1462-2920.12839

May, K. L., Lehman, K. M., Mitchell, A. M., and Grabowicz, M. (2019). A stress response monitoring lipoprotein trafficking to the outer membrane. *mBio* 10:e00618-19. doi: 10.1128/mBio.00618-19

Munch, R., Hiller, K., Grote, A., Scheer, M., Klein, J., Schobert, M., et al. (2005). Virtual footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics* 21, 4187–4189. doi: 10.1093/bioinformatics/bti635

Nakayama, S., and Watanabe, H. (1995). Involvement of cpxA, a sensor of a two-component regulatory system, in the pH-dependent regulation of expression of *Shigella sonnei* virF gene. *J. Bacteriol.* 177, 5062–5069. doi: 10.1128/JB.177.17.5062-5069.1995

Pan, X., Sun, C., Tang, M., Liu, C., Zhang, J., You, J., et al. (2019). Loss of serine-type D-Ala-D-Ala carboxypeptidase daca enhances prodigiosin production in *Serratia marcescens*. *Front. Bioeng. Biotechnol.* 7:367. doi: 10.3389/fbioe.2019.00367

Papireddy, K., Smilkstein, M., Kelly, J. X., Shweta, Salem, S. M., Alhamadsheh, M., et al. (2011). Antimalarial activity of natural and synthetic prodiginines. *J. Med. Chem.* 54, 5296–5306. doi: 10.1021/jm200543y

Redden, H., and Alper, H. S. (2015). The development and characterization of synthetic minimal yeast promoters. *Nat. Commun.* 6:7810. doi: 10.1038/ncomms8810

Santos-Zavaleta, A., Salgado, H., Gama-Castro, S., Sanchez-Perez, M., Gomez-Romero, L., Ledezma-Tejeida, D., et al. (2019). RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in *E. coli* K-12. *Nucleic Acids Res.* 47, D212–D220. doi: 10.1093/nar/gky1077

Shaw, W. V. (1975). Chloramphenicol acetyltransferase from chloramphenicol-resistant bacteria. *Methods Enzymol.* 43, 737–755. doi: 10.1016/0076-6879(75)43141-X

Simpson, B. W., and Trent, M. S. (2019). Emerging roles for NlpE as a sensor for lipoprotein maturation and transport to the outer membrane in *Escherichia coli*. *MBio* 10:e01302-19. doi: 10.1128/mBio.01302-19

Siva, R., Subha, K., Bhakta, D., Ghosh, A. R., and Babu, S. (2012). Characterization and enhanced production of prodigiosin from the spoiled coconut. *Appl. Biochem. Biotechnol.* 166, 187–196. doi: 10.1007/s12010-011-9415-8

Spinola, S. M., Fortney, K. R., Baker, B., Janowicz, D. M., Zwickl, B., Katz, B. P., et al. (2010). Activation of the CpxRA system by deletion of cpxA impairs the ability of *Haemophilus ducreyi* to infect humans. *Infect. Immun.* 78, 3898–3904. doi: 10.1128/IAI.00432-10

Tanaka, Y., Yuasa, J., Baba, M., Tanikawa, T., Nakagawa, Y., and Matsuyama, T. (2004). Temperature-dependent bacteriostatic activity of *Serratia marcescens*. *Microbes Environ.* 19, 236–240. doi: 10.1264/jsme2.19.236

Von Graevenitz, A. (1980). *The Genus Serratia*. Boca Raron, FL: CRC Press.

Wasserman, H. H., Skles, R. J., Peverada, P., Shaw, C. K., Cushley, R. J., and Lipsky, C. R. (1973). Biosynthesis of prodigiosin. Incorporation patterns of C-labeled alanine, proline, glycine, and serine elucidated by fourier transform nuclear magnetic resonance. *J. Am. Chem. Soc.* 95, 6874–6875. doi: 10.1021/ja00801a080

Williams, R. P. (1973). Biosynthesis of prodigiosin, a secondary metabolite of *Serratia marcescens*. *Appl. Microbiol.* 25, 396–402. doi: 10.1128/AEM.25.3.396-402.1973

Williams, R. P., Gott, C. L., Qadri, S. M., and Scott, R. H. (1971). Influence of temperature of incubation and type of growth medium on pigmentation in *Serratia marcescens*. *J. Bacteriol.*106, 438–443. doi: 10.1128/JB.106.2.438-443.1971

Williamson, N. R., Fineran, P. C., Leeper, F. J., and Salmond, G. P. C. (2006). The biosynthesis and regulation of bacterial prodiginines. *Nat. Rev. Microbiol.* 4, 887–899. doi: 10.1038/nrmicro1531

Williamson, N. R., Simonsen, H. T., Ahmed, R. A. A., Goldet, G., Slater, H., Woodley, L., et al. (2005). Biosynthesis of the red antibiotic, prodigiosin, in Serratia: identification of a novel 2-methyl-3-n-amyl-pyrrole (MAP) assembly pathway, definition of the terminal condensing enzyme, and implications for undecylprodigiosin biosynthesis in Streptomyces. *Mol. Microbiol.* 56, 971–989. doi: 10.1111/j.1365-2958.2005.04602.x

Yamamoto, K., and Ishihama, A. (2006). Characterization of copper-inducible promoters regulated by CpxA/CpxR in *Escherichia coli*. *Biosci. Biotechnol. Biochem.*70, 1688–1695. doi: 10.1271/bbb.60024

You, J., Sun, L., Yang, X., Pan, X., Huang, Z., Zhang, X., et al. (2018). Regulatory protein SrpA controls phage infection and core cellular processes in *Pseudomonas aeruginosa*. *Nat. Commun.* 9:1846. doi: 10.1038/s41467-018-04232-6

Zhang, L., Sun, J. A., Hao, Y., Zhu, J., Chu, J., Wei, D., et al. (2010). Microbial production of 2,3-butanediol by a surfactant (serrawettin)-deficient mutant of *Serratia marcescens* H30. *J. Indus. Microbiol. Biotechnol.* 37, 857–862. doi: 10.1007/s10295-010-0733-6

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

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

MINI REVIEW
published: 05 June 2020
doi: 10.3389/fbioe.2020.00526

Bioprospecting Through Cloning of Whole Natural Product Biosynthetic Gene Clusters

*Zhenquan Lin[1], Jens Nielsen[1,2,3,4] and Zihe Liu[1]**

[1] Beijing Advanced Innovation Center for Soft Matter Science and Engineering, College of Life Science and Technology, Beijing University of Chemical Technology, Beijing, China, [2] Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden, [3] Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark, [4] BioInnovation Institute, Copenhagen, Denmark

Since the discovery of penicillin, natural products and their derivatives have been a valuable resource for drug discovery. With recent development of genome mining approaches in the post-genome era, a great number of natural product biosynthetic gene clusters (BGCs) have been identified and these can potentially be exploited for the discovery of novel natural products that can find application as pharmaceuticals. Since many BGCs are silent or do not express in native hosts under laboratory conditions, heterologous expression of BGCs in genetically tractable hosts becomes an attractive route to activate these BGCs to discover the corresponding products. Here, we highlight recent achievements in cloning and discovery of natural product biosynthetic pathways via intact BGC capturing, and discuss the prospects of high-throughput and multiplexed cloning of rational-designed gene clusters in the future.

Keywords: natural product, biosynthetic gene cluster, heterologous expression, sequence-independent cloning, direct cloning

OPEN ACCESS

Edited by:
Jiazhang Lian,
Zhejiang University, China

Reviewed by:
Yinhua Lu,
Shanghai Normal University, China
Yajie Wang,
University of Illinois at Urbana–Champaign, United States

****Correspondence:***
Zihe Liu
zihe@mail.buct.edu.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *25 March 2020*
Accepted: *04 May 2020*
Published: *05 June 2020*

Citation:
Lin Z, Nielsen J and Liu Z (2020) Bioprospecting Through Cloning of Whole Natural Product Biosynthetic Gene Clusters. Front. Bioeng. Biotechnol. 8:526. doi: 10.3389/fbioe.2020.00526

INTRODUCTION

Natural products produced by plant, bacteria, and fungi have served as a crucial source of pharmaceuticals, therapeutic agents and industrially useful compounds, such as antibiotic, antitumor, and anti-infective drugs (Nielsen, 2019). Since the discovery of penicillin in the early 1940s, the identification and bioprospection of natural product biosynthetic gene clusters (BGCs) has attracted much attention (Mullis et al., 2019). With the development of sequencing technologies, the costs of genome sequencing has been reduced, and hereby metagenomics has emerged as a strategic approach to explore unculturable microbes through the sequencing and analysis of environmental DNA. Hereby massive DNA sequence information has become accessible. Moreover, many bioinformatic tools have been developed to uncover putative BGCs, such as antiSMASH 5.0 (Blin et al., 2019), BiG-SCAPE (Navarro-Muñoz et al., 2019), PRISM 3 (Skinnider et al., 2017), MIBiG 2.0 (Kautsar et al., 2019), RODEO (Tietz et al., 2017), and genome-scale metabolic models (Nielsen and Nielsen, 2017). However, there are many technical challenges to translate these putative BGCs into specialized chemicals, resulting in a huge gap in the natural product discover pipeline (Dejong et al., 2016).

Advances in genetics, molecular biology and synthetic biology have been successfully used for natural product discovery (Zhang J. J. et al., 2019; Zhao et al., 2019). It has been estimated that more than 99% of environmental microbes are unculturable under the defined conditions using routine techniques and hard to study using classical experimental approaches (Daniel, 2005). Moreover, a large number of BGCs are not or weakly expressed in native hosts

under laboratory conditions, known as 'silent' or 'cryptic' gene clusters. Thus, besides the traditional screening and characterization methods, such as phenotype screening, insertional mutagenesis, co-culture and elicitor screening (Cacho et al., 2015; Tomm et al., 2019; Zhang X. et al., 2019), cloning and refactoring the putative BGCs in well-defined hosts become attractive approaches for natural product discovery, achieving functional expression of uncharacterized potentially-valuable natural product biosynthetic pathways (Cook and Pfleger, 2019; Xu et al., 2020). While *E. coli*, *Streptomyces*, yeast and *Aspergillus* are often used for heterogeneous expression of BGCs, their applications are still limited by the incompatibility of different transcript regulatory systems and codon preferences among organisms, lack of post-translational protein modifications, insufficient supplies of precursors and co-factors, toxicity of intermediates or final products, and poor assembly of natural products with novel structure (Luo et al., 2016; Nielsen, 2019; Pham et al., 2019). Unlike prokaryotic gene clusters, heterogeneous expression for eukaryotic gene clusters or individual genes introduces additional challenges for heterologous expression, such as intron splicing, insertion of promoters and terminators in upstream and downstream of each coding region, etc. (Alberti et al., 2017; Harvey et al., 2018; Qiao et al., 2019). Many alternative methods have been developed and comprehensively reviewed elsewhere (Baker et al., 2018; Xiong et al., 2019; Deng et al., 2020).

Over the past decade, many approaches have been developed to clone intact BGCs for heterologous expression. However, cloning long genome segments of large gene clusters (range from 20 to ∼200 kb) remains challenging (Fayed et al., 2015). Thus, it's necessary to develop appropriate vector systems and methods for cloning large-size gene clusters and transfer these genetic segments between different hosts (Liao et al., 2019). In this review, we will focus on recent developments of cloning intact gene clusters from complex genome sequence for natural products discovery, including sequence-independent methods and direct cloning methods, and prospect on high-throughput multiplexed cloning of BGCs.

SEQUENCE-INDEPENDENT METHODS FOR HETEROGENEOUS EXPRESSION OF BGCs

Sequence-independent method constructs expression libraries on sheared genomes from a mixed population (e.g., environmental DNA) or a pure culture, and screens for natural products. Key technologies in sequence-independent methods include high-quality high-molecular-weight DNAs isolation, DNA fragmentation and library construction. This method is particularly useful for scenarios when the genomic information of native hosts is under-characterized. Sequence-independent methods have the advantage to prospect the entire genetic materials, and is possible to cover all the BGCs in the sample and discover novel structural natural products (Zhang J. J. et al., 2019). The approach does, however, require highly efficient screening assays as the library will have a very low fraction of positives. Many groups have successfully used sequence-independent library cloning based on different library construction strategies [e.g., cosmids, fosmids, bacterial artificial chromosomes (BACs), phage artificial chromosomes (PACs)] for natural product discovery (**Table 1**) (Deng et al., 2017; Nara et al., 2017).

Isolation of High-Quality High-Molecular-Weight DNAs

Biosynthetic gene clusters are often 10s of kilobase and even over 100 kb. Thus, methods for preparing high-quality and high-molecular-weight DNAs are critical for successful cloning of intact BGCs. Zhang et al. (2010) reported a method of extracting high-molecular-weight DNAs from a variety of biological materials using CTAB (cetyl trimethyl ammonium bromide) extraction buffer for extraction, followed by phenolchloroform extraction and/or ethanol precipitation. However, this method often causes long DNA molecules shearing, and is used for extracting genomic DNAs up to ∼10 kb. To prepare megabase-size genomic DNA, cellulase and pectinase were firstly used to hydrolyze the cell wall before isolating DNAs from organisms having a cell walls (Zhang et al., 2012). Unlike conventional genomic DNA isolation methods, the protoplast, cells, or the nuclei are embedded in low-melting-point agarose gel matrix to protect large DNA fragments from mechanical shearing during the isolation step (Zhang et al., 2012). Alternatively, for rapid extraction of high-molecular-weight genomic DNA (range from ∼20 to ∼130 kb) from bacteria, plants, and animals, Mayjonade et al. (2016) developed a method that grounds the cell into a fine powder in liquid nitrogen, lyses the cell with SDS-base buffer and finally uses carboxylated magnetic beads to purify the DNA. For more information on the topic of isolating high-quality DNAs, please refer to recent reviews elsewhere (Mohamad Roslan et al., 2017; Green and Sambrook, 2018). Commercial kits for extracting high-molecular-weight DNA are also available (e.g., QIAGEN, Macherey Nagel). A detailed comparison of each method can be found in **Supplementary Table S1**.

DNA Fragmentation

Methods available for DNA fragmentation in library construction include enzymatic digestion, sonication, and hydrodynamic shearing (Ignatov et al., 2019). Enzymatic digestion, such as using site-directed restriction enzyme *Sau*3AI to partial digest purified DNA (Clos and Zander-Dinse, 2019), sonication, such as using ultrasound to generate > 120 kb fragments (Bhushan et al., 2011), and hydrodynamic shearing, such as repeatedly passing DNA through a syringe needle (Liu C. et al., 2016), have been widely used for constructions of large-fragment libraries. Compared with enzymatic shearing, sonication and hydrodynamic shearing, which are mechanical fragmentation methods, are more random and enable better control of the size distribution (Li et al., 2017). After fragmentation, DNA samples can be analyzed by fragment analyzer or horizontal agarose gel electrophoresis to test the extent of the yield fragments. Desired size of fragmented DNAs can

TABLE 1 | Different strategies for intact natural product BGCs cloning.

Class	Strategies	Principles	Capacity	Advantages	Disadvantages	BGCs
Sequence-independent libraries cloning	Cosmid/ fosmid libraries	• Fragmentation, gel-fractionated, ligation and phage packaging	<50 kb	• Not requiring genome sequence data; • Capable of generating natural product with novel structure; • Capable of covering the complete genetic material; • Suitable for cloning environmental DNA.	• Untargeted; • Laborious and time consuming; • Packaging; • Large BGCs maybe spanned into separatebd clones.	Omnipeptin (Libis et al., 2019) Anisomycin (Zheng et al., 2017) Ashimides (Shi et al., 2019) Frigocyclinone (Mo et al., 2019) Locillomycins (Luo et al., 2019)
	PAC/BAC libraries	• Fragmentation, gel-fractionated, and ligation	<300 kb	• Not requiring genome sequence data; • Capable of generating natural product with novel structure; • Capable of covering the complete genetic material.	• Untargeted; • Laborious and time consuming; • Technically challenging for large fragment cloning and DNA extraction.	Atratumycin (Yang et al., 2019) Neoabyssomicin/abyssomicin (Tu et al., 2018) Avermectins (Deng et al., 2017) Murayaquinone (Peng et al., 2018)
	FAC libraries	• Random fragmentation, adaptors ligation, gel-fractionated and ligation	10–200 kb	• Unbiased library; • Not requiring genome sequence data; • Capable of generating natural product with novel structure; • Suitable for fungal BGCs cloning.	• Untargeted; • Laborious and time consuming.	Sesterterpenoid (Clevenger et al., 2017) Benzomalvin A/D (Clevenger et al., 2018) Diketomorpholines (Robey et al., 2018)
Direct cloning	TAR	• *In vivo* homologous recombination of *Saccharomyces cerevisiae*	<100 kb	• Cas9-facilitated high efficiency cloning; • Suitable for cloning large genomic regions.	• Technically challenging to use yeast spheroplasts for highly transformation efficient; • Some false positives; • Requires careful preparation and/or manipulation of gDNA.	Plipastatin (Hu et al., 2018) Scleric acid (Alberti et al., 2019) Brasiliquinones (Herisse et al., 2019).
	LLHR	• RecET-mediated linear-plus-linear homologous recombination in *E. coli*	< ∼52 kb	• Technically easier; • Suitable for cloning small- and mid-BGCs; • Simply for using short recombination homologous arms.	• False positive; • Difficult to clone large-size BGCs; • Require highly specialized capturing vectors; • Multi-rounds selection.	Luminmide A/B (Fu et al., 2012) Bacillomycin (Liu Q. et al., 2016) Streptoketides (Qian et al., 2020)
	ExoCET	• CRISPR/Cas9 digestion, T4 polymerase for *in vitro* annealing and RecET mediated homologous recombination	<∼102 kb	• Technically easier; • Simply for using short recombination homologous arms.	• Low efficiency for clone large-size BGCs; • Require pathway specialized vectors; • False positive.	Salinomycin (Wang et al., 2018) Spinosad (Song et al., 2019)
	CATCH	• Cas9-assisted site-specific cleavage and Gibson assembly	< ∼150 kb	• Suitable for cloning large genomic regions.	• Require carefully prepare the target DNA in gel.	Bacillaene (Jiang et al., 2015) Mutanocyclin/SNC1-465 (Hao et al., 2019)
	DiPaC	• Q5 hi-fidelity PCR amplication and Gibson assembly	<22 kb per round	• Technically easier; • Extremely efficient for cloning small- to mid-size BGCs.	• Introduction of new mutations during PCR; • Impractical for large BGCs.	Phenazine fontizine A5 (Greunke et al., 2018) Sodorifen (Duell et al., 2019) Hapalosin (D'Agostino et al., 2018)

(Continued)

TABLE 1 | Continued

Class	Strategies	Principles	Capacity	Advantages	Disadvantages	BGCs
	Site-specific recombinase	• Homologous recombination and circularizing the plasmid *in vivo* or *in vitro* with recombinase	< ∼200 kb	• Effectively avoid the introduction of new mutations; • Suitable for high frequency recombination hosts; • Suitable for cloning large BGCs.	• Time-consuming; • Only use in the host with high-frequency natural homologous recombination; • Usually need multi-rounds recombination.	Napsamycin, daptomycin (Du et al., 2015) Siderophore (Hu et al., 2016) Erythromycin (Dai et al., 2015)
	iCatch	• Homologous recombination and *in vitro* self-ligation	< ∼20 kb	• Suitable for high frequency recombination hosts.	• Time-consuming; • Only use in the host with high-frequency natural homologous recombination; • Usually need multi-rounds recombination; • Carefully prepare the gnome DNA.	Actinorhodin (Wang et al., 2019)

be separated and extracted using multi-rounds of pulsed field gel electrophoresis (PFGE) with different ramped pulse times (Clos and Zander-Dinse, 2019). Compared to mechanical fragmentation methods, the unevenly distributed restriction sites in the genome may cause inherently biased and incomplete library with enzyme methods.

Cloning Strategies

After fragmentation and purification, the desired size of fragmented DNAs can be separated by multi-rounds PFGE. The size-selected fragments were end-repaired and ligated to the digested and dephosphorylated vector, such as cosmid, fosmid, BAC, or PAC (**Table 1**) (Liu et al., 2018; Tu et al., 2018; Clos and Zander-Dinse, 2019). Total ligation products can be transformed into *E. coli* or packaged into a phage for infecting bacteria. The insert size of cosmid/fosmid libraries usually is limited to ∼50 kb, thus, large gene clusters are often split into multiple fragments and reassemble into the whole cluster (Wolpert et al., 2008). Alternatively, PACs can clone inserts ranging in size from 60 to 150 kb, while BACs have a capacity to accommodate and propagate DNA fragments with an average insert size ∼150 kb (Bilyk et al., 2016). Several of these technologies have been turned into products, commercialized by a variety of companies such as Agilent, Bio S&T, and Epicentre Biotechnologies. For unbiased fungal artificial chromosome (FAC) library construction, fragmented DNAs was end-repaired and ligated with *Bst*XI adaptors and after for separating desired sizes DNA by PFGE (Bok et al., 2015). Purified large DNA fragments were ligated into the *Bst*XI-digested shuttle vector-FAC. The average inserts size of FAC libraries was about 150 kb, which can cover most fungal BGCs.

Successful Applications

The above-mentioned DNA assembly methods were developed in the past decades, and there are already many successful applications (**Table 1**). For example, Libis et al. (2019) have screened a 10 million cosmid library from soil metagenomic DNA samples using Co-occurrence Network Analysis of Targeted Sequences (CONKAT-seq), and identified omnipeptin. Moreover, Bok et al. (2015) have constructed a novel *Aspergillus–E. coli* shuttle FAC expression vector coupling a BAC vector backbone with an autonomous fungal replicating element AMA1 from *Aspergillus nidulans*. Clevenger et al. (2017) had then optimized the FAC-cloning method, and developed fungal artificial chromosomes with metabolomic scoring (FAC-MS) platform for the discovery of fungal specialized metabolites. Utilizing this approach, researchers have screened fragmented genome DNA containing uncharacterized fungal BGCs from *A. terreus*, *A. aculeatus*, and *A. wentii*, and discovered 17 compounds including 15 unreported compounds (Clevenger et al., 2017), including benzomalvin A/D (Clevenger et al., 2018), diketomorpholines (Robey et al., 2018).

In summary, sequence-independent library cloning can generate libraries for both un-sequenced and sequenced DNAs, with each clone harboring 10 to ∼200 kb inserts, promoting the natural product discovery. However, sequence-independent library cloning is usually laborious and time consuming. For

example, to reliably cover the whole genome, researchers usually need to generate 10–20 folds genome coverage to obtain the clones harboring BGCs (Bok et al., 2015). This will require optimization of the whole cloning process, for example, the genome extraction should not result in too much genomic fragmentation, the assembly including the transformation step should be highly efficient to generate the required library size, etc. Moreover, desired BGCs may be split into different clones, especially when using cosmid/fosmid libraries to screen and identify large gene clusters.

DIRECT CLONING METHODS FOR HETEROGENEOUS EXPRESSION OF BGCs

Direct cloning methods rely on precise bioinformatics to predict BGCs with targeted functions and use specialized cloning method to capture target sequence for expression and/or identification of secondary metabolites. The development of sequencing technologies has resulted in a dramatic reduction of sequencing cost, thus the genome or metagenome information can be easily generated. Meanwhile, several bioinformatic tools have been developed and successfully applied to identify BGCs with potential functions, including PRISM 3 (Skinnider et al., 2017), BiG-SCAPE (Navarro-Muñoz et al., 2019), and antiSMASH (Blin et al., 2019). Direct cloning methods aim to bypass the conventional library generation and screening process and directly isolate gene clusters of interest. Several groups have developed different approaches for direct capture of the BGCs (**Figure 1** and **Table 1**) (Hu et al., 2018; Alberti et al., 2019).

DNA Isolation and Fragmentation

Methods of DNA isolation and fragmentation for direct cloning are similar with methods used in sequence independent strategies. For DNA fragmentation, physical methods maybe caused target BGCs shearing into fragments. Moreover, the target BGCs are usually too large to find appropriate restriction enzymes that are capable to digest flank homologous regions without also digesting internal targets. To simplify the capture of BGCs, Greunke et al. (2018) developed the direct pathway cloning (DiPaC) method that utilizes long-amplicon PCR to generate target region and Gibson assembly to construct expression plasmids *in vitro*, as shown in **Figure 1B**. This method is capable of direct cloning small- to mid-sized BGCs (up to < 22 kb per round), resulting in discovery of phenazine fontizine A5 (Greunke et al., 2018), sodorifen (Duell et al., 2019), and heterologous production of anabaenopeptin and erythromycin (Greunke et al., 2018).

The development of advancing genome editing tools, such as clustered regularly interspaced short palindromic repeat–CRISPR-associated protein (CRISPR-Cas) system, has substantially accelerated the process of direct cloning and made it possible to isolate the exact sequence of target BGCs *in vitro* (Lee et al., 2015; Wang et al., 2018; Tao et al., 2019). For example, Jiang et al. (2015) developed Cas9-Assisted Targeting of Chromosome (CATCH) using CRISPR/Cas9 to generate double strand breaks at both ends of target BGCs *in vitro*, and cloned a 78-kb bacillaene gene cluster from *Bacillus subtilis* using Gibson assembly. The concentration of extracted BGCs from target genome or metagenome without enrichment may be too low to yield efficient cloning, and spheroplasts can be used to increase transformation efficiency in transformation-associated recombination (TAR) (Kouprina et al., 2020).

Cloning Strategies

The development of synthetic biology tools have enhanced cloning of intact BGCs in heterologous hosts (**Table 1**). Some of these methods are based on exonucleases to "chew back" one of the strands of double-stranded DNAs, thereby exposing complementary single-stranded DNA sequences that can anneal to each other *in vitro* (**Figures 1A,B**), such as Gibson isothermal assembly (Jiang et al., 2015; Greunke et al., 2018), sequence- and ligation-independent cloning (SLIC) (D'Agostino et al., 2018). Blunt-end ligation have also been emplyed to ligate the CRISPR/Cas9 digested product into a universal vector for λ packaging into phage and transfecting into *E. coli* (Tao et al., 2019) (**Figure 1A**).

This "chew-back and repair" mechanism has also been applied to clone intact BGCs leveraging on *in vivo* homologous recombination (**Figure 1C** and **Table 1**). Several hosts have been widely used for cloning purpose, such as the TAR method in *Saccharomyces cerevisiae* (Kouprina and Larionov, 2019), linear-linear homologous recombination (LLHR) or linear plus circular homologous recombination (LCHR) in *E. coli* (Fu et al., 2012), and exonuclease combined with RecET recombination (ExoCET) (Wang et al., 2018). In these methods, partially digested or randomly sheared DNA was co-transformed into the recombinant host with linearized and pathway-specific vectors containing homology arms that flank the upstream and downstream of the target BGCs. Also, ExoCET can be used to promote homologous recombinations between a linear DNA molecule and a circular plasmid (Fu et al., 2012).

Another type of approach employs site-direct recombination to clone intact BGCs by first integrating specific-vector with the integrase recognition sites in the native host, then the targeted BGCs together with the integrated vector are captured and circularized for heterologous expression (**Figures 1D,E** and **Table 1**). This approach requires the native host to have high efficiency of homologous recombination. For example, Dai et al. (2015) intergrated plasmid pEry-up and pEry-down with the BT1 integrase recognition sites *BattP* and *BattB* via single- or double- crossover at both ends of erythromycin BGC, after which genome DNA was carefully isolated and treated with the BT1 integrase to circularize at *att* recombination sequences as a plasmid via *in vitro* site-specific recombination. Similarily, iCatch intergrates homing endonucleases I-SceI and PI-PspI recognition sites flanking the region of interest, after which the genome is isolated and digested with I-SceI or PI-PspI and then self-ligated to clone the target BGC *in vitro* (Wang et al., 2019). Moreover, several groups have developed methods that express recombinases to extract DNA fragments between two integrase recognition sites and circularize the plasmid *in vivo* (**Figure 1D**), such as phage ϕBT1 integrase-mediated site-specific

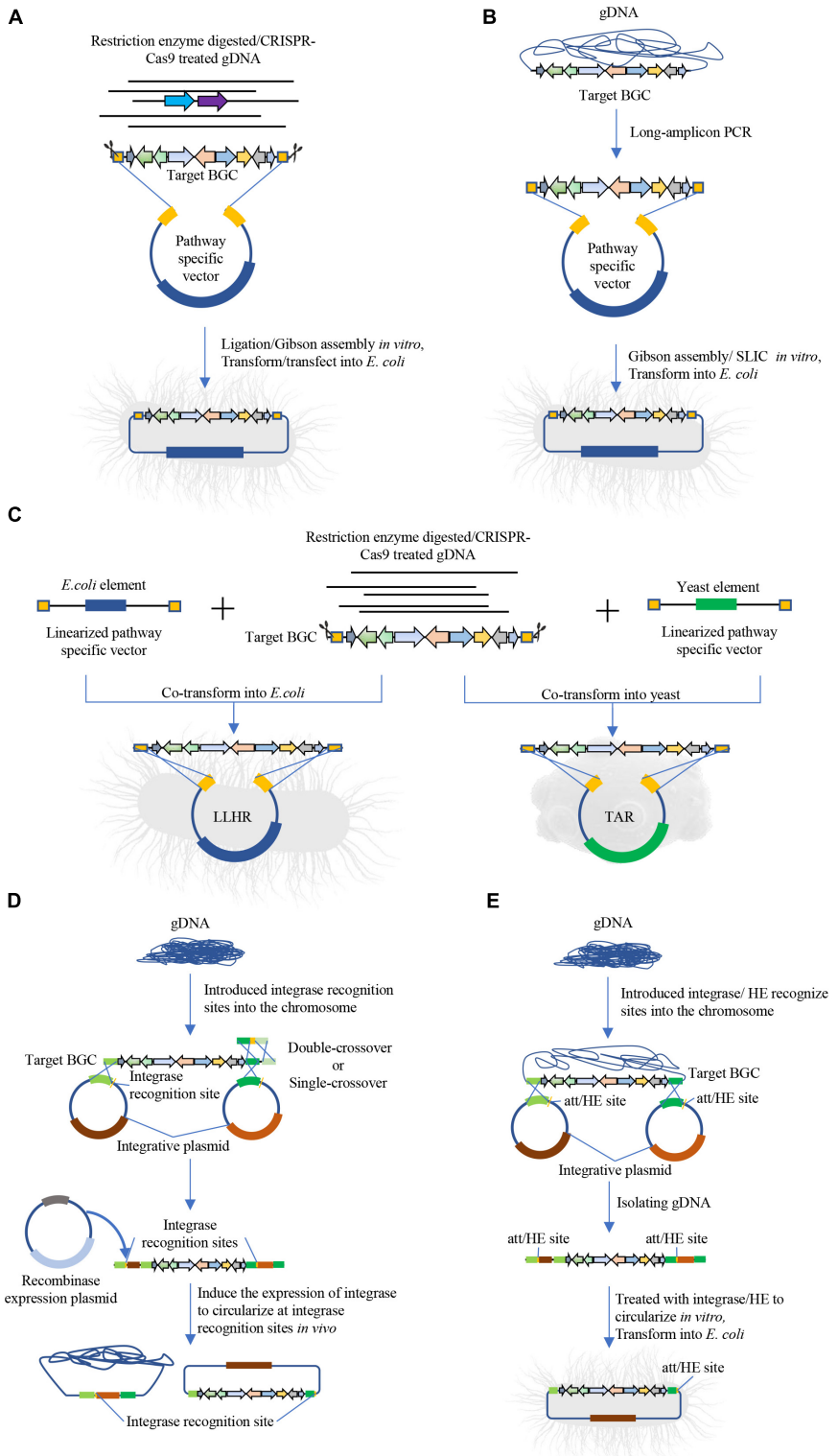


FIGURE 1 | Intact BGC capturing for natural products discovery. **(A)** Direct cloning method based on enzyme digestion and ligation, including ligation or Gibson assembly-based cloning of BGCs, such as CATCH. **(B)** Direct cloning method based on long-amplicon PCR and ligation, such as the DiPaC method. **(C)** Linear–linear homologous recombination (LLHR) mediated by full RecET in *E. coli* or transformation-associated recombination (TAR) in yeast for cloning BGCs. **(D)** Site-direct recombination for cloning BGC *in vivo,* including ϕBT1 integrase-mediated *in vivo* site-specific recombination, Cre/loxP plus BAC. **(E)** Site-direct recombination for cloning BGC *in vitro,* including iCatch, ϕBT1 integrase-mediated *in vitro* recombination.

recombination (Du et al., 2015), Cre/loxP plus BAC (Hu et al., 2016). These plasmids can then be isolated from the native host for heterologous expression.

As shown in **Table 1**, compared with *in vivo* methods (e.g., TAR, LLHR), *in vitro* cloning methods (e.g., DiPaC, CATCH, iCatch) require carefully preparation of DNA via pre-treatment and purification. Moreover, site-directed recombination methods are suitable for the host with high efficient homologous recombination system. Direct cloning are clearly valuable methods that are well-suited for mining the vast amount of genome for applications in natural product discovery.

Successful Applications

Over the past decade, direct cloning methods have made great advances and there are already many successful applications (**Table 1**). For example, Wang et al. (2018) developed a method where exonuclease was combined with ExoCET using the CRISPR/Cas9 cleavage system to digest the target genome and T4 polymerase to pre-anneal linear vector and target DNA before cotransforming into *E. coli*, resulting in cloning of the 106 kb salinomycin cluster and a 79 kb artificial gene cluster (Song et al., 2019). Moreover, TAR has been employed for identification of several novel natural products including orphan cosmomycin (Larson et al., 2017), thiostreptamide S4 (Frattaruolo et al., 2017), and scleric acid (Alberti et al., 2019).

In summary, direct cloning methods can clone intact clusters of interest accurately, and can substantially save time and efforts compared with sequence independent methods. It can also be combined with other modified methods to activated or refactor BGCs in heterologous host. However, current direct cloning methods rely heavily on the quality of genome sequencing and annotation techniques, and have been limited to capture and analyze only one or two clusters each reaction. With the rapidly developed synthetic biology tools available, it will be interesting to see whether these methods can be extended to directly clone all putative BGCs from a give genome in a single reaction.

CONCLUSION AND FUTURE PERSPECTIVES

In the past decades, the developments in synthetic biology, sequencing technology, and bioinformatics have greatly promoted the discovery of BGCs and corresponding products. We can now easily generate vast genome sequences via next-generation sequencing, and annotate them for potential BGCs using defined bioinformatic tools. These predictive BGCs, most of which are putative and do not fall in any known class of BGCs, can be cloned using sequence-independent methods to screen 1000s of clones in one round, or direct cloning methods to clone and analyze targeted BGCs one by one. However, with current technologies it is still challenging to combine the advantages of both methods, and to use direct cloning methods to test all predictive BGCs from an entire genome in a single reaction. Limitations include that high-quality and high-molecular weight DNAs for generating large BGCs is still difficult to isolate, efficient methods for seperating multi-fragments from digested DNA mixture are still missing, highly effective approaches for library construction of targeted BGCs are still limited. The cloning strategies cited in this review will need to be further optimized for BGCs identification and characterization.

In conclusion, with advancements in synthetic biology along with powerful genome mining techniques, we envision a new era of natural product discovery in which BGC cloning will be highly multiplexable, efficient and accurate in a high-throughput manner, leading to the discovery of numerous novel natural products with important biological activities.

AUTHOR CONTRIBUTIONS

ZLn, JN, and ZLu drafted the outline and wrote the manuscript. JN and ZLu supervised the research. All authors have read and approved the final manuscript.

FUNDING

This work was supported by National Key Research and Development Program (2018YFA0903000 and 2018YFA0900100), National Natural Science Foundation of China (21908004), the Novo Nordisk Foundation (NNF10CC1016517), the Knut and Alice Wallenberg Foundation and Beijing Advanced Innovation Center for Soft Matter Science and Engineering.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00526/full#supplementary-material

REFERENCES

Alberti, F., Foster, G. D., and Bailey, A. M. (2017). Natural products from filamentous fungi and production by heterologous expression. *Appl. Microbiol. Biotechnol.* 101, 493–500. doi: 10.1007/s00253-016-8034-2

Alberti, F., Leng, D. J., Wilkening, I., Song, L., Tosin, M., and Corre, C. (2019). Triggering the expression of a silent gene cluster from genetically intractable bacteria results in scleric acid discovery. *Chem. Sci.* 10, 453–463. doi: 10.1039/C8SC03814G

Baker, K. V., Takano, E., and Breitling, R. (2018). The "Three Cs" of novel antibiotic discovery and production through synthetic biology: biosynthetic gene clusters. Heterologous chassis, and synthetic microbial consortia. *Adv. Biosyst.* 2, 1–16. doi: 10.1002/adbi.201800064

Bhushan, S., Hossain, H., Lu, Y., Geisler, A., Tchatalbachev, S., Mikulski, Z., et al. (2011). Uropathogenic *E. coli* induce different immune response in testicular and peritoneal macrophages: implications for testicular immune privilege. *PLoS One* 6:e28452. doi: 10.1371/journal.pone.0028452

Bilyk, O., Sekurova, O. N., Zotchev, S. B., and Luzhetskyy, A. (2016). Cloning and heterologous expression of the grecocycline biosynthetic gene cluster. *PLoS One* 11:e0158682. doi: 10.1371/journal.pone.0158682

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., et al. (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47, W81–W87. doi: 10.1093/nar/gkz310

Bok, J. W., Ye, R., Clevenger, K. D., Mead, D., Wagner, M., Krerowicz, A., et al. (2015). Fungal artificial chromosomes for mining of the fungal secondary metabolome. *BMC Genomics* 16:343. doi: 10.1186/s12864-015-1561-x

Cacho, R. A., Tang, Y., and Chooi, Y.-H. (2015). Next-generation sequencing approach for connecting secondary metabolites to biosynthetic gene clusters in fungi. *Front. Microbiol.* 5:774. doi: 10.3389/fmicb.2014.00774

Clevenger, K. D., Bok, J. W., Ye, R., Miley, G. P., Verdan, M. H., Velk, T., et al. (2017). A scalable platform to identify fungal secondary metabolites and their gene clusters. *Nat. Chem. Biol.* 13, 895–901. doi: 10.1038/nchembio.2408

Clevenger, K. D., Ye, R., Bok, J. W., Thomas, P. M., Islam, M. N., Miley, G. P., et al. (2018). Interrogation of Benzomalvin biosynthesis using fungal artificial chromosomes with metabolomic scoring (FAC-MS): discovery of a benzodiazepine synthase activity. *Biochemistry* 57, 3237–3243. doi: 10.1021/acs.biochem.8b00076

Clos, J., and Zander-Dinse, D. (2019). Cosmid library construction and functional cloning. *Methods Mol. Biol.* 1971, 123–140. doi: 10.1007/978-1-4939-9210-2_6

Cook, T. B., and Pfleger, B. F. (2019). Leveraging synthetic biology for producing bioactive polyketides and non-ribosomal peptides in bacterial heterologous hosts. *Medchemcomm* 10, 668–681. doi: 10.1039/C9MD00055K

D'Agostino, P. M., Gulder, T. A. M. M., D'Agostino, P. M., and Gulder, T. A. M. M. (2018). Direct pathway cloning combined with sequence- and ligation-independent cloning for fast biosynthetic gene cluster refactoring and heterologous expression. *ACS Synth. Biol.* 7, 1702–1708. doi: 10.1021/acssynbio.8b00151

Dai, R., Zhang, B., Zhao, G., and Ding, X. (2015). Site-specific recombination for cloning of large DNA fragments *in vitro*. *Eng. Life Sci.* 15, 655–659. doi: 10.1002/elsc.201400267

Daniel, R. (2005). The metagenomics of soil. *Nat. Rev. Microbiol.* 3, 470–478. doi: 10.1038/nrmicro1160

Dejong, C. A., Chen, G. M., Li, H., Johnston, C. W., Edwards, M. R., Rees, P. N., et al. (2016). Polyketide and nonribosomal peptide retro-biosynthesis and global gene cluster matching. *Nat. Chem. Biol.* 12, 1007–1014. doi: 10.1038/nchembio.2188

Deng, H., Bai, Y., Fan, T., Zheng, X., and Cai, Y. (2020). Advanced strategy for metabolite exploration in filamentous fungi. *Crit. Rev. Biotechnol.* 40, 180–198. doi: 10.1080/07388551.2019.1709798

Deng, Q., Zhou, L., Luo, M., Deng, Z., and Zhao, C. (2017). Heterologous expression of Avermectins biosynthetic gene cluster by construction of a Bacterial Artificial Chromosome library of the producers. *Synth. Syst. Biotechnol.* 2, 59–64. doi: 10.1016/j.synbio.2017.03.001

Du, D., Wang, L., Tian, Y., Liu, H., Tan, H., and Niu, G. (2015). Genome engineering and direct cloning of antibiotic gene clusters via phage ϕBT1 integrase-mediated site-specific recombination in *Streptomyces*. *Sci. Rep.* 5:8740. doi: 10.1038/srep08740

Duell, E. R., D'Agostino, P. M., Shapiro, N., Woyke, T., Fuchs, T. M., and Gulder, T. A. M. (2019). Direct pathway cloning of the sodorifen biosynthetic gene cluster and recombinant generation of its product in *E. coli*. *Microb. Cell Fact.* 18:32. doi: 10.1186/s12934-019-1080-6

Fayed, B., Ashford, D. A., Hashem, A. M., Amin, M. A., El Gazayerly, O. N., Gregory, M. A., et al. (2015). Multiplexed integrating plasmids for engineering of the erythromycin gene cluster for expression in *Streptomyces* spp. and combinatorial biosynthesis. *Appl. Environ. Microbiol.* 81, 8402–8413. doi: 10.1128/AEM.02403-15

Frattaruolo, L., Lacret, R., Cappello, A. R., and Truman, A. W. (2017). A genomics-based approach identifies a thioviridamide-like compound with selective anticancer activity. *ACS Chem. Biol.* 12, 2815–2822. doi: 10.1021/acschembio.7b00677

Fu, J., Bian, X., Hu, S., Wang, H., Huang, F., Seibert, P. M., et al. (2012). Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat. Biotechnol.* 30, 440–446. doi: 10.1038/nbt.2183

Green, M. R., and Sambrook, J. (2018). Isolation and Quantification of DNA. *Cold Spring Harb. Protoc* 2018:db.to093336. doi: 10.1101/pdb.top093336

Greunke, C., Duell, E. R., D'Agostino, P. M., Glöckle, A., Lamm, K., and Gulder, T. A. M. (2018). Direct Pathway Cloning (DiPaC) to unlock natural product biosynthetic potential. *Metab. Eng.* 47, 334–345. doi: 10.1016/j.ymben.2018.03.010

Hao, T., Xie, Z., Wang, M., Liu, L., Zhang, Y., Wang, W., et al. (2019). An anaerobic bacterium host system for heterologous expression of natural product biosynthetic gene clusters. *Nat. Commun.* 10:3665. doi: 10.1038/s41467-019-11673-0

Harvey, C. J. B., Tang, M., Schlecht, U., Horecka, J., Fischer, C. R., Lin, H.-C., et al. (2018). HEx: a heterologous expression platform for the discovery of fungal natural products. *Sci. Adv.* 4:eaar5459. doi: 10.1126/sciadv.aar5459

Herisse, M., Ishida, K., Porter, J. L., Howden, B., Hertweck, C., Stinear, T. P., et al. (2019). Identification and mobilization of a cryptic antibiotic biosynthesis gene locus from a human-pathogenic Nocardia isolate. *ACS Chem. Biol.* [Epub ahead of print].

Hu, S., Liu, Z., Zhang, X., Zhang, G., Xie, Y., Ding, X., et al. (2016). "Cre/loxP plus BAC": a strategy for direct cloning of large DNA fragment and its applications in *Photorhabdus luminescens* and *Agrobacterium tumefaciens*. *Sci. Rep.* 6:29087. doi: 10.1038/srep29087

Hu, Y., Nan, F., Maina, S. W., Guo, J., Wu, S., and Xin, Z. (2018). Clone of plipastatin biosynthetic gene cluster by transformation-associated recombination technique and high efficient expression in model organism *Bacillus subtilis*. *J. Biotechnol.* 288, 1–8. doi: 10.1016/j.jbiotec.2018.10.006

Ignatov, K. B., Blagodatskikh, K. A., Shcherbo, D. S., Kramarova, T. V., Monakhova, Y. A., and Kramarov, V. M. (2019). Fragmentation Through Polymerization (FTP): A new method to fragment DNA for next-generation sequencing. *PLoS One* 14:e0210374. doi: 10.1371/journal.pone.0210374

Jiang, W., Zhao, X., Gabrieli, T., Lou, C., Ebenstein, Y., and Zhu, T. F. (2015). Cas9-Assisted Targeting of CHromosome segments CATCH enables one-step targeted cloning of large gene clusters. *Nat. Commun.* 6:8101. doi: 10.1038/ncomms9101

Kautsar, S. A., Blin, K., Shaw, S., Navarro-Muñoz, J. C., Terlouw, B. R., van der Hooft, J. J. J., et al. (2019). MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* 48, D454–D458. doi: 10.1093/nar/gkz882

Kouprina, N., and Larionov, V. (2019). TAR cloning: perspectives for functional genomics, biomedicine, and biotechnology. *Mol. Ther. Methods Clin. Dev.* 14, 16–26. doi: 10.1016/j.omtm.2019.05.006

Kouprina, N., Noskov, V. N., and Larionov, V. (2020). Selective isolation of large segments from individual microbial genomes and environmental DNA samples using transformation-associated recombination cloning in yeast. *Nat. Protoc.* 15, 734–749. doi: 10.1038/s41596-019-0280-1

Larson, C. B., Crüsemann, M., and Moore, B. S. (2017). PCR-independent method of transformation-associated recombination reveals the cosmomycin biosynthetic gene cluster in an ocean streptomycete. *J. Nat. Prod.* 80, 1200–1204. doi: 10.1021/acs.jnatprod.6b01121

Lee, N. C. O., Larionov, V., and Kouprina, N. (2015). Highly efficient CRISPR/Cas9-mediated TAR cloning of genes and chromosomal loci from complex genomes in yeast. *Nucleic Acids Res.* 43:e55. doi: 10.1093/nar/gkv112

Li, L., Jin, M., Sun, C., Wang, X., Xie, S., Zhou, G., et al. (2017). High efficiency hydrodynamic DNA fragmentation in a bubbling system. *Sci. Rep.* 7:40745. doi: 10.1038/srep40745

Liao, L., Su, S., Zhao, B., Fan, C., Zhang, J., Li, H., et al. (2019). Biosynthetic potential of a novel Antarctic Actinobacterium *Marisediminicola antarctica* ZS314^{T} revealed by genomic data mining and pigment characterization. *Mar. Drugs* 17, 1–15. doi: 10.3390/md17070388

Libis, V., Antonovsky, N., Zhang, M., Shang, Z., Montiel, D., Maniko, J., et al. (2019). Uncovering the biosynthetic potential of rare metagenomic DNA using co-occurrence network analysis of targeted sequences. *Nat. Commun.* 10:3848. doi: 10.1038/s41467-019-11658-z

Liu, C., Liu, X., Lei, L., Guan, H., and Cai, Y. (2016). Fosmid library construction and screening for the maize mutant gene *Vestigial glume* 1. *Crop J.* 4, 55–60. doi: 10.1016/j.cj.2015.09.003

Liu, Q., Shen, Q., Bian, X., Chen, H., Fu, J., Wang, H., et al. (2016). Simple and rapid direct cloning and heterologous expression of natural product biosynthetic gene cluster in *Bacillus subtilis* via Red/ET recombineering. *Sci. Rep.* 6:34623. doi: 10.1038/srep34623

Liu, Y., Zhang, B., Wen, X., Zhang, S., Wei, Y., Lu, Q., et al. (2018). Construction and characterization of a bacterial artificial chromosome library for *Gossypium mustelinum*. *PLoS One* 13:e0196847. doi: 10.1371/journal.pone.0196847

Luo, C., Chen, Y., Liu, X., Wang, X., Wang, X., Li, X., et al. (2019). Engineered biosynthesis of cyclic lipopeptide locillomycins in surrogate host *Bacillus velezensis* FZB42 and derivative strains enhance antibacterial activity. *Appl. Microbiol. Biotechnol.* 103, 4467–4481. doi: 10.1007/s00253-019-09784-1

Luo, Y., Enghiad, B., and Zhao, H. (2016). New tools for reconstruction and heterologous expression of natural product biosynthetic gene clusters. *Nat. Prod. Rep.* 33, 174–182. doi: 10.1039/c5np00085h

Mayjonade, B., Gouzy, J., Donnadieu, C., Pouilly, N., Marande, W., Callot, C., et al. (2016). Extraction of high-molecular-weight genomic DNA for long-read sequencing of single molecules. *Biotechniques* 61, 203–205. doi: 10.2144/000114460

Mo, J., Ye, J., Chen, H., Hou, B., Wu, H., and Zhang, H. (2019). Cloning and identification of the Frigocyclinone biosynthetic gene cluster from *Streptomyces griseus* strain NTK 97. *Biosci. Biotechnol. Biochem.* 83, 2082–2089. doi: 10.1080/09168451.2019.1638755

Mohamad Roslan, M. A., Naim Mohamad, M. A., and Mohd Omar, S. (2017). High-quality DNA from peat soil for metagenomic studies: a minireview on DNA extraction methods. *Sci. Herit. J.* 1, 01–06. doi: 10.26480/gws.02.2017.01.06

Mullis, M. M., Rambo, I. M., Baker, B. J., and Reese, B. K. (2019). Diversity, ecology, and prevalence of antimicrobials in nature. *Front. Microbiol.* 10:2518. doi: 10.3389/fmicb.2019.02518

Nara, A., Hashimoto, T., Komatsu, M., Nishiyama, M., Kuzuyama, T., and Ikeda, H. (2017). Characterization of bafilomycin biosynthesis in *Kitasatospora setae* KM-6054 and comparative analysis of gene clusters in Actinomycetales microorganisms. *J. Antibiot.* 70, 616–624. doi: 10.1038/ja.2017.33

Navarro-Muñoz, J. C., Selem-Mojica, N., Mullowney, M. W., Kautsar, S. A., Tryon, J. H., Parkinson, E. I., et al. (2019). A computational framework to explore large-scale biosynthetic diversity. *Nat. Chem. Biol.* 16, 60–68. doi: 10.1038/s41589-019-0400-9

Nielsen, J. (2019). Cell factory engineering for improved production of natural products. *Nat. Prod. Rep.* 36, 1233–1236. doi: 10.1039/C9NP00005D

Nielsen, J. C., and Nielsen, J. (2017). Development of fungal cell factories for the production of secondary metabolites: linking genomics and metabolism. *Synth. Syst. Biotechnol.* 2, 5–12. doi: 10.1016/j.synbio.2017.02.002

Peng, Q., Gao, G., Lü, J., Long, Q., Chen, X., Zhang, F., et al. (2018). Engineered *Streptomyces lividans* strains for optimal identification and expression of cryptic biosynthetic gene clusters. *Front. Microbiol.* 9:3042. doi: 10.3389/fmicb.2018.03042

Pham, J. V., Yilma, M. A., Feliz, A., Majid, M. T., Maffetone, N., Walker, J. R., et al. (2019). A review of the microbial production of bioactive natural products and biologics. *Front. Microbiol.* 10:1404. doi: 10.3389/fmicb.2019.01404

Qian, Z., Bruhn, T., D'Agostino, P. M., Herrmann, A., Haslbeck, M., Antal, N., et al. (2020). Discovery of the streptoketides by direct cloning and rapid heterologous expression of a cryptic PKS II gene cluster from *Streptomyces* sp. Tü 6314. *J. Org. Chem.* 85, 664–673. doi: 10.1021/acs.joc.9b02741

Qiao, Y.-M., Yu, R.-L., and Zhu, P. (2019). Advances in targeting and heterologous expression of genes involved in the synthesis of fungal secondary metabolites. *RSC Adv.* 9, 35124–35134. doi: 10.1039/C9RA06908A

Robey, M. T., Ye, R., Bok, J. W., Clevenger, K. D., Islam, M. N., Chen, C., et al. (2018). Identification of the first diketomorpholine biosynthetic pathway using FAC-MS technology. *ACS Chem. Biol.* 13, 1142–1147. doi: 10.1021/acschembio.8b00024

Shi, J., Zeng, Y. J., Zhang, B., Shao, F. L., Chen, Y. C., Xu, X., et al. (2019). Comparative genome mining and heterologous expression of an orphan NRPS gene cluster direct the production of ashimides. *Chem. Sci.* 10, 3042–3048. doi: 10.1039/C8SC05670F

Skinnider, M. A., Merwin, N. J., Johnston, C. W., and Magarvey, N. A. (2017). PRISM 3: expanded prediction of natural product chemical structures from microbial genomes. *Nucleic Acids Res.* 45, W49–W54. doi: 10.1093/nar/gkx320

Song, C., Luan, J., Cui, Q., Duan, Q., Li, Z., Gao, Y., et al. (2019). Enhanced heterologous spinosad production from a 79-kb synthetic Multioperon assembly. *ACS Synth. Biol.* 8, 137–147. doi: 10.1021/acssynbio.8b00402

Tao, W., Chen, L., Zhao, C., Wu, J., Yan, D., Deng, Z., et al. (2019). *In vitro* packaging mediated one-step targeted cloning of natural product pathway. *ACS Synth. Biol.* 8, 1991–1997. doi: 10.1021/acssynbio.9b00248

Tietz, J. I., Schwalen, C. J., Patel, P. S., Maxson, T., Blair, P. M., Tai, H.-C., et al. (2017). A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat. Chem. Biol.* 13, 470–478. doi: 10.1038/nchembio.2319

Tomm, H. A., Ucciferri, L., and Ross, A. C. (2019). Advances in microbial culturing conditions to activate silent biosynthetic gene clusters for novel metabolite production. *J. Ind. Microbiol. Biotechnol.* 46, 1381–1400. doi: 10.1007/s10295-019-02198-y

Tu, J., Li, S., Chen, J., Song, Y., Fu, S., Ju, J., et al. (2018). Characterization and heterologous expression of the neoabyssomicin/abyssomicin biosynthetic gene cluster from *Streptomyces koyangensis* SCSIO 5802. *Microb. Cell Fact.* 17:28. doi: 10.1186/s12934-018-0875-1

Wang, H., Li, Z., Jia, R., Yin, J., Li, A., Xia, L., et al. (2018). ExoCET: exonuclease in vitro assembly combined with RecET recombination for highly efficient direct DNA cloning from complex genomes. *Nucleic Acids Res.* 46:e28. doi: 10.1093/nar/gkx1249

Wang, J., Wang, J., Lu, A., Liu, J., Huang, W., Cai, Z., et al. (2019). iCatch: a new strategy for capturing large DNA fragments using homing endonucleases. *Acta Biochim. Biophys. Sin.* 51, 97–103. doi: 10.1093/abbs/gmy139

Wolpert, M., Heide, L., Kammerer, B., and Gust, B. (2008). Assembly and heterologous expression of the coumermycin a1 gene cluster and production of new derivatives by genetic engineering. *Chembiochem* 9, 603–612. doi: 10.1002/cbic.200700483

Xiong, Q., Xie, C., Zhang, Z., Liu, L., Powell, J. T., Shen, Q., et al. (2019). DNA origami post-processing by CRISPR-Cas12a. *Angew. Chem.* 132:201915555. doi: 10.1002/ange.201915555

Xu, W., Klumbys, E., Ang, E. L., and Zhao, H. (2020). Emerging molecular biology tools and strategies for engineering natural product biosynthesis. *Metab. Eng. Commun.* 10:e00108. doi: 10.1016/j.mec.2019.e00108

Yang, Z., Wei, X., He, J., Sun, C., Ju, J., and Ma, J. (2019). Characterization of the noncanonical regulatory and transporter genes in atratumycin biosynthesis and production in a heterologous host. *Mar. Drugs* 17:560. doi: 10.3390/md17100560

Zhang, J. J., Tang, X., and Moore, B. S. (2019). Genetic platforms for heterologous expression of microbial natural products. *Nat. Prod. Rep.* 36, 1313–1332. doi: 10.1039/C9NP00025A

Zhang, M., Zhang, Y., Scheuring, C. F., Wu, C.-C., Dong, J. J., and Zhang, H.-B. (2012). Preparation of megabase-sized DNA from a variety of organisms using the nuclei method for advanced genomics research. *Nat. Protoc.* 7, 467–478. doi: 10.1038/nprot.2011.455

Zhang, X., Hindra, and Elliot, M. A. (2019). Unlocking the trove of metabolic treasures: activating silent biosynthetic gene clusters in bacteria and fungi. *Curr. Opin. Microbiol.* 51, 9–15. doi: 10.1016/j.mib.2019.03.003

Zhang, Y. J., Zhang, S., Liu, X. Z., Wen, H. A., and Wang, M. (2010). A simple method of genomic DNA extraction suitable for analysis of bulk fungal strains. *Lett. Appl. Microbiol.* 51, 114–118. doi: 10.1111/j.1472-765X.2010.02867.x

Zhao, Q., Wang, L., and Luo, Y. (2019). Recent advances in natural products exploitation in Streptomyces via synthetic biology. *Eng. Life Sci.* 19, 452–462. doi: 10.1002/elsc.201800137

Zheng, X., Cheng, Q., Yao, F., Wang, X., Kong, L., Cao, B., et al. (2017). Biosynthesis of the pyrrolidine protein synthesis inhibitor anisomycin involves novel gene ensemble and cryptic biosynthetic steps. *Proc. Natl. Acad. Sci. U.S.A.* 114, 4135–4140. doi: 10.1073/pnas.1701361114

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 12 June 2020
doi: 10.3389/fbioe.2020.00551

Characterization of Context-Dependent Effects on Synthetic Promoters

Sebastian Köbbing[1], *Lars M. Blank*[1] *and Nick Wierckx*[1,2]*

[1] *Institute of Applied Microbiology - iAMB, Aachen Biology and Biotechnology – ABBt, RWTH Aachen University, Aachen, Germany,* [2] *Institute of Bio- and Geosciences (IBG-1: Biotechnology), Forschungszentrum Jülich GmbH, Jülich, Germany*

Understanding the composability of genetic elements is central to synthetic biology. Even for seemingly well-known elements such as a sigma 70 promoter the genetic context-dependent variability of promoter activity remains poorly understood. The lack of understanding of sequence to function results in highly limited *de novo* design of novel genetic element combinations. To address this issue, we characterized in detail concatenated "stacked" synthetic promoters including varying spacer sequence lengths and compared the transcription strength to the output of the individual promoters. The proxy for promoter activity, the msfGFP synthesis from stacked promoters was consistently lower than expected from the sum of the activities of the single promoters. While the spacer sequence itself had no activity, it drastically affected promoter activities when placed up- or downstream of a promoter. Single promoter-spacer combinations revealed a bivalent effect on msfGFP synthesis. By systematic analysis of promoter and spacer combinations, a semi-empirical correlation was developed to determine the combined activity of stacked promoters.

Keywords: ***Pseudomonas putida*, synthetic biology, synthetic promoter libraries, Tn7 transposon, tandem promoter, heterologous expression**

OPEN ACCESS

Edited by:
Yunzi Luo,
Sichuan University, China

Reviewed by:
Wenjing Cui,
Jiangnan University, China
Mario Andrea Marchisio,
Tianjin University, China

***Correspondence:**
Nick Wierckx
n.wierckx@fz-juelich.de

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *14 March 2020*
Accepted: *07 May 2020*
Published: *12 June 2020*

Citation:
Köbbing S, Blank LM and Wierckx N (2020) Characterization of Context-Dependent Effects on Synthetic Promoters. Front. Bioeng. Biotechnol. 8:551. doi: 10.3389/fbioe.2020.00551

INTRODUCTION

The Pseudomonads are a promising group of bacteria for industrial applications (Wierckx et al., 2005; Tiso et al., 2014; Aparicio et al., 2018). A versatile metabolism enables them to grow on several carbon sources like glucose and glycerol, but also on a wide range of aliphatics and aromatics (Jiménez et al., 2002; Nikel et al., 2014; Köhler et al., 2015). Different *Pseudomonas* strains have been engineered for the production of chemicals with industrial importance from different renewable carbon sources, like furandicarboxylic acid, rhamnolipids, and aromatics (Wierckx et al., 2005; Sun et al., 2007; Blank et al., 2008; Koopman et al., 2010; Meijnen et al., 2011; Wynands et al., 2018). *Pseudomonas* is highly tolerant to chemical stresses and can survive harmful conditions caused by oxidative stress (Isken and de Bont, 1998; Ramos et al., 2002; Wierckx et al., 2005; Wynands et al., 2018). Some strains can thrive under a second phase of toxic hydrophobic solvents such as toluene or styrene (Heipieper et al., 2007; Kusumawardhani et al., 2018). *P. putida* KT2440 is a non-pathogenic representative of this versatile group of bacteria (Nelson et al., 2002). The strain is able to produce and accumulate polyhydroxyalkanoates (PHA) as a storage polymer in granules under nitrogen depletion from different carbon sources like glycerol, glucose, ethylene glycol, 1,4-butanediol, or fatty acids (Sun et al., 2007; Wang and Nomura, 2010; Franden et al., 2018; Li et al., 2019, 2020).

Parallel to the increasing industrial interest in Pseudomonads, an ever-increasing set of synthetic biology tools is developed for this genus. These include genomic integration tools like transposon Tn5 (de Lorenzo et al., 1990; Herrero et al., 1990; Nikel and de Lorenzo, 2013) or Tn7 (Lambertsen et al., 2004; Damron et al., 2013; Silva-Rocha and de Lorenzo, 2014), as well as a suite of tools for targeted and marker-less integration (Martínez-García and de Lorenzo, 2011). To obtain gene replacements, counter-selection procedures were established, like *sacB* originating from *Bacillus subtilis* (Schweizer, 1992). New tools are based on CRISPR/Cas9 showing high potential for whole-genome engineering approaches (Jiang et al., 2013; Aparicio et al., 2018). These tools enable a deep genetic and metabolic re-factoring of different Pseudomonads as exemplified in the engineering of streamlined chassis strains (Shen et al., 2017; Wynands et al., 2018; Sánchez-Pascuala et al., 2019).

Especially when such deep engineering entails the (over-) expression of many homologous or heterologous genes, balanced and reliable gene expression is required, which doesn't unnecessarily burden the cell. In this context, calibrated synthetic promoter libraries enable modulation of enzyme expression in metabolic pathways and protein production (Rud et al., 2006; Solem et al., 2007). Two major ways to generate a promoter library are prominently used. A low degeneracy approach, where only a few random nucleotides are introduced, reduces the number of possible generated promoter sequences and thus decreases the number of sequences, which have to be tested (Mutalik et al., 2013). This allows a deeper insight into promoter sequence-activity relationships. On the other hand, high degeneracy promoter libraries based on a degenerated core promoter sequence lead to billions of different possibilities (Zobel et al., 2015; Gilman and Love, 2016; Elmore et al., 2017). While the sequence space clearly outnumbers the experimental space possible to address, a high resolution of different promoter activities is possible. The use of calibrated and standardized synthetic promoters covering a range of activities are commonly used (Zobel et al., 2015). Constitutive synthetic promoters are generally based on sigma-70 (σ^{70}) factor core promoters (Gruber and Gross, 2003). The σ^{70} factor encoded by *rpoD* guides the RNA polymerase to many promoters active during growth including the expression of housekeeping genes (Kang et al., 1997; Potvin et al., 2008). Varying the consensus sequences of the −10 and −35 elements, which are recognized by the holoenzyme as part of the core promotor (Lodge et al., 1990; McLean et al., 1997), leads to weaker expression strength in *E. coli* and *P. aeruginosa* (McLean et al., 1997). The σ^{70} factors consensus sequence of *P. putida* KT2440 and *P. aeruginosa* are identical (McLean et al., 1997; Zobel et al., 2015).

Characterization of synthetic promoters has been performed using plasmid-based expression systems or genomically integrated probes (Jensen and Hammer, 1998; Hammer et al., 2006; Zobel et al., 2015). However, varying plasmid copy numbers and high fitness costs for the host makes plasmid-based expression systems less suitable for promoter characterization in particular, and for metabolic engineering in general (Gao et al., 2014; Jahn et al., 2014; Lindmeyer et al., 2015; San Millan and MacLean, 2017). Genomic integration of the probe is preferred for characterization procedures (Zobel et al., 2015). The major difference is the fact that many of the plasmids used are multicopy, which increases the variability of the reporter output by copy number variations. In addition, an often-overlooked disadvantage of using multicopy plasmids for synthetic promoter screening is that they favor the selection of relatively weak promoters, as the combined effect of a strong constitutive promoter at high copy number may pose a too high burden. Genomic integration abolishes these copy number effects, as well as clonal variations, which have also been observed for different *Pseudomonas* strains (Friehs, 2004; Gao et al., 2014; Zobel et al., 2015). Nevertheless, the integration site in the genome must be chosen wisely and must be the same for all promoters. The expression activity differs not only for single genes, but also in larger regions on the genome ("hot" and "cold" spots). Therefore, we used a mini Tn7 transposon, which integrates in a targeted manner downstream of the *glmS* gene in the *attTn7* site of a broad range of bacteria including *P. putida* KT2440, thereby enabling reliable and stable expression (Lambertsen et al., 2004; Choi et al., 2005; Zobel et al., 2015).

Synthetic promoter libraries are described for *P. putida* KT2440 (Zobel et al., 2015; Elmore et al., 2017). However, the predictability and composability of these promoters in different genetic contexts is poorly understood. Li et al. (2012) has shown that different numbers of promoters in tandem direction result in increased activities. Several other publications feature tandem promoters, but so far without a characterization of these promoter combinations that focusses on composability and predictability of the activity of these genetic elements (Dixon, 1984; Martens et al., 2004; Tamsir et al., 2011). The combination of promoters in different contexts is also a key element in logic gate construction in synthetic biology, and sensitivity to genetic context is considered a challenge there (Stanton et al., 2013).

In this work, we stacked (concatenated) promoters of a previously published promoter library from Zobel et al. (2015) in series and analyzed the resulting activities as single genomically integrated probes by measuring *msfGFP* expression (Landgraf, 2012). The obvious assumption that the combination of two promoters would yield their summed activity proved to be false. The reasons for this are investigated and a semi-empirical correlation was developed to reliably predict stacked synthetic promoter activities. This provided insights into the context-dependent activity of promoters that may foster a better predictability and composability of this key element of synthetic biology.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Cultivation Conditions

Strains and plasmids used and generated in this study are listed in **Table 1**. For cloning chemically competent *E. coli* PIR2 (Life Technologies, Carlsbad, USA) were used (Hanahan, 1983). Cultivation of *E. coli* was performed in lysogeny broth (LB) with 5 g L^{-1} NaCl (Sambrook et al., 1989). For solid media 15 g L^{-1} agar was added to the medium before autoclaving. To

TABLE 1 | Strains and plasmids used and generated in this study.

Strain	Description	References
E. coli		
HB101	*F*$^-$ *mcrB mrr hsdS20*(*rB*$^-$ *mB*$^-$) *recA13 leuB6 ara-14 proA2 lacY1 galK2 xyl-5 mtl-1 rpsL20*(Sm^R) *gln V44*λ^-	Boyer and Roulland-Dussoix (1969)
CC118λpir	Δ(*ara-leu*) *araD* Δ*lacX74 galE galK phoA20 thi-1 rpsE rpoB argE*(*Am*) *recA1,* lysogenized with λpir phage	Herrero et al. (1990)
PIR2	F$^-$ Δ*lac169 rpoS* (Am) *robA1 creC510 hsdR514 endA reacA1 uidA* (Δ*MluI*)::*pir*	Life Technologies
E. coli DH5αλ pir	*endA1 hsdR17 glnV44* (= *supE44*) *thi-1 recA1 gyrA96 relA1* ϕ*80dlac*Δ(*lacZ*)*M15* Δ(*lacZYA-argF*)*U169 zdg-232*::Tn10 *uidA*::*pir*+	de Lorenzo lab
P. putida		
KT2440	Wild-type strain derived of *P. putida* mt-2 cured of the pWW0 plasmid	Bagdasarian et al. (1981)
BG	Gm^R, *P. putida* KT2440 with genomic insertion of pBG	Zobel et al. (2015)
BG13	Gm^R, *P. putida* KT2440 with genomic insertion of pBG13	Zobel et al. (2015)
BG14a	Gm^R, *P. putida* KT2440 with genomic insertion of pBG14a	Zobel et al. (2015)
BG14b	Gm^R, *P. putida* KT2440 with genomic insertion of pBG14b	Zobel et al. (2015)
BG14c	Gm^R, *P. putida* KT2440 with genomic insertion of pBG14c	Zobel et al. (2015)
BG14d	Gm^R, *P. putida* KT2440 with genomic insertion of pBG14d	Zobel et al. (2015)
BG14e	Gm^R, *P. putida* KT2440 with genomic insertion of pBG14e	Zobel et al. (2015)
BG14f	Gm^R, *P. putida* KT2440 with genomic insertion of pBG14f	Zobel et al. (2015)
BG14g	Gm^R, *P. putida* KT2440 with genomic insertion of pBG14g	Zobel et al. (2015)
BG14f_##_14g	Gm^R, *P. putida* KT2440 with genomic insertion of pBG14f_##_14g, spacer with varying length from ten to 100 bp	This work
BG_80i	Gm^R, *P. putida* KT2440 with genomic insertion of pBG_80i	This work
BG_80new	Gm^R, *P. putida* KT2440 with genomic insertion of pBG_80new	This work
BG14x_80i_14y	Gm^R, *P. putida* KT2440 with genomic insertion of pBG14x_80i_14y	This work
BG14f_80i_14f_80i_14g	Gm^R, *P. putida* KT2440 with genomic insertion of pBG14f_80i_14f_80i_14g	This work
BG14x_80i	Gm^R, *P. putida* KT2440 with genomic insertion of pBG14x_80i	This work
BG_80i_14y	Gm^R, *P. putida* KT2440 with genomic insertion of pBG_80i_14y	This work
BG14f_80new	Gm^R, *P. putida* KT2440 with genomic insertion of pBG14f_80new	This work
BG_80new_14g	Gm^R, *P. putida* KT2440 with genomic insertion of pBG_80new_14g	This work
BG14f_80new_14g	Gm^R, *P. putida* KT2440 with genomic insertion of pBG14f_80new_14g	This work
BG14g_SNP_PosZZn	Gm^R, *P. putida* KT2440 with genomic insertion of pBG14g_SNP_PosZZ_N	This work
BG14g_PosZZn_80i	Gm^R, *P. putida* KT2440 with genomic insertion of pBG14f_PosZZ_N_80i	This work
Plasmids		
pRK600	Cm^R, oriColE1, *tra* + *mob* + of RK2	Keen et al. (1988)
pTnS-1	Ap^R, oriR6K, *TnSABC+D* operon	Choi et al. (2005)
pBG	Km^R, Gm^R, oriR6K, Tn7L and Tn7R extremes, BCD2–*msfgfp* fusion	Zobel et al. (2015)
pBG13	Km^R, Gm^R, oriR6K, pBG-derived, promoter P_{em7}	Martínez-García et al. (2015)
pBG14a	Km^R, Gm^R, oriR6K, pBG-derived, promoter 14a	Zobel et al. (2015)
pBG14b	Km^R, Gm^R, oriR6K, pBG-derived, promoter 14b	Zobel et al. (2015)
pBG14c	Km^R, Gm^R, oriR6K, pBG-derived, promoter 14c	Zobel et al. (2015)
pBG14d	Km^R, Gm^R, oriR6K, pBG-derived, promoter 14d	Zobel et al. (2015)
pBG14e	Km^R, Gm^R, oriR6K, pBG-derived, promoter 14e	Zobel et al. (2015)
pBG14f	Km^R, Gm^R, oriR6K, pBG-derived, promoter 14f	Zobel et al. (2015)
pBG14g	Km^R, Gm^R, oriR6K, pBG-derived, promoter 14g	Zobel et al. (2015)
pBG14f_##_14g	Km^R, Gm^R, oriR6K, pBG-derived, stacked promoter 14f/14g, spacer with varying length from ten to 100 bp	This work
pBG_80i	Km^R, Gm^R, oriR6K, pBG-derived, promoter-less control, reverse complement spacer sequence with 80 bp length	This work
pBG_80new	Km^R, Gm^R, oriR6K, pBG-derived, promoter-less control, new spacer sequence with 80 bp length	This work
pBG14x_80i_14y	Km^R, Gm^R, oriR6K, pBG-derived, stacked promoter 14x/14y, inverted spacer with a length of 80 bp	This work

(Continued)

TABLE 1 | Continued

Strain	Description	References
pBG14f_80new_14g	Km^R, Gm^R, oriR6K, pBG-derived, stacked promoter 14f/14g, new spacer sequence with 80 bp length	This work
pBG14f_80i_14f_80i_14g	Km^R, Gm^R, oriR6K, pBG-derived, stacked promoter 14f_80i_14f_80i_14g, inverted spacer with a length of 80 bp	This work
pBG14x_80i	Km^R, Gm^R, oriR6K, pBG-derived, first position promoter control 14x, inverted spacer with a length of 80 bp	This work
pBG_80i_14y	Km^R, Gm^R, oriR6K, pBG-derived, second position promoter control 14y, inverted spacer with a length of 80 bp	This work
pBG14f_80new	Km^R, Gm^R, oriR6K, pBG-derived, first position promoter control 14f, new spacer sequence with 80 bp length	This work
pBG_80bp_14g_new	Km^R, Gm^R, oriR6K, pBG-derived, second position promoter control 14g, new spacer sequence with 80 bp length	This work
pBG14g_SNP_PosZZn	Km^R, Gm^R, oriR6K, pBG-derived, single nucleotide promoter library with specific positions changes, library is based on 14g	This work
pBG14g_PosZZn_80i	Km^R, Gm^R, oriR6K, pBG-derived, first position promoter control 14g_PosZZ_N with modified nucleotide at distinct position of 14g core promoter sequence, inverted spacer with a length of 80 bp	This work

*A complete list containing all strains described more in detail can be found in the **Supplementary Table 8**.*
##Distance between two promoters in bp.
x stands for promoter 14a to 14g at the first position.
y stands for promoter 14a to 14g at the second position.
ZZ stands for position 1 to 30 of 14g based SNP.
n stands for nucleotides A, C, G, or T.

maintain the mini Tn7 plasmid in *E. coli*, 50 mg L^{-1} kanamycin was added to either liquid or solid medium. For pRK600 (Keen et al., 1988) bearing strains chloramphenicol (10 mg L^{-1}) and for pTnS-1 (Choi et al., 2005) ampicillin (100 mg L^{-1}) was used. Cultivation of *E. coli* strains was carried out at 37°C. Integration of the mini Tn7 transposon was performed by patch mating on LB agar plates and subsequent cultivation overnight at 30°C (Zobel et al., 2015). This was done with a mini Tn7 suicide plasmid-bearing donor strain, acceptor strain *P. putida* KT2440, and two helper-strains. *E. coli* HB101 (Boyer and Roulland-Dussoix, 1969) bearing plasmid pRK600 (Keen et al., 1988) with mobilization genes. *E. coli* DH5αλpir bearing pTnS-1 (Choi et al., 2005) encodes a transposase for transposition of the mini Tn7 transposon. For selection and counter-selection of Tn7-bearing *P. putida* KT2440, cetrimide agar plates containing 30 mg L^{-1} gentamycin and 1 % glycerol were used.

Cultivations of *P. putida* KT2440 derivatives for promoter characterization were performed in minimal medium containing 3.88 g L^{-1} K_2HPO_4 and 1.63 g L^{-1} NaH_2PO_4 with 20 mM glucose as the sole carbon source (Hartmans et al., 1989).

DNA Techniques

All oligonucleotides used in this study are listed in **Supplementary Table 1**. For the generation of different length spacer sequences PCR was used (**Supplementary Table 2**). Up to 40 bp length a forward oligonucleotide (SK11, SK34, SK36, or SK38) and reverse oligonucleotide SK2 were used with plasmid pBG14g as template (**Supplementary Table 3**). Q5 polymerase (New England Biolabs) with proofreading activity was used for amplification and reactions were prepared as described by the manufacturer. The resulting fragments were cut out from agarose gels and purified with the DNA Gel Extraction kit from New England Biolabs. Longer spacer sequences from 50 to 100 bp were assembled by PCR with two long oligonucleotides with complementary 3′-ends and an initial annealing step in the PCR (**Figure 1**, **Supplementary Table 4**). Dimer formation of 3′-ends was checked *in silico* to ensure that stacked promoter constructs can be formed by annealing of two oligonucleotides (https://www.thermofisher.com/de/de/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/multiple-primer-analyzer.html). A detailed approach is described in the Supporting Information. PCR fragments were purified with PCR & DNA Cleanup Kit (New England Biolabs). Promoter-less plasmid pBG was used as backbone. Generated spacer fragments were digested by *Pac*I and *Avr*II (New England Biolabs) at 37°C. Cloning procedures are following the rules of SEVA and are thus compatible with other constructs (Silva-Rocha et al., 2013). Plasmid pBG was additionally treated with alkaline phosphatase (Fermentas) to circumvent self-ligation. Digested fragments were purified using PCR & DNA Cleanup Kit (New England Biolabs). DNA concentrations of eluted DNA were measured with a NanoDrop One (Thermo Scientific). Fragments and backbone with adjusted concentrations were ligated using T4 ligase (New England Biolabs) at room temperature for 30 min. Transformation of ligated plasmids was performed by heat shock into chemically competent *E. coli* PIR2 cells (Hanahan, 1983). Plasmid isolation was done with Plasmid Miniprep Kit (New England Biolabs) and sequences were confirmed by Sanger sequencing (Eurofins

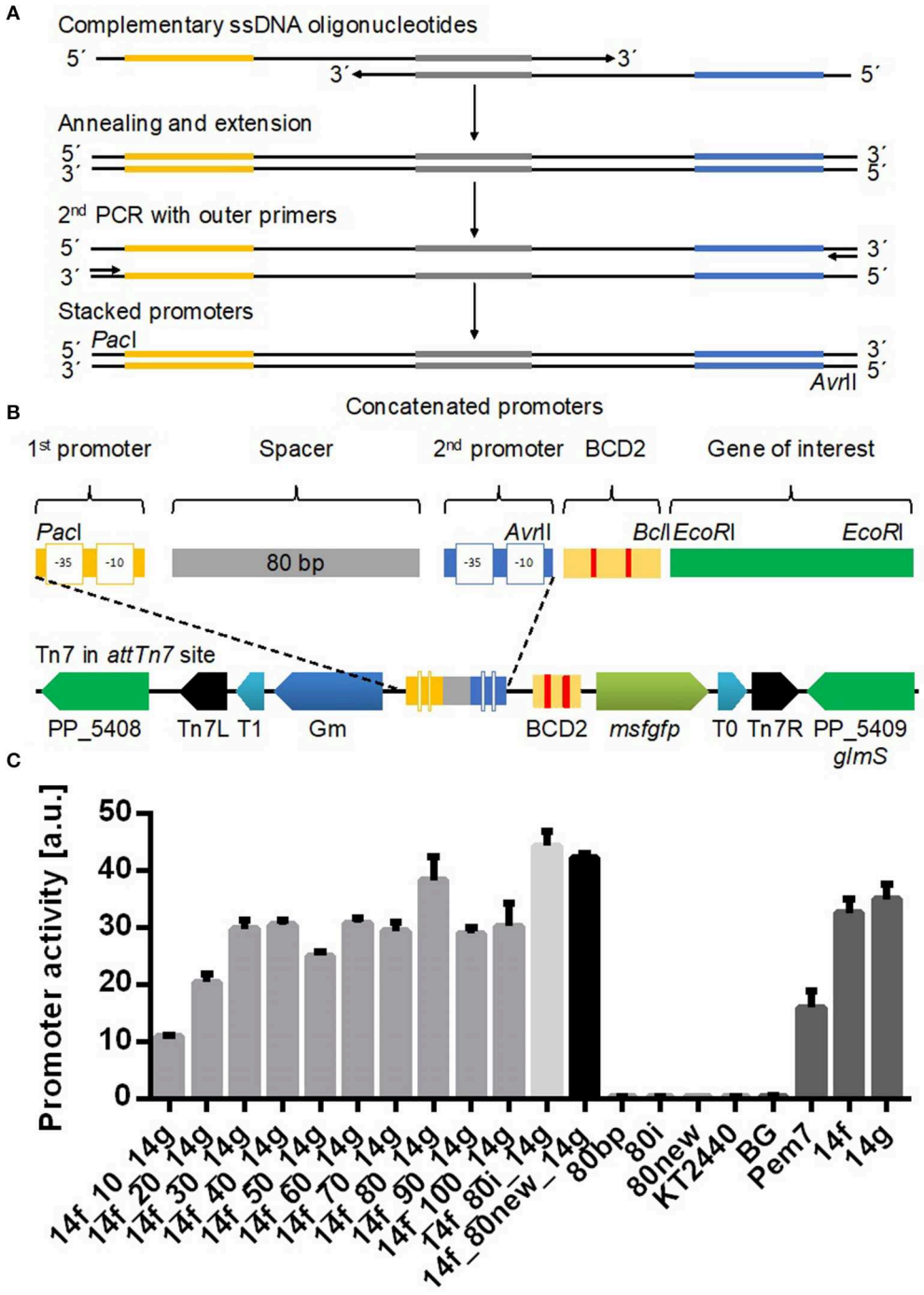


FIGURE 1 | Identification of the optimal spacer length between the two promoters 14f and 14g using a mini Tn7 vector (Zobel et al., 2015). **(A)** Two PCR reactions were performed to generate stacked promoters with longer (>45 bp) spacer sequences and promoter combinations with the 80i spacer sequence. In a first PCR reaction long single stranded DNA oligonucleotides containing one promoter sequence (yellow or blue) were annealed via complementary sequences in the spacer (gray) and extended by Q5 polymerase to double stranded DNA. The resulting dsDNA fragment was amplified in a second PCR. **(B)** Structural organization of stacked promoters and of the mini Tn7 used in this study after genomic integration. Stacked promoters consisting of promoter sequences at two positions separated by a spacer were inserted via restriction sites *Pac*I and *Avr*II. BCD2 element for translational coupling and *msfGFP* as reporter gene. Tn7 module contains a GmR marker for selection and two terminators (T0 and T1) for insulation of the probe. Tn7R and Tn7L are recognized by a transposase. **(C)** Tested spacers contained between 10 and 100 bp. *P. putida* KT2440 *attTn7::*BGf##g-mfsGFP, where ## refers to the number of nucleotides in the spacer sequence (gray bars), were cultured in a BioLector in minimal medium with 20 mM glucose in a 96 well plate. The control strains BG13 with the Pem7 promoter of average strength, the individual promoters 14f and 14g, and promoterless BG and wild type *P. putida* KT2440, as well as additional controls with two 80 bp spacers 80i and 80new are also shown. Identical strains from at least two different transformations were tested, with three biological replicates each. Error bars indicate the standard error of the mean ($n > 6$).

Genomics). Oligonucleotide combinations and detailed PCR protocols are described in **Supplementary Table 5**.

For the construction of promoter positions and spacer controls, previously cloned plasmids containing stacked promoters were used as template (**Supplementary Tables 6**, **7**).

Construction of the single nucleotide polymorphism (SNP) library was done by PCR with plasmid pBG14g as template and primers SK63-SK92 containing single degenerate nucleotides.

We used colony PCR with oligonucleotides SK4 and SK5 to verify the correct and full-length genomic integration of Tn7 at the *attTn7* site in *P. putida* KT2440. Single colonies were picked with a pipette tip or toothpick and lysed in 30 μL lysis buffer containing 60 % alkaline PEG 200 (pH adjusted to 13–13.5 with 2M KOH) for 15 min at room temperature (Wynands et al., 2018). As template one microliter was used for the PCR reaction (*Taq* 2X Master Mix, New England BioLabs).

Measuring Fluorescence and Determination of Promoter Activity

For the identification of SNPs in the obtained promoter library for the individual positions we measured GFP fluorescence. We cultivated *E. coli* PIR2 mini Tn7 plasmid-bearing strains in 0.5 mL LB medium containing 50 mg L^{-1} kanamycin at 30°C in 96 well System Duetz plates (Enzyscreen, The Netherlands). Fluorescence of msfGFP was measured in a synergyMX plate reader (Biotek, Bad Friedrichshall, Germany). Samples were measured in black bottom 96 well plates at an excitation wavelength of 488 nm and emission wavelength of 520 nm. Absorption was measured at 600 nm in clear bottom 96 well plates. From strains showing different intensities for GFP fluorescence the plasmid was isolated and sequenced. Followed by genomic integration of desired plasmids in *P. putida* KT2440 by triparental mating.

Growth and fluorescence measurements of integrated promoter constructs in *P. putida* KT2440 were performed with a Biolector (M2P Labs, Baesweiler, Germany) in 96 well plates (Greiner Bio-One) with a filling volume of 200 μL. Cultures were inoculated to an optical density at 600 nm of 0.1 for each strain from precultures cultivated at 30 °C at 300 rpm in 24 well System Duetz plates (Enzyscreen, Heemstede, The Netherlands) containing 1.5 mL of previously described minimal medium. The Biolector was set to 30°C, 900 rpm and humidity control of 85 %. Two internal filter modules of the device were used for online measurement. Fluorescence of GFP was measured at excitation wavelength at 488 nm and emission wavelength of 520 nm with gain 50. Biomass was determined at 620 nm with gain 40 as scattered light. Scattered light was correlated to OD600 with a dilution series of a stationary phase culture. Determination of promoter activity was done with Microsoft Excel by calculating the slope of GFP fluorescence to optical density during the exponential phase.

Determination of Transcript Levels by Quantitative Real Time PCR

Transcription levels of *msfGFP* was determined by quantitative real time PCR. RNA was isolated from chosen strains grown on minimal medium containing 20 mM glucose as sole carbon source in 24 well System Duetz plates at 30°C and 300 rpm (Hartmans et al., 1989). Biological duplicates of each strain were cultivated until an optical density of 1.0 was reached. One milliliter of cell cultures were harvested, supernatant discarded and the resulting pellet resuspended with 1 mL RNA*later*™ Stabilization Solution (ThermoFisher Scientific). Afterwards the cells were resuspended in 700 μL lyse solution (New England Biolabs, Monarch™ Total RNA Miniprep Kit) and transferred to bead beating tubes containing glass bead with a size of 0.5 mm (Zymo Research, Irvine, USA). Tubes were beaten for 1 min to destroy the cells (Mini-Beadbeater-16, Biospec Products, Bartlesville, USA). Cells debris were removed by centrifugation at 13.000 rpm for 2 min. The supernatant was transferred to an RNase-free tube and used for further works. RNA isolation from lysed samples followed the manual from the kit Monarch Total RNA Miniprep Kit (New England Biolabs). Elution of RNA was done with 50 μL RNase free water. RNA concentration was measured with a NanoDrop One (Thermo Scientific) at 260 nm. Samples were adjusted to a final RNA concentration of 280 ng in a total volume of 40 μL, dilution was done with RNase-free water. An additional DNase treatment was done by adding 5 μL DNaseI and 5 μL DNase I reaction buffer (New England Biolabs) to the RNA isolates. Digestion was done at 37°C for 10 min and DNase inactivation at 75°C for 10 min. For cDNA synthesis LunaScript RT SuperMix Kit (New England Biolabs) was performed as describe in the manual.

Determination of primer efficiencies was done with diluted cDNA from BG14f_80. cDNA was diluted 1:10, 1:20, 1:40, 1:80, 1:160, 1:320, and 1:640. 1.25 μL of each used for the qRT-PCR reaction. A total volume of 10 μL containing 5 μL Universal qPCR Master Mix (New England Biolabs), 0.25 μL of each oligonucleotide, 1.25 μL sample and 3.25 μL RNase-free water were used. We tested oligonucleotide combinations for the target gene *msfGFP* and housekeeping gene *rpoD* in a CFX Connect Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, USA) using a protocol described in the manual of Universal qPCR Master Mix (New England Biolabs). CFX Manager software (Bio-Rad Laboratories, Hercules, USA) was used for the calculation of resulting primer efficiencies and an online tool was used to calculated the amplification factor (https://www.thermofisher.com/de/de/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/qpcr-efficiency-calculator.html). Tested oligonucleotide combinations for *msfGFP* achieved a value of 1.97 and *rpoD* of 2.02 (Udvardi et al., 2008).

Each cDNA sample was diluted 1:10 and analyzed as technical duplicate. Volumes for each reaction are described above. As negative control the same amount of water was added to the reaction instead of cDNA. Examination from resulting Ct values was done with Microsoft Excel and a ΔCt method was applied (Pfaffl, 2001). To exclude genomic DNA in the samples, isolated RNA was used in separate reactions with oligonucleotides for *rpoD*.

Statistics

Each promoter construct was characterized in 2–3 independent transformations performed on different days. Three clones from each transformant were tested in a Biolector to determine

promoter activities, yielding a total of 6–9 biological replicates. For each construct the mean and standard error of the mean was calculated from these combined biological replicates. Significance of difference of the activity of constructs with different spacer lengths was analyzed by one-way ANOVA with Turkey's *post-hoc* comparison. Coefficient of variation, determined by dividing the absolute difference of the predicted and experimental value by the experimental value, was used to compare the accuracy of prediction of stacked promoter activities.

RESULTS AND DISCUSSION

Identification of the Optimal Distance Between Two Promoters

For the characterization of stacked promoters, we used a mini Tn7 transposon, which integrates as single copy into the *attTn7* site downstream of the *glmS* gene in the genome of *P. putida* KT2440 (Bagdasarian et al., 1981; Choi et al., 2005; Zobel et al., 2015). The transposon is designed to characterize promoters in a reliable and reproducible manner, featuring a BCD2 element to reduce GOI-based expression variability (Mutalik et al., 2013), an *msfGFP* (Landgraf, 2012) gene as reporter, and two flanking terminators to minimize genomic read-through (**Figure 1**; Zobel et al., 2015).

In order to determine the optimal distance between two promoters, we stacked the 14f and 14g promoters from a previously published synthetic promoter library (Zobel et al., 2015) with spacer sequences with increasing length from 10 to 100 bp by extension at the 3'-end. The promoters are referred to solely by their SEVA code (Zobel et al., 2015; 14a-g, with a being the weakest and g being the strongest) for ease of reference. The spacer was randomly generated (http://www.faculty.ucr.edu/~mmaduro/random.htm) and manually curated for unwanted restriction sites as well as putative ribosome binding sites, −35, and −10 like sequences, which could disturb the analysis due to intrinsic activity. The spacer was created with a GC content of 62%, similar to the genomic average of *P. putida* KT2440 (Nelson et al., 2002).

A promoterless construct BG and wildtype *P. putida* KT2440 were used as negative controls. As positive controls we used the single calibrated promoters described in Zobel et al. (2015), including P_{em7} (Martínez-García et al., 2015) which reaches half of the activity of promoter 14g. With short spacer sequences of <35 bp, the activity of the stacked promoter is lower than that of either of the single promoters (**Figure 1**). This result is likely caused by steric hindrance of the RNA polymerase holoenzyme, since the combined sigma factor and RNA polymerase cover around 80 bp upstream of the promoter (Schmitz and Galas, 1979). A spacer length of 40–70 and 90–100 bp resulted in activities of comparable strength. A significantly higher activity was observed for 14f_80_14g with an 80 bp spacer compared to all other spacer lengths (one-way ANOVA with Turkey's *post-hoc* comparison).

To exclude that this 80 bp spacer is an outlier due to possible activating sequences, the experiment was repeated with a reverse complement version of the spacer (80i) and a new, independently generated spacer sequence (80new, **Supplementary Table 2**). All three 80 bp control spacers led to comparable activities, indicating that this distance between two promoters is promoting additive activity of the two promoters. It is interesting to note that the spacer length of 80 bp matches the sequence covered by the RNA polymerase holoenzyme (Schmitz and Galas, 1979), although this correlation should not be confused with causation. Promoterless controls with only 80i and 80new also show no activity, excluding any intrinsic activity from the spacers themselves (**Figure 1**).

With the 80i spacer a cumulative effect occurred, with the total output of the stacked promoters being higher than the individual activities. However, the output was much lower than the sum of the two individual promoters for each tested spacer length. For further characterization, we used the 80i spacer since it enabled the highest promoter activity.

Characterization of Context Effects on Stacked Promoters

We hypothesized two possible ways how these stacked promoters are affected. The primary hypothesis is that of an effect of the spacer on the promoter. The alternative hypothesis is a mutual influence of one promoter on the other (Callen et al., 2004; Shearwin et al., 2005). To test these hypotheses, we constructed 14 different stacked promoter combinations and 12 controls to determine the influence of the 80i spacer on single promoter activities. Following the rules provided by SEVA (Martínez-García et al., 2015), the promoter is integrated between restriction sites *Pac*I and *Avr*II. The spacer is an additional sequence in the probe vector published by Zobel et al. (2015) and is not replacing any sequences from the original construct. The constructs are named according to their composition, i.e., in 14f_80i promoter 14f is cloned upstream of the 80i spacer. After genomic integration of the Tn7 transposon all strains were characterized in a BioLector system (**Figure 2**).

The single promoter controls without spacer reached activities that are comparable to those initially described by Zobel et al. (2015). In contrast, single promoter combined with the 80i spacer, either upstream or downstream, were strongly affected in their activity (**Figure 2**). With downstream placement of the spacer, all promoters were negatively affected, with decreases up to 70% for 14f_80i. In contrast, no clear trend could be discerned with upstream placement of the spacer, with most combinations having decreased activities up to 28%, but 80i_14c gained 12 % and 80i_14d even 50% activity.

These results show that the spacer has a drastic effect on all tested promoters. This is in spite of the fact that the spacer itself doesn't display any promoter activity (**Figure 1**), nor does it contain any discernible sequences that might affect promoter activity, such as AT-rich UP elements (Estrem et al., 1998). In addition, up- and downstream effects of the spacer are unpredictable. Most of the single promoters show a decreased activity when combined with the spacer, which could be explained by missing upstream activating elements potentially present in the original construct such as the AT-rich

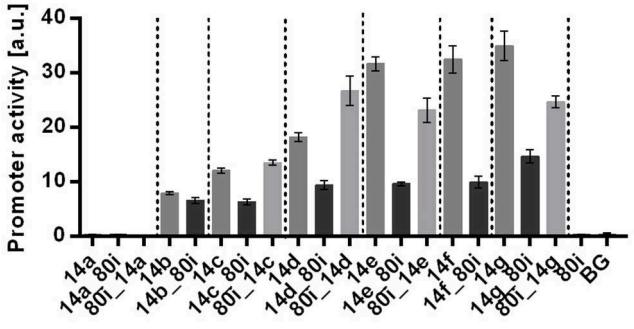


FIGURE 2 | Characterization of context-depended promoter activities using the 80i spacer. Promoters with both up- and downstream spacer were tested. Shown are promoter activities for the original promoters (dark bars) from Zobel et al. (2015) and promoter-spacer (14x_80i) as well as spacer-promoter (80i_14x) combinations (gray bars), where x stands for promoter 14a to 14g. All constructs were genomically integrated in *P. putida* KT2440. Strains were cultured in a BioLector in minimal medium with 20 mM glucose in a 96 well plate Identical strains from at least two different transformations were tested, with three biological replicates each. Vertical dotted lines are separating individual sets. Error bars indicate the standard error of the mean ($n > 6$).

*Pac*I restriction site. While no consistent correlation between spacer position and promoter activity is discernible, the results do confirm the primary hypothesis that the activity of the promoters is affected by the spacer.

To further test if, besides the effect of the spacer, the stacked promoters also affect each other, all seven calibrated promoters were stacked with 14g at the second position. As additional control, the reverse-order combination 14g_80i_14a was also included. As expected from the abovementioned results, all of these combinations led to much lower activities than expected from the sum of the individual promoters disregarding context-effects (**Figure 3**). Interestingly, the combinations 14a_80i_14g and 14g_80i_14a, which only differ in the order of the promoters, reached completely different activities. During the cloning of these stacks, a triple 'ffg' promoter consisting of two 14f and one 14g sequences separated by two spacers was accidentally created, in which the second 14f is shorter by two nucleotides between the −35 and −10 elements. Deriving sequence-function relationships from this promoter would be too complex, but the fact that it is around 45% stronger than the strongest promoter combination 14g_80i_14g makes it useful in applications where very high expression is needed (Lenzen et al., 2019; Bator et al., 2020) (**Supplementary Figure 1**).

When comparing the activities of these stacks to the single promoter-spacer controls above, it becomes apparent that the immediate context of single promoters is the major determinant for the prediction of promoter activity. The sum of the single promoter activities greatly overestimates the activities of stacked promoters by as much as 140% for the 14g_80i_14a combination (**Figure 3**). In contrast, the sum of the context-specific controls provides a much more accurate prediction, i.e., 14g_80i + 80i_14a = 14g_80i_14a. In this case, the coefficient of variance between context-specific prediction and experimental values is lower than 15% for all tested combinations. This strongly

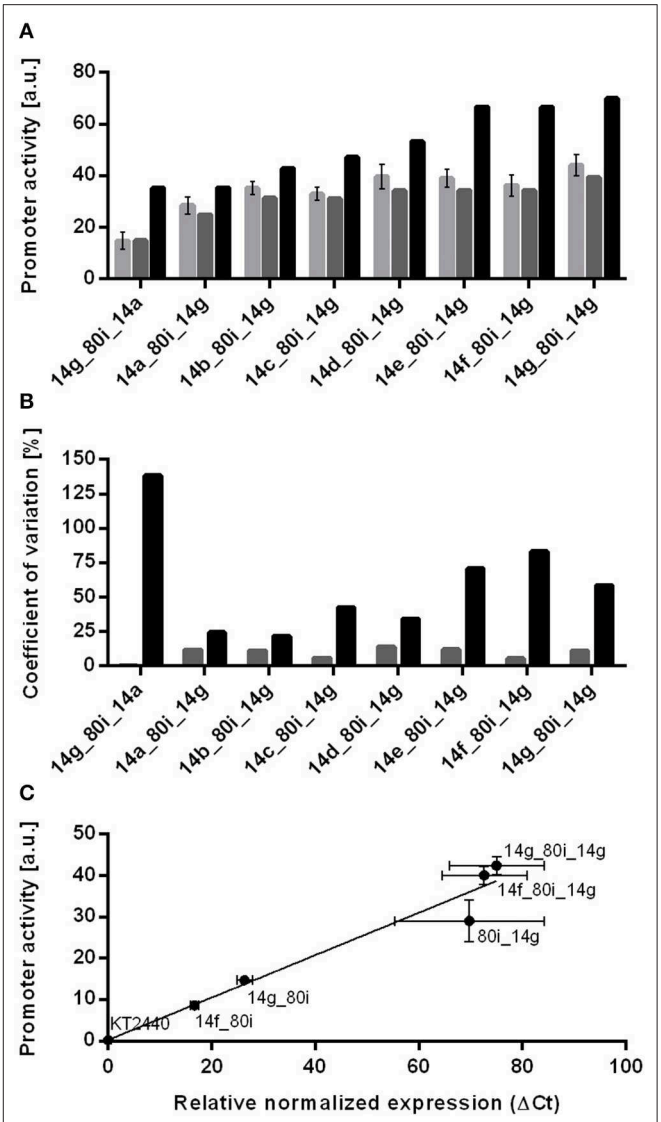


FIGURE 3 | Comparison of experimental values, context-specific and context-unspecific prediction of promoter activities for stacked promoter with the 80i spacer separating promoters 14a to 14g. **(A)** Shown are determined promoter activities for stacked promoters (light gray bars), context-specific activities calculated with context-depended values (gray bars) and context-unspecific activities using original promoter activities (black bars). All constructs were genomically integrated into *P. putida* KT2440. Cultivation was done in a BioLector in minimal medium with 20 mM glucose in a 96 well plate. Identical strains from at least three different transformations were tested, with three biological replicates each. Error bars indicate the standard error of the mean ($n > 9$). **(B)** Coefficient of variation (CV) of context-specific (light gray bars) and context-unspecific (black bars) prediction of resulting promoter activities. **(C)** Plot of *msfGFP* transcription levels normalized to *rpoD* determined by quantitative real time PCR (qRT-PCR) and promoter activities from single promoter controls and stacked promoters. Wild type *P. putida* KT2440 was used as negative control. All constructs are genomically integrated into the genome of *P. putida* KT2440. Cultivations to determine promoter activities were done in a BioLector in minimal medium with 20 mM glucose in a 96 well plate. Cultivation to determine transcription levels was done in 24 well System Duetz plates containing minimal medium with 20 mM glucose. Identical strains from at least two different transformations were tested, with three biological replicates each. Error bars indicate the standard error of the mean ($n > 6$).

suggests that, once the direct context of the individual promoters is sufficiently taken into account, the stacked promoters don't affect each other's activity.

Beyond having different promoter contexts, the abovementioned constructs also generate different 5′-terminal mRNA ends, which may cause differences in mRNA stability or translation initiation rates. In order to minimize the effect of these differences, a bicistronic design (Mutalik et al., 2013) was included in the reporter construct. To verify whether the altered expression of context-affected constructs is caused by increased transcription, we performed quantitative real time PCR (qRT-PCR) on selected constructs. Determined transcript levels correlate well with promoter activities derived from fluorescence measurements ($r^2 = 0.95$, **Figure 3**), confirming that the spacer influences transcription, rather than translation. Attempts to determine the relative contributions of the first or second promoter by qRT-PCR were inconclusive. In principle, stacked promoters generate two overlapping transcripts of different length, which might be distinguished with different primers pairs. However, longer transcripts show a shift in Ct value compared to shorter amplicons, and suitable primer pairs for similar lengths could not be found (Debode et al., 2017).

Using an SNP Promoter Library for Stacked Promoters

A change in context greatly affects promoter activitiy, and there are large quantitative differences for each tested promoter-spacer combination. Given that the main variable between these constructs is the promoter sequence, this might be due to specific DNA-DNA interactions between promoter and spacer, which influence promoter activity, or RNA-RNA interactions, which affect RNA stability. To further investigate the sequence-activity relationship, we generated a single nucleotide polymorphism (SNP) library based on promoter 14g (**Figure 4**). Such a library yields promoters with very similar sequences, but large differences in activity. If the variability of the impact of the spacer on the activity is indeed caused by DNA-DNA or RNA-RNA interactions, using promoters with more similar sequences can be expected to reduce this variability. The library contained 90 different promoter sequences, which were generated in 30 PCR reaction with one degenerate nucleotide in the sequence.

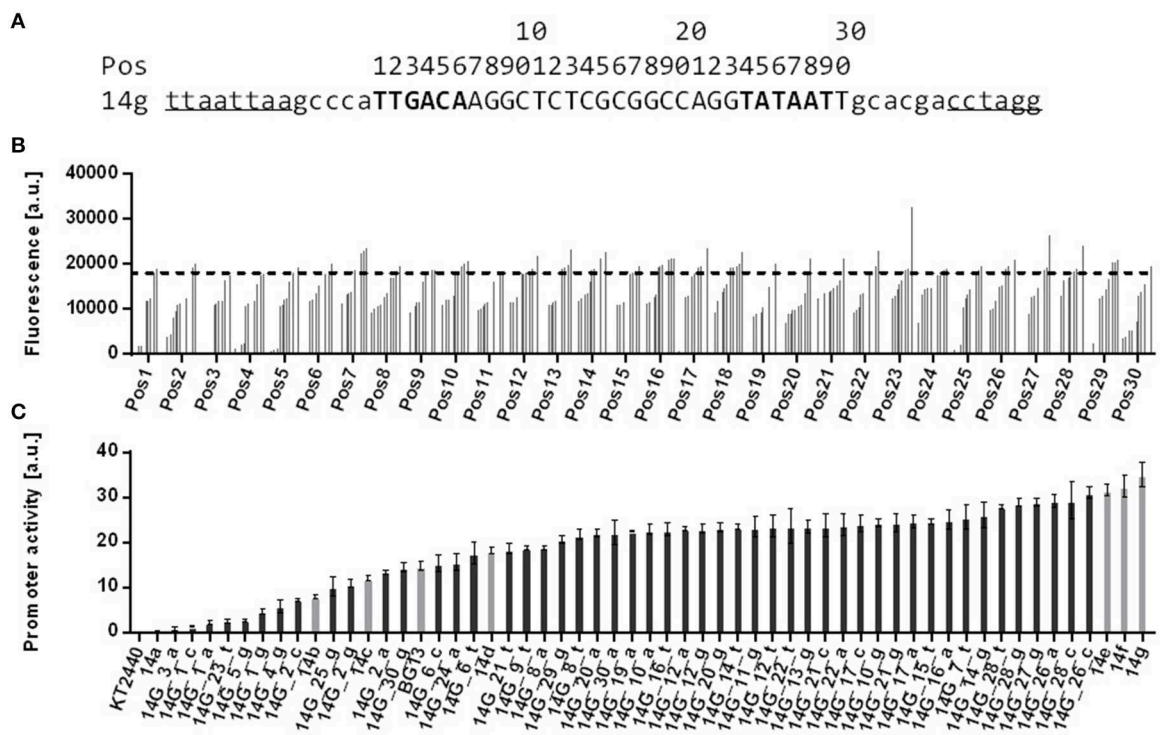


FIGURE 4 | Comparison of screened and characterized promoter sequences based on 14g with single nucleotide polymorphisms. **(A)** Original promoter sequence of 14g with highlighted−35 and−10 elements (in bold). Restriction sites *Pac*I and *Avr*II are underlined and positions of the modified core promoter sequence are given with numbers above the sequence. **(B)** msfGFP fluorescence of *E. coli* PIR2 bearing plasmid pBG14g with degenerate bases at 30 positions along the core promoter sequence. Changed position is shown on the x axis. Determined values are ranked by fluorescence intensity of 14 strains tested for each position. Strains were cultivated in 96 well System Duetz plates with LB medium supplemented with 50 mg L^{-1} kanamycin. Fluorescence and optical density were measured with a plate reader. The dotted line indicates the promoter activity of the 14g control. **(C)** Chosen SNP promoter constructs were genomically integrated into *P. putida* KT2440. Strains are named 14G_##n, whereas ## stands for the position in the promoter sequence and n for a nucleotide (A, C, G or T). Cultivation was done in a BioLector in minimal medium with 20 mM glucose in 96 well plates. Identical strains from at least two different transformations were tested, with three biological replicates each. Error bars indicate the standard error of the mean ($n > 6$).

Changes of the core promoter sequence were inserted within the −35 element (position 1–6), in the interspaced region (position 7–23, 30), and in the −10 element (position 24–29). For each position four different promoters can occur, of which one will correspond to the original 14g sequence. Initial screening of mini libraries (14 clones each) of these low degeneracy promoters was performed in plasmid-bearing *E. coli* PIR2 strains by analyzing msf GFP fluorescence. Aberrations compared to original pBG14g-bearing *E. coli* PIR2 strains were recognized (**Figure 4**). For nearly each position clones were found with either a higher or lower fluorescence signal than the original pBG14g plasmid. Variations in the interspaced region generally had a lower effect on expression strength, while changing single nucleotides in the −35 and −10 consensus sequences yielded more clones with decreased promoter activity, as expected (Lodge et al., 1990; McLean et al., 1997). We therefore focused further characterization on these elements in order to obtain a set of promoters with a range of activity that is comparable to the previously described calibrated promoter library from Zobel et al. (2015) (**Figure 4**).

After initial screening of positions in the SNP promoter sequences, we selected three such positions within the −35 or −10 elements for further characterization. Introducing a degenerate base in these three positions yields nine different promoters with a good spread of activity, and these promoters were combined with the 80i spacer in the downstream position (**Figure 5**). Changing position 26 in the −10 sequence resulted in a slightly decreased activity, while changes in positions 1 and 2 in the −35 sequence yielded larger decreases, which is in accordance with Lodge et al. (1990).

We have seen that small changes in the 14g promoter sequence can strategically affect activity in a mini-promoter-library (**Supplementary Table 10**). In spite of the relative uniformity of the promoter sequences in this library, combination of these promoters with the 80i spacer again strongly affected the activities with both in- and decreases up to 66 %. The reduced sequence variability did not reduce the quantitative variability of the spacer effect compared to the CalPro library from Zobel et al. (2015). Both have a high coefficient of variation of 40% for the SNP library and 25% for the CalPro library (**Figure 5**). This strongly suggests that the promoter sequence *per se* does not cause the context-dependent effect, suggesting that other factors such as the varying transcription-initiation rates are in play.

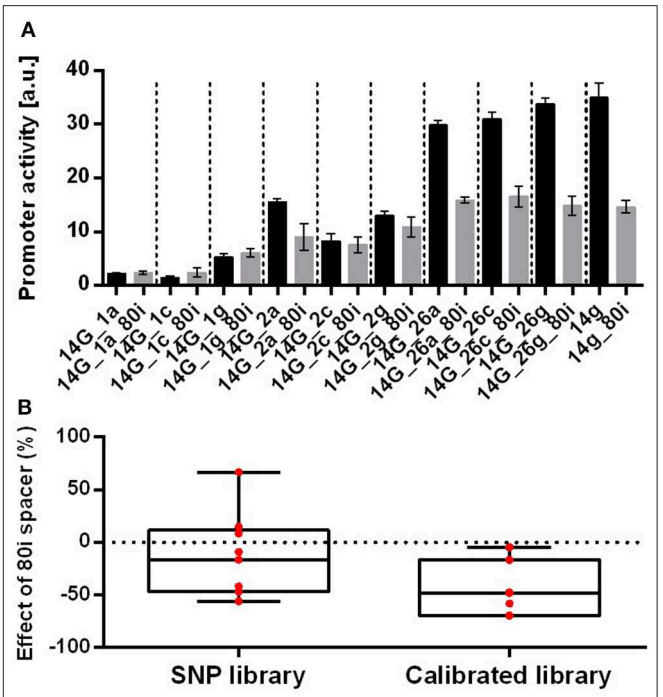


FIGURE 5 | Characterization of the effect of the 80i spacer on a single nucleotide exchange promoter library based on promoter 14g genomically integrated in *P. putida* KT2440. **(A)** Promoter activities derived from GFP fluorescence analysis of SNP promoters with (gray bars) and without (black bars) downstream 80i spacer. Cultivation was done in a BioLector in minimal medium with 20 mM glucose in 96 well plates. Identical strains from at least two different transformations were tested, with three biological replicates each. Vertical dotted lines are separating individual sets. Error bars indicate the standard error of the mean ($n > 6$). **(B)** Box and whiskers plot of the relative effect of the downstream 80i spacer on promoter activities of the SNP library and the original calibrated promoter library from Zobel et al. (2015). The effect of the spacer is calculated as the % change of the promoter with downstream 80i sequence compared to the original promoter. Data points are indicated in red.

CONCLUSION

In this work, we aimed to increase the composability and predictability of synthetic promoters by investigating the effects of differing contexts on their activity. In the combination of two promoters, the length of the spacer region is crucial for reaching higher and cumulative effects, with 80 bp being the optimal. Even though the spacer sequence has no intrinsic activity, it strongly and unpredictably influences the activity of promoters by up- and downstream effects. By accounting for this influence, the activity of two stacked promoters can be accurately predicted with coefficients of variance below 15%. A strong reduction of sequence variability was achieved using an SNP library, but this reduction did not reduce the quantitative variability of the spacer effect. This strongly indicates that nucleotide-nucleotide interactions between promoter and spacer do not play a prominent role. Clearly, context-specific effects of synthetic promoters are not yet fully understood in *Pseudomonas*. Although the semi-empirical approach for prediction of stacked promoter activities provides an accurate workaround to this, further work is needed to understand the fundamental interaction of genetic elements and their surroundings.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

NW and SK designed the experiments. SK performed all molecular engineering and experiments, prepared

figures and wrote the manuscript. NW supervised the study and edited the manuscript. LB advised on all experiments, analyzed and discussed data, and edited the manuscript.

FUNDING

We gratefully acknowledge funding for SK from the European Union's Horizon 2020 research and innovation program under grant agreement no. 633962 for the project P4SB. NW was supported by the German Research Foundation (DFG) through the Emmy Noether program (WI 4255/1-1). The laboratory of LB was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00551/full#supplementary-material

REFERENCES

Aparicio, T., de Lorenzo, V., and Martinez-Garcia, E. (2018). CRISPR/Cas9-based counterselection boosts recombineering efficiency in *Pseudomonas putida*. *Biotechnol. J.* 13:e1700161. doi: 10.1002/biot.201700161

Bagdasarian, M., Lurz, R., Ruckert, B., Franklin, F., Bagdasarian, M., Frey, J., et al. (1981). Specific-purpose plasmid cloning vectors. II. broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. *Gene* 16, 237–247. doi: 10.1016/0378-1119(81)90080-9

Bator, I., Wittgens, A., Rosenau, F., Tiso, T., and Blank, L. M. (2020). Comparison of three xylose pathways in *Pseudomonas putida* KT2440 for the synthesis of valuable products. *Front. Bioeng. Biotechnol.* 7:480. doi: 10.3389/fbioe.2019.00480

Blank, L. M., Ionidis, G., Ebert, B. E., Bühler, B., and Schmid, A. (2008). Metabolic response of *Pseudomonas putida* during redox biocatalysis in the presence of a second octanol phase. *FEBS J.* 275, 5173–5190. doi: 10.1111/j.1742-4658.2008.06648.x

Boyer, H. W., and Roulland-Dussoix, D. (1969). A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J. Mol. Biol.* 41, 459–472. doi: 10.1016/0022-2836(69)90288-5

Callen, B. P., Shearwin, K. E., and Egan, J. B. (2004). Transcriptional interference between convergent promoters caused by elongation over the promoter. *Mol. Cell.* 14, 647–656. doi: 10.1016/j.molcel.2004.05.010

Choi, K. H., Gaynor, J. B., White, K. G., Lopez, C., Bosio, C. M., Karkhoff-Schweizer, R. R., et al. (2005). A Tn7-based broad-range bacterial cloning and expression system. *Nat. Methods* 2, 443–448. doi: 10.1038/nmeth765

Damron, F. H., McKenney, E. S., Barbier, M., Liechti, G. W., Schweizer, H. P., and Goldberg, J. B. (2013). Construction of mobilizable mini-Tn7 vectors for bioluminescent detection of gram-negative bacteria and single-copy promoter lux reporter analysis. *Appl. Environ. Microb.* 79, 4149–4153. doi: 10.1128/AEM.00640-13

de Lorenzo, V., Herrero, M., Jakubzik, U., and Timmis, K. N. (1990). Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative bacteria. *J. Bacteriol.* 172, 6568–6572. doi: 10.1128/JB.172.11.6568-6572.1990

Debode, F., Marien, A., Janssen, É., Bragard, C., and Berben, G. (2017). Influence of the amplicon length on real-time PCR results. *Biotechnol. Agron. Soc.* 21, 3–11. doi: 10.25518/1780-4507.13461

Dixon, R. (1984). Tandem promoters determine regulation of the *Klebsiella pneumonia* glutamine synthetase (gln) gene. *Nucleic Acids Res.* 12, 7811–7830. doi: 10.1093/nar/12.20.7811

Elmore, J. R., Furches, A., Wolff, G. N., Gorday, K., and Guss, A. M. (2017). Development of a high efficiency integration system and promoter library for rapid modification of *Pseudomonas putida* KT2440. *Metab. Eng. Commun.* 5, 1–8. doi: 10.1016/j.meteno.2017.04.001

Estrem, S. T., Gaal, T., Ross, W., and Gourse, R. L. (1998). Identification of an UP element consensus sequence for bacterial promoters. *Proc. Natl. Acad. Sci. U.S.A.* 95, 9761–9766. doi: 10.1073/pnas.95.17.9761

Franden, M. A., Jayakody, L. N., Li, W. J., Wagner, N. J., Cleveland, N. S., Michener, W. E., et al. (2018). Engineering *Pseudomonas putida* KT2440 for efficient ethylene glycol utilization. *Metab. Eng.* 48, 197–207. doi: 10.1016/j.ymben.2018.06.003

Friehs, K. (2004). Plasmid copy number and plasmid stability. *Adv. Biochem. Eng. Biotechnol.* 86, 47–82. doi: 10.1007/b12440

Gao, Y., Liu, C., Ding, Y., Sun, C., Zhang, R., Xian, M., et al. (2014). Development of genetically stable *Escherichia coli* strains for poly(3-hydroxypropionate) production. *PLoS ONE* 9:e97845. doi: 10.1371/journal.pone.0097845

Gilman, J., and Love, J. (2016). Synthetic promoter design for new microbial chassis. *Biochem. Soc. Trans.* 44, 731–737. doi: 10.1042/BST20160042

Gruber, T. M., and Gross, C. A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.* 57, 441–466. doi: 10.1146/annurev.micro.57.030502.090913

Hammer, K., Mijakovic, I., and Jensen, P. R. (2006). Synthetic promoter libraries-tuning of gene expression. *Trends Biotechnol.* 24, 53–55. doi: 10.1016/j.tibtech.2005.12.003

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166, 557–580. doi: 10.1016/S0022-2836(83)80284-8

Hartmans, S., Smits, J. P., van der Werf, M. J., Volkering, F., and de Bont, J. A. M. (1989). Metabolism of styrene oxide and 2-phenylethanol in the styrene-degrading *Xanthobacter* strain 124X. *Appl. Environ. Microb.* 55, 2850–2855. doi: 10.1128/AEM.55.11.2850-2855.1989

Heipieper, H. J., Neumann, G., Cornelissen, S., and Meinhardt, F. (2007). Solvent-tolerant bacteria for biotransformations in two-phase fermentation systems. *Appl. Microbiol. Biotechnol.* 74, 961–973. doi: 10.1007/s00253-006-0833-4

Herrero, M., de Lorenzo, V., and Timmis, K. N. (1990). Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. *J. Bacteriol.* 172, 6557–6567. doi: 10.1128/JB.172.11.6557-6567.1990

Isken, S., and de Bont, J. A. (1998). Bacteria tolerant to organic solvents. *Extremophiles* 2, 229–238. doi: 10.1007/s007920050065

Jahn, M., Vorpahl, C., Turkowsky, D., Lindmeyer, M., Buhler, B., Harms, H., et al. (2014). Accurate determination of plasmid copy number of flow-sorted cells using droplet digital PCR. *Anal. Chem.* 86, 5969–5976. doi: 10.1021/ac501118v

Jensen, P. R., and Hammer, K. (1998). Artificial promoters for metabolic optimization. *Biotechnol. Bioeng.* 58, 191–195. doi: 10.1002/(SICI)1097-0290(19980420)58:2/3<191::AID-BIT11>3.0.CO;2-G

Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L. A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* 31, 233–239. doi: 10.1038/nbt.2508

Jiménez, J. I., Miñambres, B., García, J. L., and Díaz, E. (2002). Genomic analysis of the aromatic catabolic pathways from *Pseudomonas putida* KT2440. *Environ. Microbiol.* 4, 824–841. doi: 10.1046/j.1462-2920.2002.00370.x

Kang, J-G., Hahn, M-Y., Roe, J-H., and Ishihama, A. (1997). Identification of sigma factors for growth phase-related promoter selectivity of RNA polymerases from *Streptomyces coelicolor* A3(2). *Nucleic Acids Res.* 25, 2566–2573. doi: 10.1093/nar/25.13.2566

Keen, N. T., Tamaki, S., Kobayashi, D., and Trollinger, D. (1988). Improved broad-host-range plasmids for DNA cloning in gram-negative bacteria. *Gene* 70, 191–197. doi: 10.1016/0378-1119(88)90117-5

Köhler, K. A. K., Blank, L. M., Frick, O., and Schmid, A. (2015). D-Xylose assimilation via the weimberg pathway by solvent tolerant

Pseudomonas taiwanensis VLB120. *Environ. Microbiol.* 17, 156–170. doi: 10.1111/1462-2920.12537

Koopman, F., Wierckx, N., de Winde, J. H., and Ruijssenaars, H. J. (2010). Efficient whole-cell biotransformation of 5-(hydroxymethyl)furfural into FDCA, 2,5-furandicarboxylic acid. *Bioresour. Technol.* 101, 6291–6296. doi: 10.1016/j.biortech.2010.03.050

Kusumawardhani, H., Hosseini, R., and de Winde, J. H. (2018). Solvent tolerance in bacteria: fulfilling the promise of the biotech era? *Trends Biotechnol.* 36, 1025–1039. doi: 10.1016/j.tibtech.2018.04.007

Lambertsen, L., Sternberg, C., and Molin, S. (2004). Mini-Tn7 transposons for site-specific tagging of bacteria with fluorescent proteins. *Environ. Microbiol.* 6, 726–732. doi: 10.1111/j.1462-2920.2004.00605.x

Landgraf, D. (2012). *Quantifying Localizations and Dynamics in Single Bacterial Cells*. Doctoral Dissertation, Harvard University, Cambridge, MA, USA.

Lenzen, C., Wynands, B., Otto, M., Bolzenius, J., Mennicken, P., Blank, L. M., et al. (2019). High-yield production of 4-hydroxybenzoate from glucose or glycerol by an engineered *Pseudomonas taiwanensis* VLB120. *Front. Bioeng. Biotechnol.* 7:130. doi: 10.3389/fbioe.2019.00130

Li, M., Wang, J., Geng, Y., Li, Y., Wang, Q., Liang, Q., et al. (2012). A strategy of gene overexpression based on tandem repetitive promoters in *Escherichia coli*. *Microb. Cell Fact.* 11:19. doi: 10.1186/1475-2859-11-19

Li, W. J., Jayakody, L. N., Franden, M. A., Wehrmann, M., Daun, T., Hauer, B., et al. (2019). Laboratory evolution reveals the metabolic and regulatory basis of ethylene glycol metabolism by *Pseudomonas putida* KT2440. *Environ. Microbiol.* 21, 3669–3682. doi: 10.1111/1462-2920.14703

Li, W. J., Narancic, T., Kenny, S. T., Niehoff, P. J., O'Connor, K. E., Blank, L. M., et al. (2020). Unraveling 1,4-butanediol metabolism in Pseudomonas putida KT2440. *Front. Microbiol.* 11:382. doi: 10.3389/fmicb.2020.00382

Lindmeyer, M., Jahn, M., Vorpahl, C., Müller, S., Schmid, A., and Bühler, B. (2015). Variability in subpopulation formation propagates into biocatalytic variability of engineered *Pseudomonas putida* strains. *Front. Microbiol.* 6:1042. doi: 10.3389/fmicb.2015.01042

Lodge, J., Williams, R., Bell, A., Chan, B., and Busby, S. (1990). Comparison of promoter activities in *Escherichia coli* and *Pseudomonas aeruginosa*: use of a new broad-host-range promoter-probe plasmid. *FEMS Microbiol. Lett.* 67, 221–225. doi: 10.1111/j.1574-6968.1990.tb13867.x

Martens, J. A., Laprade, L., and Winston, F. (2004). Intergenic transcription is required to repress the *Saccharomyces cerevisiae* SER3 gene. *Nature* 429, 571–574. doi: 10.1038/nature02538

Martínez-García, E., Aparicio, T., Goñi-Moreno, A., Fraile, S., and de Lorenzo, V. (2015). SEVA 2.0: an update of the standard European vector architecture for de-/re-construction of bacterial functionalities. *Nucleic Acids Res.* 43, D1183–D1189. doi: 10.1093/nar/gku1114

Martínez-García, E., and de Lorenzo, V. (2011). Engineering multiple genomic deletions in gram-negative bacteria: analysis of the multi-resistant antibiotic profile of *Pseudomonas putida* KT2440. *Environ. Microbiol.* 13, 2702–2716. doi: 10.1111/j.1462-2920.2011.02538.x

McLean, B. W., Wiseman, S. L., and Kropinski, A. M. (1997). Functional analysis of sigma-70 consensus promoters in *Pseudomonas aeruginosa* and *Escherichia coli*. *Can. J. Microbiol.* 43, 981–985. doi: 10.1139/m97-141

Meijnen, J. P., de Winde, J. H., and Ruijssenaars, H. J. (2011). Sustainable production of fine chemicals by the solvent-tolerant *Pseudomonas putida* S12 using lignocellulosic feedstock. *Int. Sugar J.* 113, 24–30.

Mutalik, V. K., Guimaraes, J. C., Cambray, G., Lam, C., Christoffersen, M. J., Mai, Q. A., et al. (2013). Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* 10, 354–360. doi: 10.1038/nmeth.2404

Nelson, K. E., Weinel, C., Paulsen, I. T., Dodson, R. J., Hilbert, H., Martins dos Santos, V. A., et al. (2002). Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ. Microbiol.* 4, 799–808. doi: 10.1046/j.1462-2920.2002.00366.x

Nikel, P. I., and de Lorenzo, V. (2013). Implantation of unmarked regulatory and metabolic modules in gram-negative bacteria with specialised mini-transposon delivery vectors. *J. Biotechnol.* 163, 143–154. doi: 10.1016/j.jbiotec.2012.05.002

Nikel, P. I., Kim, J., and de Lorenzo, V. (2014). Metabolic and regulatory rearrangements underlying glycerol metabolism in *Pseudomonas putida* KT2440. *Environ. Microbiol.* 16, 239–254. doi: 10.1111/1462-2920.12224

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Res.* 29:e45. doi: 10.1093/nar/29.9.e45

Potvin, E., Sanschagrin, F., and Levesque, R. C. (2008). Sigma factors in *Pseudomonas aeruginosa*. *FEMS Microbiol. Rev.* 32, 38–55. doi: 10.1111/j.1574-6976.2007.00092.x

Ramos, J. L., Duque, E., Gallegos, M. T., Godoy, P., Ramos-Gonzalez, M. I., Rojas, A., et al. (2002). Mechanisms of solvent tolerance in gram-negative bacteria. *Annu. Rev. Microbiol.* 56, 743–768. doi: 10.1146/annurev.micro.56.012302.161038

Rud, I., Jensen, P. R., Naterstad, K., and Axelsson, L. (2006). A synthetic promoter library for constitutive gene expression in *Lactobacillus plantarum*. *Microbiology* 152, 1011–1019. doi: 10.1099/mic.0.28599-0

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

San Millan, A., and MacLean, R. C. (2017). Fitness costs of plasmids: a limit to plasmid transmission. *Microbiol. Spectr.* 5, 1–12. doi: 10.1128/microbiolspec.MTBP-0016-2017

Sánchez-Pascuala, A., Fernández-Cabezón, L., de Lorenzo, V., and Nikel, P. I. (2019). Functional implementation of a linear glycolysis for sugar catabolism in *Pseudomonas putida*. *Metab. Eng.* 54, 200–211. doi: 10.1016/j.ymben.2019.04.005

Schmitz, A., and Galas, D. J. (1979). The interaction of RNA polymerase and lac repressor with the lac control region. *Nucleic Acids Res.* 6, 111–137. doi: 10.1093/nar/6.1.111

Schweizer, H. P. (1992). Allelic exchange in *Pseudomonas aeruginosa* using novel ColE1-type vectors and a family of cassettes containing a portable oriT and the counter-selectable *Bacillus subtilis* sacB marker. *Mol. Microbiol.* 6, 1195–1204. doi: 10.1111/j.1365-2958.1992.tb01558.x

Shearwin, K. E., Callen, B. P., and Egan, J. B. (2005). Transcriptional interference-a crash course. *Trends Genet.* 21, 339–345. doi: 10.1016/j.tig.2005.04.009

Shen, X., Wang, Z., Huang, X., Hu, H., Wang, W., and Zhang, X. (2017). Developing genome-reduced *Pseudomonas chlororaphis* strains for the production of secondary metabolites. *BMC Genomics* 18:715. doi: 10.1186/s12864-017-4127-2

Silva-Rocha, R., and de Lorenzo, V. (2014). Chromosomal integration of transcriptional fusions. *Methods Mol. Biol.* 1149, 479–489. doi: 10.1007/978-1-4939-0473-0_37

Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., de Las Heras, A., et al. (2013). The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.* 41, D666–D675. doi: 10.1093/nar/gks1119

Solem, C., Koebmann, B., Yang, F., and Jensen, P. R. (2007). The las enzymes control pyruvate metabolism in *Lactococcus lactis* during growth on maltose. *J. Bacteriol.* 189, 6727–6730. doi: 10.1128/JB.00902-07

Stanton, B. C., Nielsen, A. A. K., Tamsir, A., Clancy, K., Peterson, T., and Voigt, C. A. (2013). Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nat. Chem. Biol.* 10, 99–105. doi: 10.1038/nchembio.1411

Sun, Z., Ramsay, J. A., Guay, M., and Ramsay, B. A. (2007). Carbon-limited fed-batch production of medium-chain-length polyhydroxyalkanoates from nonanoic acid by *Pseudomonas putida* KT2440. *Appl. Microbiol. Biot.* 74, 69–77. doi: 10.1007/s00253-006-0655-4

Tamsir, A., Tabor, J. J., and Voigt, C. A. (2011). Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'. *Nature* 469, 212–215. doi: 10.1038/nature09565

Tiso, T., Wierckx, N., Blank, L., and Grunwald, P. (2014). *Non-pathogenic Pseudomonas as Platform for Industrial Biocatalysis.* Singapore: Pan Stanford Publishing.

Udvardi, M. K., Czechowski, T., and Scheible, W-R. (2008). Eleven golden rules of quantitative RT-PCR. *Plant Cell* 20, 1736–1737. doi: 10.1105/tpc.108.061143

Wang, Q., and Nomura, C. T. (2010). Monitoring differences in gene expression levels and polyhydroxyalkanoate (PHA) production in *Pseudomonas putida* KT2440 grown on different carbon sources. *J. Biosci. Bioeng.* 110, 653–659. doi: 10.1016/j.jbiosc.2010.08.001

Wierckx, N. J. P., Ballerstedt, H., De Bont, J. A. M., and Wery, J. (2005). Engineering of solvent-tolerant *Pseudomonas putida* S12 for

bioproduction of phenol from glucose. *Appl. Environ. Microb.* 71, 8221–8227. doi: 10.1128/AEM.71.12.8221-8227.2005

Wynands, B., Lenzen, C., Otto, M., Koch, F., Blank, L. M., and Wierckx, N. (2018). Metabolic engineering of *Pseudomonas taiwanensis* VLB120 with minimal genomic modifications for high-yield phenol production. *Metab. Eng.* 47, 121–133. doi: 10.1016/j.ymben.2018.03.011

Zobel, S., Benedetti, I., Eisenbach, L., de Lorenzo, V., Wierckx, N., and Blank, L. M. (2015). Tn7-based device for calibrated heterologous gene expression in *Pseudomonas putida*. *ACS Synth. Biol.* 4, 1341–1351. doi: 10.1021/acssynbio.5b00058

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

BRIEF RESEARCH REPORT
published: 26 June 2020
doi: 10.3389/fbioe.2020.00698

CRISPR-Cas12a-Assisted Genome Editing in *Amycolatopsis mediterranei*

Yajuan Zhou[1,2,3†], Xinqiang Liu[1,2†], Jiacheng Wu[1,2,3], Guoping Zhao[1,4] and Jin Wang[5]*

[1] CAS Key Laboratory of Synthetic Biology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China, [2] University of Chinese Academy of Sciences, Beijing, China, [3] School of Life Sciences and Technology, Shanghai Tech University, Shanghai, China, [4] Department of Microbiology and Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, Hong Kong, [5] College of Life Sciences, Shanghai Normal University, Shanghai, China

Amycolatopsis mediterranei U32 is an industrial producer of rifamycin SV, whose derivatives have long been the first-line antimycobacterial drugs. In order to perform genetic modification in this important industrial strain, a lot of efforts have been made in the past decades and a homologous recombination-based method was successfully developed in our laboratory, which, however, requires the employment of an antibiotic resistance gene for positive selection and did not support convenient markerless gene deletion. Here in this study, the clustered regularly interspaced short palindromic repeat (CRISPR) system was employed to establish a genome editing system in *A. mediterranei* U32. Specifically, the *Francisella tularensis* subsp. *novicida Cas12a* (*FnCas12a*) gene was first integrated into the U32 genome to generate target-specific double-stranded DNA (dsDNA) breaks (DSBs) under the guidance of CRISPR RNAs (crRNAs). Then, the DSBs could be repaired by either the non-homologous DNA end-joining (NHEJ) system or the homology-directed repair (HDR) pathway, generating inaccurate or accurate mutations in target genes, respectively. Besides of *A. mediterranei*, the present work may also shed light on the development of CRISPR-assisted genome editing systems in other species of the *Amycolatopsis* genus.

Keywords: *Amycolatopsis mediterranei*, CRISPR-Cas12a, genome editing, NHEJ, HDR

OPEN ACCESS

Edited by:
Jiazhang Lian,
Zhejiang University, China

Reviewed by:
Feng-Qing Wang,
East China University of Science and Technology, China
Zehua Bao,
Boston University, United States
John Van Der Oost,
Wageningen University and Research, Netherlands

****Correspondence:***
Jin Wang
wangj01@hotmail.com

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *13 March 2020*
Accepted: *03 June 2020*
Published: *26 June 2020*

Citation:
Zhou Y, Liu X, Wu J, Zhao G and Wang J (2020) CRISPR-Cas12a-Assisted Genome Editing in Amycolatopsis mediterranei. Front. Bioeng. Biotechnol. 8:698. doi: 10.3389/fbioe.2020.00698

INTRODUCTION

Amycolatopsis mediterranei U32 is an industrial strain for production of rifamycin SV (Zhao et al., 2010), the first-line drug for anti-mycobacterial therapy till now (Rothstein, 2016). Due to the great importance of rifamycin, extensive efforts such as optimization of the fermentation conditions had been made to improve the yield of the antibiotics in the last century (Jiao et al., 1979; Lee et al., 1983; Mejia et al., 1998). Later, to facilitate the study of rifamycin biosynthesis as well as the molecular bioengineering of the producer, a genetic manipulation method based on native homologous recombination was developed for gene knockout in *A. mediterranei* (Ding et al., 2003). However, due to the relatively low efficiency of both DNA transformation and homologous recombination in *A. mediterranei*, an antibiotic cassette is usually employed to replace the target gene and the transformants are grown under antibiotic selection. To remove the antibiotic resistance cassette in the knockout mutant, the cassette should be flanked by site-specific recombination sequences such as the loxP sites or homologous arms, and a second cross-over recombination event is required. However, due to the

relatively low genetic engineering efficiency, there are very few reports of successful construction of a markerless mutant in *A. mediterranei*. What is worse, since there are only a limited number of antibiotics applicable in *A. mediterranei*, it is difficult to perform continuous genetic engineering operations. Therefore, although the *Amycolatopsis* genus is well known to produce a huge diversity of secondary metabolites (Xu et al., 2014; Kumari et al., 2016; Adamek et al., 2018), the lack of efficient genome editing technology has severely impeded the research progress in this genus.

The clustered regularly interspaced short palindromic repeat (CRISPR) system is an adaptive immune system in bacteria and archaea (Horvath and Barrangou, 2010; Jinek et al., 2012; Mohanraju et al., 2016), where the CRISPR-associated (Cas) protein complex utilizes guide RNA for specific recognition, binding, and cutting of target nucleic acids with proper protospacer adjacent motifs (PAM) (Jinek et al., 2012). The CRISPR systems can be divided into class 1 and class 2 (Koonin et al., 2017), where the crRNA ribonucleoprotein (crRNP) effector of the class 1 system complexes are composed of multiple Cas proteins as subunits (Makarova et al., 2011, 2017a), whereas the class 2 system crRNP complexes contain single Cas protein such as the types II, V, and VI Cas proteins (Makarova et al., 2017b). With CRISPR-Cas-assisted accurate cleavage in target DNA sequences and thus introducing double-stranded DNA (dsDNA) breaks (DSBs), the genome engineering efficiency can be greatly improved. Up to date, both the type II Cas9 system and type V Cas12a system have been widely applied in genome editing in a large number of species (Cong et al., 2013; Jiang et al., 2014, 2015; Cobb et al., 2015; Huang et al., 2015; Matsu-Ura et al., 2015; Low et al., 2016; Jia et al., 2017; Harrison and Hart, 2018; Hu et al., 2019). Compared to Cas9, Cas12a has several distinct features, including the preference of T-rich PAM sequences and the staggered cleavage pattern against target dsDNA (Zetsche et al., 2015; Yamano et al., 2016). Besides, unlike Cas9, Cas12a only requires the CRISPR RNA (crRNA) but not the *trans*-activating RNA (tracrRNA), and is able to mature precursor crRNAs, thereby enabling Cas12a in multiple gene editing and regulation with much convenience (Fonfara et al., 2016; Zetsche et al., 2017).

Bacteria have evolved two mechanisms to efficiently repair DSB damage, including both homology-directed repairing (HDR) (Sung and Klein, 2006) and non-homologous DNA end-joining (NHEJ) (Lieber, 2010). Combined with the CRISPR system, HDR provides accurate and markerless target gene deletion, mutation, and insertion of foreign DNA sequences (Cobb et al., 2015). Alternatively, in some bacterial species such as *Mycobacterium smegmatis*, DSB can be repaired by the NHEJ system, which comprises an ATP-dependent DNA ligase and a Ku protein (Wright et al., 2017; Zheng et al., 2017). Unlike HDR, the NHEJ system does not require homologous DNA sequences for recombination, but directly joins the breaks, facilitating convenient gene deletion and insertion (Babynin, 2007). However, as the NHEJ repair may introduce errors at the joining site, it is inappropriate for accurate gene editing.

Here in this study, we successfully established the CRISPR-Cas12a-based genome editing system in *A. mediterranei* U32. We first demonstrated the existence of the NHEJ system in U32, and then combined NHEJ with Cas12a to construct site-specific markerless gene deletion mutants. Moreover, we also used the endogenous HDR to repair the Cas12a-introduced DSBs, facilitating efficient genome editing in U32.

METHODOLOGY

Strains, Media, and Growth Conditions

Strains and plasmids used in this study are listed in **Supplementary Table S1**. *Escherichia coli* DH10B was used for DNA cloning and was grown at 37°C in LB medium. *A. mediterranei* U32 was grown at 30°C in Bennet medium (yeast extract, 1 g/l; glycerol, 10 g/l; glucose, 10 g/l; beef powder, 1 g/l; tryptone, 2 g/l; and agar, 15 g/l; pH 7.0). To prepare the competent cells, *A. mediterranei* strains were cultured at 30°C in MYM medium (yeast extract, 3 g/l; malt extract, 3 g/l; peptone, 5 g/l; glucose, 10 g/l; sucrose, 170 g/l; KNO_3, 6 g/l; glycine, 10 g/l; $CaCl_2{\cdot}2H_2O$, 0.735 g/l; $MgCl_2{\cdot}6H_2O$, 1.015 g/l; pH 7.0). For analyses of the growth phenotypes, *A. mediterranei* strains were grown in minimal medium (4% glucose, 0.05% NaCl, 0.2% K_2HPO_4, 0.1% $MgSO_4$, 0.001% $FeSO_4{\cdot}7H_2O$, 0.0001% $MnCl_2{\cdot}4H_2O$, and 0.001% $ZnSO_4{\cdot}7H_2O$) supplemented with either $(NH_4)_2SO_4$ or KNO_3 as the sole nitrogen sources. When necessary, appropriate antibiotics were added in the medium mentioned above at the following concentrations: 100 μg/ml ampicillin, 50 μg/ml kanamycin, 34 μg/ml chloramphenicol, 50 μg/ml apramycin, and 100 μg/ml hygromycin.

Construction of CRISPR-Cas12a-Based Genome Editing Plasmids

Primers for plasmid construction, mutant verification, and Sanger sequencing are listed in **Supplementary Table S2**. The *Francisella tularensis* subsp. *novicida Cas12a* gene (*FnCas12a*, previously known as *FnCpf1*) was PCR amplified from the plasmid pJV53-Cpf1 (Yan et al., 2017) using primer of *FnCas12a*-F and *FnCas12a*-R. Then, the linearized vector was obtained through PCR amplification of pDZL803 (Li et al., 2017b) with primers of pDZL803-apr-F and pDZL803-apr-R, which contained the promoter region of the apramycin resistance gene (P_{apr}). *Cas12a* gene and the linearized pDZL803 vector were then assembled using the Ezmax seamless assembly kit (Tolo Biotech., Shanghai, China), and the obtained recombinant plasmid pDZLCas12a was further confirmed by Sanger sequencing (**Supplementary Figure S1**).

The crRNA guide sequences are listed in **Supplementary Table S3**. First, the BpmI and HindIII restriction sites in pCR-Hyg were replaced by the BbsI and AseI sites, obtaining the plasmid pCR1. In detail, the plasmid pCR-Hyg (Yan et al., 2017) was employed as the template for PCR amplification with primers of pCR-HYG-F and pCR-HYG-R, followed by template removal with DpnI and then self-assembly with the Ezmax seamless assembly kit (Tolo Biotech.). Then, the crRNA expressing cassette, containing the hsp60 promoter, two crRNA Direct Repeats (DR) sequences, the BbsI and AseI sites for insertion of crRNA guide sequences and the *rrnB* T1 terminator,

was further amplified from pCR1 with paired primers of hsp60-rrnB-F and hsp60-rrnB-R, and the amplicon was then inserted into plasmid pULVK2A (Kumar et al., 1994), generating plasmid pULcrRNA (**Supplementary Figure S1**). Plasmid pULVK2A was generated from pRL1 (Lal et al., 1991) by spontaneous deletion of DNA sequences during passage, and can stably self-replicate in *A. mediterranei*. The Cas12a-expressing plasmid pDZLCas12a and the crRNA-expressing plasmid pULcrRNA were used to test the effectiveness of the CRISPR/Cas12a system in U32.

Deletion of *rifZ* and *glnR* Genes in U32

The CRISPR/Cas12a-assisted genome editing plasmids were constructed on the basis of plasmid pULcrRNA. First, 20-nt crRNA guide sequences for targeting *rifZ* and *glnR* were designed, synthesized, and individually annealed, and were then inserted into pULcrRNA that was digested by BbsI and AseI, generating pULrifZ and pULglnR, respectively. Then, 1.5-kb upstream and downstream sequences of the target genes (e.g., *rifZ* and *glnR*) were PCR amplified from U32 genome with paired primers (rifZL-F/rifZL-R, rifZR-F/rifZR-R, glnRL-F/glnRL-R, and glnRR-F/glnRR-R), the amplicons of which were used as homologous arms for HDR. The apramycin resistance gene was released from pBCAm plasmid by PstI digestion. Then, the upstream and downstream arms as well as the apramycin resistance cassette were assembled (designated as LAR donor fragment) by Ezmax seamless assembly kit (Tolo Biotech.) and introduced into the NdeI-treated plasmids of pULrifZ and pULglnR, obtaining the knout-out plasmids of pULrifZ-LAR and pULglnR-LAR, respectively (**Supplementary Figures S1, S3**). Alternatively, plasmids for markerless deletion of target genes were constructed through direct assembly of the upstream and downstream homologous arms and the crRNA expression cassette for guiding target-specific cleavage, and the obtained plasmids for *rifZ* and *glnR* markerless deletion were named pULrifZ-LR and pULglnR-LR, respectively (**Supplementary Figure S1**).

The U32 competent cells for electroporation were prepared as previously described (Ding et al., 2003). The Cas12a expression vector (pDZLCas12a) was electroporated into U32 competent cells and transformants were cultured on selective plates containing hygromycin. Specifically, about 500-ng pDZLCas12a was electroporated into 75-μl U32 competent cells with the following electroporation parameters: 1760 V, 1000 Ω, 25 μF, and 2 mm cuvette. Transformants were cultivated at 30°C for 7 days, and the clones were counted, analyzed, and verified by both PCR amplification and subsequent Sanger sequencing. The transformant expressing Cas12a was then employed for preparation of competent cells for subsequent gene editing. To test the NHEJ activities in U32, 300-ng crRNA-expressing plasmids of pULrifZ and pULglnR, targeting *rifZ* and *glnR*, respectively, were electroporated into the competent cells expressing Cas12a, and the transformants were then cultured and analyzed. Noticeably, there were no donor arms on plasmids pULrifZ and pULglnR for homologous recombination.

Similarly, to precisely delete target genes via HDR-mediated repair of DSBs, 300-ng plasmids of pULrifZ-LAR, pULrifZ-LR, pULglnR-LAR, and pULglnR-LR were individually electroporated into the Cas12a-expressing competent cells to delete the target gene *rifZ* and *glnR*, respectively. The transformants were cultivated on Bennet plate supplemented with apramycin at 30°C for 7 days, and the colonies were confirmed by colony PCR and Sanger sequencing.

Phenotypic Analyses of Gene Deletion Mutants

To analyze the growth phenotypes of the *glnR* deletion mutants, mutants were cultured with minimal medium with 20 mM KNO_3 or 10 mM $(NH4)_2SO_4$ as the sole nitrogen sources. Specifically, mutants were first grown in Bennet medium and the cells were then washed with nitrogen-free medium. After that, a 10-fold serial dilution was made from the starting OD_{600} density of 1–1/400, and diluted cells were then spotted onto minimal medium plate (Dadura et al., 2017) before being further incubated at 30°C for 5 days. Three independent experiments were performed.

RESULTS

As *A. mediterranei* is an important industrial strain for rifamycin production, many efforts have been made to study its genetic operation system, including the characterization of endogenous plasmids. However, up to now, there is only one stable replicon (namely, the pA-rep) characterized from the endogenous plasmid pA387 in *Amycolatopsis* sp. DSM 43387, and all self-replicable plasmids (e.g., pRL1 and pULVK2A) in *A. mediterranei* are generated from this plasmid (Lal et al., 1991; Kumar et al., 1994). Due to the plasmid incompatibility, it is hard to stably transform two plasmids with the same replication origin inside one cell. Therefore, to develop a CRISPR-based genetic engineering system in *A. mediterranei* U32, we decided to clone the *Cas* gene in an integrative plasmid and the crRNA expression cassette in a self-replicable plasmid. We ever tested the dead *SpCas9* gene from *Streptococcus pyogenes* (Jinek et al., 2012), and cloned it into an integrative plasmid, which was then electroporated into U32 competent cells. However, no transformants were obtained (data not shown) after repeated electroporation experiments, which indicated that the expression of *dCas9* alone was toxic to U32.

Then, instead of testing the wild-type Cas9, we tested the FnCas12a from *F. tularensis* (Zetsche et al., 2015), an alternative to Cas9 for CRISPR-mediated genome editing and has been successfully used in *M. smegmatis* (Yan et al., 2017). Similarly, the codon optimized *FnCas12a* gene was cloned in an integrative plasmid and was further electroporated into U32 competent cells to allow for integration into the *attB* site in the genome (**Figure 1**). Transformants were successfully obtained and the integrated *FnCas12a* gene was further confirmed by colony PCR verification and subsequent Sanger DNA sequencing; however, the transformation efficiency was much lower than that of the control plasmid with no *Cas12a* gene. Further phenotypic analysis showed that both bacterial growth and the rifamycin production of the transformant expressing FnCas12a were similar to those of the wild type U32 (**Supplementary Figure S2**), which implied that the constructed strain was a qualified system for genome editing analysis.

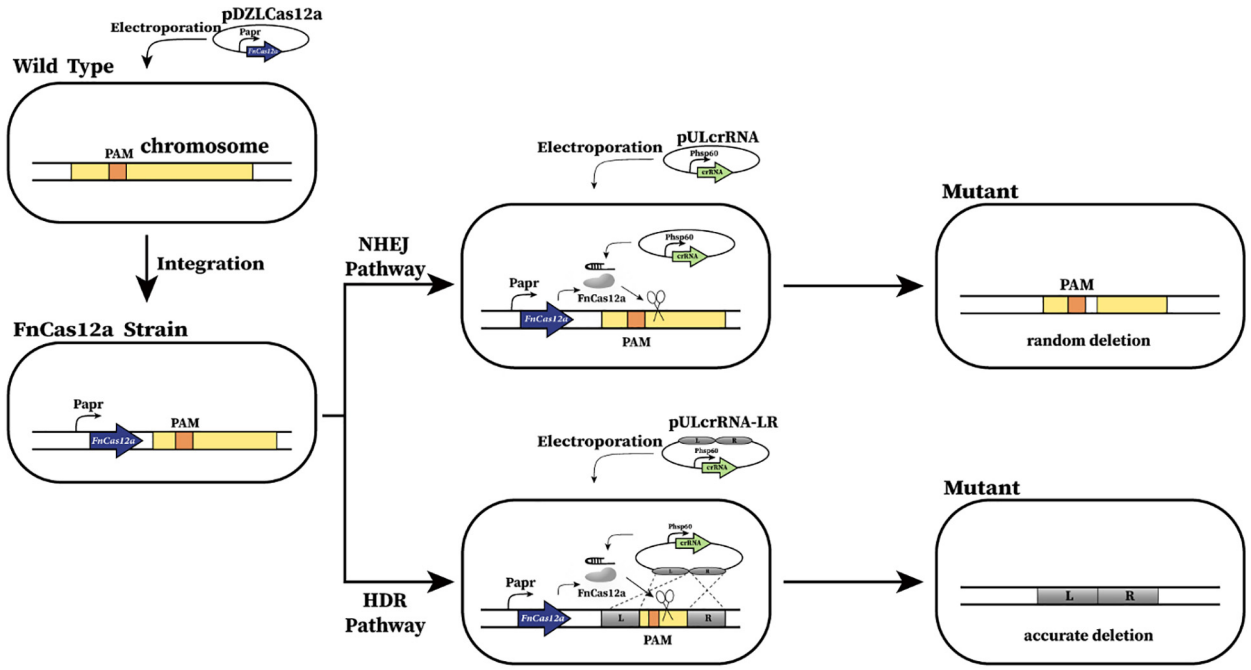


FIGURE 1 | Schematic of CRISPR-Cas12a-assisted genome editing in *A. mediterranei* U32. First, *FnCas12a* was integrated into U32 chromosome and constitutively expressed under the promoter of P_{apr}. Then, the crRNA transcribing plasmid was electroporated into the U32 competent cells that harbored FnCas12a. Guided by a target-specific crRNA, Cas12a specifically cleaved target dsDNA and generated DSBs on the chromosome, which could be repaired by either NHEJ or HDR, generating desired mutants. When DSBs were repaired by NHEJ, the breaks were directly jointed with no homologous donor DNA required; however, DNA sequences of ranging length around the Cas12a cleavage site will be deleted during the repair.

To test the whether Cas12a could introduce site-specific DSBs in U32, the *glnR* gene, which encodes the global regulator for nitrogen metabolisms, was chosen as the target gene. We designed three crRNAs targeting different coding regions in *glnR*, where crRNA1 and crRNA2 targeted the non-template strand (NT) and the crRNA-3 targeted the template strand (T) (**Supplementary Table S3**). The above three crRNAs were individually cloned into self-replicable pULVK2A, and the obtained plasmids were then electroporated into the U32 competent cells that constitutively expressed FnCas12a. In comparison with thousands of colonies obtained with the transformation of the control plasmid with no crRNA expression cassette, only 0, 9, and 1 transformants were obtained for plasmids pULglnR1, pULglnR2, and pULglnR3, respectively (**Figure 2A**), suggesting that the CRISPR-Cas12a system could efficiently cleave the genomic DNA.

On the other hand, although no homologous DNA sequences were introduced for homologous recombination, we still obtained some transformants (**Figure 2A**), and the nine transformants with pULglnR2 were further verified by PCR amplification of the target regions. The PCR results showed that the amplicons were of different sizes (**Figure 2B**), which indicated that there might be random deletions inside the target gene. We further confirmed this hypothesis by Sanger sequencing of the PCR amplicons and found five of nine transformants contained deletions (ranging from 5 to 71 bps) at the glnRcrRNA2 targeting site (**Figure 2C**). With the identification of conserved homologues of ATP-dependent DNA ligase and Ku protein (**Supplementary Tables S4, S5**), one may conclude that the DSBs were probably repaired by NHEJ in U32, although the possibility of other template-independent repair such as alternative end-joining (A-EJ) (Chayot et al., 2010) and microhomology-mediated end-joining (MMEJ) (Sfeir and Symington, 2015) cannot be completely excluded. Meanwhile, there might also exist unknown mechanisms to inactivate the CRISPR-Cas12a system in U32 as no mutations were found in the target region among the rest four clones. Collectively, above findings not only demonstrated that Cas12a generated site-specific DSBs but also suggested that there was NHEJ in U32. As a consequence, with the combination of CRISPR-Cas12a-induced site-specific DSBs and NHEJ-mediated repair, markerless gene mutations can be easily acquired in this bacterium.

Besides, we also combined the CRISPR-Cas12a-assisted target cleavage with endogenous HDR activity to precisely delete target genes. To measure the efficiency of the HDR-mediated precise genome editing, we next knocked out the *rifZ* gene encoding the rifamycin pathway-specific activator by replacing it with the apramycin resistance cassette (**Figure 3A**). Three crRNAs were designed to target both the T strand (rifZcrRNA-1 and rifZcrRNA-2) and the NT strand (rifZcrRNA-3) of *rifZ* (**Supplementary Table S3**). The resistance cassette was in fusion assembled with both upstream and downstream homologous arms of *rifZ*, and the obtained donor fragment

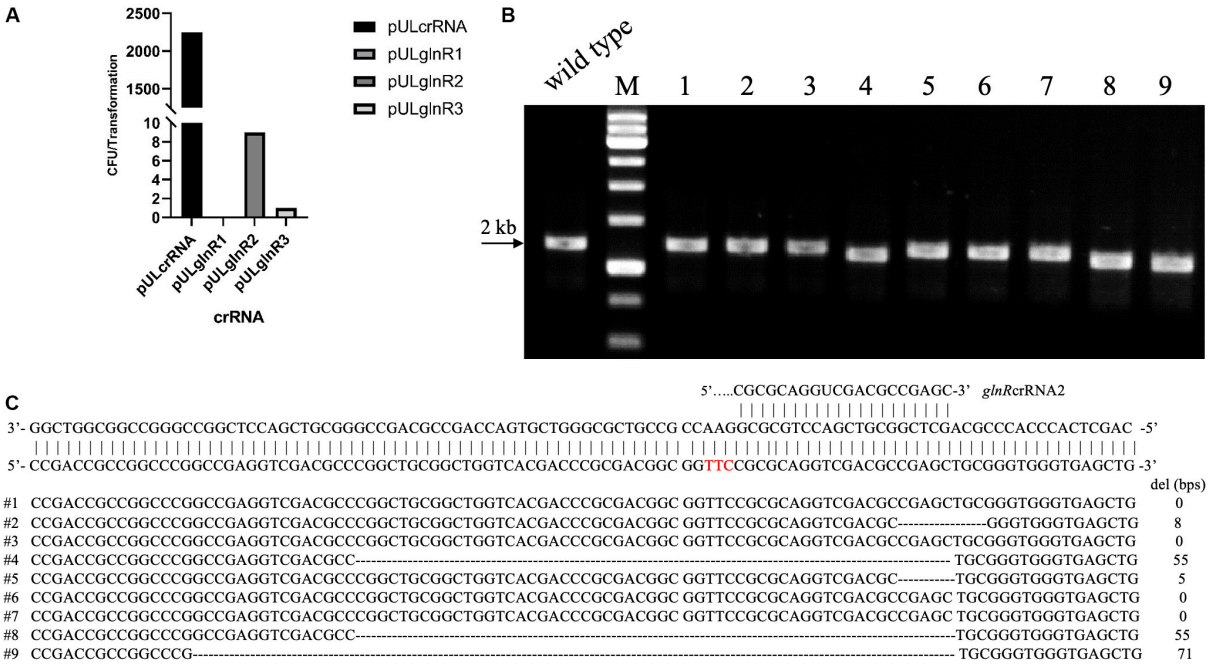


FIGURE 2 | CRISPR-Cas12a-assisted *glnR* deletion in combination with the NHEJ repair in *A. mediterranei* U32. **(A)** The number of transformants obtained with different crRNAs transformed. Plasmid pULcrRNA, which contained no crRNA guide sequences, was employed as a control plasmid. The number stood for the total colonies obtained from three transformation experiments, where no transformants were obtained with crRNA1. The guide sequences are listed in **Supplementary Table S3**. **(B)** Characterization of pULglnR2 transformants by colony PCR amplification. M, GeneRuler 1 kb DNA Ladder (Thermo Scientific). Lanes 1–9, nine colonies transformed with CRISPR-Cas12a and pULglnR2 targeting *glnR* gene. **(C)** Sanger sequencing results of the PCR amplicons of the nine colonies in **Figure 2B**. Five colonies contained random deletion at the target site within *glnR* gene, which was repaired by the NHEJ pathway, and the deleted DNA sequences as well as the length were indicated. The PAM sequence was highlighted in red.

was then introduced into the plasmids expressing *rifZ* targeting crRNAs, generating plasmids pULrifZ1-LAR, pULrifZ2-LAR, and pULrifZ3-LAR, respectively. The obtained three plasmids as well as a control plasmid were then individually electroporated into the U32 competent cells harboring *FnCas12a*. In comparison with the more than 1000 colonies obtained from the transformation of the control plasmid, less than 10 colonies on average were obtained with the three plasmids with *rifZ*-specific crRNAs (**Figure 3B**). Subsequently, eight colonies from the transformation of pULrifZ1-LAR were verified by both PCR amplification and Sanger sequencing, and the results unambiguously showed that the *rifZ* gene was precisely replaced with the apramycin resistance cassette in all tested colonies (**Figures 3C,D**).

After confirmation of the effectiveness of HDR-mediated repair of CRISPR-Cas12a-generated DSBs, we then attempted to combine the CRISPR-Cas12a system and the endogenous HDR pathway to construct precise markerless mutants of both *rifZ* and *glnR*. The upstream and downstream homologous arms of the target genes were in fusion assembled and then inserted into the crRNA expressing plasmid, followed by electroporation into the U32 competent cells that constitutively expressed Cas12a protein (**Figure 4A**). For both target genes, i.e., *glnR* and *rifZ*, a dozen transformants were successfully obtained, which were further verified by both PCR amplification and Sanger sequencing. Among the four tested *rifZ* mutants, three had the HDR-assisted accurate *rifZ* gene deletion and one had inaccurate 692-bp deletion within *rifZ* gene, which was obviously repaired by NHEJ (**Figures 4B–D**). Phenotypic analysis showed that all four *rifZ* mutants produced no golden pigment (**Figure 4E** and **Supplementary Figure S4a**) and much reduced rifamycin SV yield (**Supplementary Figure S4b**) as indicated by the bactericidal test, which were consistent with the previous findings that RifZ functions as the pathway-specific activator for the whole *rif* cluster (Li et al., 2017a). Similarly, both the colony PCR and Sanger sequencing results demonstrated that the *glnR* gene, which encodes the central governor for nitrogen metabolisms, was precisely and markerlessly deleted (**Supplementary Figures S5a,b**) in all four tested transformants. Subsequent growth phenotypic analysis showed that all these *glnR* mutants grew poorly on minimal medium with nitrate as the sole nitrogen source (**Supplementary Figure S5c**). Collectively, above results clearly demonstrated that endogenous HDR pathway can be employed to efficiently repair the CRISPR-Cas12a-generated DSBs and engineer precise and markerless mutants.

DISCUSSION

In this study, we successfully employed CRISPR-Cas12a system to develop a genome editing system in *A. mediterranei*. To test

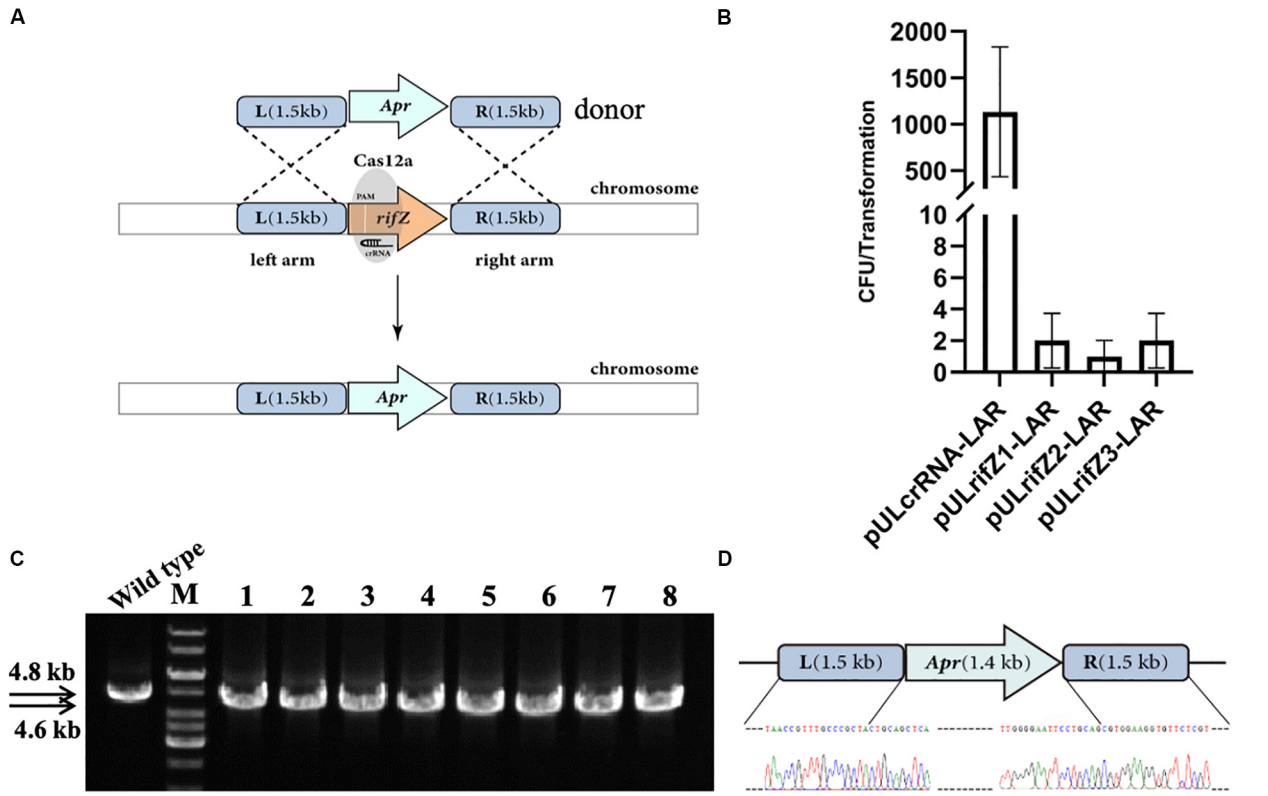


FIGURE 3 | Replacement of *rifZ* by an antibiotic cassette with CRISPR-Cas12a-assisted precise cutting and HDR-mediated precise recombination in *A. mediterranei* U32. **(A)** Schematic of *rifZ* deletion. With the existence of a crRNA targeting *rifZ*, Cas12a introduced DSBs within *rifZ* gene, which could then be precisely repaired by a donor DNA fragment containing an apramycin resistance cassette and the flanking upstream and downstream homologous sequences. **(B)** The number of transformants obtained with three different crRNAs targeting different loci within *rifZ*. Plasmid pULcrRNA, which contained no crRNA guide sequences, was employed as a control plasmid. The data were obtained from three independent transformation experiments. **(C)** PCR amplification analysis of *rifZ* mutant with paired primers of *rifZ*-KO-F and *rifZ*-KO-R. The amplicon of the wild type was 4.8 kb in length, while was smaller in *rifZ* mutants. M, GeneRuler 1 kb DNA Ladder (Thermo Scientific); lanes 1 and 2, colonies obtained with pUL*rifZ*1-LAR; lanes 3 and 4, colonies obtained with pUL*rifZ*2-LAR; lanes 5–8, colonies obtained with pUL*rifZ*3-LAR. **(D)** Sanger sequencing results of the amplicon from **Figure 3C**. All eight amplicons were sequenced and found to be correct, and only the results of the first colony were presented.

the effectiveness of Cas12a-mediated site-specific DSBs, Cas12a and crRNAs were co-expressed in U32. And to our great surprise, even no homologous recombination arms were introduced, several transformants were obtained, leading to the identification of the endogenous NHEJ activities. Then, we showed that Cas12a-introduced DSBs could be efficiently repaired by either NHEJ or HDR, which therefore facilitates convenient genome editing in *A. mediterranei*.

We first tested dCas9 but found the protein was toxic to U32. As Cas9 has been demonstrated to be toxic in many other species (Jiang et al., 2017; Cho et al., 2018), we here directly used Cas12a to construct the genome editing system in U32 instead of testing the wild type Cas9. As the transformation efficiency of the plasmids containing *Cas12a* was much lower than that of the control plasmid, the *Cas12a* gene might also be harmful to U32 cells. However, once transformants were obtained, both the growth rate and the rifamycin yield of these transformants expressing Cas12a were similar to those of the wild type U32, and Cas12a was therefore employed to develop the genome editing system in U32.

Because of the relatively low transformation and recombination efficiencies in the genus *Amycolatopsis*, it is difficult to construct markerless mutants in this genus. Although an electroporation transformation system has been established in U32 years ago (Ding et al., 2002), the restricted condition for bacterial growth and the complex procedure for preparation of electro-competent cells make it difficult to prepare U32 competent cells of high transformation efficiency. Many factors have been known to affect bacterial transformation efficiency, including the restriction systems. With the CRISPR-Cas12a genome editing system established in this study, these factors can be efficiently modified to improve the U32 transformation efficiency. Furthermore, with the availability of the CRISPR-Cas12a system, precise DSBs can be introduced by the crRNA-guided Cas12a cleavage, then either NHEJ or HDR can be employed to repair the DNA damage, generating desired mutants with no markers left. Furthermore, CRISPR-Cas12a-assisted markerless mutagenesis makes it possible to perform continuous genome editing operations. As there is only one available self-replicating plasmid origin in U32, the

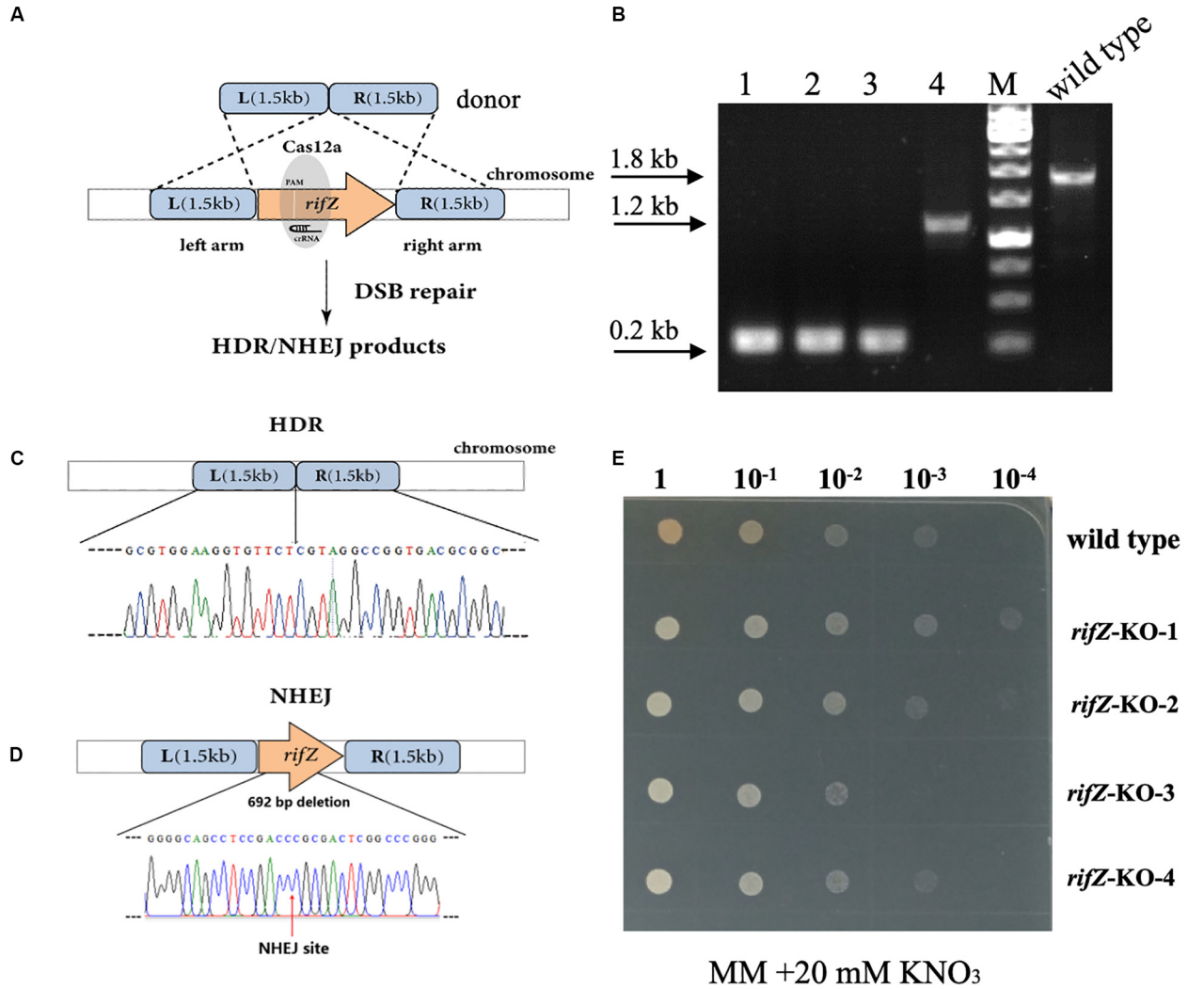


FIGURE 4 | Markerless deletion of *rifZ* gene with the Cas12a and HDR systems in U32. **(A)** Schematic of markerless deletion of *rifZ* gene. Cas12a-generated DSBs in *rifZ* could be repaired by HDR. **(B)** Verification of *rifZ* deletion mutants by colony PCR with paired primers of rifZ+P-F and rifZ+P-R. M, GeneRuler 1 kb DNA Ladder (Thermo Scientific). Lanes 1–4 represented *rifZ*-KO-1 to *rifZ*-KO-4, respectively. **(C)** Sanger sequencing results from the amplicon of *rifZ*-KO-1 in **Figure 4B**. **(D)** Schematic of the *rifZ*-KO-4 mutant in **Figure 4B**. The mutant was obtained via NHEJ-mediated repair of the Cas12a-generated DSBs in *rifZ* gene. Confirmed by Sanger sequencing, 692-bp sequences were found to be deleted at the Cas12a cleavage site. **(E)** The growth phenotype of *rifZ* mutants on minimal medium. Serially diluted liquid culture was spotted on plates and cultured at 30°C for 5 days, and only the wild type produced pigmented rifamycin SV.

original plasmid must be eliminated before a new plasmid can be transformed, expressing a new crRNA to target a new locus. Fortunately, plasmid curing experiments showed that nearly all original plasmids could be eliminated after one or two generations of passage in medium without selective pressure of antibiotics (**Supplementary Figure S6**). Alternatively, a new plasmid carrying a different antibiotic resistance cassette can be directly transformed and the transformants can be cultured on plates with the new antibiotic. Moreover, the new plasmid may also express a crRNA targeting the original antibiotic resistance cassette to help clear the original plasmid.

Cas12a is so far the most minimalistic of CRISPR systems and can process precursor crRNAs (Fonfara et al., 2016). Based on this characteristic, multiple gene editing can be easily achieved through simply constructing a crRNA array, expressing multiple precursor crRNAs driven under one promoter, which can be further processed by Cas12a to generate multiple mature crRNAs for multiple gene editing or gene regulation *in vivo* (Zetsche et al., 2017; Zhang et al., 2017b). After mutation of the RuvC domain, the DNase-dead Cas12a (namely, ddCas12a) can be employed for gene regulation. Similarly, multiple gene regulation can be achieved with the co-expression of both ddCas12a and a crRNA array (Zhang et al., 2017b). Moreover, with the mutagenesis of the crRNA DR, the ddCas12a binding affinities against mutant crRNAs can be precisely determined and hence the regulatory strength of target genes' transcription (Wang et al., 2019). Although gene regulation was not tested in this study, one

may easily perform transcriptional regulation through simply changing the wild type Cas12a to ddCas12a.

There are vast majority of biosynthetic clusters in the genus *Amycolatopsis*, demonstrating the genus has great potential to produce diverse secondary metabolites (Adamek et al., 2018). Without efficient genome editing tools, heterologous expression is the main way to produce and characterize these metabolites, which could be of low efficiency. As CRISPR-based genome editing has been demonstrated as an efficient approach to discover unique metabolites in *Streptomyces* (Zhang et al., 2017a), the present work will certainly shed light on the development of CRISPR-assisted genome editing systems in other species in *Amycolatopsis* genus and further facilitates the genome mining in this genus.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

YZ and XL performed most of the experiments. YZ prepared the draft. JWu drew the schematic maps. JWa and GZ designed the study. JWa revised the manuscript and supervised the whole project.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (31770057, 31670058, and 31430004). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

ACKNOWLEDGMENTS

The authors acknowledge the valuable assistance of Professor Yi-Cheng Sun for kindly providing plasmids of pJV53-Cpf1 and pCR-HYG. The authors thank Professor Hua Yuan for his careful revision of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00698/full#supplementary-material

REFERENCES

Adamek, M., Alanjary, M., Sales-Ortells, H., Goodfellow, M., Bull, A. T., Winkler, A., et al. (2018). Comparative genomics reveals phylogenetic distribution patterns of secondary metabolites in *Amycolatopsis* species. *BMC Genomics* 19:426. doi: 10.1186/s12864-018-4809-4

Babynin, E. V. (2007). [Molecular mechanism of homologous recombination in meiosis: origin and biological significance]. *Tsitologiia* 49, 182–193.

Chayot, R., Montagne, B., Mazel, D., and Ricchetti, M. (2010). An end-joining repair mechanism in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2141–2146. doi: 10.1073/pnas.0906355107

Cho, S., Choe, D., Lee, E., Kim, S. C., Palsson, B., and Cho, B. K. (2018). High-level dCas9 expression induces abnormal cell morphology in *Escherichia coli*. *ACS Synth. Biol.* 7, 1085–1094. doi: 10.1021/acssynbio.7b00462

Cobb, R. E., Wang, Y., and Zhao, H. (2015). High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS Synth. Biol.* 4, 723–728. doi: 10.1021/sb500351f

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823.

Dadura, K., Plocinska, R., Rumijowska-Galewicz, A., Plocinski, P., Zaczek, A., Dziadek, B., et al. (2017). PdtaS deficiency affects resistance of mycobacteria to ribosome targeting antibiotics. *Front. Microbiol.* 8:2145. doi: 10.3389/fmicb.2017.02145

Ding, X., Tian, Y., Chiao, J., Zhao, G., and Jiang, W. (2003). Stability of plasmid pA387 derivatives in *Amycolatopsis* mediterranei producing rifamycin. *Biotechnol. Lett.* 25, 1647–1652.

Ding, X. M., Zhang, N., Tian, Y. Q., Jiang, W. H., Zhao, G. P., and Jiao, R. S. (2002). [Establishment of gene replacement/disruption system through homologous recombination in *Amycolatopsis* mediterranei U32]. *Sheng Wu Gong Cheng Xue Bao* 18, 431–437.

Fonfara, I., Richter, H., Bratovic, M., Le Rhun, A., and Charpentier, E. (2016). The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* 532, 517–521. doi: 10.1038/nature17945

Harrison, P., and Hart, S. (2018). Gene editing and gene regulation with CRISPR. *Exp. Physiol.* 103, 437–438. doi: 10.1113/ep086864

Horvath, P., and Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327, 167–170. doi: 10.1126/science.1179555

Hu, X. F., Zhang, B., Liao, C. H., and Zeng, Z. J. (2019). High-efficiency CRISPR/Cas9-mediated gene editing in honeybee (*Apis mellifera*) embryos. *G3* 9, 1759–1766. doi: 10.1534/g3.119.400130

Huang, H., Zheng, G., Jiang, W., Hu, H., and Lu, Y. (2015). One-step high-efficiency CRISPR/Cas9-mediated genome editing in *Streptomyces*. *Acta Biochim. Biophys. Sin.* 47, 231–243. doi: 10.1093/abbs/gmv007

Jia, H., Zhang, L., Wang, T., Han, J., Tang, H., and Zhang, L. (2017). Development of a CRISPR/Cas9-mediated gene-editing tool in *Streptomyces rimosus*. *Microbiology* 163, 1148–1155. doi: 10.1099/mic.0.000501

Jiang, W., Yang, B., and Weeks, D. P. (2014). Efficient CRISPR/Cas9-mediated gene editing in *Arabidopsis thaliana* and inheritance of modified genes in the T2 and T3 generations. *PLoS One* 9:e99225. doi: 10.1371/journal.pone.0099225

Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., and Yang, S. (2015). Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl. Environ. Microbiol.* 81, 2506–2514. doi: 10.1128/aem.04023-14

Jiang, Y., Qian, F., Yang, J., Liu, Y., Dong, F., Xu, C., et al. (2017). CRISPR-Cpf1 assisted genome editing of *Corynebacterium glutamicum*. *Nat. Commun.* 8:15179.

Jiao, R. S., Chen, Y. M., Wu, M. G., and Gu, W. L. (1979). Studies on the metabolic regulation of biosynthesis of rifamycin by Norcadia (*Amycolatopsis*) mediterranei I. The stimulative effect of nitrate on biosynthesis of rifamycin SV by Nocardia (*Amycolatopsis*) mediterranei. *Acta Phytophysiol. Sin.* 05, 395–402.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10.1126/science.1225829

Koonin, E. V., Makarova, K. S., and Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.* 37, 67–78. doi: 10.1016/j.mib.2017.05.008

Kumar, C. V., Coque, J. J., and Martin, J. F. (1994). Efficient transformation of the cephamycin C producer nocardia lactamdurans and development of shuttle and promoter-probe cloning vectors. *Appl. Environ. Microbiol.* 60, 4086–4093. doi: 10.1128/aem.60.11.4086-4093.1994

Kumari, R., Singh, P., and Lal, R. (2016). Genetics and genomics of the Genus *Amycolatopsis*. *Indian J Microbiol.* 56, 233–246. doi: 10.1007/s12088-016-0590-8

Lal, R., Lal, S., Grund, E., and Eichenlaub, R. (1991). Construction of a hybrid plasmid capable of replication in *Amycolatopsis* mediterranei. *Appl. Environ. Microbiol.* 57, 665–671. doi: 10.1128/aem.57.3.665-671.1991

Lee, J. G., Choi, C. Y., Seong, B. L., and Han, M. H. (1983). Optimal Ph Profile in Rifamycin-B fermentation. *J. Ferment. Technol.* 61, 49–53.

Li, C., Liu, X., Lei, C., Yan, H., Shao, Z., Wang, Y., et al. (2017a). RifZ (AMED_0655) is a pathway-specific regulator for rifamycin biosynthesis in *Amycolatopsis mediterranei*. *Appl. Environ. Microbiol.* 83, e3201–e3216.

Li, C., Zhou, L., Wang, Y., Zhao, G., and Ding, X. (2017b). Conjugation of varphiBT1-derived integrative plasmid pDZL802 in *Amycolatopsis mediterranei* U32. *Bioengineered* 8, 549–554. doi: 10.1080/21655979.2016.1270808

Lieber, M. R. (2010). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* 79, 181–211. doi: 10.1146/annurev.biochem.052308.093131

Low, B. E., Kutny, P. M., and Wiles, M. V. (2016). Simple, efficient CRISPR-Cas9-mediated gene editing in mice: strategies and methods. *Methods Mol. Biol.* 1438, 19–53. doi: 10.1007/978-1-4939-3661-8_2

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., et al. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9, 467–477.

Makarova, K. S., Zhang, F., and Koonin, E. V. (2017a). SnapShot: class 1 CRISPR-Cas systems. *Cell* 168, 946–946.

Makarova, K. S., Zhang, F., and Koonin, E. V. (2017b). SnapShot: class 2 CRISPR-Cas systems. *Cell* 168, 328–328.

Matsu-Ura, T., Baek, M., Kwon, J., and Hong, C. (2015). Efficient gene editing in *Neurospora crassa* with CRISPR technology. *Fungal Biol. Biotechnol.* 2:4.

Mejia, A., Barrios-Gonzalez, J., and Viniegra-Gonzalez, G. (1998). Overproduction of rifamycin B by *Amycolatopsis* mediterranei and its relationship with the toxic effect of barbital on growth. *J. Antibiot.* 51, 58–63. doi: 10.7164/antibiotics.51.58

Mohanraju, P., Makarova, K. S., Zetsche, B., Zhang, F., Koonin, E. V., and van der Oost, J. (2016). Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* 353:aad5147. doi: 10.1126/science.aad5147

Rothstein, D. M. (2016). Rifamycins, alone and in combination. *Cold Spring Harbor Perspect. Med.* 6:a027011. doi: 10.1101/cshperspect.a027011

Sfeir, A., and Symington, L. S. (2015). Microhomology-mediated end joining: a back-up survival mechanism or dedicated pathway? *Trends Biochem Sci.* 40, 701–714. doi: 10.1016/j.tibs.2015.08.006

Sung, P., and Klein, H. (2006). Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat. Rev. Mol. Cell Biol.* 7, 739–750. doi: 10.1038/nrm2008

Wang, J., Lu, A., Bei, J., Zhao, G., and Wang, J. (2019). CRISPR/ddCas12a-based programmable and accurate gene regulation. *Cell Discovery* 5:15.

Wright, D. G., Castore, R., Shi, R., Mallick, A., Ennis, D. G., and Harrison, L. (2017). Mycobacterium tuberculosis and *Mycobacterium marinum* non-homologous end-joining proteins can function together to join DNA ends in *Escherichia coli*. *Mutagenesis* 32, 245–256.

Xu, L., Huang, H., Wei, W., Zhong, Y., Tang, B., Yuan, H., et al. (2014). Complete genome sequence and comparative genomic analyses of the vancomycin-producing *Amycolatopsis* orientalis. *BMC Genomics* 15:363. doi: 10.1186/1471-2164-15-363

Yamano, T., Nishimasu, H., Zetsche, B., Hirano, H., Slaymaker, I. M., Li, Y., et al. (2016). Crystal structure of cpf1 in complex with guide RNA and target DNA. *Cell* 165, 949–962.

Yan, M. Y., Yan, H. Q., Ren, G. X., Zhao, J. P., Guo, X. P., and Sun, Y. C. (2017). CRISPR-Cas12a-assisted recombineering in bacteria. *Appl. Environ. Microbiol.* 83, e00947-17.

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163, 759–771. doi: 10.1016/j.cell.2015.09.038

Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E. M., et al. (2017). Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat. Biotechnol.* 35, 31–34. doi: 10.1038/nbt.3737

Zhang, M. M., Wong, F. T., Wang, Y., Luo, S., Lim, Y. H., Heng, E., et al. (2017a). CRISPR-Cas9 strategy for activation of silent *Streptomyces* biosynthetic gene clusters. *Nat. Chem. Biol.* 13, 607–609. doi: 10.1038/nchembio.2341

Zhang, X., Wang, J., Cheng, Q., Zheng, X., Zhao, G., and Wang, J. (2017b). Multiplex gene regulation by CRISPR-ddCpf1. *Cell Discovery* 3:17018.

Zhao, W., Zhong, Y., Yuan, H., Wang, J., Zheng, H., Wang, Y., et al. (2010). Complete genome sequence of the rifamycin SV-producing *Amycolatopsis* mediterranei U32 revealed its genetic characteristics in phylogeny and metabolism. *Cell Res.* 20, 1096–1108. doi: 10.1038/cr.2010.87

Zheng, X., Li, S. Y., Zhao, G. P., and Wang, J. (2017). An efficient system for deletion of large DNA fragments in *Escherichia coli* via introduction of both Cas9 and the non-homologous end joining system from *Mycobacterium smegmatis*. *Biochem. Biophys. Res. Commun.* 485, 768–774. doi: 10.1016/j.bbrc.2017.02.129

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 30 June 2020
doi: 10.3389/fbioe.2020.00726

Efficient Multiplex Genome Editing in *Streptomyces* via Engineered CRISPR-Cas12a Systems

Jun Zhang[1†], Dan Zhang[1,2†], Jie Zhu[1], Huayi Liu[1], Shufang Liang[1,2] and Yunzi Luo[1,3*]*

[1] Department of Gastroenterology, State Key Laboratory of Biotherapy, West China Hospital, Sichuan University and Collaborative Innovation Center of Biotherapy, Chengdu, China, [2] State Key Laboratory of Biotherapy and Cancer Center, West China Hospital, Sichuan University and Collaborative Innovation Center of Biotherapy, Chengdu, China, [3] Key Laboratory of Systems Bioengineering (Ministry of Education), Frontiers Science Center of Synthetic Biology, School of Chemical Engineering and Technology, Tianjin University, Tianjin, China

OPEN ACCESS

Edited by:
Yuan Lu,
Tsinghua University, China

Reviewed by:
Yiping Qi,
University of Maryland, United States
Dae-Hee Lee,
Korea Research Institute of Bioscience & Biotechnology (KRIBB), South Korea
Steven Lin,
Institute of Biological Chemistry, Academia Sinica, Taiwan

***Correspondence:**
Shufang Liang
zizi2006@scu.edu.cn
Yunzi Luo
yunzi.luo@tju.edu.cn;
luoyunzi827@aliyun.com

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: 07 January 2020
Accepted: 09 June 2020
Published: 30 June 2020

Citation:
Zhang J, Zhang D, Zhu J, Liu H, Liang S and Luo Y (2020) Efficient Multiplex Genome Editing in Streptomyces via Engineered CRISPR-Cas12a Systems. Front. Bioeng. Biotechnol. 8:726. doi: 10.3389/fbioe.2020.00726

Streptomyces strains produce a great number of valuable natural products. With the development of genome sequencing, a vast number of biosynthetic gene clusters with high potential for use in the discovery of valuable clinical drugs have been revealed. Therefore, emerging needs for tools to manipulate these biosynthetic pathways are presented. Although the clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 (CRISPR/Cas 9) system has exhibited great capabilities for gene editing in multiple *Streptomyces* strains, it has failed to work in some newly discovered strains and some important industrial strains. Additionally, the protospacer adjacent motif (PAM) recognition scope of this system sometimes limits its applications for generating precise site mutations and insertions. Here, we developed three efficient CRISPR-*Fn*Cas12a systems for multiplex genome editing in several *Streptomyces* strains. Each system exhibited advantages for different applications. The CRISPR-*Fn*Cas12a1 system was efficiently applied in the industrial strain *Streptomyces hygroscopicus*, in which *Sp*Cas9 does not work well. The CRISPR-*Fn*Cas12a2 system was used to delete large fragments ranging from 21.4 to 128 kb. Additionally, the CRISPR-*Fn*Cas12a3 system employing the engineered *Fn*Cas12a mutant EP16, which recognizes a broad spectrum of PAM sequences, was used to precisely perform site mutations and insertions. The CRISPR-*Fn*Cas12a3 system addressed the limitation of TTN PAM recognition in *Streptomyces* strains with high GC contents. In summary, all the CRISPR-*Fn*Cas12a systems developed in this study are powerful tools for precise and multiplex genome editing in *Streptomyces* strains.

Keywords: CRISPR, *Fn*Cas12a, *Streptomyces*, genome editing, PAM recognition

INTRODUCTION

Streptomyces, the largest genus of actinobacteria, has been well studied, as it contains the most prolific producers of a vast array of bioactive natural products, including antibiotics, antifungals, and anticancer agents (Baltz, 2008; Zhu et al., 2011; Cho et al., 2017; Frattaruolo et al., 2017). Over the past decades, large-scale genome sequencing efforts have revealed that great potential still remains for the discovery of new natural products produced by members of this genus. Thus,

increasing numbers of genetic engineering tools have been developed to explore these products (Luo et al., 2015b, 2016; Zhao et al., 2019). In particular, with the rapid development of the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein (Cas) system (Jiang et al., 2013; Mali et al., 2013; Yan et al., 2018), effective genome editing has become increasingly easy and convenient, paving the way for us to assemble or activate uncharacterized gene clusters. However, CRISPR/Cas9 has been reported to not work in some *Streptomyces* strains, such as *Streptomyces* sp. KY 40-1 (Salem et al., 2017), *Streptomyces* sp. NRRL S-244 (Yeo et al., 2019) and *Streptomyces hygroscopicus* SIPI-KF (Li et al., 2018), because of its toxicity to the hosts.

Cas12a, a Class 2 CRISPR effector, is a single RNA-guided endonuclease (Zetsche et al., 2015). Due to their simplicity, *As*Cas12a from *Acidaminococcus* sp. BV3L6 and *Lb*Cas12a from *Lachnospiraceae bacterium* ND2006 have generally been applied to mammalian cells (Toth et al., 2016) and plants (Tang et al., 2017), while *Fn*Cas12a from *Francisella novicida* U112 has successfully been applied to yeast (Swiat et al., 2017), *Corynebacterium glutamicum* (Jiang et al., 2017) and other bacteria (Ungerer and Pakrasi, 2016). Wild-type (WT) *Fn*Cas12a has been used in *Streptomyces*, particularly in some strains in which the CRISPR/Cas9 system does not function properly, such as *S. hygroscopicus* SIPI-KF (Li et al., 2018) and *Streptomyces* sp. NRRL S-244 (Yeo et al., 2019).

However, the TTN protospacer adjacent motif (PAM) requirement restricts the application of *Fn*Cas12a in *Streptomyces* strains, which are GC rich. In particular, target site selection is limited for site mutations or insertions in the genome. Thus, there remains a demand for highly efficient and versatile genome editing tools for *Streptomyces*. Recently, we constructed an engineered *Fn*Cas12a variant, EP16, which shows broad PAM recognition abilities *in vitro*, including YN (Y = C or T), TAC and CAA (Wang et al., 2019). Therefore, EP16 provides additional opportunities to precisely edit the high-GC-contents genomes of *Streptomyces* strains.

In this study, three optimized CRISPR-*Fn*Cas12a systems, CRISPR-*Fn*Cas12a1, CRISPR-*Fn*Cas12a2 and CRISPR-*Fn*Cas12a3 (**Figure 1** and **Supplementary Table 2**), were developed to edit the genomes of different *Streptomyces* strains based on homology-directed repair (HDR), as non-homologous end joining (NHEJ) does not occur in most streptomycetes. (i) The CRISPR-*Fn*Cas12a1 system exhibited a higher transformation efficiency than the CRISPR-*Fn*Cas12a2 system, as it contains a *traJ* gene encoding an activator of the transfer operon (Will and Frost, 2006). (ii) A higher editing efficiency was observed for the CRISPR-*Fn*Cas12a2 system than the CRISPR-*Fn*Cas12a1 system when *Fn*Cas12a was driven by the constitutive promoters *kasO*p* and rpsLp(XC). (iii) The *Fn*Cas12a3 system containing the *Fn*Cas12a mutant EP16 (N607R/K613V/N617R/K180S/K660R/D616N) with expanded PAM recognition ability was used to generate site mutation, insertion and subsequent biosynthetic gene cluster activation (**Figure 1**). Overall, our findings describe powerful tools for precise genome editing and subsequently for discovering and activating valuable natural products in *Streptomyces* strains.

MATERIALS AND METHODS

Strains, Cultivation and Reagents

Streptomyces coelicolor A3(2), *Streptomyces griseus*, *Streptomyces lividans*, *E. coli* DH5α and *E. coli* ET12567/pUZ8002 were gifts from Professor Huimin Zhao at the University of Illinois at Urbana-Champaign. *S. hygroscopicus* NRRL5491 and *Streptomyces roseosporus* NRRL 11379 were purchased from the American Type Culture Collection (ATCC, Manassas, VA, United States). All *Streptomyces* strains were grown in MYG liquid medium (10 g/L malt extract broth, 4 g/L yeast extract, and 4 g/L glucose) or on M-ISP4 agar medium (1 g/L yeast extract, 2 g/L tryptone, 5 g/L soluble starch, 5 g/L D-mannitol, 5 g/L soya flour, 1 g/L NaCl, 2 g/L $(NH_4)_2SO_4$, 1 g/L K_2HPO_3, 2 g/L $CaCO_3$, 20 g/L agar, 1 g/L $FeSO_4$, 1 g/L $MnCl_2$, and 1 g/L $ZnSO_4$, pH 7.0) at 30°C. Next, 2 × YT broth (1% yeast extract, 1% tryptone, and 0.5% NaCl, pH 7.0) was used for strain washing and spore germination before conjugation. M-ISP4 agar medium supplemented with 25 mM $MgCl_2$ or 20 μg/mL apramycin and 40 μg/mL nalidixic acid was used for conjugation from *E. coli* ET12567/pUZ8002 to *Streptomyces* (Kieser et al., 2000) or for plasmid selection in *Streptomyces*. R2YE medium supplemented with 20 μg/mL apramycin and 40 μg/mL nalidixic acid was used to screen potentially edited *Streptomyces*. *E. coli* DH5α was grown in Luria-Bertani (LB) broth with 50 μg/mL apramycin and used for plasmid cloning and maintenance. *E. coli* ET12567/pUZ8002 was grown in LB broth supplemented with 25 μg/mL apramycin, 12.5 μg/mL chloramphenicol and 25 μg/mL kanamycin.

D-Mannitol was obtained from Sigma-Aldrich (St. Louis, MO, United States). All media components of LB broth were purchased from Oxoid (Basingstoke, Hampshire, United Kingdom), and other reagents added to media were purchased from Sangon Biotech Co., Ltd. (Shanghai, China). PCR was performed with Q5 DNA polymerase (New England Biolabs, Ipswich, MA, United States) or T5 mix (Tsingke, Beijing, China). All PCR products were purified using a Wizard Genomic DNA Purification Kit (Promega, Madison, WI, United States). The restriction endonucleases *Xba*I, *Bbs*I, and CIP were purchased from New England Biolabs. A one-step cloning kit for two or more fragment assemblies was purchased from Vazyme Biotech Co., Ltd. (Nanjing, Jiangsu, China).

Plasmid Construction

Three CRISPR-*Fn*Cas12a systems (CRISPR-*Fn*Cas12a1, CRISPR-*Fn*Cas12a2 and CRISPR-*Fn*Cas12a3) were constructed in this study. The CRISPR-*Fn*Cas12a1 system was constructed based on pWHU2653 (Zeng et al., 2015). The CRISPR-*Fn*Cas12a2 system and the CRISPR-*Fn*Cas12a3 system were constructed based on pCRISPomyces-2 (Cobb et al., 2015). The CRISPR-*Fn*Cas12a1 system consisted of a counterselection marker *codA(sm)*. The CRISPR-*Fn*Cas12a systems were constructed in several steps. The *kasO*p* promoter (Wang et al., 2013), rpsLp(XC) promoter (Shao et al., 2013), *ermE*p* promoter (Luo et al., 2015a), and Potr* system (Wang et al., 2016) were amplified by PCR from template plasmids. *Streptomyces* codon-optimized *Fn*Cas12a was

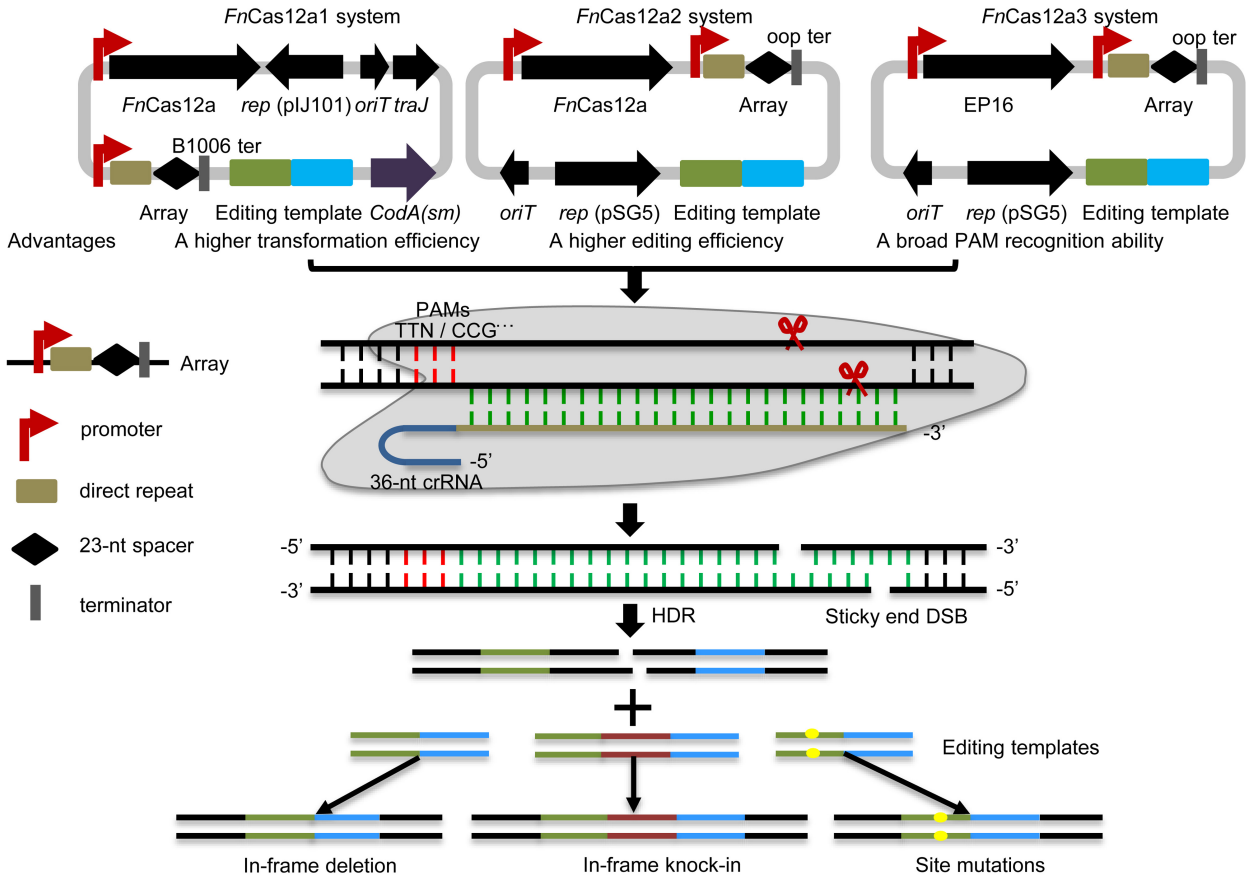


FIGURE 1 | Overview of the *Fn*Cas12a systems. Three *Fn*Cas12a systems were developed. The *Fn*Cas12a1 system carries a crRNA terminated by the B1006 terminator, *Fn*Cas12a, *codA(sm)*, *traJ*, *oriT*, and the pIJ101 replicon. The *Fn*Cas12a2 system carries a crRNA terminated by the oop terminator, *Fn*Cas12a, and partial sequences of *traJ*, oriT, and the pSG5 replicon. The *Fn*Cas12a3 system contains the same elements as the *Fn*Cas12a2 system, except that it contains the *Fn*Cas12a mutant EP16. *Fn*Cas12a1 and *Fn*Cas12a2 cleave DNA with TTN PAM recognition sites, while the *Fn*Cas12a3 system targets the YN (Y = C or T), TAC and CAA PAM recognition sites. These systems were applied to generate gene deletions, site mutations and insertions via HDR in *Streptomyces* strains. DSB, double-strand break.

chemically synthesized by Genewiz (Suzhou, Jiangsu, China). Other elements, including the *lacZ* cassette, unique *Bbs*I and *Xba*I restriction sites, temperature-sensitive pSG5 *rep* region, *aac(3)IV* coding sequence, *colE1* origin, and transfer *oriT* region, were amplified from the pWHU2653 or pCRISPomyces-2 plasmid. The yeast helper fragment was amplified from pRS416, which was a gift from Professor Huimin Zhao. All fragments were assembled into a plasmid with a one-step assembly kit or DNA assembler (Shao et al., 2009). A specific 23-nt spacer with a TTN PAM sequence located at the 5′ end of the coding strand was chosen. The PAM sequence together with the nearby 12-nt spacer sequence was analyzed using BLAST to confirm its specificity. The 19-nt or 36-nt direct repeat (DR) sequences and the 23-nt spacer targeting genes and gene clusters were synthesized and inserted into pYL-*Fn*Cas12a plasmids through Golden Gate assembly. Next, two 1- or 2-kb homologous arms corresponding to the upstream and downstream regions of the target genes or gene clusters were amplified from purified genomic DNA and fused into the *Xba*I site of the desired plasmid by one-step assembly. The presence of the correct plasmids was confirmed by sequencing (Tsingke, Beijing, China). All primers used in this study are listed in **Supplementary Table 1**.

Transformation

DNA was transformed into *E. coli* DH5α or ET12567/pUZ8002 using the heat shock method. Yeast transformation was performed by electroporation following a protocol described in a previous study (Shao et al., 2009). The conjugation of plasmids from *E. coli* ET12567/pUZ8002 to *Streptomyces* strains was performed following a previously described protocol (Kieser et al., 2000).

Screening of Potentially Edited Streptomyces Strains

Seven days after conjugation, seven single exconjugants were randomly picked, restreaked on R2YE or M-ISP4 agar plates supplemented with 20 μg/mL apramycin and 40 μg/mL nalidixic acid, and grown at 30°C for 7 days. Then, mycelia were collected from the plates for genomic DNA isolation using a bacterial

genomic DNA extraction kit (Tiangen, Beijing, China). Deletions were identified using PCR amplification of the genomic DNA or spores with primers located upstream and downstream of the target genes or primers annealing within or outside of the target gene clusters. Then, sequencing of the PCR products was performed to confirm the sequences (Tsingke, Beijing, China). For CRISPR-*Fn*Cas12a1 system clearance, M-ISP4 plates containing 800 μg/mL 5FC were used to cultivate the strains in the dark at 28°C for 3 or 4 days. For the CRISPR-*Fn*Cas12a2 system, plasmid clearance was carried out through high-temperature cultivation at 37°C for 2–3 days after normal cultivation at 30°C for 1 day.

Transcription Analysis of *Fn*Cas12a by Real-Time PCR (RT-qPCR)

Spores from every *Streptomyces* strain were separately inoculated into MYG medium for mycelium growth. After cultivation for 72 h, total RNA was extracted using TRIzol (Thermo Fisher, Waltham, MA, United States). Reverse transcription was completed using a first-strand cDNA synthesis kit (Bio-Rad, Carlsbad, CA, United States). SYBR Green PCR Master Mix (Bio-Rad) was used for RT-qPCR. Primers were designed with an online tool.[1] The reaction mixtures for RT-qPCR included 1 μL of cDNA templates, 1 μL of primers, 3 μL of ddH_2O and 5 μL of SYBR Green PCR Master Mix and were assayed using the following program: 2 min at 50°C and 3 min at 95°C for one cycle; 15 s at 95°C, 30 s at 62°C and 30 s at 72°C for 30 cycles; and 10 min at 72°C for a final cycle. The endogenous *hrdB* gene was used as an internal control. The transcription levels of the other genes were normalized to the control.

Fermentation and HPLC Analysis

After plasmid clearance, the culture of vegetative mycelium was spread over M-ISP4 plates and cultivated at 30°C for 4–7 days. The spores were then scraped and counted by spreading appropriate dilutions on plates and counting single colonies that had grown for 3 days. For daptomycin fermentation, 1×10^9 *Streptomyces* spores were inoculated into 15 mL of MYG medium and grown at 30°C for 48 h as the seed culture. Then, the seed culture was inoculated in 500 mL of fermentation medium (40 g/L dextrin, 5 g/L casein, 80 g/L glucose, and 5 g/L $MgSO_4$, pH 7.5) at 30°C for 9 days. After 48 h of fermentation, decanoate (final concentration 0.05%, w/v) was added every 12 h until the end of the fermentation period. After fermentation, the broth was centrifuged at 10,000 rpm for 30 min. The supernatant was then mixed with an equal volume of ethyl acetate to extract the metabolites. Next, the organic phase was dried under a vacuum. Finally, the metabolites were redissolved in 4 mL of methanol and filtered through a 0.22-μm membrane before the HPLC analysis.

The metabolites were analyzed on an HPLC system (Agilent Technologies Inc., Carpinteria, CA, United States) with an Agilent C18 reverse-phase column (internal diameter 4.6 mm × 250 mm, 5-μm particle size, Agilent Technologies, Inc.) at room temperature. The flow rate was 1 mL/min, and the absorbance of the eluate was monitored at 223 nm. The mobile phase, which was buffered with 0.01% formic acid, was initially maintained at a 75:25 water/acetonitrile for 5 min, followed by a linear gradient of 100% acetonitrile for 20 min.

[1]https://www.idtdna.com/scitools/Applications/RealTimePCR/

Statistical Analysis

The transformation frequency was calculated based on the ratio of the number of exconjugants to the number of spores used for conjugation. The experiments were performed in triplicate. The editing efficiency was calculated with the equation for editing efficiency (EF) = number of edited colonies evaluated by PCR/total number of evaluated colonies × 100%. All the colonies that showed double bands of WT and mutant products were calculated as wild type, and they were not included in the edited colonies. The editing efficiencies were calculated based on three replicates. Significant differences between the two groups were analyzed using *t* test. $p < 0.05$ was considered to indicate statistical significance.

RESULTS

Design and Construction of Three CRISPR-*Fn*Cas12a Systems

In previous studies, the pCRISPomyces-2 system (Cobb et al., 2015) and the pWHU2653 system (Zeng et al., 2015) have exhibited high editing efficiencies in model *Streptomyces* strains. To develop efficient and versatile genome editing tools in multiple *Streptomyces* strains, two *Fn*Cas12a systems were constructed: one based on the pWHU2653 system (Zeng et al., 2015) (CRISPR-*Fn*Cas12a1) and the other based on the pCRISPomyces-2 system (Cobb et al., 2015) (CRISPR-*Fn*Cas12a2). The *Fn*Cas12a1 system carries the selection markers *codA(sm)* (Dubeau et al., 2009) and *aac(3)IV*, the *rep*(pIJ101) replicon, and the B1006 terminator that terminates the transcription of the crRNA cassette, while the *Fn*Cas12a2 system carries only the selection marker *aac(3)IV*, the *pSG5* origin of replication, and the fd terminator that terminates the transcription of the crRNA cassette (**Supplementary Table 2**). Three constitutive promoters, *kasO*p* (Wang et al., 2013), rpsLp(XC) (Shao et al., 2013), and *ermE*p* (Luo et al., 2015a), and an inducible Potr* system were selected to control the expression of *Fn*Cas12a in each system. Overall, 8 plasmids were constructed: pYL-*kasO*p*-*Fn*Cas12a1, pYL-rpsLp(XC)-*Fn*Cas12a1, pYL-*ermE*p*-*Fn*Cas12a1, pYL-Potr*-*Fn*Cas12a1, pYL-*kasO*p*-*Fn*Cas12a2, pYL-rpsLp(XC)-*Fn*Cas12a2, pYL-*ermE*p*-*Fn*Cas12a2, and pYL-Potr*-*Fn*Cas12a2 (**Figures 2A,D**, **Supplementary Figure 1**, and **Table 3**). The CRISPR-*Fn*Cas12a3 system was constructed based on the *Fn*Cas12a2 system by replacing the WT *Fn*Cas12a protein with the *Fn*Cas12a variant EP16.

All the above plasmids were transformed into different *Streptomyces* strains. Over 10,000 exconjugants were generated by the transformation of *S. coelicolor* with pYL-Potr*-*Fn*Cas12a1 with or without addition of oxytetracycline (OTC), a value that was significantly greater than the number of exconjugants generated by plasmids pYL-*kasO*p*-*Fn*Cas12a1,

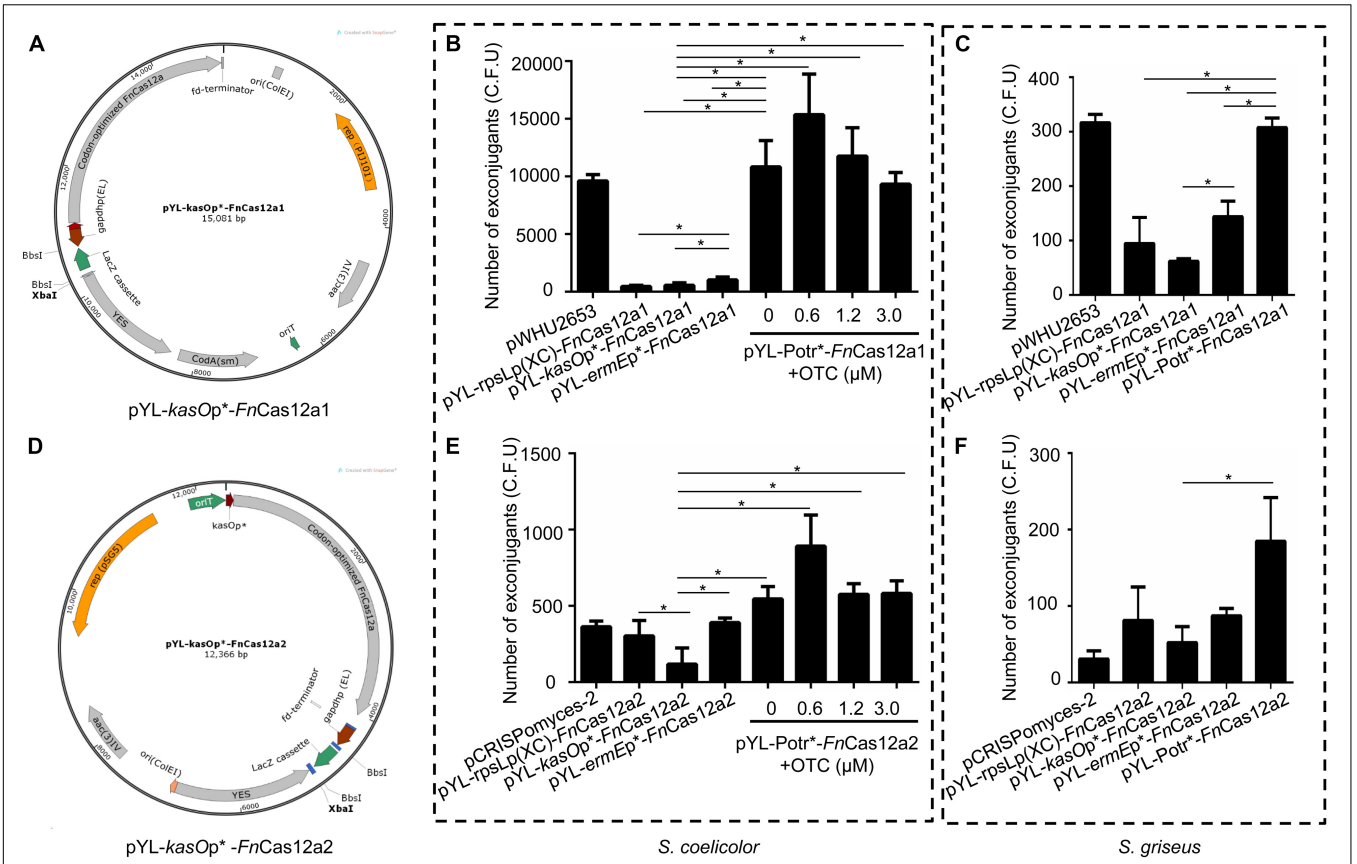


FIGURE 2 | *Fn*Cas12a1 and *Fn*Cas12a2 systems and their transformation frequencies. **(A)** pYL-*kasOp**-*Fn*Cas12a1 system. Notable features include the *kasOp**-driven *Fn*Cas12a, the pIJ101 replicon, the *aac(3)IV* selection marker, the *codA(sm)* selection marker, a yeast helper fragment, the *Bbs*I-flanked *lacZ* cassette controlled by the gapdhp(EL) promoter and the B1006 terminator. **(B,C)** The numbers of exconjugants on plates after transfer of *Fn*Cas12a1 system plasmids to *S. coelicolor* and *S. griseus* for 7 days. **(D)** pYL-*kasOp**-*Fn*Cas12a2 system. Most of the features of the pYL-*kasOp**-*Fn*Cas12a2 system are the same as those of the pYL-*kasOp**-*Fn*Cas12a1 system, excluding the temperature-sensitive pSG5 origin and the oop terminator of the *lacZ* cassette. The pYL-*kasOp**-*Fn*Cas12a2 system lacks a *traJ* gene. **(E–F)** The numbers of exconjugants on plates after transfer of *Fn*Cas12a2 system plasmids to *S. coelicolor* and *S. griseus* for 7 days. The experiments were performed in triplicate. * Indicates a significant difference.

pYL-rpsLp(XC)-*Fn*Cas12a1, and pYL-*ermEp**-*Fn*Cas12a1 (**Figure 2B**). Meanwhile, 0.6 μM OTC in M-ISP4 medium induced the greatest number of exconjugants (12453 ± 4582). Higher OTC concentrations resulted in fewer exconjugants. Similarly, introducing pYL-Potr*-*Fn*Cas12a1 into *S. griseus* generated 308 ± 13 exconjugants. Other *Fn*Cas12a1 plasmids were also transformed into *S. griseus* and generated 95 ± 39 (pYL-rpsLp(XC)-*Fn*Cas12a1), 62 ± 4 (pYL-*kasOp**-*Fn*Cas12a1), and 143 ± 23 (pYL-*ermEp**-*Fn*Cas12a1) exconjugants, respectively (**Figure 2C**).

A small number of *S. coelicolor* exconjugants (178 ± 27) were generated after introducing pYL-*kasOp**-*Fn*Cas12a2, while more exconjugants (303 ± 83 and 391 ± 21, respectively) were generated after introducing pYL-rpsLp(XC)-*Fn*Cas12a2, pYL-*ermEp**-*Fn*Cas12a2, and pYL-Potr*-*Fn*Cas12a2. The transformation of the plasmid pYL-Potr*-*Fn*Cas12a2 in the absence of OTC resulted in more exconjugants (546 ± 66) than the transformation of other *Fn*Cas12a2 system plasmids with constitutive promoters. Moreover, in the presence of OTC at a final concentration of 0.6 μM, the transformation of pYL-Potr*-*Fn*Cas12a2 generated the greatest number of exconjugants (892 ± 166) (**Figure 2E**). Similarly, the introduction of pYL-Potr*-*Fn*Cas12a2 into *S. griseus* generated the most exconjugants in the absence of OTC (**Figure 2F**). A comparison of editing efficiencies between two systems showed that all *Fn*Cas12a1 system plasmids exhibited significantly higher transformation frequencies than *Fn*Cas12a2 system plasmids. The transformation frequencies of the *Fn*Cas12a1 system plasmids were higher than those of the *Fn*Cas12a2 system plasmids in *S. griseus*. However, there were no statistically significant differences in *S. griseus* unlike *S. coelicolor* (**Supplementary Figure 2**).

Additionally, *Fn*Cas12a systems controlled by the promoters *ermEp** and Potr* showed significantly higher transformation frequencies than those controlled by *kasOp** (**Figures 2B,C,E,F**). These results are consistent with those of a study by Ungerer and Pakrasi, which revealed that a vector carrying *Fn*Cas12a yielded fewer colonies than an empty vector. *Fn*Cas12a is toxic to the host, and host toxicity increases with increasing *Fn*Cas12a promoter strength. Therefore, we chose the *Fn*Cas12a1 system

driven by the Potr* or *ermE*p* promoter to perform genome editing in *Streptomyces* strains that are difficult to be transformed.

Optimization of the CRISPR-*Fn*Cas12a Systems

To optimize the CRISPR-*Fn*Cas12a systems, we assessed the impacts of the DR length and the *Fn*Cas12a transcription level on genome editing efficiency. In some hosts, a shorter DR length resulted in a higher editing efficiency, as observed for *Fn*Cas12a in *Saccharomyces cerevisiae* (Swiat et al., 2017) and for *As*Cas12a in mammalian cells (Zetsche et al., 2017). In some other hosts, such as *Synechococcus* UTEX 2973 (Ungerer and Pakrasi, 2016), *Streptomyces* sp. NRRL S−244 (Yeo et al., 2019), and *Streptomyces albus* J1074 (Yeo et al., 2019), 36-nt DRs have shown high editing efficiency. However, there have been no reports focusing on the impact of the DR length on the editing efficiency in *Streptomyces* strains.

To confirm the impact of DR length on *Fn*Cas12a-mediated genome editing efficiency in *Streptomyces* strains, we chose two crRNA arrays with different DR lengths to guide *actII-orf4* gene editing in *S. coelicolor*. Array 1 contained a 19-nt DR and a 23-nt spacer, while array 2 contained a 36-nt DR and a 23-nt spacer (**Figure 3A**). CRISPR array 1 or 2 and an editing template were introduced into every pYL-*Fn*Cas12a plasmid to obtain a series of *actII-orf4* deletion plasmids (**Figure 3B** and **Supplementary Table 3**). When controlled by *kasO*p*, *Fn*Cas12a pairing with array 2 showed a significantly higher deletion efficiency than that pairing with array 1 for both *Fn*Cas12a1 and *Fn*Cas12a2 systems in *S. coelicolor* (**Table 1** and **Supplementary Figures 3A,B**). Specifically, plasmid pYL-*kasO*p*-*Fn*Cas12a2-*actII-orf4*-DR36 led to a loss of blue pigment and complete *actII-orf4* deletion with the highest editing efficiency of 100% (**Figure 3C** and **Table 1**). In addition, when driven by *kasO*p* or rpsLp(XC), the *Fn*Cas12a2 system showed significantly higher editing efficiencies than *Fn*Cas12a1 system for both 19-nt and 36-nt DRs (**Supplementary Figure 4**). In summary, pYL-*kasO*p*-*Fn*Cas12a2 carrying the 36-nt DR showed the highest and most stable editing efficiency among all *Fn*Cas12a plasmid combinations (**Table 1** and **Supplementary Figure 4**). Therefore, pYL-*kasO*p*-*Fn*Cas12a2 with a 36-nt DR was deemed the most suitable system for genome editing.

To explore the effects of *Fn*Cas12a expression level on genome editing efficiency, we measured the transcription levels of *Fn*Cas12a in *S. coelicolor*. We found that the transcription

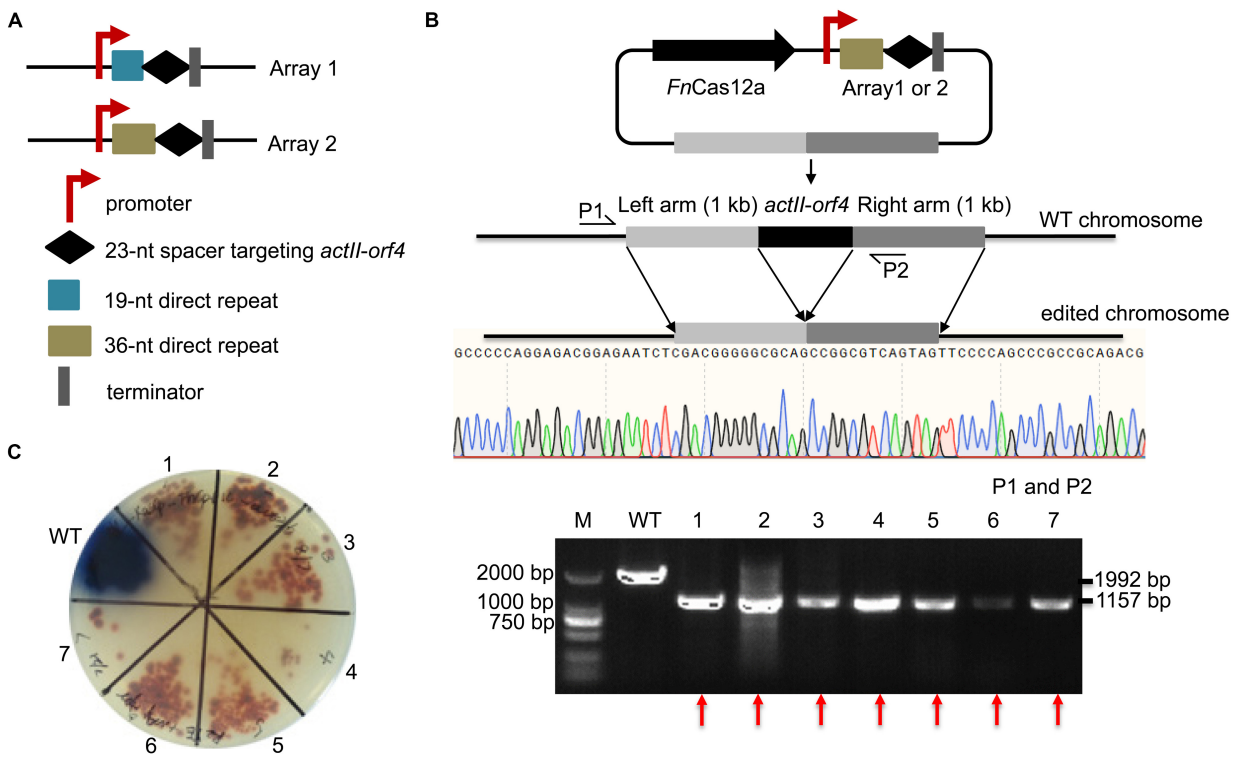


FIGURE 3 | Evaluation of the single-gene deletion efficiency. **(A)** Two arrays were designed for deletion of the *actII-orf4* gene. Array 1 contained the gapdhp(EL) promoter, a 19-nt DR sequence, a 23-nt spacer and the oop terminator. Array 2 contained similar elements, except that it contained a 36-nt DR. **(B)** Schematic of gene editing using *Fn*Cas12a with a single DR-spacer cassette and a single editing template. The editing template repaired the DSB cut induced by *Fn*Cas12a via HDR. Sequencing of PCR products amplified from the edited strains indicated the complete deletion of the *actII-orf4* gene with no undesired mutations. P1 and P2 indicate the locations of the primers used to validate the deletion. **(C)** Phenotypic evaluation and PCR evaluation of the WT strain *S. coelicolor* A3(2) and 7 randomly selected exconjugants. Red pigmentation was produced by the *actII-orf4*-deletion strains, but not by WT strains. The 1,992- and 1,157-bp replicons were amplified from the WT and *actII-orf4*-deletion strains, respectively. The experiments were performed in triplicate. Red arrows indicate the successfully deleted colonies.

TABLE 1 | Deletion efficiencies in the *actII-orf4* gene deletion experiments by using two *Fn*Cas12a systems.

Plasmid	Editing efficiency
pYL-*kasO*p*-*Fn*Cas12a1-*actII-orf4*-DR19	28.6% ± 11.7%
pYL-*kasO*p*-*Fn*Cas12a1-*actII-orf4*-DR36	76.2% ± 6.7%
pYL-rpsLp(XC)-*Fn*Cas12a1-*actII-orf4*-DR19	33.3% ± 6.7%
pYL-rpsLp(XC)-*Fn*Cas12a1-*actII-orf4*-DR36	71.4% ± 11.7%
pYL-*ermE*p*-*Fn*Cas12a1-*actII-orf4*-DR19	14.3% ± 11.7%
pYL-*ermE*p*-*Fn*Cas12a1-*actII-orf4*-DR36	23.8% ± 6.7%
pYL-Potr*-*Fn*Cas12a1-*actII-orf4*-DR19	9.5% ± 6.7%
pYL-Potr*-*Fn*Cas12a1-*actII-orf4*-DR36	28.6% ± 14.3%
pYL-*kasO*p*-*Fn*Cas12a2-*actII-orf4*-DR19	80.2% ± 6.3%
pYL-*kasO*p*-*Fn*Cas12a2-*actII-orf4*-DR36	100.0% ± 0.0%
pYL-rpsLp(XC)-*Fn*Cas12a2-*actII-orf4*-DR19	81.0% ± 17.8%
pYL-rpsLp(XC)-*Fn*Cas12a2-*actII-orf4*-DR36	94.4% ± 7.9%
pYL-*ermE*p*-*Fn*Cas12a2-*actII-orf4*-DR19	14.3% ± 0.0%
pYL-*ermE*p*-*Fn*Cas12a2-*actII-orf4*-DR36	28.6% ± 11.7%
pYL-Potr*-*Fn*Cas12a2-*actII-orf4*-DR19	23.8% ± 6.7%
pYL-Potr*-*Fn*Cas12a2-*actII-orf4*-DR36	21.4% ± 7.1%

levels of *Fn*Cas12a were significantly higher under the control of *kasO*p* than under the control of other promoters or the inducible system for both *Fn*Cas12a1 and *Fn*Cas12a2 systems (**Supplementary Figure 5**). With regard to editing efficiency, the plasmids carrying the 36-nt crRNA and *Fn*Cas12a driven by the strong promoter *kasO*p* resulted in higher editing efficiencies than those driven by a weak constitutive promoter (*ermE*p*) or an inducible system (Potr*) (**Table 1**, **Supplementary Figure 6**, and **Tables 3**, **4**). Notably, the inducible system has been reported to be fully induced by OTC at a final concentration of 3 μM in *S. coelicolor* M1146 (Wang et al., 2016). Thus, we performed transformations on M-ISP4 plates supplemented with OTC at final concentrations of 0, 0.6, 1.2, and 3.0 μM to maintain the Potr* system at low, medium and high activity levels. Among the four OTC concentrations, 3.0 or 1.2 μM OTC was sufficient to induce a high editing efficiency of pYL-Potr*-*Fn*Cas12a1-*actII-orf*-DR19 (23.8% ± 13.5%) or pYL-Potr*-*Fn*Cas12a1-*actII-orf*-DR36 (28.6% ± 11.7%). In the *Fn*Cas12a2 system, 1.2 μM OTC supplementation was sufficient to induce a high editing efficiency of pYL-Potr*-*Fn*Cas12a2-*actII-orf4*-DR19 (35.7% ± 7.1%) or pYL-Potr*-*Fn*Cas12a2-*actII-orf4*-DR36 (69.1% ± 2.4%) (**Supplementary Table 4**). Then, we measured the transcription levels of *Fn*Cas12a driven by the inducible Potr* system in the presence of 0, 0.6, 1.2 and 3.0 μM OTC. Significantly higher transcription levels of the *Fn*Cas12a were produced by both the *Fn*Cas12a1 and *Fn*Cas12a2 systems in the presence of 0.6, 1.2 and 3.0 μM OTC than in the absence of OTC (**Supplementary Figures 5A,B**). Pearson's correlation coefficients were calculated to explore the correlations between *Fn*Cas12a transcription levels and editing efficiencies. The analysis revealed that *Fn*Cas12a expression and the editing efficiency were positively correlated with the *Fn*Cas12a1 (correlation coefficient: $p = 0.050$, $r = 0.750$) and *Fn*Cas12a2 systems (correlation coefficient: $p = 0.036$, $r = 0.790$) when pairing with the 36-nt DR-containing crRNA (**Supplementary Figures 5C,D**). Briefly, elevated *Fn*Cas12a expression leads to increased genome editing efficiency of *Fn*Cas12a pairing with the 36-nt DR-containing crRNA.

The Ability of *Fn*Cas12a to Delete Large Chromosomal Fragments

pYL-*kasO*p*-*Fn*Cas12a2 carrying a 36-nt DR has been proven to be efficient for genome editing in *Streptomyces* strains. Thus, the ability of *Fn*Cas12a to delete large DNA fragments was evaluated. The actinorhodin biosynthesis gene cluster (ACT, 21.4 kb) and Ca^{2+}-dependent antibiotic biosynthesis gene cluster (CDA, 82.8 kb) in *S. coelicolor* and the daptomycin biosynthesis gene cluster (DAP, 127.6 kb) in *S. roseosporus* were selected. For single cuts, one to three spacers were selected for each cluster. The ACT-deletion strains produced red pigment on R2YE plates, while the WT strains produced blue pigment (**Figure 4A**). Subsequently, the PCR results showed that the 577-bp bands were amplified from the genomic DNA of WT strains but not from that of the edited strains. In addition, the 1,176-bp bands were amplified from the genomic DNA of successfully edited strains but were not amplified from the genomic DNA of WT strains (**Figure 4A**). The sequencing of the 1,176-bp fragments indicated that the ACT gene cluster was completely deleted from the genome of *S. coelicolor*. The deletion efficiency of the 21.4-kb gene fragment was 92.9% ± 7.2% (**Supplementary Table 5**). The efficiencies decreased with increasing deletion fragment sizes. The deletion efficiencies of the CDA gene cluster were 55.6% ± 7.9% (sp1), 18.8% ± 6.3% (sp2), and 25.0% ± 0% (sp3) (**Supplementary Table 5** and **Figure 4B**). The deletion efficiencies of the DAP gene cluster were 25.0% ± 0% (sp1) and 0% (sp2) (**Supplementary Table 5** and **Figure 4C**). To increase editing efficiency, we also attempted to introduce double cuts in the CDA gene cluster using an *Fn*Cas12a2 system carrying two crRNA cassettes (36-nt DR+2sp-1+36-nt DR+2sp-2). Two spacers (2sp-1 and 2sp-2) flanking the edges of the CDA gene cluster were selected to cut the two loci in the cluster. Two 2-kb arms homologous to the corresponding upstream and downstream sequences of the target gene clusters were introduced into the plasmids to repair the double-strand breaks at the edges of the CDA gene cluster. However, no increase in the editing efficiency was observed (25.0% ± 0%) (2sp) (**Supplementary Table 5** and **Figure 4B**).

Application of the CRISPR-*Fn*Cas12a System for Multiplex Genome Editing in Streptomyces Species

Editing multiple genes step by step in *Streptomyces* is labor-intensive and time-consuming since the *Streptomyces* growth cycle is relatively long. Thus, an efficient tool is needed to perform multigene editing. pYL-*kasO*p*-*Fn*Cas12a2 was used to assess the possibility of CRISPR/*Fn*Cas12a-mediated multiplex gene deletion. We constructed an *actII-orf4*/*redD* double-deletion mutant and an *actI-orf1*/*redX* double-deletion mutant based on the pYL-*kasO*p*-*Fn*Cas12a2 system. Both double-deletion constructs contained two arrays. Array 2-1 consisted of the gapdhp(EL) promoter, DR-spacer1 and the T7 terminator. Array 2-2 consisted of the rpsLp(XC) promoter, DR-spacer2 and oop

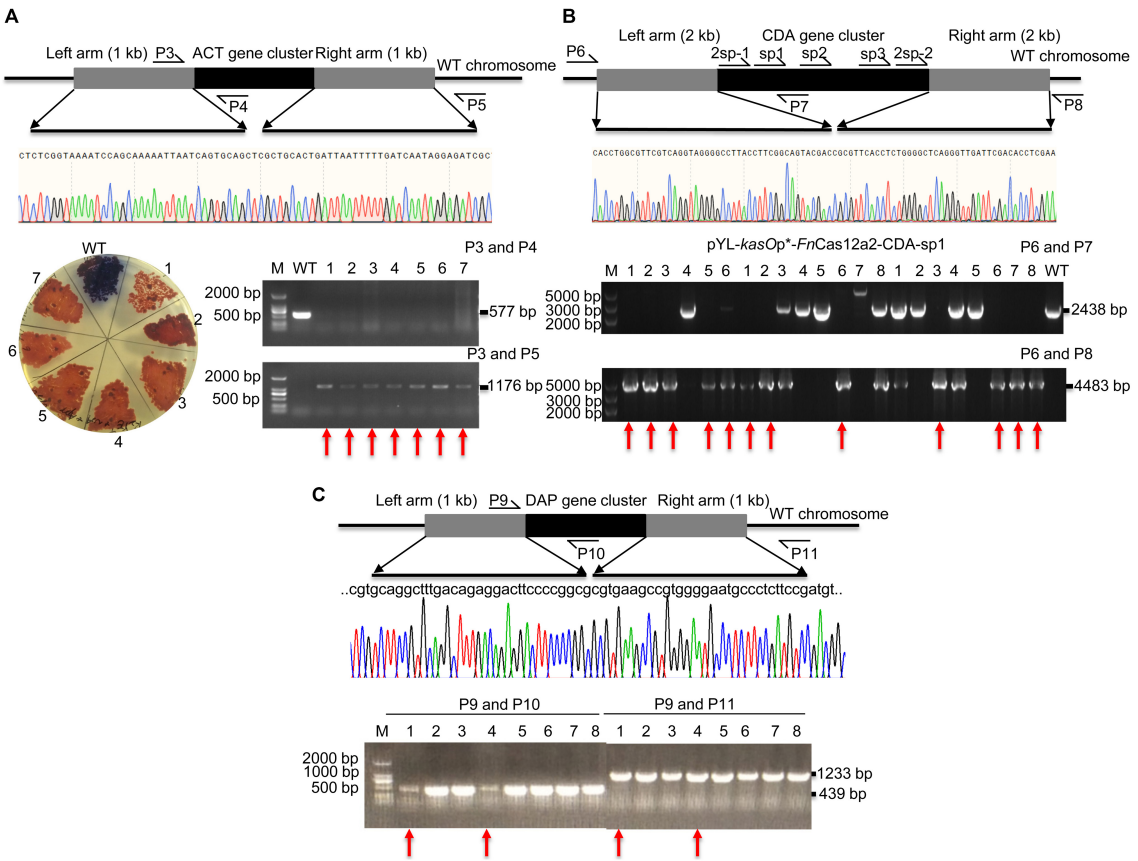


FIGURE 4 | Evaluation of a large-fragment deletion via the *Fn*Cas12a2 system. **(A)** Identification of the ACT gene cluster deletion. Red pigment was produced by the ACT-deletion strains but not by the WT strains. The primers P3 and P4 produced a 577-bp amplicon for the WT strains and no products for the ACT-deletion strains. The primers P3 and P5 generated an 1,176-bp amplicon for the ACT-deletion strains and no products for the WT strains. The sequencing data for the 1,176-bp amplicon showed the complete deletion of the ACT gene cluster. **(B)** Evaluation of CDA gene cluster deletion. For single cuts, three single spacers, sp1, sp2, and sp3, were selected. For double cuts, the paired spacers 2sp-1 and 2sp-2 were selected. The primers P6 and P7 produced a 2,438-bp amplicon for the WT strains and no products for the CDA-deletion strains. The primers P6 and P8 generated a 4,483-bp amplicon for the CDA-deletion strains, but no products for the WT strains. **(C)** Identification of DAP gene cluster deletion. The primers P9 and P10 generated a 439-bp amplicon for the WT strains, but no products for the DAP-deletion strains. The primers P9 and P11 generated a 1,233-bp amplicon for the DAP-deletion strains but no products for the WT strains. The sequencing data for the 1,233-bp amplicon showed the complete deletion of the DAP gene cluster. The experiments were performed in triplicate. Red arrows indicate the successfully deleted colonies.

terminator (**Figure 5A**). *ActII-orf4* and *redD* are pathway-specific regulatory genes, and *actI-orf1* and *redX* are beta-ketoacyl synthase genes. *ActII-orf4* and *actII-orf1* are responsible for actinorhodin biosynthesis, while *redD* and *redX* contribute to undecylprodigiosin biosynthesis in *S. coelicolor*. *S. coelicolor* lacking *actII-orf4* or *actI-orf1* failed to synthesize actinorhodin and produced a red pigment on R2YE medium. *RedD* or *redX* deletion abolished undecylprodigiosin synthesis and resulted in a blue color on R2YE medium (**Figures 5B,C**). The success of double deletion of *actII-orf4* and *redD* was evaluated using PCR. For the double deletions of *actII-orf4* and *redD*, the generation of an 1,157-bp band indicated the deletion of *actII-orf4*, and the amplification of an 1,102-bp band indicated the deletion of *redD*. Among the seven randomly selected exconjugants, exconjugant 4 harbored the correct double deletion (**Figure 5B**). For the double deletion of *actI-orf1* and *redX*, the amplification of an 1,129-bp band indicated the deletion of *actI-orf1*, and the amplification of a 2,283-bp band indicated the deletion of *redX*. Among the seven randomly selected exconjugants, exconjugant 1 harbored the correct double deletion, and colonies 3, 4, 6, and 7 showed mixtures of the *redX*-deletion mutant and the WT strain (**Figure 5C**). This phenomenon is common in microbes and has been called "incomplete genome editing" (Huang et al., 2015b; Ungerer and Pakrasi, 2016). The colonies that showed double bands with WT and mutant sizes were not classified as mutant colonies. The double-deletion efficiencies of *actII-orf4*/*redD* and *actI-orf1*/*redX* were 14.3% ± 0% and 14.3% ± 0%, respectively (**Supplementary Table 5** and **Supplementary Figure 7**).

Precise Genome Editing With the CRISPR-*Fn*Cas12a Systems

pYL-*ermE*p*-*Fn*Cas12a1 and pYL-Potr*-*Fn*Cas12a1 have been indicated to have increased transformation frequencies in

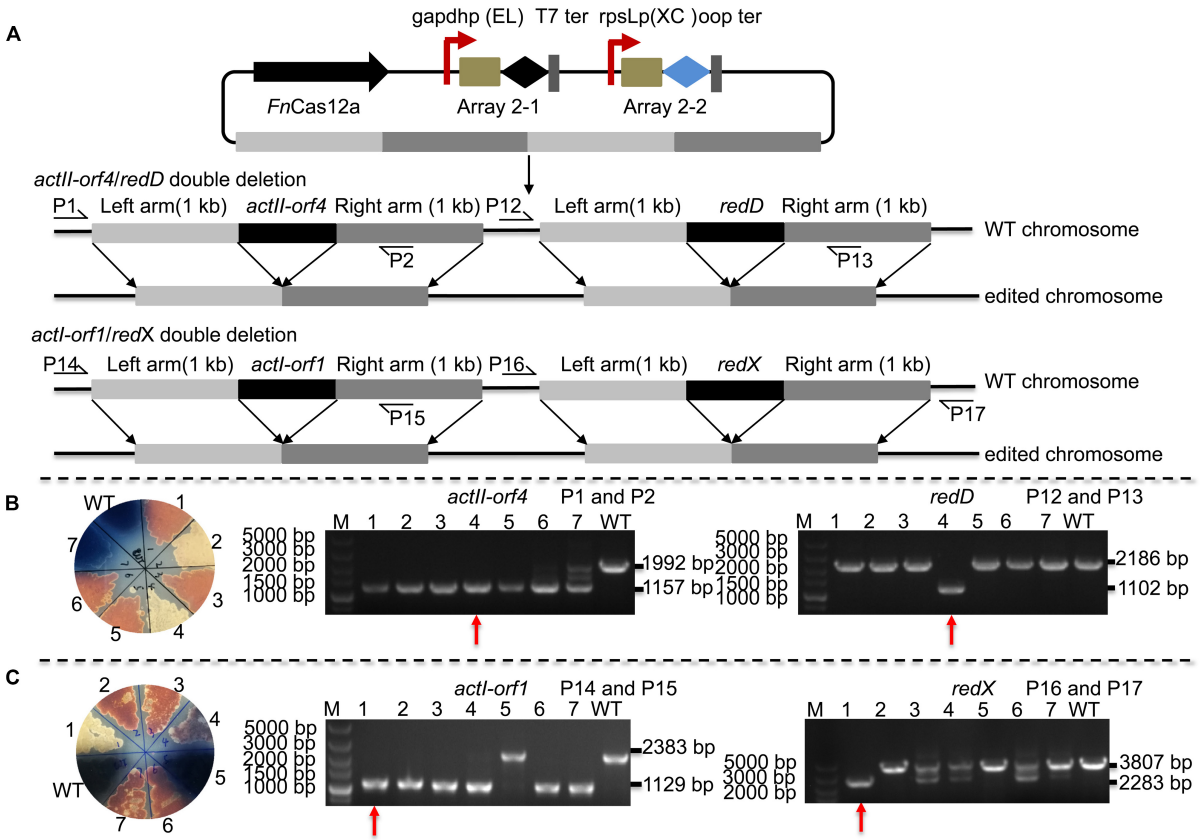


FIGURE 5 | Strategy and evaluation of double deletion via the *Fn*Cas12a2 system. **(A)** Strategy used for double deletion. Each plasmid for *actII-orf4/redD* double deletion or *actI-orf1/redX* double deletion carried two arrays and two editing templates. Array 2-1, including a 36-nt DR and spacer 1, was controlled by gapdhp(EL) and terminated by the T7 terminator; array 2-2, including a 36-nt DR and spacer 2, was controlled by rpsLp(XC) and terminated by the oop terminator. The template for every target consisted of left and right arms corresponding to upstream and downstream sequences of the target gene, respectively. P1, P2, P12, P13, P14, P15, P16, and P17 indicate the locations of the primers used to assess the deletions. **(B)** Seven randomly selected exconjugants carrying the *actII-orf4/redD* double-deletion plasmid were selected for verification by phenotype screening and PCR. The double-deletion mutant strains produced no pigment on R2YE. For the *actII-orf4/redD* double-deletion mutant strains, the primers P1 and P2 produced an 1,157-bp amplicon, while the primers P12 and P13 produced an 1,102-bp amplicon; in contrast, for the strains in which double deletion failed, the primers produced a 2,188-bp amplicon or a 1,992-bp amplicon. **(C)** Similarly, seven randomly selected exconjugants carrying the *actII-orf1/redX* double-deletion plasmid were selected for verification. For the *actI-orf1/redX* double-deletion mutant strains, the primers P14 and P15 produced an 1,129-bp amplicon, while the primers P16 and P17 produced a 2,283-bp amplicon; for the strains in which double deletion failed, the primers produced a 2,383-bp amplicon and a 3,807-bp amplicon. The double-deletion mutants are indicated by red arrows. The experiments were performed in triplicate.

Streptomyces. Thus, the genome editing efficiencies of these constructs were tested in *S. hygroscopicus* NRRL5491, in which it is difficult to perform genetic editing. Rapamycin is a macrolide immunosuppressant produced by *S. hygroscopicus* NRRL5491 and *Actinoplanes* sp. N902-109 (Huang et al., 2015a), and it has been approved as a treatment for select conditions by the FDA (Arriola Apelo and Lamming, 2016). In addition, rapamycin exhibits antifungal and anticancer activities. According to a previous study, *Actinoplanes* sp. N902-109 carries an additional *rapTH* gene and exhibits higher production of rapamycin than *S. hygroscopicus* NRRL5491. *RapTH* encodes a homolog of type II thioesterase (Huang et al., 2015a). The *rapTH* gene has been proposed to play an important role in rapamycin generation. To introduce *rapTH* into *S. hygroscopicus* NRRL5491, we inserted the rpsLp(CF) promoter in front of *rapTH* and inserted the *ermE*p* promoter in front of *rapQ* in constructs based on pYL-*ermE*p*-*Fn*Cas12a1 and pYL-Potr*-*Fn*Cas12a1 (**Supplementary Figure 8A**). The insertion fragment contained the *ermE*p* promoter (282 bp), the *rapTH* gene (756 bp) and the rpsLp(CF) promoter (302 bp) (a total of 1,340 bp) (**Supplementary Figure 8A**). Eight to twelve single colonies were selected and evaluated. A 2,277-bp band was amplified from the genomic DNA of WT strains, while a 3,668-bp band was amplified from genomic DNA of successfully edited strains (**Supplementary Figure 8B**). The maximum insertion efficiency of the two promoters and the *rapTH* gene was 25.0% when pYL-Potr*-*Fn*Cas12a1 was used in the presence of 1.2 μM OTC (**Supplementary Table 5**).

WT *Fn*Cas12a requires a TTN PAM sequence, which may limit its application in *Streptomyces* strains with high-GC-content genomes. For gene deletion, a TTN or NGG PAM can easily be selected in the whole gene reading frame. Thus,

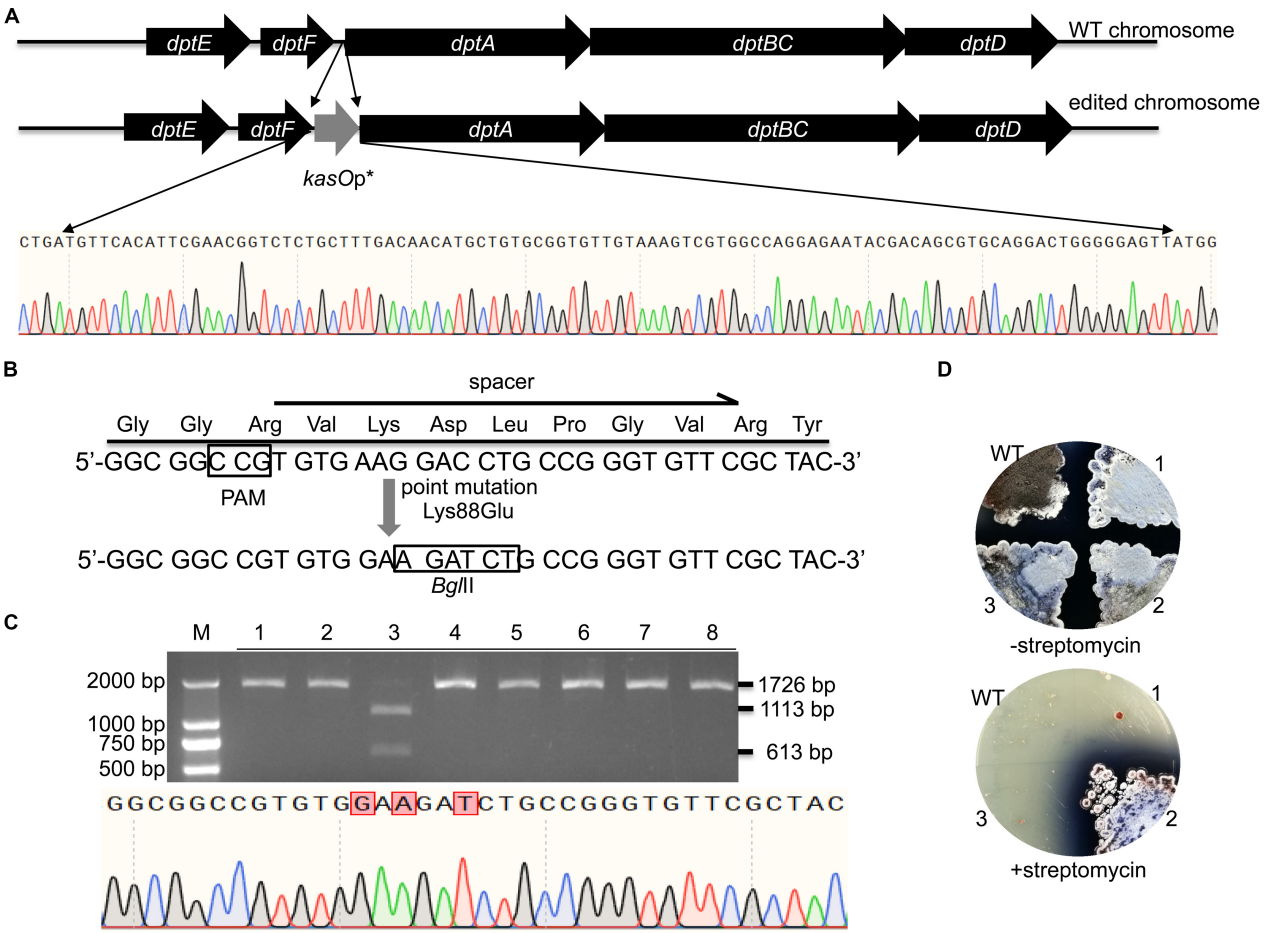


FIGURE 6 | Evaluation of CRISPR-*Fn*Cas12a3-mediated precise genome editing. **(A)** Strategy and evaluation of the *kasOp** insertion by CRISPR-*Fn*Cas12a3. The sequencing results showed the successful insertion of the promoter *kasOp**. **(B)** Evaluation of the precise mutations caused by CRISPR-*Fn*Cas12a3 with the CCG PAM. The Lys88Glu mutation and GAC-to-GAT mutations were introduced into the spacer region in the DSB repair template. **(C)** The 1,726-bp PCR products amplified from the mutants were digested into 1,113-bp and 613-bp fragments by *Bgl*II, while the WT amplicons were not digested. The sequencing data for the PCR product of the *rpsL* gene revealed the three mutations that are highlighted in red. **(D)** Three randomly selected exconjugants and one WT *S. coelicolor* strain were picked and streaked onto R2YE plates with or without 100 μg/mL streptomycin. Three days later, the WT strains were sensitive to 100 μg/mL streptomycin, while the mutant strains were resistant to 100 μg/mL streptomycin. The experiments were performed in triplicate.

there are no significant restrictions in the use of CRISPR-Cas9 or CRISPR-Cas12a. However, the selection region is sometimes limited for insertions or site mutations. Therefore, the use of an NGG or TTN PAM may be restricted. Daptomycin, a cyclic lipopeptide produced in *S. roseosporus*, shows significant activity against Gram-positive pathogens, such as methicillin-resistant *Staphylococcus aureus* (MRSA) (Vilhena and Bettencourt, 2012). The daptomycin biosynthetic pathway contains three nonribosomal peptide synthetase (NRPS) genes, *dptA*, *dptBC* and *dptD* (Miao et al., 2005). To increase the production of daptomycin in *S. roseosporus*, we planned to introduce *kasOp** in front of the *dptA* gene, which encodes the first subunit of NRPS (**Figure 6A**). However, no TTN PAM sequences suitable for *Fn*Cas12a recognition were located in front of the NRPS genes of daptomycin. Therefore, it was impossible for us to introduce a strong promoter in front of the NRPS genes via the *Fn*Cas12a system with WT *Fn*Cas12a. To overcome this limitation, we applied a CRISPR-Cas12a3 system containing the *Fn*Cas12a mutant EP16, which was previously generated by our laboratory (Wang et al., 2019). Among the tested variants, the *Fn*Cas12a mutant EP16 (N607R/K613V/N617R/K180S/K660R/D616N) exhibited the best recognition capabilities *in vitro*, as it could recognize YN (Y = C or T), TAC and CAA PAMs. We selected three spacers adjacent to CCG, CCA and ATC PAMs with which the *Fn*Cas12a variant EP16, but not WT *Fn*Cas12a, worked well *in vitro*. The *kasOp** promoter insertion efficiencies of CRISPR-*Fn*Cas12a3 with CCG, CCA and ATC PAMs were 50.0% ± 12.5%, 40.0% ± 17.4%, and 23.6% ± 8.6%, respectively (**Supplementary Table 5**). In summary, the CRISPR-*Fn*Cas12a3 system provides opportunities to select suitable PAMs for precise genome editing in GC-rich organisms.

After we introduced *kasOp** in front of the daptomycin biosynthesis genes, we performed plasmid clearance to generate

the strain *S. roseosporus*/PkasA (**Figure 9A**). Introduction of *kasOp** increased the transcription levels of the *dptA*, *dptBC* and *dptD* genes by 1. 83-, 1.12- and 2.04-fold, respectively (**Supplementary Figure 9A**). Compared with WT *S. roseosporus*, *S. roseosporus*/PkasA showed increased production of daptomycin (**Supplementary Figures 9B,C**).

In addition to insertions, CRISPR-*Fn*Cas12a3 was also applied to generate precise site mutations. *RpsL*(SCO4659) encodes the ribosomal protein S12 in *S. coelicolor* A3(2), and its specific site mutation (K88E) confers resistance to high concentrations of streptomycin (100 μg/mL) (Shima et al., 1996). Here, we introduced a GAA mutation (K88E) to ensure the resistance of *S. coelicolor* A3(2) to streptomycin (**Figure 6B**). Moreover, a specific C-to-T mutation was introduced at the 267th nucleotide of *rpsL* to introduce a *Bgl*II restriction site for rapid identification of the correct exconjugants (**Figure 6B**). To avoid undesirable re-editing, the double-strand break (DSB) was repaired using a template containing altered nucleotides at the 5′ end of the spacer sequence (**Figure 6B**). PCR amplicons of 1,726 bp were amplified from the genomes of eight exconjugants and digested separately with the *Bgl*II restriction enzyme. One 1,726-bp PCR amplicon was successfully digested into 1,113-bp and 613-bp fragments. Subsequent sequencing of the 1,726-bp PCR amplicon from exconjugant 3 indicated that the site mutations were successfully created (**Figure 6C**). The restriction digestion and DNA sequencing results showed a site mutation efficiency 12.5% ± 0% (**Figure 6C** and **Supplementary Table 5**). However, phenotype screening indicated that 33.3% of the *S. coelicolor* A3(2) colonies harboring pYL-*kasOp**-*Fn*Cas12a3-rpsL were resistant to streptomycin (**Figure 6D**). These results suggest that some strains failed to be mutated.

DISCUSSION

In the present study, we developed three useful genome editing tools with different advantages and applied them to examine their transformation frequencies and genome editing efficiencies in several *Streptomyces* species. The successful application of the CRISPR-*Fn*Cas12a3 system with expanded PAM recognition ability for precise insertions and site mutations overcomes the restricted applications of the TTN PAM in organisms with high GC contents. Its flexibility in PAM selection will promote the application of *Fn*Cas12a in *Streptomyces* species. The *Fn*Cas12a1 system worked well in *S. hygroscopicus*, in which CRISPR-Cas9 is ineffective. The *Fn*Cas12a2 system efficiently deleted large chromosomal fragments (∼128 kb) and was useful for deleting multiple genes. Altogether, these three systems have different advantages and are complementary to each other.

Due to the high GC content of the *Streptomyces* genome, it will not be easy for us to select a TTN PAM with high prediction scores. In particular, the number of suitable sequence region that can be targeted for insertions and site mutations is very limited. Thus, it is inconvenient to select a TTN PAM for large-scale genome engineering in *Streptomyces*. The *Fn*Cas12a mutant EP16 has a wide range of PAM recognition sites (60/64 sites), including YN (Y = C or T), TAC and CAA sites. EP16 cleaves target DNA by the PAMs CCG, CCA and ATC *in vitro* with high efficiencies of 97, 94 and 96%, respectively. However, the efficiencies of CRISPR-*Fn*Cas12a3 harboring EP16 with CCG, CCA and ATC recognition sites were only 50.0% ± 12.5%, 40.0% ± 17.4%, and 23.6% ± 8.6%, respectively, *in vivo*. This phenomenon is common. As shown in the studies by Zetsche et al. and Tu et al., *Fn*Cas12a recognizes TTN sites *in vitro* but frequently fails to recognize TTN sites in human cells (Zetsche et al., 2015; Tu et al., 2017). *Fn*Cas12a has been reported to prefer KYTV in human cells (Sun et al., 2018), but prefer TTTV PAMs in rice (Zhong et al., 2018). Moreover, our unpublished data from human HEK293T cells also show a similar phenomenon. This phenomenon may be attributable to the complex microenvironments in living organisms. For instance, post-translational modifications, such as acetylation (Ishigaki et al., 2017) and methylation (Huang et al., 2018), are common in *Streptomyces*. *Moraxella bovoculi* (Mb) Cas12a has been demonstrated to lose its cleavage functions upon the acetylation of the critical PAM recognition residue Lys635 (Dong et al., 2019) *in vitro*. Taken together, results indicate that the *Fn*Cas12a3 system overcomes the restricted applications of WT *Fn*Cas12a, which requires a TTN PAM for editing in *Streptomyces*. Although the editing efficiencies of the *Fn*Cas12a3 system were low, the identification of a new Cas12a variant (EP16) that works on a broad range of PAMs in *Streptomyces* is a very important step forward. This powerful tool will enable researchers to generate desired insertions and precise site mutations and will also enable the activation of biosynthetic pathways to generate valuable natural products.

In a previous study, conservation was discovered at the 3′ end of the DR sequence among all *Fn*Cas12a family proteins (Zetsche et al., 2015). In addition, 19-nt DR-containing crRNA cassettes have been proven to exhibit good editing efficiencies *in vitro* (Wang et al., 2019) and in some host cells, such as human HEK293T cells (Zetsche et al., 2017) and *S. cerevisiae* (Swiat et al., 2017). In another study, a 36-nt DR was applied for markerless editing in *Cyanobacteria* species (Ungerer and Pakrasi, 2016). Although *Fn*Cas12a has been used in many hosts, to the best of our knowledge, only one report has focused on the impacts of the DR length on the editing efficiency in *Saccharomyces cerevisiae* (Swiat et al., 2017). In the current study, a crRNA containing a 19-nt DR led to much higher editing efficiency than that containing a 36-nt DR in *Saccharomyces cerevisiae*. In *Streptomyces* species, when *Fn*Cas12a was controlled by strong constitutive promoters, the crRNA containing a 36-nt DR led to high editing efficiency. However, when *Fn*Cas12a was controlled by a weak promoter or an inducible system, significant differences in the editing efficiencies resulting from 36-nt DR-containing crRNA and 19-nt DR-containing crRNA were not observed (**Table 1** and **Supplementary Figure 3**). Therefore, the impact of the DR length is different for different plasmid systems or hosts. Thus, the impact of the DR length should be evaluated whenever Cas12a-mediated genome editing is performed for the first time in a specific organism.

The transformation frequencies and editing efficiencies of the *Fn*Cas12a1 and *Fn*Cas12a2 systems were significantly different. The plasmids constructed from the *Fn*Cas12a1 system

induced markedly higher transformation frequencies than those constructed from the *Fn*Cas12a2 system. A comparison of the elements related to transformation revealed that the *Fn*Cas12a1 system carries a *traJ* gene. This gene encodes an activator of the transfer (*tra*) operon, which is crucial for the transfer region of the fertility factor (Will and Frost, 2006). After the *traJ* gene was deleted from the *Fn*Cas12a1 system, the transformation frequency was significantly decreased (**Supplementary Figure 10**). In this case, the *traJ* gene in *Fn*Cas12a1 system was required for its high transformation frequency. Moreover, the *Fn*Cas12a1 system contains the counterselection marker *codA(sm)*, which saves a substantial amount of time during plasmid elimination (Zeng et al., 2015). Thus, the *Fn*Cas12a1 system is more suitable than the *Fn*Cas12a2 system for applications in some strains with low transformation frequencies. On the other hand, Wei et al. (2017) found that strong terminators are responsible for high gene expression. To explore whether the different terminators of crRNA arrays contribute to the different editing efficiencies between *Fn*Cas12a1 system and *Fn*Cas12a2 system, we exchanged the terminators of crRNA arrays of the two systems to construct the plasmids pYL-*kasO*p*-*Fn*Cas12a1-*actII-orf4*-DR36-oop and pYL-*kasO*p*-*Fn*Cas12a2-*actII-orf4*-DR36-B1006. The pYL-*kasO*p*-*Fn*Cas12a1-*actII-orf4*-DR36-oop plasmid led to significantly higher editing efficiency (95.2% ± 6.7%) than the pYL-*kasO*p*-*Fn*Cas12a1-*actII-orf4*-DR36 plasmid (76.2% ± 6.7%), and pYL-*kasO*p*-*Fn*Cas12a2-*actII-orf4*-DR36-B1006 led to a significantly lower efficiency (47.6% ± 17.8%) than pYL-*kasO*p*-*Fn*Cas12a2-*actII-orf4*-DR36 (100%) (**Supplementary Figure 11**). Thus, the oop terminator of the crRNA array was responsible for the higher editing efficiency of the *Fn*Cas12a system. In conclusion, the two systems described in the present study provide researchers with additional choices for manipulation of different *Streptomyces* strains.

Genome editing and transcriptional repression in *Streptomyces* with the CRISPR-Cas12a system have also recently been reported by Li et al. (2018). Those researchers obtained an editing efficiency for single-gene deletion with a CRISPR-Cas12a system of 95%, which was lower than that obtained with our *Fn*Cas12a2 system (100%). However, the editing efficiencies of pYL-*ermE*p*-*Fn*Cas12a1-*actII-orf4-DR19* and pYL-*ermE*p*-*Fn*Cas12a2-*actII-orf4-DR19* were lower than those of their CRISPR-*Fn*Cas12a system carrying Cas12a controlled by the same promoter, *ermE*p*. Several factors may explain this discrepancy. First, the promoter controlling the crRNA cassette in the CRISPR-Cas12a system described by Lei Li et al. was *kasO*p*, which is stronger than the *gapdh*(EL) (Shao et al., 2013; Myronovskyi and Luzhetskyy, 2016) promoter used in our study. The high crRNA expression level may have increased the editing efficiency. Second, the strains were different. The authors of the previous study used *S. hygroscopicus* SIPI-KF and *E. coli* S17-1. *E. coli* S17-1 contains a chromosomally integrated derivative of RP4, which stimulates the integration of DNA from the donor strain into the recipient genome (Simon et al., 1983; Voeykova et al., 1998; Kieser et al., 2000). Thus, the status of the *Fn*Cas12a system in *Streptomyces* may be influenced by different donor strains. Importantly, *Fn*Cas12a was successfully applied in the current study to accurately delete large chromosomal DNA fragments ranging from 24 to 128 kb with efficiencies ranging from 25% to 92.9% ± 7.2% based on HDR, and the editing efficiency obtained using HDR was much higher than that obtained using NHEJ repair (10%, 27.6 kb). Thus, HDR was more suitable than NHEJ for deleting large DNA fragments. Moreover, the *Fn*Cas12a3 system has an expanded PAM recognition ability, which might provide increased opportunities for researchers to conduct precise genome editing. This system will be particularly useful for generating insertions and site mutations, as it facilitates selection of PAMs in the limited DNA sequence regions in *Streptomyces* strains with a high GC content. Therefore, the three *Fn*Cas12a systems we developed are versatile tools for precise genome editing in different *Streptomyces* strains.

In summary, the engineering tools developed in the present study are applicable for biosynthetic pathway reconstruction, metabolic engineering, and chassis cell construction in different *Streptomyces* strains. These tools will also be beneficial for natural product discovery and overproduction.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

JZha, HL, SL, and YL designed the experiments. JZha, DZ, and JZhu performed the experiments. DZ, JZha, and YL wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Natural Science Foundation of Tianjin City (19JCYBJC24200), the National Key R&D Program of China (2018YFA0903300), the National Natural Science Foundation of China (81502966) the Projects of International Cooperation and Exchanges of NSFC (31961143005), and the International Cooperation of Sichuan Province Scientific Committee (20GJHZ0190).

ACKNOWLEDGMENTS

We are grateful to Professor Huimin Zhao from the University of Illinois at Urbana-Champaign for providing the plasmid pCRISPomyces-2 and Professor Yuhui Sun from Wuhan University for kindly sharing the plasmid pWHU2653.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00726/full#supplementary-material

REFERENCES

Arriola Apelo, S. I., and Lamming, D. W. (2016). Rapamycin: an inhibiTOR of aging emerges from the soil of Easter island. *J. Gerontol. A. Biol. Sci. Med. Sci.* 71, 841–849. doi: 10.1093/gerona/glw090

Baltz, R. H. (2008). Renaissance in antibacterial discovery from actinomycetes. *Curr. Opin. Pharmacol.* 8, 557–563. doi: 10.1016/j.coph.2008.04.008

Cho, G., Kim, J., Park, C. G., Nislow, C., Weller, D. M., and Kwak, Y. S. (2017). Caryolan-1-ol, an antifungal volatile produced by *Streptomyces* spp., inhibits the endomembrane system of fungi. *Open Biol.* 7:170075. doi: 10.1098/rsob.170075

Cobb, R. E., Wang, Y., and Zhao, H. (2015). High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS Synth. Biol.* 4, 723–728. doi: 10.1021/sb500351f

Dong, L., Guan, X., Li, N., Zhang, F., Zhu, Y., Ren, K., et al. (2019). An anti-CRISPR protein disables type V Cas12a by acetylation. *Nat. Struct. Mol. Biol.* 26, 308–314. doi: 10.1038/s41594-019-0206-1

Dubeau, M. P., Ghinet, M. G., Jacques, P. E., Clermont, N., Beaulieu, C., and Brzezinski, R. (2009). Cytosine deaminase as a negative selection marker for gene disruption and replacement in the genus *Streptomyces* and other actinobacteria. *Appl. Environ. Microbiol.* 75, 1211–1214. doi: 10.1128/AEM.02139-08

Frattaruolo, L., Lacret, R., Cappello, A. R., and Truman, A. W. (2017). A genomics-based approach identifies a thioviridamide-like compound with selective anticancer activity. *ACS Chem. Biol.* 12, 2815–2822. doi: 10.1021/acschembio.7b00677

Huang, H., Ren, S. X., Yang, S., and Hu, H. F. (2015a). Comparative analysis of rapamycin biosynthesis clusters between *Actinoplanes* sp. N902-109 and *Streptomyces hygroscopicus* ATCC29253. *Chin. J. Nat. Med.* 13, 90–98. doi: 10.1016/S1875-5364(15)60012-7

Huang, H., Zheng, G., Jiang, W., Hu, H., and Lu, Y. (2015b). One-step high-efficiency CRISPR/Cas9-mediated genome editing in *Streptomyces*. *Acta Biochim. Biophys. Sin.* 47, 231–243. doi: 10.1093/abbs/gmv007

Huang, T., Duan, Y., Zou, Y., Deng, Z., and Lin, S. (2018). NRPS protein MarQ catalyzes flexible adenylation and specific S-methylation. *ACS Chem. Biol.* 13, 2387–2391. doi: 10.1021/acschembio.8b00364

Ishigaki, Y., Akanuma, G., Yoshida, M., Horinouchi, S., Kosono, S., and Ohnishi, Y. (2017). Protein acetylation involved in streptomycin biosynthesis in *Streptomyces griseus*. *J. Proteomics* 155, 63–72. doi: 10.1016/j.jprot.2016.12.006

Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L. A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* 31, 233–239. doi: 10.1038/nbt.2508

Jiang, Y., Qian, F., Yang, J., Liu, Y., Dong, F., Xu, C., et al. (2017). CRISPR-Cpf1 assisted genome editing of *Corynebacterium glutamicum*. *Nat. Commun.* 8:15179. doi: 10.1038/ncomms15179

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2000). *Practical Streptomyces Genetics*. Norwich: John Innes Foundation.

Li, L., Wei, K., Zheng, G., Liu, X., Chen, S., Jiang, W., et al. (2018). CRISPR-Cpf1-assisted multiplex genome editing and transcriptional repression in *Streptomyces*. *Appl. Environ. Microbiol.* 84:e00827-18. doi: 10.1128/AEM.00827-18

Luo, Y., Enghiad, B., and Zhao, H. (2016). New tools for reconstruction and heterologous expression of natural product biosynthetic gene clusters. *Nat. Prod. Rep.* 33, 174–182. doi: 10.1039/c5np00085h

Luo, Y., Zhang, L., Barton, K. W., and Zhao, H. (2015a). Systematic identification of a panel of strong constitutive promoters from *Streptomyces albus*. *ACS Synth. Biol.* 4, 1001–1010. doi: 10.1021/acssynbio.5b00016

Luo, Y., Li, B. Z., Liu, D., Zhang, L., Chen, Y., Jia, B., et al. (2015b). Engineered biosynthesis of natural products in heterologous hosts. *Chem. Soc. Rev.* 44, 5265–5290. doi: 10.1039/c5cs00025d

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., et al. (2013). RNA-guided human genome engineering via Cas9. *Science* 339, 823–826. doi: 10.1126/science.1232033

Miao, V., Coeffet-Legal, M. F., Brian, P., Brost, R., Penn, J., Whiting, A., et al. (2005). Daptomycin biosynthesis in *Streptomyces roseosporus*: cloning and analysis of the gene cluster and revision of peptide stereochemistry. *Microbiology* 151(Pt 5), 1507–1523. doi: 10.1099/mic.0.27757-0

Myronovskyi, M., and Luzhetskyy, A. (2016). Native and engineered promoters in natural product discovery. *Nat. Prod. Rep.* 33, 1006–1019. doi: 10.1039/c6np00002a

Salem, S. M., Weidenbach, S., and Rohr, J. (2017). Two cooperative glycosyltransferases are responsible for the sugar diversity of saquayamycins isolated from *Streptomyces* sp. KY 40-1. *ACS Chem. Biol.* 12, 2529–2534. doi: 10.1021/acschembio.7b00453

Shao, Z., Rao, G., Li, C., Abil, Z., Luo, Y., and Zhao, H. (2013). Refactoring the silent spectinabilin gene cluster using a plug-and-play scaffold. *ACS Synth. Biol.* 2, 662–669. doi: 10.1021/sb400058n

Shao, Z., Zhao, H., and Zhao, H. (2009). DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res.* 37:e16. doi: 10.1093/nar/gkn991

Shima, J., Hesketh, A., Okamoto, S., Kawamoto, S., and Ochi, K. (1996). Induction of actinorhodin production by rpsL (encoding ribosomal protein S12) mutations that confer streptomycin resistance in *Streptomyces lividans* and *Streptomyces coelicolor* A3(2). *J. Bacteriol.* 178, 7276–7284. doi: 10.1128/jb.178.24.7276-7284.1996

Simon, R., Priefer, U., and Pühler, A. (1983). A broad host range mobilization system for *in vivo* genetic engineering transposon mutagenesis in Gram negative bacteria. *Nat. Biotechnol.* 1, 784–791. doi: 10.1038/nbt1183-784

Sun, H. H., Li, F. F., Liu, J., Yang, F. Y., Zeng, Z. H., Lv, X. J., et al. (2018). A single multiplex crRNA array for FnCpf1-mediated human genome editing. *Mol. Ther.* 26, 2070–2076. doi: 10.1016/j.ymthe.2018.05.021

Swiat, M. A., Dashko, S., den Ridder, M., Wijsman, M., van der Oost, J., Daran, J. M., et al. (2017). FnCpf1: a novel and efficient genome editing tool for *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 45, 12585–12598. doi: 10.1093/nar/gkx1007

Tang, X., Lowder, L. G., Zhang, T., Malzahn, A. A., Zheng, X., Voytas, D. F., et al. (2017). A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nat. Plants* 3:17018. doi: 10.1038/nplants.2017.18

Toth, E., Weinhardt, N., Bencsura, P., Huszar, K., Kulcsar, P. I., Talas, A., et al. (2016). Cpf1 nucleases demonstrate robust activity to induce DNA modification by exploiting homology directed repair pathways in mammalian cells. *Biol. Direct.* 11:46. doi: 10.1186/s13062-016-0147-0

Tu, M. J., Lin, L., Cheng, Y. L., He, X. B., Sun, H. H., Xie, H. H., et al. (2017). A 'new lease of life': FnCpf1 possesses DNA cleavage activity for genome editing in human cells. *Nucleic Acids Res.* 45, 11295–11304. doi: 10.1093/nar/gkx783

Ungerer, J., and Pakrasi, H. B. (2016). Cpf1 is a versatile tool for CRISPR genome editing across diverse species of Cyanobacteria. *Sci. Rep.* 6:39681. doi: 10.1038/srep39681

Vilhena, C., and Bettencourt, A. (2012). Daptomycin: a review of properties, clinical use, drug delivery and resistance. *Mini Rev. Med. Chem.* 12, 202–209. doi: 10.2174/138955711209030202

Voeykova, T., Emelyanova, L., Tabakov, V., and Mkrtumyan, N. (1998). Transfer of plasmid pTO1 from *Escherichia coli* to various representatives of the order Actinomycetales by intergeneric conjugation. *FEMS Microbiol. Lett.* 162, 47–52. doi: 10.1016/s0378-1097(98)00100-1

Wang, L., Wang, H., Liu, H., Zhao, Q., Liu, B., Wang, L., et al. (2019). Improved CRISPR-Cas12a-assisted one-pot DNA editing method enables seamless DNA editing. *Biotechnol. Bioeng.* 116, 1463–1474. doi: 10.1002/bit.26938

Wang, W., Yang, T., Li, Y., Li, S., Yin, S., Styles, K., et al. (2016). Development of a synthetic oxytetracycline-inducible expression system for *Streptomycetes* using de novo characterized genetic parts. *ACS. Synth. Biol.* 5, 765–773. doi: 10.1021/acssynbio.6b00087

Wang, W. L., Wang, J., Xiang, S., Feng, X., and Yang, K. (2013). An engineered strong promoter for *Streptomycetes*. *Appl. Environ. Microbiol.* 79, 4484–4492. doi: 10.1128/aem.00985-13

Wei, L. N., Wang, Z. X., Zhang, G. L., and Ye, B. E. (2017). Characterization of terminators in *Saccharomyces cerevisiae* and an exploration of factors affecting their strength. *Chembiochem* 18, 2422–2427. doi: 10.1002/cbic.201700516

Will, W. R., and Frost, L. S. (2006). Characterization of the opposing roles of H-NS and TraJ in transcriptional regulation of the F-plasmid tra operon. *J. Bacteriol.* 188, 507–514. doi: 10.1128/JB.188.2.507-514.2006

Yan, X. T., An, Z., Huangfu, Y., Zhang, Y. T., Li, C. H., Chen, X., et al. (2018). Polycyclic polyprenylated acylphloroglucinol and phenolic metabolites from the aerial parts of Hypericum elatoides and their neuroprotective and anti-neuroinflammatory activities. *Phytochemistry* 159, 65–74. doi: 10.1016/j.phytochem.2018.12.011

Yeo, W. L., Heng, E., Tan, L. L., Lim, Y. W., Lim, Y. H., Hoon, S., et al. (2019). Characterization of Cas proteins for CRISPR-Cas editing in streptomycetes. *Biotechnol. Bioeng.* 116, 2330–2338. doi: 10.1002/bit.27021

Zeng, H., Wen, S., Xu, W., He, Z., Zhai, G., Liu, Y., et al. (2015). Highly efficient editing of the actinorhodin polyketide chain length factor gene in *Streptomyces coelicolor* M145 using CRISPR/Cas9-CodA(sm) combined system. *Appl. Microbiol. Biotechnol.* 99, 10575–10585. doi: 10.1007/s00253-015-6931-4

Zetsche, B., Gootenberg, J., Abudayyeh, O., Slaymaker, M., Makarova, K., Essletzbichler, P., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a Class 2 CRISPR-Cas system. *Cell* 163, 759–771. doi: 10.1016/j.cell.2015.09.038

Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E. M., et al. (2017). Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat. Biotechnol.* 35, 31–34. doi: 10.1038/nbt.3737

Zhao, Y., Wang, L. P., and Luo, Y. Z. (2019). Recent advances in natural products exploitation in *Streptomyces* via synthetic biology. *Eng. Life Sci.* 19, 452–462. doi: 10.1002/elsc.201800137

Zhong, Z. H., Zhang, Y. X., You, Q., Tang, X., Ren, Q. R., Liu, S. S., et al. (2018). Plant genome editing using FnCpf1 and LbCpf1 nucleases at redefined and altered PAM sites. *Mol. Plant* 11, 999–1002. doi: 10.1016/j.molp.2018.03.008

Zhu, F., Qin, C., Tao, L., Liu, X., Shi, Z., Ma, X., et al. (2011). Clustered patterns of species origins of nature-derived drugs and clues for future bioprospecting. *Proc. Natl. Acad. Sci. U.S.A.* 108, 12943–12948. doi: 10.1073/pnas.1107336108

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

REVIEW
published: 30 June 2020
doi: 10.3389/fbioe.2020.00711

Development and Application of CRISPR/Cas in Microbial Biotechnology

*Wentao Ding[1,2], Yang Zhang[1] and Shuobo Shi[1]**

[1] Beijing Advanced Innovation Center for Soft Matter Science and Engineering, Beijing University of Chemical Technology, Beijing, China, [2] Key Laboratory of Food Nutrition and Safety, Ministry of Education, College of Food Engineering and Biotechnology, Tianjin University of Science and Technology, Tianjin, China

The clustered regularly interspaced short palindromic repeats (CRISPR)-associated (Cas) system has been rapidly developed as versatile genomic engineering tools with high efficiency, accuracy and flexibility, and has revolutionized traditional methods for applications in microbial biotechnology. Here, key points of building reliable CRISPR/Cas system for genome engineering are discussed, including the Cas protein, the guide RNA and the donor DNA. Following an overview of various CRISPR/Cas tools for genome engineering, including gene activation, gene interference, orthogonal CRISPR systems and precise single base editing, we highlighted the application of CRISPR/Cas toolbox for multiplexed engineering and high throughput screening. We then summarize recent applications of CRISPR/Cas systems in metabolic engineering toward production of chemicals and natural compounds, and end with perspectives of future advancements.

Keywords: CRISPR/Cas, guide RNA, genome editing, gene regulation, microbial biotechnology

OPEN ACCESS

Edited by:
Yi Wang,
Auburn University, United States

Reviewed by:
Yuan Qiao,
Nanyang Technological University, Singapore
Mingfeng Cao,
University of Illinois at Urbana-Champaign, United States

****Correspondence:***
Shuobo Shi
shishuobo@mail.buct.edu.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *16 April 2020*
Accepted: *08 June 2020*
Published: *30 June 2020*

Citation:
Ding W, Zhang Y and Shi S (2020) Development and Application of CRISPR/Cas in Microbial Biotechnology. Front. Bioeng. Biotechnol. 8:711. doi: 10.3389/fbioe.2020.00711

INTRODUCTION

Microbial cell factories producing fuels, chemicals, and pharmaceutics are perspective production mode to replace petrol relied methods because microbial methods are usually clean and renewable. One restriction to the development of microbial producer is the slow, inefficient and arduous genomic engineering processes. The emerging toolbox based on clustered regularly interspaced short palindromic repeats (CRISPR) system have largely improved genome editing efficiency, simplified steps of multi-loci editing, and enabled fast disturbance of metabolic network. The CRISPR system is prokaryotic adaptive immune system against intruded heterologous DNA/RNA from virus or other organisms (Grissa et al., 2007a; Sorek et al., 2013). So far, the CRISPR/Cas system has been intensively adopted as toolbox for both fundamental studies and biotechnological applications for genome editing, molecular diagnosis, metabolic engineering, gene function mining, etc., in microorganisms, plants and mammals (Sander and Joung, 2014; Zhang et al., 2014; Wang H. et al., 2016; Tang and Fu, 2018; Tarasava et al., 2018; Armario Najera et al., 2019; Moon et al., 2019; Xu and Oi, 2019). In the field of microbial biotechnology, the CRISPR/Cas system has been applied for numerous model and non-model microorganisms, e.g., *Escherichia coli* (Jiang et al., 2013), *Saccharomyces cerevisiae* (DiCarlo et al., 2013), *Bacillus* (Westbrook et al., 2016), *Clostridium* (Li et al., 2016; Joseph et al., 2018), *Corynebacterium* (Jiang et al., 2017), *Lactobacillus* (Oh and van Pijkeren, 2014), *Mycobacterium* (Choudhary et al., 2015), *Pseudomonas* (Tan S. Z. et al., 2018), *Streptomyces* (Cobb et al., 2015). However, there still remains interested microorganisms that CRISPR system has not been applied, and some weakness of existing CRISPR/Cas systems needs

to be overcome. This review focuses on the establishment and development of CRISPR toolbox for genome editing and gene regulation, and applications of these techniques in metabolic engineering and synthetic biology in microorganisms.

THE CRISPR/CAS SYSTEM FOR GENOME EDITING

The CRISPR systems are adaptive evolved for counteracting foreign DNA or RNAs, and the systems are present in nearly half of bacteria and almost all archaea (Grissa et al., 2007b; Zetsche et al., 2015a), but absent from eukaryotes or viruses (Jansen et al., 2002). The CRISPR/Cas systems have been categorized into two classes and six major types based on the constitution of effector protein and signature genes, protein sequence conservation, and organization of the respective genomic loci (Koonin et al., 2017; Tang and Fu, 2018). Among these CRISPR systems, the Cas9 (Type II), Cas12a (previously known as Cpf1, type V) and their mutant variants are most investigated effectors, and have shown broad applicational potentials in genome editing, gene regulation, DNA detection, DNA imaging, etc. (Tang and Fu, 2018; Miao et al., 2019).

The CRISPR/Cas system can introduce a double-strand DNA break (DSB) at the specific DNA target (also called protospacer) binding by a guide RNA (gRNA) and harboring a short protospacer adjacent motif (PAM) flanked at the 3′ end of protospacer (**Figures 1A,B**; Garneau et al., 2010; Gasiunas et al., 2012; Jinek et al., 2012; Wang H. et al., 2016). A DSB triggers DNA repair through intrinsic cellular mechanisms, mainly including non-homologous end joining (NHEJ), which direct ligates two breaking ends with small insertions or deletions (indels); and homology-directed repair (HDR), which repair DSB according to a homologous template (Hsu et al., 2014; Doetschman and Georgieva, 2017). Considering the guide RNAs are easy to design and expressed, Cas protein can be programmed to introduce DSBs at one or more DNA targets, making CRISPR/Cas an convenient and precise platform for genome editing (Doetschman and Georgieva, 2017). Compared with similar genome editing tools such as zinc-finger nucleases (ZFNs) (Kim et al., 1996; Urnov et al., 2010) and TAL effector nucleases (TALENs) (Boch et al., 2009; Christian et al., 2010), CRISPR/Cas shows a significant advantage that it is easier to target a specific region by adjusting a 20 nt spacer sequence of gRNA, rather than producing target-specific proteins (Doetschman and Georgieva, 2017).

Selection and Expression of Cas Protein

The CRISPR/Cas systems have been reported to have two classes and six major types, and among these types, the class 2 type II CRISPR system (CRISPR/Cas9) is currently most studied and developed as toolbox for gene editing and other applications. As shown in **Figure 1A**, the effector (Cas9) is activated when forming a complex with single guide RNA [sgRNA, a fusion RNA of CRISPR targeting RNA (crRNA) and trans-activating CRISPR RNA (tracrRNA) (Jinek et al., 2012)], and triggers DSB at DNA target near PAM (Mougiakos et al., 2016). The spacer part is responsible for DNA target (also called protospacer) binding, and guides the Cas9 complex for sequence specific DNA cleavage. PAM flanks the 3′ end of the protospacer, and is required for Cas9-mediated cleavage (Deveau et al., 2008; Mojica et al., 2009). The PAM of the most commonly used SpCas9 (Cas9 from *Streptococcus pyogenes*) is 'NGG,' which occurs once every 8 bp on average within the genome, allowing targeting on most genes of interest (Doudna and Charpentier, 2014; Hsu et al., 2014). Cas9s from different resources recognize different PAM sequences, which further expands the application of CRISPR for various genomic sequence [e.g., Cas9 from *Staphylococcus aureus* (Kleinstiver et al., 2015; Ran et al., 2015), *Streptococcus thermophiles* (Esvelt et al., 2013; Kleinstiver et al., 2015), *Neisseria meningitides* (Esvelt et al., 2013; Hou et al., 2013)]. Cas9 'nickase' variant (nCas9), with mutations deactivating one nickase activity and converting the endonuclease activity of wildtype Cas9 to nickase activity, introduces a single stranded break (SSB) rather than DSB (Jinek et al., 2012; Cong et al., 2013). Generally, SSBs are repaired by HDR, not by NHEJ, thus nCas9 can be applied for precise genome editing (Standage-Beier et al., 2015). Another Cas9 mutant, the nuclease-deactivated Cas9 (dCas9), has been fused with a variety of effectors, including transcriptional activators, repressors, and epigenetic modifiers to enable sequence specific genomic regulation (Gilbert et al., 2013, 2014; Qi et al., 2013).

In 2013, the application of CRISPR/Cas9 system for genome editing was originally reported in human cells (Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013b), mouse cells (Cong et al., 2013), Zebrafish (Hwang et al., 2013), *Saccharomyces cerevisiae* (DiCarlo et al., 2013), *Streptococcus pneumoniae,* and *Escherichia coli* (Jiang et al., 2013). In following studies, the CRISPR/Cas9 system has been widely applied for genome editing in numerous microorganisms, plants and animals.

As an eukaryotic model microorganism, *S. cerevisiae* was one of the earliest hosts for CRISPR/Cas9 mediated genome editing (DiCarlo et al., 2013). In order to improve genome editing efficiency, the Cas9 protein is usually highly expressed by a strong constitutive promoter [e.g., *TEF1* promoter (DiCarlo et al., 2013; Gilbert et al., 2013; Bao et al., 2015), *TDH3* promoter (Gilbert et al., 2013; Laughery et al., 2015; Jensen et al., 2017)] in a episomal CEN low copy plasmid (DiCarlo et al., 2013; Gilbert et al., 2013) or episomal 2 μ high copy plasmid (Ryan and Cate, 2014; Bao et al., 2015; Shi et al., 2016; Jensen et al., 2017). However, in some researches, expression of Cas9 with strong promoter (e.g., promoter of *TEF1*, *HXT7*, and *TDH3*) showed toxic effect to cell growth (Ryan and Cate, 2014; Generoso et al., 2016). Nevertheless, medium strength or weak promoters showed similar editing efficiency, and no significant negative impact on the strain's growth rate. For efficient CRISPR editing rate, codon usage in heterologous organisms should be also considered to guarantee sufficient Cas9 abundance *in vivo*. In eukaryotes, Cas9 protein should be transported to nuclei to facilitate genome editing, and thus the nuclear localization sequence (NLS) should be fused to the Cas9 protein (**Figure 1B**). In *S. cerevisiae*, the SV40 NLS ('PKKKRKV') is typically fused to the N- or C-terminus of

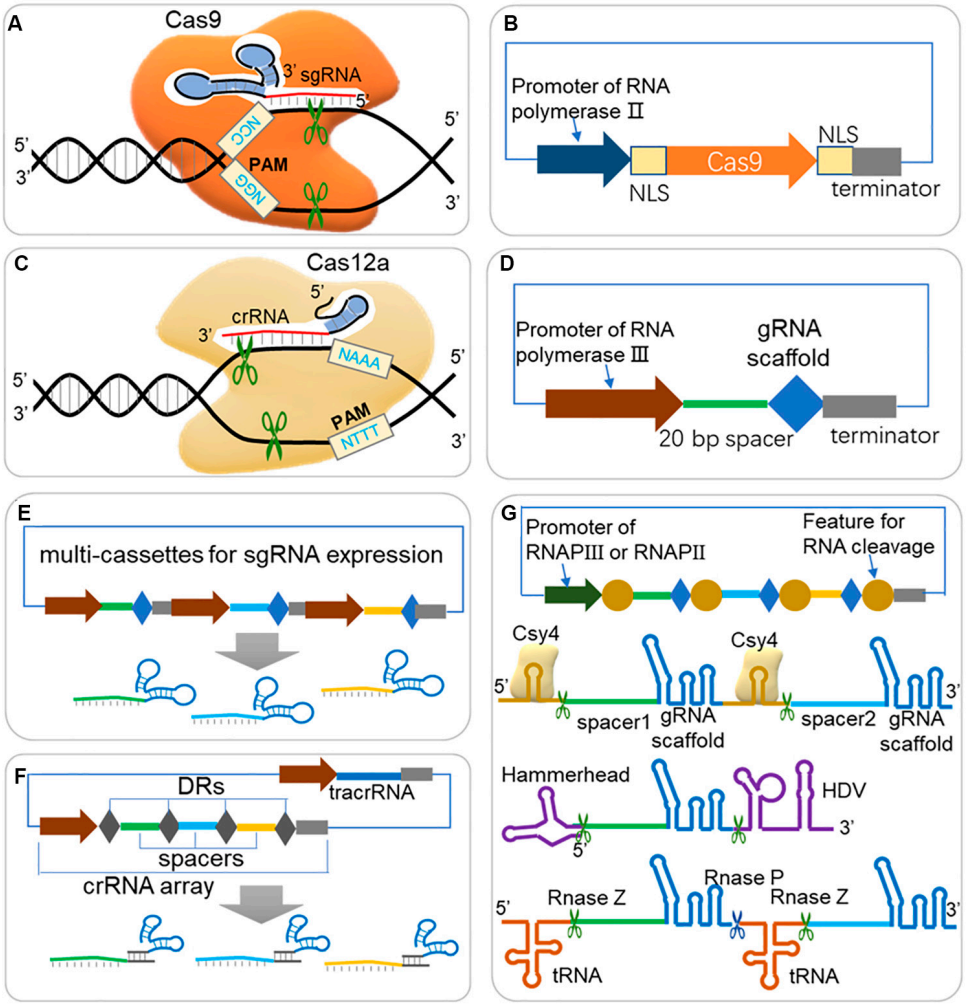


FIGURE 1 | Guidelines for expression of Cas protein and sgRNA in CRISPR/Cas system. **(A)** Scheme of CRISPR/Cas9 system. The Cas9-sgRNA (or Cas9-crRNA-tracrRNA) complex binds to DNA target arising from Watson-Crick base pairing of spacer sequence, and triggers double strand break (DSB) when next to a short protospacer adjacent motif (PAM, 'NGG' for Cas9 from *S. pyogenes*). **(B)** Expression cassette for Cas9. For efficient targeting to nucleus in eukaryotes, the Cas9 should be fused to NLS (nuclear localization sequence) at one end or both ends. **(C)** Scheme of CRISPR/Cas12a (Cpf1) system. Cas12a triggers DSB through a similar scheme of Cas9, but depends on different PAM ('NTTT') and less folded crRNA, and creates a sticky end at 18–23 bases away from the PAM. **(D)** Expression cassette for sgRNA. A promoter of RNA polymerase III (RNAP III) is usually required for directing sgRNA in nucleus and with less modification. A 20 bp spacer should be well designed according to target DNA sequence for efficient editing rates and avoiding off-target effects. **(E)** Multi-sgRNA expression through multi-cassettes. Repeated elements, such as promoters, gRNA scaffold and terminators are repeated for different spacer sequences. **(F)** Multi-sgRNA expression through crRNA array and tracrRNA (HI-CRISPR system). Different spacers are separated with direct repeats (DRs) and expressed by one promoter of RNAP III. The pre-crRNA is transcribed and processed into mature crRNA by RNase III and unknown nuclease(s). The tracrRNA and Cas9 protein are complexed with mature crRNA to form the dual-RNA-guided nuclease. **(G)** gRNA multiplexing strategies. Both RNAP II and RNAP III promoter can be used for expression the sgRNA array, where sgRNAs are separated by features for RNA cleavage. RNA endonuclease Csy4 recognizes a 28 nucleotide sequence flanking the sgRNA sequence and cleaves after the 20th nucleotide. The hammerhead ribozyme and HDV ribozyme flanked the 5′ and 3′ of the sgRNA, respectively, allowing for self-cleaving production of sgRNAs, which are not dependent on the presence of an exogenous protein. Polycistronic tRNA-gRNA architecture allows the production of multiple sgRNAs by endogenous RNase P and RNase Z.

the Cas9, and two NLSs fused to one terminus or both were also applicable.

The model bacteria *E. coli* has also been intensively researched as a host for CRISPR/Cas9 mediated genome editing. However, *E. coli* lacks the NHEJ mechanism for DSB repair (Chayot et al., 2010), and is highly reliant on a native homology-directed repair system with low efficiency, challenging the DSB producing CRISPR/Cas9 system (Jiang et al., 2015). Thus, co-expression of heterologous phage-derived recombinase to improve the frequency of homologous recombination showed significant improved survival rates when CRISPR/Cas9 and gRNA expressed (Jiang et al., 2015; Pyne et al., 2015; Bassalo et al., 2016). In *E. coli*, inducible promoters were mostly used for both Cas9 and gRNA expression.

The CRISPR/Cas9 system has also been constructed with similar strategy for non-model microorganisms (Yan and Fong, 2017; Cho et al., 2018; Raschmanova et al., 2018; Wang and Coleman, 2019). Generally, species-specific strong promoters should be used for Cas9 expression, either constitutively or inducible expressed. Codon optimization should also be conducted when the Cas9 protein cannot be efficiently expressed. In eukaryotic microorganisms, NLS should be fused to Cas9 at one or both termini for cell nucleus localization. The NLS of SV40 from *S. cerevisiae* has been proven effective and applied in other yeast species. Native DNA repair types and efficiency also largely determined genome editing rate, because DSB induced by CRISPR/Cas9 can be repaired by NHEJ, resulting in indels and gene inactivation, or be repaired by HDR, resulting in precise genome editing by supplying proper DNA donors. Thus, in some organisms with both NHEJ and HDR pathways, deletion of *KU70*/*KU80* often repressed NHEJ and increased CRISPR mediate genome editing rate through HDR (Gao S. et al., 2016; Schwartz et al., 2016; Cao et al., 2018; Bae et al., 2020). However, in some organisms lacking HDR, phage-derived recombinases (RecET and λ-Red) should be co-expressed with Cas9, similar to the approaches adopted *in E. coli* (Jiang et al., 2015; Wang B. et al., 2018).

In addition to widely applied Cas9, Cas12a (also known as Cpf1) is a newly emerging Cas protein that is currently under evaluation for gene editing potential (Zetsche et al., 2015a). Cas12a is a crRNA-guided endonuclease, lacking tracrRNA compared with Cas9, and cleaves DNA at 18 nucleotides away from the PAM, resulting in a DSB with 4- to 5-nucleotide overhangs (**Figure 1C**; Zetsche et al., 2015a). Besides, Shmakov et al. (2015) further classified three class 2 CRISPR systems, including C2c1, C2c3, and C2c2, which further expands CRISPR toolbox for genome editing.

Design and Expression of Guide RNA

The efficient expression of guide RNA is also critical to a CRISPR system because the spacer sequence of guide RNA is responsible for DNA target binding and thus decides the editing loci, and is closely related to on-target and off-target efficiency. Generally, one or more single guide RNAs (sgRNAs) are expressed in a CRISPR/Cas system (**Figures 1D–G**); but in some other cases, a crRNA matrix and a tracrRNA, instead of sgRNAs, are expressed separately for efficient CRISPR editing (Bao et al., 2015). The spacer sequence should be carefully designed, which binds to a DNA target close to a PAM sequence, and to promote editing efficiency and reduce off-target rate. A serial of studies have suggested that mismatches at the 5′ end of spacer sequence are generally better tolerated than those at the 3′ end, and especially the 8–12 bps at the 3′ end of the spacer sequence are crucial for target recognition (Cong et al., 2013; Fu et al., 2013; Hsu et al., 2013; Jiang et al., 2013; Sander and Joung, 2014). It is crucial to design gRNAs for CRISPR system, and a well-selected gRNA would minimize the risk of CRISPR-mediated DSBs at unwanted sites in genome (off-target effects) and maximize the editing efficiency at the selected site (on-target activity) (Stovicek et al., 2017). Several rules and algorithms have been proposed, and web-tools for gRNA design can help to choose best gRNAs in various species (shown in **Table 1**). The rules for gRNA scoring includes possible binding sites with mismatches in the spacer sequence or in the seed sequence, the GC content and poly T presence and self-complementarity (Heigwer et al., 2014; Liu et al., 2015; Naito et al., 2015; Labun et al., 2019). Except for gene editing, CRISPR-ERA and CHOPCHOP also help to design gRNAs for gene activation and repression (Liu et al., 2015; Labun et al., 2019).

Generally, a strong expression of gRNA is recommended for an efficient target binding and CRISPR complex activation. To express RNA without modifications added by the RNA polymerase II (RNAPII) transcription system, RNA polymerase III (RNAPIII) regulatory elements have been used for transcription of functional gRNA (**Figure 1D**), e.g., the *SNR52* promoter has been used in yeast (Raschmanova et al., 2018) and *U6* promoter has been used in human cells (Zhang et al., 2014; Wang H. et al., 2016). However, it is noted that some promoters require special rules of gRNA sequence, e.g., the U6 promoter or the T7 promoter require a 'G' or 'GG,' respectively, at the 5′ end of the RNA to be transcribed (Sander and Joung, 2014; Wang H. et al., 2016). Despite RNAPIII promoters are suitable for gRNA transcription, in some organisms, however, these promoters are poorly characterized. On the other hand, RNAPII promoters can also be used to express gRNAs when proper strategies are adopted (Nowak et al., 2016). A RNAPII promoter of *rrk1* and its leader RNA was used to express sgRNA by flanking a Hammerhead ribozyme on the 3′ end of gRNA (**Figure 1G**) in fission yeast (Jacobs et al., 2014). Another research also used RNAPII promoter but flanked the sgRNA with a 28 nucleotide hairpin at each end that is recognized by the endoribonuclease Csy4 (**Figure 1G**; Nissim et al., 2014). Fusion gRNAs with a hammerhead (HH) ribozyme on their 5′ end and a hepatitis delta virus (HDV) ribozyme on their 3′ end was also reported functional for RNAPII promoter (**Figure 1G**; Nissim et al., 2014; Weninger et al., 2016). Interestingly, fusion of sgRNA with special RNA scaffold (e.g., HDV, RNA triplex) would increase *in vivo* RNA stability and thus promote engineering efficiency (Nissim et al., 2014; Ryan and Cate, 2014).

When CRISPR/Cas system is constructed for multi-loci editing (**Figures 1E–G**), several strategies have been proposed to enable an efficient expression of multiple gRNAs. Multi-sgRNA expression could be achieved through multi-expression cassettes using individual promoters to control each gRNA (**Figure 1E**). This method was successfully demonstrated to enable multiple editing (Jakociunas et al., 2015). For another strategy, the crRNA matrix and tracrRNA were expressed separately by RNAPIII promoters, and processed into mature crRNA by RNase III and unknown nuclease(s) (**Figure 1F**), which also showed high gene disruption efficiency in *S. cerevisiae* (Bao et al., 2015). The tRNA-processing system, which precisely cleaves both ends of the tRNA precursor by RNase P and RNase Z (or RNase E in bacterium, **Figure 1G**), exists in virtually all organisms and can be broadly used to boost the targeting capability and editing efficiency of CRISPR/Cas systems (Xie et al., 2015; Port and Bullock, 2016;

TABLE 1 | List of selected Web-sites for gRNA design in multi-species.

Name	Link	PAM	Organism	Function	References
CHOPCHOP v3	http://chopchop.cbu.uib.no	Most reported PAMs or a self-defined sequence	Over 200 genomes	Knock out/knock in/activation/repression/ Nanopore enrichment	Labun et al., 2019
E-CRISPR	http://www.e-crisp.org/	Most reported PAMs	Over 50 genomes	Single design/paired designs	Heigwer et al., 2014
ATUM	https://www.atum.bio/eCommerce/cas9/input	NGG/NAG	*Homo sapiens/Mus musculus/Saccharomyces cerevisiae/Escherichia coli/Arabidopsis thaliana*		
CRISPRdirect	https://crispr.dbcls.jp/	Self-defined PAM	Over 200 species		Naito et al., 2015
CRISPR-ERA	http://crisprera.stanford.edu/	NGG	Human/mouse/rat/ zebrafish/ *D. melanogaster/ C. elegans/S. cerevisiae/ E. coli/B. subtilis*	Gene editing/ activation/repression	Liu et al., 2015
CC TOP	https://crispr.cos.uni-heidelberg.de	Most reported PAMs	102 species	gRNA and off-target prediction	Stemmer et al., 2015

Qi et al., 2016; Ding et al., 2018; Zhang et al., 2019). Single or multiple gRNAs can be expressed by one promoter but separated by tRNA scaffolds [e.g., a 71 bp long pre-tRNA^{Gly} (Xie et al., 2015; Zhang et al., 2019)].

It is costly and time consuming for the sub-cloning of plasmids used for multi-loci editing, and some strategies could be taken for saving cloning time or improving editing efficiency. Gibson assembly, Golden gate cloning and USER cloning have showed high rates in multi DNA fragments assembly, which simplifies cloning steps for multiple gRNA expression cassettes, and thus saves the processing time (Bao et al., 2015; Shi et al., 2016; Smith et al., 2016; Jensen et al., 2017; Zhang et al., 2019). Meanwhile, *in vivo* homologous recombination has been reported for rapid assembling a certain plasmid backbone and PCR cassettes bearing sgRNAs in some yeast species (*S. cerevisiae* and *K. lactis*), thus saving cloning steps for high-efficiency engineering (Horwitz et al., 2015; Generoso et al., 2016; Reider Apel et al., 2017).

DNA Repair and Donor Design for DNA Deletion, Insert and Mutation

The CRISPR/Cas mediated precise genome editing relies on intrinsic DNA repair mechanisms after a DSB or SSB was introduced to genome by a Cas protein, e.g., Cas9 nuclease or a Cas9 mutant (Cas9 nickase, nCas9) (**Figures 2A,B**). There are two main pathways for DSB repair in nearly all organisms: non-homologous end-joining (NHEJ), direct ligation of two break ends with little or no sequence homology required; and homology-directed repair (HDR), repairing DSB according to a DNA template with homology sequence (**Figure 2A**; Ceccaldi et al., 2016; Ranjha et al., 2018). Despite alternative end joining [alt-EJ, also termed microhomology-mediated end-joining (MMEJ)] and single-strand annealing (SSA) may also repair DSBs in some organisms, NHEJ and HDR remain dominant pathways in most organisms (Ceccaldi et al., 2016; Ranjha et al., 2018). NHEJ is a fast, template independent and mutagenic pathway for DSB repair that occurs in whole cell cycle (Chang et al., 2017); whereas HDR is a slow, accurate, template dependent pathway for both DSB and SSB repair, but only occurs in S/G2 phase (Ranjha et al., 2018). NHEJ introduces unpredictable patterns of insertions and deletions, but if multiple DSBs are present, large deletions or chromosomal rearrangements may occur (Chang et al., 2017; Ranjha et al., 2018). On the other hand, CRISPR/Cas mediated precise genome editing relies on DSB or SSB repairing through HDR pathway and DNA template (donor DNA).

Both single strand DNA (ssDNA) and double strand DNA (dsDNA) fragments can be used as donors for genome editing. Despite ssDNA donors showed higher editing efficiency than dsDNA donors in several researches (Ran et al., 2013b; Miura et al., 2015; Singh et al., 2015), dsDNA donors (linear or circular) showed comparable efficiency but higher flexibility and have been widely adopted for gene deleting, mutation and insertion (DiCarlo et al., 2013; Zerbini et al., 2017; Zhang et al., 2019).

A DNA donor could be provided as HDR template to destroy the target open reading frame, and change or eliminate gRNA binding sequence and PAM to avoid repeated cleavage by Cas protein (**Figures 2A,C**; Raschmanova et al., 2018; Zhang et al., 2019). A short dsDNA donor, with ∼50 bp homologous sequence at each end, is usually viable and can be prepared by PCR of two oligonucleotide primers (Jakociunas et al., 2015; Zhang et al., 2019). Such short donors can also be used for introduction of single-nucleotide mutations within different gene loci (Wang Y. et al., 2016), if the locus to be edited is within the "GG" loci of a PAM or the 20 nt protospacer. Long dsDNA donors can be used for insertion (**Figure 2D**). Expression cassettes or other inserts can be carried by long donors and inserted to genome through HDR pathway (**Figure 2D**). These donors should have long homology arms (0.1–3 kb) for efficient HDR (Doetschman and Georgieva, 2017), and up to 24 kb fragments have been integrated to yeast genome through CRISPR/Cas9 system (Shi et al., 2016, 2019). Recently, transposons were proposed as an alternative

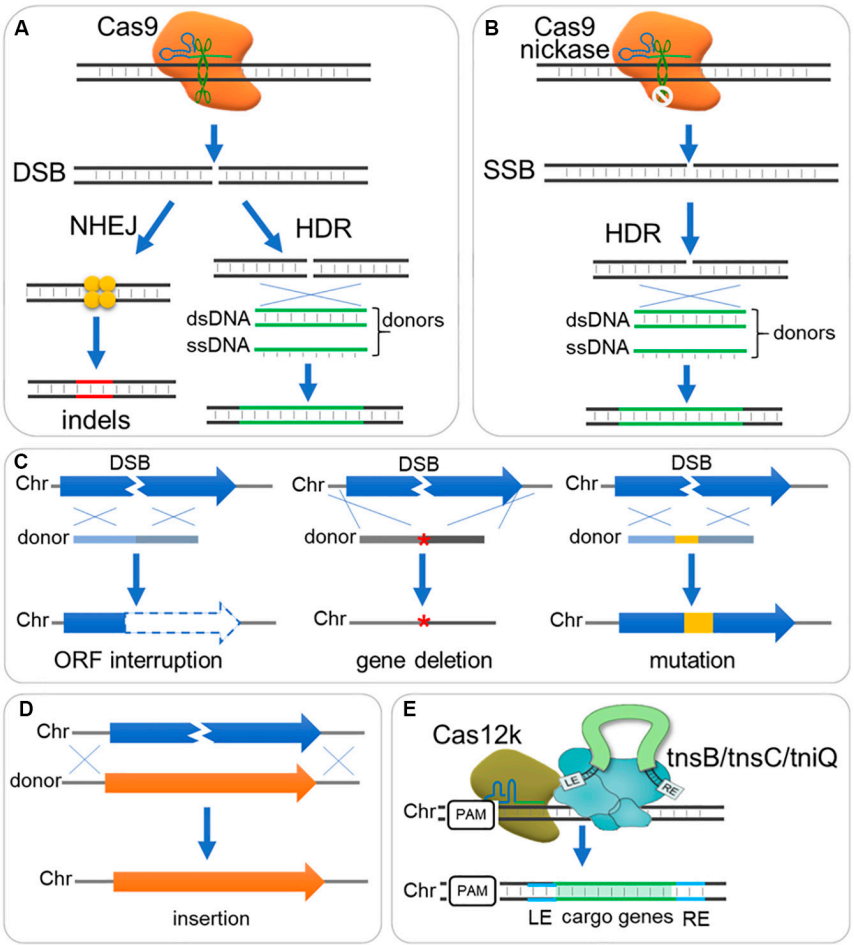


FIGURE 2 | DNA repair and donor design for DNA deletion, insert and mutation. **(A)** The Cas9-sgRNA complex binds to DNA target and triggers a double strand break (DSB), which is subsequently repaired generally through non-homologous end joining (NHEJ) or homology-directed (HDR) pathway. NHEJ directs ligation of two break ends with little or no sequence homology required, resulting in small insertions or deletions (indels); while HDR repairs DSB according to a DNA template with homology sequence, resulting in precise editing when supplemented with ds- or ss-DNA donors. **(B)** A Cas9 nickase mutant with HNH or RuvC inactive domain introduces a single strand break (SSB), which can be repaired by HDR rather than NHEJ pathway. **(C)** Donor designs for gene interruption, deletion and mutation. Gene interruption: Small deletion (e.g., 8 bp) or insertion is integrated to shift reading frame, or stop codon is introduced to interrupt gene translation. Gene deletion: A donor fused with sequence upstream and downstream ORF is sued for gene deletion ('*' indicate the deleted gene). Gene mutation: Sequence mutations can be introduced by a donor, where seed sequence and PAM should be destroyed to avoid cutting again by Cas9-sgRNA complex. Chr, chromosome. **(D)** Donor design for sequence insertion. A donor contain long sequence is integrated through HDR, and longer homology arms are required when inserting long sequence. **(E)** Another strategy employing CRISPR I-F or V-K (e.g., Cas12k) mediates DNA integration with Tn7-like transposons (e.g., *tnsB/tnsC/tniQ*).

tool to mediate DNA integration via a HDR independent way (Klompe et al., 2019; Strecker et al., 2019), which depends on type I–F or V-K CRISPR effectors (e.g., Cas12k) and interacts with Tn7-like transposons (e.g., *tnsB*/*tnsC*/*tniQ*) (**Figure 2E**).

Adaption of CRISPR/Cas System to Non-model Microorganisms

As a powerful toolbox for genome editing and regulation, CRISPR systems are highly valued not only for model microorganisms (e.g., *E. coli*, *S. cerevisiae*), but also provide more applicable perspectives for non-model microorganisms that are difficult to be processed through traditional methods. Despite CRISPR/Cas systems have already been applied in plenty of microbial hosts (Freed et al., 2018; Raschmanova et al., 2018; Palazzotto et al., 2019; Wang and Coleman, 2019; Ng et al., 2020), it is still challenging to construct CRISPR system with high editing efficiency, and/or apply various CRISPR strategies in non-model microorganisms. In particular, lessons have also been learned that several limitations should be overcome to enable the multiplexed /genome-scale processing of CRISPR in non-model microorganisms, such as the delivery of gRNAs or Cas proteins, the genotoxic stress, etc.

One dominant challenge is active, reliable and sufficient expression of Cas protein and gRNAs in a non-model host. Due to the limited knowledge of non-conventional organisms, it is necessary to identify expression architectures ahead of CRISPR system construction. Constitutive or inducible RNAPII

promoters are used for expression of Cas proteins, but RNAP III promoters should be used for sgRNA expression. In some organisms without identified RNAPIII promoters, RNAPII promoters can also be used to express gRNAs when proper strategies adopted when fusing sgRNA with special elements at each end, e.g., Hammerhead ribozyme, HDV ribozyme, and Csy4 cutting site (Jacobs et al., 2014; Nissim et al., 2014; Nowak et al., 2016; Weninger et al., 2016). Some architectures for stable episomal expression could also largely improve CRISPR efficiency, such as centromeric sequence (Cao et al., 2017, 2020) and autonomously replicating sequences (ARSs) (Gu et al., 2019).

Usually, the Cas9 form *S. pyogenes* (SpCas9) is efficient enough for genome editing in different organisms. Codon optimization is occasionally needed when the wildtype SpCas9 was not actively expressed. In some organisms, however, SpCas9 showed low efficiency or toxic effect, and repressed cell growth significantly (Ungerer and Pakrasi, 2016; Wendt et al., 2016; Jiang et al., 2017). To solve this issue, different CRISPR systems or effector variants (e.g., Cas12a) showed high editing efficiency but lower toxicity, and were applied in those organisms (Ungerer and Pakrasi, 2016; Jiang et al., 2017; Yeo et al., 2019).

On the other hand, the CRISPR aided precise, time-saving and markerless genome editing relays on introducing DSBs at DNA targets and repairing process thereafter. Thus the intrinsic DNA repairing system largely determinates editing efficiency in non-model microorganisms. DSB repairing through NHEJ pathway results in small random deletions or inserts at the site of DSB, rather than precise repairing according to a template through HDR pathway. Thus, in those NHEJ dominant species, CRISPR/Cas system can be used for just gene inactivation, but very low efficiency in precise DNA insertion, unless NHEJ is blocked, e.g., by knocking out KU70 and/or KU80 as mentioned before (Gao S. et al., 2016; Schwartz et al., 2016; Cao et al., 2018; Bae et al., 2020). In some species lacking HDR pathway, phage-derived recombinases (RecET and λ-Red) should be expressed to assist genome editing (Jiang et al., 2015; Wang B. et al., 2018). In addition, some chemical reagents can be supplemented to increase HDR efficiency, such as SCR7 (Maruyama et al., 2015), RS-1(Song et al., 2016), KU0060648, and NU7441 (Robert et al., 2015). The HDR pathway is the dominant mechanism for DSB repair in most bacteria, and NHEJ is present in some bacteria including *Mycobacterium*, *Pseudomonas*, and *Bacillus (*Weller et al., 2002; Shuman and Glickman, 2007). In most eukaryote, however, NHEJ is the dominant mechanism for DNA repairing. It is recently reported that expression of T4 DNA ligase provides efficient *in vivo* NHEJ repairing pathway in bacteria (Su et al., 2019). Donors also vary between organisms. In some cases, short (∼50 bp) homologous arms (HAs) are sufficient for HDR (Jakociunas et al., 2015; Zhang et al., 2019); while in other cases, long (∼1–3 kb) HAs are preferred (Doetschman and Georgieva, 2017).

Efforts to Reduce Off-Target Effects

Despite Cas9 cleavages DNA target depending on a 20 nt spacer sequence of gRNA and PAM, it still potentially introduces an undesired DSB at an unintended chromosomal locus (off-target), possibly because of gRNA binding to a similar sequence elsewhere on chromosome (Fu et al., 2013; Hsu et al., 2013; O'Geen et al., 2015). The off-target effect may lead to unexpected DNA mutations, which limits the application of CRISPR in various organisms. Efforts to address this issue have been made to increase CRISPR specificity and to predict possible off-target loci on genome. A well designed gRNA would largely reduce the crisis of off-target (Wang and Coleman, 2019), and the "seed" sequence of gRNA (10–12 bp adjacent to the PAM) highly decides the Cas9 cleavage specificity (Jinek et al., 2012). To reduce the off-target risk and protect binding and cleavage activity, bioinformatic tools or websites have been developed for gRNA design, such as Cas-OFFinder[1] (Bae et al., 2014) and CCTop[2] (Stemmer et al., 2015). Using truncated sgRNAs (17-18 bp) showed reduced off-target effect with Cas9 nuclease and paired Cas9 nickases in human cells (Fu et al., 2014). sgRNAs with two unpaired Gs on the 5′ end also showed more sensitive to mismatches in human cells (Kim et al., 2015). Engineering of the Cas9 protein for fidelity or specificity improvement also largely reduces off-target effects: e.g., Kleinstiver et al. (2016) reported a high-fidelity variant, SpCas9-HF1 (N497A/R661A/Q695A/Q926A); Slaymaker et al. (2016) engineered several SpCas9 variants with high efficiency and specificity, e.g., eSpCas9(1.0) (K810A/K1003A/R1060A), and eSpCas9(1.1) (K848A/K1003A/R1060A); Chen J. S. et al. (2017) reported a new hyper-accurate Cas9 variant, HypaCas9 (N692A/M694A/Q695A/H698A), which demonstrated high genome-wide specificity without compromising on-target activity; Hu et al. (2018) reported an expanded PAM SpCas9 variant, xCas9 (xCas9-3.7: A262T, R324L, S409I, E480K, E543D, M694I, and E1219V), which showed much improved specificity and more broad PAM sequence, e.g., 'NG,' 'GAA,' and 'GAT.' A Cas9 nickase mutant (nCas9) system can also reduce off-target effect, in which a pair of guide RNAs is designed to bind to a narrow target region and thus nCas9 complexes introduce two SSBs on both strand of DNA, forming a DSB with sticky ends (Mali et al., 2013a; Ran et al., 2013a; Shen et al., 2014). Similarly, Guilinger et al. fused catalytically inactive Cas9 (dCas9) and FokI nuclease (fCas9), which produces DSB by simultaneous binding of two fCas9 monomers to the DNA target sites ∼15 or 25 base pairs apart, and resulted in at least 4-fold higher specificity than that of paired nickases (Guilinger et al., 2014; Tsai et al., 2014; Wyvekens et al., 2015).

Till now, the CRISPR/Cas system has already become the most commonly used gene editing tool for numerous species. It has become a precise, convenient and portable platform for genome editing and beyond.

REGULATION OF GENE EXPRESSION BY CRISPR/CAS TOOLBOX

In addition to site-specific gene editing, the catalytically dead Cas protein (e.g., dCas9, with H840A and D10A mutation) that retained its capability to recognize and bind a target DNA

[1]http://www.rgenome.net/cas-offinder/

[2]https://crispr.cos.uni-heidelberg.de

sequence (Qi et al., 2013) has been developed as a multi-functional platform based on its DNA recognizing and binding properties. The CRISPR/dCas9 system has been intensively researched and applied for transcription regulation, complex metabolic engineering, directed revolution, gene target screening and activation of silent gene clusters (Lino et al., 2018; Tarasava et al., 2018; Xu and Oi, 2019). Especially, the CRISPR interference (CRISPRi) (Qi et al., 2013) and the CRISPR activation (CRISPRa) (Tanenbaum et al., 2014) that allow programmed controlling of gene expression without altering the genome, are effective tools for metabolic engineering, and are highlighted here.

Repression of Gene Expression by dCas9 (CRISPRi)

CRISPR interference (CRISPRi) represses expression of targeted genes in a simple and reversible way without altered DNA sequence or off-target effects (Qi et al., 2013). Especially for those organisms lacking the RNA interference pathway, CRISPRi system offers an easy and efficient approach for targeted gene knockdown (Li et al., 2016; Peters et al., 2016). The CRISPR/dCas9 system was first used for repressing transcription by sterically hindering the RNA polymerase recruiting (**Figure 3A**) or RNA polymerase processivity along the coding sequence (**Figure 3B**; Qi et al., 2013). Ni et al. (2019) developed a CRISPRi method in which multi-gRNA plasmid was constructed that could down-regulate 7 genes simultaneously in *S. cerevisiae*. However, this 'road blocker' strategy using dCas9 alone is not always efficient in some organisms (Qi et al., 2013). Gilbert et al. compared different repressive effector domains, including the KRAB (Krüppel associated box) domain, the WRPW domain and the CS (Chromo Shadow) domain, and found that dCas9-KRAB was the best repressor when targeting to a window of -50 to +300 bp relative to the transcription start site (TSS), or 0–100 bp region just downstream of the TSS (**Figure 3C**; Gilbert et al., 2013, 2014). Another dCas9 fusion domain, Mxi1, a mammalian transcriptional repressor domain that is reported to interact with the histone deacetylase Sin3 homolog in yeast, also showed effective repression in yeast (Gilbert et al., 2013; Jensen et al., 2017; Schwartz et al., 2017; Geller et al., 2019; Wensing et al., 2019). In another research, KRAB was fused to RNA-binding domains (COM-KRAB) and achieved similar repression effects when targeting DNA sites overlapped the TSS using a scaffold RNA (scRNA) (**Figure 3D**; Zalatan et al., 2015). Kearns et al. fused NmdCas9 with the histone demethylase LSD1, which suppressed the expression of genes controlled by the targeted enhancers (**Figure 3E**; Kearns et al., 2015).

Activation of Gene Expression by dCas9 (CRISPRa)

When dCas9 is fused with transcriptional activator and binds to the specific genomic locus, it can efficiently activate transcription via recruitment of RNA polymerase (RNAP). This CRISPR mediated transcriptional activation (CRISPRa) strategy has been applied in both prokaryotic and eukaryotic cells, and several transcriptional activators have been reported.

Bikard et al. (2013) reported a fusion protein between dCas9 and the omega subunit (ω) of RNA Polymerase (dCas9-ω) that can activate transcription by binding at an optimal distance from the promoter in *E. coli*. However, this activation effect varied depending on the binding position and the innate promoter strength, with highest activation observed for weak promoters (Bikard et al., 2013). In *S. cerevisiae*, one commonly used activator domain is VP64, consisting of four tandem copies of Herpes Simplex Viral Protein 16. dCas9-VP64 (**Figure 3F**) increased target gene expression by 2.5-fold, and when multiple operators were targeted, the expression reached up to 70-fold improvement (Farzadfard et al., 2013). Chavez et al. fused dCas9 with a tripartite activator VP64-p65-Rta (VPR, **Figure 3G**), which showed higher activating effect (∼10-fold) than dCas9-VP64 counterparts (Chavez et al., 2015). Zalatan et al. (2015) tested "scaffold RNAs" (scRNA) that encode both target locus and MS2, PP7, or com RNA hairpins, recruiting their cognate RNA-binding proteins fusing with VP64 for transcriptional activation. When the scRNA with two RNA hairpins connected by a double-stranded linker was used, stronger activation effects were observed (Zalatan et al., 2015). In a recent research, Dong et al. found that an activating effector, SoxS showed the highest effect among *E. coli* regulators (SoxS, MarA, Rob, and CAP), Hijackers (TetD, λcII, GP33, and $N4_{SSB}$) and RNAP subunits (αNTD, RpoZ, and RpoD) in a CRISPRa system with gRNA scaffold MS2-MCP interaction in *E. coli* (Dong et al., 2018). Especially, a SoxS mutant $SoxS^{R93A}$ and 5 aa linker further increased the activation activity (Dong et al., 2018). Konermann et al. (2015) reported a synergistic activation mediator (SAM) system for transcriptional activation (**Figure 3H**), which combined dCas9-VP64 with a modified scRNA system. The activator domain of p65 and the human heat shock factor 1 (HSF1) were fused with MS2 coat protein (MCP), and bound to MS2 hairpins on sgRNA for transcription activation (Konermann et al., 2015). Tanenbaum et al. developed a dCas9-SunTag system with strong activation of endogenous gene expression (**Figure 3I**), where the dCas9 was fused to a multimeric peptide (GCN4) array (SunTag), which can recruit multiple copies of scFv-VP64 for gene activation (Tanenbaum et al., 2014). Zhou et al. designed a new activation system, named as SunTag-p65-HSF1 (SPH), by combining the peptide array of SunTag and P65-HSF of SAM, which showed the highest level of activation compared to SAM, VPR, VP64 and SunTag in HEK293T and N2a cells (Zhou et al., 2018). Hilton et al. reported another strategy that fused dCas9 to the catalytic core of the human acetyltransferase p300 (**Figure 3J**). This fusion protein binds to upstream of a gene target, and catalyzes acetylation of histone H3 lysine 27 at its target sites, resulting in transcriptional activation (Hilton et al., 2015).

Orthogonal CRISPR Systems for Comprehensive Engineering

In metabolic engineering and synthetic biology, complex engineering, e.g., overexpression, dynamic regulation, knock-down, and knock-out of multiple gene targets, is often required. Unfortunately, such engineering processes are often carried out sequentially and with low throughput. The development

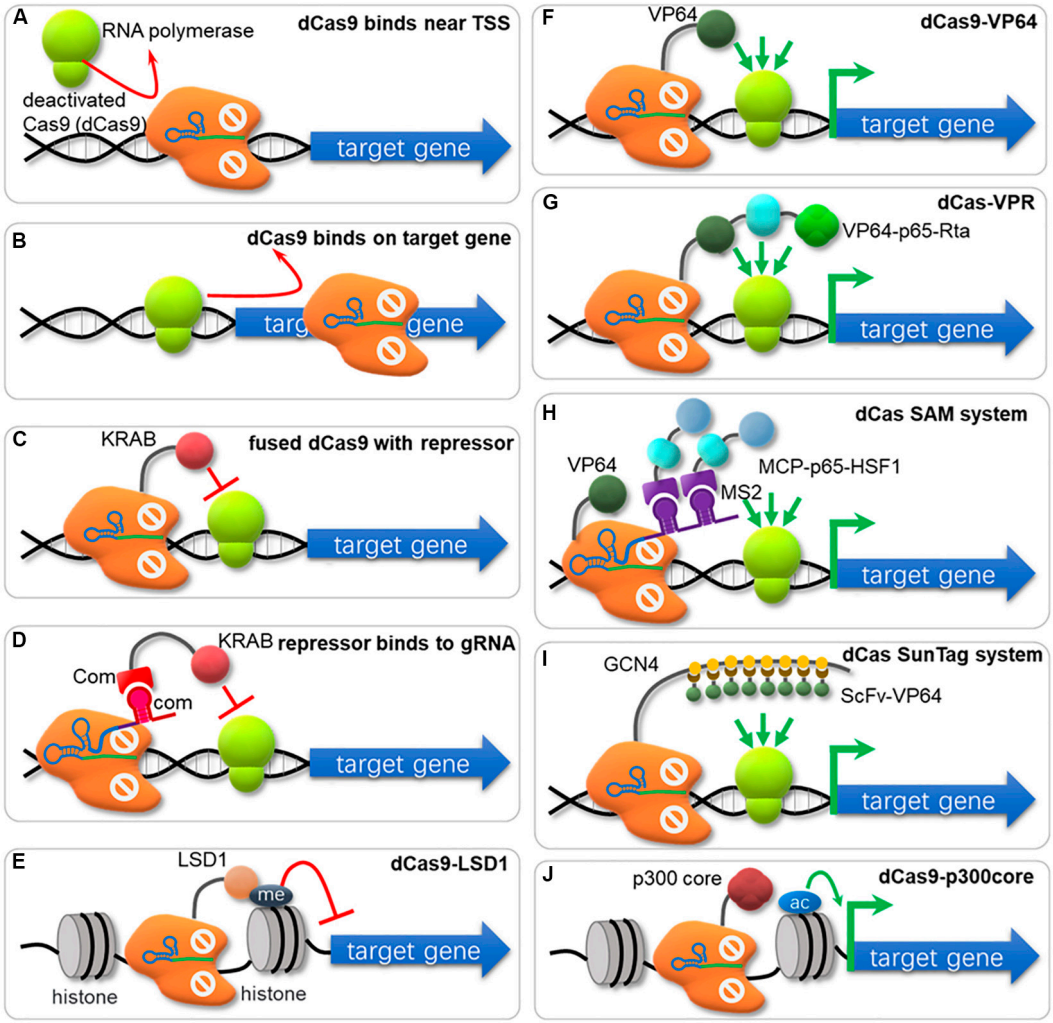


FIGURE 3 | The nuclease-deactivated Cas9 (dCas9) mediated CRISPRi and CRISPRa system. **(A)** dCas9 blocks recruiting of RNA polymerase (RNAP). **(B)** dCas9 can sterically block the transcriptional elongation of RNAP. **(C)** dCas9 fuses repressors (e.g., KRAB, Mxi1) to repress gene transcription. **(D)** KRAB is fused to RNA-binding domains (e.g., COM-KRAB) and achieves gene repression when targeting DNA sites overlaps the TSS using an scaffold RNA. **(E)** Fusion of dCas9 with the histone demethylase LSD1 suppresses gene expression. **(F)** Fusion of dCas9 with activators (e.g., VP64, ω-subuint of RNAP) activates gene transcription. **(G)** The VPR strategy for gene activation. The dCas9 has been fused to the combinatory transcriptional activator VP64-p65-Rta (VPR) to amplify the activation effects. **(H)** The SAM system. The dCas9 is fused to VP64 and the sgRNA has been modified to contain two MS2 RNA aptamers to recruit the MS2 bacteriophage coat protein (MCP), which was fused to the transcriptional activators p65 and heat shock factor 1 (HSF1). **(I)** The SunTag system. The tandem repeats of a small peptide GCN4 are utilized to recruit multiple copies of scFv (single-chain variable fragment) in fusion with the transcriptional activator VP64. **(J)** dCas9 is fused with the catalytic core of the human acetyltransferase p300, which catalyzes acetylation of histone H3 lysine 27 at its target sites, corresponding with robust transcriptional activation.

of CRISPR toolbox enables nearly all engineering types, and comprehensive applications of various CRISPR tools could solve this problem. Vanegas et al. (2017) developed a CRISPR/CRISPRi system termed SWITCH, where the Cas9 cassette was integrated into genome for genetic engineering as stage 1; and then the dCas9 cassette was integrated and replaced the Cas9 cassette for transcriptional regulation as stage 2 in *S. cerevisiae*. However, the SWITCH system does not enable genomic engineering and regulation control simultaneously. Lian et al. (2017) developed an orthogonal tri-functional CRISPR system that combines transcriptional activation, transcriptional interference, and gene deletion (CRISPR-AID, **Figure 4A**) in the yeast *S. cerevisiae*. This orthogonal tri-functional CRISPR system employed dLbCpf1-VP for CRISPRa, dSpCas9-RD1152 for CRISPRi, and SaCas9 for CRISPRd (gene deletion), which recognize different type of sgRNA and PAMs (Lian et al., 2017). By combining array-synthesized oligo pools, CRISPR-AID was further developed as a genome-wide system (MAGIC) to generate diversified genomic libraries to identify genetic determinants of complex phenotypes in yeast (Lian et al., 2019). This system was highlighted for complex engineering (gene interference, activation and deletion), high coverage (nearly 100% ORFs

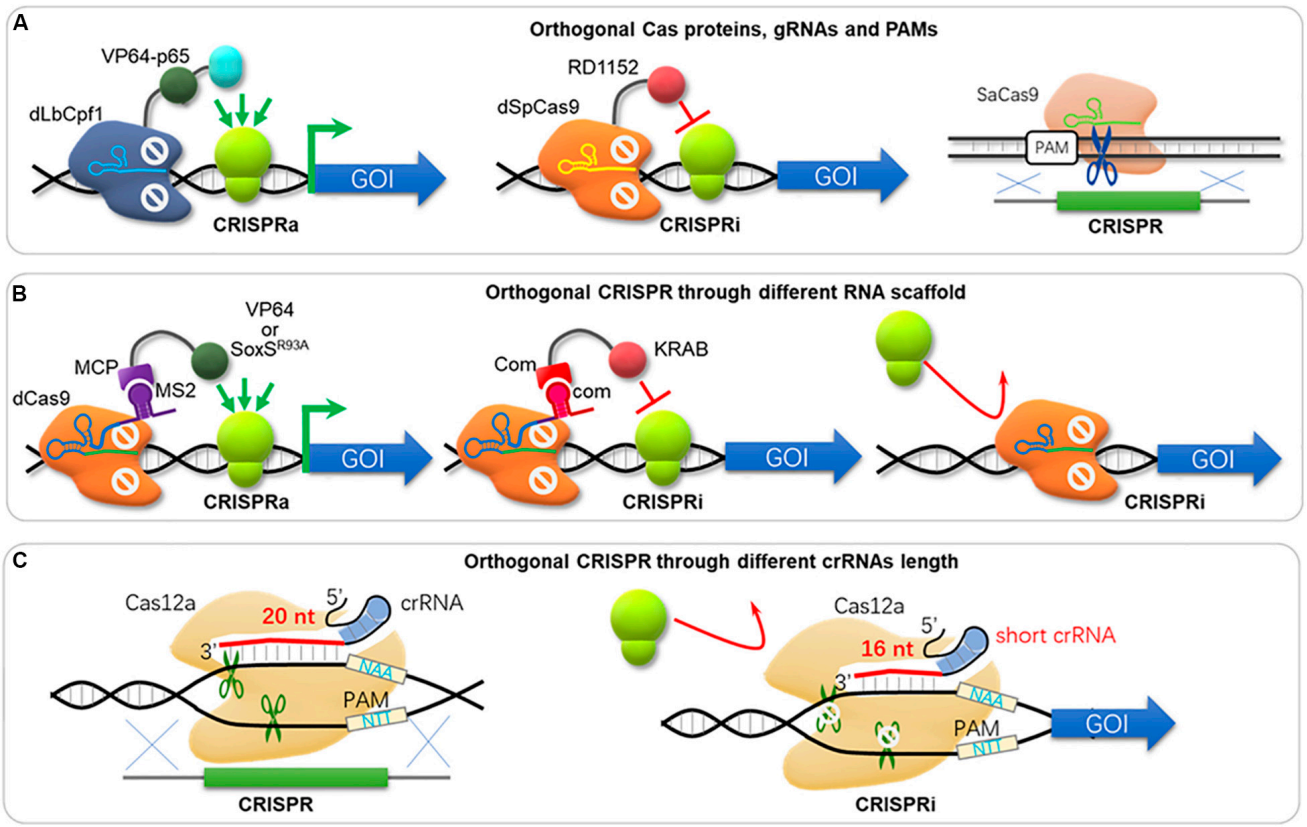


FIGURE 4 | Orthogonal CRISPR systems. **(A)** The orthogonal tri-functional CRISPR system that combines transcriptional activation, transcriptional interference, and gene deletion (CRISPR-AID). This orthogonal tri-functional CRISPR system employed dLbCpf1-VP for CRISPRa, dSpCas9-RD1152 for CRISPRi, and SaCas9 for CRISPRd (gene deletion), which recognized different type of sgRNA and PAMs. dLbCpf1, dCpf1 from *Lachnospiraceae bacterium* ND2006; dSpCas9, dCas9 from *S. pyogenes*; SaCas9, Cas9 from *S. aureus*. **(B)** The orthogonal CRISPR system with different RNA scaffolds. The gRNA fused with MS2 is used for activation through binding of MCP-VP64 or MCP-SoxS^{R93A}. The gRNA fused with com is used for repression through binding of Com-KRAB. Alternatively, a sgRNA without MS2 or com scaffold can hinder gene expression either. **(C)** CRISPR and CRISPRi via different crRNA length. Cas12a triggers DSB and genome editing with 20 bp-spacer in crRNA, while it blocks transcription with a short crRNA (16 bp-spacer).

and RNA genes) and iterative/simultaneous construction, which enabled identification of new gene targets and interactions for furfural tolerance as a demonstration (Lian et al., 2019). Combining orthogonal CRISPR and CRISPRi enables genome engineering and transcriptional regulation in *E. coli*, where orthogonal Cas protein candidates were expressed for CRISPR and CRISPRi separately and simultaneously (Sung et al., 2019). Sung et al. (2019) harnessed the St1Cas9 (from *Streptococcus thermophilus*) for DNA cleavage and insertion, and the SpdCas9 for CRISPRi. In addition to orthogonal effectors, RNA scaffold and binding protein can also be used for CRISPRi and CRISPRa simultaneously. Zalatan et al. used "scaffold RNAs" (scRNA) to recruit activators or repressors (e.g., using MS2 to recruit MCP-VP64 and com to recruit Com-KRAB, **Figure 4B**; Zalatan et al., 2015). Thus, genes are activated or repressed depending on the scRNA features instead of Cas9 orthologs. Another strategy of simultaneous activation and interference was achieved by using one dCas9 protein but MS2 scRNAs for activation by recruiting MCP-(5aa)-SoxS^{R93A}, while an unmodified gRNAs for repression (**Figure 4B**; Dong et al., 2018). On the other hand, one Cas12a was used for both gene editing and repression simultaneously by supplemented crRNA with different length (**Figure 4C**), where a 20 bp-crRNA triggers DSB and genome editing, but a 16 bp-crRNA results in gene repression without DNA cleavage (Liu W. et al., 2019).

CRISPR system can also be dynamically controlled by chemical or light with specific wavelength (ligand). Generally, a ligand induces dimerizing of two ligand binding domains (LBDs), and each domain can be fused to dCas9 and transcription effector (e.g., VPR for activation, and KRAB for repression), respectively. In such a ligand inducible CRISPRa/CRISPRi system, the presence of ligand will induce the binding of dCas9 and effector, and thus activate or repress the downstream gene expression. Several ligands have been reported for development of inducible CRISPR systems, including abscisic acid (inducing dimerization of ABI-PYL1) (Gao Y. et al., 2016; Bao et al., 2017; Chen T. et al., 2017), gibberellin (inducing dimerization of GID1-GAI24) (Gao Y. et al., 2016), rapamycin (inducing dimerization of FKBP–FRB) (Zetsche et al., 2015b; Bao et al., 2017), magnet (inducing dimerization of pMag–nMag) (Nihongaki et al.,

2015a,b, 2017; Polstein and Gersbach, 2015), blue light (inducing dimerization of CRY2-CIB1), and phytochrome-based red light (inducing dimerization of PhyB–PIF (Levskaya et al., 2009). When orthogonal dCas proteins are used to response to different ligands and effector-LBDs, the CRISPR system is expected for complex, dynamic, and programmable regulations (Gao Y. et al., 2016; Bao et al., 2017; Hill et al., 2018; Xu and Oi, 2019).

Precise Single Base Editing With CRISPR

Since Cas9 can tolerate mismatches in the 20 bp gRNA binding region, single-nucleotide mutations in this region could be bound and cleaved again. Thus, single-nucleotide mutations become difficult for CRISPR system. Such repeated cleavage can be avoided by introduction of additional mutations to eliminate the gRNA target site or the PAM sequence (DiCarlo et al., 2013; Jakociunas et al., 2015; Laughery et al., 2015). However, extra mutations are introduced for avoiding repeated cleavage. A two-step strategy (**Figure 5A**) was developed for precise single mutation by introducing the CRISPR/Cas9 twice (Biot-Pelletier and Martin, 2016; Paquet et al., 2016; Wang Y. et al., 2016). In the first step, the target was eliminated by insertion of a 20 nucleotide heterologous stuffer sequence via CRISPR/Cas9 system; and in the second step, this stuffer was eliminated by the original sequence with desired point mutation via CRISPR/Cas9 system (Biot-Pelletier and Martin, 2016; Paquet et al., 2016; Wang Y. et al., 2016).

Another method for processing precise base editing is to use dCas9 fused deaminase, which hydrolyzes the amine group of 'C' and 'A,' and enables 'C' to 'T' and 'A' to 'G' conversions without dsDNA cleavage (**Figures 5B,C**). Cytidine base editors (CBEs) and adenine base editors (ABEs) were developed to convert 'C' to 'T' (Komor et al., 2016) and 'A' to 'G' (Gaudelli et al., 2017) separately. Typically, BE3 (the mainly used CBE, cytidine deaminase-nCas9-UGI) and ABE7.10 (the most widely used ABE, wtTadA-mutantTadA-nCas9) showed the highest editing efficiency within the protospacer position 4–8 and 4–7 (counting the PAM as positions 21–23) (Komor et al., 2016; Gaudelli et al., 2017). CBEs using LbCpf1 showed an editing window preference of positions 10-12 (Li et al., 2018). In a recent research, a single-base editing termed CRISPR-BEST was developed by fusing Cas9 nickase (D10A) to cytidine and adenosine deaminase as editors. The CRISPR-BEST enabled 'C→T' and 'A→G' conversion within a window of approximately 7 and 6 nucleotides, respectively, with high efficiency in *Streptomyces* species (Tong et al., 2019). In another research, Zhao et al. fused dCas9 with PmCDA1 (the cytidine deaminase from *Petromyzon marinus*) and UGI (the uracil DNA glycosylase inhibitor), which enabled point mutations from 'C' to 'T' ('C→T') in *Streptomyces coelicolor*, and the efficiency reached up to 100%, 60%, and 20% for one, two and three loci, respectively (Zhao et al., 2019). Wang et al. fused Cas9 nickase (D10A) with activation-induced cytidine deaminase, which enabled precise 'C→T' conversion at one,

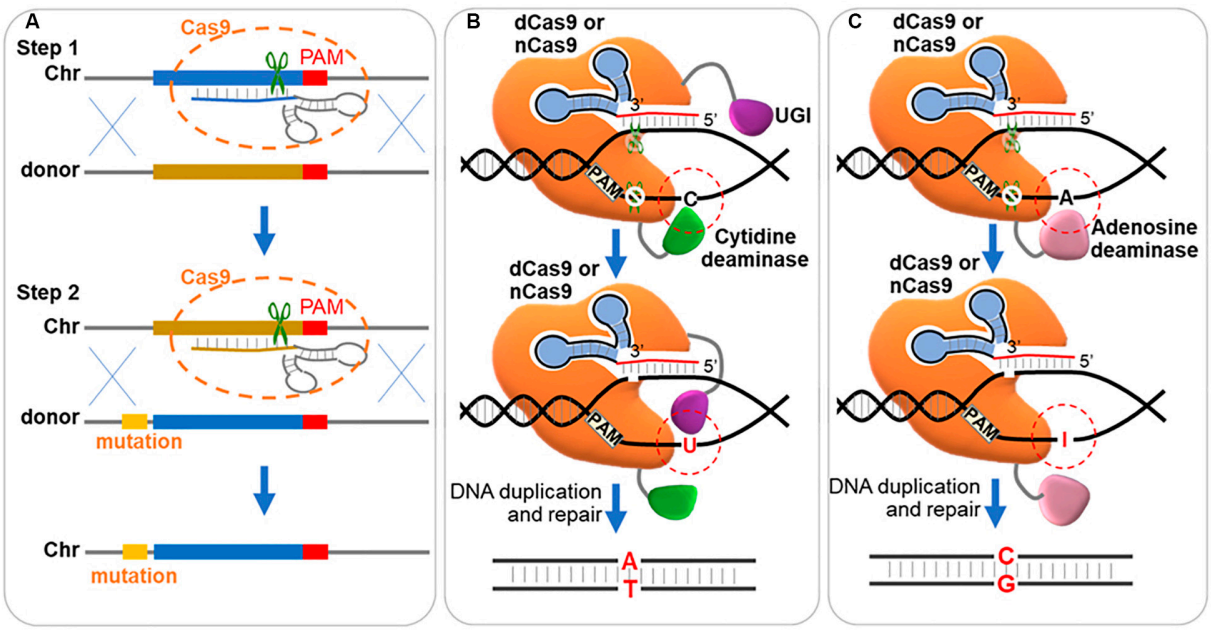


FIGURE 5 | Strategies for precise single base editing. **(A)** Two-step stuffer-assisted point mutation. In the first step, a 20-nucleotide target genome sequence close to the target is replaced by a heterologous stuffer fragment via homologous recombination. In the second step, the stuffer fragment acts as the target sequence, recognized by a second gRNA, and the original sequence with arbitrary mutation is inserted back. Chr, chromosome. **(B)** 'C→T' mutation through DSB independent pathway. The dCas9 or Cas9 nickase (nCas9, D10A) is fused with cytidine deaminase and uracil DNA glycosylase inhibitor (UGI), and binds to a DNA target. Cytidine deaminase converts the cytidine ('C') to uracil ('U') in the non-targeted strand, which is protected by UGI from the nucleotide excision repair (NER) pathway. And in the next replication cycle, the 'G:C' base pair is repaired to 'T:A'. **(C)** 'A→G' mutation through DSB independent pathway. The dCas9 or nCas9 is fused with adenosine deaminase and binds to a DNA target. Adenosine deaminase converts the adenosine ('A') to hypoxanthine ('I') in the non-targeted strand. And in the next replication cycle, the 'T:A' base pair is repaired to 'C:G'.

two and three loci with an efficiency of 100%, 87%, and 23%, respectively, in *Corynebacterium glutamicum*, and built a library of 14154 unique gRNAs for inactivation of 2726 genes (Wang Y. et al., 2018). The low efficiency of triple-site editing could be possibly caused by the lower amount of the base editor at each locus than those targeting single loci. And a developed system with expanding targeting scope, editing window, and base transition capability was further constructed in *C. glutamicum* by the same group (Wang et al., 2019).

APPLICATION OF CRISPR/CAS SYSTEM IN MICROBIAL BIOTECHNOLOGY

The fast developed and multiple functioned CRISPR system enables versatile, systematic and automatic applications in microbial technology. Especially, the CRISPR/Cas9 system has been developed for fast, efficient, precise and concise multi-loci editing and metabolic engineering. These researches imploring CRISPR/Cas system for hyper or wider applications in recent two years are shown in **Table 2**. And efforts for promoting CRISPR system for multi-loci editing and metabolic engineering are highlighted.

Promotion of CRISPR/Cas System for Multi-Loci Editing

One bias of CRISPR/Cas system is that the Cas/sgRNA complex can bind to more than one loci when proper sgRNAs are provided, which enables multi-loci editing simultaneously. Several groups have developed the CRISPR/Cas9 system for more efficient multi-loci editing, which makes genomic engineering more efficient, simple and convenient. *E. coli* and *S. cerevisiae* are typical model strains for prokaryotic and eukaryotic organisms, respectively, and multi-loci editing strategies are well illustrated thereby, enlightening adapted multi-loci editing strategies in other organisms (Gao S. et al., 2016; Wang J. et al., 2018; Zhang et al., 2018; Liu D. et al., 2019; Schultz et al., 2019; Tran et al., 2019; Zheng et al., 2019; Yang et al., 2020).

E. coli is the most intensively researched prokaryotic model microorganism, and multi-loci editing mediated by CRISPR/Cas is typical in *E. coli*. Jiang et al. expressed SpCas9 and λ-Red in *E. coli*, and achieved 3 genes disruption at an efficiency of 47% (Jiang et al., 2015). Ronda et al. expressed tracrRNA and crRNA separately, and achieved 2 genes disruption at an efficiency higher than 70% in *E. coli* (Ronda et al., 2016). Bassalo et al. (2016) developed a rapid and efficient one-step engineering method, and engineered 7 targets simultaneously with efficiencies ranging from 70 to 100%. Ao et al. (2018) expressed Cas12a instead of Cas9, resulting in the efficiency of integration of 2 loci at 40%, and the efficiency of integration of 3 loci at 20%. Sung et al. developed a method that combined orthogonal CRISPR and CRISPRi and enabled constitutive knockdown of three genes, knock-in of *pyc* and knockout of *adhE*, without compromising the CRISPRi knockdown efficiency (Sung et al., 2019).

Saccharomyces cerevisiae is the most intensively researched eukaryotic model microorganism, which enables highly efficient multi-loci editing because of the high HDR rate. Several multi-loci editing systems have been developed, including CRISPRm, HI-CRISPR, CasEMBLR, GTR-CRISPR (Ryan and Cate, 2014; Bao et al., 2015; Jakociunas et al., 2015, 2018a; Zhang et al., 2019). Bao et al. expressed crRNA and tracrRNA separately, and *CAN1*, *ADE2* and *LYP1* were simultaneously disrupted in 4 days with an efficiency ranging from 27 to 87%. Furthermore, another three genes were simultaneously disrupted in 6 days with 100% efficiency (Bao et al., 2015). Ryan and Cate developed a CRISPRm system, where 1–3 sgRNAs were expressed by a tRNA promoter and fused to the 3′ end of the self-cleaving HDV ribozyme for protecting the sgRNA from 5′-exonucleolytic activities, and achieved modifications of 1–3 targets with 81–100% efficiency (Ryan and Cate, 2014; Ryan et al., 2014). In another research, sgRNAs were separated by a 28 nt stem-loop sequence and cleaved by Csy4 (a bacterial endoribonuclease from *Pseudomonas aeruginosa*) to generate multiple gRNAs from a single transcript for multiple gene deletion in *S. cerevisiae* (Ferreira et al., 2018). This strategy enabled a deletion of 4 genes simultaneously with an efficiency of 96% (Ferreira et al., 2018). Jakociunas et al. (2015) developed a strategy, termed CasEMBLR, for *in vivo* assembly of gene cassettes and integrated to genome at up to 3 cleavage loci by CRISPR with high efficiency (30.6%, when optimized gRNAs were used). By using this method, 15 exogenous DNA parts were correctly assembled and integrated into 3 genomic loci for carotenoid production in one transformation (Jakociunas et al., 2015, 2018a). Kildegaard adopted similar strategy for multi-architecture assembly and insertion (Kildegaard et al., 2019). Kuivanen et al. (2018) reported a high-throughput workflow for CRISPR/Cas9 mediated combinatorial promoter replacements, and successfully edited 3 loci simultaneously with a frequency of 50%. Mans et al. used *in vitro* assembly for gRNAs expression and achieved simultaneous deletion of up to 6 genes in a single transformation step with a high efficiency at 65% (Mans et al., 2015). Zhang et al. report a gRNA-tRNA array for CRISPR-Cas9 (GTR-CRISPR) for multiplexed engineering, and simultaneously disrupted 8 genes with 87% efficiency, where gRNAs were fused with $tRNA^{GLY}$ scaffolds and expressed in 2 quadruple arrays (Zhang et al., 2019). Besides, Zhang et al. also reported an accelerated Lightning GTR-CRISPR strategy, which saving the cloning step in *E. coli* by directly transforming the Golden Gate reaction mix (the successfully assembled plasmid contained sgRNA expression cassettes and a Cas9 expression cassette) to yeast (Zhang et al., 2019). Bao et al. developed a CRISPR-Cas9- and homology-directed-repair-assisted genome-scale engineering method named CHAnGE, to construct genetic variant libraries in yeast (Bao et al., 2018). In CHAnGE, guide sequence and the homologous recombination (HDR) template were arranged and synthesized in a single oligonucleotide, and a oligonucleotide library of 24,765 unique guide sequences targeting 6,459 ORFs was synthesized on a chip and then assembled into a vector [pCRCT, harboring iCas9, tracrRNA expression cassettes and a promoter for sgRNA expression, as reported in HI-CRISPR system (Bao et al., 2015)] to build a pool of plasmids. This plasmid pool was then used to create a genome-wide gene disruption collection, in which more than 98% of target sequences were efficiently edited with an average

TABLE 2 | Selected recent CRISPR mediated metabolic engineering works.

Host	Toolbox	Product	Engineering by CRISPR	Achievements	References
Bacillus subtilis	CRISPRi	Hyaluronic acid (HA)	Reduce the expression of *pfkA* or *zwf*	Increased HA titer of up to 108% at 2.26 g/L and enhanced molecular weight	Westbrook et al., 2018
Bacillus subtilis	Xylose-induced CRISPRi	N-acetylglucosamine	Reduced the expression of *zwf*, *pfkA*, *glmM*	103.1 g/L in fed-batch fermentation	Wu et al., 2018b
Clostridium ljungdahlii	CRISPR/Cas9	Butyric acid	A butyric acid production pathway was integrated	1.01 g/L of butyric acid within 3 days by fermenting synthesis gas (CO_2/CO)	Huang et al., 2019
Clostridium ljungdahlii	CRISPRi	3-Hydroxybutyrate (3HB)	Repression of *pta* and *aor2*	Downregulation of *pta* increases 3HB production 2.3-fold with a titer at 21 mM	Woolston et al., 2018
Clostridium tyrobutyricum	Type I-B CRISPR-Cas	n-Butanol	Deletion of *spo0A* and *pyrF*, and integration of *adhE1* or *adhE2* to replace *cat1*	26.2g/L	Zhang et al., 2018
Corynebacterium glutamicum	Cas9 nickase (D10A) with activation-induced cytidine deaminase	Glutamate	Construction of a combinatorial gene inactivation library, and *pyk*/*ldhA* double inactivation for glutamate production	Increased production by 3-fold	Wang Y. et al., 2018
Synechocystis sp. PCC 6803	Inducible CRISPRi	n-Butanol	Repression of *gltA*	5-fold increase of carbon partitioning to n-butanol relative to a non-repression strain	Shabestary et al., 2018
Escherichia coli	CRISPR-Cas12a	5-Aminolevulinic acid	Integrating the *T7 RNAP* cassette and pT7-*hem1* cassette into the *lacZ* site and the *torS* site, respectively	1.55 g/L	Ao et al., 2018
Escherichia coli	Orthogonal CRISPR and CRISPRi systems	Succinate	Knock in *pyc*, knockout *adhE* and knockdown of *ptsG*, *ldhA*, and *pflB*	Increased by 178% with a titer at 2.5 g/L, and the titer increased to 15 g/L in a fermenter	Sung et al., 2019
Escherichia coli	CRISPRi	Naringenin 7-sulfate		Increased bioconversion rate by 2.83-fold (48.67%)	Chu et al., 2018
Escherichia coli	Iterative CRISPR EnAbled Trackable genome Engineering (iCREATE)	3HP	13 rounds of editing using iCREATE	Increased by up to 60-fold with a titer at 30 g/L	Liu et al., 2018a,b
Escherichia coli	PS-Brick assembly and CRISPR/Cas9	1-Propanol	*ppc*, *aspA*, *aspC*, *asd*, *pntAB*, *thrA*^{443}*BC*, *rhtC* were overexpressed and *tdh* and *ilvA* were deleted for threonine production; *kivD* and *ADH2* were expressed in *A*^{443}*BC* and *asd* expressed strain for 1-propanol production.	1.35 g/L in fed-batch fermentation	Liu S. et al., 2019
Escherichia coli	CRISPR/Cas9	Uridine	Expression of pyrimidine operon of Bacillus subtilis and *prs*, and deletion of *lacI*, *rihC*, *argF*, *thrA*, *iclR*, *purr*, *nupC* and *nupG*	70.3 g/L in fed-batch fermentation	Wu et al., 2018a

(Continued)

TABLE 2 | Continued

Host	Toolbox	Product	Engineering by CRISPR	Achievements	References
Escherichia coli	CRISPR/Cas9	Octanoic acid	Overexpression of *fabZ* and deletion of *fadE*, *fumAC* and *ackA*	Increased by 61% with a titer at 442 mg/L and further optimized to 1 g/L in fed-batch fermentation	Tan Z. et al., 2018
Escherichia coli	CRISPR/Cas9	Itaconic acid	Deletion of *ldhA*, *poxB* and *pflB*	3.06 g/L	Yang et al., 2018
Escherichia coli	CRISPRi	Isopentenol	Reduced expression of *asnA*, *prpE* and *gldA*	Increased by 98%	Tian et al., 2019
Escherichia coli	CRISPR/Cas9	Uridine		5.6 g/L	Li Y. et al., 2019
Halomonas bluephagenesis	CRISPR/Cas9	Poly(3-hydroxybutyrate-co-3-hydroxyvalerate) (PHBV)	Deletion of *sdhE* and *icl*	6.3 g/L cell dry weight (CDW), 65% PHBV in CDW and 25mol% 3HV in PHBV	Chen et al., 2019
Halomonas spp.	CRISPR/Cas9	P(3HB-co-3HV) consisting of 3-hydroxybutyrate (3HB) and 3-hydroxyvalerate (3HV)	Deletion of *prpC*	Increased 3HV fraction in the copolymers by approximately 16-folds with a fraction at 11.81 mol%	Qin et al., 2018
Klebsiella pneumoniae	CRISPRi	3HP	Deletion of *pmd*, *ldhA*, *aldA and mgsA*	Increased 3HP titer by 37% by reducing lactic acid synthesis, and further enhanced to 36.7 g/L 3-HP in fed-batch cultivation	Wang J. et al., 2018
Saccharomyces cerevisiae	GTR-CRISPR	Fatty acids	Knocking out of *FAA1*, *FAA4*, *POX1*, *ARE2*, *PAH1*, *LPP1*, *DPP1,* and *ARE1*	Increased free fatty acids by 30-fold with a titer at 559.52 mg/L, and increased total fatty acids by 1.8-fold with a titer at 943.92 mg/L	Zhang et al., 2019
Saccharomyces cerevisiae	CRISPRi	β-amyrin	down-regulating *ADH1*, *ADH4*, *ADH5*, *ADH6*, *CIT2*, *MLS2,* and *ERG7*	156.7 mg/L	Ni et al., 2019
Saccharomyces cerevisiae	CRISPRa		CRISPR/Cas-based gene activation library	Developed a CRISPR/Cas-based gene activation library, and improved thermotolerance	Li P. et al., 2019
Saccharomyces cerevisiae	CRISPR mediated genome shuffling			improved thermotolerance	Mitsui et al., 2019
Saccharomyces cerevisiae	CRISPRi and *in vivo* assembly	*cis*, *cis*-Muconic acid	Integration of multiple expression cassettes and down-regulating of *ZWF1*	Increased the titer by 5–21%	Kildegaard et al., 2019
Saccharomyces cerevisiae	CRISPRi, construction of tRNA-sgRNA operons using LEGO	2,3-Butanediol (BDO)	Knocking down *ADH1/3/5* and *GPD1*, and overexpression of *BDH1*	Increased BDO titer by 2-fold	Deaner et al., 2018
Saccharomyces cerevisiae	CRISPRa	3HP	A gRNA library targeting 168 genes	increased by 15 - 36%	Ferreira et al., 2019
Synechocystis sp.	CRISPRi	Fatty alcohols	Repression of aar, ado, sll1848, sll1752, slr2060, and slr1510	Increased by 3-fold with a specific titer of octadocanol at 10.3 mg/g DCW	Kaczmarzyk et al., 2018
Ustilago maydis	CRISPR/Cas9	Itaconic acid	Δ*cyp3*, ΔMEL, ΔUA, and ΔP_{ria1}::P_{etef}	Increased by 10.2-fold with a yield at 19.4 g/L and further enhanced to 53.5 g/L under optimized medium	Becker et al., 2019

frequency of 82% (Bao et al., 2018). In parallel, Jakociunas et al. employed error-prone PCR to generate DNA mutant libraries as donor, and used Cas9-mediated genome integration to introduce mutations at single- or multi-loci with efficiencies reaching 98-99%, for robust directed evolution (Jakociunas et al., 2018b). Besides, large chromosomal fragment deletion methods were developed based on CRISPR/Cas9 system. Easmin developed a guide RNA-transient expression system (gRNA-TES), where two sgRNA expression fragments (locating to each end of target region on genome) and DNA donor containing CgLEU2 were co-transformed into host for a replacement of up to 500-kb regions with efficiencies of 67-100% (Easmin et al., 2019).

In multi-loci editing using CRISPR systems, co-expression of many sgRNAs often requires repetitive DNA sequences (e.g., repeated promoters/terminators and guide RNA scaffolds), which possibly triggers genetic instability and phenotype loss. Reis et al. reported a non-repetitive extra-long sgRNA arrays (ELSAs) strategy, where different promoters, terminators and the sgRNAs' 61-nucleotide handle sequences were characterized for multiplex sgRNA expression (Reis et al., 2019). Through ELSAs, 22 sgRNAs within non-repetitive extra-long sgRNA arrays are simultaneously expressed for CRISPRi system, and repressed up to 13 genes by up to 3,500-fold in *E. coli* (Reis et al., 2019). The design of ELSAs and the identified 28 sgRNA handles that bind Cas9 can be adopted for CRISPR mediated multi-loci editing for metabolic engineering and synthetic biology applications in other organisms.

The CRISPR/Cas Mediated Metabolic Engineering

The developing powerful CRISPR toolbox enables advanced genome editing and transcription regulation, and has become the ideal strategy for metabolic engineering, because of its advantage of ease of use, modularity, and scalability. Metabolic engineering rewrites the metabolic network through single or multiple gene manipulation, to create or improve microbial cell factories for the production of fuels, chemicals, pharmaceutics, etc. CRISPR systems have been increasingly used in metabolic engineering field for construction of microbial cell factories (Yan and Fong, 2017; Mougiakos et al., 2018; Tarasava et al., 2018), and those recent works are summarized in **Table 2**.

One advantage of CRISPR/Cas system is that it realizes precise genome editing at multi-loci in one transformation, without integrating a marker gene on genome for selection, and thus it would largely simplify operation steps and save time and labor in metabolic engineering works. As a proof of concept, Zhang et al. employed GTR-CRISPR to engineer lipid metabolism in *S. cerevisiae* for free fatty acid (FFA) production (Zhang et al., 2019). 8 genes in lipid metabolism were deleted through two rounds operation: *FAA1*, *FAA4*, *POX1,* and *ARE2* were deleted in the first round; and after losing the plasmid through anti-selection on 5-FOA medium, *PAH1*, *LPP1*, *DPP1,* and *ARE1* were knocked out in the second round transformation (Zhang et al., 2019). Thus, the final strain with 8-gene deletion was constructed in 10 days, which produced 559.52 mg/L FFA with 30-fold increase compared with wildtype.

Application of orthogonal CRISPR systems would also make complex metabolic engineering work simpler and more efficient, and knocking-in, knocking-out, interference and activation could be simultaneously processed for multiplex target genes. Several excellent examples for orthogonal CRISPR aided metabolic engineering were demonstrated recently (**Table 2**). For example, Sung et al. employed a Cas9 protein from *Streptococcus thermophilus* CRISPR1 (St1Cas9) to deliver DNA cleavage, and used the common dSpCas9 for gene interference (Sung et al., 2019). Each Cas9 recognized its cognate sgRNA, and worked orthogonally. Thus, St1Cas9 was harnessed to integrate SpdCas9 and sgRNA arrays, as well as knock in *pyc* and knockout *adhE*; whereas SpdCas9 was applied for constitutive knockdown of *ptsG*, *ldhA*, and *pflB* to eliminate competing pathways for lactate, formate, and ethanol synthesis. The final engineered strain produced 2.5 g/L succinate with 178% improvement (Sung et al., 2019).

Other CRISPR Applications

The fast development of CRISPR tools enable various applications beyond genome editing and transcriptional regulation. One application is building activated and/or interfered gene libraries to screen phenotype related genes. Gilbert et al. applied genomic libraries of CRISPRi and CRISPRa to screen gene targets related to the sensitivity to a cholera-diphtheria toxin (Gilbert et al., 2014). Li et al. build a CRISPR/Cas-based gene activation library, and used it to screen gene targets for improved thermotolerance in *S. cerevisiae* (Li P. et al., 2019). Lee et al. used a CRISPRi system, targeting 4,565 (99.7%) genes to identify a minimal set of genes required for rapid growth of *Vibrio natriegens* (Lee et al., 2019). Bassalo et al. applied CRISPR/Cas9 to perform a parallel and high-resolution interrogation of over 16,000 mutations to identify proteins associated to lysine metabolism in *E. coli* (Bassalo et al., 2018). While Wang et al. built a larger guide RNA library of ∼60,000 members for coding and non-coding targets in *E. coli*, and applied CRISPRi system to associate genes with phenotypes at the genome level (Wang T. et al., 2018).

CRISPR system can also be used to discover novel compounds by activating the expression of silent gene or gene cluster, which may code enzymes for novel or undetectable nature products synthesis. Zhang et al. reported an one-step CRISPR/Cas9 knock-in strategy to activate biosynthetic gene cluster expression and trigger metabolite production by insertion of strong promoters upstream biosynthetic operons in *Streptomyces* species (Zhang et al., 2017; Lim et al., 2018). Grijseels et al. (2018) implemented the CRISPR/Cas9 technology to identify the decumbenone biosynthetic gene cluster in *Penicillium decumbens*, and evaluated the importance of targets for production of calbistrin. Similarly, Lee et al. (2018) adopted the CRISPRi system for rapid identification of unknown carboxyl esterase activity in *C. glutamicum*. Naseri et al. (2019) employed orthogonal, plant-derived artificial transcription factors (ATFs) for the balanced expression of multiple genes in *S. cerevisiae*, and generated CRISPR/Cas9-mediated cell

libraries for producing β-carotene and co-producing β-ionone and biosensor-responsive naringenin.

CRISPR/Cas9 system also amplified the power of evolutionary engineering for industrial microorganisms. Mitsui et al. developed CRISPR/Cas9 system as a genome shuffling method for evolutionary engineering to obtain a thermotolerant mutant strain (Mitsui et al., 2019). Halperin et al. (2018) proposed a new method called EvolvR that can accelerate mutagenesis up to 7,770,000-fold within a tunable window length via CRISPR-guided nickases. Jakoèiûnas et al. reported a method named Cas9-mediated Protein Evolution Reaction (CasPER) for efficient mutagenesis of nucleotides by combining error-prone PCR and Cas9-mediated genome integration (Jakociunas et al., 2018b). Garst et al. (2017) constructed CRISPR-enabled trackable genome engineering (CREATE) method, where a library of targets was built and transformed for multiplex editing *in vivo*, followed by screening and mutation identification. Through CREATE, a library of 10^4–10^6 individual members was built, and an average mutation rate of 75% was reached for site saturation mutagenesis for protein engineering and adaptive laboratory evolution (Garst et al., 2017). Based on CREATE, Liu et al. developed an iterative CRISPR EnAbled Trackable genome Engineering (iCREATE) strategy for the rapid construction of combinatorially modified genomes, and used it for 3-hydroxypropionate (3HP) production improvement (Liu et al., 2018a,b).

CONCLUSION AND FUTURE PERSPECTIVES

The intrinsic advantage of CRISPR enables an evolutionary and versatile platform for genotypic, metabolic and phenotypic engineering in microbial biotechnology. The CRISPR based tools are generally with higher efficiency, more convenience, more efficient multiplex targets editing/regulation and time-saving compared with traditional ones. However, challenges and weaknesses still exist. Despite the CRISPR/Cas system has been used for a broad range of microorganisms, the genome editing efficiency varies between species to species and even between cell to cell, indicating cellular intrinsic process impacts CRISPR/Cas system. More reliable, inducible and widely applicable expression architectures, e.g., RNAPII- and RNAPIII-promoters, ARSs and centromere sequences can be developed for multi-hosts, which would enable the expression of Cas effectors and gRNAs in different organisms with simple modification, especially in non-model microorganism, which would make CRISPR a portable platform and transplant CRISPR strategies from model microorganism to those non-model ones. The efficiency of CRISPR system (both for genome editing and transcriptional regulation tools) showed a gRNA position reliable phenomenon, which means high-efficiency on some gRNAs, but low-efficiency or even non-work on others. Therefore, more than one gRNAs should be tested when editing a new target, especially for efficient CRISPRa and CRISPRi. Thus, it remains important for developing more powerful effectors for robust activation/inference, engineering the Cas protein for better performance, and developing algorithms that can predict and design efficient gRNAs for CRISPRa and CRISPRi at single nucleotide level. Another highlighted direction is a comprehensive application of different CRISPR/Cas systems to facilitate insertion, deletion and transcriptional regulation simultaneously. Lian et al. developed a such strategy in *S. cerevisiae* (Lian et al., 2017), enabling perturbation of the metabolic and regulatory networks in a modular, parallel, and high-throughput manner, which is worthy to adapt such strategy in other organisms. Besides, with the genome wide application of CRISPR and array-synthesized oligo pools, it is more easier to generate large libraries containing millions and even billions of variants (Lian et al., 2019). Therefore, developing high throughput techniques, e.g., high efficient transformation methods, robotic platforms and microfluidic systems remain necessary and challenging. Furthermore, with the aid of automated robotic systems (HamediRad et al., 2019), CRISPR system could become more powerful for functional mapping and multiplex optimization of strains in an unprecedented scale.

On the other hand, types VI and III CRISPR systems were reported to have specialized or pluralistic for RNA targeting activity (Shmakov et al., 2017), which enabled direct RNA engineering by CRISPR systems (Abudayyeh et al., 2016, 2017). Despite limitations in those RNA-targeting CRISPR systems [reviewed in Smargon et al. (2020)], it has showed capabilities in RNA imaging (Abudayyeh et al., 2017), RNA interference (Abudayyeh et al., 2016), RNA mutation (Abudayyeh et al., 2017) and RNA detection (Gootenberg et al., 2017, 2018). Thus CRISPR aided RNA manipulation shows bright prospect as an emerging tool in fundamental research and bioengineering.

AUTHOR CONTRIBUTIONS

WD and SS outlined this manuscript. WD drafted the manuscript. SS and YZ revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Key Research and Development Program of China (2018YFA0901800 and 2018YFA0900100), the National Natural Science Foundation of China (21878013), the China Postdoctoral Science Foundation (2019M650450), the Fundamental Research Funds for the Central Universities (ZY1933), the Foundation of Key Laboratory of Biomass Chemical Engineering of Ministry of Education, Zhejiang University (No. 2018BCE004), the Fundamental Research Funds for the Central Universities, and the Beijing Advanced Innovation Center for Soft Matter Science and Engineering.

ACKNOWLEDGMENTS

We thank Prof. Huimin Zhao (University of Illinois at Urbana-Champaign) for his very enlightening comments.

REFERENCES

Abudayyeh, O. O., Gootenberg, J. S., Essletzbichler, P., Han, S., Joung, J., Belanto, J. J., et al. (2017). RNA targeting with CRISPR-Cas13. *Nature* 550, 280–284. doi: 10.1038/nature24049

Abudayyeh, O. O., Gootenberg, J. S., Konermann, S., Joung, J., Slaymaker, I. M., Cox, D. B., et al. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* 353:aaf5573. doi: 10.1126/science.aaf5573

Ao, X., Yao, Y., Li, T., Yang, T. T., Dong, X., Zheng, Z. T., et al. (2018). A multiplex genome editing method for *Escherichia coli* based on CRISPR-Cas12a. *Front. Microbiol.* 9:2307. doi: 10.3389/fmicb.2018.02307

Armario Najera, V., Twyman, R. M., Christou, P., and Zhu, C. (2019). Applications of multiplex genome editing in higher plants. *Curr. Opin. Biotechnol.* 59, 93–102. doi: 10.1016/j.copbio.2019.02.015

Bae, S., Park, J., and Kim, J. S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 30, 1473–1475. doi: 10.1093/bioinformatics/btu048

Bae, S. J., Park, B. G., Kim, B. G., and Hahn, J. S. (2020). Multiplex gene disruption by targeted base editing of *Yarrowia lipolytica* genome using cytidine deaminase combined with the CRISPR/Cas9 system. *Biotechnol. J.* 15:e1900238. doi: 10.1002/biot.201900238

Bao, Z., HamediRad, M., Xue, P., Xiao, H., Tasan, I., Chao, R., et al. (2018). Genome-scale engineering of *Saccharomyces cerevisiae* with single-nucleotide precision. *Nat. Biotechnol.* 36, 505–508. doi: 10.1038/nbt.4132

Bao, Z., Jain, S., Jaroenpuntaruk, V., and Zhao, H. (2017). Orthogonal genetic regulation in human cells using chemically induced CRISPR/Cas9 activators. *ACS Synth. Biol.* 6, 686–693. doi: 10.1021/acssynbio.6b00313

Bao, Z., Xiao, H., Liang, J., Zhang, L., Xiong, X., Sun, N., et al. (2015). Homology-integrated CRISPR-Cas (HI-CRISPR) system for one-step multigene disruption in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* 4, 585–594. doi: 10.1021/sb500255k

Bassalo, M. C., Garst, A. D., Choudhury, A., Grau, W. C., Oh, E. J., Spindler, E., et al. (2018). Deep scanning lysine metabolism in *Escherichia coli*. *Mol. Syst. Biol.* 14:e8371. doi: 10.15252/msb.20188371

Bassalo, M. C., Garst, A. D., Halweg-Edwards, A. L., Grau, W. C., Domaille, D. W., Mutalik, V. K., et al. (2016). Rapid and efficient one-step metabolic pathway integration in *E. coli*. *ACS Synth. Biol.* 5, 561–568. doi: 10.1021/acssynbio.5b00187

Becker, J., Hosseinpour Tehrani, H., Gauert, M., Mampel, J., Blank, L. M., and Wierckx, N. (2019). An *Ustilago maydis* chassis for itaconic acid production without by-products. *Microb. Biotechnol.* 13, 350–362. doi: 10.1111/1751-7915.13525

Bikard, D., Jiang, W. Y., Samai, P., Hochschild, A., Zhang, F., and Marraffini, L. A. (2013). Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* 41, 7429–7437. doi: 10.1093/nar/gkt520

Biot-Pelletier, D., and Martin, V. J. (2016). Seamless site-directed mutagenesis of the *Saccharomyces cerevisiae* genome using CRISPR-Cas9. *J. Biol. Eng.* 10:6. doi: 10.1186/s13036-016-0028-1

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., et al. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326, 1509–1512. doi: 10.1126/science.1178811

Cao, M., Fatma, Z., Song, X., Hsieh, P. H., Tran, V. G., Lyon, W. L., et al. (2020). A genetic toolbox for metabolic engineering of Issatchenkia orientalis. *Metab. Eng.* 59, 87–97. doi: 10.1016/j.ymben.2020.01.005

Cao, M., Gao, M., Lopez-Garcia, C. L., Wu, Y., Seetharam, A. S., Severin, A. J., et al. (2017). Centromeric DNA facilitates nonconventional yeast genetic engineering. *ACS Synth. Biol.* 6, 1545–1553. doi: 10.1021/acssynbio.7b00046

Cao, M., Gao, M., Ploessl, D., Song, C., and Shao, Z. (2018). CRISPR-mediated genome editing and gene repression in *Scheffersomyces stipitis*. *Biotechnol. J.* 13:e1700598. doi: 10.1002/biot.201700598

Ceccaldi, R., Rondinelli, B., and D'Andrea, A. D. (2016). Repair pathway choices and consequences at the double-strand break. *Trends Cell Biol.* 26, 52–64. doi: 10.1016/j.tcb.2015.07.009

Chang, H. H. Y., Pannunzio, N. R., Adachi, N., and Lieber, M. R. (2017). Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell Biol.* 18, 495–506. doi: 10.1038/nrm.2017.48

Chavez, A., Scheiman, J., Vora, S., Pruitt, B. W., Tuttle, M., et al. (2015). Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* 12, 326–328. doi: 10.1038/nmeth.3312

Chayot, R., Montagne, B., Mazel, D., and Ricchetti, M. (2010). An end-joining repair mechanism in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2141–2146. doi: 10.1073/pnas.0906355107

Chen, J. S., Dagdas, Y. S., Kleinstiver, B. P., Welch, M. M., Sousa, A. A., Harrington, L. B., et al. (2017). Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* 550, 407–410. doi: 10.1038/nature24268

Chen, T., Gao, D., Zhang, R., Zeng, G., Yan, H., Lim, E., et al. (2017). Chemically controlled epigenome editing through an inducible dCas9 system. *J. Am. Chem. Soc.* 139, 11337–11340. doi: 10.1021/jacs.7b06555

Chen, Y., Chen, X. Y., Du, H. T., Zhang, X., Ma, Y. M., Chen, J. C., et al. (2019). Chromosome engineering of the TCA cycle in Halomonas bluephagenesis for production of copolymers of 3-hydroxybutyrate and 3-hydroxyvalerate (PHBV). *Metab. Eng.* 54, 69–82. doi: 10.1016/j.ymben.2019.03.006

Cho, S., Shin, J., and Cho, B. K. (2018). Applications of CRISPR/Cas System to bacterial metabolic engineering. *Int. J. Mol. Sci.* 19:1089. doi: 10.3390/ijms19041089

Choudhary, E., Thakur, P., Pareek, M., and Agarwal, N. (2015). Gene silencing by CRISPR interference in mycobacteria. *Nat. Commun.* 6:6267. doi: 10.1038/ncomms7267

Christian, M., Cermak, T., Doyle, E. L., Schmidt, C., Zhang, F., Hummel, A., et al. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186, 757–761. doi: 10.1534/genetics.110.120717

Chu, L. L., Dhakal, D., Shin, H. J., Jung, H. J., Yamaguchi, T., and Sohng, J. K. (2018). Metabolic engineering of *Escherichia coli* for enhanced production of naringenin 7-sulfate and its biological activities. *Front. Microbiol.* 9:1671. doi: 10.3389/fmicb.2018.01671

Cobb, R. E., Wang, Y., and Zhao, H. (2015). High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS Synth. Biol.* 4, 723–728. doi: 10.1021/sb500351f

Cong, L., Ran, F. A., Cox, D., Lin, S. L., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823. doi: 10.1126/science.1231143

Deaner, M., Holzman, A., and Alper, H. S. (2018). Modular ligation extension of guide RNA operons (LEGO) for multiplexed dCas9 regulation of metabolic pathways in *Saccharomyces cerevisiae*. *Biotechnol. J.* 13:e1700582. doi: 10.1002/biot.201700582

Deveau, H., Barrangou, R., Garneau, J. E., Labonte, J., Fremaux, C., Boyaval, P., et al. (2008). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190, 1390–1400.

DiCarlo, J. E., Norville, J. E., Mali, P., Rios, X., Aach, J., and Church, G. M. (2013). Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* 41, 4336–4343. doi: 10.1093/nar/gkt135

Ding, D., Chen, K., Chen, Y., Li, H., and Xie, K. (2018). Engineering introns to express RNA guides for Cas9- and Cpf1-mediated multiplex genome editing. *Mol. Plant* 11, 542–552. doi: 10.1016/j.molp.2018.02.005

Doetschman, T., and Georgieva, T. (2017). Gene editing with CRISPR/Cas9 RNA-directed nuclease. *Circ. Res.* 120, 876–894. doi: 10.1161/CIRCRESAHA.116.309727

Dong, C., Fontana, J., Patel, A., Carothers, J. M., and Zalatan, J. G. (2018). Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nat. Commun.* 9:2489. doi: 10.1038/s41467-018-04901-6

Doudna, J. A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346:1258096. doi: 10.1126/science.1258096

Easmin, F., Hassan, N., Sasano, Y., Ekino, K., Taguchi, H., and Harashima, S. (2019). gRNA-transient expression system for simplified gRNA delivery in CRISPR/Cas9 genome editing. *J. Biosci. Bioeng.* 128, 373–378. doi: 10.1016/j.jbiosc.2019.02.009

Esvelt, K. M., Mali, P., Braff, J. L., Moosburner, M., Yaung, S. J., and Church, G. M. (2013). Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods* 10, 1116–1121. doi: 10.1038/nmeth.2681

Farzadfard, F., Perli, S. D., and Lu, T. K. (2013). Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. *ACS Synth. Biol.* 2, 604–613. doi: 10.1021/sb400081r

Ferreira, R., Skrekas, C., Hedin, A., Sanchez, B. J., Siewers, V., Nielsen, J., et al. (2019). Model-assisted fine-tuning of central carbon metabolism in yeast through dCas9-based regulation. *ACS Synth. Biol.* 8, 2457–2463. doi: 10.1021/acssynbio.9b00258

Ferreira, R., Skrekas, C., Nielsen, J., and David, F. (2018). Multiplexed CRISPR/Cas9 genome editing and gene regulation using Csy4 in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* 7, 10–15. doi: 10.1021/acssynbio.7b00259

Freed, E., Fenster, J., Smolinski, S. L., Walker, J., Henard, C. A., Gill, R., et al. (2018). Building a genome engineering toolbox in nonmodel prokaryotic microbes. *Biotechnol. Bioeng.* 115, 2120–2138. doi: 10.1002/bit.26727

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., et al. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 31, 822–826. doi: 10.1038/nbt.2623

Fu, Y. F., Sander, J. D., Reyon, D., Cascio, V. M., and Joung, J. K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* 32, 279–284. doi: 10.1038/nbt.2808

Gao, S., Tong, Y., Wen, Z., Zhu, L., Ge, M., Chen, D., et al. (2016). Multiplex gene editing of the *Yarrowia lipolytica* genome using the CRISPR-Cas9 system. *J. Ind. Microbiol. Biotechnol.* 43, 1085–1093. doi: 10.1007/s10295-016-1789-8

Gao, Y., Xiong, X., Wong, S., Charles, E. J., Lim, W. A., and Qi, L. S. (2016). Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. *Nat. Methods* 13, 1043–1049. doi: 10.1038/nmeth.4042

Garneau, J. E., Dupuis, M. E., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., et al. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71. doi: 10.1038/nature09523

Garst, A. D., Bassalo, M. C., Pines, G., Lynch, S. A., Halweg-Edwards, A. L., Liu, R., et al. (2017). Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nat. Biotechnol.* 35, 48–55. doi: 10.1038/nbt.3718

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 109, E2579–E2586. doi: 10.1073/pnas.1208507109

Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., et al. (2017). Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551, 464–471. doi: 10.1038/nature24644

Geller, S. H., Antwi, E. B., Di Ventura, B., and McClean, M. N. (2019). Optogenetic repressors of gene expression in yeasts using light-controlled nuclear localization. *Cell. Mol. Bioeng.* 12, 511–528. doi: 10.1007/s12195-019-00598-9

Generoso, W. C., Gottardi, M., Oreb, M., and Boles, E. (2016). Simplified CRISPR-Cas genome editing for *Saccharomyces cerevisiae*. *J. Microbiol. Methods* 127, 203–205. doi: 10.1016/j.mimet.2016.06.020

Gilbert, L. A., Horlbeck, M. A., Adamson, B., Villalta, J. E., Chen, Y., Whitehead, E. H., et al. (2014). Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* 159, 647–661. doi: 10.1016/j.cell.2014.09.029

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442–451. doi: 10.1016/j.cell.2013.06.044

Gootenberg, J. S., Abudayyeh, O. O., Kellner, M. J., Joung, J., Collins, J. J., and Zhang, F. (2018). Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. *Science* 360, 439–444. doi: 10.1126/science.aaq0179

Gootenberg, J. S., Abudayyeh, O. O., Lee, J. W., Essletzbichler, P., Dy, A. J., Joung, J., et al. (2017). Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science* 356, 438–442. doi: 10.1126/science.aam9321

Grijseels, S., Pohl, C., Nielsen, J. C., Wasil, Z., Nygard, Y., Nielsen, J., et al. (2018). Identification of the decumbenone biosynthetic gene cluster in *Penicillium decumbens* and the importance for production of calbistrin. *Fungal Biol. Biotechnol.* 5:18. doi: 10.1186/s40694-018-0063-4

Grissa, I., Vergnaud, G., and Pourcel, C. (2007a). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 35, W52–W57. doi: 10.1093/nar/gkm360

Grissa, I., Vergnaud, G., and Pourcel, C. (2007b). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* 8:172. doi: 10.1186/1471-2105-8-172

Gu, Y., Gao, J., Cao, M., Dong, C., Lian, J., Huang, L., et al. (2019). Construction of a series of episomal plasmids and their application in the development of an efficient CRISPR/Cas9 system in *Pichia pastoris*. *World J. Microbiol. Biotechnol.* 35:79.

Guilinger, J. P., Thompson, D. B., and Liu, D. R. (2014). Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* 32, 577–582. doi: 10.1038/nbt.2909

Halperin, S. O., Tou, C. J., Wong, E. B., Modavi, C., Schaffer, D. V., and Dueber, J. E. (2018). CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. *Nature* 560, 248–252. doi: 10.1038/s41586-018-0384-8

HamediRad, M., Chao, R., Weisberg, S., Lian, J., Sinha, S., and Zhao, H. (2019). Towards a fully automated algorithm driven platform for biosystems design. *Nat. Commun.* 10:5150. doi: 10.1038/s41467-019-13189-z

Heigwer, F., Kerr, G., and Boutros, M. (2014). E-CRISP: fast CRISPR target site identification. *Nat. Methods* 11, 122–123. doi: 10.1038/nmeth.2812

Hill, Z. B., Martinko, A. J., Nguyen, D. P., and Wells, J. A. (2018). Human antibody-based chemically induced dimerizers for cell therapeutic applications. *Nat. Chem. Biol.* 14, 112–117. doi: 10.1038/nchembio.2529

Hilton, I. B., D'Ippolito, A. M., Vockley, C. M., Thakore, P. I., Crawford, G. E., Reddy, T. E., et al. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* 33, 510–517. doi: 10.1038/nbt.3199

Horwitz, A. A., Walter, J. M., Schubert, M. G., Kung, S. H., Hawkins, K., Platt, D. M., et al. (2015). Efficient multiplexed integration of synergistic alleles and metabolic pathways in yeasts via CRISPR-Cas. *Cell Syst.* 1, 88–96. doi: 10.1016/j.cels.2015.02.001

Hou, Z., Zhang, Y., Propson, N. E., Howden, S. E., Chu, L. F., Sontheimer, E. J., et al. (2013). Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 15644–15649. doi: 10.1073/pnas.1313587110

Hsu, P. D., Lander, E. S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157, 1262–1278. doi: 10.1016/j.cell.2014.05.010

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832. doi: 10.1038/nbt.2647

Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., et al. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556, 57–63. doi: 10.1038/nature26155

Huang, H., Chai, C., Yang, S., Jiang, W., and Gu, Y. (2019). Phage serine integrase-mediated genome engineering for efficient expression of chemical biosynthetic pathway in gas-fermenting *Clostridium ljungdahlii*. *Metab. Eng.* 52, 293–302. doi: 10.1016/j.ymben.2019.01.005

Hwang, W. Y., Fu, Y. F., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., et al. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 227–229. doi: 10.1038/nbt.2501

Jacobs, J. Z., Ciccaglione, K. M., Tournier, V., and Zaratiegui, M. (2014). Implementation of the CRISPR-Cas9 system in fission yeast. *Nat. Commun.* 5:5344. doi: 10.1038/ncomms6344

Jakociunas, T., Jensen, E. D., Jensen, M. K., and Keasling, J. D. (2018a). Assembly and multiplex genome integration of metabolic pathways in yeast using CasEMBLR. *Methods Mol. Biol.* 1671, 185–201. doi: 10.1007/978-1-4939-7295-1_12

Jakociunas, T., Pedersen, L. E., Lis, A. V., Jensen, M. K., and Keasling, J. D. (2018b). CasPER, a method for directed evolution in genomic contexts using mutagenesis and CRISPR/Cas9. *Metab. Eng.* 48, 288–296. doi: 10.1016/j.ymben.2018.07.001

Jakociunas, T., Rajkumar, A. S., Zhang, J., Arsovska, D., Rodriguez, A., Jendresen, C. B., et al. (2015). CasEMBLR: Cas9-facilitated multiloci genomic integration of in vivo assembled DNA parts in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* 4, 1226–1234. doi: 10.1021/acssynbio.5b00007

Jansen, R., van Embden, J. D. A., Gaastra, W., and Schouls, L. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* 43, 1565–1575. doi: 10.1046/j.1365-2958.2002.02839.x

Jensen, E. D., Ferreira, R., Jakociunas, T., Arsovska, D., Zhang, J., Ding, L., et al. (2017). Transcriptional reprogramming in yeast using dCas9 and combinatorial gRNA strategies. *Microb. Cell Fact.* 16, 1–16. doi: 10.1186/s12934-017-0664-2

Jiang, W. Y., Bikard, D., Cox, D., Zhang, F., and Marraffini, L. A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* 31, 233–239. doi: 10.1038/nbt.2508

Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., and Yang, S. (2015). Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl. Environ. Microbiol.* 81, 2506–2514. doi: 10.1128/AEM.04023-14

Jiang, Y., Qian, F., Yang, J., Liu, Y., Dong, F., Xu, C., et al. (2017). CRISPR-Cpf1 assisted genome editing of *Corynebacterium glutamicum*. *Nat. Commun.* 8:15179. doi: 10.1038/ncomms15179

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A progrzaammable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10.1126/science.1225829

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-programmed genome editing in human cells. *eLife* 2:e00471. doi: 10.7554/eLife.00471

Joseph, R. C., Kim, N. M., and Sandoval, N. R. (2018). Recent developments of the synthetic biology toolkit for *Clostridium*. *Front. Microbiol.* 9:154. doi: 10.3389/fmicb.2018.00154

Kaczmarzyk, D., Cengic, I., Yao, L., and Hudson, E. P. (2018). Diversion of the long-chain acyl-ACP pool in *Synechocystis* to fatty alcohols through CRISPRi repression of the essential phosphate acyltransferase PlsX. *Metab. Eng.* 45, 59–66. doi: 10.1016/j.ymben.2017.11.014

Kearns, N. A., Pham, H., Tabak, B., Genga, R. M., Silverstein, N. J., Garber, M., et al. (2015). Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nat. Methods* 12, 401–403. doi: 10.1038/nmeth.3325

Kildegaard, K. R., Tramontin, L. R. R., Chekina, K., Li, M., Goedecke, T. J., Kristensen, M., et al. (2019). CRISPR/Cas9-RNA interference system for combinatorial metabolic engineering of *Saccharomyces cerevisiae*. *Yeast* 36, 237–247. doi: 10.1002/yea.3390

Kim, D., Bae, S., Park, J., Kim, E., Kim, S., Yu, H. R., et al. (2015). Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat. Methods* 12, 237–243. doi: 10.1038/nmeth.3284

Kim, Y. G., Cha, J., and Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U.S.A.* 93, 1156–1160. doi: 10.1073/pnas.93.3.1156

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., et al. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529, 490–495. doi: 10.1038/nature16526

Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Topkar, V. V., Nguyen, N. T., Zheng, Z., et al. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* 523, 481–485. doi: 10.1038/nature14592

Klompe, S. E., Vo, P. L. H., Halpin-Healy, T. S., and Sternberg, S. H. (2019). Transposon-encoded CRISPR-Cas systems direct RNA-guided DNA integration. *Nature* 571, 219–225. doi: 10.1038/s41586-019-1323-z

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi: 10.1038/nature17946

Konermann, S., Brigham, M. D., Trevino, A. E., Joung, J., Abudayyeh, O. O., Barcena, C., et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 517, 583–588. doi: 10.1038/nature14136

Koonin, E. V., Makarova, K. S., and Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.* 37, 67–78. doi: 10.1016/j.mib.2017.05.008

Kuivanen, J., Holmstrom, S., Lehtinen, B., Penttila, M., and Jantti, J. (2018). A High-throughput workflow for CRISPR/Cas9 mediated combinatorial promoter replacements and phenotype characterization in yeast. *Biotechnol. J.* e1700593. doi: 10.1002/biot.201700593 [Epub ahead of print].

Labun, K., Montague, T. G., Krause, M., Torres Cleuren, Y. N., Tjeldnes, H., and Valen, E. (2019). CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* 47, W171–W174. doi: 10.1093/nar/gkz365

Laughery, M. F., Hunter, T., Brown, A., Hoopes, J., Ostbye, T., Shumaker, T., et al. (2015). New vectors for simple and streamlined CRISPR-Cas9 genome editing in *Saccharomyces cerevisiae*. *Yeast* 32, 711–720. doi: 10.1002/yea.3098

Lee, H. H., Ostrov, N., Wong, B. G., Gold, M. A., Khalil, A. S., and Church, G. M. (2019). Functional genomics of the rapidly replicating bacterium Vibrio natriegens by CRISPRi. *Nat. Microbiol.* 4, 1105–1113. doi: 10.1038/s41564-019-0423-8

Lee, S. S., Shin, H., Jo, S., Lee, S. M., Um, Y., and Woo, H. M. (2018). Rapid identification of unknown carboxyl esterase activity in *Corynebacterium glutamicum* using RNA-guided CRISPR interference. *Enzyme Microb. Technol.* 114, 63–68. doi: 10.1016/j.enzmictec.2018.04.004

Levskaya, A., Weiner, O. D., Lim, W. A., and Voigt, C. A. (2009). Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature* 461, 997–1001. doi: 10.1038/nature08446

Li, P., Fu, X., Zhang, L., and Li, S. (2019). CRISPR/Cas-based screening of a gene activation library in *Saccharomyces cerevisiae* identifies a crucial role of OLE1 in thermotolerance. *Microb. Biotechnol.* 12, 1154–1163. doi: 10.1111/1751-7915.13333

Li, Q., Chen, J., Minton, N. P., Zhang, Y., Wen, Z., Liu, J., et al. (2016). CRISPR-based genome editing and expression control systems in Clostridium acetobutylicum and Clostridium beijerinckii. *Biotechnol. J.* 11, 961–972. doi: 10.1002/biot.201600053

Li, X., Wang, Y., Liu, Y., Yang, B., Wang, X., Wei, J., et al. (2018). Base editing with a Cpf1-cytidine deaminase fusion. *Nat. Biotechnol.* 36, 324–327. doi: 10.1038/nbt.4102

Li, Y., Yan, F., Wu, H., Li, G., Han, Y., Ma, Q., et al. (2019). Multiple-step chromosomal integration of divided segments from a large DNA fragment via CRISPR/Cas9 in *Escherichia coli*. *J. Ind. Microbiol. Biotechnol.* 46, 81–90. doi: 10.1007/s10295-018-2114-5

Lian, J., HamediRad, M., Hu, S., and Zhao, H. (2017). Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. *Nat. Commun.* 8:1688. doi: 10.1038/s41467-017-01695-x

Lian, J., Schultz, C., Cao, M., HamediRad, M., and Zhao, H. (2019). Multi-functional genome-wide CRISPR system for high throughput genotype-phenotype mapping. *Nat. Commun.* 10:5794. doi: 10.1038/s41467-019-13621-4

Lim, Y. H., Wong, F. T., Yeo, W. L., Ching, K. C., Lim, Y. W., Heng, E., et al. (2018). Auroramycin: a potent antibiotic from *Streptomyces roseosporus* by CRISPR-Cas9 activation. *Chembiochem* doi: 10.1002/cbic.201800266 [Epub ahead of print].

Lino, C. A., Harper, J. C., Carney, J. P., and Timlin, J. A. (2018). Delivering CRISPR: a review of the challenges and approaches. *Drug Deliv.* 25, 1234–1257. doi: 10.1080/10717544.2018.1474964

Liu, D., Huang, C., Guo, J., Zhang, P., Chen, T., Wang, Z., et al. (2019). Development and characterization of a CRISPR/Cas9n-based multiplex genome editing system for *Bacillus subtilis*. *Biotechnol. Biofuels* 12:197. doi: 10.1186/s13068-019-1537-1

Liu, H., Wei, Z., Dominguez, A., Li, Y., Wang, X., and Qi, L. S. (2015). CRISPR-ERA: a comprehensive design tool for CRISPR-mediated gene editing, repression and activation. *Bioinformatics* 31, 3676–3678. doi: 10.1093/bioinformatics/btv423

Liu, R., Liang, L., Choudhury, A., Bassalo, M. C., Garst, A. D., Tarasava, K., et al. (2018a). Iterative genome editing of *Escherichia coli* for 3-hydroxypropionic acid production. *Metab. Eng.* 47, 303–313. doi: 10.1016/j.ymben.2018.04.007

Liu, R., Liang, L., Garst, A. D., Choudhury, A., Nogue, V. S. I., Beckham, G. T., et al. (2018b). Directed combinatorial mutagenesis of *Escherichia coli* for complex phenotype engineering. *Metab. Eng.* 47, 10–20. doi: 10.1016/j.ymben.2018.02.007

Liu, S., Xiao, H., Zhang, F., Lu, Z., Zhang, Y., Deng, A., et al. (2019). A seamless and iterative DNA assembly method named PS-Brick and its assisted metabolic engineering for threonine and 1-propanol production. *Biotechnol. Biofuels* 12:180. doi: 10.1186/s13068-019-1520-x

Liu, W., Tang, D. D., Wang, H. J., Lian, J. Z., Huang, L., and Xu, Z. N. (2019). Combined genome editing and transcriptional repression for metabolic pathway engineering in *Corynebacterium glutamicum* using a catalytically active Cas12a. *Appl. Microbiol. Biotechnol.* 103, 8911–8922. doi: 10.1007/s00253-019-10118-4

Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., et al. (2013a). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* 31, 833–838. doi: 10.1038/nbt.2675

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., et al. (2013b). RNA-guided human genome engineering via Cas9. *Science* 339, 823–826. doi: 10.1126/science.1232033

Mans, R., van Rossum, H. M., Wijsman, M., Backx, A., Kuijpers, N. G., van den Broek, M., et al. (2015). CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 15:fov004. doi: 10.1093/femsyr/fov004

Maruyama, T., Dougan, S. K., Truttmann, M. C., Bilate, A. M., Ingram, J. R., and Ploegh, H. L. (2015). Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* 33, 538–542. doi: 10.1038/nbt.3190

Miao, C., Zhao, H., Qian, L., and Lou, C. (2019). Systematically investigating the key features of the DNase deactivated Cpf1 for tunable transcription regulation in prokaryotic cells. *Synth. Syst. Biotechnol.* 4, 1–9. doi: 10.1016/j.synbio.2018.11.002

Mitsui, R., Yamada, R., and Ogino, H. (2019). Improved stress tolerance of *Saccharomyces cerevisiae* by CRISPR-Cas-mediated genome evolution. *Appl. Biochem. Biotechnol.* 189, 810–821. doi: 10.1007/s12010-019-03040-y

Miura, H., Gurumurthy, C. B., Sato, T., Sato, M., and Ohtsuka, M. (2015). CRISPR/Cas9-based generation of knockdown mice by intronic insertion of artificial microRNA using longer single-stranded DNA. *Sci. Rep.* 5:12799. doi: 10.1038/srep12799

Mojica, F. J. M., Diez-Villasenor, C., Garcia-Martinez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155, 733–740. doi: 10.1099/mic.0.023960-0

Moon, S. B., Kim, D. Y., Ko, J. H., and Kim, Y. S. (2019). Recent advances in the CRISPR genome editing tool set. *Exp. Mol. Med.* 51, 1–11. doi: 10.1038/s12276-019-0339-7

Mougiakos, I., Bosma, E. F., de Vos, W. M., van Kranenburg, R., and van der Oost, J. (2016). Next generation prokaryotic engineering: the CRISPR-Cas Toolkit. *Trends Biotechnol.* 34, 575–587. doi: 10.1016/j.tibtech.2016.02.004

Mougiakos, I., Bosma, E. F., Ganguly, J., van der Oost, J., and van Kranenburg, R. (2018). Hijacking CRISPR-Cas for high-throughput bacterial metabolic engineering: advances and prospects. *Curr. Opin. Biotechnol.* 50, 146–157. doi: 10.1016/j.copbio.2018.01.002

Naito, Y., Hino, K., Bono, H., and Ui-Tei, K. (2015). CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics* 31, 1120–1123. doi: 10.1093/bioinformatics/btu743

Naseri, G., Behrend, J., Rieper, L., and Mueller-Roeber, B. (2019). COMPASS for rapid combinatorial optimization of biochemical pathways based on artificial transcription factors. *Nat. Commun.* 10:2615. doi: 10.1038/s41467-019-10224-x

Ng, I. S., Keskin, B. B., and Tan, S. I. (2020). A critical review of genome editing and synthetic biology applications in metabolic engineering of microalgae and cyanobacteria. *Biotechnol. J.* e1900228. doi: 10.1002/biot.201900228

Ni, J., Zhang, G., Qin, L., Li, J., and Li, C. (2019). Simultaneously down-regulation of multiplex branch pathways using CRISPRi and fermentation optimization for enhancing beta-amyrin production in *Saccharomyces cerevisiae*. *Synth. Syst. Biotechnol.* 4, 79–85. doi: 10.1016/j.synbio.2019.02.002

Nihongaki, Y., Furuhata, Y., Otabe, T., Hasegawa, S., Yoshimoto, K., and Sato, M. (2017). CRISPR-Cas9-based photoactivatable transcription systems to induce neuronal differentiation. *Nat. Methods* 14, 963–966. doi: 10.1038/nmeth.4430

Nihongaki, Y., Kawano, F., Nakajima, T., and Sato, M. (2015a). Photoactivatable CRISPR-Cas9 for optogenetic genome editing. *Nat. Biotechnol.* 33, 755–760. doi: 10.1038/nbt.3245

Nihongaki, Y., Yamamoto, S., Kawano, F., Suzuki, H., and Sato, M. (2015b). CRISPR-Cas9-based photoactivatable transcription system. *Chem. Biol.* 22, 169–174. doi: 10.1016/j.chembiol.2014.12.011

Nissim, L., Perli, S. D., Fridkin, A., Perez-Pinera, P., and Lu, T. K. (2014). Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells. *Mol. Cell* 54, 698–710. doi: 10.1016/j.molcel.2014.04.022

Nowak, C. M., Lawson, S., Zerez, M., and Bleris, L. (2016). Guide RNA engineering for versatile Cas9 functionality. *Nucleic Acids Res.* 44, 9555–9564. doi: 10.1093/nar/gkw908

O'Geen, H., Yu, A. S., and Segal, D. J. (2015). How specific is CRISPR/Cas9 really? *Curr. Opin. Chem. Biol.* 29, 72–78. doi: 10.1016/j.cbpa.2015.10.001

Oh, J. H., and van Pijkeren, J. P. (2014). CRISPR-Cas9-assisted recombineering in *Lactobacillus reuteri*. *Nucleic Acids Res.* 42:e131. doi: 10.1093/nar/gku623

Palazzotto, E., Tong, Y., Lee, S. Y., and Weber, T. (2019). Synthetic biology and metabolic engineering of actinomycetes for natural product discovery. *Biotechnol. Adv.* 37:107366. doi: 10.1016/j.biotechadv.2019.03.005

Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., et al. (2016). Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 533, 125–129. doi: 10.1038/nature17664

Peters, J. M., Colavin, A., Shi, H., Czarny, T. L., Larson, M. H., Wong, S., et al. (2016). A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell* 165, 1493–1506. doi: 10.1016/j.cell.2016.05.003

Polstein, L. R., and Gersbach, C. A. (2015). A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol.* 11, 198–200. doi: 10.1038/nchembio.1753

Port, F., and Bullock, S. L. (2016). Augmenting CRISPR applications in *Drosophila* with tRNA-flanked sgRNAs. *Nat. Methods* 13, 852–854. doi: 10.1038/nmeth.3972

Pyne, M. E., Moo-Young, M., Chung, D. A., and Chou, C. P. (2015). Coupling the CRISPR/Cas9 system with Lambda Red recombineering enables simplified chromosomal gene replacement in *Escherichia coli*. *Appl. Environ. Microbiol.* 81, 5103–5114. doi: 10.1128/AEM.01248-15

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., et al. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173–1183. doi: 10.1016/j.cell.2013.02.022

Qi, W., Zhu, T., Tian, Z., Li, C., Zhang, W., and Song, R. (2016). High-efficiency CRISPR/Cas9 multiplex gene editing using the glycine tRNA-processing system-based strategy in maize. *BMC Biotechnol.* 16:58. doi: 10.1186/s12896-016-0289-2

Qin, Q., Ling, C., Zhao, Y., Yang, T., Yin, J., Guo, Y., et al. (2018). CRISPR/Cas9 editing genome of extremophile *Halomonas* spp. *Metab. Eng.* 47, 219–229. doi: 10.1016/j.ymben.2018.03.018

Ran, F. A., Cong, L., Yan, W. X., Scott, D. A., Gootenberg, J. S., Kriz, A. J., et al. (2015). In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* 520, 186–191. doi: 10.1038/nature14299

Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., et al. (2013a). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380–1389. doi: 10.1016/j.cell.2013.08.021

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., and Zhang, F. (2013b). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281–2308. doi: 10.1038/nprot.2013.143

Ranjha, L., Howard, S. M., and Cejka, P. (2018). Main steps in DNA double-strand break repair: an introduction to homologous recombination and related processes. *Chromosoma* 127, 187–214.

Raschmanova, H., Weninger, A., Glieder, A., Kovar, K., and Vogl, T. (2018). Implementing CRISPR-Cas technologies in conventional and non-conventional yeasts: current state and future prospects. *Biotechnol. Adv.* 36, 641–665. doi: 10.1016/j.biotechadv.2018.01.006

Reider Apel, A., d'Espaux, L., Wehrs, M., Sachs, D., Li, R. A., Tong, G. J., et al. (2017). A Cas9-based toolkit to program gene expression in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 45, 496–508. doi: 10.1093/nar/gkw1023

Reis, A. C., Halper, S. M., Vezeau, G. E., Cetnar, D. P., Hossain, A., Clauer, P. R., et al. (2019). Simultaneous repression of multiple bacterial genes using nonrepetitive extra-long sgRNA arrays. *Nat. Biotechnol.* 37, 1294–1301. doi: 10.1038/s41587-019-0286-9

Robert, F., Barbeau, M., Ethier, S., Dostie, J., and Pelletier, J. (2015). Pharmacological inhibition of DNA-PK stimulates Cas9-mediated genome editing. *Genome Med.* 7:93. doi: 10.1186/s13073-015-0215-6

Ronda, C., Pedersen, L. E., Sommer, M. O., and Nielsen, A. T. (2016). CRMAGE: CRISPR optimized MAGE recombineering. *Sci. Rep.* 6:19452. doi: 10.1038/srep19452

Ryan, O. W., and Cate, J. H. D. (2014). Multiplex engineering of industrial yeast genomes using CRISPRm. *Methods Enzymol.* 546, 473–489. doi: 10.1016/B978-0-12-801185-0.00023-4

Ryan, O. W., Skerker, J. M., Maurer, M. J., Li, X., Tsai, J. C., Poddar, S., et al. (2014). Selection of chromosomal DNA libraries using a multiplex CRISPR system. *eLife* 3:e03703. doi: 10.7554/eLife.03703

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

Schultz, J. C., Cao, M. F., and Zhao, H. M. (2019). Development of a CRISPR/Cas9 system for high efficiency multiplexed gene deletion in *Rhodosporidium toruloides*. *Biotechnol. Bioeng.* 116, 2103–2109. doi: 10.1002/bit.27001

Schwartz, C., Frogue, K., Ramesh, A., Misa, J., and Wheeldon, I. (2017). CRISPRi repression of nonhomologous end-joining for enhanced genome engineering via homologous recombination in *Yarrowia lipolytica*. *Biotechnol. Bioeng.* 114, 2896–2906. doi: 10.1002/bit.26404

Schwartz, C. M., Hussain, M. S., Blenner, M., and Wheeldon, I. (2016). Synthetic RNA polymerase III promoters facilitate high-efficiency CRISPR-Cas9-mediated genome editing in *Yarrowia lipolytica*. *ACS Synth. Biol.* 5, 356–359. doi: 10.1021/acssynbio.5b00162

Shabestary, K., Anfelt, J., Ljungqvist, E., Jahn, M., Yao, L., and Hudson, E. P. (2018). Targeted repression of essential genes to arrest growth and increase carbon partitioning and biofuel titers in *Cyanobacteria*. *ACS Synth. Biol.* 7, 1669–1675. doi: 10.1021/acssynbio.8b00056

Shen, B., Zhang, W., Zhang, J., Zhou, J., Wang, J., Chen, L., et al. (2014). Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat. Methods* 11, 399–402. doi: 10.1038/nmeth.2857

Shi, S., Liang, Y., Ang, E. L., and Zhao, H. (2019). Delta integration CRISPR-Cas (Di-CRISPR) in *Saccharomyces cerevisiae*. *Methods Mol. Biol.* 1927, 73–91. doi: 10.1007/978-1-4939-9142-6_6

Shi, S., Liang, Y., Zhang, M. M., Ang, E. L., and Zhao, H. (2016). A highly efficient single-step, markerless strategy for multi-copy chromosomal integration of large biochemical pathways in *Saccharomyces cerevisiae*. *Metab. Eng.* 33, 19–27. doi: 10.1016/j.ymben.2015.10.011

Shmakov, S., Abudayyeh, O. O., Makarova, K. S., Wolf, Y. I., Gootenberg, J. S., Semenova, E., et al. (2015). Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. *Mol. Cell* 60, 385–397. doi: 10.1016/j.molcel.2015.10.008

Shmakov, S., Smargon, A., Scott, D., Cox, D., Pyzocha, N., Yan, W., et al. (2017). Diversity and evolution of class 2 CRISPR-Cas systems. *Nat. Rev. Microbiol.* 15, 169–182. doi: 10.1038/nrmicro.2016.184

Shuman, S., and Glickman, M. S. (2007). Bacterial DNA repair by non-homologous end joining. *Nat. Rev. Microbiol.* 5, 852–861. doi: 10.1038/nrmicro1768

Singh, P., Schimenti, J. C., and Bolcun-Filas, E. (2015). A mouse geneticist's practical guide to CRISPR applications. *Genetics* 199, 1–15. doi: 10.1534/genetics.114.169771

Slaymaker, I. M., Gao, L., Zetsche, B., Scott, D. A., Yan, W. X., and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science* 351, 84–88. doi: 10.1126/science.aad5227

Smargon, A. A., Shi, Y. J., and Yeo, G. W. (2020). RNA-targeting CRISPR systems from metagenomic discovery to transcriptomic engineering. *Nat. Cell Biol.* 22, 143–150. doi: 10.1038/s41556-019-0454-7

Smith, J. D., Suresh, S., Schlecht, U., Wu, M. H., Wagih, O., Peltz, G., et al. (2016). Quantitative CRISPR interference screens in yeast identify chemical-genetic interactions and new rules for guide RNA design. *Genome Biol.* 17:45. doi: 10.1186/s13059-016-0900-9

Song, J., Yang, D., Xu, J., Zhu, T., Chen, Y. E., and Zhang, J. (2016). RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat. Commun.* 7:10548. doi: 10.1038/ncomms10548

Sorek, R., Lawrence, C. M., and Wiedenheft, B. (2013). CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu. Rev. Biochem.* 82, 237–266. doi: 10.1146/annurev-biochem-072911-172315

Standage-Beier, K., Zhang, Q., and Wang, X. (2015). Targeted large-scale deletion of bacterial genomes using CRISPR-nickases. *ACS Synth. Biol.* 4, 1217–1225. doi: 10.1021/acssynbio.5b00132

Stemmer, M., Thumberger, T., Del Sol Keyer, M., Wittbrodt, J., and Mateo, J. L. (2015). CCTop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. *PLoS One* 10:e0124633. doi: 10.1371/journal.pone.0124633

Stovicek, V., Holkenbrink, C., and Borodina, I. (2017). CRISPR/Cas system for yeast genome engineering: advances and applications. *FEMS Yeast Res.* 17:fox030. doi: 10.1093/femsyr/fox030

Strecker, J., Ladha, A., Gardner, Z., Schmid-Burgk, J. L., Makarova, K. S., Koonin, E. V., et al. (2019). RNA-guided DNA insertion with CRISPR-associated transposases. *Science* 365, 48–53. doi: 10.1126/science.aax9181

Su, T., Liu, F., Chang, Y., Guo, Q., Wang, J., Wang, Q., et al. (2019). The phage T4 DNA ligase mediates bacterial chromosome DSBs repair as single component non-homologous end joining. *Synth. Syst. Biotechnol.* 4, 107–112. doi: 10.1016/j.synbio.2019.04.001

Sung, L. Y., Wu, M. Y., Lin, M. W., Hsu, M. N., Truong, V. A., Shen, C. C., et al. (2019). Combining orthogonal CRISPR and CRISPRi systems for genome engineering and metabolic pathway modulation in *Escherichia coli*. *Biotechnol. Bioeng.* 116, 1066–1079. doi: 10.1002/bit.26915

Tan, S. Z., Reisch, C. R., and Prather, K. L. J. (2018). A robust CRISPR interference gene repression system in *Pseudomonas*. *J. Bacteriol.* 200:e00575-17. doi: 10.1128/JB.00575-17

Tan, Z., Yoon, J. M., Chowdhury, A., Burdick, K., Jarboe, L. R., Maranas, C. D., et al. (2018). Engineering of *E. coli* inherent fatty acid biosynthesis capacity to increase octanoic acid production. *Biotechnol Biofuels* 11:87. doi: 10.1186/s13068-018-1078-z

Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S., and Vale, R. D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* 159, 635–646. doi: 10.1016/j.cell.2014.09.039

Tang, Y., and Fu, Y. (2018). Class 2 CRISPR/Cas: an expanding biotechnology toolbox for and beyond genome editing. *Cell Biosci.* 8:59. doi: 10.1186/s13578-018-0255-x

Tarasava, K., Oh, E. J., Eckert, C. A., and Gill, R. T. (2018). CRISPR-enabled tools for engineering microbial genomes and phenotypes. *Biotechnol. J.* 13:e1700586. doi: 10.1002/biot.201700586

Tian, T., Kang, J. W., Kang, A., and Lee, T. S. (2019). Redirecting metabolic flux via combinatorial multiplex CRISPRi-mediated repression for isopentenol production in *Escherichia coli*. *ACS Synth. Biol.* 8, 391–402. doi: 10.1021/acssynbio.8b00429

Tong, Y., Whitford, C. M., Robertsen, H. L., Blin, K., Jorgensen, T. S., Klitgaard, A. K., et al. (2019). Highly efficient DSB-free base editing for streptomycetes with CRISPR-BEST. *Proc. Natl. Acad. Sci. U.S.A.* 116, 20366–20375. doi: 10.1073/pnas.1913493116

Tran, V. G., Cao, M., Fatma, Z., Song, X., and Zhao, H. (2019). Development of a CRISPR/Cas9-based tool for gene deletion in *Issatchenkia orientalis*. *mSphere* 4:e00345-19. doi: 10.1128/mSphere.0034519

Tsai, S. Q., Wyvekens, N., Khayter, C., Foden, J. A., Thapar, V., Reyon, D., et al. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* 32, 569–576. doi: 10.1038/nbt.2908

Ungerer, J., and Pakrasi, H. B. (2016). Cpf1 is a versatile tool for CRISPR genome editing across diverse species of *Cyanobacteria*. *Sci. Rep.* 6:39681. doi: 10.1038/srep39681

Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., and Gregory, P. D. (2010). Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* 11, 636–646. doi: 10.1038/nrg2842

Vanegas, K. G., Lehka, B. J., and Mortensen, U. H. (2017). SWITCH: a dynamic CRISPR tool for genome engineering and metabolic pathway control for cell factory construction in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* 16:25. doi: 10.1186/s12934-017-0632-x

Wang, B., Hu, Q., Zhang, Y., Shi, R., Chai, X., Liu, Z., et al. (2018). A RecET-assisted CRISPR-Cas9 genome editing in *Corynebacterium glutamicum*. *Microb. Cell Fact.* 17:63. doi: 10.1186/s12934-018-0910-2

Wang, H., La Russa, M., and Qi, L. S. (2016). CRISPR/Cas9 in genome editing and beyond. *Annu. Rev. Biochem.* 85, 227–264. doi: 10.1146/annurev-biochem-060815-014607

Wang, J., Zhao, P., Li, Y., Xu, L., and Tian, P. (2018). Engineering CRISPR interference system in *Klebsiella pneumoniae* for attenuating lactic acid synthesis. *Microb. Cell Fact.* 17:56. doi: 10.1186/s12934-018-0903-1

Wang, Q., and Coleman, J. J. (2019). Progress and Challenges: development and implementation of CRISPR/cas9 technology in filamentous fungi. *Comput. Struct. Biotechnol. J.* 17, 761–769. doi: 10.1016/j.csbj.2019.06.007

Wang, T., Guan, C., Guo, J., Liu, B., Wu, Y., Xie, Z., et al. (2018). Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance. *Nat. Commun.* 9:2475. doi: 10.1038/s41467-018-04899-x

Wang, Y., Liu, Y., Li, J., Yang, Y., Ni, X., Cheng, H., et al. (2019). Expanding targeting scope, editing window, and base transition capability of base editing in *Corynebacterium glutamicum*. *Biotechnol. Bioeng.* 116, 3016–3029. doi: 10.1002/bit.27121

Wang, Y., Liu, Y., Liu, J., Guo, Y., Fan, L., Ni, X., et al. (2018). MACBETH: multiplex automated *Corynebacterium glutamicum* base editing method. *Metab. Eng.* 47, 200–210. doi: 10.1016/j.ymben.2018.02.016

Wang, Y., Zhang, Z. T., Seo, S. O., Lynn, P., Lu, T., Jin, Y. S., et al. (2016). Bacterial genome editing with CRISPR-Cas9: deletion, integration, single nucleotide modification, and desirable "clean" mutant selection in *Clostridium beijerinckii* as an example. *ACS Synth. Biol.* 5, 721–732. doi: 10.1021/acssynbio.6b00060

Weller, G. R., Kysela, B., Roy, R., Tonkin, L. M., Scanlan, E., Della, M., et al. (2002). Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* 297, 1686–1689. doi: 10.1126/science.1074584

Wendt, K. E., Ungerer, J., Cobb, R. E., Zhao, H., and Pakrasi, H. B. (2016). CRISPR/Cas9 mediated targeted mutagenesis of the fast growing cyanobacterium *Synechococcus elongatus* UTEX 2973. *Microb. Cell Fact.* 15:115. doi: 10.1186/s12934-016-0514-7

Weninger, A., Hatzl, A. M., Schmid, C., Vogl, T., and Glieder, A. (2016). Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast *Pichia pastoris*. *J. Biotechnol.* 235, 139–149. doi: 10.1016/j.jbiotec.2016.03.027

Wensing, L., Sharma, J., Uthayakumar, D., Proteau, Y., Chavez, A., and Shapiro, R. S. (2019). A CRISPR interference platform for efficient genetic repression in *Candida albicans*. *mSphere* 4:e00002-19. doi: 10.1128/mSphere.00002-19

Westbrook, A. W., Moo-Young, M., and Chou, C. P. (2016). Development of a CRISPR-Cas9 tool kit for comprehensive engineering of *Bacillus subtilis*. *Appl. Environ. Microbiol.* 82, 4876–4895. doi: 10.1128/AEM.01159-16

Westbrook, A. W., Ren, X., Oh, J., Moo-Young, M., and Chou, C. P. (2018). Metabolic engineering to enhance heterologous production of hyaluronic acid in *Bacillus subtilis*. *Metab. Eng.* 47, 401–413. doi: 10.1016/j.ymben.2018.04.016

Woolston, B. M., Emerson, D. F., Currie, D. H., and Stephanopoulos, G. (2018). Rediverting carbon flux in *Clostridium ljungdahlii* using CRISPR interference (CRISPRi). *Metab. Eng.* 48, 243–253. doi: 10.1016/j.ymben.2018.06.006

Wu, H., Li, Y., Ma, Q., Li, Q., Jia, Z., Yang, B., et al. (2018a). Metabolic engineering of *Escherichia coli* for high-yield uridine production. *Metab. Eng.* 49, 248–256. doi: 10.1016/j.ymben.2018.09.001

Wu, Y., Chen, T., Liu, Y., Lv, X., Li, J., Du, G., et al. (2018b). CRISPRi allows optimal temporal control of N-acetylglucosamine bioproduction by a dynamic coordination of glucose and xylose metabolism in *Bacillus subtilis*. *Metab. Eng.* 49, 232–241. doi: 10.1016/j.ymben.2018.08.012

Wyvekens, N., Topkar, V. V., Khayter, C., Joung, J. K., and Tsai, S. Q. (2015). Dimeric CRISPR RNA-guided FokI-dCas9 nucleases directed by truncated grnas for highly specific genome editing. *Hum. Gene Ther.* 26, 425–431. doi: 10.1089/hum.2015.084

Xie, K., Minkenberg, B., and Yang, Y. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3570–3575. doi: 10.1073/pnas.1420294112

Xu, X. S., and Oi, L. S. (2019). A CRISPR-dCas toolbox for genetic engineering and synthetic biology. *J. Mol. Biol.* 431, 34–47. doi: 10.1016/j.jmb.2018.06.037

Yan, Q., and Fong, S. S. (2017). Challenges and advances for genetic engineering of non-model bacteria and uses in consolidated bioprocessing. *Front. Microbiol.* 8:2060. doi: 10.3389/fmicb.2017.02060

Yang, Z., Edwards, H., and Xu, P. (2020). CRISPR-Cas12a/Cpf1-assisted precise, efficient and multiplexed genome-editing in *Yarrowia lipolytica*. *Metab. Eng. Commun.* 10:e00112. doi: 10.1016/j.mec.2019.e00112

Yang, Z., Wang, H., Wang, Y., Ren, Y., and Wei, D. (2018). Manufacturing multienzymatic complex reactors in vivo by self-assembly to improve the biosynthesis of itaconic acid in *Escherichia coli*. *ACS Synth. Biol.* 7, 1244–1250. doi: 10.1021/acssynbio.8b00086

Yeo, W. L., Heng, E., Tan, L. L., Lim, Y. W., Lim, Y. H., Hoon, S., et al. (2019). Characterization of Cas proteins for CRISPR-Cas editing in streptomycetes. *Biotechnol. Bioeng.* 116, 2330–2338. doi: 10.1002/bit.27021

Zalatan, J. G., Lee, M. E., Almeida, R., Gilbert, L. A., Whitehead, E. H., La Russa, M., et al. (2015). Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* 160, 339–350. doi: 10.1016/j.cell.2014.11.052

Zerbini, F., Zanella, I., Fraccascia, D., Konig, E., Irene, C., Frattini, L. F., et al. (2017). Large scale validation of an efficient CRISPR/Cas-based multi gene editing protocol in *Escherichia coli*. *Microb. Cell Fact.* 16:68. doi: 10.1186/s12934-017-0681-1

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., et al. (2015a). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163, 759–771. doi: 10.1016/j.cell.2015.09.038

Zetsche, B., Volz, S. E., and Zhang, F. (2015b). A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nat. Biotechnol.* 33, 139–142. doi: 10.1038/nbt.3149

Zhang, F., Wen, Y., and Guo, X. (2014). CRISPR/Cas9 for genome editing: progress, implications and challenges. *Hum. Mol. Genet.* 23, R40–R46. doi: 10.1093/hmg/ddu125

Zhang, J., Zong, W., Hong, W., Zhang, Z. T., and Wang, Y. (2018). Exploiting endogenous CRISPR-Cas system for multiplex genome editing in *Clostridium tyrobutyricum* and engineer the strain for high-level butanol production. *Metab. Eng.* 47, 49–59. doi: 10.1016/j.ymben.2018.03.007

Zhang, M. M., Wong, F. T., Wang, Y., Luo, S., Lim, Y. H., Heng, E., et al. (2017). CRISPR-Cas9 strategy for activation of silent *Streptomyces* biosynthetic gene clusters. *Nat. Chem. Biol.* 13, 607–609. doi: 10.1038/nchembio.2341

Zhang, Y., Wang, J., Wang, Z., Zhang, Y., Shi, S., Nielsen, J., et al. (2019). A gRNA-tRNA array for CRISPR-Cas9 based rapid multiplexed genome editing in *Saccharomyces cerevisiae*. *Nat. Commun.* 10:1053. doi: 10.1038/s41467-019-09005-3

Zhao, Y. W., Tian, J. Z., Zheng, G. S., Chen, J., Sun, C. W., Yang, Z. Y., et al. (2019). Multiplex genome editing using a dCas9-cytidine deaminase fusion in *Streptomyces*. *Sci. China Life Sci.* doi: 10.1007/s11427-019-1559-y [Epub ahead of print].

Zheng, Y., Han, J., Wang, B., Hu, X., Li, R., Shen, W., et al. (2019). Characterization and repurposing of the endogenous Type I-F CRISPR-Cas system of *Zymomonas mobilis* for genome engineering. *Nucleic Acids Res.* 47, 11461–11475. doi: 10.1093/nar/gkz940

Zhou, H. B., Liu, J. L., Zhou, C. Y., Gao, N., Rao, Z. P., Li, H., et al. (2018). In vivo simultaneous transcriptional activation of multiple genes in the brain using CRISPR-dCas9-activator transgenic mice. *Nat. Neurosci.* 21, 440–446. doi: 10.1038/s41593-017-0060-6

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

BRIEF RESEARCH REPORT
published: 30 June 2020
doi: 10.3389/fbioe.2020.00653

Construction of a Stable and Temperature-Responsive Yeast Cell Factory for Crocetin Biosynthesis Using CRISPR-Cas9

Tengfei Liu[1,2], *Chang Dong*[1,2], *Mingming Qi*[2,3], *Bei Zhang*[2], *Lei Huang*[1], *Zhinan Xu*[1] *and Jiazhang Lian*[1,2]*

[1] Key Laboratory of Biomass Chemical Engineering of Ministry of Education, College of Chemical and Biological Engineering, Zhejiang University, Hangzhou, China, [2] Center for Synthetic Biology, College of Chemical and Biological Engineering, Zhejiang University, Hangzhou, China, [3] School of Bioengineering, Dalian University of Technology, Dalian, China

OPEN ACCESS

Edited by:
Yuan Lu,
Tsinghua University, China

Reviewed by:
Jing Fu,
Chalmers University of Technology, Sweden
Dae-Hee Lee,
Korea Research Institute of Bioscience and Biotechnology (KRIBB), South Korea
Fengxue Xin,
Nanjing Tech University, China
Patrick K. H. Lee,
City University of Hong Kong, Hong Kong
Dennis Dienst,
Uppsala University, Sweden

***Correspondence:**
Jiazhang Lian
jzlian@zju.edu.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: 13 March 2020
Accepted: 27 May 2020
Published: 30 June 2020

Citation:
Liu T, Dong C, Qi M, Zhang B, Huang L, Xu Z and Lian J (2020) Construction of a Stable and Temperature-Responsive Yeast Cell Factory for Crocetin Biosynthesis Using CRISPR-Cas9. Front. Bioeng. Biotechnol. 8:653. doi: 10.3389/fbioe.2020.00653

Crocetin is a plant natural product with broad medicinal applications, such as improvement of sleep quality and attenuation of physical fatigue. However, crocetin production using microbial cell factories is still far from satisfaction, probably due to the conflict between cell growth and product accumulation. In the present work, a temperature-responsive crocetin-producing *Saccharomyces cerevisiae* strain was established to coordinate cell growth, precursor (zeaxanthin) generation, and product (crocetin) biosynthesis. The production of crocetin was further enhanced via increasing the copy numbers of *CCD2* and *ALDH* genes using the CRISPR-Cas9 based multiplex genome integration technology. The final engineered strain TL009 produced crocetin up to 139.67 ± 2.24 μg/g DCW. The advantage of the temperature switch based crocetin production was particularly demonstrated by much higher zeaxanthin conversion yield. This study highlights the potential of the temperature-responsive yeast platform strains to increase the production of other valuable carotenoid derivatives.

Keywords: crocetin, temperature switch, copy number, genome integration, CRISPR-Cas9

INTRODUCTION

Crocetin ($C_{20}H_{24}O_4$) has been found in the stigmas of *Crocus sativus* L. and the fruit of *Gardenia jasminoides* (Sheu and Hsin, 1998; Frusciante et al., 2014) and contributes to the most important therapeutic effects of saffron (Hashemi and Hosseinzadeh, 2019). Crocetin has different pharmacological effects on a large number of cancer cells: liver, ovarian, breast, prostate, leukemia, colorectal, bladder, lung, tongue carcinoma, and esophageal (Colapietro et al., 2019). The retail price of the red stigmas of *C. sativus* ranges from 2,000 to 7,000 £/kg, because 1 kg of dry saffron requires the manual harvest of around 110,000–170,000 flowers (Frusciante et al., 2014). As an apocarotenoid, crocetin is isolated from the saffron stigmas and large-scale plantation of saffron crocus is required for commercial applications. Alternatively, microbial production of carotenoids and their derivatives have been demonstrated as a promising solution (Niu et al., 2017; Wang C. et al., 2019). Therefore, *de novo* biosynthesis of crocetin from carbohydrates using microbial cell factories would be a more sustainable and economic way.

The biosynthesis of crocetin in *C. sativus* stigmas starting from β-carotene contains three major steps, catalyzed by a β-carotene hydroxylase (CrtZ), a carotenoid-cleaving dioxygenase (CCD2),

and an aldehyde dehydrogenase (ALDH), respectively (**Figure 1A**; Frusciante et al., 2014). The introduction of these three genes together with the carotenogenic genes enabled the production of crocetin in engineered *Saccharomyces cerevisiae* (Chai et al., 2017; Tan et al., 2019) and *Escherichia coli* (Wang W. et al., 2019) strains. Chai et al. found that the production of crocetin was much higher at lower temperature than at 30°C (the optimal temperature for cell growth), probably due to the higher enzymatic activity of CCD2 from *C. sativus* L (Ahrazem et al., 2016). To address the dilemma between cell growth and product formation, a general strategy is to perform two-stage fermentation, shifting culture temperature from 30°C to a lower level (i.e., 25 or 20°C) when the cell density reaches to a relatively high level. Although high-level production of crocetin was achieved using such a strategy, the conversion yield of zeaxanthin to crocetin remained at a low level. As a lipophilic and water-insoluble compound (Murill et al., 2019), zeaxanthin was synthesized and stored in the cellular membranes of microorganisms (Shen et al., 2016; Sun et al., 2018). In this case, due to low CCD2 enzymatic activity at 30°C, zeaxanthin was synthesized at high efficiency and mainly accumulated in the cellular membranes in *S. cerevisiae* (Chai et al., 2017). Although the activity of CCD2 was significantly enhanced by shifting to low temperature, the physical separation of the enzyme (present in the cytoplasm) and the substrate (storage in the cellular membranes) in two compartments led to low zeaxanthin cleavage and conversion efficiency (**Figure 1B**). Therefore, how to balance zeaxanthin accumulation and temperature-regulated CCD2 activity became a key question to enhance zeaxanthin conversion and crocetin production.

The use of fermentation temperature as a general input signal for gene expression regulation has a number of advantages including ready controllability, fast temporal response, high reversibility, and wide applicability (Chakshusmathi et al., 2004). Therefore, the construction of a temperature-responsive cell factory is a promising approach for crocetin production. In previous studies, a few temperature-induced protein expression systems have been established in yeast, such as those based on the mutation of the acid phosphatase regulatory genes *PHO80* and *PHO4*^{ts} (Kramer et al., 1984), mating type control involving *SIR3* mutation and *MATα*2-hybrid promoters (Sledziewski et al., 1990), as well as the modified GAL regulation system (Xie et al., 2014; Zhou et al., 2018). Among these nicely designed systems, the GAL regulon based system demonstrated the advantages of easy manipulation (single point mutation of the GAL4 activator) and high expression level of heterologous genes (>1,000-fold induction). The GAL based temperature switch was established by knocking out *GAL80* encoding the *GAL4* inhibitor and replacing the wild-type *GAL4* with the temperature-sensitive mutant *GAL4M9*. The evolved GAL4M9 (a single point mutation of GAL4) enabled the expression of the GAL regulon to be only turned on at lower temperatures, with 24°C determined to be the optimal induction temperature (Zhou et al., 2018). The application of the temperature switch was demonstrated by the production of lycopene and astaxanthin with temperature as an input signal for metabolic pathway regulation in yeast cell factories (Zhou et al., 2018, 2019).

In the present study, the temperature-responsive yeast cell factory was evaluated for the application in the biosynthesis of crocetin, by coordinately regulated the biosynthesis and cleavage of zeaxanthin in a temperature-dependent manner. Firstly, the temperature-responsive yeast strain was reconstructed by knocking out *GAL4* and *GAL80*, followed by the introduction of *GAL4M9* expression cassette (**Figure S1**). Then the crocetin biosynthetic pathway genes under the control of GAL promoters were integrated into the chromosome, using the CRISPR-Cas9 technology (Lian et al., 2018a,b), to create a stable and temperature-responsive yeast strain for crocetin production. Thanks to the high efficiency of genome integration, crocetin biosynthesis was further optimized by integrating different copy numbers of *CCD2* and *ALDH* genes. Finally, a temperature-responsive and stable yeast strain was established to produce crocetin with a titer of 139.67 ± 2.24 μg/g DCW and a zeaxanthin conversion yield of up to 77%.

MATERIALS AND METHODS

Strains, Media, and Chemicals

E. coli Trans T1 (Transgen Biotech Co., Ltd., Beijing, China) was use for gene cloning and plasmid amplification. Recombinant *E. coli* strains were cultured in LB medium (OXIOD Biotech Co., Ltd., London, England) supplemented with 100 mg/L ampicillin. *S. cerevisiae* BY4741 strain was used as the host for genome engineering and crocetin production. Yeast strains were routinely cultivated in YPD medium (OXIOD). Synthetic complete medium (SCD) containing 5 g/L ammonium sulfate, 1.7 g/L yeast nitrogen base without ammonium and amino acids (BD Diagnostics), 0.6 g/L CMS missing the appropriate nutrients, and 20 g/L glucose. When necessary, 200 mg/L G418 sulfate (Sangon Bio-tech Co., Ltd., Shanghai, China) was supplemented. All chemicals were bought from Sigma-Aldrich (St. Louis Missouri, USA), unless specifically mentioned. Crocetin (**Figure S2A**) and zeaxanthin (**Figure S2B**) standards were purchased from Yuanye Bio-tech Co., Ltd. (Shanghai, China) and Chemface Bio-tech Co., Ltd. (Wuhan, China), respectively.

Plasmid and Strain Construction

KOD-Plus-Neo DNA Polymerase (TOYOBO Biotech Co., Ltd., Tokyo, Japan) was used for gene amplification and PCR products were purified by the Gene JET PCR Purification Kit (Thermofisher Scientific, Shanghai, China). Restriction enzymes and T4 DNA ligase were purchased from NEB (Beijing, China). Plasmids were extracted from *E. coli* using the AxyPrep Plasmid Miniprep Kit (Axygen) according to manufacturer's instructions. DNA sequencing was performed by Tsingke Biotech Co., Ltd. (Hangzhou, China).

CrtI (accession number: Y15007.1), *CrtYB* (accession number: KJ783314.1), and *CrtE* (accession number: DQ016502.1) were amplified from the genomic DNA of the yeast strain CEN-Crt and *CrtZ* (accession number: D90087.2) was amplified from pRS426-Zea (Lian et al., 2016, 2017). CCD2 from *C. sativus* L. (*CsCCD2*; accession number: KJ541749.1) and *ALDH*s (Trautmann et al., 2013; Costantina et al., 2018)

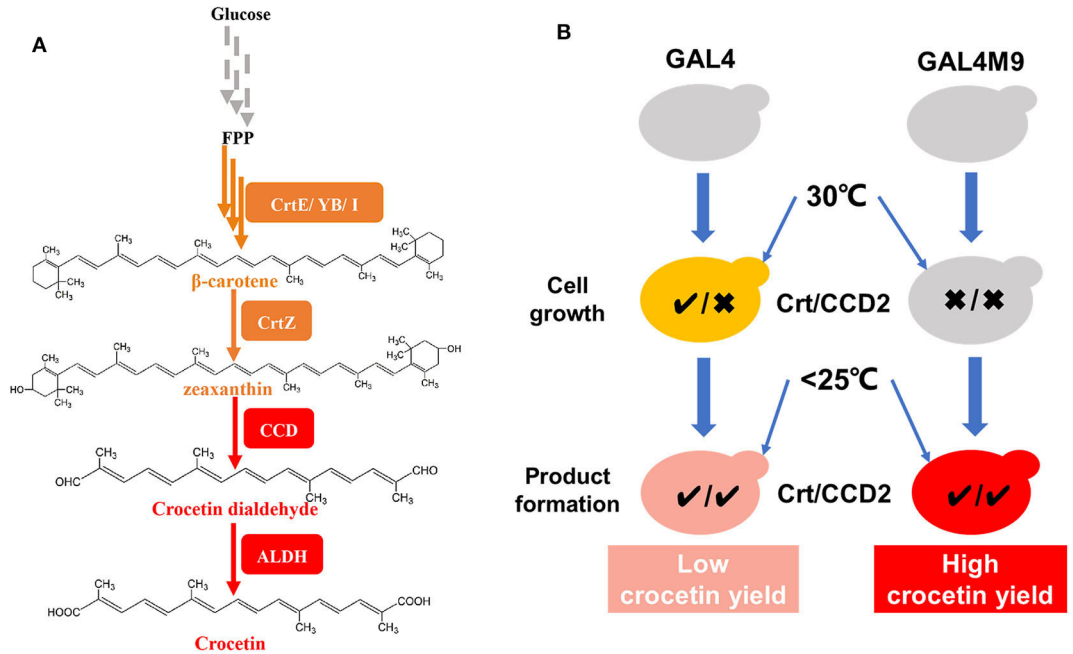


FIGURE 1 | Overview on the construction of a temperature-responsive yeast cell factory for crocetin biosynthesis. **(A)** *De novo* crocetin biosynthetic pathways from glucose. The carotenogenic genes were shown in orange and the crocetin biosynthetic pathway genes were shown in red. **(B)** The strategy of GAL4M9 based temperature switch to synchronize zeaxanthin biosynthesis (denoted by the activity of Crt) and cleavage (denoted by the activity of CCD2, equivalent to crocetin biosynthesis). The traditional two-stage fermentation strategy (left panel) resulted in physical separation of the substrate zeaxanthin (efficient synthesis at 30°C and storage at the cellular membranes) the CCD2 enzyme (functional at lower temperatures), accordingly low zeaxanthin conversion yield. In the temperature-responsive strain, zeaxanthin biosynthesis and CCD2 functional expression were synchronized at the lower temperature stage and higher zeaxanthin conversion yield could be achieved. CrtE, lycopene ε-cyclase; CrtYB, bifunctional phytoene synthase and lycopene cyclase; CrtI, phytoene desaturase; CrtZ, β-carotene hydroxylase; CCD, carotenoid cleavage dioxygenase; ALDH, aldehyde dehydrogenase; GAL4, transcription factor for activating GAL regulon genes; GAL4M9, GAL4 mutant showing low temperature induction.

from *Synechocystis* sp. PCC6803 (*syaldh*; accession number: WP_010873792) and *C. sativus* L. (*CsALDH*; accession number: MF596165.1) were codon optimized and synthesized by Tsingke Biotech (**Table S1**). All these genes were cloned into the multiple cloning sites (MSCs) of the pESC vectors (**Figure S3**), pESC-URA, pESC-LEU, and pESC-LEU2d by restriction digestion/ligation (MCS1: *Bam*HI/*Xho*I; MCS2: *Not*I) or Gibson Assembly.

Gene deletion and integration in *S. cerevisiae* were performed using the CRISPR-Cas9 method (Lian et al., 2017) and the schematic overview was briefly demonstrated in **Figure S4**. The guide RNA (gRNA) sequences were designed using the Benchling CRISPR-Cas9 tool (https://www.benchling.com/crispr) and cloned into p423-SpSgH and p426-SpSgH, constructed in our previous studies (Lian et al., 2017, 2019). For the construction of strains TL001-TL014, the expression cassettes were amplified by PCR containing 40 bp homology arms to the target chromosomal locus and co-transformed with the corresponding gRNA plasmid to the yeast strains using the LiAc/SS carrier DNA/PEG method (Gietz and Schiestl, 2007). The strains and plasmids used in this study were listed in **Table S2**, the corresponding primers were listed in **Table S3**, and the gRNA sequences as well as the chromosome loci for the integration of the heterologous gene expression cassettes were listed in **Table S4**.

Fermentation Conditions

For crocetin quantification, a single colony was picked from YPD or SCD agar plates and subcultured in tubes at 30°C and 250 rpm until saturation. Then, 300 μL seed culture was inoculated into a 250 mL flask containing 30 mL YPD (for the chromosome integrated strains) or SCD (for the plasmid bearing yeast strains). After culturing at 30°C for 24 h, temperature was shifted to 24 or 20°C and fermentation was continued for additional 7–8 days. All the experiments were performed in biological triplicates.

Carotenoid Extraction

Carotenoids were extracted from yeast cells according to the previous protocol (Chai et al., 2017). Three milliliter cells were harvested by centrifugation at 12,000 rpm for 3 min, washed with 3 mL distilled water, and suspended in 0.5 mL of 3M HCl. The suspension was boiled for 2 min and chilled on ice immediately for 3 min. The cell pellet was resuspended in 100 μL of 50:50 MeOH: acetone containing 1% (w/v) butylated hydroxytoluene

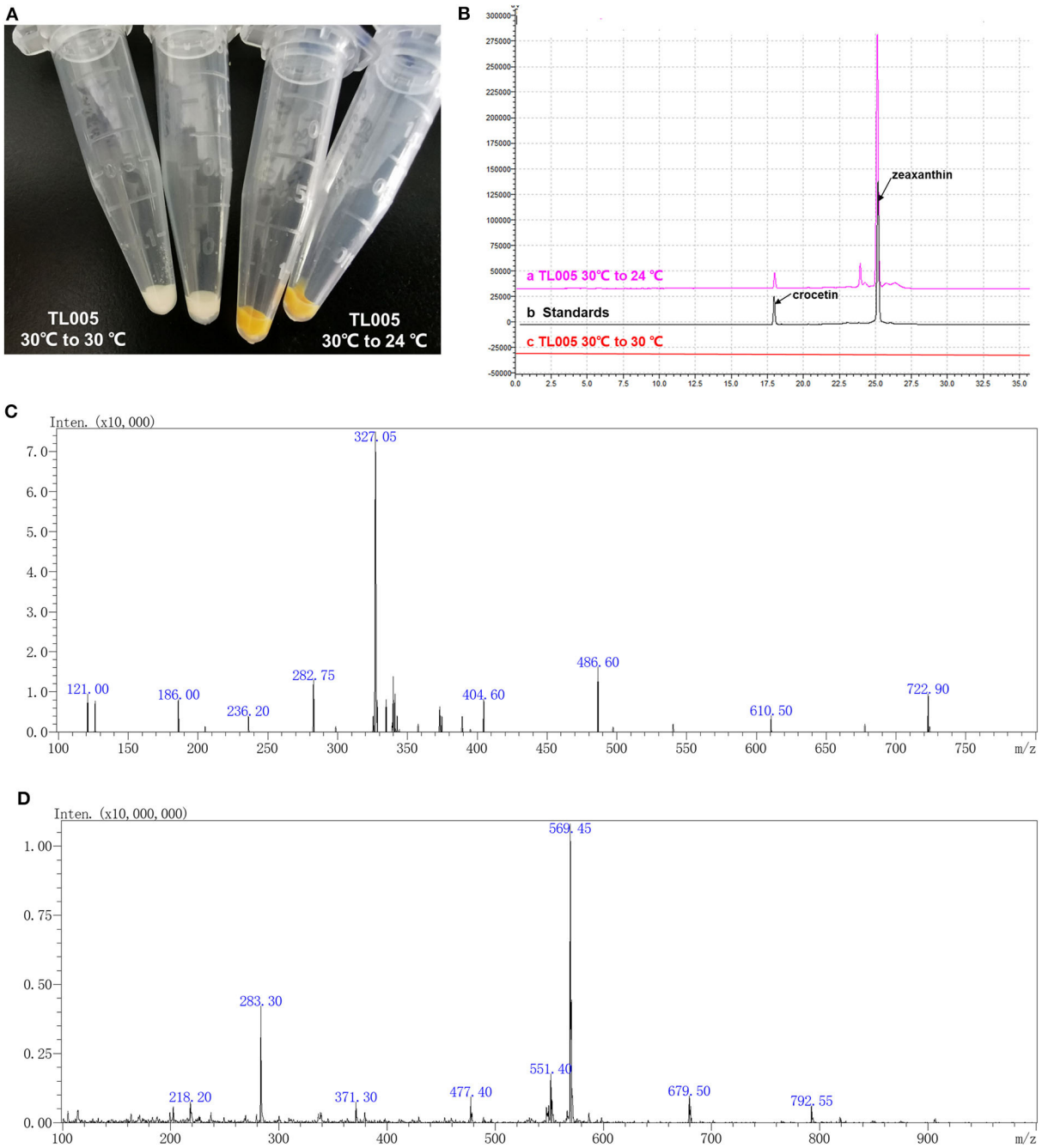


FIGURE 2 | Temperature dependence of zeaxanthin and crocetin biosynthesis in the engineered yeast cell factory. **(A)** Color of the cell pellets (strain TL005) with (right two tubes, yellow-to-orange color) or without (left two tubes, no visible color) shifting temperature to 24°C. **(B)** HPLC chromatograms of the fermentation profiles of TL005, together with the crocetin and zeaxanthin standards. **(C)** MS spectra of crocetin produced by TL005. **(D)** MS spectra of zeaxanthin produced by TL005.

and each sample was extracted twice with the same amount of solvent.

HPLC-MS Quantification

The supernatant was passed through a 0.22 μm membrane filter and directly analyzed using a SHIMADZU Liquid chromatography-tandem mass spectrometry (LC-MS/MS 8045) equipped with a UV detector at 420 nm and 30°C. Separation of compounds were performed on a HyPURITY^{TM} C18 HPLC column (150 mm × 4.6 mm, 3 μm, Thermo Scientific) with a flow rate of 0.5 mL/min. The mobile phase consisted of 10 mM ammonium formate solution (solvent A) and methanol (solvent B). The following gradient elution program was used: 60–2% solvent A over 20 min and returned to 60% solvent A over 20 min.

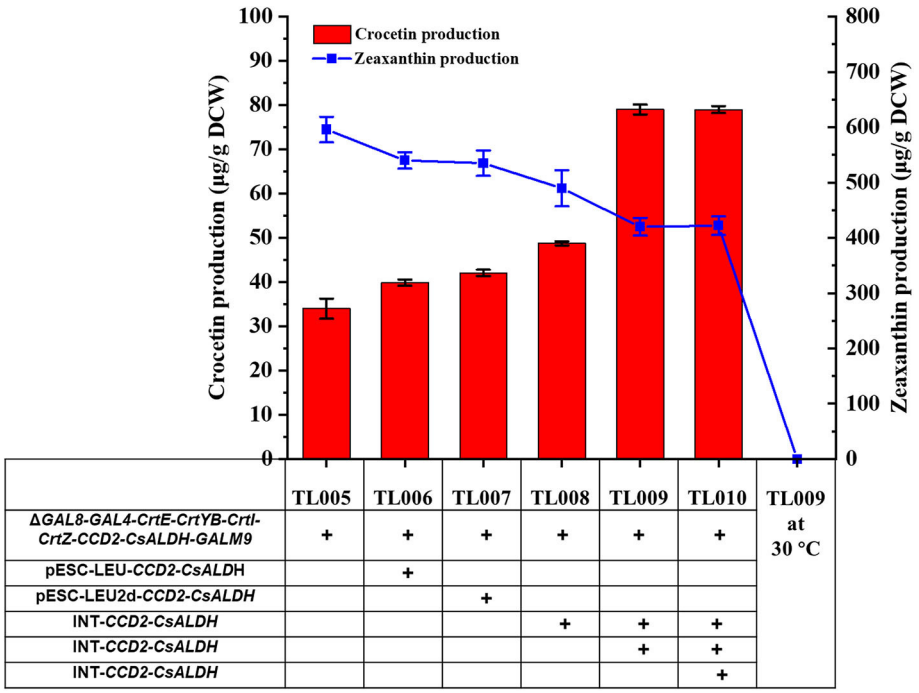


FIGURE 3 | Enhancement of crocetin biosynthesis via optimizing the copy numbers of *CCD2* and Cs*ALDH* genes. With TL005 at 30°C as the reference strain, additional copies of *CCD2*-Cs*ALDH* were introduced either by multi-copy plasmids (TL006 and TL007) or multi-copy genome integration (TL008, TL009, and TL010). Crocetin (red bars) and zeaxanthin (blue squares) produced by the engineered strains were quantified by HPLC-MS. The strains were cultured at 30°C for 24 h and then switched to 24°C fermentation for additional 168 h. Error bars represent SD of biological triplicates.

The mass spectrometer was an APCI ion source equipped with a triple quadrupole mass analyzer and the negative ionization mode was used for carotenoid and crocetin analysis. The mass spectrometer was scanned from m/z 50 to 800. The desolvation line (DL) temperature was held at 200°C, with a spray voltage of 1.8 kV and an atomizing gas flow rate of 6 L/min.

RESULTS

Construction of a Temperature-Regulated Crocetin Biosynthetic Pathway

According to the previous study (Zhou et al., 2018), a temperature-responsive yeast strain with an engineered GAL regulon should be reconstructed by knocking out *GAL80* and *GAL4* and introducing the *GAL4M9* expression cassette. In order to construct a temperature-responsive crocetin producing strain (TL005), four carotenogenic genes (*CrtE*, *CrtYB*, *CrtI,* and *CrtZ*), crocetin biosynthesis genes (Cs*CCD2* and Cs*ALDH*), and *GAL4M9* was integrated to the yeast genome, together with the inactivation of *GAL4* and *GAL80* (**Figure S1**). More specifically, *GAL80* locus was replaced by the Cs*CCD2* and *CrtE* expression cassettes, *GAL4* locus was replaced by the *CrtZ* and Cs*ALDH* expression cassettes, followed by the integration of *CrtYB* and *CrtI* expression cassettes as well as the *GAL4M9* expression cassette. To verify the temperature-dependence of zeaxanthin and crocetin biosynthesis, strain TL005 was cultured with or without a temperature shift to 24°C. As shown in **Figure 2A**, no visible color was observed when the strain was constantly maintained at 30°C (**Table S2**, strain TL005), while yellow-to-orange pigment was formed after temperature switch, indicating a temperature-dependent biosynthesis of carotenoids. The biosynthesis of zeaxanthin and crocetin was further confirmed using HPLC-MS analysis (**Figure 2B**). Consistent with the pigment formation results, crocetin, and zeaxanthin production was only detected in strain TL005 after the temperature shift. A peak with a retention time (t_R) of 17.85 min was identified as crocetin (m/z = 327.05) by MS (**Figure 2C**, **Figure S2A** for crocetin standard). Similarly, the biosynthesis of zeaxanthin (t_R = 25.02 min, m/z = 569.25) was also found to be temperature-dependent (**Figure 2D**, **Figure S2B** for zeaxanthin standard). Therefore, the temperature-responsive crocetin-producing yeast cell factory was successfully constructed.

Optimization of Crocetin Biosynthesis by Adjusting Copy Numbers of *CCD2-ALDH* Genes

Although the production of crocetin was achieved, zeaxanthin was accumulated to high levels, indicating CCD2 as a rate-limiting enzyme for crocetin biosynthesis. Integrating multiple copies of the biosynthetic genes or pathways into the yeast genome has been reported to benefit the production of the target compounds (Li et al., 2015). In addition, in several cases, the chromosome integrated strains were found to demonstrate

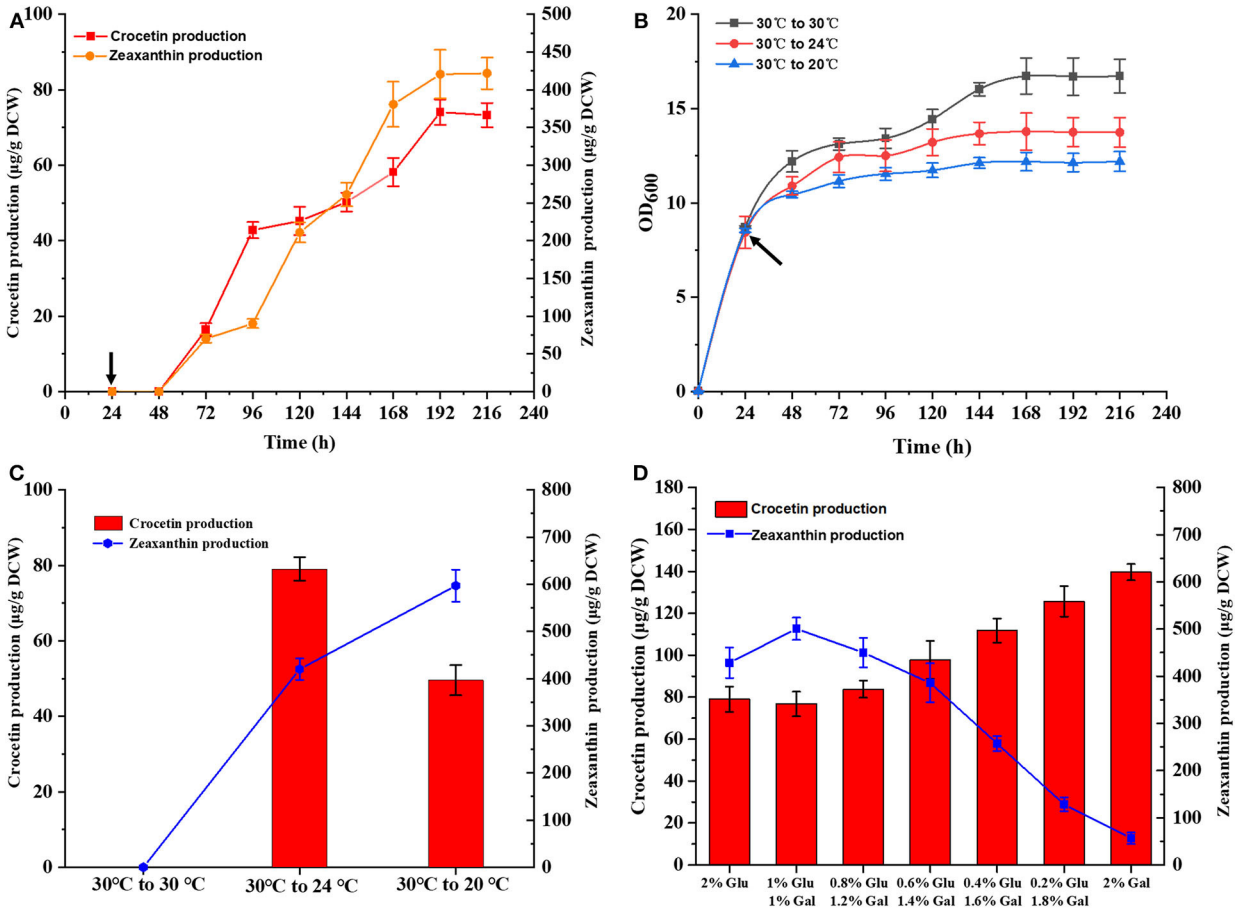


FIGURE 4 | Optimization of fermentation conditions for crocetin production in the engineered strain TL009. **(A)** Time course of crocetin production with TL009 at 24°C. **(B)** Growth curves of the engineered yeast strain TL009 at 30°C (black squares), shifting to 24°C (red circles), and shifting to 20°C (blue triangles). Arrows indicated the timing of temperature shift. Error bars represent SD of biological triplicates. **(C)** Effect of fermentation temperatures (30°C, 30–24°C, 30–20°C) for zeaxanthin (blue line) and crocetin (red bars) biosynthesis. **(D)** Crocetin and zeaxanthin production of TL009 in the media with various ratios of glucose and galactose. Error bars represent SD of biological triplicates.

higher stability and accordingly higher production than multi-copy plasmid bearing yeast strains (Lee and Silva, 1997; Shi et al., 2014, 2016). SyAldh was found to enable efficient production of crocetin in a previous report (Chai et al., 2017) and was included for evaluation in the present study as well.

To explore whether the production of crocetin could be improved via further increasing the copy number of *CCD2* and *ALDH* genes, plasmids pESC-LEU-*CCD2-CsALDH*, pESC-LEU-*CCD2-syaldh*, pESC-LEU2d-*CCD2-CsALDH*, and pESC-Leu2d-*CCD2-syaldh* were constructed (LEU: 20–30 copies per cell; LEU2d: 90–100 copies per cell; Erhart and Hollenberg, 1983) and transformed into TL005 and TL010 strains, respectively. Compared with the reference strains TL005 (33.09 ± 3.34 μg/g DCW) and TL0010 (29.47 μg/g DCW), the introduction of *CCD2-ALDH* on multi-copy plasmids only marginally increased crocetin production, with the highest (1.27-fold) achieved via the cloning of *CCD2-CsALDH* on the pESC-LEU2d plasmid. These results indicated that the benefits of multi-copy overexpression might be limited by the stability of the episomal plasmid system.

Alternatively, *CCD2/CsALDH* and *CCD2/syaldh* expression cassettes were iteratively integrated into the chromosome of TL005 and TL011, respectively. Two chromosomal copies of *CCD2* and *ALDH* resulted in the construction of TL008 and TL0014, three copies for TL009 and TL015, and four copies for TL0010 and TL016, respectively (**Table S2**). Different with the plasmid bearing system, the production of crocetin in TL009 was significantly increased to 79.03 ± 1.78 μg/g DCW, which was approximately 2.38-fold higher than that in the reference strain TL005 (**Figure 3**). Interestingly, the introduction of *syaldh* failed to contribute to increasing crocetin production significantly (**Figure S5**). The discrepancy with the previous report (Chai et al., 2017) was probably due to the difference in the genetic background of the yeast host and the codon optimization algorithm. To further investigate the role of *syaldh* overexpression, strain TL017 without any heterologous *ALDH* was constructed and evaluated for crocetin production. Beyond expectation, strain TL017 and the *syaldh* overexpression strains produced comparable amount of crocetin, indicating

that the endogenous ALDHs were active enough and the overexpression of *syaldh* didn't further increase the ALDH activities. Nevertheless, these results indicated the benefits of higher copy numbers and genetic stability in improving crocetin production.

Optimization of Fermentation Conditions for Crocetin Production

Subsequently, time-course studies on crocetin production was performed using strain TL009, containing three chromosomal copies of *CCD2* and *CsALDH*. TL009 was cultured in shake flasks at 30°C for 24 h and then continuously cultured for additional 192 h after shifting the temperature to 24°C. The target product crocetin started to accumulate at 72 h and reached the maximal production level (74.34 ± 2.31 μg/g DCW) in 192 h (**Figures 4A,B**). In accordance with above results, TL009 only accumulated crocetin at 24°C. The absence of any crocetin accumulation at 30°C (**Figures 2B**, **3**) demonstrated the high sensitivity and stringency of the GAL4M9-mediated temperature-responsive regulation.

As mentioned above, the performance of the GAL regulon and enzymatic activity of CCD2 were largely dependent on temperature. Thus, the temperature for crocetin fermentation should be optimized. Strain TL009 was cultured in shake flasks at 30°C for 24 h and sub-cultured at 24 or 20°C for additional 192 h. Different with the previous result that 20°C was the optimal temperature for crocetin biosynthesis, strain TL009 produced a much higher level of crocetin at 24°C (79.03 ± 1.78 μg/g DCW) than that at 20°C (49.55 ± 1.24 μg/g DCW) (**Figure 4C**). Therefore, 24°C was more beneficial for converting zeaxanthin to crocetin in the temperature-responsive crocetin-producing yeast cell factory and used for the subsequent studies for crocetin production.

Finally, as the GAL regulon is tightly regulated by the composition of carbon sources, the ratio glucose and galactose supplemented to the fermentation media was optimized to further increase crocetin production by strain TL009. Generally, higher level production of crocetin was achieved with the increase of galactose composition (**Figure 4D**). More importantly, more zeaxanthin was converted and the zeaxanthin conversion yield was dramatically improved (**Figure S6**), indicating a balanced pathway for zeaxanthin biosynthesis and cleavage with galactose as the carbon source. Under the optimal conditions, strain TL009 produced crocetin at a level up to 139.67 ± 2.24 μg/g DCW in the present study, with a zeaxanthin conversion yield higher than 75%.

DISCUSSION

Crocetin have attracted many researcher's interests due to its biological activities, i.e., anti-tumor, enhancement of the rate of oxygen transport, and inhibition of pro-inflammatory mediators (Nam et al., 2010). With the goal of balancing zeaxanthin accumulation with temperature-regulated CCD2 activity, a temperature-responsive yeast cell factory for crocetin biosynthesis was established. The pigment formation and product analysis of the engineered strains after temperature shift indicated the high stringency and sensitivity of the temperature-regulated system. The introduction of additional copies of *CCD2-ALDH* further improved crocetin biosynthesis, with the chromosome-integrated strain (TL009) worked much better than the plasmid-bearing (TL006 and TL007) strains. These results demonstrated the advantages of using CRISPR-Cas9 technology to construct genetically stable and multi-copy integrated strains for efficient production of the desired products (Tyo et al., 2009; Shi et al., 2014).

Through the optimization of *CCD2-ALDH* copy numbers as well as shake-flask fermentation conditions, an efficient and stable crocetin-producing strain was established, with a titer of up to 139.67 ± 2.24 μg/g DCW and the highest zeaxanthin conversion yield of 77%. The crocetin production level was comparable to those reported in previous studies, ∼160 μg/g DCW in 5-L bioreactors (Chai et al., 2017) and 62.79 μg/g DCW in shake flasks (Tan et al., 2019). Nevertheless, much higher zeaxanthin conversion yield (up to 77%) was achieved in the present study (**Figure S6**), highlighting the significance of synchronizing zeaxanthin biosynthesis and conversion for efficient crocetin biosynthesis as well as the advantage of the temperature-responsive system to achieve such a challenging goal. In addition, as the upper stream mevalonate pathway was systematically engineered to enhance terpenoid biosynthesis in the previous studies (Chai et al., 2017; Tan et al., 2019), the reported zeaxanthin production level was at least 10-fold higher than that achieved in the present study. As the direct precursor for crocetin biosynthesis, zeaxanthin biosynthesis should be strengthened, such as overexpression of the mevalonate pathway genes and the repression of the competing pathway genes, to further enhance crocetin production in the temperature-responsive yeast strain.

Although the production of crocetin was significantly improved via pathway engineering and/or metabolic engineering in both the present study and previous studies (Chai et al., 2017; Tan et al., 2019), zeaxanthin was still accumulated to relatively high levels, indicating zeaxanthin cleavage catalyzed by CCD2 as a rate-limiting step and room for further improving crocetin production. In other words, although the timing of zeaxanthin biosynthesis and cleavage was synchronized using the temperature switch, the pathway efficiency of zeaxanthin biosynthesis and crocetin biosynthesis should be further coordinated. Frusciante et al. found that CCD2 showed low affinity, low activity, and poor substrate specificity for zeaxanthin, which remained the biggest challenge for efficient production of crocetin using microbial cell factories. Increasing the catalytic activity of CCD2 by protein engineering or screening novel CCD2 enzymes, the fusion of CrtZ with CCD2 to channel the flux from zeaxanthin toward crocetin, and the engineering of the upper mevalonate pathway to enhance carotenogenesis are effective strategies to further improve crocetin production.

In summary, the present study reported the construction of a temperature-regulated crocetin production in *S. cerevisiae* for the first time, where the biosynthesis and cleavage of zeaxanthin was coordinated using an engineered GAL4-based temperature

switch. Using the CRISPR-Cas9 based facile genome engineering technology, crocetin biosynthesis was optimized by integrating three copies of *CCD2* and Cs*ALDH* genes, with a final titer of 139.67 μg/g DCW and a zeaxanthin conversion yield of up to 77%. Our study provides a versatile platform for facilitating the biosynthesis of crocetin and other valuable epoxycarotenoids in yeast, such as violaxanthin, neoxanthin, and fucoxanthin (Cataldo et al., 2020).

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

TL, CD, MQ, and BZ performed the experiments. JL and TL conceived the study and wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Key Research and Development Program of China (2018YFA0901800), the Natural Science Foundation of China (21808199), the Natural Science Foundation of Zhejiang Province (R20B060006), and the Fundamental Research Funds for the Zhejiang Provincial Universities (2019XZZX003-12).

ACKNOWLEDGMENTS

We would like to express our gratitude to Prof. Hongwei Yu and Lidan Ye for providing the plasmid pUMRI-P_{ACT}-GAL4M9.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00653/full#supplementary-material

REFERENCES

Ahrazem, O., Rubio-Moraga, A., Argandoña, J., Castillo-López, R., and Gómez-Gómez, L. (2016). Intron retention andrhythmic diel pattern regulation of carotenoid cleavage dioxygenase 2 during crocetin biosynthesis in saffron. *Plant Mol. Biol.* 91, 355–374. doi: 10.1007/s11103-016-0473-8

Cataldo, V. F., Arenas, N., Salgado,V., Camilo, C., Ibáñez, F., and Agosin, E. (2020). Heterologous production of the epoxycarotenoid violaxanthin in *S. cerevisiae*. *Metab. Eng.* 59, 53–63. doi: 10.1016/j.ymben.2020.01.006

Chai, F., Wang, Y., Mei, X., Yao, M., Chen, Y., Liu, H., et al. (2017). Heterologous biosynthesis and manipulation of crocetin in *S. cerevisiae*. *Microb. Cell Factories*. 16:54. doi: 10.1186/s12934-017-0665-1

Chakshusmathi, G., Mondal, K., Lakshmi, G. S., Singh, G., Roy, A., Ch, R. B., et al. (2004). Design of temperature-sensitive mutants solely from amino acid sequence. *Pro. Natl. Acad. Sci. U.S.A.* 101, 7925–7930. doi: 10.1073/pnas.0402222101

Colapietro, A., Mancini, A., D'Alessandro, A. M., and Festuccia, C. (2019). Crocetin and crocin from saffron in cancer chemotherapy and chemoprevention. *Anticancer. Agents Med. Chem.* 19, 38–47. doi: 10.2174/1871520619666181231112453

Costantina, D. O., Sarah, F., Paola, F., Gianfranco, D., Hosseinpour, A. N., Marco, P., et al. (2018). Candidate enzymes for saffron crocin biosynthesis are localized in multiple cellular compartments. *Plant Physiol.* 177, 990–1006. doi :10.1104/pp.17.01815

Erhart, E., and Hollenberg, C. (1983). The presence of a defective LEU2 gene on 2μDNA recombinant plasmids of *S. cerevisiae* is responsible for curing and high copy number. *J. Bacteriol.* 156, 625–635. doi: 10.1128/JB.156.2.625-635.1983

Frusciante, S., Diretto, G., Bruno, M., Ferrante, P., Pietrella, M., Prado-Cabrero, A., et al. (2014). Novel carotenoid cleavage dioxygenase catalyzes the first dedicated step in saffron crocin biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.*111, 12246–12251. doi: 10.1073/pnas.1404629111

Gietz, R. D., and Schiestl, R. H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* 2, 31–34. doi: 10.1038/nprot.2007.13

Hashemi, M., and Hosseinzadeh, H. (2019). A comprehensive review on biological activities and toxicology of crocetin. *Food Chem. Toxicol.* 130, 44–60. doi: 10.1016/j.fct.2019.05.017

Kramer, R. A., DeChiara, T. M., Schaber, M. D., and Hilliker, S. (1984). Regulated expression of a human interferon gene in yeast: control by phosphate concentration or temperature. *Proc. Natl. Acad. Sci. U.S.A.* 81, 367–370. doi: 10.1073/pnas.81.2.367

Lee, F. W. F., and Silva, N. A. D. (1997). Improved efficiency and stability of multiple cloned gene insertions at the δ sequences of *S. cerevisiae*. *Appl. Microbiol. Biotechnol.* 48, 339–345. doi: 10.1007/s002530051059

Li, M., Kildegaard, K. R., Chen, Y., Rodriguez, A., Borodina, I., and Nielsen, J. (2015). *De novo* production of resveratrol from glucose or ethanol by engineered *S. cerevisiae*. *Metab. Eng.* 32, 1–11. doi: 10.1016/j.ymben.2015.08.007

Lian, J., Hamedirad, M., Hu, S., and Zhao, H. (2017). Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. *Nat. Commun.* 8:1688. doi: 10.1038/s41467-017-01695-x

Lian, J., Hamedirad, M., and Zhao, H. (2018a). Advancing metabolic engineering of *Saccharomyces cerevisiae* using the CRISPR/Cas system. *Biotechnol. J.* 13:e1700601. doi: 10.1002/biot.201700601

Lian, J., Jin, R., and Zhao, H. (2016). Construction of plasmids with tunable copy numbers in *S. cerevisiae* and their applications in pathway optimization and multiplex genome integration. *Biotechnol. Bioeng.*113, 2463–2473. doi: 10.1002/bit.26004

Lian, J., Mishra, S., and Zhao, H. (2018b). Recent advances in metabolic engineering of, *Saccharomyces cerevisiae*: new tools and their applications. *Metab. Eng.* 50, 85–108. doi: 10.1016/j.ymben.2018.04.011

Lian, J., Schultz, C., Cao, M., Hamedirad, M., and Zhao, H. (2019). Multi-functional genome-wide CRISPR system for high throughput genotype–phenotype mapping. *Nat. Commun.* 10:5794. doi: 10.1038/s41467-019-13621-4

Murill, G., Hu, S., and Fernandez, M. (2019). Zeaxanthin: metabolism, properties, and antioxidant protection of eyes, heart, liver, and skin. *Antioxidants* 8:390. doi: 10.3390/antiox8090390

Nam, K. N., Park, Y. M., Jung, H. J., Lee, J. Y., Min, B. D., Park, S. U., et al. (2010). Anti-inflammatory effects of crocin and crocetin in rat brain microglial cells. *Eur. J. Pharmacol.* 648, 110–116. doi: 10.1016/j.ejphar.2010.09.003

Niu, F. X., Lu, Q., Bu, Y. F., and Liu, J. Z. (2017). Metabolic engineering for the microbial production of isoprenoids: carotenoids and isoprenoid-based biofuels. *Synth. Syst. Biotechnol.* 2, 167–175. doi: 10.1016/j.synbio.2017.08.001

Shen, H., Cheng, B., Zhang, Y., Tang, L., Li, Z., and Bu, Y. (2016). Dynamic control of the mevalonate pathway expression for improved zeaxanthin production in *E. coli* and comparative proteome analysis. *Metab. Eng.* 38, 180–190. doi: 10.1016/j.ymben.2016.07.012

Sheu, S. J., and Hsin, W. C. (1998). HPLC separation of the major constituents of Gardeniae fructus. *J. High. Resol. Chromatogr.* 21, 523–526. doi: 10.1002/SICI1521-41681998090121:9523::AID-JHRC5233.0.CO;2-B

Shi, S., Liang, Y., Zhang, M. M., Ang, E. L., and Zhao, H. (2016). A highly efficient single-step, markerless strategy for multi-copy chromosomal integration of large biochemical pathways in *Saccharomyces cerevisiae*. *Metab. Eng.* 33, 19–27. doi: 10.1016/j.ymben.2015.10.011

Shi, S., Valle-Rodríguez, J. O., Siewers, V., and Nielsen, J. (2014). Engineering of chromosomal wax ester synthase integrated *Saccharomyces cerevisiae* mutants for improved biosynthesis of fatty acid ethyl esters. *Biotechnol. Bioeng.* 111, 1740–1747. doi: 10.1002/bit.25234

Sledziewski, A. Z., Bell, A., Yip, C., Kelsay, K., Grant, F. J., and MacKay, V. L. (1990). Superimposition of temperature regulation on yeast promoters. *Methods Enzymol.* 185, 351–366. doi: 10.1016/0076-6879908 5031-I

Sun, T., Yuan, H., Cao, H., Yazdani, M., Tadmor, Y., and Li, L. (2018). Carotenoid metabolism in plants: the role of plastids. *Mol. Plant.* 11, 58–74. doi: 10.1016/j.molp.2017.09.010

Tan, H., Chen, X., Liang, N., Chen, R., Chen, J., Hu, C., et al. (2019). Transcriptome analysis reveals novel enzymes for apo-carotenoid biosynthesis in saffron and allows construction of a pathway for crocetin synthesis in yeast. *J. Exp. Bot.* 70, 4819–4834. doi: 10.1093/jxb/erz211

Trautmann, D., Beyer, P., and Al-Babili, S. (2013). The orf slr0091 of *Synechocystis* sp. pcc6803 encodes a high-light induced aldehyde dehydrogenase converting apocarotenals and alkanals. *FEBS J.* 280, 3685–3696. doi: 10.1111/febs. 12361

Tyo, K. E., Ajikumar, P. K., and Stephanopoulos, G. (2009). Stabilized gene duplication enables long-term selection-free heterologous pathway expression. *Nat. Bio.* 27, 760–765. doi: 10.1038/nbt.1555

Wang, C., Zhao, S., Shao, X., Park, J. B., Jeong, S. H., Kwak, W. J., et al. (2019). Challenges and tackles in metabolic engineering for microbial production of carotenoids. *Microb. Cell. Fact.*18:55. doi: 10.1186/s12934-019-1 105-1

Wang, W., He, P., Zhao, D., Ye, L., Dai, L., Zhang, X., et al. (2019). Construction of *Escherichia coli* cell factories for crocin biosynthesis. *Microb. Cell. Fact.* 18:120. doi: 10.1186/s12934-019-1166-1

Xie, W., Liu, M., Lv, X., Lu, W., Gu, J., and Yu, H. (2014). Construction of a controllable beta-carotene biosynthetic pathway by decentralized assembly strategy in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* 111, 125–133. doi: 10.1002/bit.25002

Zhou, P., Li, M., Shen, B., Yao, Z., Bian, Q., Ye, L., et al. (2019). Directed co-evolution of β-carotene ketolase and hydroxylase and its application in temperature-regulated biosynthesis of astaxanthin. *J. Agric. Food Chem.* 67, 1072–1080. doi: 10.1021/acs.jafc.8b05003

Zhou, P., Xie, W., Yao, Z., Zhu, Y., Ye, L., and Yu, H. (2018). Development of a temperature-responsive yeast cell factory using engineered GAL4 as a protein switch. *Biotechnol. Bioeng.* 115, 1321–1330. doi: 10.1002/bit.26544

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

MINI REVIEW
published: 02 July 2020
doi: 10.3389/fbioe.2020.00731

Advances in RNAi-Assisted Strain Engineering in *Saccharomyces cerevisiae*

*Yongcan Chen, Erpeng Guo, Jianzhi Zhang and Tong Si**

CAS Key Laboratory of Quantitative Engineering Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

Saccharomyces cerevisiae is a widely used eukaryotic model and microbial cell factory. RNA interference (RNAi) is a conserved regulatory mechanism among eukaryotes but absent from *S. cerevisiae*. Recent reconstitution of RNAi machinery in *S. cerevisiae* enables the use of this powerful tool for strain engineering. Here we first discuss the introduction of heterologous RNAi pathways in *S. cerevisiae*, and the design of various expression cassettes of RNAi precursor reagents for tunable, dynamic, and genome-wide regulation. We then summarize notable examples of RNAi-assisted functional genomics and metabolic engineering studies in *S. cerevisiae*. We conclude with the future challenges and opportunities of RNAi-based approaches, as well as the potential of other regulatory RNAs in advancing yeast engineering.

Keywords: RNAi, *Saccharomyces cerevisiae*, strain engineering, functional genomics, synthetic biology

OPEN ACCESS

Edited by:
Yunzi Luo,
Sichuan University, China

Reviewed by:
Zihe Liu,
Beijing University of Chemical Technology, China
Xinqing Zhao,
Shanghai Jiao Tong University, China

***Correspondence:**
Tong Si
tong.si@siat.ac.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *17 April 2020*
Accepted: *10 June 2020*
Published: *02 July 2020*

Citation:
Chen Y, Guo E, Zhang J and Si T (2020) Advances in RNAi-Assisted Strain Engineering in Saccharomyces cerevisiae. Front. Bioeng. Biotechnol. 8:731. doi: 10.3389/fbioe.2020.00731

INTRODUCTION

As a unicellular eukaryotic microorganism, *S. cerevisiae* has become one of the most widely used microbial cell factory in the production of value-added chemicals, biofuels and biopharmaceuticals (Liu et al., 2013; Nielsen, 2019). *S. cerevisiae* has many desirable traits for industrial fermentation. With a long history in baking and brewing, *S. cerevisiae* is generally recognized as safe (GRAS). Compared with prokaryotes such as *E. coli*, *S. cerevisiae* possesses multiple organelles providing physical compartments for diverse biochemical reactions; it is also capable of post-translational modifications, which are required for heterologous synthesis of complex proteins (Tokmakov et al., 2012). *S. cerevisiae* is physiologically stable and highly robust toward harsh industrial conditions such as low pH, high osmotic pressure, and toxic inhibitors (Hong and Nielsen, 2012; Kavscek et al., 2015). Due to its popularity in fundamental and applied research, versatile tools have been developed in *S. cerevisiae* for genetic manipulation (Si et al., 2015b; Lian et al., 2018; Jiang et al., 2019) and bioprocess development (Hasunuma and Kondo, 2012; Tripathi and Shrivastava, 2019).

RNA interference (RNAi) is a post-transcriptional, gene-silencing mechanism broadly distributed in eukaryotic organisms. RNAi mediates many essential biological processes, including defense against viruses and transposons, maintenance of chromosome and genome integrity, and cellular differentiation and development (Castel and Martienssen, 2013; Gutbrod and Martienssen, 2020). Various organisms comprise different mechanisms for RNAi, but the basic process shared three common steps. First, small interfering RNA (siRNA) duplexes of 21–25 nucleotides are generated from long double-stranded RNA (dsRNA) precursors by a ribonuclease enzyme called Dicer. Second, siRNAs are loaded into Argonaute proteins to form a protein-RNA complex known as RNA-induced silencing complex (RISC), where the guide and passenger strands of siRNA

are dissociated. Third, RISC finds and cleaves the cognate mRNA molecule, whose sequence is homologous to the siRNA loaded in the complex (Hannon, 2002; Wilson and Doudna, 2013; Ipsaro and Joshua-Tor, 2015). Thanks to its facile implementation and high specificity, RNAi has been widely used to knock down individual genes of interest, as well as to perform genome-wide reduction-of-function screening, since its first discovery in *Caenorhabditis elegans* (Fire et al., 1998; Agrawal et al., 2003; Boutros and Ahringer, 2008).

RNA interference is conserved in almost all eukaryotic organisms including fungi, plants and animals (Agrawal et al., 2003; Gutbrod and Martienssen, 2020), but it is evolutionarily lost in *S. cerevisiae*, possibly due to the incompatibility with a beneficial dsRNA "killer virus" (Drinnenberg et al., 2009, 2011). This theory is supported by the ecological significance of "killer viruses" (Boynton, 2019) and the importance of RNAi in antiviral defense (Li et al., 2013; Waldron et al., 2018). However, a non-canonical RNAi pathway exists in other budding yeasts, including *Saccharomyces castellii*, *Candida albicans,* and *Kluyveromyces polysporus* (Drinnenberg et al., 2009). Unlike canonical Dicers, which generate siRNAs of regular sizes from dsRNA termini, budding-yeast Dicers start processing in the interior of a dsRNA and work outward, with product size determined by the distance between neighboring active sites (Weinberg et al., 2011; Wilson and Doudna, 2013). On the other hand, other yeasts contain canonical RNAi machinery, including the fission yeast *Schizosaccharomyces pombe* (Volpe et al., 2002) and a human pathogenic yeast *Cryptococcus neoformans* (Loftus et al., 2005).

Since the discovery and characterization of the budding yeast pathways, RNAi has become an attractive platform for gene regulation in *S. cerevisiae*. Here we summarize the reconstitution of heterologous RNAi pathways in baker's yeast and the design of RNAi cassettes for desirable potency and specificity. We also provide notable examples in implementing RNA-based engineering tools for tunable, dynamic, and genome-scale gene modulation in *S. cerevisiae*. To conclude, the future directions of RNAi-based tools and the potentials of other regulatory RNAs in *S. cerevisiae* strain engineering are discussed.

TOOL DEVELOPMENT

RNAi Pathway Reconstitution

Heterologous RNAi pathways from *S. castelli* and human were successfully reconstituted in *S. cerevisiae*. To evaluate RNAi effectiveness and efficiency, fluorescent proteins were commonly utilized as reporters (Drinnenberg et al., 2009; Suk et al., 2011; Crook et al., 2014; Si et al., 2014; Purcell et al., 2018; Kildegaard et al., 2019). For the *S. castelli* pathway, robust repression of green fluorescent protein (GFP) signals were observed when both *DCR1* and *AGO1* were present, but *DCR1* alone was sufficient to generate GFP siRNAs in *S. cerevisiae* (Drinnenberg et al., 2009). On the other hand, all three components of the human RISC complex, Dicer, Argonaute-2 (Ago2), and HIV-1 transactivating response (TAR) RNA-binding protein (TRBP) (Gregory et al., 2005; Maniataki and Mourelatos, 2005), are necessary to reconstitute functional RNAi in *S. cerevisiae* (Suk et al., 2011). This difference between *S. castelli* and human pathways is probably because human TRBP is required to mediate interaction of Ago2 and siRNA bounded by Dicer, whereas budding yeast Dicers act as homodimers and did not require additional dsRBD domains (Chendrimada et al., 2005; Weinberg et al., 2011).

Due to evolutionary closeness and simplicity, the *S. castellii* RNAi pathway is commonly employed in *S. cerevisiae* (Drinnenberg et al., 2009; Crook et al., 2014; Si et al., 2014; Williams et al., 2015b; Purcell et al., 2018; Wang et al., 2019; **Figure 1A**). *S. castellii* Dicer and Argonaute were expressed either from a low-copy (centromeric) plasmid (Crook et al., 2014, 2016; Lee et al., 2016) or a single, integrated locus in the yeast genome (Drinnenberg et al., 2009; Si et al., 2014, 2017; Xiao and Zhao, 2014; Williams et al., 2015a,b; Purcell et al., 2018; Wang et al., 2019; **Figure 1A** and **Table 1**). To enable efficient gene silencing, gene expression of Dicer and Argonaute was driven by strong constitutive promoters, such as P_{TEF1}, P_{GPD1}, P_{TPI1}, and P_{PGK1}. Although no comprehensive comparisons have been performed among different expression formats, it is suggested that constitutive expression driven by a strong promoter from a genomic locus is sufficient for efficient RNAi silencing.

RNAi Reagent Cassette Design

In addition to RNAi pathways, RNAi precursor is another essential component for efficient gene silencing. We will limit our discussions to the design of RNAi reagent cassettes for the *S. castellii* RNAi pathway, which is widely employed for RNAi reconstitution in *S. cerevisiae*. In its native host, *S. castellii* siRNAs are originated from hairpin RNAs, which are transcribed from inverted genomic repeats and fold back on their own paired regions (Drinnenberg et al., 2009). These RNA hairpins were presumed to be composed of stems of 100–400 bp and loops ranging from 19 to >1600 nt (Drinnenberg et al., 2009). As expected, when co-expressed with *S. castellii* Dcr1 and Ago1 in *S. cerevisiae*, the hairpin RNA construct with inverted repeats of a 275 bp GFP fragment induced strong reduction of GFP mRNA and fluorescence levels (Drinnenberg et al., 2009). In addition to hairpin RNAs (Drinnenberg et al., 2009; Crook et al., 2014; Purcell et al., 2018), other forms of RNAi precursors, such as dsRNAs (Si et al., 2014; Crook et al., 2016) and antisense RNAs (Suk et al., 2011; Si et al., 2014, 2017), can also be accepted by the *S. castellii* RNA machinery for gene repression in *S. cerevisiae*. Notably, although a growing collection of computational tools is available to design RNAi targeting sequences, these tools are primarily developed for canonical RNAi pathways as recently summarized (Lagana et al., 2014; Jain and Wadhwa, 2018). Given the mechanistic differences of siRNA generation between canonical and budding-yeast RNAi machinery, it remains elusive if existing algorithms may help to identify RNAi targets for efficient gene silencing in *S. cerevisiae*. Here, we will focus on how different designs of expression cassettes may result in various levels of RNAi silencing (**Table 1**).

Hairpin RNA

The design of hairpin RNA cassettes, such as promoter, length of hairpin RNAs, and plasmid copy number, has a major impact

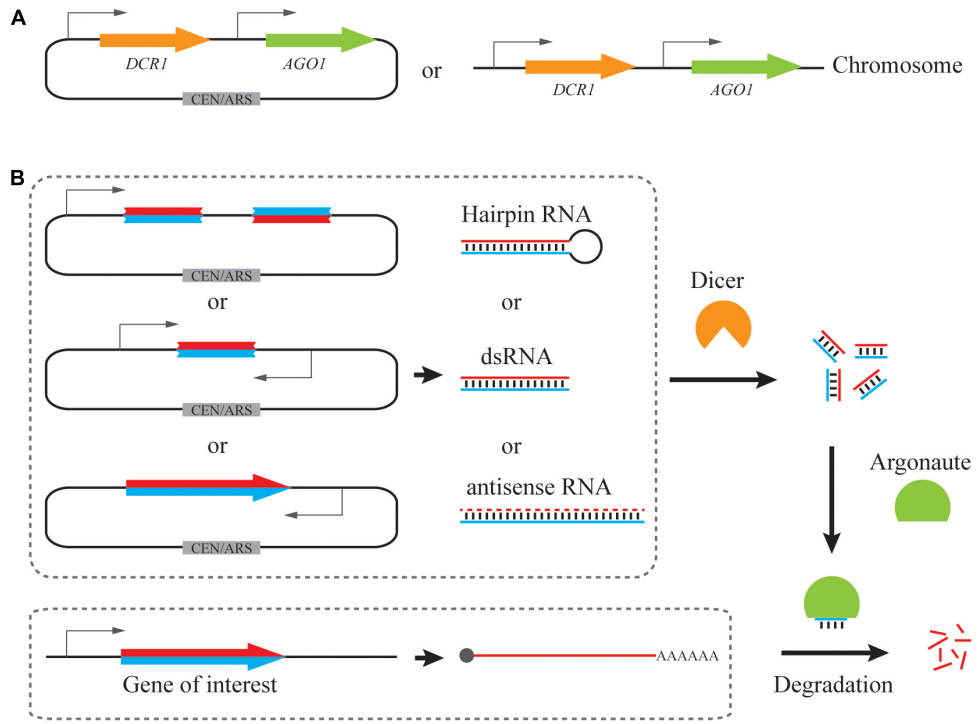


FIGURE 1 | Reconstitution of a synthetic RNAi pathway in *S. cerevisiae*. **(A)** Expression cassettes of RNAi pathway. *DCR1* and *AGO1* from *S. castellii* are more commonly utilized, being cloned into a low-copy plasmid or integrated into the genome. **(B)** Expression cassettes of siRNA precursors. Hairpin RNA, dsRNA and full-length antisense RNA have been used for generating siRNAs by Dicer. siRNA duplexes mediate mRNA degradation and hence repression of a target gene with the help of Argonaute.

TABLE 1 | Expression cassettes and efficiency of RNAi in *S. cerevisiae*.

siRNA precursor	Cassettes for RNAi		Silencing efficiency	References
	Pathway	siRNA precursor		
hairpin RNA	pRS404-TEF1p-Ago1 and pRS405-TEF1p-Dcr1	pRS403-GAL1p-GFP (275 bp)-*RAD9* intron 1-GFPrc	strong silencing	Drinnenberg et al., 2009
	p415-TDH3p-Dicer-TEF1p-Ago2	pRS414-TDH3p-YFP (200 bp)-intron-YFPrc	93% and 80% for highly and weakly expressed genome-integrated YFP	Crook et al., 2014
	pRS404-PTEF-Ago1(x2)-pRS405-PTEF-Dcr1(x2)	PFUS1J2- ARO7sense-rad9linker-ARO7antisense-pRS413; PFUS1J2-ZWF1/CDC19sense-rad9linker-ZWF1/CDC19antisense-pRS406	97% for *ARO7*, 87.8% for *CDC19*, insignificant for *ZWF1*	Williams et al., 2015a
	pX-3-AGO1-DCR1 TDH3p:ZWF1↓:can1; pX-3-AGO1-DCR1 RNR2p:ZWF1↓:can1	pX-3-AGO1-DCR1 TDH3p:ZWF1↓:can1; pX-3-AGO1-DCR1 RNR2p:ZWF1↓:can1	95% or 80% for TDH3p or RNR2p-drived shRNA expression, respectively	Kildegaard et al., 2019
dsRNA	pRS404-TEF1p-Ago1 and pRS405-TEF1p-Dcr1	pRS403-GAL1p-GFP (275 bp)-URA3prc	intermediate silencing	Drinnenberg et al., 2009
	pRS-delta-KanMX-LoxP-TEF1p-AGO1-PGK1t-TPI1p-DCR1-GPD1t	pRS416-GPDtrc-TEF1p-GFP (1–180 bp region)-TPI1prc- PGK1t;	~80%	Si et al., 2014
	p415-TDH3p-Dicer-TEF1p-Ago2	pRS414-TDH3p-*RAD9* intron-YFP (200 bp)-*RAD9* intron rc-TEF1prc	84% for strongly expressed YFP, 94% for weakly expressed YFP	Crook et al., 2016
antisense RNA	pAG413Gal-Ago2, pAG416Gal-Dicer, pAG415Gal-TRBP	pAG424Gal-AS-GFP	80%	Suk et al., 2011
	pRS-delta-KanMX-LoxP-TEF1p-AGO1 -PGK1t-TPI1p-DCR1-GPD1t	pRS406-TEF1p-AS GFP-PGK1t	~95%	Si et al., 2014

on RNAi silencing efficiency (**Table 1**). To state a general observation, a low-copy auxotrophic plasmid expressing long hairpin by a strong promoter produces a potent RNAi reagent.

To investigate the impact of promoter strength, Crook et al. engineered reporter strains expressing yellow fluorescence protein (YFP) from a strong or weak promoter, which could mimic the different expression levels of native genes. Only when the hairpin RNA was expressed from the strong P_{TDH3} promoter, efficient repression (up to 80%) was obtained for strongly expressed YFP (Crook et al., 2014), indicating the amount of RNAi precursor was limited in this case. Moreover, three promoters with different strength (P_{CYC1}, P_{TEF1} and P_{TDH3}) were used to express a hairpin RNA targeting *ADE3*, the deletion of which can improve heterologous production of itaconic acid (IA). Various IA production levels were observed with different levels of hairpin RNA expression in three *S. cerevisiae* strains (BY4741, CEN.PK2-a, and Sigma 10560-4A), demonstrating the necessity of promoter screening to fine-tune RNAi efficacy (Crook et al., 2014). Similarly, the strong P_{TDH3} promoter elicited higher repression levels relative to the weak P_{RNR2} promoter when targeting the endogenous *ZWF1* gene using hairpin RNAi (Kildegaard et al., 2019). The observation that the promoter activity of a hairpin RNA cassette modulates its gene silencing efficiency was exploited for dynamic, context-dependent gene knockdown, whereby an inducible promoter was used to drive hairpin RNA expression (Williams et al., 2015a).

Hairpin length also affects RNAi efficiency. When the length of hairpin increased from 100 bp to 200 bp, the downregulation efficiency was improved by 30% when YFP is strongly expressed (Crook et al., 2014). Notably, when YFP was weakly expressed, the inhibition level of 200 bp hairpin was 6-fold higher than that of the 100 bp hairpin, indicating that longer hairpin length is especially important for efficient silencing of low-abundance transcripts. The potency of long (∼250 bp) RNAi hairpins in *S. cerevisiae* was also confirmed in later studies (Williams et al., 2015a; Kildegaard et al., 2019).

For plasmid copy number, when using an auxotrophic marker (*TRP1*), a low-copy, centromeric plasmid achieved higher knockdown efficiency relative to a high-copy, 2-micron vector, for both strongly and weakly expressed YFP (Crook et al., 2014). This effect, however, cannot be observed when using an antibiotic resistance marker (KanMX). Improved RNAi efficiency by low-copy, centromeric plasmid may be related to reduced cell-to-cell variability.

It is also possible to introduce synthetic RNAi targets to a native transcript. A single, invariant target sequence was inserted into the 3′ untranslated region (UTR) of a yERFP reporter gene, and repression of fluorescence signal was achieved by introducing a corresponding hairpin RNA in an RNAi capable *S. cerevisiae* (Purcell et al., 2018). Gene silencing efficiency can be modulated via variations in the repeat number or sequence length of targeting siRNA sequences placed in the stem structure. For example, more repeats induced stronger gene repression. The authors also demonstrated that nucleus-residing non-coding RNA can be targeted by hairpin RNAs, but it was unlikely mediated by RNAi, because RNAi effector proteins are normally expressed in cytosol. In this way, any gene can be theoretically regulated by the same RNAi precursor upon insertion of a synthetic target sequence in the UTRs, offering a graded and scalable module of gene regulation.

Long dsRNA

Long dsRNA molecules without hairpin loop structures can also be used as siRNA precursors. The dsRNA molecules can be transcribed by convergent promoters, whereby the sense and antisense transcripts are individually expressed under the control of two opposing promoters. However, it was showed that the amount of siRNA from dsRNA was much lower than from hairpin RNA precursors, leading to weaker GFP repression (Drinnenberg et al., 2009). Like hairpin RNA, strong promoters achieved higher repression than weaker promoters when driving dsRNA expression (Crook et al., 2016). To investigate the influence of dsRNA length, we generated various dsRNA constructs on a low-copy, centromeric plasmid (pRS416) targeting different regions of the *GFP* gene, driven by two strong, convergent promoters P_{TEF1} and P_{TPI1}. Among the dsRNAs corresponding to the 1–180, 1–360, 1–540, and 1–717 bp (full length) region of the *GFP* gene, knockdown efficiency was inversely related to dsRNA length, with the shortest (1–180 bp) dsRNA exhibited the strongest GFP repression (80%) (Si et al., 2014). As to the vector of dsRNA cassette, although efficient downregulation was observed from a centromeric plasmid, the silencing efficiency was even higher when integrated into the genome (Si et al., 2014; Crook et al., 2016). In addition, with the introduction of an intron from *Schizosaccharomyces pombe RAD9* directly downstream of convergent promoters to enhance transcriptional strength, the extent of downregulation was significantly improved for weakly expressed YFP, but insignificant for strongly expressed YFP (Crook et al., 2016). Together, similar to hairpin RNA, the use of strong promoters and low-copy vectors is desirable for dsRNA. The length of dsRNAs should be carefully optimized, and ∼200 bp was frequently used (Si et al., 2014; Crook et al., 2016; **Table 1**).

Antisense RNA

Antisense RNAs can act as siRNA precursors by hybridizing with endogenous transcripts to form dsRNA substrates of Dicer. We expressed the full-length, antisense GFP transcript under control of the P_{TEF1} promoter on an episomal plasmid. Silencing effect is negligible without a RNA pathway, and a 95% repression of GFP signal was observed in an RNAi capable strain (Si et al., 2014; **Table 1**). When integrated into the genome, an even higher silencing efficiency and less population heterogeneity in fluorescence profile were observed (Si et al., 2014).

Taken together, hairpin RNA, dsRNA, and antisense RNA have been employed for RNAi silencing in *S. cerevisiae* (**Figure 1B**). Upon optimization, all three forms of siRNA precursor could induce strong gene knockdown efficiency (up to 95%). However, the design of hairpin RNA and dsRNA expression cassettes need more optimization, including the length and region of the targeting sequence. In addition, hairpin RNA structures with longer inverted repeats are difficult to build and unstable during cloning, probably due to interference with DNA replication (Voineagu et al., 2008). An intron-containing gap region could

help to improve plasmid stability (Drinnenberg et al., 2009; Crook et al., 2014). Alternatively, full-length antisense RNA represents a simple yet effective RNAi cassette design.

Genome-Scale Library Design

A remarkable advantage of RNAi technology is genome-wide reduction-of-function screening (Boutros and Ahringer, 2008; Mohr et al., 2014). Genome-scale coverage, knockdown efficiency, and off-target effects are important considerations when designing large-scale RNAi experiments (Boutros and Ahringer, 2008; Horn et al., 2010). For heterologous RNAi in *S. cerevisiae*, these characteristics are related to the form of siRNA precursors, the DNA source of RNAi cassettes, the length of inserted fragments, and the library size (**Figure 2**). Due to the difficulty in construction of long inverted repeats and the issue of plasmid stability, it is challenging to generate genome-wide RNAi libraries based on hairpin RNAs. Alternatively, dsRNA and antisense RNA are commonly utilized to create comprehensive RNAi libraries in *S. cerevisiae*.

Pooled long-dsRNA libraries can be created by inserting genomic DNA or complementary DNA (cDNA) fragments between convergent promoters (Cheng and Jian, 2010; Si et al., 2014). Genomic DNA or cDNA fragments can be generated either through endonuclease-mediated, biochemical digestion, such as *Sau*3AI (Si et al., 2014), or using physical shearing method, such as ultrasonication (Crook et al., 2016). Fragmentation parameters should be carefully tuned to achieve a proper length distribution. For example, we previously constructed a strain library of $>3 \times 10^5$ independent clones, where each clone contained ~200 bp genomic DNA fragments. This library achieved a >5-fold redundancy of the 12 Mb yeast genome and hence >99% probability of a full genomic coverage (Si et al., 2014). Also, double-stranded cDNA were fragmented by ultrasonication to an average of 200 or 400 bp to minimize possible translation of open reading frames within dsRNA constructs (Crook et al., 2016). The resultant library contained over 10^5 distinct members and achieved a good coverage of yeast transcriptome (Lee et al., 2016). Notably, due to the heterogeneity of transcript abundance, cDNA normalization via selective degradation of abundant transcripts is necessary to improve genome-scale representation (Zhulidov et al., 2004; Si et al., 2017). On the other hand, genomic DNA-derived dsRNA libraries exhibit an even coverage of each gene, but such libraries will also contain non-transcribed regions that may be inefficient RNAi reagents in most cases. However, this feature may reveal unexpected RNAi targets such as tRNAs (Si et al., 2014), which are not commonly included in cDNA libraries.

Alternatively, antisense RNA libraries can be generated by cloning of cDNA molecules in the reverse direction downstream of a promoter. Based on the SMART (Switching Mechanism At 5′-end of RNA Template) mechanism (Zhu et al., 2001), two different adaptor sequences are included specifically to each end of a full-length cDNA molecule, and hence insertion direction can be controlled by arranging the homologous adaptor sequences in a desirable order in an expression cassette (Si et al., 2017). Upon cDNA normalization based on transcript abundance, a strain library of $>10^6$ independent clones achieved >92% coverage of all yeast genes, which was confirmed via next generation sequencing (NGS) analysis (Si et al., 2017). Taken together, genomic DNA and cDNA libraries can be readily sourced for generating RNAi reagents for genome-wide, reduction-of-function screening in *S. cerevisiae*.

APPLICATION

As one of the most widely used cell factory, *S. cerevisiae* is intensively engineered with different industrially relevant traits, including broadened substrate scope, improved titer, yield and productivity, and enhanced robustness (Kavscek et al., 2015; Lian et al., 2018). Due to the complexity of cellular networks, strain engineering requires concerted identification of genetic targets on a genome scale and fine-tuning of their expression with temporal precision (Si et al., 2015a; Lian et al., 2018). As a modular, scalable, and predictable tool for gene regulation, RNAi has been increasingly applied in yeast engineering, offering unique advantages such as dynamic regulation, genome-scale engineering, and multigene optimization.

Dynamic Regulation

Introduction of synthetic constructs and heterologous pathways typically causes metabolic burden, due to competition for cellular resources and accumulation of toxic intermediates and products (Keasling, 2008; Borkowski et al., 2016). To address this problem, the growth and production phases can be separated via dynamic gene regulation, which can be achieved by inducible RNAi expression.

GAL and Tet are well-characterized inducible promoters in *S. cerevisiae*, modulated by galactose and doxycycline, respectively (Blount et al., 2012). As a proof of concept, when a hairpin RNA cassette targeting *URA3* was constructed under the control of P_{GAL1}, addition of galactose reduced cell growth in the absence of uracil, and enabled cell growth in the presence of 5-fluoroorotic acid (5-FOA), which can be converted to a toxic product by Ura3p (Drinnenberg et al., 2009). Inducible expression of RNAi hairpins were also demonstrated using P_{TPGI}, a promoter derived from P_{GAL1} and induced by galactose and anhydrotetracycline (aTc) (Purcell et al., 2018). Although dynamic RNAi regulation was achieved, the use of GAL promoters requires expensive inducers and suffers from leaky expression. Therefore, it is desirable to utilize promoters that are strongly repressed by glucose and activated by a cheap carbon source. Williams et al. found that during long-term fermentation, GFP expression from the *SUC2* promoter was strongly repressed during the growth phase on glucose; once glucose concentration dropped below 5 nM, GFP expression was rapidly and strongly activated, exhibiting a dynamic range of 12.5-fold when the carbon source was switched to sucrose. Notably, dynamic repression of GFP expression was also achieved by coupling a full-length, antisense GFP cassette with the *SUC2* promoter in an RNAi capable strain (Williams et al., 2015b).

Quorum sensing (QS) describes a widespread phenomenon that couples gene expression with population density (Miller and Bassler, 2001; Choudhary and Schmidt-Dannert, 2010),

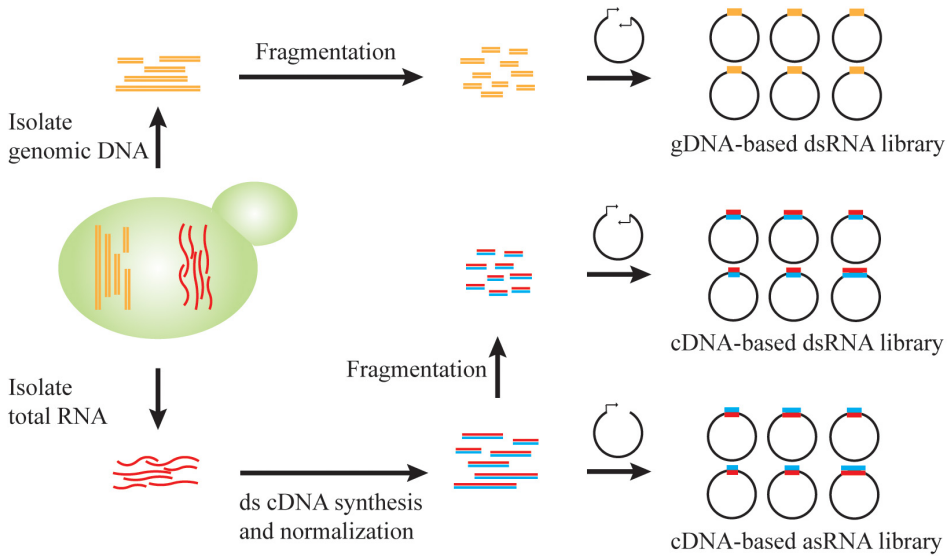


FIGURE 2 | Genome-scale RNAi library construction in *S. cerevisiae*. RNAi library can be constructed from genomic DNA (gDNA) or total RNA. By cloning the fragmented genomic DNA into a vector with convergent promoters, a gDNA-based dsRNA library enables the generation of dsRNA precursors for RNAi pathway. After normalization of transcript abundance, the double-stranded cDNA could be fragmented and cloned into a vector with convergent promoters for a dsRNA library or be directionally cloned into a vector for antisense RNA transcription.

commonly achieved via cell-cell communications mediated by extracellular signaling molecules. QS can help to separate growth phase and production phase, when activation of target pathway genes and repression of competing metabolic activities are only initiated after sufficient biomass is accumulated. As a proof of concept, Williams et al. designed a synthetic QS circuit (Williams et al., 2013, 2015a), where an aromatic amino acid (i.e., tryptophan) induces expression and subsequent secretion of α-pheromone peptide. When extracellular pheromone accumulates to a certain threshold due to increasing cell densities, gene expression from the *FUS1J2* promoter is triggered and results in cell cycle arrest at the G1 phase (**Figure 3**). To achieve QS-regulated gene repression, RNAi constructs of *ARO7*, *PYK*, and *CDC19*, whose knockdown improves production of promoter para-hydroxybenzoic acid (PHBA) but impairs cell fitness, was expressed under the control of *FUS1J2* promoter (**Figure 3**). In an RNAi capable strain, QS-coupled RNAi regulation achieved 1.07 mM PHBA titer, which represents a 37-fold increase over the base strain without the QS circuit (Williams et al., 2015a). This work demonstrates the capability of QS-linked RNAi circuit in dynamic and multiplex gene repression for metabolic engineering.

Genome-Scale Regulation

Genome-scale engineering creates microbial strain libraries for large-scale genotype-phenotype mapping (Si et al., 2015b). For *S. cerevisiae*, non-essential genes have been individually deleted for functional genomic studies (Winzeler et al., 1999; Giaever et al., 2002). But gene-knockout libraries are only available in certain laboratory strains, because it is traditionally tedious to replace each gene sequence with a selection marker via homologous recombination (Scherens and Goffeau, 2004). It is necessary to perform genome-wide screening in custom genetic background due to substantial phenotypic variations among different strains (Van Dijken et al., 2000). Thanks to its *trans-acting* nature, genome-scale knockdown is readily achieved by delivery of RNAi reagents to a select strain on episomal plasmids or via genomic integration. RNAi screening is therefore widely applied to improve a range of target traits in *S. cerevisiae*.

Using a long dsRNA library constructed with digested genomic DNA and convergent promoters, we performed the first genome-scale RNAi screening in *S. cerevisiae* (Si et al., 2014). We first validated the screening process by identification of known and new knockdown suppressors (*ret1*, *say1*, *ssa1*, *cst6*, and *mlp2*) of *yku70*Δ mutation, which causes yeast growth arrest at elevated temperature. Notably, *RET1* is an essential gene and hence cannot be identified using gene deletion collections. Moreover, inspired by directed protein evolution, we performed iterative rounds of RNAi screening for continuous improvement of acetic acid tolerance via accumulation of beneficial silencing cassettes, one at a time in each round (**Figure 4A**). Using this RNAi-assisted genome evolution (RAGE) method, we identified three knockdown targets (*ptc6*, *ypr084w*, and $tRNA^{Val(AAC)}$) that synergistically enable yeast to grow in the presence of 1.0% (v/v) HAc. Similarly, the genomic DNA-derived RAGE library was applied to identify *siz1* as a new gene knockdown target to improve furfural tolerance in *S. cerevisiae* (Xiao and Zhao, 2014).

To achieve multi-mode, genome-scale engineering, we designed RAGE 2.0 libraries that permitted genome-wide overexpression and knockdown screening in a single step (Si et al., 2017; **Figure 4B**). Briefly, a normalized, full-length-enriched cDNA library was first constructed to cover >92% of yeast genes. To the 5′ and 3′ ends of cDNA molecules, we attached two different adaptor sequences that allowed directional

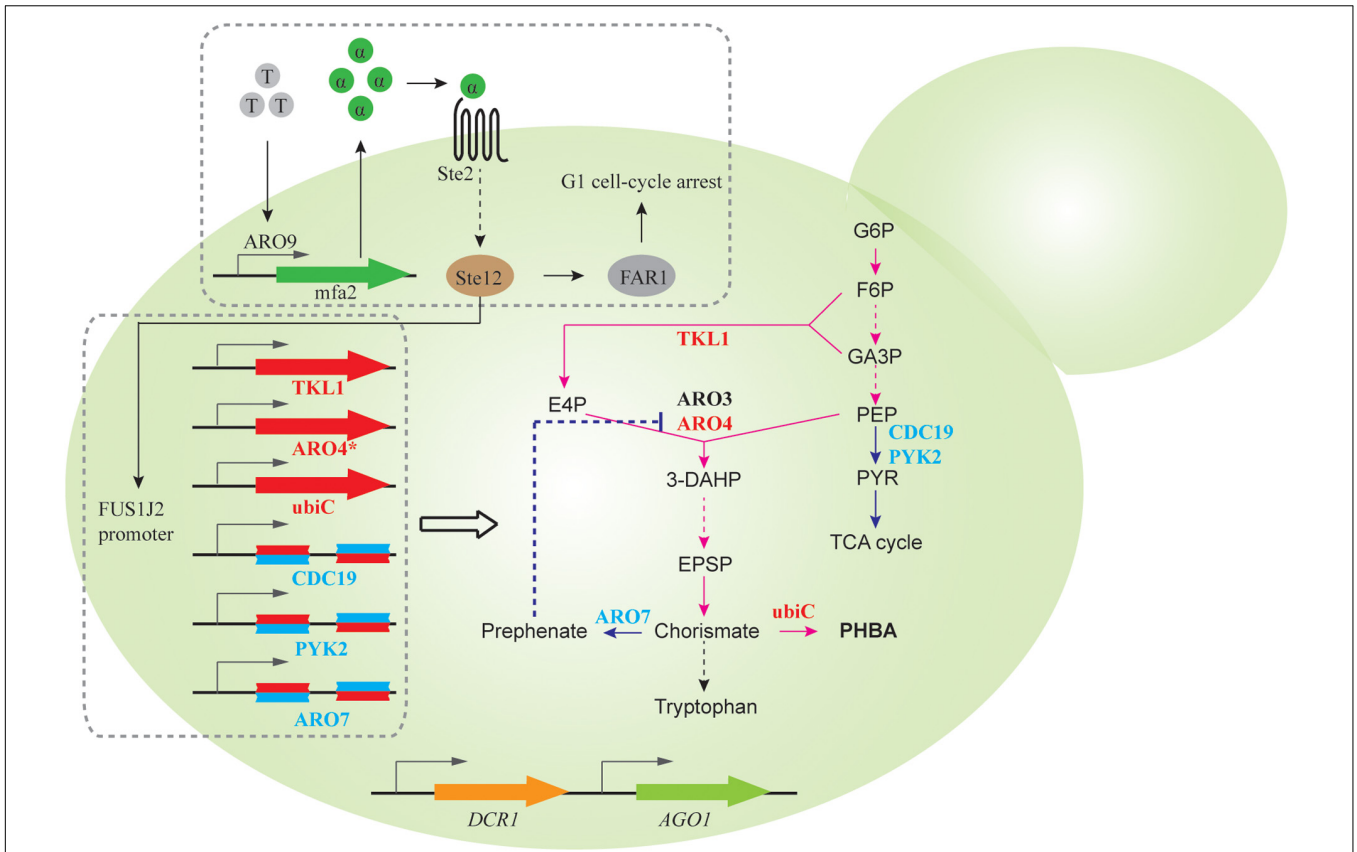


FIGURE 3 | QS-linked, dynamic RNAi regulation during PHBA production. When the population density is high, the extracellular concentration of tryptophan-induced ARO9 promoter-dependent α-pheromone is high, which triggers the mitogen activated protein kinase (MAPK) phosphorylation cascade. The activated Ste12 transcription factor promotes the expression of FAR1, which leads to G1 cell-cycle arrest, and the expression of FUS1J2 promoter-mediated expression of *TKL1*, *ARO4*^{K229L} and *ubiC* gene, and RNAi cassettes targeting *CDC19*, *PYK2*, and *ARO7* gene, which enables the optimization of the PHBA pathway in the presence of a synthetic RNAi pathway.

cloning of cDNAs under a strong constitutive promoter P_{TEF1}. Therefore, gene overexpression was achieved when cDNA molecules were transcribed in the sense direction, whereas gene repression was achieved through transcription of antisense RNAs in the presence of a heterologous RNAi pathway. Using this collection of modulation parts, both gene overexpression and knockdown targets were successfully identified in laboratory and industrial yeast strains for improved cellulase expression, isobutanol production, and xylose utilization (Si et al., 2017; HamediRad et al., 2018). Notably, in addition to isolating individual mutant strains, genome-wide mapping of gene–trait relations can be also achieved. We targeted glycerol substrate utilization, and improved mutants were enriched during serial transfer in glycerol media. Before and after growth enrichment, NGS was applied to track frequency changes of individual modulation parts, which was used to calculate enrichment ratios of corresponding mutants. In this way, the functional impact of every RAGE2.0 overexpression or knockdown mutation on glycerol utilization was quantified (Si et al., 2017).

Furthermore, the standardized and modular design of RAGE2.0 parts enabled accumulation of multiplex, genome-wide mutations in an evolving yeast genome (Si et al., 2017). We utilized the yeast genomic δ sites as landing pads for integrating multiple parts, where genomic integration efficiency was greatly enhanced by introduction of specific double-stranded breaks within δ sequences using the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) system (**Figure 4B**). We then designed a standardized process where strain library creation and screening can be iteratively and robotically performed. After three rounds of selection under increasing HAc concentrations of 0.9, 1.0, and 1.1% (v/v), a mutant strain harboring at least 26 different modulation parts was isolated, indicating the success in creating multiplex genomic diversity. This strain was able to ferment glucose into ethanol in 1.1% (v/v) HAc media, which is the highest reported HAc tolerance to the best of our knowledge.

Alternatively, the Alper group pioneered in creating cDNA-derived RNAi libraries for tunable, genome-wide screening in *S. cerevisiae* (Crook et al., 2016). Using this library, Hsp70 family proteins were identified as key regulators of isobutanol tolerance. Reduced expression of these proteins increased biomass accumulation by 64% under 12 g/L isobutanol stress. Moreover, two rounds of screening identified two

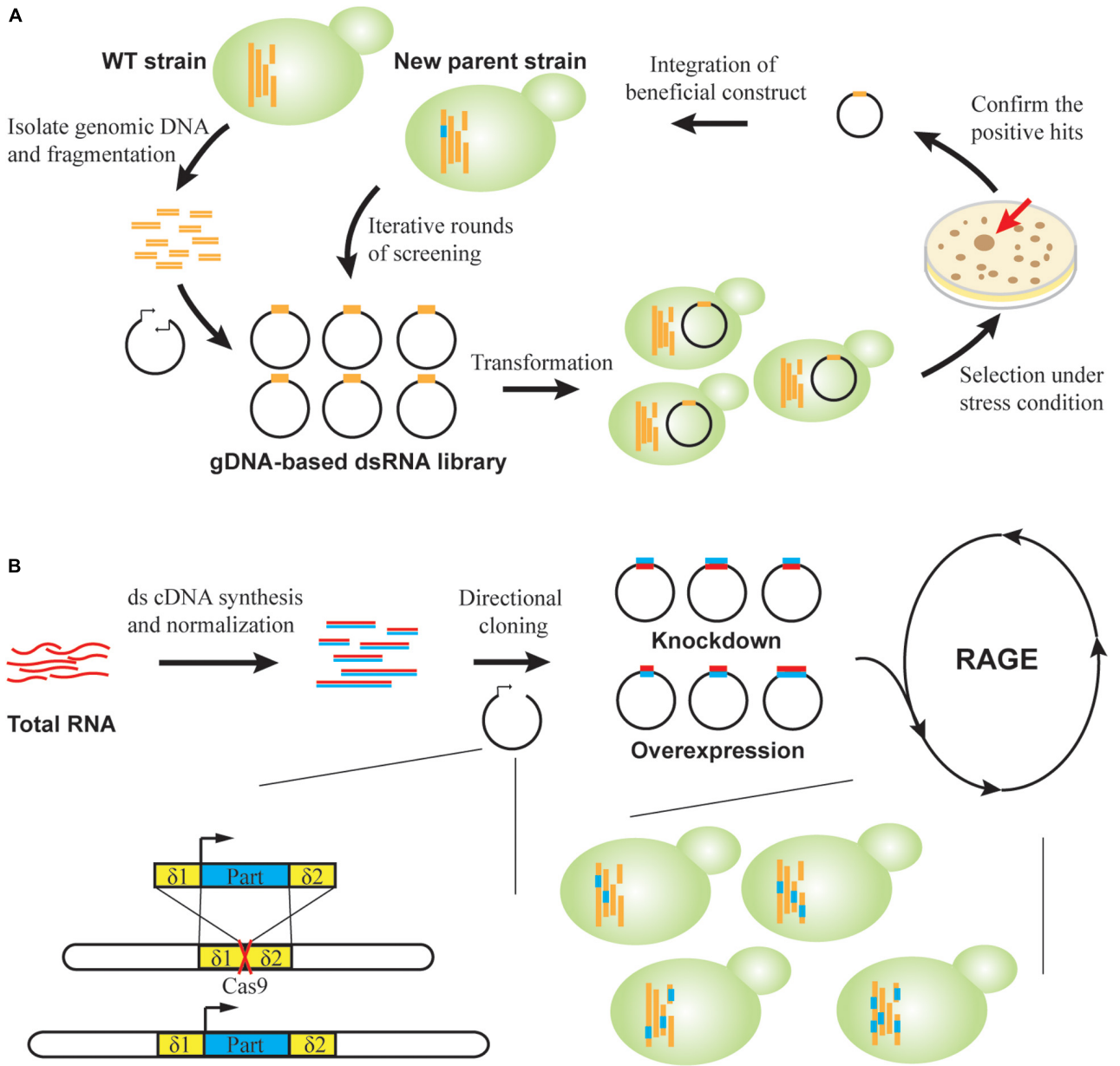


FIGURE 4 | RNAi-assisted genome evolution in *S. cerevisiae*. **(A)** RAGE1.0. A dsRNA-encoding library is constructed by inserting genomic DNA fragments into convergent promoters. Strain libraries are created and screened on agar under selective pressure. For confirmed positive clones, corresponding beneficial RNA cassettes are genomically integrated to generate new parents for a subsequent round of selection. **(B)** RAGE2.0. Directional cloning creates overexpression and knockdown part libraries in a single step from cDNA libraries. Modulation parts are flanked by homologous δ sequences to direct multiplex genomic integration, which is facilitated by CRISPR-Cas. A standardized process of library creation and screening permits robotic iterations for directed genome evolution.

RNAi reagents targeting alcohol dehydrogenase and enolase. When combined, these two mutations exhibited a 31-fold improvement in cell growth in the presence of 10 g/L 1-butanol. Furthermore, reduced expression of the ribosome-associated chaperone Ssb1p was discovered to confer a 52% increase in lactic acid tolerance using the same RNAi library (Lee et al., 2016).

In a recent study, the Nielsen group integrated RNAi screening, microfluidic sorting, and Cas9-mediated genome recombineering to establish a workflow for identification and optimization of multiple gene targets (Wang et al., 2019). Two RNAi libraries were constructed, whereas the genomic DNA-derived library and full-length cDNA-derived library exhibited low and high knockdown efficiencies, respectively. Using droplet microfluidics, 1,500,000 cells covering ∼243,000 knockdown effectors were sorted and enriched in three rounds for improved α-amylase secretion, and genes involved in cellular metabolism, protein modification and degradation, and cell

cycle were found to exert substantial impact on recombinant protein production.

Taken together, a variety of genome-scale RNAi libraries were created to screen a wide range of industrially relevant phenotypes in *S. cerevisiae*. These studies demonstrated several unique advantages. First, unlike gene deletion collections that are available only in certain laboratory strains (Scherens and Goffeau, 2004), genome-wide RNAi libraries can be readily introduced in a custom strain background, such as CEN.PK (Si et al., 2014, 2017; Wang et al., 2019), BY4741 (Xiao and Zhao, 2014; Crook et al., 2016; Lee et al., 2016), and industrial strains (Si et al., 2017; HamediRad et al., 2018). Notably, this feature enables the use of directed evolution strategies in genome-wide engineering, whereby beneficial mutations can be continuously accumulated in an evolving yeast genome (Si et al., 2014). Second, the multiplex nature of RNAi silencing permits simultaneous modulation of many gene targets, whose synergistic actions may lead to superior traits (Si et al., 2014, 2017; Wang et al., 2019). Third, mediating mRNA degradation, RNAi is particularly suited to screen gene silencing targets in polyploid industrial strains, whereby multiple gene copies often produce the same transcript (Si et al., 2017; HamediRad et al., 2018). Fourth, the use of RNAi expands the range of gene regulation in genome-scale screening, whereby not only various repression levels can be achieved by using different RNAi designs (Wang et al., 2019), but also inclusion of both overexpression and knockdown mutations in a single genome-wide library has been realized (Si et al., 2017).

On the other hand, there are certain limitations of genome-scale RNAi screening in *S. cerevisiae*. Due to the random nature of genomic DNA and cDNA-derived libraries, it is essential to perform necessary quality control to ensure sufficient genomic coverage and representation are achieved (Si et al., 2017). Also, due to the prevalence of off-target effects (Jackson et al., 2003), RNAi targets identified from large-scale screening need to be further confirmed by complementary approaches. For example, one can investigate whether the observed phenotype can be generated by a second RNAi reagent designed to target a distinct region of the same gene relative to the isolated RNAi (Si et al., 2014), as it is highly unlikely that two independent RNAi sequences exhibit the same off-target effect. Moreover, most RNAi screening studies in *S. cerevisiae* target "easy-to-screen" phenotypes such as chemical tolerance (Si et al., 2014, 2017; Xiao and Zhao, 2014; Crook et al., 2016; Lee et al., 2016) and substrate utilization (Si et al., 2017; HamediRad et al., 2018), because desirable mutant strains can be facilely identified by growth-based selection or enrichment. The integration of robotics and microfluidics will help to expand the scope of phenotypes that can be engineered using genome-scale RNAi screening (Si et al., 2017; Wang et al., 2019). Finally, for multiplex gene silencing (Si et al., 2017), simultaneous introduction of many RNAi reagents may saturate the RNAi machinery and resulted in inefficient silencing of individual targets, but such complication has not been rigorously assessed. Also, it is inherent challenging to speculate how multiple mutations may confer an improved trait, so that mechanistic studies are necessary to analyze the functional impact of single mutations and their interactions.

CONCLUSION AND FUTURE DIRECTIONS

RNA interference has long been a versatile tool for studying eukaryotic functional genomics studies but is absent from *S. cerevisiae*. Recent advances have greatly expanded the toolkit and application of RNAi in this model yeast, exhibiting many advantages including efficiency, specificity, graded and dynamic repression, genome-scale coverage, and multiplex targeting. Currently, a few groups have been routinely utilizing RNAi as a functional genomics and strain engineering tool in *S. cerevisiae*, although other potential applications such as genetic circuit design (Purcell et al., 2018) and gene delivery (Duman-Scheel, 2019) are also explored. One major factor limiting a wider use of RNAi by the *S. cerevisiae* communities is the availability of ample gene regulation tools in this model organism. Other gene repression methods include but are not limited to promoter engineering (Nevoigt et al., 2006), mRNA destabilization via 3′ insertion (Breslow et al., 2008), and CRISPR interference (CRISPRi) (Gilbert et al., 2013). However, the ease of creating genome-wide libraries directly from genomic DNA and cDNA molecules makes RNAi an attractive tool even in the era of CRISPR/Cas. Creating genome-wide deletion libraries by CRISPR/Cas requires genome sequencing data, algorithm to predict and rank guide RNA (gRNA) sequences, as well as introduction of homologous repair donors for efficient genome editing (Bao et al., 2018; Sharon et al., 2018). Using protein fusions between nuclease-dead mutants of Cas9 (dCas9) and transcriptional repressors, CRISPRi was successfully developed for genome-scale reduction-of-function screening in *S. cerevisiae* (Lian et al., 2019). Many prerequisites of CRISPR/Cas are not always available in polyploid industrial strains, but they do not limit the use of RNAi (HamediRad et al., 2018). Nonetheless, the ability to prototype gRNA collections based on array-synthesized oligo pools confers a greater potential of CRISPRi over RNAi, because the coverage and efficacy of gRNA libraries may be improved in "design-build-test" iterations. Notably, because both genome-scale CRISRPi and RNAi screening are currently performed in pooled formats, so that the silencing efficiency and off-target effect of individual reagents (gRNA or siRNA) remains unknown, which warrants detailed characterization in future studies.

Certain cautions should be taken during the design and implementation of RNAi in *S. cerevisiae*. The length and targeting position of hairpin RNA or dsRNA cassettes exhibit substantial impact on RNAi efficiency (Crook et al., 2014, 2016; Si et al., 2014), but there is currently no design rules. In addition, the expression level and mRNA structure of target gene may affect RNAi efficacy. For example, inefficient repression of *ZWF1* was attributed to the interference from mRNA secondary structure (Williams et al., 2015a). Moreover, RNAi often suffers from off-target effects, which may be particularly severe due to longer dsRNA precursors in budding yeasts for siRNA

production (Weinberg et al., 2011; Wilson and Doudna, 2013; Ipsaro and Joshua-Tor, 2015). Systematic experimental evaluation and computer-aided design may help to understand and avoid the abovementioned issues (Lagana et al., 2014).

To further improve RNAi tools in *S. cerevisiae*, a more potent RNAi pathway can be engineered. For example, Xrn1 orthologs were reported to be involved in multiple steps of RNAi in *Arabidopsis thaliana*, *Caenorhabditis elegans*, *Drosophila melanogaster*, and *Homo sapiens* (Souret et al., 2004; Orban and Izaurralde, 2005; Chatterjee et al., 2011; Lima et al., 2016). In a genetic selection to discover novel components of RNAi in *S. castellii*, Xrn1p was found to play multiple roles to enhance silencing efficiency, including affecting the ratio of different types of siRNA precursors and duplexes, increasing siRNA loading into Ago1p, and enhancing degradation of the passenger strand (Getz et al., 2019). It is interesting to see if introduction of *S. castellii* Xrn1p can improve RNAi processing capabilities in *S. cerevisiae*, which may be limiting when multiplex RNAi regulation is performed (Si et al., 2017).

In addition to RNAi, other regulatory RNAs have been or can be potentially implemented for yeast strain engineering, such as CRISPR RNA, antisense non-coding RNAs, and tRNAs (David et al., 2006; Samanta et al., 2006; Donaldson and Saville, 2012; Si et al., 2015a; Bakowska-Zywicka et al., 2016). For example, antisense non-coding RNAs have been revealed for gene-specific transcriptional regulation, targeting *SER3* (Martens et al., 2004), *IME4* (Hongay et al., 2006), *PHO5* (Uhler et al., 2007), *PHO84* (Camblong et al., 2007, 2009; Castelnuovo et al., 2013), *GAL* (Houseley et al., 2008), *IME1* (Van Werven et al., 2012; Moretto et al., 2018), and *CDC28* (Nadal-Ribelles et al., 2014). Distinct mechanisms are employed, including mRNA interference, competition with transcriptional machinery, histone modification, and gene loop (Yamashita et al., 2016; Niederer et al., 2017; Till et al., 2018; Novačić et al., 2020). Systematic examination of antisense transcriptome will provide comprehensive insights into the mechanisms of antisense RNA-dependent gene regulation to guide engineering tool development (Huber et al., 2016; Nevers et al., 2018). With a deeper understanding, we envision the development of versatile RNA tools that are complementary to RNAi for facilitating strain engineering in *S. cerevisiae*.

AUTHOR CONTRIBUTIONS

YC and TS conceived the scope of this work. YC and EG drafted the manuscript. JZ and TS involved in revising and editing the manuscript. All authors approved the submitted version and agreed both to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work.

ACKNOWLEDGMENTS

We apologize to those whose findings are relevant but not cited in this review due to space limitation. We acknowledged the financial support from Shenzhen Institute of Synthetic Biology Scientific Research Program JCHZ20190003.

REFERENCES

Agrawal, N., Dasaradhi, P. V., Mohmmed, A., Malhotra, P., Bhatnagar, R. K., and Mukherjee, S. K. (2003). RNA interference: biology, mechanism, and applications. *Microbiol. Mol. Biol. Rev.* 67, 657–685. doi: 10.1128/mmbr.67.4.657-685.2003

Bakowska-Zywicka, K., Mleczko, A. M., Kasprzyk, M., Machtel, P., Zywicki, M., and Twardowski, T. (2016). The widespread occurrence of tRNA-derived fragments in *Saccharomyces cerevisiae*. *FEBS Open Bio.* 6, 1186–1200. doi: 10.1002/2211-5463.12127

Bao, Z., HamediRad, M., Xue, P., Xiao, H., Tasan, I., Chao, R., et al. (2018). Genome-scale engineering of *Saccharomyces cerevisiae* with single-nucleotide precision. *Nat. Biotechnol.* 36, 505–508. doi: 10.1038/nbt.4132

Blount, B. A., Weenink, T., and Ellis, T. (2012). Construction of synthetic regulatory networks in yeast. *FEBS Lett.* 586, 2112–2121. doi: 10.1016/j.febslet.2012.01.053

Borkowski, O., Ceroni, F., Stan, G. B., and Ellis, T. (2016). Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Curr. Opin. Microbiol.* 33, 123–130. doi: 10.1016/j.mib.2016.07.009

Boutros, M., and Ahringer, J. (2008). The art and design of genetic screens: RNA interference. *Nat. Rev. Genet.* 9, 554–566. doi: 10.1038/nrg2364

Boynton, P. J. (2019). The ecology of killer yeasts: Interference competition in natural habitats. *Yeast* 36, 473–485. doi: 10.1002/yea.3398

Breslow, D. K., Cameron, D. M., Collins, S. R., Schuldiner, M., Stewart-Ornstein, J., Newman, H. W., et al. (2008). A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nat. Methods* 5, 711–718. doi: 10.1038/nmeth.1234

Camblong, J., Beyrouthy, N., Guffanti, E., Schlaepfer, G., Steinmetz, L. M., and Stutz, F. (2009). Trans-acting antisense RNAs mediate transcriptional gene cosuppression in *S. cerevisiae*. *Genes Dev.* 23, 1534–1545. doi: 10.1101/gad.522509

Camblong, J., Iglesias, N., Fickentscher, C., Dieppois, G., and Stutz, F. (2007). Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in *S. cerevisiae*. *Cell* 131, 706–717. doi: 10.1016/j.cell.2007.09.014

Castel, S. E., and Martienssen, R. A. (2013). RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat. Rev. Genet.* 14, 100–112. doi: 10.1038/nrg3355

Castelnuovo, M., Rahman, S., Guffanti, E., Infantino, V., Stutz, F., and Zenklusen, D. (2013). Bimodal expression of PHO84 is modulated by early termination of antisense transcription. *Nat. Struct. Mol. Biol.* 20, 851–858. doi: 10.1038/nsmb.2598

Chatterjee, S., Fasler, M., Bussing, I., and Grosshans, H. (2011). Target-mediated protection of endogenous microRNAs in *C. elegans*. *Dev. Cell* 20, 388–396. doi: 10.1016/j.devcel.2011.02.008

Chendrimada, T. P., Gregory, R. I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., et al. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* 436, 740–744. doi: 10.1038/nature03868

Cheng, X., and Jian, R. (2010). Construction and application of random dsRNA interference library for functional genetic screens in embryonic stem cells. *Methods Mol. Biol.* 650, 65–74. doi: 10.1007/978-1-60761-769-3_5

Choudhary, S., and Schmidt-Dannert, C. (2010). Applications of quorum sensing in biotechnology. *Appl. Microbiol. Biotechnol.* 86, 1267–1279. doi: 10.1007/s00253-010-2521-7

Crook, N., Sun, J., Morse, N., Schmitz, A., and Alper, H. S. (2016). Identification of gene knockdown targets conferring enhanced isobutanol and 1-butanol tolerance to *Saccharomyces cerevisiae* using a tunable RNAi screening approach. *Appl. Microbiol. Biotechnol.* 100, 10005–10018. doi: 10.1007/s00253-016-7791-2

Crook, N. C., Schmitz, A. C., and Alper, H. S. (2014). Optimization of a yeast RNA interference system for controlling gene expression and enabling rapid metabolic engineering. *ACS Synth. Biol.* 3, 307–313. doi: 10.1021/sb4001432

David, L., Huber, W., Granovskaia, M., Toedling, J., Palm, C. J., Bofkin, L., et al. (2006). A high-resolution map of transcription in the yeast genome. *Proc. Natl. Acad. Sci. U.S.A.* 103, 5320–5325. doi: 10.1073/pnas.0601091103

Donaldson, M. E., and Saville, B. J. (2012). Natural antisense transcripts in fungi. *Mol. Microbiol.* 85, 405–417. doi: 10.1111/j.1365-2958.2012.08125.x

Drinnenberg, I. A., Fink, G. R., and Bartel, D. P. (2011). Compatibility with killer explains the rise of RNAi-deficient fungi. *Science* 333:1592. doi: 10.1126/science.1209575

Drinnenberg, I. A., Weinberg, D. E., Xie, K. T., Mower, J. P., Wolfe, K. H., Fink, G. R., et al. (2009). RNAi in Budding Yeast. *Science* 326, 544–550. doi: 10.1126/science.1176945

Duman-Scheel, M. (2019). *Saccharomyces cerevisiae* (Baker's Yeast) as an Interfering RNA expression and delivery system. *Curr. Drug Targets* 20, 942–952. doi: 10.2174/1389450120666181126123538

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806–811. doi: 10.1038/35888

Getz, M. A., Weinberg, D. E., Drinnenberg, I. A., Fink, G. R., and Bartel, D. P. (2019). Xrn1p acts at multiple steps in the budding-yeast RNAi pathway to enhance the efficiency of silencing. *bioRxiv*. [Preprint]. doi: 10.1101/2019.12.12.873604

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Veronneau, S., et al. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418, 387–391. doi: 10.1038/nature00935

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442–451. doi: 10.1016/j.cell.2013.06.044

Gregory, R. I., Chendrimada, T. P., Cooch, N., and Shiekhattar, R. (2005). Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* 123, 631–640. doi: 10.1016/j.cell.2005.10.022

Gutbrod, M. J., and Martienssen, R. A. (2020). Conserved chromosomal functions of RNA interference. *Nat. Rev. Genet.* 21, 311–331. doi: 10.1038/s41576-019-0203-6

HamediRad, M., Lian, J., Li, H., and Zhao, H. (2018). RNAi assisted genome evolution unveils yeast mutants with improved xylose utilization. *Biotechnol. Bioeng.* 115, 1552–1560. doi: 10.1002/bit.26570

Hannon, G. J. (2002). RNA interference. *Nature* 418, 244–251. doi: 10.1038/418244a

Hasunuma, T., and Kondo, A. (2012). Development of yeast cell factories for consolidated bioprocessing of lignocellulose to bioethanol through cell surface engineering. *Biotechnol. Adv.* 30, 1207–1218. doi: 10.1016/j.biotechadv.2011.10.011

Hong, K. K., and Nielsen, J. (2012). Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries. *Cell. Mol. Life Sci.* 69, 2671–2690. doi: 10.1007/s00018-012-0945-1

Hongay, C. F., Grisafi, P. L., Galitski, T., and Fink, G. R. (2006). Antisense transcription controls cell fate in *Saccharomyces cerevisiae*. *Cell* 127, 735–745. doi: 10.1016/j.cell.2006.09.038

Horn, T., Sandmann, T., and Boutros, M. (2010). Design and evaluation of genome-wide libraries for RNA interference screens. *Genome Biol.* 11:R61. doi: 10.1186/gb-2010-11-6-r61

Houseley, J., Rubbi, L., Grunstein, M., Tollervey, D., and Vogelauer, M. (2008). A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. *Mol. Cell* 32, 685–695. doi: 10.1016/j.molcel.2008.09.027

Huber, F., Bunina, D., Gupta, I., Khmelinskii, A., Meurer, M., Theer, P., et al. (2016). Protein abundance control by non-coding antisense transcription. *Cell Rep.* 15, 2625–2636. doi: 10.1016/j.celrep.2016.05.043

Ipsaro, J. J., and Joshua-Tor, L. (2015). From guide to target: molecular insights into eukaryotic RNA-interference machinery. *Nat. Struct. Mol. Biol.* 22, 20–28. doi: 10.1038/nsmb.2931

Jackson, A. L., Bartz, S. R., Schelter, J., Kobayashi, S. V., Burchard, J., Mao, M., et al. (2003). Expression profiling reveals off-target gene regulation by RNAi. *Nat. Biotechnol.* 21, 635–637. doi: 10.1038/nbt831

Jain, C. K., and Wadhwa, G. (2018). "Computational Tools: RNA Interference in Fungal Therapeutics," in *Current Trends in Bioinformatics: An Insight*, eds G. Wadhwa, P. Shanmughavel, A. K. Singh, and J. R. Bellare (Singapore: Springer), 207–225. doi: 10.1007/978-981-10-7483-7_12

Jiang, S., Si, T., and Dai, J. (2019). Whole-Genome Regulation for Yeast Metabolic Engineering. *Small Methods* 4:1900640. doi: 10.1002/smtd.201900640

Kavscek, M., Strazar, M., Curk, T., Natter, K., and Petrovic, U. (2015). Yeast as a cell factory: current state and perspectives. *Microb. Cell Fact.* 14:94. doi: 10.1186/s12934-015-0281-x

Keasling, J. D. (2008). Synthetic biology for synthetic chemistry. *ACS Chem. Biol.* 3, 64–76. doi: 10.1021/cb7002434

Kildegaard, K. R., Tramontin, L. R. R., Chekina, K., Li, M., Goedecke, T. J., Kristensen, M., et al. (2019). CRISPR/Cas9-RNA interference system for combinatorial metabolic engineering of *Saccharomyces cerevisiae*. *Yeast* 36, 237–247. doi: 10.1002/yea.3390

Lagana, A., Shasha, D., and Croce, C. M. (2014). Synthetic RNAs for Gene Regulation: Design Principles and Computational Tools. *Front. Bioeng. Biotechnol.* 2:65. doi: 10.3389/fbioe.2014.00065

Lee, J. J., Crook, N., Sun, J., and Alper, H. S. (2016). Improvement of lactic acid production in *Saccharomyces cerevisiae* by a deletion of ssb1. *J. Ind. Microbiol. Biotechnol.* 43, 87–96. doi: 10.1007/s10295-015-1713-7

Li, Y., Lu, J., Han, Y., Fan, X., and Ding, S. W. (2013). RNA interference functions as an antiviral immunity mechanism in mammals. *Science* 342, 231–234. doi: 10.1126/science.1241911

Lian, J., Mishra, S., and Zhao, H. (2018). Recent advances in metabolic engineering of *Saccharomyces cerevisiae*: New tools and their applications. *Metab. Eng.* 50, 85–108. doi: 10.1016/j.ymben.2018.04.011

Lian, J., Schultz, C., Cao, M., HamediRad, M., and Zhao, H. (2019). Multi-functional genome-wide CRISPR system for high throughput genotype-phenotype mapping. *Nat. Commun.* 10:5794. doi: 10.1038/s41467-019-13621-4

Lima, W. F., De Hoyos, C. L., Liang, X. H., and Crooke, S. T. (2016). RNA cleavage products generated by antisense oligonucleotides and siRNAs are processed by the RNA surveillance machinery. *Nucleic Acids Res.* 44, 3351–3363. doi: 10.1093/nar/gkw065

Liu, L., Redden, H., and Alper, H. S. (2013). Frontiers of yeast metabolic engineering: diversifying beyond ethanol and *Saccharomyces*. *Curr. Opin. Biotechnol.* 24, 1023–1030. doi: 10.1016/j.copbio.2013.03.005

Loftus, B. J., Fung, E., Roncaglia, P., Rowley, D., Amedeo, P., Bruno, D., et al. (2005). The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. *Science* 307, 1321–1324. doi: 10.1126/science.1103773

Maniataki, E., and Mourelatos, Z. (2005). A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes Dev.* 19, 2979–2990. doi: 10.1101/gad.1384005

Martens, J. A., Laprade, L., and Winston, F. (2004). Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. *Nature* 429, 571–574. doi: 10.1038/nature02538

Miller, M. B., and Bassler, B. L. (2001). Quorum sensing in bacteria. *Annu. Rev. Microbiol.* 55, 165–199. doi: 10.1146/annurev.micro.55.1.165

Mohr, S. E., Smith, J. A., Shamu, C. E., Neumuller, R. A., and Perrimon, N. (2014). RNAi screening comes of age: improved techniques and complementary approaches. *Nat. Rev. Mol. Cell Biol.* 15, 591–600. doi: 10.1038/nrm3860

Moretto, F., Wood, N. E., Kelly, G., Doncic, A., and van Werven, F. J. (2018). A regulatory circuit of two lncRNAs and a master regulator directs cell fate in yeast. *Nat. Commun.* 9:780. doi: 10.1038/s41467-018-03213-z

Nadal-Ribelles, M., Sole, C., Xu, Z., Steinmetz, L. M., de Nadal, E., and Posas, F. (2014). Control of Cdc28 CDK1 by a stress-induced lncRNA. *Mol. Cell* 53, 549–561. doi: 10.1016/j.molcel.2014.01.006

Nevers, A., Doyen, A., Malabat, C., Néron, B., Kergrohen, T., Jacquier, A., et al. (2018). Antisense transcriptional interference mediates condition-specific gene repression in budding yeast. *Nucleic Acids Res.* 46, 6009–6025. doi: 10.1093/nar/gky342

Nevoigt, E., Kohnke, J., Fischer, C. R., Alper, H., Stahl, U., and Stephanopoulos, G. (2006). Engineering of promoter replacement cassettes for fine-tuning of gene expression in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* 72, 5266–5273. doi: 10.1128/AEM.00530-06

Niederer, R. O., Hass, E. P., and Zappulla, D. C. (2017). Long Noncoding RNAs in the Yeast *S. cerevisiae*. *Adv. Exp. Med. Biol.* 1008, 119–132. doi: 10.1007/978-981-10-5203-3_4

Nielsen, J. (2019). Yeast Systems Biology: Model Organism and Cell Factory. *Biotechnol. J.* 14:e1800421. doi: 10.1002/biot.201800421

Novačić, A., Vučenović, I., Primig, M., and Stuparević, I. (2020). Non-coding RNAs as cell wall regulators in *Saccharomyces cerevisiae*. *Crit. Rev. Microbiol.* 46, 1–11. doi: 10.1080/1040841X.2020.1715340

Orban, T. I., and Izaurralde, E. (2005). Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. *RNA* 11, 459–469. doi: 10.1261/rna.7231505

Purcell, O., Cao, J., Muller, I. E., Chen, Y. C., and Lu, T. K. (2018). Artificial repeat-structured siRNA precursors as tunable regulators for *Saccharomyces cerevisiae*. *ACS Synth. Biol.* 7, 2403–2412. doi: 10.1021/acssynbio.8b00185

Samanta, M. P., Tongprasit, W., Sethi, H., Chin, C. S., and Stolc, V. (2006). Global identification of noncoding RNAs in *Saccharomyces cerevisiae* by modulating an essential RNA processing pathway. *Proc. Natl. Acad. Sci. U.S.A.* 103, 4192–4197. doi: 10.1073/pnas.0507669103

Scherens, B., and Goffeau, A. (2004). The uses of genome-wide yeast mutant collections. *Genome Biol.* 5:229. doi: 10.1186/gb-2004-5-7-229

Sharon, E., Chen, S. A., Khosla, N. M., Smith, J. D., Pritchard, J. K., and Fraser, H. B. (2018). Functional genetic variants revealed by massively parallel precise genome editing. *Cell* 175, 544–557.e16. doi: 10.1016/j.cell.2018.08.057

Si, T., Chao, R., Min, Y., Wu, Y., Ren, W., and Zhao, H. (2017). Automated multiplex genome-scale engineering in yeast. *Nat. Commun.* 8:15187. doi: 10.1038/ncomms15187

Si, T., HamediRad, M., and Zhao, H. (2015a). Regulatory RNA-assisted genome engineering in microorganisms. *Curr. Opin. Biotechnol.* 36, 85–90. doi: 10.1016/j.copbio.2015.08.003

Si, T., Xiao, H., and Zhao, H. (2015b). Rapid prototyping of microbial cell factories via genome-scale engineering. *Biotechnol. Adv.* 33, 1420–1432. doi: 10.1016/j.biotechadv.2014.11.007

Si, T., Luo, Y., Bao, Z., and Zhao, H. (2014). RNAi-assisted genome evolution in *Saccharomyces cerevisiae* for complex phenotype engineering. *ACS Synth. Biol.* 4, 283–291. doi: 10.1021/sb500074a

Si, T., and Zhao, H. (2016). RNAi-assisted genome evolution (RAGE) in *Saccharomyces cerevisiae*. *Methods Mol. Biol.* 1470, 183–198. doi: 10.1007/978-1-4939-6337-9_15

Souret, F. F., Kastenmayer, J. P., and Green, P. J. (2004). AtXRN4 degrades mRNA in *Arabidopsis* and its substrates include selected miRNA targets. *Mol. Cell* 15, 173–183. doi: 10.1016/j.molcel.2004.06.006

Suk, K., Choi, J., Suzuki, Y., Ozturk, S. B., Mellor, J. C., Wong, K. H., et al. (2011). Reconstitution of human RNA interference in budding yeast. *Nucleic Acids Res.* 39:e43. doi: 10.1093/nar/gkq1321

Till, P., Mach, R. L., and Mach-Aigner, A. R. (2018). A current view on long noncoding RNAs in yeast and filamentous fungi. *Appl. Microbiol. Biotechnol.* 102, 7319–7331. doi: 10.1007/s00253-018-9187-y

Tokmakov, A. A., Kurotani, A., Takagi, T., Toyama, M., Shirouzu, M., Fukami, Y., et al. (2012). Multiple post-translational modifications affect heterologous protein synthesis. *J. Biol. Chem.* 287, 27106–27116. doi: 10.1074/jbc.M112.366351

Tripathi, N. K., and Shrivastava, A. (2019). Recent developments in bioprocessing of recombinant proteins: expression hosts and process development. *Front. Bioeng. Biotechnol.* 7:420. doi: 10.3389/fbioe.2019.00420

Uhler, J. P., Hertel, C., and Svejstrup, J. Q. (2007). A role for noncoding transcription in activation of the yeast PHO5 gene. *Proc. Natl. Acad. Sci. U.S.A.* 104, 8011–8016. doi: 10.1073/pnas.0702431104

Van Dijken, J. P., Bauer, J., Brambilla, L., Duboc, P., Francois, J. M., Gancedo, C., et al. (2000). An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. *Enzyme Microb. Technol.* 26, 706–714. doi: 10.1016/s0141-0229(00)00162-9

Van Werven, F. J., Neuert, G., Hendrick, N., Lardenois, A., Buratowski, S., van Oudenaarden, A., et al. (2012). Transcription of two long noncoding RNAs mediates mating-type control of gametogenesis in budding yeast. *Cell* 150, 1170–1181. doi: 10.1016/j.cell.2012.06.049

Voineagu, I., Narayanan, V., Lobachev, K. S., and Mirkin, S. M. (2008). Replication stalling at unstable inverted repeats: interplay between DNA hairpins and fork stabilizing proteins. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9936–9941. doi: 10.1073/pnas.0804510105

Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I., and Martienssen, R. A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* 297, 1833–1837. doi: 10.1126/science.1074973

Waldron, F. M., Stone, G. N., and Obbard, D. J. (2018). Metagenomic sequencing suggests a diversity of RNA interference-like responses to viruses across multicellular eukaryotes. *PLoS Genet.* 14:e1007533. doi: 10.1371/journal.pgen.1007533

Wang, G., Björk, S. M., Huang, M., Liu, Q., Campbell, K., Nielsen, J., et al. (2019). RNAi expression tuning, microfluidic screening, and genome recombineering for improved protein production in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 116, 9324–9332. doi: 10.1073/pnas.1820561116

Weinberg, D. E., Nakanishi, K., Patel, D. J., and Bartel, D. P. (2011). The inside-out mechanism of Dicers from budding yeasts. *Cell* 146, 262–276. doi: 10.1016/j.cell.2011.06.021

Williams, T. C., Averesch, N. J. H., Winter, G., Plan, M. R., Vickers, C. E., Nielsen, L. K., et al. (2015a). Quorum-sensing linked RNA interference for dynamic metabolic pathway control in *Saccharomyces cerevisiae*. *Metab. Eng.* 29, 124–134. doi: 10.1016/j.ymben.2015.03.008

Williams, T. C., Espinosa, M. I., Nielsen, L. K., and Vickers, C. E. (2015b). Dynamic regulation of gene expression using sucrose responsive promoters and RNA interference in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* 14:43. doi: 10.1186/s12934-015-0223-7

Williams, T. C., Nielsen, L. K., and Vickers, C. E. (2013). Engineered quorum sensing using pheromone-mediated cell-to-cell communication in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* 2, 136–149. doi: 10.1021/sb300110b

Wilson, R. C., and Doudna, J. A. (2013). Molecular mechanisms of RNA interference. *Annu. Rev. Biophys.* 42, 217–239. doi: 10.1146/annurev-biophys-083012-130404

Winzeler, E. A., Shoemaker, D. D., Astromoff, A., Liang, H., Anderson, K., Andre, B., et al. (1999). Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. *Science* 285, 901–906. doi: 10.1126/science.285.5429.901

Xiao, H., and Zhao, H. (2014). Genome-wide RNAi screen reveals the E3 SUMO-protein ligase gene SIZ1 as a novel determinant of furfural tolerance in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* 7:78. doi: 10.1186/1754-6834-7-78

Yamashita, A., Shichino, Y., and Yamamoto, M. (2016). The long non-coding RNA world in yeasts. *BBA Gene Regul. Mech.* 1859, 147–154. doi: 10.1016/j.bbagrm.2015.08.003

Zhu, Y. Y., Machleder, E. M., Chenchik, A., Li, R., and Siebert, P. D. (2001). Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. *Biotechniques* 30, 892–897. doi: 10.2144/01304pf02

Zhulidov, P. A., Bogdanova, E. A., Shcheglov, A. S., Vagner, L. L., Khaspekov, G. L., Kozhemyako, V. B., et al. (2004). Simple cDNA normalization using kamchatka crab duplex-specific nuclease. *Nucleic Acids Res.* 32:e37. doi: 10.1093/nar/gnh031

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

REVIEW
published: 08 July 2020
doi: 10.3389/fbioe.2020.00691

Regulating Strategies for Producing Carbohydrate Active Enzymes by Filamentous Fungal Cell Factories

Teng Zhang[1], Hu Liu[1], Bo Lv[1] and Chun Li[1,2,3]**

[1] Institute for Synthetic Biosystem/Department of Biochemical Engineering, School of Chemistry and Chemical Engineering, Beijing Institute of Technology, Beijing, China, [2] Key Laboratory of Systems Bioengineering (Ministry of Education), Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), School of Chemical Engineering and Technology, Tianjin University, Tianjin, China, [3] Key Lab for Industrial Biocatalysis, Ministry of Education, Department of Chemical Engineering, Tsinghua University, Beijing, China

Filamentous fungi are important eukaryotic organisms crucial in substrate degradation and carbon cycle on the earth and have been harnessed as cell factories for the production of proteins and other high value-added products in recent decades. As cell factories, filamentous fungi play a crucial role in industrial protein production as both native hosts and heterologous hosts. In this review, the regulation strategies of carbohydrate active enzyme expression at both transcription level and protein level are introduced, and the transcription regulations are highlighted with induction mechanism, signaling pathway, and promoter and transcription factor regulation. Afterward, the regulation strategies in protein level including suitable posttranslational modification, protein secretion enhancement, and protease reduction are also presented. Finally, the challenges and perspectives in this field are discussed. In this way, a comprehensive knowledge regarding carbohydrate active enzyme production regulation at both transcriptional and protein levels is provided with the particular goal of aiding in the practical application of filamentous fungi for industrial protein production.

Keywords: filamentous fungal cell factory, protein expression, carbohydrate active enzymes, transcription factors, carbon catabolite repression, signal pathway, regulation strategies

OPEN ACCESS

Edited by:
Yuan Lu,
Tsinghua University, China

Reviewed by:
Shuobo Shi,
Beijing University of Chemical Technology, China
Konstantinos Vavitsas,
National and Kapodistrian University of Athens, Greece

****Correspondence:***
Bo Lv
lv-b@bit.edu.cn
Chun Li
lichun@tsinghua.edu.cn;
lichun@bit.edu.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *18 April 2020*
Accepted: *03 June 2020*
Published: *08 July 2020*

Citation:
Zhang T, Liu H, Lv B and Li C (2020) Regulating Strategies for Producing Carbohydrate Active Enzymes by Filamentous Fungal Cell Factories. Front. Bioeng. Biotechnol. 8:691. doi: 10.3389/fbioe.2020.00691

INTRODUCTION

In recent years, green biomanufacturing is getting increasing attention because of the serious energy crisis and environment pollution. Searching for green and environmentally friendly ways for industrial production has gradually become a new theme for human beings, such as converting hardly degradable biomass into fermentable sugars and ethanol, as well as replacing fossil energy with environmentally friendly clean energy. However, the transformation of biomass requires a huge number of industrial enzymes, especially carbohydrate active enzymes (CAZymes), such as cellulases and amylases. It is crucial to attain better cost performance or improved properties for these industrial enzymes. Microorganisms show irreplaceable advantages in bioeconomy and green biomanufacturing for their fast growth, short culture period, and low culture cost. Filamentous fungi, which widely exist on earth, play a crucial role in global carbon cycle for their excellent performance in degrading organic matters and converting plant biomass into cost-effective fermentable sugar. In addition, filamentous fungi can also produce a lot of secondary

metabolites that could be used as antibiotics, such as penicillin, or some other drugs that are applied in tumor therapy. Moreover, they can produce many organic acids and other chemical materials.

Altogether, the application of fungal biotechnology enables the development of many industrial fields, such as enzyme and pharmaceutical production, biofuels and biochemistry, food, agriculture, pulp and paper, detergents, textiles, and crop protection. As a consequence, filamentous fungi are getting increasing attention for their major role in industrial production (Dunford, 2012; Deacon, 2013; Benocci et al., 2017; Fang and Qu, 2018; Jiang et al., 2018). As eukaryotic cell factory, filamentous fungi can presumably serve as ideal hosts with rapid growth rate on simple and inexpensive media.

There are several kinds of hosts for protein expression, including both prokaryotic and eukaryotic protein expression system (Altmann et al., 1999; Balbas and Lorence, 2004; Kantardjieff and Zhou, 2013; Schmoll and Dattenböck, 2016; Vega, 2016), and their comparison is listed in **Table 1**. As eukaryotic cell factory, filamentous fungi have numerous advantages that cannot be replaced by other organisms. For instance, filamentous fungi can grow rapidly on simple and cheap media and even fermented or unfermented agroindustrial wastes. They also have a strong survivability, which make them an ideal cell factory for producing drugs, antibiotics, industrial enzymes, and other substances. Most importantly, filamentous fungi have a strong ability in protein expression and perform various posttranslational processing correctly, including glycosylation, peptide chain shearing, and disulfide bond formation, which are similar to mammal cells (Bergquist et al., 2002). Besides, filamentous fungi have a powerful secretory pathway, including signal recognition particle signaling and efficient function of the endoplasmic reticulum in protein modification, as well as rapid clearance of misfolded proteins, fusion between vesicles and target membranes, and apical secretion of the proteins, which conferred them the ability to produce eukaryotic proteins correctly (Kavanagh, 2011; Karagiosis and Baker, 2012; Fang and Qu, 2018).

The common hosts of filamentous fungi are *Aspergillus* species, *Trichoderma* species, and *Penicillium* species, such as *Aspergillus niger*, *Aspergillus oryzae*, *Aspergillus nidulans*, *Trichoderma reesei*, *Penicillium oxalicum*, and other model fungi such as *Neurospora crassa*, all of which can be used for both mechanism investigation and protein expression. Besides, *T. reesei* and *P. oxalicum* are likely to be applied in cellulases expression and plant biomass degradation, whereas *A. niger* and *A. oryzae* are often applied in food industry. What's more, the genetic toolboxes of model filamentous fungi *A. nidulans* and *N. crassa* have been fully developed and used for the investigations of various mechanisms in filamentous fungi, especially for *N. crassa*, as the *N. crassa* Gene Knockout Library is available and brought great convenience for further studies. Filamentous fungi are not suitable for those proteins that are easily produced by other hosts with a considerable yield due to their complex and time-consuming genetic manipulations. As saprophytic fungi, most of the filamentous fungi possess the advantages of biomass degradation. Therefore, filamentous fungi are often used for the expression of CAZymes, which are responsible for the degradation of plant biomass in industrial field, thus providing clean energy by green manufacture.

This review elucidates the regulating strategies in enzyme expression at both transcription level and protein level. Filamentous fungal cell factories produce both endogenous and heterologous enzymes. The filamentous fungi–derived CAZymes, such as cellulases, are often expressed in their native hosts. The expression of these endogenous proteins is regulated at the transcription level to a large extent under the control of carbon catabolite repression (CCR). The promoter and transcription factor regulation mechanisms as well as signal pathways of protein expression in transcription level will be highlighted in detail, which brings a better understanding of the transcription regulation and further applications in the improvement of CAZyme expression. Furthermore, the conventional regulation strategies for improving heterologous expression, such as increasing copy number, codon optimization, protein fusion expressing, and protease reduction, are also introduced briefly. The summarized regulation strategies for enhancing protein expression in filamentous fungi are shown in **Figure 1**.

TRANSCRIPTIONAL REGULATION FOR EFFICIENT PROTEIN PRODUCTION

Most of protein production systems for filamentous fungi cell factories require transformation methods and gene editing strategies, such as vectors, insertion manner, and selection markers, and the present studies and advances of genetic manipulations of filamentous fungi are listed in **Table 2**. Meanwhile, the efficiency of mRNA production and protein expression mostly depends on transcriptional regulation, which includes promoter regulation, CCR regulation, and transcription factor regulation. Herein, various regulation strategies in transcription level for protein expression in filamentous fungi are introduced.

Promoter Regulation

The strategy of using known regulatory elements for protein production is widely applied in fungi cell factory, and the utilization of strong promoters in filamentous fungi can efficiently improve the transcription level of target genes. The strong promoters such as the promoters of glucoamylase gene (*glaA*) in *A. oryzae*, glyceraldehyde-3-phosphate dehydrogenase gene (*gpdA*) in *A. nidulans*, α-amylase gene (*amyB*) in *A. oryzae*, and cellobiose hydrolysis enzyme gene (*cbh1*) in *T. reesei* are most commonly used strong promoters and have been successfully applied in the efficient expression of recombinant proteins.

The original promoters of CAZymes are often inducing promoters, which called for some specific induction conditions. Thus, the strategy of converting the inducing promoter to a strong constitutive promoter is efficient sometimes and has been commonly used in the transcriptional regulation of CAZymes (Su et al., 2012). However, constitutive promoters with strong

TABLE 1 | Comparison of common expression hosts.

Organisms	Growth and culture condition	Genetic transformation	Posttranslational modification	Expression efficiency	Cost
Prokaryote					
E. coli	Fast and high efficiency, simple media requirement	Well-defined, simple, and high efficiency	No posttranslational	High without efficient secretion	Low cost
Bacillus subtilis	Fast, high efficiency, and safe	Convenient for gene modification	Almost none	High yield with secretory expression and produces no lipopolysaccharide	Low cost
Eukaryote					
S. cerevisiae	Fast and high efficiency, easy scale-up	Well-established manipulation	Yes but hyperglycosylation	Moderate and mannosylation of secreted proteins	Low cost
Komagataella pastoris	High cell density, easy scale-up	Well-established manipulation	Yes but hyper-mannosylation	Moderate of secreted proteins	Low cost
Filamentous fungi	Fast and high efficiency	Complex manipulation and lower transformation efficiency	Typical eukaryotic posttranslational modifications	High and efficiency secretion	Low cost
Plant cell	Safe and efficacious	Complex manipulation, long period, and lower transformation efficiency	Tailor-made glycans	High expressing	Cost and potential contamination with microorganisms
Insects	Safe for vertebrates, more demanding culture conditions	Excellent tool for recombinant glycoprotein production	Glycosylation of protein terminal with mannose glycans	High expressing but cannot be expressed continuously	High cost
Mammal cells	Slow growth and expensive nutrient requirement, limited large-scale industrial production	Complicated technology	Proper protein folding, posttranslational modifications	Moderate	High cost and potential contamination with animal viruses

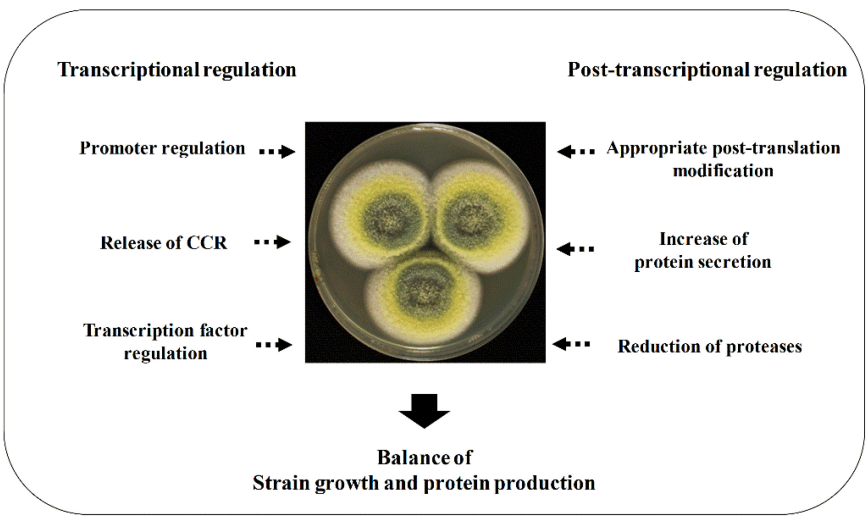


FIGURE 1 | Regulation strategies for enhancing protein expression in filamentous fungi. The enhancement of protein expression at both transcriptional and translational levels is required, and a fermentation optimization is needed to balance the strain growth and protein production.

transcription abilities are not always suitable for the enhancement of proteins as recombinant proteins may be toxic for the growth of the host strain. In this case, inducible promoters would be more preferable as their controllable ability to transfer to the protein expression phase from strain growth phase (Weinhandl et al., 2014).

TABLE 2 | The advances of genetic manipulations of filamentous fungi.

Manipulations	Categories	Introduction	Host strain	References
Vectors	Autonomously replicating vector	Heterologous genes inserting outside chromosome of host cells and replicate independently in an extranuclear manner	Almost all the *Aspergillus species, Rosellinia necatrix*, *Ceriporiopsis subvermispora*	Gupta et al., 2012; Shimizu et al., 2012; Istvan et al., 2017; Li D. et al., 2017; Honda et al., 2019
	Integrated vectors	Foreign genes would be integrated into the genome and be maintained and expressed stably during mitotic and meiotic cell divisions	All the reported filamentous fungi	Istvan et al., 2017; Honda et al., 2019
Transformation methods	PEG/$CaCl_2$-mediated transformation	In the presence of Ca^{2+}, exogenous DNA entered into the host strain by mediation of PEG when protoplasts served as recipient cells	All the reported filamentous fungi	Jain et al., 1992; Zhang et al., 2009
	Agrobacterium tumefaciens mediate transformation	Exogenous genes entered into any recipient cells of the host strain such as protoplasts, mycelium, and even spores in the mediation of *A. tumefaciens*	All the reported filamentous fungi	Michielse et al., 2008; Xu and Bluhm, 2011
Selection marker	Nutrition selection	Genetic transformation selection via the remedy of exogenous substances due to the abnormal synthesis or metabolism pathway of nutrition-deficient strains, such as genes of *niaD*, *glmS*, *argB*, *amdS*, *pyrG*/*pyrF*/*pyr4*/*ura3*/*ura5*	Strain with corresponding deficiency phenotype	Navarrete et al., 2009; Liu et al., 2015; Niu et al., 2016
	Resistance selection	Selection method of strain for their growth under a certain drug concentration and show resistance when resistance gene was transferred into the host strain, such as *Bar*, *Neo*, *Hph*, *BenA*, *Ble*	Strain without corresponding resistance phenotype	Niu et al., 2016; Liu et al., 2019; He et al., 2020

As a consequence, the modification of the existing promoters and new promoters mining would be better choices. The promoter series were also applied for the overexpression of the target gene (Zhang and Xia, 2016). The modification of promoters greatly enhanced the expression of proteins when the binding sites of repressors were replaced by those of activators (Zou et al., 2012; Sun et al., 2020). Native promoters with different strength were mined and used when there were no suitable or enough promoters applied in some non-model strain with industrial value (Liu et al., 2018).

Carbon Catabolite Repression of Filamentous Fungi

As we know, the factors that affect protein expression in filamentous fungi are mainly related to the hierarchy of carbon source utilization of the strain, the signaling sensing and transduction pathways that regulate catabolites, and expression of corresponding enzymes. CCR widely exists in various microorganisms with a regulation system of carbon source utilization, which determines the utilization hierarchy of a huge variety of carbon substrates. CCR ensures the utilization of preferential carbon sources, such as glucose, and inhibits the utilization of less preferred carbon sources by repressing the expression of CAZymes required for the catabolism of a wide range of alternative carbon sources (Kiesenhofer et al., 2016). A large number of genes involved in carbon catabolism are under the control of CCR, including the industrially important CAZymes such as cellulase, amylase, and xylanase (Kunitake et al., 2019). CCR energetically selects the preferential carbon sources, helps microorganisms adapt to the environment by absorbing favorable nutrients at maximum, and reduces the cost of CAZymes synthesis when the preferred carbon source is available, which represents an economical manner for carbon catabolism regulation (Adnan et al., 2017). However, CCR is caused not only by glucose, but also by other monosaccharides. It is reported that xylose also acts as a carbon catabolite repressor, whereas the expression of enzymes for xylose utilization can also be repressed in the presence of glucose (Prathumpai et al., 2004). Besides, the expression of alcohol dehydrogenase (ADH2) of *Saccharomyces cerevisiae* was shown to be inhibited by both glucose and the acetate (Simpson-Lavy and Kupiec, 2019).

The CCR regulation system exists in various fungi and involves several regulators. CCR is mediated by the Mig1 repressor (Kayikci and Nielsen, 2015) in yeast carbon metabolism, while it is often mediated by CreA/Cre1 in most of filamentous fungi. The transcriptional repressor CreA is a C_2H_2 finger domain DNA-binding protein and found to mediate CCR in *A. nidulans* with a transcript of 1.8 kb in length (Dowzer and Kelly, 1989), and the orthologs were identified as CRE1 in *N. crassa* and *T. reesei* with similar functions (Strauss et al., 1995; Serna et al., 1999). CreA mediates CCR with the help of CreB–CreC deubiquitination complex, which plays a crucial role in the function and stability of CreA (Todd et al., 2000;

Lockington and Kelly, 2002; Ries et al., 2016). The subcellular localization of CreA is crucial for derepression of CCR and utilization of carbon source in *A. nidulans* (Ries et al., 2016), and the detailed introduction can be found in next part. CreA functions in repressing transcription of CAZymes via directly binding to 5'-SYGGRG-3' on the promoters of target genes or the transcription activators (Benocci et al., 2017). Besides CCR mediation, CreA also functions in hyphal growth and metabolism in *Aspergillus* species, *T. reesei*, and *Humicola insolens* (Ries et al., 2016; Xu et al., 2019).

As CreA is a repressor for CAZyme expression, strategies of CreA deletion or modification have been used for improving the expression of genes related to carbon utilization. Xylose catabolism was activated in the CreA deletion strain even under high glucose concentration, whereas the major enzymes for xylose utilization were expressed only when glucose repression was relieved in the wild-type strain of *A. nidulans* (Prathumpai et al., 2004). The significantly enhanced expression of cellulase and hemicellulase in *T. reesei* was obtained when the *cre1* gene of the mutant strain was either completely removed or partly truncated, which resulted in the derepressed cellulase expression even in the presence of glucose under both inducing and non-inducing conditions (Nakari-Setälä et al., 2009). However, persistent nuclear localization was obtained when domains of CreA in *A. nidulans* were deleted, which led to a repression of cellulase coding genes under cellulase-inducing conditions (Ries et al., 2016). More interestingly, a truncated Cre1 could turn into an activator, which functioned to activate and enhance the expression of cellulase and xylanase in *T. reesei* without causing any growth deficiencies. The truncated CreA, which served as an activator, exerted its function by locating to the nucleus and directly binding to the upstream regulatory regions of target genes under both inducing and repressing conditions, especially of the main transcription activator of the cellulases, Xyr1 (Rassinger et al., 2018).

As a consequence, the investigation of CCR is important for the regulation and enhancement of endogenous protein expression, whereas for foreign protein expression, choosing a pathway independent of CCR and a strong promoter to make the target protein expression in a constitutive manner might be a more efficient strategy for enhancing its expression level (Zhang et al., 2008).

Signal Pathway in Filamentous Fungi

There are two prerequisites for the expression of endogenous CAZymes in filamentous fungi; one is derepression of CCR, whereas the other one is the presence of inducers. Derepression would be introduced in the following paragraphs, whereas the mechanism of how inducers function in the signaling pathway and the interactions between inducers and elements or factors of the target genes still lack of full investigation. The derepression of CCR is an essential condition for the expression of CAZymes as they can only be induced by corresponding inducers when glucose was depleted in the wild-type strain without any genetic modifications (Gancedo, 1998; Sarma et al., 2007; de Souza et al., 2013; Brown et al., 2015; Wang et al., 2017).

Derepression of CCR is achieved by depletion of favorable carbon sources or a deletion of the crucial factors related to the CCR pathway, such as CreA and other repressors that mediated CCR or PKA pathway [cyclic adenosine monophosphate (cAMP); cAMP-dependent protein kinase A (PKA)], which are responsible for the nuclear location of repressors. The repression and derepression of CCR consist of two crucial pathways, which are the AMP-activated protein kinase (AMPK) pathway and the PKA pathway (Hardie, 2010; Adnan et al., 2017; Lin and Hardie, 2017; Kunitake et al., 2019).

AMPK is regarded as a sensor and regulator of nutritious conditions of the extracellular environment as it can switch on the expression of CAZymes for alternative carbon source utilization. It is activated in the condition of low energy by sensing the cellular adenine nucleotide level, and it was first discovered in mammalian cells for its crucial role in energy sensing to regulate the energy balance of the whole body in a way of inhibiting ATP consumption and accelerating ATP production by switching on alternative catabolic pathways (Lin and Hardie, 2017). G protein–coupled receptors (GPCRs), hexokinases, and hexose transporters are all involved in glucose sensing. When high concentration of glucose was sensed by the corresponding receptors, the glucose was transported into the cell and participated in glycolysis, leading to an increased ATP level and a reduced AMP/ATP ratio, which resulted in a silenced mode of AMPK. With the consumption of glucose and ATP, the AMP/ATP ratio was increased, and the AMPK pathway was activated, which further affected the subcellular localization of CAZyme repressors directly or indirectly by exporting repressors from the nucleus, followed by the degradation in cytoplasm (Tanaka et al., 2018). Thus, the regulatory sequences of target genes were released and further bound by transcription activators, which initiated the transcription of the corresponding CAZyme coding genes (Rubenstein, 2008; Brown et al., 2013).

AMPK always exists as heterotrimeric complexes and consists of the catalytic α subunit and the β and γ regulatory subunit. The α subunit of AMPK in *S. cerevisiae*, which is encoded by *snf1* gene with the target of Ser/Thr site, is required for the derepression of CCR (Gancedo, 1998). The repressor Mig1 is regulated by phosphorylation with the protein kinase Snf1 in *S. cerevisiae*, which is the functional homolog of CreA. Snf1 is activated at a low glucose level by glucose sensing and signaling cascades. Mig1 is phosphorylated by Snf1 and removed to cytoplasm during glucose starvation, which led to a liberation of the regulatory sequences of target genes, whereas Snf1 is inactivated by a high extracellular glucose concentration and results in nucleus localization of Mig1. And as a consequence, Mig1 binds to the regulatory element upstream of the target genes and represses their transcription (Adnan et al., 2017). Similar to yeast, the derepression of CCR in filamentous fungi is controlled dominantly by the function of MAPK. Studies show that *snfA* in *A. nidulans*, which encodes Snf1 homolog protein kinases, is required for CreA derepression and cellulase production. The absence of SnfA led to an inactivation of removing CreA away from the nucleus under the inducing condition of growth on cellulose, which indicated that the subcellular localization was an important process responding to the nutrients in the

environment (Brown et al., 2013). Besides, the presence of inducers (less favorable carbon sources) and carbon starvation are also crucial for the activation of nutrient sensing kinase pathways and the release of inducer binding sites for gene induction. And CreA derepression is necessary for the induction of CAZymes (Brown et al., 2013).

The PKA pathway, which plays an antagonistic role in the regulation of CAZymes compared with AMPK pathway, is also involved in the CCR. G protein–coupled receptors, the cellular second messenger cAMP, and adenylate cyclase are all involved in the PKA pathway and affected the downstream targets by phosphorylation cascades (Nogueira et al., 2015; Yang et al., 2018). PKA is essential for CCR as its deletion leads to a misfunction of CCR with impaired nucleus location effect of repressors and partial derepression of CAZymes, which is consistent with the result obtained by the PkaA deletion strain with a varied degree of derepression of cellulase genes under different culture conditions (Kunitake et al., 2019). PKAs are involved in the regulation of various physiological processes including growth, virulence, and metabolism, and responding to the extracellular nutrient status by a series of phosphorylation cascades (Ribeiro et al., 2019). PkaA in *A. nidulans* plays a curial role in the glucose signaling pathway, and the deletion of this protein kinase resulted in an increased secretion of CAZymes but a decreased growth in the presence of inducing carbon sources (Assis et al., 2015). A mutation of amino acid site of Cre1 led to a loss of phosphorylation and the function of DNA binding, which caused a carbon catabolite derepression, and suggested phosphorylation is required for recovering the function of Cre1 from an inactive conformation (Cziferszky et al., 2002).

When a high glucose level was sensed by GPCRs or other related receptors in Ras/cAMP pathway (Wang et al., 2004), Gα subunit (catalytic subunit of G protein) was activated, and further, adenylate cyclase was activated; cAMP level was thus increased via elevated adenylate cyclase activity. The generated cAMP would bind to the regulatory subunit of PKA, which led to the liberation of PKA catalytic subunit. PKA was activated by the release of catalytic subunit and then initiated phosphorylation cascades and transmitted signals to the downstream targets, leading to the nucleus location of repressors such as CreA (Brown et al., 2014; Ziv, 2020). Phosphorylation is essential for the subcellular localization of repressors, which is crucial for repression/derepression and the transcription of the majority of CAZymes. That is, repressors such as CreA would be imported to the nucleus and repress the expression of target genes when a high extracellular glucose level was sensed, while it would also be exported from nucleus into cytoplasm and finally degraded, which led to the derepression of CCR when the energy depletion is sensed by the cell (Tanaka et al., 2018).

The receptors sense the nutrients status and transmit the signal by phosphorylation cascades to the protein kinases including AMPK and PKA pathway, followed by the direct or indirect phosphorylation effects on the specific sites of the repressors with their final subcellular localization. Some results indicated that dephosphorylation of CreA was essential for CCR in *T. reesei* (Cziferszky et al., 2002), whereas other studies declared the dephosphorylation of CreA is also found in derepressing conditions in *A. nidulans* (Alam et al., 2016). Therefore, it cannot be concluded that the role of phosphorylation is repression or derepression of CAZymes. The derepression of CCR mediated by AMPK pathway plays a positive role in the expression of CAZymes by direct or indirect phosphorylation in different hosts (Gancedo, 1998; Rubenstein, 2008; Adnan et al., 2017), but it certainly functioned via mediating the nucleus exportation of repressors. However, PKA pathway–mediated CCR has negative effects on CAZyme expression, which also regulates the catabolism via nucleus location by indirect phosphorylation of repressors as no direct interaction was found between PKA and CreA, and it was deduced that the CreA was phosphorylated via the kinases controlled by PKA, which further mediated the nucleus import of CreA (Ribeiro et al., 2019). Both PKA and AMPK pathways are indispensably involved in nutrients sensing, signal transduction, catabolism, and growth and function synergistically with the counteract effects in the regulation of CAZyme expression. The possible pathway of carbon catabolism regulation is shown in **Figure 2**.

Transcription Factors Regulation

The transcription factors involved in CAZymes regulatory network are important regulatory elements in filamentous fungi. Besides CreA, which mediated CCR, there are many other transcription factors functioning as activators or repressors and playing a crucial role in the transcription regulation of CAZyme expression, most of which are Zn2Cys6-type transcription factors consisting of zinc finger DNA-binding domain and transcription activation/repression domain. The crucial activators functioning in CAZymes regulation are Xyr1 (xylanase regulator 1)/XlnR, CLR1/ClrA, and CLR2/ClrB, ClrC, and CLR-4 and also ACE2, ACE3, and AmyR, whereas the common repressors are ACE1, Rce1, and Xpp1 (xylanase promoter-binding protein 1) (Coradetti et al., 2012).

Xyr1 and its homologs are the main activators of cellulase gene expression, which play important roles in the expression of cellulolytic enzymes in many filamentous fungi, with a general consensus binding motif of the sequence GGCTRR (Castro et al., 2016; Benocci et al., 2017; Jiang et al., 2018). The overexpression of Xyr1 caused a relief from the CCR and thus produced cellulases in a constitutive expression manner, which resulted in the full expression of cellulases even on the non-inducing carbon sources in *T. reesei*, such as glucose and glycerol (Xinxing et al., 2015). AmyR is a key regulator in activating amylase expression and repressing cellulase gene expression in the meantime (Tani et al., 2001). The other activators also functioned together in the regulation of cellulases induction and positively regulated the expression of cellulases, such as CLR1/ClrA and CLR2/ClrB, ClrC, and CLR-4 and also ACE2 and ACE3 (Craig et al., 2015; Raulo et al., 2016; Liu et al., 2018; Zhang et al., 2019).

The factors such as ACE1, Rce1, and Xpp1 (xylanase promoter-binding protein 1) serve as transcriptional repressors of xylanase expression in *T. reesei* (*Hypocrea jecorina*), ACE1 can bind to the *chb1* promoter of main cellulase gene, whose deletion led to an enhancement in the cellulase and xylanase expression (Saloheimo et al., 2003; Portnoy et al., 2011). The

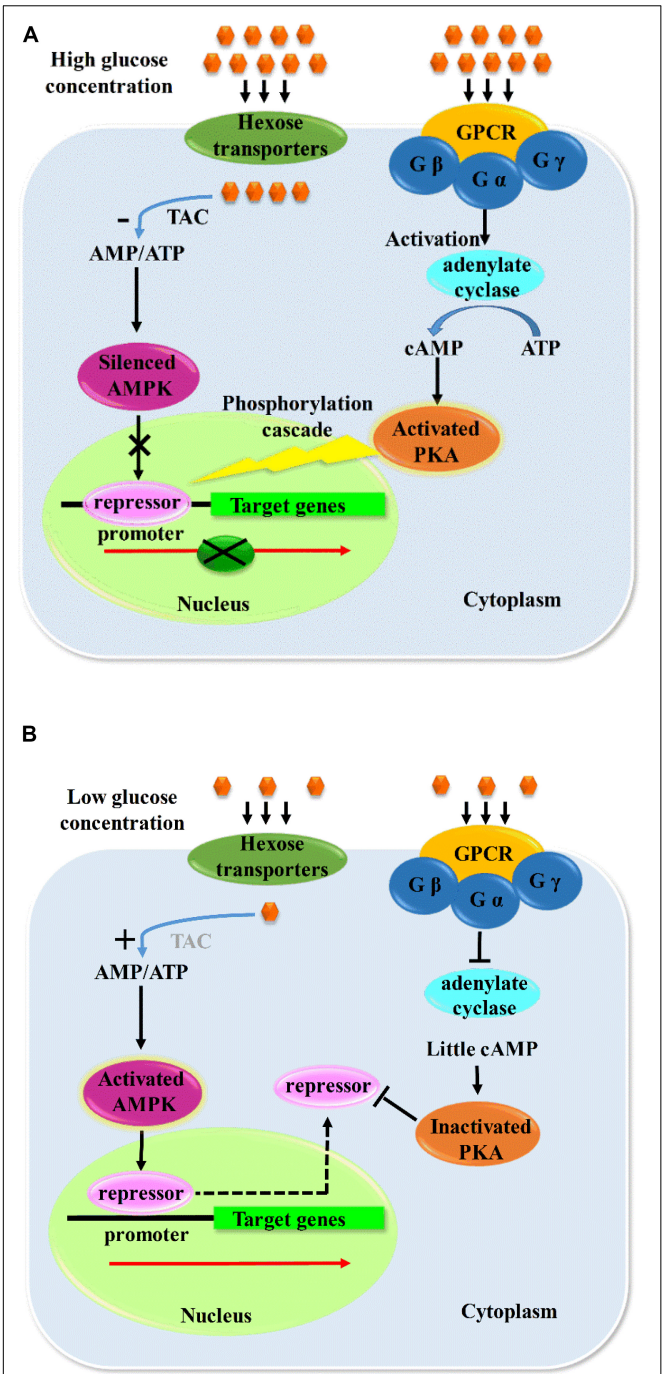


FIGURE 2 | Probable model of carbon metabolism regulation. When the concentration of glucose is high outside the cell **(A)**, G protein–coupled receptor would receive and activate G protein, and catalytic subunit of which would activate adenylate cyclase and result in the increased level of cAMP, and further activate PKA and a series of phosphorylation cascade, which would cause a nucleus location of repressors. In the same time, the high concentration of glucose would accelerate the tricarboxylic acid cycle and increase the concentration of ATP, and the decreased ratio of AMP/ATP would silence the AMPK and make it unable to mediate the export of repressors to cytoplasm from nucleus, thus forming a repression model of CAZymes, and *vice versa* **(B)**. Gα, Gβ, and Gγ are the 3 subunits of G protein; Gα is the catalytic subunit of G protein. TAC, tricarboxylic acid cycle.

promotion of cellulase induction and extension of induction expression process were obtained by the disruption of repressor Rce1 coding gene in *T. reesei* (Cao et al., 2017). Xpp1 could only bind to the promoters of xylanases and regulate the expression of main xylanase, without affecting the expression of cellulases (Mach-Aigner et al., 2010; Derntl et al., 2015, 2017).

Transcription regulation based on transcription factors is a useful tool in improving the CAZyme expression, such as overexpressing the transcription activators or modulating the activators into constitutive expression, and deleting the repressors if their absence does not affect the growth of strain, or downregulating when they do affect the growth. What's more, the use of fused transcription factors to release or attenuate CCR inhibition in cellulase transcription and the modification of the existing transcription repressors to inactive mode, which lose the repression function of target genes and have no influence on the strain growth in the meantime, are promising strategies for enhancing the production of CAZymes (Alazi and Ram, 2018; Wang et al., 2019).

Artificial transcriptional factors that are fused by DNA-binding domain and effectors of different transcription factors were used in cellulase production. The strategy of fusing the ACE2 effector domain with the DNA-binding domains of CRE1 and ACE1 was used to regulate the expression of cellulase (Su et al., 2009). The binding domain of CRE1 was fused to the effector and the binding domain of XYR1, which formed a constituted expression of cellulase based on glucose serving as the sole carbon source (Zhang X. et al., 2016). Enhanced cellulase production was also obtained when a strong transcriptional activation domain was fused to the C-terminus of the natural transcription factors (XYR1, ACE2, and ACE1), followed by the transfection into hypercellulolytic strain and the replacement of natural transcription factors by homologous recombination (Zhang et al., 2018b) or by replacing natural transcription factors with minimal transcriptional activators (Zhang et al., 2018a). Randomized artificial zinc finger protein library, which was constructed via linking multiple zinc finger domain by random shuffling (Park et al., 2003), was used in cellulase expression in *T. reesei* with a significant enhancement of cellulase expression (Zhang F. et al., 2016). An elevated cellulase production was obtained in *T. reesei* when 11 amino acids of the activator ACE3 were truncated, which was probably caused by increasing the interaction with the activator XYR1 (Chen et al., 2020). The modification of ClrB with middle region removal and fusing of DNA-binding/transcriptional activation domains together led to a derepression of CCR; as a consequence, that induction of cellulase in the presence of repression carbon sources such as glucose and glycerol was obtained in *P. oxalicum* without cellulose addition (Gao et al., 2019b). The transcription factor regulation strategies are shown in **Table 3**.

As mentioned previously, various transcription factors are involved in the complicated regulation network of CAZyme expression, but their mechanism investigations are still far from enough. The omics techniques play a crucial role in the investigation of transcription factors, followed by gene cloning and characterization, and their functions were finally identified by the knockout, complementation, overexpression, truncation,

TABLE 3 | Strategies for transcription factors regulation.

CAZymes	Crucial factors	Regulation strategy	Host strain	Achievement	References
Cellulases	CreA, PKA	Double deletion of *creA* and *pkaA*	*A. nidulans*	Elevated cellulases but slow growth	Kunitake et al., 2019
Cellulases	Xyr1	Overexpression of *xyr1*	*T. reesei*	Full expression of cellulases on the non-inducing carbon sources	Xinxing et al., 2015
Cellulases	Cre1 and Xyr1	Overexpression of artificial activator, which fuses Cre1 binding domain to the effector and binding domain of XYR1	*T. reesei*	Constitute expression of cellulase based on glucose	Zhang X. et al., 2016
Cellulases	XYR1, ACE2, and ACE1	Fusing strong activation domain to the C-terminus of the natural transcription factors	*T. reesei*	Enhanced cellulase production	Zhang et al., 2018b
Cellulases	ACE2, Cre1and ACE1	Fusing ACE2 effector domain with the DNA-binding domains of CRE1 and ACE1	*T. reesei*	Elevated cellulases expression	Su et al., 2009
Cellulases	ACE3	Truncation of activator ACE3	*T. reesei*	Elevated cellulases production	Chen et al., 2020
Cellulases	ClrB	ClrB with middle region removal	*P. oxalicum*	Derepression of CCR and induction of cellulase under repression carbon sources	Gao et al., 2019b
Cellulase and hemicellulase	Cre1	Completely removal or partly truncation of *cre1*	*T. reesei*	Cellulases expressed under both inducing or non-inducing condition, even in the presence of glucose	Nakari-Setälä et al., 2009
Cellulases and xylanases	ACE1	Deletion of *ace1*	*T. reesei*	Increased expression of main cellulases genes and two xylanase genes	Saloheimo et al., 2003
Cellulases and xylanases	Cre1	Partially truncation of *cre1*	*T. reesei*	Cre1 turned into an activator of cellulases and xylanases by truncation	Rassinger et al., 2018
Xylanase	Xpp1	Deletion of *xpp1*	*T. reesei*	Elevated expression of xylanase and β-xylosidase	Derntl et al., 2015
Xylanase	CreA	Deletion of *creA*	*A. nidulans*	Xylanases expressed at high glucose concentration in the presence of xylose	Prathumpai et al., 2004
Amylase	AmyR	Overexpression of *amyR*	*Myceliophthora thermophila*	Increase of amylase activity by 30%	Xu et al., 2018
Lignocellulase		Deletion of *amyR*		Relief of CCR and threefold increase of lignocellulase activities	

and even binding properties (Mäkinen et al., 2014; Li et al., 2015). Furthermore, transcriptome data are also necessary for novel regulator screening, and candidate genes were selected based on the combination of genome data and transcriptome data, which were further determined by experimental results, and finally, the underlying regulators that function in regulation of protein expression were uncovered and applied in the production of CAZymes (Derntl et al., 2017; Liao et al., 2018; Zhang et al., 2019). The flowchart of transcription factor mining is shown in **Figure 3**.

Increasing Copy Number

The transcription level of exogenous genes depends more on the transcription efficiency of insertion sites in genome than copy numbers in filamentous fungi when the genome integration strategy was applied (Zoglowek et al., 2014). Unlike heterologous gene expression with high copy number plasmids in *Escherichia coli*, the selective stress is required to ensure their stable inheritance when the extrachromosomal expression manner is used in fungi (Kavanagh, 2011), which caused a less application of increasing copy number in filamentous fungi.

PROTEIN LEVEL REGULATION

The protein regulation includes translation and posttranslational modification, and protein secretion. The strategies during this process are mainly directed to heterologous expression, which are introduced in detail in previous studies (Sharma et al., 2009; Ward, 2011). Here are several strategies for improving protein expression in heterologous host strain.

Increase the Translation Efficiency and Protein Secretion

Filamentous fungi can secrete protein out of the cell efficiently when endogenous genes were expressed, but the secretory ability significantly decreased when some heterologous genes were expressed. The strategy of codon optimization is crucial

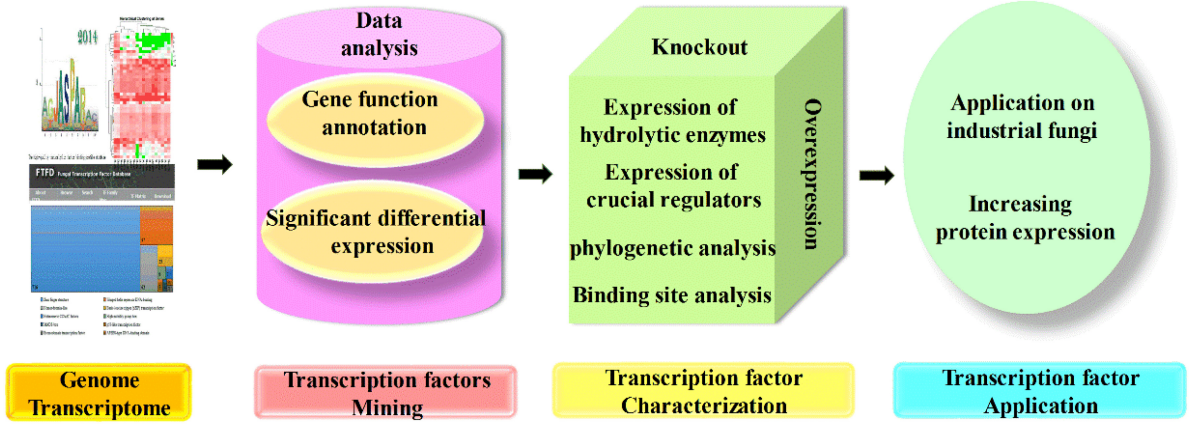


FIGURE 3 | Flowchart of transcription factor mining. The mining of transcription factors often started by genome and transcriptome analysis, and candidate genes could be chosen for further characterization based on gene function and different expression level. The mutants with candidate genes knocked out; complementation and overexpression were constructed, and characterized by the expression of target regulated genes and crucial regulators, also the phylogenetic analysis and binding site analysis, which finally applied on other industrial fungi for the hyperproduction of proteins.

and useful for improvement of translation efficiency. Codon optimizing of heterologous genes according to the codon preference of host strain is an effective way and powerful tool to improve the expression efficiency of heterologous protein in filamentous fungi (Tanaka et al., 2014), which improves the translation efficiency of heterologous genes by the increase of steady-state mRNA level via the elimination of premature polyadenylation and avoidance of mRNA degradation (Li et al., 2007; Tokuoka et al., 2008). The improvement for heterologous protein focuses on the conventional strategies, which are gene fusion with a well-secreted protein and overexpression of foldases and chaperones (Archer et al., 1994). Constructing and selecting host strains with high secretory ability are also a way to improve the expression ability of foreign genes of filamentous fungi (Gombert et al., 2016). Inserting foreign genes into the downstream of the endogenous carrier protein with hyperproduction and high secretion ability by molecular manipulation, constructing the fusion expression system, and expressing the heterologous protein in a fusing manner followed by a cleavage of resulting proteins are also an effective way for improving the level of foreign gene secretion (Gustavsson et al., 2001). Human antibody fragment fused to a truncated endogenous enzyme was expressed in a fourfold protease deletion strain of *N. crassa*, which obtained a good secretion of 3 mg/L (Havlik et al., 2017). For facilitating the subsequent cleavage of the two fused proteins, a linker with proteolytic processing sites should be contained between the carrier protein and the target protein, which should be designed to allow the independent folding of the catalytic domain and the fused protein (Ward, 2011). Besides, the secretion of heterologous protein could also be increased by overexpression of foldases, signal peptides, and chaperones via facilitated protein folding (Goedegebuur et al., 2014; Nevalainen and Peterson, 2014). The application of heterologous signal peptides led to a high level of protein secretion; when enhanced green fluorescent protein served as a model, the heterologous protein human interferon β was finally successfully expressed in *Aspergillus unguis* (Madhavan et al., 2017). The porin B signal peptide in *Corynebacterium glutamicum* was used to improve the production of model protein endoxylanase with activity in high efficiency with a performance of 615 mg/L (An et al., 2013).

Appropriate Posttranslation Modification

It is far from enough to express heterologous protein just in high transcription level and translation efficiency. The industrial and commercial values of the proteins are their structures and functional activities; therefore, appropriate posttranslation modification is crucial for the utilization of the protein produced by heterologous host strains. Unlike the simple prokaryotic expression systems, which are unable to carry out many of the posttranslation processing and lead to the secretion of inactive inclusion bodies, fungal cells have an outstanding performance of these processes. Among all the posttranslation modifications, glycosylation is very crucial for the biological activities of proteins; for example, cellobiohydrolases are typical glycoproteins, which undergo both N-linked glycosylation and O-linked glycosylation by the attachment of oligosaccharides (Stals et al., 2004; Ward, 2011). However, the overglycosylation of proteins would negatively influence enzyme activities, such as enzyme binding and protein stability. Choosing appropriate hosts for the expression of recombination protein is important because the level of glycosylation depends on host strain in a large extent. Generally speaking, the overglycosylation of filamentous fungi is less extensive than that in yeast, especially *S. cerevisiae* (Zoglowek et al., 2014). In some cases, strategies for reducing some extent of glycosylation are necessary for increasing the biology activities of proteins by site mutation, enzymatic deglycosylation, or even using glycosylation deficient strain (Kavanagh, 2011). In a word, an appropriate host strain and a reliable control system of glycosylation are both crucial during activation of heterologous protein.

TABLE 4 | Landmark studies of CAZymes.

CAZymes	Expression strategy	Expression level	Host strain	References
Cellulase	Truncation of functional allele of homolog catabolite repressor, Mig1	Maximum secreted protein titer were more than 14 g/L	*Penicillium funiculosum*	Randhawa et al., 2018
Cellulase	Deleting coding genes of β-glucosidase and repressor, along with overexpressing activator for blocking intracellular inducer hydrolysis and relieving the repression	Filter paper activity and extracellular protein concentration increased by up to more than 10- to 20-fold	*P. oxalicum*	Yao et al., 2015
Cellulase	Simultaneously disrupting the cellulase regulators such as repressor and protease coding genes	Extracellular secreted protein increased fivefold and lignocellulase activities significantly increased up to 13-fold,compared with the parental strain	*M. thermophila*	Liu et al., 2017
Cellulase	By the truncation of cellulase activator ACE3	Increased cellulase productivity with a maximum filter paper activity titer of 102.63 IU/mL	*T. reesei*	Chen et al., 2020
Cellulase	Deleting the serine/threonine protein kinase Stk12	Sevenfold higher of total cellulase production than that of wild type	*N. crassa*	Lin et al., 2019
Cellulase	Overexpressing endogenous β-glucosidase coding gene with two copies insertion into the chromosome of host strain	Filter paper activity of 47.0 IU/mL and pNPGase (*p*-nitrophenyl-β-glucosidase) activity of 144.0 IU/mL in fed-batch culture on lactose	*T. reesei*	Li C. et al., 2017
β-Glucosidase	Deleting CCR factor Cre1 and overexpressing a heterologous β-glucosidase coding genes	51.3-fold enhancement of β-glucosidase activity with 103.9 IU/mL	*T. reesei*	Gao, 2017
β-Glucosidase	Overexpressing β-glucosidase encoding genes	β-glucosidase activity was improved up to 65-fold with a level of 150 U/mL	*P. oxalicum*	Yao et al., 2016
Xylanase and β-glucosidase	Double deleting of the repressors CreA and CreB	Increased xylanase and β-glucosidase activities of more than 100-fold	*A. oryzae*	Ichinose et al., 2017
Thermostable xylanaseB	Heterologous expressing thermostable xylanase B in *A. niger* with endogenous strong promoter, signal sequence, and prosequence	Maximal enzymatic activity is 625 U/mL fermentation supernatant when Remazol Brilliant Blue R-D-Xylan was used as substrate	*A. niger*	Zhang et al., 2008
Amylase	Quintuple mutant modifying with the strategy of overexpressing and deleting several amylase regulators	The protein productivity and amylase activity of mutant strain were increased by 12.0- and 8.2-fold compared with wild type	*M. thermophila*	Li et al., 2020
Xylanase	Deregulating the expression of xylanase transcriptional activator XlnR and modulating the activity of the pH regulator PacC	200-fold increased xylanase activity	*A. nidulans*	Tamayo-Ramos and Orejas, 2014
α-L-Rhamnosidase		Increased α-L-rhamnosidase activity by 19-fold than that of control		
α-L-Arabinofuranosidases	Mutating alanine residue to valine of arabinose regulator AraR and overexpressing the mutant regulator	54.1-fold increase of expression	*P. oxalicum*	Gao et al., 2019a
Trehalase	Heterologous expressing high active trehalase coding gene in *A. niger* with an expression strategy of multi-copy knock-in	Titer of trehalase was up to 1698.83 U/mL	*A. niger*	Dong et al., 2019

Reduction of Proteases

The occurrence of some incorrect posttranslational processing in the heterologous protein expression during the cell growth of filamentous fungi is unavoidable, such as misfolding, impairment of intracellular transport, and proteolytic degradation. Proteolytic degradation is one of the most obvious reasons for the low yields of heterologous proteins (Chung et al., 2002). Proteolytic degradation occurs not only in intracellular but also extracellular by the function of endogenous proteases produced during fungal growth (Zoglowek et al., 2014).

There are several strategies in reducing the proteolysis in host cells. Deletion of the identified protease coding genes could bring a satisfactory cellulase production compared with the parent strain (Li et al., 2019). The application of protease-deficient strains for improving the production of heterologous proteins is efficient and commonly used, although it would also bring a reduction of industrial protein production in the meantime (Wang et al., 2005). In this case, the strategy of partially inactivating some of the more prominent extracellular proteases, such as alkaline proteases or metalloproteases, which were determined in the strain, could be chosen for improving the protein production (Ward, 2011). Protease inhibitors were also used to control protease activity and reduce the proteolysis, but they could only be applied in a small-scale protein expression system (Wiebe, 2003). Besides, disruption of the coding genes of some protease regulators could also work in reducing protease activity in *Aspergillus* species. What's more, adjusting the pH of fungal culture away from optimal pH for proteases activities and inhibiting the proteolysis could reduce the degradation of recombination protein (Braaksma and Punt, 2008).

CHALLENGES AND PROSPECTS FOR FILAMENTOUS FUNGI CELL FACTORY

As stated above, there are so many advantages of filamentous fungal cell factory in producing industrial enzymes, and some achievements were obtained with a large amount of protein expression, which indicates filamentous fungi are a kind of efficient hosts for the production of CAZymes. The landmark studies of CAZymes are listed in **Table 4**. However, the mining of genome and the efforts for simplifying genetic technology are far from enough of non-model strains, such as lower transformation efficiency, multinuclear cells, lack of knowledge for available genetic engineering elements, exogenous gene insertion sites, and so on (Gupta et al., 2012; Nevalainen, 2020). It is difficult to improve the level of protein expression by a single genetic modification method. Fungal cells are always regulated by the CCR and could not express most of the CAZymes with the existence of readily available carbon sources such as glucose, whereas the lack of glucose cannot enrich the biomass of the strain, which leads to a low production of target proteins. Besides, the complex protease system of filamentous fungal cells makes them difficult to accumulate and secrete the expression products of foreign genes, resulting in low yields of target products. When some non-fungal genes from bacteria, mammals, or plants are expressed in fungal hosts, the level of protein expression would be much lower than that of fungi sourced genes (Broekhuijsen et al., 1993).

In addition, the process of protein expression often includes two stages of strain growth and protein induction, and the strain needs to reach the maximum biomass in the growth stage by utilizing some favorable carbon sources, whereas the favorable carbon sources should be depleted, and the corresponding inducers should be added for the initiation of protein induction stage. It is obvious that monitoring the status of carbon source consumption and control of fermentation process are inconvenient. There are also some filamentous fungi that form large mycelium pellets during fermentation, which would reduce the contact between fungal cells and the surrounding medium, hindering the efficient expression and secretion of proteins (Nevalainen, 2020). The transcription regulation of CAZymes in filamentous fungi mainly focuses on the modulation of signaling pathway or modification of transcription factors, to make CAZymes express in the presence of glucose and achieve a better growth of strain and substantial biomass, thus obtaining an elevated production of CAZymes.

The strategies such as downregulation or even deletion of genes involved in PKA pathway and repressors, upregulation of AMPK pathway, and overexpressing activators are often efficient in the enhancement of CAZyme expression. However, it should be noted that PKA functions in various physical process including cell growth and metabolism, as well as the repressor CreA; their downregulation might cause deficiency in strain growth and metabolism, which must lead to a weak production of enzymes far from best. The better choice for enhancing the expression of proteins is to keep the balance between cell growth and protein expression, such as modifying a repressor into factors without negative impacts or even positive effects for the target genes or modifying the promoter regions upstream of target genes by replacing the binding sites of repressors to those of activators, and modulate the PKA and AMPK pathway to make a perfect subcellular location of repressors and incapability for repressing the gene transcription without any negative impacts on cell growth.

Moreover, the growth of the strain would change the fluidity of the medium, resulting in changes in the conditions of nutrients, oxygen, and pH; in turn, the ability of secreting foreign protein also varied with the changes of growth status and conditions (Braaksma and Punt, 2008). Therefore, the improvement with multiple tolerances of the strain by metabolic engineering strategies is essential to enhance the stability of strain in various culture environments by the relief of multiple growth stresses and so to increase the production of proteins, which needs further efforts. A further conceptual point with regard to simplifying the genetic procedures and shortening the transformation time with an acceptable transformation efficiency, multiple tolerance host strains, and the balance between cell growth and protein expression would be the main goal for the development of fungi cell factory.

CONCLUSION

In summary, the present article systematically elucidated the strategies of protein expression in both transcription level and protein levels. Transcription regulations with both mechanism and signal pathway of protein expression were illustrated in detail. The strategies for producing industrial CAZymes in filamentous fungal cell factory of current studies were introduced, which are promoters and transcription factor regulation, protein expression and secretion regulation, and the balance of strain growth and protein expression. Although the catabolism regulatory network in filamentous fungal cells is really complex, our understanding of signaling pathways and mechanisms for CAZymes induction and also strategies of protein expression regulation improved the cognition of protein expression in filamentous fungi, which would benefit the investigations of CAZymes or filamentous fungi protein expression system.

AUTHOR CONTRIBUTIONS

TZ was responsible for the literature survey and the writing of whole manuscript. HL was responsible for the revise of manuscript. BL and CL were responsible for supervision. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (Nos. 21736002 and 21706012) and the National Key Research and Development Program of China (2018YFA0901800).

REFERENCES

Adnan, M., Zheng, W., Islam, W., Arif, M., Abubakar, Y., Wang, Z.-H., et al. (2017). Carbon catabolite repression in filamentous fungi. *Int. J. Mol. Sci.* 19:48. doi: 10.3390/ijms19010048

Alam, M. A., Kamlangdee, N., and Kelly, J. (2016). The CreB deubiquitinating enzyme does not directly target the CreA repressor protein in *Aspergillus nidulans*. *Curr. Genet.* doi: 10.1007/s00294-016-0666-3[Epub ahead of print].

Alazi, E., and Ram, A. (2018). Modulating transcriptional regulation of plant biomass degrading enzyme networks for rational design of industrial fungal strains. *Front. Bioeng. Biotechnol.* 6:133. doi: 10.3389/fbioe.2018.00133

Altmann, F., Staudacher, E., Wilson, I., and Maerz, L. (1999). Insect cells as hosts for the expression of recombinant glycoproteins. *Glycoconj. J.* 16, 109–123. doi: 10.1023/A:1026488408951

An, S. J., Yim, S. S., and Jeong, K. J. (2013). Development of a secretion system for the production of heterologous proteins in *Corynebacterium glutamicum* using the Porin B signal peptide. *Protein Expr. Purif.* 89, 251–257. doi: 10.1016/j.pep.2013.04.003

Archer, D., Jeenes, D., and MacKenzie, D. (1994). Strategies for improving heterologous protein production from filamentous fungi. *Antonie Van Leeuwenhoek* 65, 245–250. doi: 10.1007/BF00871952

Assis, L., Ries, L., Savoldi, M., Reis, T., Brown, N., and Goldman, G. (2015). *Aspergillus nidulans* protein kinase A plays an important role in cellulase production. *Biotechnol. Biofuels* 8:213. doi: 10.1186/s13068-015-0401-1

Balbas, P., and Lorence, A. (eds) (2004). *Recombinant gene expression: Reviews and Protocols*, 2nd Edn. Totowa, NJ: Humana Press. doi: 10.1385/1592597742

Benocci, T., Aguilar-Pontes, M. V., Zhou, M., Seiboth, B., and Vries, R. P. (2017). Regulators of plant biomass degradation in ascomycetous fungi. *Biotechnol. Biofuels* 10:152. doi: 10.1186/s13068-017-0841-x

Bergquist, P., Te'o, V., Gibbs, M., Cziferszky, A., de Faria, F., Azevedo, M., et al. (2002). Production of recombinant bleaching enzymes from thermophilic microorganisms in fungal hosts. *Appl. Biochem. Biotechnol.* 98-100, 165–176. doi: 10.1385/abab:98-100:1-9:165

Braaksma, M., and Punt, P. J. (2008). *Aspergillus as a Cell Factory for Protein Production: Controlling Protease Activity in Fungal Production*. Boca Raton, FL: CRC Press.

Broekhuijsen, M., Mattern, I., Contreras, R., Kinghorn, J., and Hondel, C. (1993). Secretion of heterologous proteins by *Aspergillus niger*: production of active human interleukin-6 in a protease-deficient mutant by KEX2-like processing of a glucoamylase-hIL6 fusion protein. *J. Biotechnol.* 31, 135–145. doi: 10.1016/0168-1656(93)90156-H

Brown, N., Gouvea, P., Krohn, N., Savoldi, M., and Goldman, G. (2013). Functional characterisation of the non-essential protein kinases and phosphatases regulating *Aspergillus nidulans* hydrolytic enzyme production. *Biotechnol. Biofuels* 6:91. doi: 10.1186/1754-6834-6-91

Brown, N., Reis, T., Reis, L., Caldana, C., Mah, J.-H., Yu, J.-H., et al. (2015). G-protein coupled receptor mediated nutrientsensing and developmental control in *Aspergillus nidulans*. *Mol. Microbiol.* 98, 420–439. doi: 10.1111/mmi.13135

Brown, N., Ries, L., and Goldman, G. (2014). How nutritional status signalling coordinates metabolism and lignocellulolytic enzyme secretion. *Fungal Genet. Biol.* 72, 48–63. doi: 10.1016/j.fgb.2014.06.012

Cao, Y., Zheng, F., Wang, L., Zhao, G., Chen, G., Zhang, W., et al. (2017). Rce1, a novel transcriptional repressor, regulates cellulase gene expression by antagonizing the transactivator Xyr1 in *Trichoderma reesei*. *Mol. Microbiol.* 105, 65–83. doi: 10.1111/mmi.13685

Castro, L., De Paula, R., Antoniêto, A., Persinoti, G., Silva-Rocha, R., and Silva, R. (2016). Understanding the role of the master regulator XYR1 in *Trichoderma reesei* by global transcriptional analysis. *Front. Microbiol.* 7:175. doi: 10.3389/fmicb.2016.00175

Chen, Y., Wu, C., Fan, X., Zhao, X., Zhao, X., Shen, T., et al. (2020). Engineering of *Trichoderma reesei* for enhanced degradation of lignocellulosic biomass by truncation of the cellulase activator ACE3. *Biotechnol. Biofuels* 13:62. doi: 10.1186/s13068-020-01701-3

Chung, H.-J., Park, S.-M., and Kim, D.-H. (2002). Characterization of *Aspergillus niger* mutants deficient of a protease. *Mycobiology* 30, 160–165. doi: 10.4489/MYCO.2002.30.3.160

Coradetti, S., Craig, J., Xiong, Y., Shock, T., Tian, C., and Glass, N. (2012). Conserved and essential transcription factors for cellulase gene expression in ascomycete fungi. *Proc. Natl. Acad. Sci. U.S.A.* 109, 7397–7402. doi: 10.1073/pnas.1200785109

Craig, J., Coradetti, S., Starr, T., and Glass, N. (2015). Direct target network of the *Neurospora crassa* plant cell wall deconstruction regulators CLR-1, CLR-2, and XLR-1. *mBio* 6:e1452-15. doi: 10.1128/mBio.01452-15

Cziferszky, A., Mach, R., and Kubicek, C. (2002). Phosphorylation positively regulates DNA binding of the carbon catabolite repressor Cre1 of *Hypocrea jecorina* (*Trichoderma reesei*). *J. Biol. Chem.* 277, 14688–14694. doi: 10.1074/jbc.M200744200

de Souza, W., Morais, E., Krohn, N., Savoldi, M., Goldman, M. H., Rodrigues, F., et al. (2013). Identification of metabolic pathways influenced by the G-protein coupled receptors GprB and GprD in *Aspergillus nidulans*. *PLoS One* 8:e62088. doi: 10.1371/journal.pone.0062088

Deacon, J. (2013). *Fungal Biology*, 4th Edn. Chichester: Wiley-Blackwell, 142–157.

Derntl, C., Kluger, B., Bueschl, C., Schuhmacher, R., Mach, R., and Mach-Aigner, A. (2017). Transcription factor Xpp1 is a switch between primary and secondary fungal metabolism. *Proc. Natl. Acad. Sci. U.S.A.* 114, E560–E569. doi: 10.1073/pnas.1609348114

Derntl, C., Rassinger, A., Srebotnik, E., Mach, R., and Mach-Aigner, A. (2015). Xpp1 regulates the expression of xylanases, but not of cellulases in *Trichoderma reesei*. *Biotechnol. Biofuels* 8:112. doi: 10.1186/s13068-015-0298-8

Dong, L., Lin, X., Yu, D., Huang, L., Wang, B., and Pan, L. (2019). High-level expression of highly active and thermostable trehalase from *Myceliophthora thermophila* in *Aspergillus niger* by using the CRISPR/Cas9 tool and its application in ethanol fermentation. *J. Ind. Microbiol. Biotechnol.* 47, 133–144. doi: 10.1007/s10295-019-02252-9

Dowzer, C., and Kelly, J. (1989). Cloning of the creA gene from *Aspergillus nidulans*: a gene involved in carbon catabolite repression. *Curr. Genet.* 15, 457–459. doi: 10.1007/BF00376804

Dunford, N. T. (2012). *Food and Industrial Bioproducts and Bioprocessing*. Hoboken, NJ: Wiley-Blackwell.

Fang, X., and Qu, Y. (2018). *Fungal Cellulolytic Enzymes Microbial Production and Application: Microbial Production and Application*. Berlin: Springer.

Gancedo, J. (1998). Yeast carbon catabolite repression. *Microbiol. Mol. Biol. Rev.* 62, 334–361. doi: 10.1128/MMBR.62.2.334-361.1998

Gao, J. (2017). Production of the versatile cellulase for cellulose bioconversion and cellulase inducer synthesis by genetic improvement of *Trichoderma reesei*. *Biotechnol. Biofuels* 10:272.

Gao, L., Li, S., Xu, Y., Xia, C., Xu, J., Liu, J., et al. (2019a). Mutation of a conserved alanine residue in transcription factor AraR leads to hyper-production of α-L-Arabinofuranosidases in *Penicillium Oxalicum*. *Biotechnol. J.* 14:e1800643. doi: 10.1002/biot.201800643

Gao, L., Xu, Y., Song, X., Li, S., Xia, C., Xu, J., et al. (2019b). Deletion of the middle region of the transcription factor ClrB in *Penicillium oxalicum* enables cellulase production in the presence of glucose. *J. Biol. Chem.* 294, 18685–18697. doi: 10.1074/jbc.RA119.010863

Goedegebuur, F., Neef-Kruithof, P., Pucci, J. P., and Ward, M. (2014). Over expression of foldases and chaperones improves protein production. U.S. Patent No 6,090,051. Washington, DC: U.S. Patent and Trademark Office. doi: 10.1074/jbc.ra119.010863

Gombert, A., Junior, J., Cerdán, M. E., and González-Siso, M.-I. (2016). *Kluyveromyces marxianus* as a host for heterologous protein synthesis. *Appl. Microbiol. Biotechnol.* 100, 6193–6208. doi: 10.1007/s00253-016-7645-y

Gupta, V. K., Tuohy, M., Ayyachamy, M., Turner, K., and O'Donovan, A. (2012). *Laboratory Protocols in Fungal Biology: Current Methods in Fungal Biology*. Berlin: Springer.

Gustavsson, M., Lehtiö, J., Denman, S., Teeri, T., Hult, K., and Martinelle, M. (2001). Stable linker peptides for a cellulose-binding domain-lipase fusion protein expressed in *Pichia pastoris*. *Protein Eng.* 14, 711–715. doi: 10.1093/protein/14.9.711

Hardie, D. (2010). "AMP-activated protein kinase," in *Handbook of Cell Signaling*. Amsterdam: Elsevier.

Havlik, D., Brandt, U., Bohle, K., and Fleißner, A. (2017). Establishment of *Neurospora crassa* as a host for heterologous protein production using a human antibody fragment as a model product. *Microb. Cell Fact.* 16:128.

He, L., Guo, W., Li, J., Meng, Y., Wang, Y., Lou, H., et al. (2020). Two dominant selectable markers for genetic manipulation in *Neurospora crassa*. *Curr. Genet.* doi: 10.1007/s00294-020-01063-1 [Epub ahead of print].

Honda, Y., Tanigawa, E., Tsukihara, T., Nguyen Xuan, D., Kawabe, H., Sakatoku, N., et al. (2019). Stable and transient transformation, and a promoter assay in the selective lignin-degrading fungus, *Ceriporiopsis subvermispora*. *AMB Express* 9:92. doi: 10.1186/s13568-019-0818-1

Ichinose, S., Tanaka, M., Shintani, T., and Gomi, K. (2017). Increased production of biomass-degrading enzymes by double deletion of creA and creB genes involved in carbon catabolite repression in *Aspergillus oryzae*. *J. Biosci. Bioeng.* 125, 141–147. doi: 10.1016/j.jbiosc.2017.08.019

Istvan, W., Yang, L., Vang, J., Ahring, B., Lübeck, M., and Lübeck, P. S. (2017). A comparison of Agrobacterium-mediated transformation and protoplast-mediated transformation with CRISPR-Cas9 and bipartite gene targeting substrates, as effective gene targeting tools for *Aspergillus carbonarius*. *J. Microbiol. Methods* 135, 26–34. doi: 10.1016/j.mimet.2017.01.015

Jain, S., Durand, H., and Tiraby, G. (1992). Development of a transformation system for the thermophilic fungus *Talaromyces* sp. CL240 based on the use of phleomycin resistance as a dominant selectable marker. *Mol. Gen. Genet.* 234, 489–493. doi: 10.1007/bf00538710

Jiang, Y., Liu, K., Guo, W., Zhang, R., Liu, F., Zhang, N., et al. (2018). "Lignocellulase formation, regulation, and secretion mechanisms in *Hypocrea jecorina* (*Trichoderma reesei*) and other filamentous fungi: microbial production and application," in *Fungal Cellulolytic Enzymes*, eds X. Fang, and Y. Qu, (Singapore: Springer), 43–59. doi: 10.1007/978-981-13-0749-2_3

Kantardjieff, A., and Zhou, W. (2013). Mammalian cell cultures for biologics manufacturing preface. *Adv. Biochem. Eng. Biotechnol.* 139, 1–9. doi: 10.1007/10_2013_255

Karagiosis, S. A., and Baker, S. E. (2012). "Fungal cell factories," in *Food and Industrial Bioproducts and Bioprocessing*, ed. N. T. Dunford, (Oxford: Wiley-Blackwell), 205–219

Kavanagh, K. (2011). *Fungi: Biology and Applications*. Hoboken, NJ: Wiley. doi: 10.1002/9781119976950

Kayikci, Ö., and Nielsen, J. (2015). Glucose repression in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 15:fov068. doi: 10.1093/femsyr/fov068

Kiesenhofer, D., Mach-Aigner, A., and Mach, R. (2016). "Understanding the mechanism of carbon catabolite repression to increase protein production in filamentous fungi," in *Gene Expression Systems in Fungi: Advancements and Applications. Fungal Biology*, eds M. Schmoll, and C. Dattenböck, (Cham: Springer), 275–288 doi: 10.1007/978-3-319-27951-0_12

Kunitake, E., Li, Y., Uchida, R., Nohara, T., Asano, K., Hattori, A., et al. (2019). CreA-independent carbon catabolite repression of cellulase genes by trimeric G-protein and protein kinase A in *Aspergillus nidulans*. *Curr. Genet.* 65, 941–952. doi: 10.1007/s00294-019-00944-4

Li, C., Lin, F., Zhou, L., Qin, L., Li, B.-Z., Zhou, Z., et al. (2017). Cellulase hyper-production by *Trichoderma reesei* mutant SEU-7 on lactose. *Biotechnol. Biofuels* 10:228. doi: 10.1186/s13068-017-0915-9

Li, D., Tang, Y., Lin, J., and Cai, W. (2017). Methods for genetic transformation of filamentous fungi. *Microb. Cell Fact.* 16:168. doi: 10.1186/s12934-017-0785-7

Li, F., Liu, Q., Li, X., Zhang, C., Li, J., Sun, W., et al. (2020). Construction of a new thermophilic fungus *Myceliophthora thermophila* platform for enzyme production using a versatile 2A peptide strategy combined with efficient CRISPR-Cas9 system. *Biotechnol. Lett.* 42, 1181–1191. doi: 10.1007/s10529-020-02882-5

Li, X., Liu, Q., Sun, W., He, Q., and Tian, C. (2019). Improving cellulases production by *Myceliophthora thermophila* through disruption of protease genes. *Biotechnol. Lett.* 42, 219–229. doi: 10.1007/s10529-019-02777-0

Li, X.-L., Skory, C., Ximenes, E., Jordan, D., Dien, B., and Hughes, S. (2007). Expression of an AT-rich xylanase gene from the anaerobic fungus *Orpinomyces* sp. *strain PC*-2 in and secretion of the heterologous enzyme by *Hypocrea jecorina*. *Appl. Microbiol. Biotechnol.* 74, 1264–1275. doi: 10.1007/s00253-006-0787-6

Li, Z., Yao, G., Wu, R., Gao, L., Kan, Q., Liu, M., et al. (2015). Synergistic and dose-controlled regulation of cellulase gene expression in *Penicillium oxalicum*. *PLoS Genet.* 11:e1005509. doi: 10.1371/journal.pgen.1005509

Liao, G.-Y., Zhao, S., Zhang, T., Li, C.-X., Liao, L.-S., Zhang, F.-F., et al. (2018). The transcription factor TpRfx1 is an essential regulator of amylase and cellulase gene expression in *Talaromyces pinophilus*. *Biotechnol. Biofuels* 11:276. doi: 10.1186/s13068-018-1276-8

Lin, L., Wang, S., Li, X., He, Q., Benz, J. P., and Tian, C. (2019). STK-12 acts as a transcriptional brake to control the expression of cellulase-encoding genes in *Neurospora crassa*. *PLoS Genet.* 15:e1008510. doi: 10.1371/journal.pgen.1008510

Lin, S.-C., and Hardie, D. (2017). AMPK: sensing glucose as well as cellular energy status. *Cell Metab.* 27, 299–313. doi: 10.1016/j.cmet.2017.10.009

Liu, Q., Gao, R., Li, J., Lin, L., Zhao, J., Sun, W., et al. (2017). Development of a genome-editing CRISPR/Cas9 system in thermophilic fungal Myceliophthora species and its application to hyper-cellulase production strain engineering. *Biotechnol. Biofuels* 10:1. doi: 10.1186/s13068-016-0693-9

Liu, Q., Li, J., Gao, R., Li, J., Ma, G., and Tian, C. (2018). CLR-4, a novel conserved transcription factor for cellulase gene expression in ascomycete fungi. *Mol. Microbiol.* 111, 373–394. doi: 10.1111/mmi.14160

Liu, Q., Zhang, Y., Li, F., Li, J., Sun, W., and Tian, C. (2019). Upgrading of efficient and scalable CRISPR–Cas-mediated technology for genetic engineering in thermophilic fungus *Myceliophthora thermophila*. *Biotechnol. Biofuels* 12:293. doi: 10.1186/s13068-019-1637-y

Liu, R., Chen, L., Jiang, Y., Zhou, Z., and Zou, G. (2015). Efficient genome editing in filamentous fungus *Trichoderma reesei* using the CRISPR/CAS9 system. *Cell Discov.* 1:15007. doi: 10.1038/celldisc.2015.7

Lockington, R., and Kelly, J. (2002). The WD40-repeat protein CreC interacts with and stabilizes the deubiquitinating enzyme CreB in vivo in *Aspergillus nidulans*: CreB and CreC interact *in vivo*. *Mol. Microbiol.* 43, 1173–1182. doi: 10.1046/j.1365-2958.2002.02811.x

Mach-Aigner, A., Grosstessner-Hain, K., Poças-Fonseca, M., Mechtler, K., and Mach, R. (2010). From an electrophoretic mobility shift assay to isolated transcription factors: a fast genomic-proteomic approach. *BMC Genomics* 11:644. doi: 10.1186/1471-2164-11-644

Madhavan, A., Pandey, A., and Sukumaran, R. (2017). Expression system for heterologous protein expression in the filamentous fungus *Aspergillus unguis*. *Bioresour. Technol.* 245, 1334–1342. doi: 10.1016/j.biortech.2017.05.140

Mäkinen, M., Valkonen, M., Westerholm-Parvinen, A., Aro, N., Arvas, M., Vitikainen, M., et al. (2014). Screening of candidate regulators for cellulase and hemicellulase production in *Trichoderma reesei* and identification of a factor essential for cellulase production. *Biotechnol. Biofuels* 7:14. doi: 10.1186/1754-6834-7-14

Michielse, C., Hooykaas, P., Hondel, C., and Ram, A. (2008). Agrobacterium-mediated transformation of the filamentous fungus *Aspergillus awamori*. *Nat. Protoc.* 3, 1671–1678. doi: 10.1038/nprot.2008.154

Nakari-Setälä, T., Paloheimo, M., Kallio, J., Vehmaanperä, J., Penttilä, M., and Saloheimo, M. (2009). Genetic modification of carbon catabolite repression in *Trichoderma reesei* for improved protein production. *Appl. Environ. Microbiol.* 75, 4853–4860. doi: 10.1128/AEM.00282-09

Navarrete, K., Roa, A., Vaca, Y. I, Espinoza, Navarro, C., and Chavez, R. (2009). Molecular characterization of the niaD and pyrG genes from *Penicillium camemberti*, and its uses as transformation markers. *Cell. Mol. Biol.* 14, 692–702.

Nevalainen, H. (ed.) (2020). *Grand Challenges in Fungal Biotechnology*. Cham: Springer.

Nevalainen, H., and Peterson, R. (2014). Making recombinant proteins in filamentous fungi- Are we expecting too much? *Front. Microbiol.* 5:75. doi: 10.3389/fmicb.2014.00075

Niu, J., Arentshorst, M., Seelinger, F., Ram, A., and Ouedraogo, J.-P. (2016). A set of isogenic auxotrophic strains for constructing multiple gene deletion mutants and parasexual crossings in *Aspergillus niger*. *Arch. Microbiol.* 198, 861–868. doi: 10.1007/s00203-016-1240-6

Nogueira, K., Nogueira, V., Costa, M., De Paula, R., Flávia, C., Mendonça Natividade, F., et al. (2015). Evidence of cAMP involvement in cellobiohydrolase expression and secretion by *Trichoderma reesei* in presence of the inducer sophorose. *BMC Microbiol.* 15:195. doi: 10.1186/s12866-015-0536-z

Park, K.-S., Lee, D.-K., Lee, H., Lee, Y., Jang, Y.-S., Kim, Y., et al. (2003). Corrigendum: phenotypic alteration of eukaryotic cells using randomized libraries of artificial transcription factors. *Nat. Biotechnol.* 21, 1208–1214. doi: 10.1038/nbt868

Portnoy, T., Margeot, A., Seidl-Seiboth, V., Le Crom, S., Chaabane, F., Linke, R., et al. (2011). Differential regulation of the cellulase transcription factors XYR1, ACE2, and ACE1 in *Trichoderma reesei* strains producing high and low levels of cellulase. *Eukaryot. Cell* 10, 262–271. doi: 10.1128/EC.00208-10

Prathumpai, W., Workman, M., and Nielsen, J. (2004). The effect of CreA in glucose and xylose catabolism in. *Appl. Microbiol. Biotechnol.* 63, 748–753. doi: 10.1007/s00253-003-1409-1

Randhawa, A., Ogunyewo, O., Eqbal, D., Gupta, M., and Yazdani, S. S. (2018). Disruption of zinc finger DNA binding domain in catabolite repressor Mig1 increases growth rate, hyphal branching, and cellulase expression in hypercellulolytic fungus *Penicillium funiculosum* NCIM1228. *Biotechnol. Biofuels* 11:15. doi: 10.1186/s13068-018-1011-5

Rassinger, A., Gacek- Matthews, A., Strauss, J., Mach, R., and Mach-Aigner, A. (2018). Truncation of the transcriptional repressor protein Cre1 in *Trichoderma reesei* Rut-C30 turns it into an activator. *Fungal Biol. Biotechnol.* 5:15. doi: 10.1186/s40694-018-0059-0

Raulo, R., Kokolski, M., and Archer, D. (2016). The roles of the zinc finger transcription factors XlnR, ClrA and ClrB in the breakdown of lignocellulose by *Aspergillus niger*. *AMB Express* 6:5. doi: 10.1186/s13568-016-0177-0

Ribeiro, L., Chelius, C., Boppidi, K., Naik, N., Hossain, S., Ramsey, J., et al. (2019). Comprehensive analysis of *Aspergillus nidulans* PKA phosphorylome identifies a novel mode of CreA regulation. *mBio* 10:e02825-18. doi: 10.1128/mBio.02825-18

Ries, L., Beattie, S., Espeso, E., Cramer, R., and Goldman, G. (2016). Diverse regulation of the CreA carbon catabolite repressor in *Aspergillus nidulans*. *Genetics* 203, 335–352. doi: 10.1534/genetics.116.187872

Rubenstein, E. (2008). *Glucose Sensing and the Regulation of the AMP-Activated Protein Kinase in Yeast*. Ph.D. thesis, University of Pittsburgh, Pittsburgh, PA.

Saloheimo, A., Aro, N., Ilmén, M., and Penttilä, M. (2003). ACEI of *Trichoderma reesei* is a repressor of cellulase and xylanase expression. *Appl. Environ. Microbiol.* 69, 56–65. doi: 10.1128/AEM.69.1.56-65.2003

Sarma, N., Haley, T., Haley, K., Buford, T., Willis, K., and Santangelo, G. (2007). Glucose-responsive regulators of gene expression in *Saccharomyces cerevisiae* function at the nuclear periphery via a reverse recruitment mechanism. *Genetics* 175, 1127–1135. doi: 10.1534/genetics.106.068932

Schmoll, M., and Dattenböck, C. (eds) (2016). *Gene Expression Systems in Fungi: Advancements and Applications*, 1st Edn. Berlin: Springer.

Serna, I., Ng, D., and Tyler, B. (1999). Carbon regulation of ribosomal genes in *Neurospora crassa* occurs by a mechanism which does not require Cre-1, the homologue of the aspergillus carbon catabolite repressor, CreA. *Fungal Genet. Biol.* 26, 253–269. doi: 10.1006/fgbi.1999.1121

Sharma, R., Katoch, M., Srivastava, P., and Qazi, G. (2009). Approaches for refining heterologous protein production in filamentous fungi. *World J. Microbiol. Biotechnol.* 25, 2083–2094. doi: 10.1007/s11274-009-0128-x

Shimizu, T., Ito, T., and Kanematsu, S. (2012). Transient and multivariate system for transformation of a fungal plant pathogen, *Rosellinia necatrix*, using autonomously replicating vectors. *Curr. Genet.* 58, 129–138. doi: 10.1007/s00294-012-0370-x

Simpson-Lavy, K., and Kupiec, M. (2019). Carbon catabolite repression: not only for glucose. *Curr. Genet.* 65, 1321–1323. doi: 10.1007/s00294-019-00996-6

Stals, I., Sandra, K., Geysens, S., Contreras, R., Beeumen, J., and Claeyssens, M. (2004). Factors influencing glycosylation of *Trichoderma reesei* cellulases. I: postsecretorial changes of the O- and N-glycosylation pattern of *Ce*17A. *Glycobiology* 14, 713–724. doi: 10.1093/glycob/cwh080

Strauss, J., Mach, R., Zeilinger, S., Hartler, G., Stöffler, G., Wolschek, M., et al. (1995). Cre1, the carbon catabolite repressor protein from *Trichoderma reesei*. *FEBS Lett.* 376, 103–107. doi: 10.1016/0014-5793(95)01255-5

Su, X., Chu, X., and Dong, Z. (2009). Identification of elevated transcripts in a *Trichoderma reesei* strain expressing a chimeric transcription activator using suppression subtractive hybridization. *World J. Microbiol. Biotechnol.* 25, 1075–1084. doi: 10.1007/s11274-009-9993-6

Su, X., Schmitz, G., Zhang, M., Mackie, R., and Cann, I. (2012). Heterologous gene expression in filamentous fungi. *Adv. Appl. Microbiol.* 81, 1–61. doi: 10.1016/B978-0-12-394382-8.00001-0

Sun, X., Zhang, X., Huang, H., Wang, Y., Tu, T., Bai, Y., et al. (2020). Engineering the *cbh1* promoter of *Trichoderma reesei* for enhanced protein production by replacing the binding sites of a transcription repressor ACE1 to those of the activators. *J. Agric. Food Chem.* 68, 1337–1346. doi: 10.1021/acs.jafc.9b05452

Tamayo-Ramos, J. A., and Orejas, M. (2014). Enhanced glycosyl hydrolase production in *Aspergillus nidulans* using transcription factor engineering approaches. *Biotechnol. Biofuels* 7:103. doi: 10.1186/1754-6834-7-103

Tanaka, M., Ichinose, S., Shintani, T., and Gomi, K. (2018). Nuclear export-dependent degradation of the carbon catabolite repressor CreA is regulated by a region located near the C-terminus in *Aspergillus oryzae*. *Mol. Microbiol.* 110, 176–190. doi: 10.1111/mmi.14072

Tanaka, M., Tokuoka, M., and Gomi, K. (2014). Effects of codon optimization on the mRNA levels of heterologous genes in filamentous fungi. *Appl. Microbiol. Biotechnol.* 98, 3859–3867. doi: 10.1007/s00253-014-5609-7

Tani, S., Katsuyama, Y., Hayashi, T., Suzuki, H., Kato, M., Gomi, K., et al. (2001). Characterization of the amyR gene encoding a transcriptional activator for the amylase genes in *Aspergillus nidulans*. *Curr. Genet.* 39, 10–15. doi: 10.1007/s002940000175

Todd, R., Lockington, R., and Kelly, J. (2000). The *Aspergillus nidulans* creC gene involved in carbon catabolite repression encodes a WD40 repeat protein. *Mol. Gen. Genet.* 263, 561–570. doi: 10.1007/s004380051202

Tokuoka, M., Tanaka, M., Ono, K., Takagi, S., Shintani, T., and Gomi, K. (2008). Codon Optimization increases steady-state mRNA levels in *Aspergillus oryzae* heterologous gene expression. *Appl. Environ. Microbiol.* 74, 6538–6546. doi: 10.1128/AEM.01354-08

Vega, M. C. (2016). Advanced technologies for protein complex production and characterization. *Anticancer Res.* 36:4375.

Wang, B., Li, J., Gao, J., Cai, P., Han, X., and Tian, C. (2017). Identification and characterization of the glucose dual-affinity transport system in *Neurospora crassa*: pleiotropic roles in nutrient transport, signaling, and carbon catabolite repression. *Biotechnol. Biofuels* 10:17. doi: 10.1186/s13068-017-0705-4

Wang, F., Zhang, R., Han, L., Guo, W., Du, Z., Niu, K., et al. (2019). Use of fusion transcription factors to reprogram cellulase transcription and enable efficient cellulase production in *Trichoderma reesei*. *Biotechnol. Biofuels* 12:244. doi: 10.1186/s13068-019-1589-2

Wang, L., Ridgway, D., Gu, T., and Moo-Young, M. (2005). Bioprocessing strategies to improve heterologous protein production in filamentous fungal fermentations. *Biotechnol. Adv.* 23, 115–129. doi: 10.1016/j.biotechadv.2004.11.001

Wang, Y., Pierce, M., Schneper, L., Guldal, C., Zhang, X., Tavazoie, S., et al. (2004). Ras and Gpa2 mediate one branch of a redundant glucose signaling pathway in yeast. *PLoS Biol.* 2:E128. doi: 10.1371/journal.pbio.0020128

Ward, O. (2011). Production of recombinant proteins by filamentous fungi. *Biotechnol. Adv.* 30, 1119–1139. doi: 10.1016/j.biotechadv.2011.09.012

Weinhandl, K., Winkler, M., Glieder, A., and Camattari, A. (2014). Carbon source dependent promoters in yeasts. *Microb. Cell Fact.* 13:5. doi: 10.1186/1475-2859-13-5

Wiebe, M. (2003). Stable production of recombinant proteins in filamentous fungi - Problems and improvements. *Mycologist* 17, 140–144. doi: 10.1017/S0269915X03003033

Xinxing, L., Zheng, F., Li, C., Zhang, W., Chen, G., and Liu, W. (2015). Characterization of a copper responsive promoter and its mediated overexpression of the xylanase regulator 1 results in an induction-independent production of cellulases in *Trichoderma reesei*. *Biotechnol. Biofuels* 8:67. doi: 10.1186/s13068-015-0249-4

Xu, G., Li, J., Liu, Q., Sun, W., Jiang, M., and Tian, C. (2018). Transcriptional analysis of *Myceliophthora thermophila* on soluble starch and role of regulator AmyR on polysaccharide degradation. *Bioresour. Technol.* 265, 558–562. doi: 10.1016/j.biortech.2018.05.086

Xu, J.-R., and Bluhm, B. (eds) (2011). *Fungal Genomics: Methods and Protocols*, Vol. 722. Berlin: Springer.

Xu, X., Fan, C., Song, L., Li, J., Chen, Y., Zhang, Y., et al. (2019). A novel CreA-mediated regulation mechanism of cellulase expression in the thermophilic fungus *Humicola insolens*. *Int. J. Mol. Sci.* 20:3693. doi: 10.3390/ijms20153693

Yang, H., Liu, Y., Hao, X., Wang, D., Akhberdi, O., Xiang, B., et al. (2018). Regulation of the Gα-cAMP/PKA signaling pathway in cellulose utilization of *Chaetomium globosum*. *Microb. Cell Fact.* 17:160. doi: 10.1186/s12934-018-1008-6

Yao, G., Li, Z., Gao, L., Wu, R., Kan, Q., Liu, G., et al. (2015). Redesigning the regulatory pathway to enhance cellulase production in *Penicillium oxalicum* David Wilson. *Biotechnol. Biofuels* 8:71. doi: 10.1186/s13068-015-0253-8

Yao, G., Wu, R., Kan, Q., Gao, L., Liu, M., Yang, P., et al. (2016). Production of a high-efficiency cellulase complex via β-glucosidase engineering in *Penicillium oxalicum*. *Biotechnol. Biofuels* 9:78. doi: 10.1186/s13068-016-0491-4

Zhang, F., Bai, F., and Zhao, X. (2016). Enhanced cellulase production from *Trichoderma reesei* Rut-C30 by engineering with an artificial zinc finger protein library. *Biotechnol. J.* 11, 1282–1290. doi: 10.1002/biot.201600227

Zhang, G., Seiboth, B., Wen, C., Yaohua, Z., Xian, L., and Wang, T. (2009). A novel carbon source-dependent genetic transformation system for the versatile cell factory *Hypocrea jecorina* (anamorph *Trichoderma reesei*). *FEMS Microbiol. Lett.* 303, 26–32. doi: 10.1111/j.1574-6968.2009.01851.x

Zhang, J., Chen, Y., Wu, C., Liu, P., Wang, W., and Wei, D. (2019). The transcription factor ACE3 controls cellulase activities and lactose metabolism via two additional regulators in the fungus *Trichoderma reesei*. *J. Biol. Chem.* 294, 18435–18450. doi: 10.1074/jbc.RA119.008497

Zhang, J., Guoxiu, Z., Wang, W., and Wei, D. (2018a). Enhanced cellulase production in *Trichoderma reesei* RUT C30 via constitution of minimal transcriptional activators. *Microb. Cell Fact.* 17:75. doi: 10.1186/s12934-018-0926-7

Zhang, J., Pan, J., Guan, G., Ying, L., Xue, W., Tang, G., et al. (2008). Expression and high-yield production of extremely thermostable bacterial xylanaseB in *Aspergillus niger*. *Enzyme Microb. Technol.* 43, 513–516. doi: 10.1016/j.enzmictec.2008.07.010

Zhang, J., Wu, C., Wang, W., and Wei, D. (2018b). Construction of enhanced transcriptional activators for improving cellulase production in *Trichoderma reesei* RUT C30. *Bioresour. Bioprocess.* 5:40. doi: 10.1186/s40643-018-0226-4

Zhang, X., and Xia, L. (2016). Expression of Talaromyces thermophilus lipase gene in *Trichoderma reesei* by homologous recombination at the cbh1 locus. *J. Ind. Microbiol. Biotechnol.* 44, 377–385. doi: 10.1007/s10295-016-1897-5

Zhang, X., Li, Y., Zhao, X., and Bai, F. (2016). Constitutive cellulase production from glucose using the recombinant *Trichoderma reesei* strain overexpressing an artificial transcription activator. *Bioresour. Technol.* 223, 317–322. doi: 10.1016/j.biortech.2016.10.083

Ziv, C. (2020). Carbon source affects PKA-dependent polarity of *Neurospora crassa* in a CRE-1-dependent and independent manner. *Fungal Genet. Biol.* 45, 103–116. doi: 10.1016/j.fgb.2007.05.005

Zoglowek, M., Lübeck, P. S., Ahring, B., and Lübeck, M. (2014). Heterologous expression of cellobiohydrolases in filamentous fungi – An update on the current challenges, achievements and perspectives. *Process Biochem.* 50, 211–220. doi: 10.1016/j.procbio.2014.12.018

Zou, G., Shi, S., Jiang, Y., Brink, J., Vries, R. P., Chen, L., et al. (2012). Construction of a cellulase hyper-expression system in *Trichoderma reesei* by promoter and enzyme engineering. *Microb. Cell Fact.* 11:21. doi: 10.1186/1475-2859-11-21

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

BRIEF RESEARCH REPORT
published: 30 July 2020
doi: 10.3389/fbioe.2020.00856

CRISPR/Cas13d-Mediated Microbial RNA Knockdown

Kun Zhang[1,2†], Zhihui Zhang[1,2†], Jianan Kang[3], Jiuzhou Chen[1], Jiao Liu[1], Ning Gao[1,2], Liwen Fan[1,4], Ping Zheng[1,2,4], Yu Wang[1,2*] and Jibin Sun[1,2]*

[1] Key Laboratory of Systems Microbial Biotechnology, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China, [2] University of Chinese Academy of Sciences, Beijing, China, [3] College of Life Engineering, Shenyang Institute of Technology, Fushun, China, [4] School of Life Sciences, University of Science and Technology of China, Hefei, China

RNA-guided and RNA-targeting type IV-D CRISPR/Cas systems (CRISPR/Cas13d) have recently been identified and employed for efficient and specific RNA knockdown in mammalian and plant cells. Cas13d possesses dual RNase activities and is capable of processing CRISPR arrays and cleaving target RNAs in a protospacer flanking sequence (PFS)-independent manner. These properties make this system a promising tool for multiplex gene expression regulation in microbes. Herein, we aimed to establish a CRISPR/Cas13d-mediated RNA knockdown platform for bacterial chassis. CasRx, Cas13d from *Ruminococcus flavefaciens* XPD3002, was selected due to its high activity. However, CasRx was found to be highly toxic to both *Escherichia coli* and *Corynebacterium glutamicum*, especially when it cooperated with its guide and target RNAs. After employing a low copy number vector, a tightly controlled promoter, and a weakened ribosome binding site, we successfully constructed an inducible expression system for CasRx and applied it for repressing the expression of a green fluorescent protein (GFP) in *E. coli*. Despite our efforts to optimize inducer usage, guide RNA (gRNA) architecture and combination, and target gene expression level, the highest gene repression efficiency was 30–50% at the protein level and ∼70% at the mRNA level. The moderate RNA knockdown is possibly caused by the collateral cleavage activity toward bystander RNAs, which acts as a mechanism of type IV-D immunity and perturbs microbial metabolism. Further studies on cellular response to CRISPR/Cas13d and improvement in RNA knockdown efficiency are required prior to practical application of this system in microbes.

Keywords: RNA knockdown, CRISPR, Cas13d, CasRx, type IV-D CRISPR effector

OPEN ACCESS

Edited by:
Jiazhang Lian,
Zhejiang University, China

Reviewed by:
Nathan Crook,
North Carolina State University, United States
Shuobo Shi,
Beijing University of Chemical Technology, China

****Correspondence:***
Ping Zheng
zheng_p@tib.cas.cn
Yu Wang
wang_y@tib.cas.cn

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *19 January 2020*
Accepted: *02 July 2020*
Published: *30 July 2020*

Citation:
Zhang K, Zhang Z, Kang J, Chen J, Liu J, Gao N, Fan L, Zheng P, Wang Y and Sun J (2020) CRISPR/Cas13d-Mediated Microbial RNA Knockdown. Front. Bioeng. Biotechnol. 8:856. doi: 10.3389/fbioe.2020.00856

INTRODUCTION

Clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein (Cas) systems that endow microbes with diverse mechanisms for adaptive immunity have been widely engineered to facilitate gene editing even in some genetically intractable microbes (Shapiro et al., 2018; Zheng et al., 2020). However, to facilitate rapid mapping of gene expression levels to metabolic outputs, targeted gene regulation techniques are also in demand. To this end, catalytically dead versions of RNA-guided and DNA-targeting type II (Cas9) and type V (Cas12) systems have been repurposed for CRISPR interference (CRISPRi) in microbes (Lian et al., 2017;

Otoupal and Chatterjee, 2018; Li et al., 2020). Such techniques require a protospacer adjacent motif (PAM) for target recognition. Another concern of DNA-targeting CRISPRi is that targeting a gene in an operon will cause a collateral effect, which silences transcription of downstream genes (Cui et al., 2018). In addition to gene perturbation at transcription stage, RNA interference (RNAi) that can modulate gene expression at the translation stage has also been developed using synthetic RNAs (Na et al., 2013; Laganà et al., 2014). However, functioning of RNAi depends on proper host machinery and thus is limited to certain organisms. It is also suggested that RNAi sometimes exhibit significant off-target effects (Qi et al., 2013).

Newly discovered type IV CRISPR systems are RNA-guided and RNA-targeting systems, which include a single protein effector (Cas13) that can target and cleave a specific RNA with a single guide RNA (gRNA) (Shmakov et al., 2015; Abudayyeh et al., 2016; Smargon et al., 2017). Cas13 is also capable to process a CRISPR repeat array into mature gRNAs via a HEPN domain-independent mechanism (**Figure 1A**). These properties facilitate rapid development of a new generation of RNA targeting and editing tools for multiplex gene regulation. Type IV systems can be divided into four subtypes (A–D) based on the phylogeny of effector complexes (Wang F. et al., 2019). Cas13a/b/c were initially discovered and their applications in RNA knockdown in mammalian cells exhibited high efficiency and specificity (Abudayyeh et al., 2017; Cox et al., 2017). Instead of a preferred PAM sequence, Cas13a requires a 3′ protospacer flanking sequence (PFS) of H, while Cas13b requires both a 3′ PFS of NAN or NNA and a 5′ PFS of D for effective RNA cleavage (Abudayyeh et al., 2016; Smargon et al., 2017). Type IV-D CRISPR effectors, known as Cas13d, were recently discovered and employed for RNA knockdown in mammalian cells. In contrast to other RNA-targeting systems, target RNA cleavage by CRISPR/Cas13d is PFS-independent (Konermann et al., 2018; Yan et al., 2018; Zhang C. et al., 2018). Notably, Cas13d from *Ruminococcus flavefaciens* XPD3002 (CasRx) was reported to mediate more efficient and specific RNA knockdown in plants than frequently used Cas13a and Cas13b variants (Mahas et al., 2019). Although *Escherichia coli* was used as a bacterial host for functional screening for CRISPR/Cas13d (Yan et al., 2018), this system has not been characterized in detail in microbes and employed as a microbial gene regulation tool yet.

In this study, we aimed to develop a microbial RNA knockdown technique based on the most promising CRISPR/Cas13d system, CRISPR/CasRx. Unexpectedly, we found that CRISPR/CasRx system was highly toxic to two important platform microbes *E. coli* and *Corynebacterium glutamicum*. After lowering the plasmid copy number and weakening the leaky expression of the *casRx* gene, a plasmid-based expression system for CRISPR/CasRx was constructed. Using green fluorescent protein (GFP) as a reporter, moderate gene repression was achieved in *E. coli* with this system, which was verified at both mRNA and protein levels. Various RNA architectures and combinations were also tested for their effects on RNA knockdown efficiency.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

Bacterial strains used in this study are listed in **Supplementary Table S1**. *E. coli* Trans1-T1 (Transgen, China) was used as the host strain for general cloning and RNA knockdown test. *E. coli* DB 3.1 was used for cloning plasmids harboring the *ccdB* gene. *E. coli* strains were cultivated in a Luria–Bertani (LB) medium at 37°C and with shaking at 220 rpm. Ampicillin (Amp, 100 μg/mL), chloramphenicol (Cm, 20 μg/mL), or kanamycin (Kan, 50 μg/mL) was added when only one antibiotic was used. In the case in which multiple antibiotics were added to maintain multiple plasmids, antibiotics were used at lower concentrations (Amp, 75 μg/mL; Cm, 10 μg/mL; Kan, 25 μg/mL). Tetracycline (Tc) was added at the beginning of cultivation to induce *casRx* expression. A *C. glutamicum* ATCC 13032 derivative harboring a chromosomal *gfp* expression cassette was used to test the transformation of *casRx* expression plasmid and cultivated aerobically at 30°C in an LB medium supplemented with 5 g/L glucose. Electro-competent cells were prepared as described previously (Ruan et al., 2015).

Plasmid Construction

Plasmids used in this study are listed in **Supplementary Table S1**. The *casRx* gene was synthesized by GenScript (China) and first cloned to an *E. coli–C. glutamicum* shuttle plasmid pXMJ19 under the control of isopropyl-β-D-thiogalactopyranoside (IPTG) inducible promoter P_{tac}, producing plasmid pCasRx-1. To correct the unexpected mutations in *casRx* and lower the copy number of pCasRx-1, two fragments were amplified from pCasRx-1 with primer pairs CasRx-F1/CasRx-R1 and CasRx-F2/CasRx-R2 to correct the mutations and remove the original *pUC* replicon. *pSC101* replicon was amplified from plasmid pSB4K5-I52002 (Shetty et al., 2008) with a primer pair pSC101-F/pSC101-R and ligated with the aforementioned two fragments using a ClonExpress MultiS One Step Cloning Kit (Vazyme, China), producing pCasRx-2. To achieve Tc inducible expression of *casRx* and change the original strong RBS (AAAGGAGTTGAGA) to a weaker one (AAAGGCACCCGAT), *casRx* was amplified from pCasRx-2 with a primer pair CasRx-F3/CasRx-R3. A weaker RBS was simultaneously added with the forward primer. A laboratory stock plasmid pZSA harboring a P_{tet} promoter and a TetR effector encoding gene was linearized by PCR with a primer pair pZSA-F/pZSA-R. The *casRx* and pZSA fragments were ligated to produce pCasRx-3.

An *E. coli–C. glutamicum* shuttle plasmid pgRNA-*ccdB* (pEC-XK99E backbone) (Wang et al., 2018b) for expressing gRNA of Cas9 was used as a template to construct plasmids pgRNA-*ccdB*-1 and pgRNA-*ccdB*-2, which were designed to express unprocessed pre-gRNA [a 30-nt spacer flanked by two 36-nt direct repeats (DRs)] and mature gRNA (a 30-nt DR with a 22-nt spacer) of CasRx, respectively (**Supplementary Figure S1**). gRNA transcription was controlled by a constitutive promoter

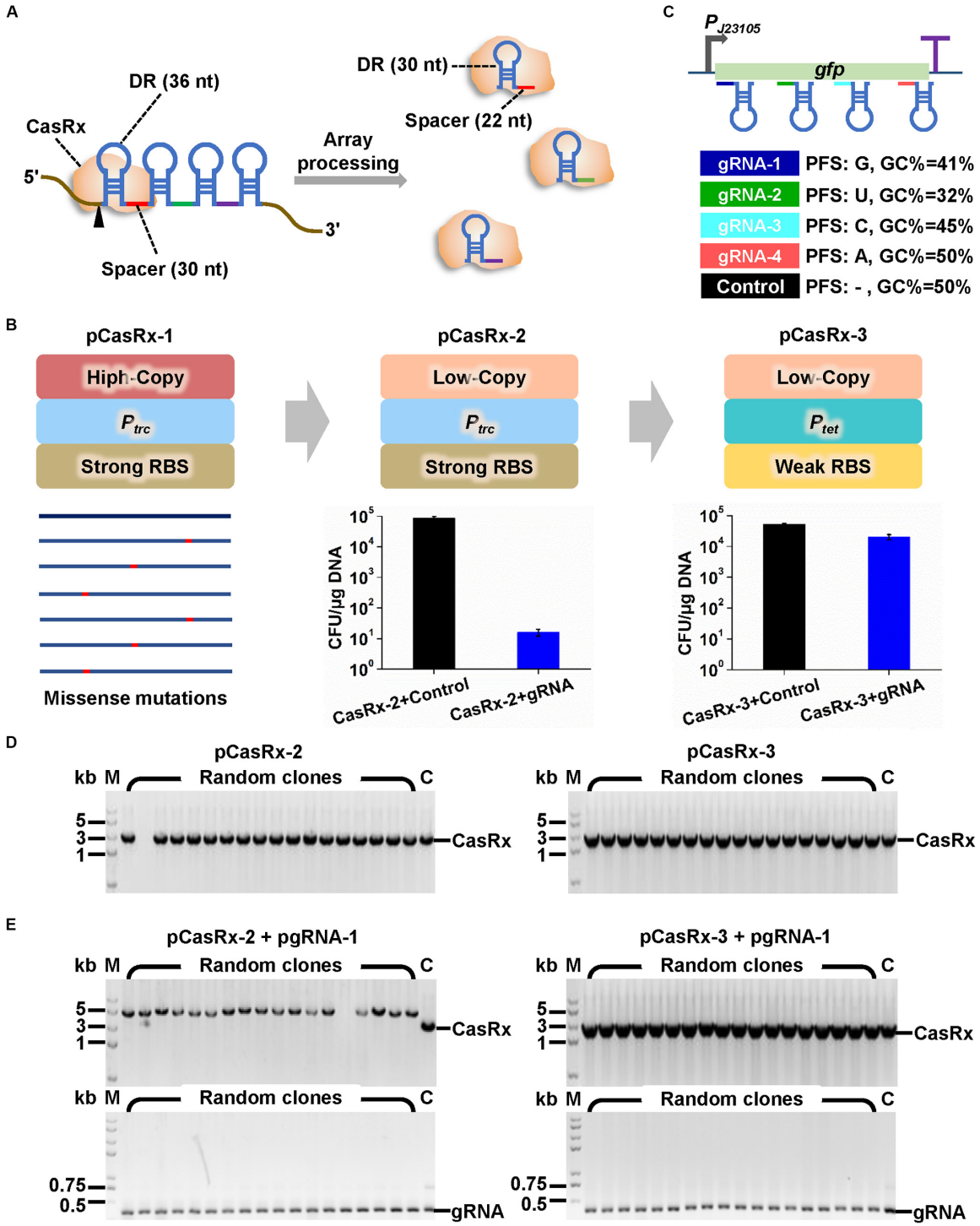


FIGURE 1 | Construction of microbial expression systems for CRISPR/CasRx. **(A)** Multiple gRNAs containing a 30-nt spacer flanked by two 36-nt DRs can be expressed as a single array and processed by CasRx into individual mature gRNAs containing 30-nt DR and 22-nt spacer. **(B)** Strategies to construct *casRx* expression plasmid for *E. coli*. The high copy number *pUC* replicon was first replaced with a low copy number *pSC101* replicon. An IPTG inducible promoter P_{tac} and a strong RBS (AAAGGAGTTGAGA) were then replaced with a Tc inducible promoter P_{tet} and a weak RBS (AAAGGCACCCGAT). To determine plasmid transformation efficiency, *E. coli* harboring pGFP was co-transformed with a *casRx* expression plasmid (pCasRx-2 or pCasRx-3) and an empty plasmid with no gRNA cassette (control) or a *gfp*-targeting gRNA expression plasmid (pgRNA-1). Cells were plated on LB solid plates supplemented with Amp + Cm + Kan. Clones were counted after 24 h cultivation. **(C)** Design of four *gfp*-targeting pre-gRNAs with different characteristics and a non-targeting control with no target site on the whole *E. coli* and *C. glutamicum* transcripts. **(D)** Clone PCR verification of transformants of pCasRx-2 or pCasRx-3. M, DNA marker; C, plasmid positive control. **(E)** Clone PCR verification of transformants of pCasRx-2 + pgRNA-1 or pCasRx-3 + pgRNA-1. M, DNA marker; C, plasmid positive control.

P_{11F} that functions in both *E. coli* and *C. glutamicum* (Liu et al., 2017). The backbone of pgRNA-*ccdB* was amplified with a primer pair pgRNA-F/pgRNA-R. *ccdB* fragment was amplified and DRs were simultaneously added with primer pairs DR36-*ccdB*-F/DR36-*ccdB*-R (or DR30-*ccdB*-F/DR30-*ccdB*-R). The plasmid backbone and *ccdB* fragment were ligated to produce pgRNA-*ccdB*-1 (or pgRNA-*ccdB*-2). To construct an expression plasmid for *gfp*-targeting gRNA, a primer pair was used to generate double-stranded DNA (dsDNA) containing a *gfp*-targeting spacer. The dsDNA was assembled with pgRNA-*ccdB*-1 (or pgRNA-*ccdB*-2) to replace *ccdB* with the spacer by Golden Gate assembly described previously (Wang et al., 2018b; **Supplementary Figure S1**). To insert additional nucleotides between the second 36-nt DR and a terminator, an *ldhA* fragment was amplified from *E. coli* genomic DNA with primer pair adn-F/adn-100-R, adn-F/adn-300-R, or adn-F/adn-1000-R. The *ldhA* fragment and a dsDNA containing spacer and 36-nt DR were assembled with pgRNA-*ccdB*-1 by Golden Gate assembly.

To construct a GFP reporter system that is compatible with the CRISPR/CasRx system, pGFP was constructed with the pTrc99A backbone, the *p15A* replicon, and the *gfp* gene. *gfp* was first amplified from pTRCmob-*egfp* (Wang et al., 2018a) with a primer pair GFP-F/GFP-R, and a constitutive promoter P_{J23105} was added to the PCR product via the forward primer. Two fragments were amplified from pTrc99A by PCR with primer pairs pTrc99A-F1/pTrc99A-R1 and pTrc99A-F2/pTrc99A-R2 to remove the original *pUC* replicon. p15A replicon was amplified from pACYCDuet-1 by PCR with a primer pair p15A-F/p15A-R. The three PCR products were ligated to produce pGFP. To construct a reporter system with a lower GFP expression level, pGFP-w was constructed by PCR using pGFP as a template and the primer pair J23117-F/J23117-R. The promoter P_{J23105} was replaced with a weaker promoter P_{J23117}. Primers used for plasmid construction, spacers, and additional nucleotides in gRNAs are listed in **Supplementary Tables S2–S4**, respectively. The full sequences for all the plasmids constructed in this study were uploaded as **Supplementary Material** in GenBank format.

Assay of RNA Knockdown Efficiency by Determining GFP Fluorescence

E. coli harboring pGFP or pGFP-w was co-transformed with pCasRx-3 and a gRNA expression plasmid. Cells were plated in LB solid plates containing Amp, Cm, and Kan. After 24-h cultivation at 37°C, clones were verified by PCR, and three correct clones were picked and incubated in a 5 mL LB medium supplemented with antibiotics. The cultures were used as seeds to inoculate a 10 mL fresh LB medium in 100 mL shake flasks supplemented with antibiotics and inducer. The initial cell density (OD_{600nm}) was set as 0.05. During cultivation, cells were collected periodically by centrifugation at 6,000 × *g* for 5 min, washed once with phosphate buffered saline (PBS) buffer (pH 7.4), and resuspended in PBS buffer in Corning 3603 96-well microplates (Corning Incorporated, United States). GFP fluorescence was measured using an Infinite 200 PRO plate reader (Tecan Trading AG, Switzerland) (λ excitation = 488 nm, λ emission = 520 nm), and OD_{600nm} was determined simultaneously.

Transcription Level Assay by Real-Time Quantitative PCR (RT-qPCR)

Total RNAs were extracted from cells of exponential phase using an RNAprep Pure Cell/Bacteria Kit, treated with DNase I, and used to synthesize cDNAs using random primers and a Fast Quant RT Kit. The resultant cDNAs were used as templates for RT-qPCR analysis. The total RNA samples were also used as templates for RT-qPCR to confirm that genomic DNA contamination during total RNA extraction was minimal. Specific primers for RT-qPCR were designed using Beacon Designer software v7.7 (PREMIER Biosoft International, United States) (**Supplementary Table S2**). RT-qPCR was performed using a SuperReal Premix SYBR Green Kit and Applied Biosystems® 7500 Real-Time PCR System (Thermo Fisher Scientific, United States) according to the manufacturers' instructions. The relative transcription level of *gfp* was calculated using the $2^{-\Delta\Delta CT}$ method. All the kits used for RT-qPCR were purchased from Tiangen Biotech, China.

RESULTS

Construction of Bacterial Expression Systems for CasRx and gRNA

Development of a gene regulation technique based on CRISPR/Cas13d requires functional and controllable expression of a Cas13d effector and transcription of a gRNA with proper architecture. CRISPR/CasRx system (Konermann et al., 2018) that has been successfully applied in RNA knockdown in mammalian and plant cells was used here. We intended to develop a versatile CRISPR/CasRx-mediated RNA knockdown platform for two widely used microbial chassis, *E. coli* and *C. glutamicum*. Therefore, *casRx* was synthesized and cloned to pXMJ19, an *E. coli–C. glutamicum* shuttle vector with a high copy number *E. coli* replicon (*pUC*, ∼75 copies/cell) and a low copy number *C. glutamicum* replicon (*pBL1*, ∼8 copies/cell) (Jakoby et al., 1999). In the recombinant plasmid pCasRx-1, *casRx* was controlled by an IPTG inducible promoter P_{tac}. However, few transformants of ligation products were obtained, and gene sequencing suggested existence of missense mutations in *casRx* in all sequenced plasmids (**Figure 1B**). Considering the possible leaky expression of P_{tac} and the cellular toxicity of Cas effectors to microbes (Liu et al., 2017; Wang K. et al., 2019), the high copy number *pUC* replicon was replaced by *pSC101* replicon with a low copy number in *E. coli* (∼8 copies/cell). This modification produced recombinant plasmid pCasRx-2 harboring the correct *casRx* gene.

Next, gRNA expression plasmids were constructed based on another *E. coli–C. glutamicum* shuttle vector pEC-XK99E with a high copy number *E. coli* replicon (*pUC*, ∼75 copies/cell) and a medium copy number *C. glutamicum* replicon (*pGA1*, ∼30 copies/cell) (Kirchner and Tauch, 2003). The gRNA was initially designed as a 30-nt spacer flanked

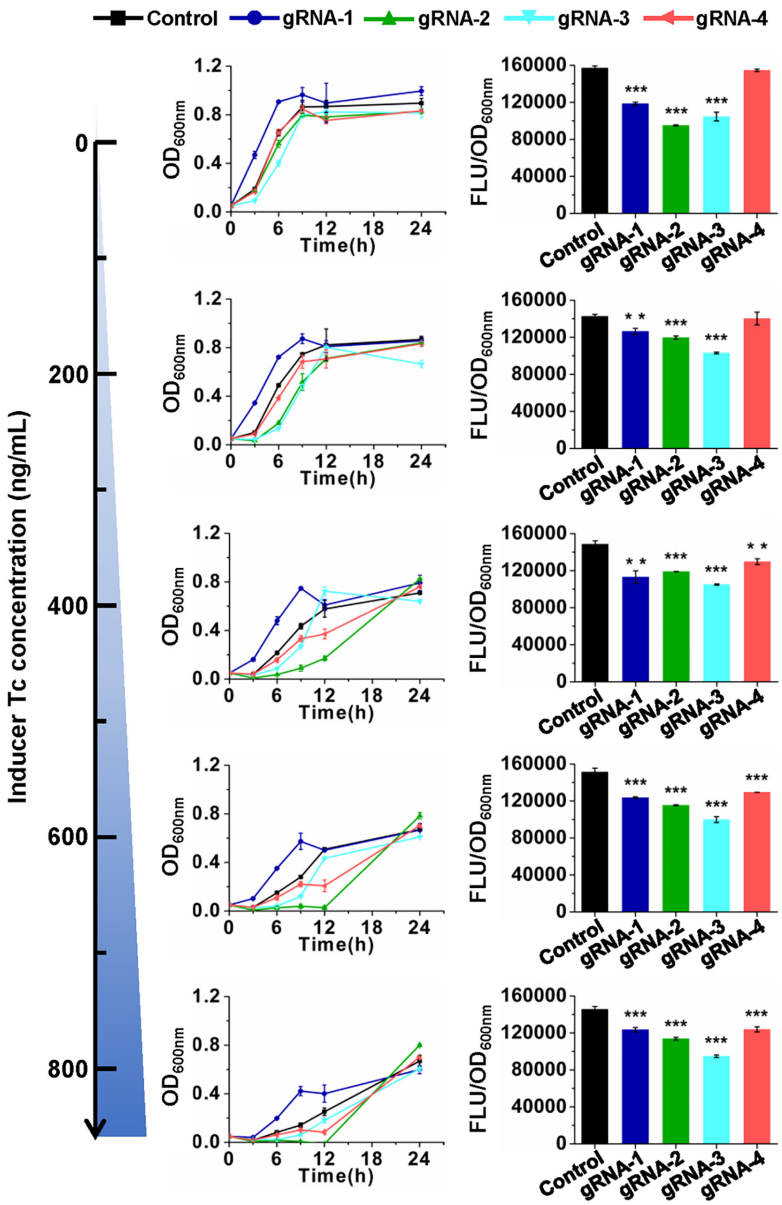


FIGURE 2 | CRISPR/CasRx-mediated gene repression with pre-gRNAs. Design of gRNA-1, gRNA-2, gRNA-3, and gRNA-4 are shown in **Figure 1B**. The spacer sequences are listed in **Supplementary Table S3**. pCasRx-3 and different gRNA expression plasmids were co-transformed into *E. coli* harboring pGFP for gene repression. Inducer was added at the beginning of cultivation to induce *casRx* expression. GFP fluorescence and OD_{600nm} were determined after 24-h cultivation. Error bars indicate standard deviations from three parallel experiments. All *t*-tests compare the GFP fluorescence per OD_{600nm} using *gfp*-targeting gRNAs against non-targeting gRNA control (**$P < 0.01$, ***$P < 0.001$).

by two 36-nt DRs to mimic an unprocessed pre-gRNA. The secondary structure of GFP reporter gene transcript was predicted by using the Vienna RNA websuite[1] (Gruber et al., 2008; **Supplementary Figure S2**). Four gRNAs whose spacers target distinct regions of *gfp* transcript and possess different G + C contents and PFSs were designed (**Figure 1C** and **Supplementary Figure S2**). A non-targeting gRNA with no target site on the whole transcripts of *E. coli* and *C. glutamicum* was designed as a control. Then, plasmid pCasRx-2 was transformed into *E. coli* harboring pGFP together with plasmid pgRNA-1 or an empty plasmid with no gRNA cassette. Unexpectedly, co-transformation of pCasRx-2 and pgRNA-1 significantly reduced the number of transformants by approximately 1,000-fold, compared to the transformation of pCasRx-2 and empty plasmid (**Figure 1B**). Clone PCR verification suggested that when pCasRx-2 and pgRNA-1 were co-transformed, the resultant transformants possessed incorrect *casRx* gene (**Figures 1D,E**).

[1]http://rna.tbi.univie.ac.at/

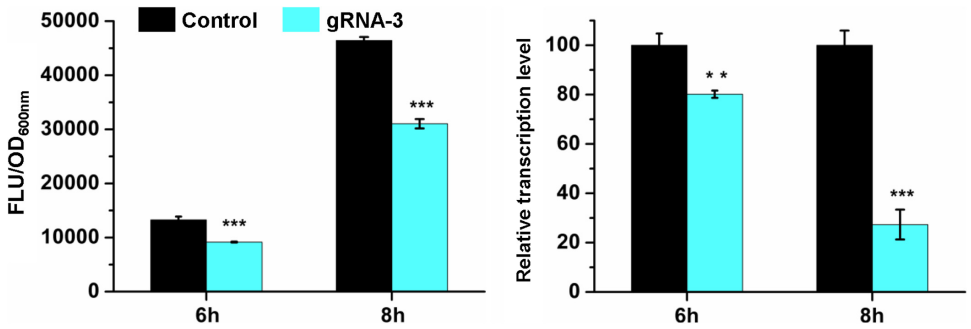


FIGURE 3 | CRISPR/CasRx-mediated mRNA knockdown. pCasRx-3 and pgRNA-3 were co-transformed into *E. coli* harboring pGFP for gene repression. Inducer (200 ng/mL Tc) was added at the beginning of cultivation to induce *casRx* expression. Cells were collected at 6 and 8 h for GFP fluorescence and mRNA determination. Error bars indicate standard deviations from three parallel experiments. All *t*-tests compare the GFP fluorescence per OD_{600nm} or *gfp* transcription level using *gfp*-targeting gRNA-3 against non-targeting gRNA control (**$P < 0.01$, ***$P < 0.001$).

According to previous literatures, when the Cas13d effector combines with gRNA and target mRNA to form a ternary Cas13d-gRNA-target mRNA complex, the complex is capable of cleaving non-specific bystander RNAs (Konermann et al., 2018; Yan et al., 2018). Such bystander cleavage may disturb cell metabolism or even be lethal to cells. We speculated that a tighter regulation of *casRx* transcription and a weaker translation initiation may be essential for successful development of a functional CRISPR/CasRx system in microbes. Therefore, the IPTG inducible P_{tac} promoter and original strong RBS (AAAGGAGTTGAGA) in pXMJ19 were replaced with a Tc inducible P_{tet} promoter with better tightness (Lutz and Bujard, 1997) and a weaker RBS (AAAGGCACCCGAT). This modification significantly increased co-transformation efficiency of the resultant plasmid pCasRx-3 and pgRNA-1 (**Figure 1B**). Clone PCR verification suggested that the transformants of pCasRx-3 and pgRNA-1 all harbored correct *casRx* gene (**Figures 1D,E**).

With efforts to recruit a low copy number replicon, a tightly controlled promoter, and a weak RBS, a CRISPR/CasRx expression system was established in *E. coli*. The pCasRx-2 plasmid that only has ∼8 copies per *C. glutamicum* cell was also tested for its maintenance in a *gfp* expressing *C. glutamicum*. However, non-transformants were obtained even though pCasRx-2 was transformed individually, suggesting that *C. glutamicum* might be more sensitive to CasRx than *E. coli* since the ternary Cas13d-gRNA-target mRNA complex was not generated due to the non-involvement of a gRNA expression plasmid.

CRISPR/CasRx-Mediated RNA Knockdown in *E. coli*

The developed CRISPR/CasRx system was then tested for its application in RNA knockdown using GFP as a reporter. pCasRx-3 was transformed into *E. coli* harboring pGFP with a gRNA plasmid harboring the *gfp*-targeting gRNA or non-targeting control gRNA, respectively (**Figure 1C**). Transformants were cultured in media containing different concentrations of Tc to explore the optimal inducer concentration for CRISPR/CasRx-mediated RNA knockdown. In the absence of an inducer, lowered GFP fluorescence was already detected for *gfp*-targeting gRNAs, except for gRNA-4 targeting the end of the *gfp* transcript (**Figure 2**). The result suggested that leaky expression of *casRx* still happened. With the increase in the Tc usage, growth inhibition and prolonged lag phase were observed, which may be caused by a reported Cas13-induced cellular dormancy (Meeske et al., 2019). However, the *gfp* repression efficiency was not significantly improved. Overall, the *gfp* repression efficiency fluctuated between 11.4 and 23.8%, 16.1 and 23.7%, and 27.7 and 34.8% for gRNA-1, gRNA-2, and gRNA-3, respectively. Up to 14.9% repression of *gfp* expression was observed for gRNA-4 (**Figure 2**).

To investigate whether the repression was caused by CRISPR/CasRx-mediated mRNA knockdown, RT-qPCR was conducted for the *gfp* repression test using gRNA-3 and CasRx induced by 200 ng/mL Tc due to relatively high repression efficiency and slight growth inhibition. Cells at the exponential phase were collected and analyzed. At 6 h when *gfp* was expressed at a relatively low level, *gfp* was repressed by 30.8% at the protein level and 19.8% at the mRNA level. At 8 h when *gfp* was expressed at a relatively high level, the repression efficiency increased to 33.1%, and the corresponding mRNA knockdown efficiency increased to 72.6% (**Figure 3**). The results suggest that the developed CRISPR/CasRx system can repress the expression of a target gene via knocking down the target mRNA in *E. coli*.

Further Attempts to Improve RNA Knockdown Efficiency

The knockdown effects of the CRISPR/CasRx system in *E. coli* are much lower compared to those in mammalian and plant cells (over 80%) (Konermann et al., 2018; Mahas et al., 2019). To improve the knockdown efficiency, several attempts were made. gRNA architecture has been proven to be essential for functioning of the CRISPR/Cas system (Liu et al., 2017). In eukaryotic transcription of gRNAs with the RNA Polymerase III promoter U6, the gRNA transcripts usually have a short

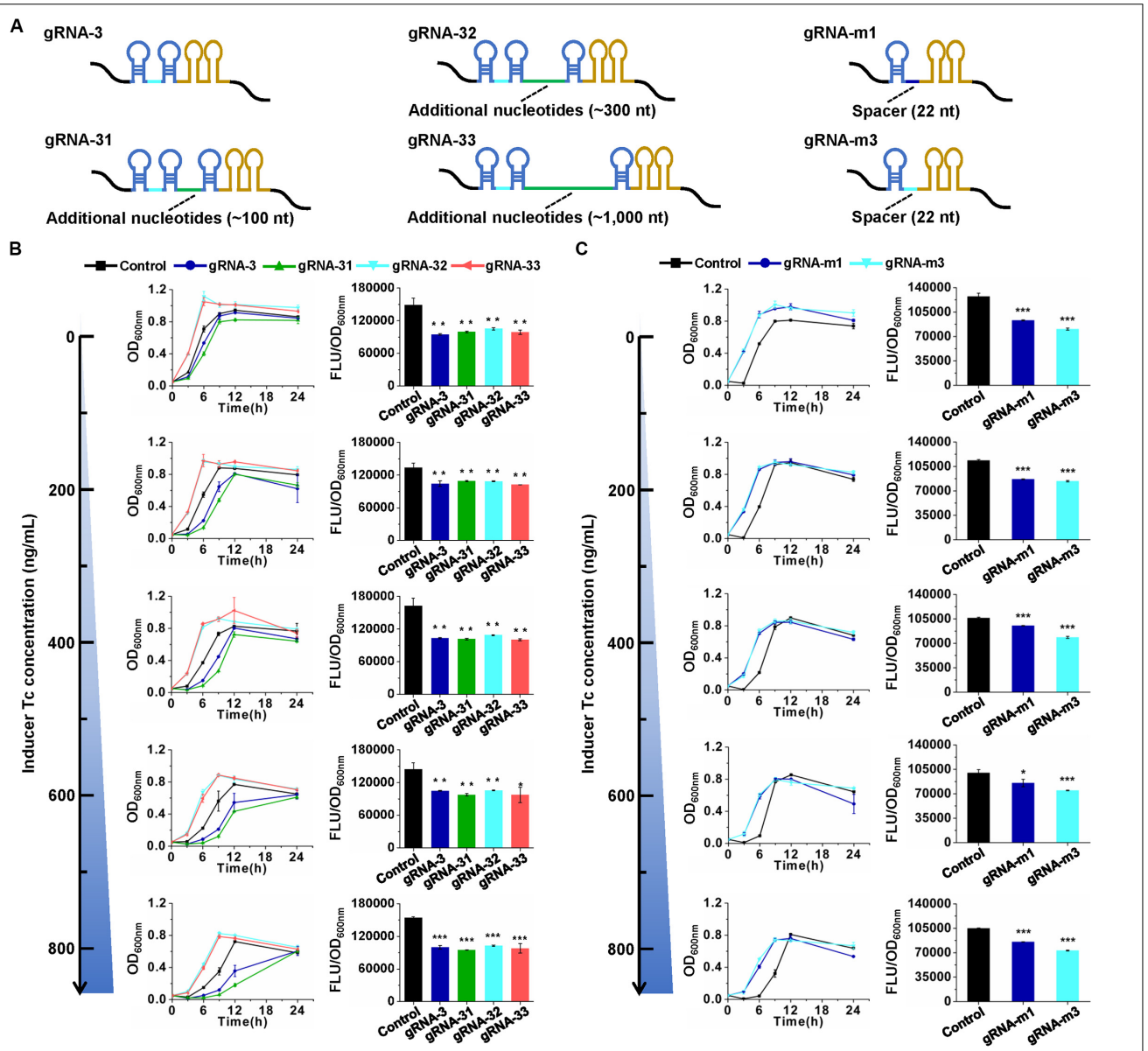


FIGURE 4 | Effects of gRNA architecture on gene repression efficiency. **(A)** Architectures of pre-gRNAs containing 30-nt spacers and additional nucleotides and mature gRNAs containing 22-nt spacers. gRNA-31, gRNA-32, and gRNA-33 have the same spacer with gRNA-3 but additional nucleotides downstream of the second DR. gRNA-m1 and gRNA-m3 have 22-nt spacers, which are the same with the CasRx-processed gRNA-1 and gRNA-3, respectively. The spacer sequences are listed in **Supplementary Table S3**. **(B)** Growth and GFP fluorescence with pre-gRNAs and inducer at different concentrations. **(C)** Growth and GFP fluorescence with mature gRNAs and inducer at different concentrations. pCasRx-3 and different gRNA expression plasmids were co-transformed into *E. coli* harboring pGFP for gene repression. Inducer was added at the beginning of cultivation to induce *casRx* expression. GFP fluorescence and OD_{600nm} were determined after 24 h cultivation. Error bars indicate standard deviations from three parallel experiments. All *t*-tests compare the GFP fluorescence per OD_{600nm} using *gfp*-targeting gRNAs against non-targeting gRNA control (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$).

stretch of four to eight Us at their 3′ end (Zhang C. et al., 2018). However, in the present prokaryotic transcription cases, intrinsic transcription terminator T_{rrnB} forms two consecutive self-annealing hairpin structures next to the second DR on the elongating transcript (**Figure 4A**). As a type IV-D CRISPR effector, CasRx first processes pre-gRNAs into mature gRNAs via a HEPN domain-independent mechanism (Zhang C. et al., 2018). We speculated that the hairpin structures of the prokaryotic gRNA terminator might disturb the binding of CasRx to pre-gRNA and processing of pre-gRNAs into mature gRNAs. Two strategies were applied to test our hypothesis. First, an additional sequence (approximately 100, 300, or 1,000 nt) was inserted between the second DR and the terminator in pre-gRNA. Second, the two 36-nt DRs with a 30-nt spacer in pre-gRNA were replaced by a 30-nt DR with a 22-nt spacer to mimic a mature gRNA (**Figure 4A**). Unfortunately, such

modifications brought no positive effects to gene knockdown efficiency (**Figures 4B,C**). We further expanded the spacer length in mature gRNAs from 22 to 30 nt to enhance the potential binding between gRNA and target mRNA. However, this trial did not improve gene knockdown efficiency either (**Supplementary Figure S3**).

In some cases of dCas9- and dCas12a-mediated CRISPRi systems, the gene repression efficiency could be improved by a multiplex gRNAs strategy of making more than one gRNA toward the target gene simultaneously (Zhang J.L. et al., 2018). To test the effects of this strategy in this CRISPR/CasRx system, the most effective gRNA-3 was combined with gRNA-2 and gRNA-4, respectively. The resultant plasmids expressing a gRNA array were co-transformed with pCasRx-3 into *E. coli* harboring pGFP. The repression efficiencies observed were no higher than those obtained with an individual gRNA (**Supplementary Figure S4**). Considering the stoichiometry between CasRx-gRNA complex and the target mRNA, more weakly expressed genes may be repressed to a greater extent. We then constructed a pGFP-w plasmid using a weaker P_{J23117} promoter than the previously used P_{J23105}. By expressing CasRx and gRNA-3 in *E. coli* harboring pGFP-w, a knockdown efficiency of 56.1% was observed, which was slightly higher than the previously obtained ∼30% efficiency using the pGFP reporter system (**Supplementary Figure S5**).

DISCUSSION

Considering the importance of gene regulation in understanding and manipulating cellular metabolism, development of robust and efficient gene regulation techniques is drawing increasing attention. CRISPRi and RNAi that function at the transcription and translation stages, respectively, are well established and widely used in biological research and disease treatment (Laganà et al., 2014; Liu et al., 2018). With the advent of the type IV-D CRISPR effector (Cas13d), new programmable RNA-guided and RNA-targeting techniques are available, which have largely advanced the field of gene regulation (Wang F. et al., 2019). In this study, we demonstrate that the CRISPR/Cas13d system can be used for repressing gene expression in bacteria by knocking down target mRNA. However, the gene knockdown efficiency is far from satisfactory (30–50%) compared to those obtained in mammalian and plant cells (over 80%) (Konermann et al., 2018; Mahas et al., 2019). Although we optimized the gRNA architectures and combinations, the knockdown efficiency still hardly meets the requirements of practical application. Moreover, the CRISPR/Cas13d system seems highly inhibitory to microbial growth and metabolism.

The recently uncovered defense mechanism of type IV CRISPR systems against bacteriophages provides a possible explanation to the different performances of CRISPR/Cas13d in eukaryotic and prokaryotic cells (Jackson and Fineran, 2019; Meeske et al., 2019). Among the discovered CRISPR systems, type IV systems are intriguing because they are unique ones that cleave viral RNA rather than DNA. It was found that Cas13d induced cellular dormancy or even triggered cell death by degrading bacterial RNA, which aborted the infectious cycle and thus provided robust defense against bacteriophages (Meeske et al., 2019). Such collateral cleavage activity of the ternary Cas13d:gRNA:target RNA complex was verified *in vitro*, whereas it did not cause observable off-target transcriptome perturbation in mammalian and plant cells (Konermann et al., 2018; Mahas et al., 2019). Since these CRISPR systems are originated from bacterial adaptive immune systems, their distinct performances in eukaryotic and prokaryotic cells may be reasonable.

In conclusion, we developed an RNA knockdown technique based on the CRISPR/Cas13d system from *R. flavefaciens* XPD3002 (CRISPR/CasRx) and obtained moderate gene repression in *E. coli*. Considering that CRISPR/Cas13d directly acts on RNA and can process a CRISPR repeat array into multiple mature gRNAs without the involvement of additional RNases, this system holds promise for multiplex gene regulation in microbes. However, the collateral cleavage activity on bacterial RNA and relatively low knockdown efficiency need to be circumvented prior to practical applications. Recent successes in rational engineering and directed evolution of Cas proteins for increased activity, elevated targeting specificity, and released PAM requirements have set good examples for improving Cas13d (Kleinstiver et al., 2019; Miller et al., 2020).

DATA AVAILABILITY STATEMENT

The sequences for pCasRx-3, pgRNA-ccdB-1, and pgRNA-ccdB-2 were deposited into the GenBank database under the accession numbers MN934322, MN934323, and MN934324, respectively, at the National Center for Biotechnology Information. The full sequences for all the plasmids constructed in this study were uploaded as supplementary materials in GenBank format. All data and materials that support the findings of this study are available from the corresponding author on reasonable request.

AUTHOR CONTRIBUTIONS

YW, PZ, and JS conceived and initiated the project. YW and KZ designed the experiments. KZ, ZZ, JK, NG, and LF carried out the experiments. KZ, ZZ, YW, JC, and JL analyzed the data. YW, KZ, PZ, and JS wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

This research was supported by the National Key R&D Program of China (2018YFA0903600 and 2018YFA0900300), the National Natural Science Foundation of China (31700044 and 31870044), and the International Partnership Program of Chinese Academy of Sciences (153D31KYSB20170121).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00856/full#supplementary-material

REFERENCES

Abudayyeh, O. O., Gootenberg, J. S., Essletzbichler, P., Han, S., Joung, J., Belanto, J. J., et al. (2017). RNA targeting with CRISPR-Cas13. *Nature* 550, 280–284. doi: 10.1038/nature24049

Abudayyeh, O. O., Gootenberg, J. S., Konermann, S., Joung, J., Slaymaker, I. M., Cox, D. B., et al. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* 353:aaf5573. doi: 10.1126/science.aaf5573

Cox, D. B. T., Gootenberg, J. S., Abudayyeh, O. O., Franklin, B., Kellner, M. J., Joung, J., et al. (2017). RNA editing with CRISPR-Cas13. *Science* 358, 1019–1027. doi: 10.1126/science.aaq0180

Cui, L., Vigouroux, A., Rousset, F., Varet, H., Khanna, V., and Bikard, D. (2018). A CRISPRi screen in *E. coli* reveals sequence-specific toxicity of dCas9. *Nat. Commun.* 9:1912. doi: 10.1038/s41467-018-04209-5

Gruber, A. R., Lorenz, R., Bernhart, S. H., Neubock, R., and Hofacker, I. L. (2008). The Vienna RNA websuite. *Nucleic Acids Res.* 36, W70–W74. doi: 10.1093/nar/gkn188

Jackson, S. A., and Fineran, P. C. (2019). Bacterial dormancy curbs phage epidemics. *Nature* 570, 173–174. doi: 10.1038/d41586-019-01595-8

Jakoby, M., Ngouoto-Nkili, C.-E., and Burkovski, A. (1999). Construction and application of new *Corynebacterium glutamicum* vectors. *Biotechnol. Tech.* 13, 437–441.

Kirchner, O., and Tauch, A. (2003). Tools for genetic engineering in the amino acid-producing bacterium *Corynebacterium glutamicum*. *J. Biotechnol.* 104, 287–299. doi: 10.1016/S0168-1656(03)00148-2

Kleinstiver, B. P., Sousa, A. A., Walton, R. T., Tak, Y. E., Hsu, J. Y., Clement, K., et al. (2019). Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat. Biotechnol.* 37, 276–282. doi: 10.1038/s41587-018-0011-0

Konermann, S., Lotfy, P., Brideau, N. J., Oki, J., Shokhirev, M. N., and Hsu, P. D. (2018). Transcriptome engineering with RNA-targeting type VI-D CRISPR effectors. *Cell* 173, 665.e14–676.e14. doi: 10.1016/j.cell.2018.02.033

Laganà, A., Shasha, D., and Croce, C. M. (2014). Synthetic RNAs for gene regulation: design principles and computational tools. *Front. Bioeng. Biotechnol.* 2:65. doi: 10.3389/fbioe.2014.00065

Li, M., Chen, J., Wang, Y., Liu, J., Huang, J., Chen, N., et al. (2020). Efficient multiplex gene repression by CRISPR-dCpf1 in *Corynebacterium glutamicum*. *Front. Bioeng. Biotechnol.* 8:357. doi: 10.3389/fbioe.2020.00357

Lian, J., HamediRad, M., Hu, S., and Zhao, H. (2017). Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. *Nat. Commun.* 8:1688. doi: 10.1038/s41467-017-01695-x

Liu, H., Wang, L., and Luo, Y. (2018). Blossom of CRISPR technologies and applications in disease treatment. *Synth. Syst. Biotechnol.* 3, 217–228. doi: 10.1016/j.synbio.2018.10.003

Liu, J., Wang, Y., Lu, Y., Zheng, P., Sun, J., and Ma, Y. (2017). Development of a CRISPR/Cas9 genome editing toolbox for *Corynebacterium glutamicum*. *Microb. Cell Fact.* 16:205. doi: 10.1186/s12934-017-0815-5

Lutz, R., and Bujard, H. (1997). Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res.* 25, 1203–1210. doi: 10.1093/nar/25.6.1203

Mahas, A., Aman, R., and Mahfouz, M. (2019). CRISPR-Cas13d mediates robust RNA virus interference in plants. *Genome Biol.* 20:263. doi: 10.1186/s13059-019-1881-2

Meeske, A. J., Nakandakari-Higa, S., and Marraffini, L. A. (2019). Cas13-induced cellular dormancy prevents the rise of CRISPR-resistant bacteriophage. *Nature* 570, 241–245. doi: 10.1038/s41586-019-1257-5

Miller, S. M., Wang, T., Randolph, P. B., Arbab, M., Shen, M. W., Huang, T. P., et al. (2020). Continuous evolution of SpCas9 variants compatible with non-G PAMs. *Nat. Biotechnol.* 38, 471–481. doi: 10.1038/s41587-020-0412-8

Na, D., Yoo, S. M., Chung, H., Park, H., Park, J. H., and Lee, S. Y. (2013). Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. *Nat. Biotechnol.* 31, 170–174. doi: 10.1038/nbt.2461

Otoupal, P. B., and Chatterjee, A. (2018). CRISPR gene perturbations provide insights for improving bacterial biofuel tolerance. *Front. Bioeng. Biotechnol.* 6:122. doi: 10.3389/fbioe.2018.00122

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., et al. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173–1183. doi: 10.1016/j.cell.2013.02.022

Ruan, Y., Zhu, L., and Li, Q. (2015). Improving the electro-transformation efficiency of *Corynebacterium glutamicum* by weakening its cell wall and increasing the cytoplasmic membrane fluidity. *Biotechnol. Lett.* 37, 2445–2452. doi: 10.1007/s10529-015-1934-x

Shapiro, R. S., Chavez, A., and Collins, J. J. (2018). CRISPR-based genomic tools for the manipulation of genetically intractable microorganisms. *Nat. Rev. Microbiol.* 16, 333–339. doi: 10.1038/s41579.018.0002-7

Shetty, R. P., Endy, D., and Knight, T. F. Jr. (2008). Engineering BioBrick vectors from BioBrick parts. *J. Biol. Eng.* 2:5. doi: 10.1186/1754-1611-2-5

Shmakov, S., Abudayyeh, O. O., Makarova, K. S., Wolf, Y. I., Gootenberg, J. S., Semenova, E., et al. (2015). Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. *Mol. Cell* 60, 385–397. doi: 10.1016/j.molcel.2015.10.008

Smargon, A. A., Cox, D. B. T., Pyzocha, N. K., Zheng, K., Slaymaker, I. M., Gootenberg, J. S., et al. (2017). Cas13b is a type VI-B CRISPR-associated RNA-guided RNase differentially regulated by accessory proteins Csx27 and Csx28. *Mol. Cell* 65, 618.e7–630.e7. doi: 10.1016/j.molcel.2016.12.023

Wang, F., Wang, L., Zou, X., Duan, S., Li, Z., Deng, Z., et al. (2019). Advances in CRISPR-Cas systems for RNA targeting, tracking and editing. *Biotechnol. Adv.* 37, 708–729. doi: 10.1016/j.biotechadv.2019.03.016

Wang, K., Zhao, Q.-W., Liu, Y.-F., Sun, C.-F., Chen, X.-A., Burchmore, R., et al. (2019). Multi-layer controls of Cas9 activity coupled with ATP synthase over-expression for efficient genome editing in *Streptomyces*. *Front. Bioeng. Biotechnol.* 7:304. doi: 10.3389/fbioe.2019.00304

Wang, Y., Cao, G., Xu, D., Fan, L., Wu, X., Ni, X., et al. (2018a). A novel *Corynebacterium glutamicum* L-glutamate exporter. *Appl. Environ. Microbiol.* 84:e02691-17. doi: 10.1128/aem.02691-17

Wang, Y., Liu, Y., Liu, J., Guo, Y., Fan, L., Ni, X., et al. (2018b). MACBETH: multiplex automated *Corynebacterium glutamicum* base editing method. *Metab. Eng.* 47, 200–210. doi: 10.1016/j.ymben.2018.02.016

Yan, W. X., Chong, S., Zhang, H., Makarova, K. S., Koonin, E. V., Cheng, D. R., et al. (2018). Cas13d is a compact RNA-targeting type VI CRISPR effector positively modulated by a WYL-domain-containing accessory protein. *Mol. Cell* 70, 327.e5–339.e5. doi: 10.1016/j.molcel.2018.02.028

Zhang, C., Konermann, S., Brideau, N. J., Lotfy, P., Wu, X., Novick, S. J., et al. (2018). Structural basis for the RNA-guided ribonuclease activity of CRISPR-Cas13d. *Cell* 175, 212.e7–223.e7. doi: 10.1016/j.cell.2018.09.001

Zhang, J.-L., Peng, Y.-Z., Liu, D., Liu, H., Cao, Y.-X., Li, B.-Z., et al. (2018). Gene repression via multiplex gRNA strategy in *Y. lipolytica*. *Microb. Cell Fact.* 17:62. doi: 10.1186/s12934-018-0909-8

Zheng, Y., Li, J., Wang, B., Han, J., Hao, Y., Wang, S., et al. (2020). Endogenous type I CRISPR-Cas: from foreign DNA defense to prokaryotic engineering. *Front. Bioeng. Biotechnol.* 8:62. doi: 10.3389/fbioe.2020.00062

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 31 July 2020
doi: 10.3389/fbioe.2020.00905

CRISPR-Assisted Multiplex Base Editing System in *Pseudomonas putida* KT2440

*Jun Sun, Li-Bing Lu, Tian-Xin Liang, Li-Rong Yang and Jian-Ping Wu**

Institute of Bioengineering, College of Chemical and Biological Engineering, Zhejiang University, Hangzhou, China

Pseudomonas putida (*P. putida*) KT2440 is a paradigmatic environmental-bacterium that possesses significant potential in synthetic biology, metabolic engineering and biodegradation applications. However, most genome editing methods of *P. putida KT2440* depend on heterologous repair proteins and the provision of donor DNA templates, which is laborious and inefficient. In this report, an efficient cytosine base editing system was established by using cytidine deaminase (APOBEC1), enhanced specificity Cas9 nickase (eSpCas9pp^{D10A}) and the uracil DNA glycosylase inhibitor (UGI). This constructed base editor converts C-G into T-A in the absence of DNA strands breaks and donor DNA templates. By introducing a premature stop codon in target spacers, we successfully applied this system for gene inactivation with an efficiency of 25–100% in various *Pseudomonas* species, including *P. putida* KT2440, *P. aeruginosa* PAO1, *P. fluorescens* Pf-5 and *P. entomophila* L48. We engineered an eSpCas9pp^{D10A}-NG variant with a NG protospacer adjacent motif to expand base editing candidate sites. By modifying the APOBEC1 domain, we successfully narrowed the editable window to increase gene inactivation efficiency in cytidine-rich spacers. Additionally, multiplex base editing in double and triple loci was achieved with mutation efficiencies of 90–100% and 25–35%, respectively. Taken together, the establishment of a fast, convenient and universal base editing system will accelerate the pace of future research undertaken with *P. putida* KT2440 and other *Pseudomonas* species.

Keywords: *Pseudomonas putida* KT2440, cytidine deaminase, base editing, gene inactivation, multiplex genome editing, Cas9 nickase

OPEN ACCESS

Edited by:
Chun Li,
Tsinghua University, China

Reviewed by:
Steven Lin,
Institute of Biological Chemistry, Academia Sinica, Taiwan
Mingfeng Cao,
University of Illinois at Urbana–Champaign, United States

****Correspondence:***
Jian-Ping Wu
wjp@zju.edu.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *12 March 2020*
Accepted: *14 July 2020*
Published: *31 July 2020*

Citation:
Sun J, Lu L-B, Liang T-X, Yang L-R and Wu J-P (2020) CRISPR-Assisted Multiplex Base Editing System in Pseudomonas putida KT2440. Front. Bioeng. Biotechnol. 8:905. doi: 10.3389/fbioe.2020.00905

INTRODUCTION

Pseudomonas spp. are well-known gram-negative environmental bacteria, which contain more than 200 species (Nikel et al., 2014). *Pseudomonas* spp. can inhabit in a large diversity of niches, including the surface of plants, rhizosphere, insects and even humans (Silby et al., 2011). The strong environmental adaptability and great metabolic versatility of *Pseudomonas* spp. not only contribute to its survival under harsh conditions, but also attract more research into the field (Stover et al., 2000; Moreno and Rojo, 2014). The research area of *Pseudomonas* spp. can be divided mainly into three fields: Non-pathogenic *Pseudomonas putida* (*P. putida*) KT2440 is used as a chassis for synthetic biology, metabolic engineering and biocatalysis (Poblete-Castro et al., 2012);

the opportunistic pathogen *P. aeruginosa* is regarded as a model strain for investigation of antibiotics resistance and disinfectants (Stover et al., 2000); the plant commensal *P. fluorescens* is well known for its biological control properties (Paulsen et al., 2005).

Targeted genome editing is an essential approach to exploit the physiological character of bacteria and in metabolic engineering and synthetic biology applications. The invention of counter-selection markers (*sacB* and *upp*) (Quénée et al., 2005; Graf and Altenbuchner, 2011) and heterologous recombinases (Flp and Cre) (Hoang et al., 1998; Luo et al., 2016) have increased the application of allelic exchange methods in *Pseudomonas* spp. Bacteriophage-based recombination proteins (λ-Red, Red/ET and Ssr) (Wenzel et al., 2005; Liang and Liu, 2010) are used to enhance the recombination efficiencies in *Pseudomonas* spp. The I-SceI homing endonuclease based system is a marker-free genome editing approach (Martínez-García and de Lorenzo, 2011) that has been used to delete a large DNA fragment in *P. putida* KT2440. However, genome editing using these approaches is inefficient and manipulation is time-consuming and tedious. In recent years, Cas9 assisted genome editing systems have revolutionarily accelerated the development of genetic studies in different organisms (Jakoèiunas et al., 2016), including *Pseudomonas* spp. (Aparicio et al., 2018; Sun et al., 2018; Wu et al., 2019a). Nevertheless, the developed *P. putida* CRISPR/Cas9 systems require the introduction of a heterologous recombination system because of inefficient homology-directed repair (HDR), and the provision of donor DNA templates or single-stranded DNA is also not straightforward (Aparicio et al., 2018; Sun et al., 2018). Additionally, the superior characteristics of the CRISPR/Cas9 system in multiplex genome editing has only been achieved in two loci with extremely low efficiency (Aparicio et al., 2018).

The CRISPR-assisted cytidine deaminase system is a newly developed base editing approach that uses cytidine deaminase (APOBEC1 or AID) with a catalytically impaired Cas9 to enable C:G to T:A mutations (Komor et al., 2016; Nishida et al., 2016). In the absence of double-stranded DNA breaks (DSBs) or donor DNA templates, this base editing system can directly target single-stranded DNA because Cas9 binding facilitates the recruitment of cytidine deaminases (Komor et al., 2016). By targeting CGA (Arg), CAG (Gln), and CAA (Gln) in the coding strand or ACC (Trp) in the non-coding strand, the cytidine deaminase base editor introduces a premature stop codon for gene inactivation (Wang Y.et al., 2018). However, the efficiency of cytosine base editing *in vivo* is potentially affected by the endogenous base excision repair enzyme uracil *N*-glycosylase (UNG) (Rees and Liu, 2018). To circumvent reversal of this base conversion by UNG, a bacteriophage derived uracil DNA glycosylase inhibitor (UGI) is usually fused to the C-terminus of the APOBEC1-nCas9^{D10A} complex to inhibit UNG and improve base editing efficiencies. Currently, cytidine deaminase-based base editing system have been extended to *Escherichia coli* (Nishida et al., 2016), *Clostridium beijerinckii* (Li et al., 2019), *Klebsiella pneumonia* (Yu et al., 2018), *Corynebacterium glutamicum* (Wang et al., 2019), and some *Pseudomonas* spp. (Chen et al., 2018). However, this *Pseudomonas* cytosine base editing system is highly dependent on a NGG protospacer-adjacent motif (PAM). The potential editable window is located between position −18 and −13 upstream of the PAM sequence, and base editing efficiencies are influenced by nucleotides neighboring C bases and the size of the editable window (Komor et al., 2016; Chen et al., 2018).

In this study, we have established a CRISPR-assisted templates-free base editing system (pSEVA6BE) in *P. putida* KT2440 (**Figure 1A**). This system (**Figure 1B**) comprises an enhanced specificity Cas9 nickase variant [eSpCas9pp^{D10A} containing the mutations K848A/K1003A/R1060A (Slaymaker et al., 2016)], a cytidine deaminase (APOBEC1) and a uracil DNA glycosylase inhibitor (UGI), and converted specific C nucleotides to T. By introducing a premature stop codon, this base editor was used for gene inactivation in *P. putida* KT2440 with maximum efficiencies of 100%. The cytosine base editing system was successfully extended into *P. aeruginosa* PAO1, *P. fluorescens* Pf-5 and *P. entomophila* L48. To expand the base editing scope, we introduced seven mutations [L1111R/D1135V/G1218R/E1219F/A1322R/R1335A/T1337R (Endo et al., 2019)] into eSpCas9pp^{D10A} to modify PAM specificities from NGG to NG (**Figure 1C**). To obtain precise base editing in cytidine-rich spacers, we engineered an APOBEC1 variant [mutations W90Y and R126E (Kim et al., 2017)] with a narrow editing window (**Figure 1C**). Multiplex base editing of double-locus and triple-locus in *P. putida* KT2440 was proved effective by one-plasmid and two-plasmid systems. By utilizing the base editor, gene knockout and amino acid substitution were implemented in *P. putida* KT2440 for enhancing the production of protocatechuic acid.

MATERIALS AND METHODS

Strains, Growth Conditions and Reagent

All of the strains used in this study are listed in **Supplementary Table S1**. *Escherichia coli* DH5α was used for cloning and maintenance. Luria Broth (LB) medium was used for cell growth. King's medium (18 g/L Glycerol, 20 g/L Tryptone, 1.498 g/L $MgSO_4.7H_2O$ and 0.673 g/L $K_2HPO_4.3H_2O$) was used for shaking flask fermentation. *P. putida* KT2440, *P. entomophila* L48 and *P. fluorescens* Pf-5 were grown at 30°C, while *P. aeruginosa* PAO1 and *Escherichia coli* were cultivated at 37°C. Antibiotics were added as the following concentrations: gentamicin (Gm), 50 μg/mL (*E. coli*), 100 μg/mL (*P. putida* KT2440, *P. entomophila* L48, and *P. protegens* Pf-5) and 30 μg/mL (*P. aeruginosa* PAO1); kanamycin (Km), 50 μg/mL (*E. coli*) and 100 μg/mL (*P. putida* KT2440).

Phanta Max Super-Fidelity DNA Polymerase was used for DNA amplification, and Green Taq mix was applied to colony PCR. The One Step Cloning Kit was used for seamless cloning. All of these reagents were purchased from Vazyme Biotech Co., Ltd (Nanjing, China).

Plasmid Construction

All of the plasmids and primers used in this study are listed in **Supplementary Tables S2, S3**. The genes eSpCas9pp

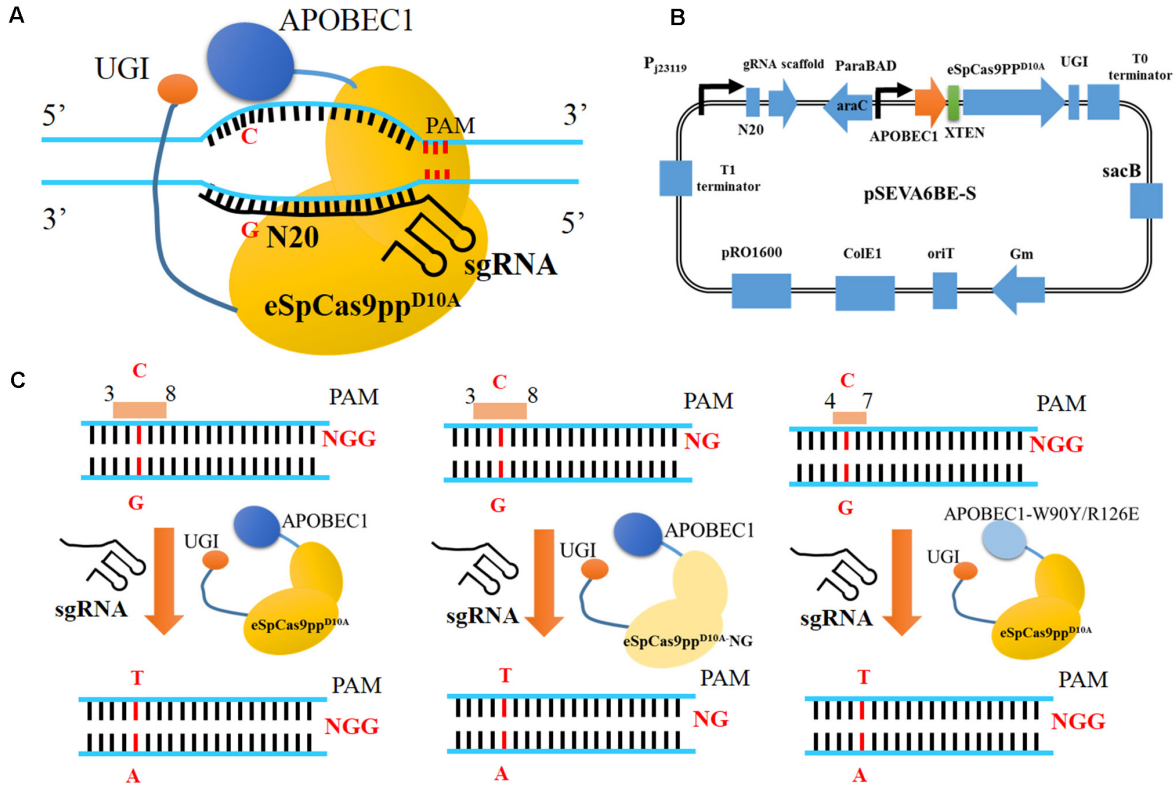


FIGURE 1 | CRISPR-Assisted cytosine base editing in *Pseudomonas putida* KT2440. **(A)** Schematic representation of CRISRP/Cas9 system-assisted base editing. **(B)** The plasmid containing APOBEC1, eSpCas9pp^{D10A}, sgRNA, UGI and *sacB* was named as pSEVA6BE-S. **(C)** Three strategies for base editing in target nucleotides (cytidines). APOBEC1, rat cytosine deaminase; N20, the 20-bp sequence at the 5′ end upstream of PAM; PAM, protospacer adjacent motif; UGI, uracil DNA glycosylase inhibitor.

[enhanced specificity Cas9 containing three mutations K848A/K1003A/R1060A (Slaymaker et al., 2016)] and APOBEC1 [cytidine deaminase (Komor et al., 2016)] were synthesized by Genscript (Nanjing, China) with codon optimized according to the codon preference of *Pseudomonas putida* KT2440. Gene SpCas9 was amplified from pCAS-RK2T (Sun et al., 2018) using primers S9-F/R. Uracil DNA glycosylase inhibitor (UGI) was cloned from pCMV-BE3 using primers U-F/R. To construct Cas9 nickase variants, the mutation of D10A in eSpCas9pp and SpCas9 was introduced by point mutations using primers eC9D10A-F/R and C9D10A-F/R, respectively. The NG PAM recognizing eSpCas9pp^{D10A}-NG was obtained by introducing seven mutations [L1111R, D1135V, G1218R, E1219F, A1322R, R1335A, and T1337R (Endo et al., 2019)] into eSpCas9pp^{D10A} using primers 11-F/R, 12-F/R, and 13-F/R. The variant APOBEC1-YE1 with a narrow editing window was obtained by site-directed mutagenesis at W90Y and R126E (Kim et al., 2017) using primers YE-1F/R.

Plasmid pSEVA-gRNAF (Sun et al., 2018) was used as template for plasmid construction. pSEVA-TtgA was modified from pSEVA-gRNAF, which was obtained by introducing a TtgA spacer. Here, we constructed 5 expression modules for cytidine deaminase-mediated base editing on the basis of pSEVA-TtgA. Among these five modules, XTEN linker (Komor et al., 2016) was used to link APOBEC1 to the N-terminus of SpCas9^{D10A} or eSpCas9pp^{D10A}. The cassettes APOBEC1-XTEN-SpCas9^{D10A} and APOBEC1-XTEN-eSpCas9pp^{D10A} under the control of the constitutive promoter Pbs (Yang et al., 2018) were inserted into pSEVA-TtgA, which generates pSEVA-Module 1 and pSEVA-Module 3, respectively. Inducible promoters Xyls-Pm and AraC-ParaBAD were amplified from pSEVA258 and pCAS-RK2T, respectively. The constitutive promoter Pbs in pSEVA-Module 1 and pSEVA-Module 3 was replaced by AraC-ParaBAD, generating pSEVA-Module 2 and pSEVA-Module 4, respectively. pSEVA-Module 5 was constructed by substituting Xyls-Pm for AraC-ParaBAD in pSEVA-Module 4. The element UGI was fused to the C-terminus of APOBEC1-XTEN- eSpCas9pp^{D10A} in pSEVA-Module 4, giving rise to pSEVA-Module 6. pSEVA-Module 6 was named pSEVA6BE, which was used as the template for the construction of following base editing plasmids (**Supplementary Table S2**). All of the base editing spacers (**Supplementary Table S4**) used in this study were designed and analyzed by CasOT (Xiao et al., 2014), gBIG (Wang et al., 2019), or BE-Designer (Hwang et al., 2018).

By replacing eSpCas9pp^{D10A} with eSpCas9pp^{D10A}-NG, a PAM-altering plasmid pSEVA6BE-NG was derived from pSEVA6BE using primers eC9NG-1F/1R and eC9NG-2F/2R. In a similar way, plasmid pSEVA6BE-YE1 was obtained by the

substitution of APOBEC1-YE1 for APOBEC1 (using primers BE-1F/1R and BE-2F/2R) in pSEVA6BE. By the combination of the backbone T1 terminator-sgRNA cassette-cytosine base editor Module 6-T0 terminator from pSEVA6BE and the fragment RSF1010 replicon-kanamycin resistance marker from pVLT33 using primers 62-1F/1R and 62-2F/2R, pSEVA2BE was constructed by seamless cloning.

The sequence of counter-selection marker *sacB* was cloned from pCAS-RK2T, and inserted into the plasmids pSEVA2BE, pSEVA6BE-YE1, pSEVA6BE-NG and pSEVA6BE via seamless cloning using primers Sac-1F/1R and Sac-2F/2R, which generated pSEVA2BE-S, pSEVA6BE-YE1-S, pSEVA6BE-NG-S, and pSEVA6BE-S, respectively.

Base Editing in *Pseudomonas*

pSEVA6BE, pSEVA6BE-NG, pSEVA6BE-YE1 or other derivative plasmids was transformed into *Pseudomonas* according to a previous electroporation method (Sun et al., 2018). After electroporation, the cells were recovered at 30°C (*P. aeruginosa* PAO1 at 37°C) for 2 h. Next, the recovered cells were concentrated and then plated onto LB agar containing antibiotics to select recombinants. The individual transformant was inoculated into a 5 mL LB tube together with the addition of antibiotics and inducers (6 mg/mL L-arabinose or 2 mg/mL *m*-toluic acid) and then cultivated at 30°C (*P. aeruginosa* PAO1 at 37°C) for 24 h. After the cultivation process, the base editing cells were streaked onto the selective plates and cultured until the appearance of visible colonies. Colony PCR was used to amplify the target PCR products and base editing results were confirmed by DNA sequencing.

In particular, pSEVA6BE-PobA (or pSEVA6BE-PobA-TrpE) and pSEVA2BE-QuiC were co-transformed into *P. putida* KT2440 by electroporation for achieving multiplex base editing. In the coelectroporation recovery process, the cultivation time was extended to 3 h. To enhance the editing efficiency in multiple loci by one-plasmid or two-plasmid system, individual transformants colonies were cultured for 36 h.

Plasmid Curing

To facilitate the loss of sacB-containing plasmids in cells, the mutant *Pseudomonas* strains were inoculated into a 5 mL LB tube containing 10 g/L sucrose and 5 g/L glucose, and then cultivated at the optimum growth temperature for 24 h. Next, the cultivated cells were streaked onto LB agar and plasmid-curing was identified by colony PCR using primers C9-F/R.

Analytical Methods

Protocatechuic acid (PCA) was purchased from Aladdin Chemistry (Shanghai, China) and used as standards. The edited *P. putida* KT2440 strains were inoculated into LB medium without adding antibiotics. The cultivated strains were used as seed cultures and then transferred into 250 mL flasks containing 50 mL King's medium under agitation rate of 250 rpm at 30°C for 66 h. The titer of PCA in *P. putida* KT2440 fermentation cultures was analyzed by HPLC (SHIMADZU, Prominence LC-20A) with a reverse phase C18 column (Pntulips^{TM} QS-C18, 5 μm, 250 × 4.6 mm). The flowing phase comprised 70% solvent A (water with 0.1% formic acid) and 30% solvent B (methanol). The optimal ultraviolet absorbance for PCA was set to 270 nm. The temperature of column oven was controlled at 40°C and the flowing rate was adjusted to 0.5 mL/min.

RESULTS

Establishment of a Cytosine Base Editing System in *P. putida* KT2440

Six expression combinations (**Figure 2A**) of APOBEC1 and Cas9^{D10A} nickase (a codon non-optimized SpCas9^{D10A} or a codon-optimized and enhanced specificity eSpCas9pp^{D10A}) were evaluated to establish an efficient cytidine base editing system in *P. putida* KT2440. Using the plasmid pSEVA-gRNAF (Sun et al., 2018) as the template, a N20 sequence (TtgA spacer targeting the *ttgA* gene in KT2440) was used as the target site and inserted into pSEVA-gRNAF to give pSEVA-TtgA.

Initially, we constructed two different Cas9^{D10A} nickase expression modules, Module 1 (containing SpCas9^{D10A}) and Module 3 (containing eSpCas9pp^{D10A}). Both of these modules were under the control of the strong constitutive promoter Pbs (Yang et al., 2018). Performance of the base editing process revealed that none of the ten selected colonies achieved cytidine mutations in the target-editing window. The constitutive promoter Pbs was exchanged with the inducible arabinose promoter, AraC-ParaBAD to generate Modules 2 and Modules 4. Modules 4 converted C to T successfully in one of the five colonies, whereas no base editing was detected for Module 2. Thus, cytosine base editing was achieved by overexpression of Module 4; however, the mutation efficiency was low. We hypothesized that cellular cytidine base editing could be reversed by base excision repair, which reduces the base editing efficiency. To verify our hypothesis, the uracil DNA glycosylase inhibitor (UGI) was fused to the C-terminus of APOBEC1-eSpCas9pp^{D10A}to yield the cassette APOBEC1-eSpCas9pp^{D10A}-UGI. Two common inducible promoters Xyls-Pm and AraC-ParaBAD, were introduced to control this cassette to give Modules 5 and 6. After the base editing procedure, DNA sequencing showed that three C-to-T mutants were identified from five random colonies harboring Module 5, and all five randomly selected strains containing Modules 6 converted C to T, which indicates that Module 6 exhibited a higher efficiency than Module 5. To verify the effects of adding UGI, we designed another two spacers GllA and MexE as target sites by using pSEVA-Module 4 and pSEVA-Module 6. DNA sequencing results (**Supplementary Figures S1, S2**) showed that the mutation rate of cytosine to guanine for spacers GllA and MexE increased from 60 and 40% in Module 4 to 100 and 80% in Module 6, which proves that the addition of UGI improved the editing efficiency.

DNA sequencing results of colonies containing Module 6 detected three different cytidine substitutions (**Figure 2B**). Among the five C→T mutations, 3/5 base editing substitution occurred at positions 3 and 4, 1/5 occurred at position 8 and 1/5 at position 4, with the PAM sequence at position 21–23. The editing efficiency of cytidines (TC ≥ CC ≥ AC > GC) was in consistent with previous reports in mammalian cells (Komor et al., 2016).

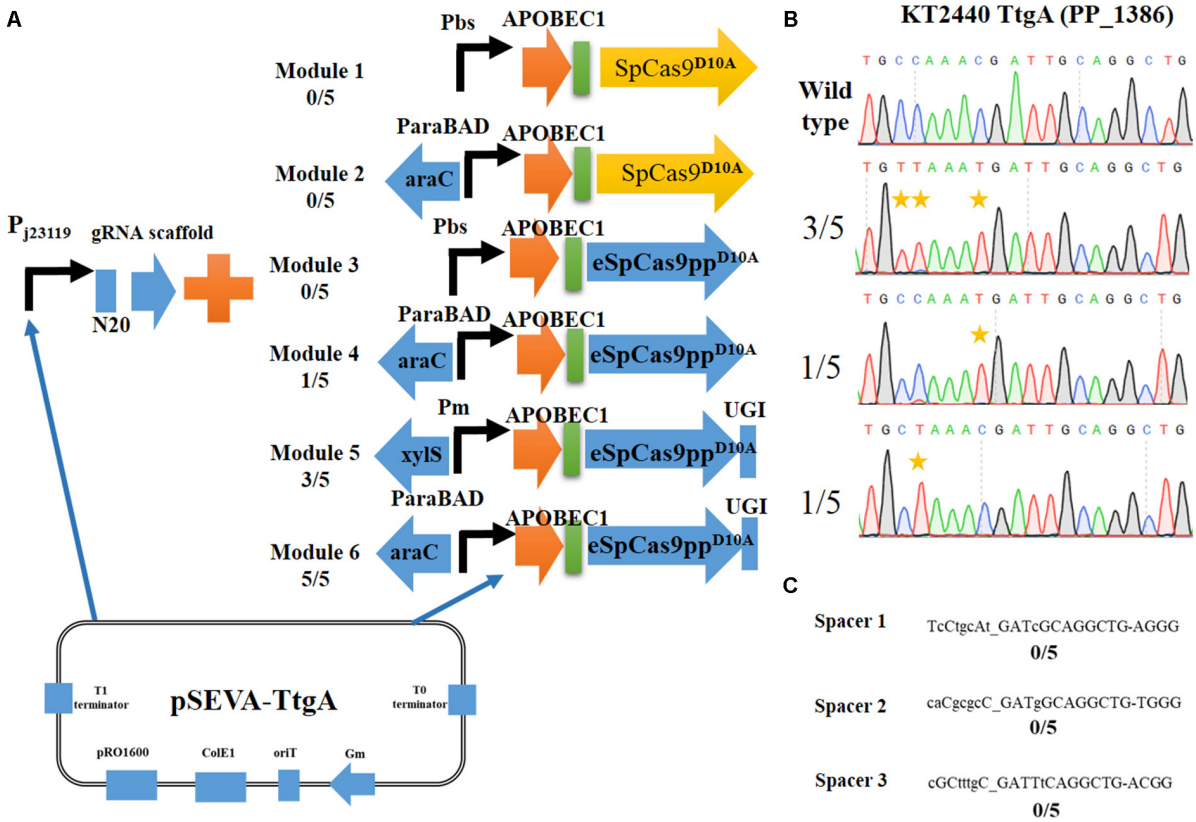


FIGURE 2 | Construction of an efficient cytosine base editing system in *Pseudomonas putida* KT2440. **(A)** Six different modules for cytosine base editing designed in this study. **(B)** Determination of pSEVA-Module 6 (pSEVA6BE) mediated cytosine base editing in the TtgA spacer region. **(C)** Potential off-target loci of TtgA in *P. putida* KT2440 genome.

Mutations at potential off-target loci by the Module 6-containing base editing system were evaluated. With the help of CasOT (Xiao et al., 2014), three top-similar DNA sequences (**Figure 2C**) of this *ttgA* N20 sequence were identified in the *P. putida* KT2440 genome and DNA sequencing of these spacer sites showed that none of them had point mutations (**Supplementary Figure S3**). These results confirmed that the Module 6-containing cytosine base editing system exhibited high specificity. Thus, Module 6 was identified as the optimal expression cassette. The plasmid containing Module 6 was named pSEVA6BE and used as the backbone for the following base editing experiments.

Application of the Cytosine Base Editing System for Gene Inactivation in *P. putida* KT2440

Having demonstrated the feasibility of pSEVA6BE for cytosine base editing in *P. putida* KT2440, we sought to apply this system for gene inactivation by introducing a premature stop codon. Because of its capability to convert codons CAA, CAG, CGA, and TGG into a stop codon, the cytidine base editing system can be used to inactivate genes without generating DSBs and the provision of DNA repairing templates (Arazoe et al., 2018). To assess the efficiencies of gene inactivation, we selected *hmgA*, *pobA*, *quiC*, and *ttgA* as target sites.

HmgA, encoding homogentisate dioxygenase, is a key gene of the homogentisate pathway (**Figure 3A**) in *P. putida* KT2440 (Arias-Barrau et al., 2004). The deletion of *hmgA* can disable the ring-cleavage reaction of homogentisate, which leads to the accumulation of homogentisate. As homogentisate can be oxidized into dark brown by oxygen, we used *hmgA* as a reporter to assess the efficiency of gene inactivation in *P. putida* KT2440. After the plasmid pSEVA6BE-HmgA targeting *hmgA* was introduced into *P. putida* KT2440, the transformant colonies were inoculated, cultivated and streaked on selective LB agar plates. As shown in **Figure 3C**, mutant cells produced the dark brown pigment when compared with that of the control wild-type strain. Five random mutant colonies were selected for colony PCR and DNA sequencing. The sequencing results (**Figure 3B**) showed that the codon CAG representing residue Gln32 was successfully mutated to TAG in four out of five colonies, thus disabling the activity of HmgA. In the case of PobA, the N20 sequence harbors two potential editable cytidines (Cs) at positions 5 and 7. To achieve gene inactivation of *pobA*, substitution of C→T should occur at position 7, or positions 5 and 7. DNA sequencing of the base editing results (**Figure 3D**) showed that all five randomly picked colonies carried the C to T

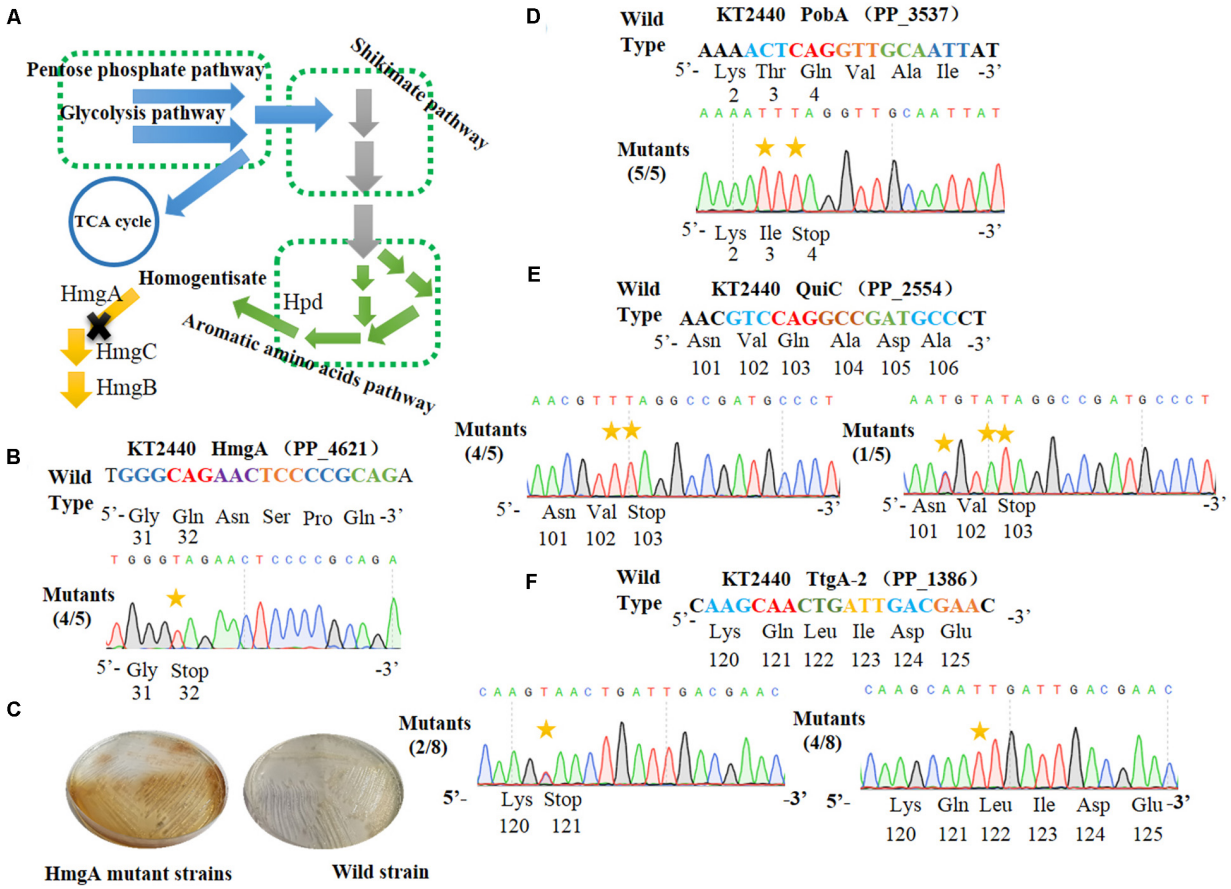


FIGURE 3 | Cytosine base editor-mediated gene inactivation in *Pseudomonas putida* KT2440. **(A)** Overview of the homogentisate pathway in *P. putida* KT2440. **(B)** Mutation alignment of the targeted HmgA spacer. **(C)** Phenotypes of *hmgA* knockout mutants. The *hmgA* mutant strains of *P. putida* KT2440 were streaked on selective plates. **(D)** The PobA spacer in *pobA* gene was selected for base editing to achieve gene inactivation. **(E)** Base editing results of Quic spacer in *quiC* gene. **(F)** Sequence alignment of the TtgA mutants after cytosine base editing.

mutation, yielding an editing efficiency for *pobA* of 100%. The N20 sequence from *quiC* contains three editable Cs at positions of 3, 6, and 7. Base editing results (**Figure 3E**) showed that all five identified strains had successfully performed gene inactivation by introduction of a stop codon at amino acid position 103. Among these mutants, 80% of the strains had the C→T substitution at positions of 6 and 7, and the remaining 20% harbored two C→T mutations (at positions 3 and 7) and one C→A mutation (at position 6) within the base editing window. Next, the plasmid pSEVA6BE-TtgA-2 was transformed into *P. putida* KT2440 to test the editing efficiency of the TtgA-2 spacer. The possible editable Cs in the TtgA-2 spacer are located at positions of 5 and 8. The results of colony PCR (**Figure 3F**) showed that half of the eight selected strains exhibited the C→T substitution at position 8, and only 25% of the strains were identified to achieve gene inactivation by mutating CAA to TAA at position 5. The high mutation efficiencies in *hmgA*, *pobA*, and *quiC* demonstrated that the cytosine base editing system is an efficient tool for gene inactivation in *P. putida* KT2440. The editing results of TtgA-2 revealed the discrepancy of base editing toward cytosines with different adjacent nucleotides, which indicates that the AC motif had a higher base editing efficiency than GC. This is in agreement with the previously reported mutation preference of APOBEC1 (Chen et al., 2018).

Cytosine Base Editing in *P. aeruginosa* PAO1, *P. fluorescens* Pf-5 and *P. entomophila* L48

To test the universality of our cytosine base editor in *Pseudomonas* species, we tried to extend the pSEVA6BE system into *P. aeruginosa* PAO1, *P. fluorescens* Pf-5 and *P. entomophila* L48. *PA1236* (encoding probable major facilitator superfamily transporter) and *PA2018* (encoding resistance-nodulation-cell division multidrug efflux transporter) were selected as target sites for investigation of knock-out efficiency in *P. aeruginosa* PAO1. After base editing, DNA sequencing (**Figure 4A**) of PCR products showed that TAG and TAA could be generated with efficiencies of 100% and 80%, respectively. In *P. fluorescens* Pf-5, *PFL_0054* (encoding gluconate 2-dehydrogenase) and *PFL_0556* (encoding flagellar motor protein MotB) were designed as target base editing sites, respectively. As shown in **Figure 4B**, the codons CAG

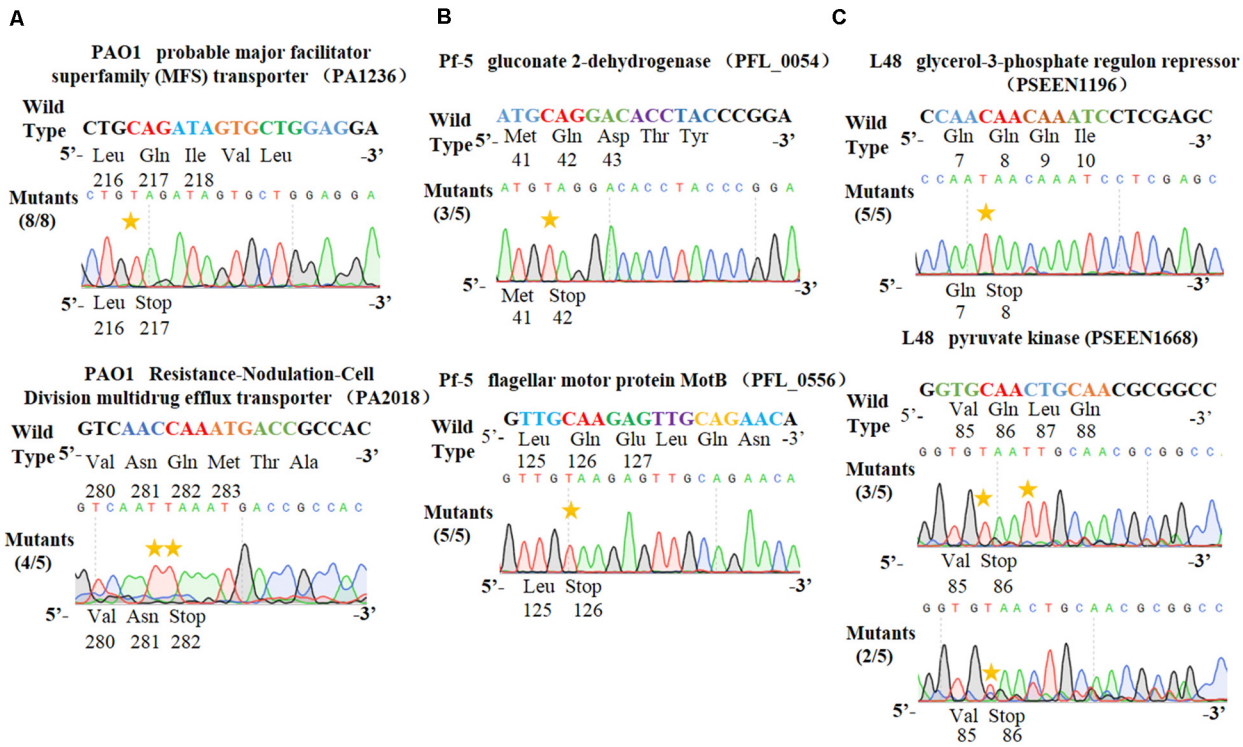


FIGURE 4 | Application of the PSEVA6BE system into *Pseudomonas aeruginosa* PAO1, *Pseudomonas fluorescens* Pf-5 and *Pseudomonas entomophila* L48. **(A)** Examples of gene activation in *P. aeruginosa* PAO1 by cytosine base editor. **(B)** Results of cytosine base editing experiments in *P. fluorescens* Pf-5. **(C)** Examination of pSEVA6BE-mediated gene knockout in *P. entomophila* L48.

(Gln) and CAA (Gln) could be modified to the stop codons TAG and TAA with efficiencies of 3/5 and 5/5, respectively. Plasmids pSEVA6BE-L48glpR and pSEVA6BE-L48pyrF were designed to target the *PSEEN1196* and *PSEEN1668* genes in *P. entomophila* L48, respectively. Base editing results (**Figure 4C**) showed that C→T substitution could be achieved at position 5 in spacers PSEEN1196 and PSEEN1668 which resulted in knockout of these two genes. The application of the cytosine base editor in these three *Pseudomonas* species demonstrated the pSEVA6BE system can be a convenient and highly efficient genome editing tool in a wide range of *Pseudomonas* species.

Expansion of the Base Editing Candidate Sites by Modification of PAM Specificities

The cytosine base editor pSEVA6BE requires a strict PAM motif (SpCas9 recognizing NGG) and the target nucleotides (cytidines) must be distributed within the editable window. To expand the cytosine base editing candidate sites, the plasmid pSEVA6BE-NG containing an eSpCas9pp^{D10A}-NG was constructed, which is capable of recognizing NG PAM. To that end the mutations [L1111R/D1135V/G1218R/E1219F/A1322R/R1335A/T1337R (Endo et al., 2019)] were introduced into the eSpCas9pp^{D10A} of pSEVA6BE (**Figure 5A**). To assess the application of pSEVA6BE-NG in *Pseudomonas* species, we selected two target spacers (HexR-2 and HexR-3 were inserted into pSEVA6BE-NG) with NG PAM in *hexR* using the online tool gBIG, and pSEVA6BE derivative plasmids were used as control. The results of HexR-2 and HexR-3 showed the conversion of C to T was present at the target sites with a CG or AG PAM (**Figure 5B**). Gene inactivation of both sites was achieved with an efficiency of 100%. Conversely, the pSEVA6BE-derived plasmids could not recognize non-NGG PAM target sites, thus all of these control strains failed to mutate C to T (**Supplementary Figure S4**). The application of pSEVA6BE-NG expands the candidate sites of the cytosine base editor in *P. putida* KT2440 and is equally applicable to other *Pseudomonas* species.

Narrowing the Editable Window for Precision Editing

The CRISPR-assisted cytosine base editor exhibits a wide editable window width mainly from position 3 to position 8 (Chen et al., 2018; Rees and Liu, 2018), counting the first 5′ end nucleotide in the N20 sequence as position 1. The wide editing window allows substitution of multiple cytidines into thymines, but can also cause unpredictable editing effects when the desired target cytidine mutation is located among multiple cytidines. For example, the spacer HexR (5′-GCCCGGCAGATCCACTTCTT-3′) derived from *P. putida* KT2440 was selected as the target site for gene inactivation by the cytosine base editor. To achieve gene inactivation of *hexR*, the desired cytidine mutation should

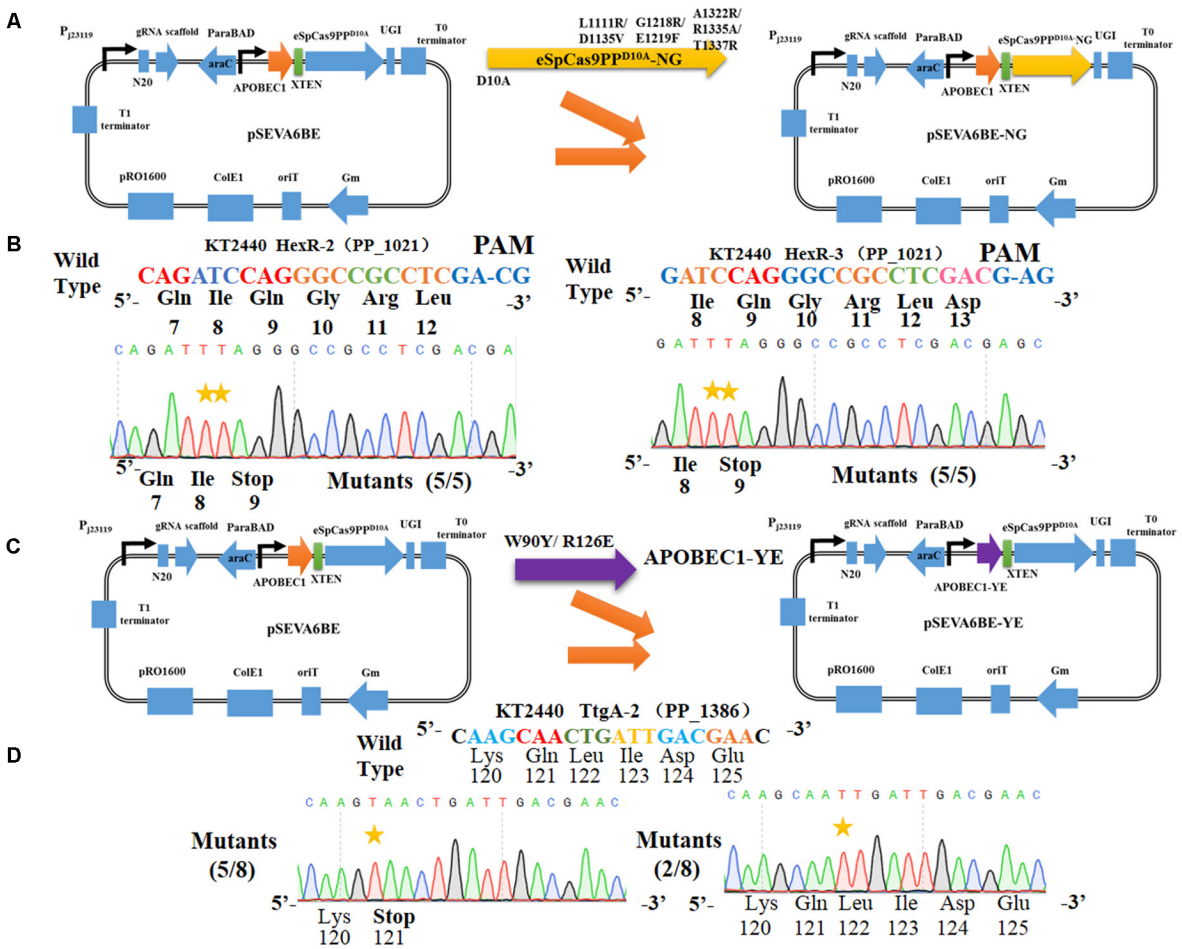


FIGURE 5 | Altering the PAM specificity of eSpCas9PP^{D10A} and narrowing the editable window of APOBEC1. **(A)** The eSpCas9pp^{D10A}–NG gene was cloned into pSEVA6BE, generating pSEVA6BE-NG. **(B)** C-to-T substitution was successfully achieved in target spacers HexR2 and HexR3 with NG PAM sequences. **(C)** The APOBEC1 gene in pSEVA6BE was replaced by the APOBEC1-YE1 gene, thus creating pSEVA6BE-YE. **(D)** pSEVA6BE-YE1-mediated base editing of TtgA-2 spacer tend to have a narrow editing windows compared to using pSEVA6BE (**Figure 3F**).

occur at position 7 by mutating CAG to TAG. However, DNA sequencing (**Supplementary Figure S5**) showed that six in ten edited colonies possessed C→T mutations at positions 3, 4, and 12, one colony had cytidines mutation at positions 4 and 18, and only three of the ten colonies displayed gene inactivation with cytidine substitutions at positions 3, 4, and 7. In another example presented in **Figure 3F**, the knockout efficiency of *ttgA* was only 25% when the editing window contained multiple cytidines. We hypothesize that a wide editing window can decrease the efficiency of gene inactivation by the cytosine base editor within a cytidine-rich target site due to the catalytic preference of APOBEC1 toward different NC motifs (in the order of TC ≥ CC ≥ AC > GC).

To obtain precise gene inactivation efficiency, we sought to narrow the editable window by introducing two mutations [W90Y and R126E (Kim et al., 2017)] into the APOBEC1 domain, generating pSEVA6BE-YE1 from pSEVA6BE (**Figure 5C**). Next, the spacer TtgA-2 was inserted into pSEVA6BE-YE1 to yield pSEVA6BE-YE1-TtgA-2. After base editing of pSEVA6BE-YE1-TtgA-2 in *P. putida* KT2440, we observed that the stop codon TAA was introduced successfully into five out of eight tested colonies (**Figure 5D**). By using pSEVA6BE-YE1, the knockout efficiency of the TtgA-2 spacer containing multiple cytidines increased from 25 to 62.5% in compared with using pSEVA6BE-TtgA. The construction of a cytosine base editor with a narrower editing window provides precision base editing of cytidine-rich sites.

A Multiplex Base Editing System in *Pseudomonas putida* KT2440

A unique characteristic of the CRISPR/Cas9 system is the feasibility of concurrent multiplex genome editing, yet genome editing of two loci or more in eukaryotes has been the primary focus (Jakoèiunas et al., 2016) and few studies of multiplex genome editing have been reported for prokaryotes, which can be ascribed to the weak HDR and poor or lack of non-homologous end joining (NHEJ) repair (Wu et al., 2019b). Although

multiplex genome editing of two loci by the CRISPR/Cas9 system had been tested in *P. putida* KT2440, the efficiency was extremely low and could not be applied readily in experiments (Aparicio et al., 2018).

Two multiplex base editing systems were constructed, that is, a one-plasmid system and two-plasmid system, to explore multiplex genome editing in *P. putida* KT2440. By connecting the base-editing cassette from pSEVA6BE with the broad-host-replicon RSF1010 and kanamycin-resistance marker from PVLT33, a kanamycin version of the base editing system pSEVA2BE (**Supplementary Figure S6**) was constructed and is compatible with pSEVA6BE in a two-plasmid system

In this study, multiplex base editing of a double-locus was tested initially for genomic loci *pobA* and *quiC*. Spacers PobA and QuiC-2 were inserted into pSEVA6BE and pSEVA2BE, respectively to generate the two-plasmid systems pSEVA6BE-PobA and pSEVA2BE-QuiC-2. An expression cassette containing the J23119 promoter, QuiC-2 spacer and sgRNA scaffold was inserted into pSEVA6BE-PobAto give the one-plasmid system pSEVA6BE-PobA-QuiC-2. After electroporation of the resulting plasmids, transformants were selected and cultivated according to base editing procedure. To confirm the substitution of C→T in the two loci (PobA and QuiC-2), ten randomly picked colonies in each system were used as DNA templates for colony PCR. The editing efficiency of the double-locus was 100 and 90% for the one-plasmid system and two-plasmid system (**Figure 6A** and **Supplementary Figure S7**), respectively.

Next, simultaneous base editing at three loci was tested using spacers TrpE, PobA and QuiC-2. *TrpE* encoding anthranilate

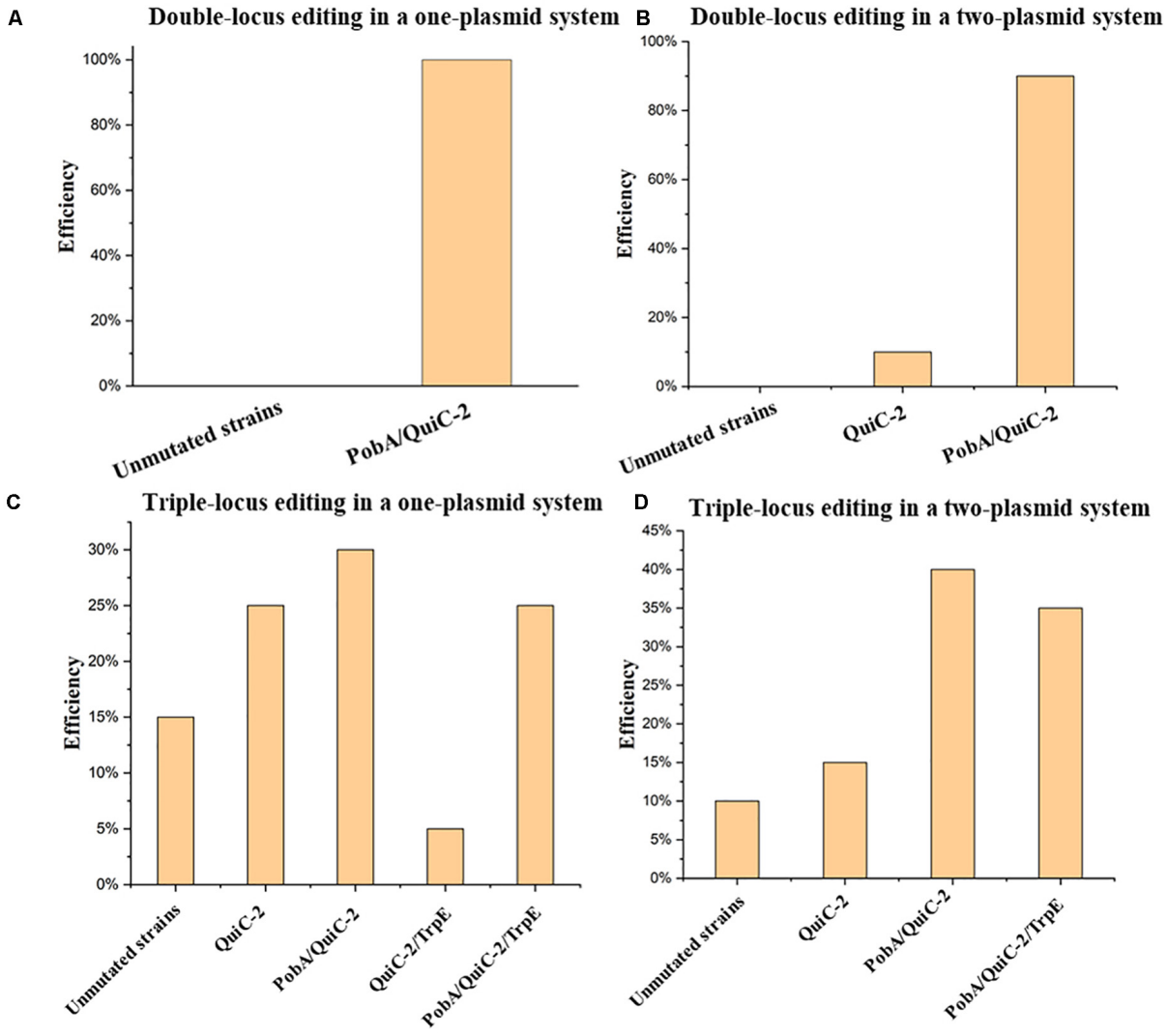


FIGURE 6 | Assessment of multiplex base editing system in *Pseudomonas. putida* KT2440. **(A)** Double-locus editing of spacers PobA and QuiC-2 in the one-plasmid system pSEVA6BE-PobA-QuiC-2. **(B)** CRISPR-mediated double-locus base editing toward spacers PobA and QuiC-2 using a two-plasmid system. **(C)** Triple-locus base editing profiles (PobA, QuiC-2 and TrpE) using the one-plasmid system pSEVA6BE- PobA-QuiC-2-TrpE. **(D)** The two-plasmid system pSEVA6BE-PobA-TrpE/pSEVA2BE-QuiC-2-mediated triple-locus base editing in the *P. putida* KT2440.

synthase (TrpE) was selected as the third target site for multiple base editing. The TrpE editing cassette was inserted into pSEVA6BE-PobA-QuiC-2 and pSEVA6BE-PobA by using the same construction strategy to generate pSEVA6BE-PobA-QuiC-2-TrpE and pSEVA6BE-PobA-TrpE. After transformation of the one-plasmid (pSEVA6BE-PobA-QuiC-2-TrpE) or two-plasmid system (pSEVA6BE-PobA-TrpE and pSEVA2BE-QuiC-2) into *P. putida* KT2440 cells, the transformant colonies were identified, inoculated and cultivated in accordance with the previous cytosine base editing procedure. The simultaneous triple-locus gene inactivation efficiency was 25% (2/10 + 3/10) and 35% (4/10 + 3/10) for the one-plasmid and two-plasmid systems, respectively (**Figure 6B** and **Supplementary Figure S8**). The above results indicate that editing efficiency negatively correlates with the number of targeted spacers.

In our study, we investigated two kinds of system in multiplex base editing. The editing efficiency of double-locus and triple-locus showed no apparent difference between the one-plasmid and two-plasmid systems. The one-plasmid system gave a higher plasmid transformation efficiency when compared with that of the two-plasmid system. The two-plasmid system has the advantage of expressing different Cas9 proteins to recognize spacers with different PAM (**Supplementary Figure S9**). To test the feasibility of the two-plasmid system, the NG-recognize pSEVA6BE-NG-HexR3 and NGG-recognize pSEVA2BE-QuiC-2 were coelectroporated into *P. putida* KT2440. The sequencing results (**Supplementary Figure S10**) showed base editing can be achieved simultaneously in NG-spacer HexR3 and NGG-spacer QuiC-2 with an efficiency of 100%. The combination of different PAM-specificity Cas9 proteins in the two-plasmid system enriched the base editing scope of the one-step experiment.

Together, a fast and convenient multiplex base editing system has been established, which should facilitate construction of mutant libraries for metabolic engineering and synthetic biology in *P. putida* KT2440.

Modification of *P. putida* KT2440 for Production of Protocatechuic Acid Using the Cytosine Base Editor

To construct a universal plasmid curing strategy for a base editing system in *Pseudomonas* spp., we inserted a counter-selection marker *sacB* into pSEVA6BE, pSEVA6BE-NG, pSEVA6BE-YE1 and pSEVA2BEto give pSEVA6BE-S, pSEVA6BE-NG-S, pSEVA6BE-YE1-S and pSEVA2BE-S (**Supplementary Figure S11**), respectively.

In addition to gene knockout, this base editing system can be used to make point mutations that relieve feedback inhibition in metabolic engineering. The application of our base editing system for metabolic engineering was tested by modifying *P. putida* KT2440 for the production of protocatechuic acid (PCA). The genes encoding protocatechuate 3,4-dioxygenase subunit beta (pcaH) and pyruvate kinase (pykA) (**Figure 7A**) were selected as knockout sites for enhancing the accumulation of PCA (Wang W.et al., 2018). The introduction of the G136E mutation was used to relieve feedback inhibition of the 3-deoxy-D-arabinoheptulosonate-7-phosphate (DAHP) synthase isozyme AroF-2 (Wynands et al., 2018). Cassettes targeting PcaH and PykA were inserted into pSEVA6BE-NG-S to create a double-locus editing plasmid pSEVA6BE-NG-pcaHpykA-S. The NG-PAM spacer AF1, which is adjacent G136 in AroF-2, was inserted into pSEVA6BE-NG-S to give pSEVA6BE-NG-AF1-S.

The *P. putida* KT2440 strain KT2440::AroF2-P_{J23119} (containing a strong constitutive promoter P_{J23119} inserted in the upstream region of AroF-2) was used as the starting strain for the production of PCA. We obtained the mutant strain KT2440::AroF2-P_{J23119}ΔpcaHΔpykA after double-locus base editing using pSEVA6BE-NG-pcaHpykA-S. After curing of the double-locus plasmid, pSEVA6BE-NG-AF1- S was transformed into KT2440::AroF2-P_{J23119}ΔpcaHΔpykA for achieving point mutations. The base editing results (**Figure 7B**) showed that mutations G136E, G136K, and G136R were introduced into genomic loci with efficiencies of 20, 20, and 10%, respectively. Next, the edited, unedited and starting strains were cultivated and used as seed cultures to transfer into King's medium for shaking-flask experiments. After 66 h cultivation, the shaking flask fermentation showed that the production of PCA in the G136E mutation strain was 264.87 mg/L, which is an increase by 69.01 and 611.17% when compared with that of the AroF-2 unedited strain and the starting strain (**Figure 7C**). This study proved that the base editing system is a convenient tool for genomic modification in metabolic engineering.

DISCUSSION

Pseudomonas putida KT2440 is a potential chassis for industrial production of bio-based materials, pharmaceuticals and chemicals (Nikel et al., 2016). However, traditional genetic tools in *P. putida* KT2440 are difficult to manipulate and inefficient. The emergence of the CRISPR/Cas9 genome editing system has greatly simplified genetic engineering of *P. putida* KT2440, but still requires the provision of heterologous repair proteins and donor DNA templates (Aparicio et al., 2018; Sun et al., 2018). Additionally, simultaneous genome editing of two loci remains extremely challenging using the CRISPR/Cas9 system.

In the absence of templates and heterologous recombinases, the CRISPR-assisted cytosine deaminase system can achieve gene knockout or amino acid substitution, which is a quick and easy approach. Initially, the feasibility of cytosine base editing by expression of different cytidine deaminase modules was tested. After electroporation, strains containing pSEVA-Module 2 or pSEVA-Module 4 were cultivated at 30°C for 2–4 h after adding an inducer and then spread on LB agar containing antibiotics and inducer. However, all of the picked colonies from the plates were identified as wild-type strains. Based on this result and previous studies (Wang Y.et al., 2018; Li et al., 2019), we hypothesized that cytosine base editing could not fully perform its function in the initial transformant colonies, and that the cytosine base editor mutated only a fraction of the strains. By subculturing, we observed the C→T substitution of the target site, which demonstrated that a catalytic process for gradual accumulation was required for the generation of base editing mutants using the pSEVA6BE system in *P. putida* KT2440 cells.

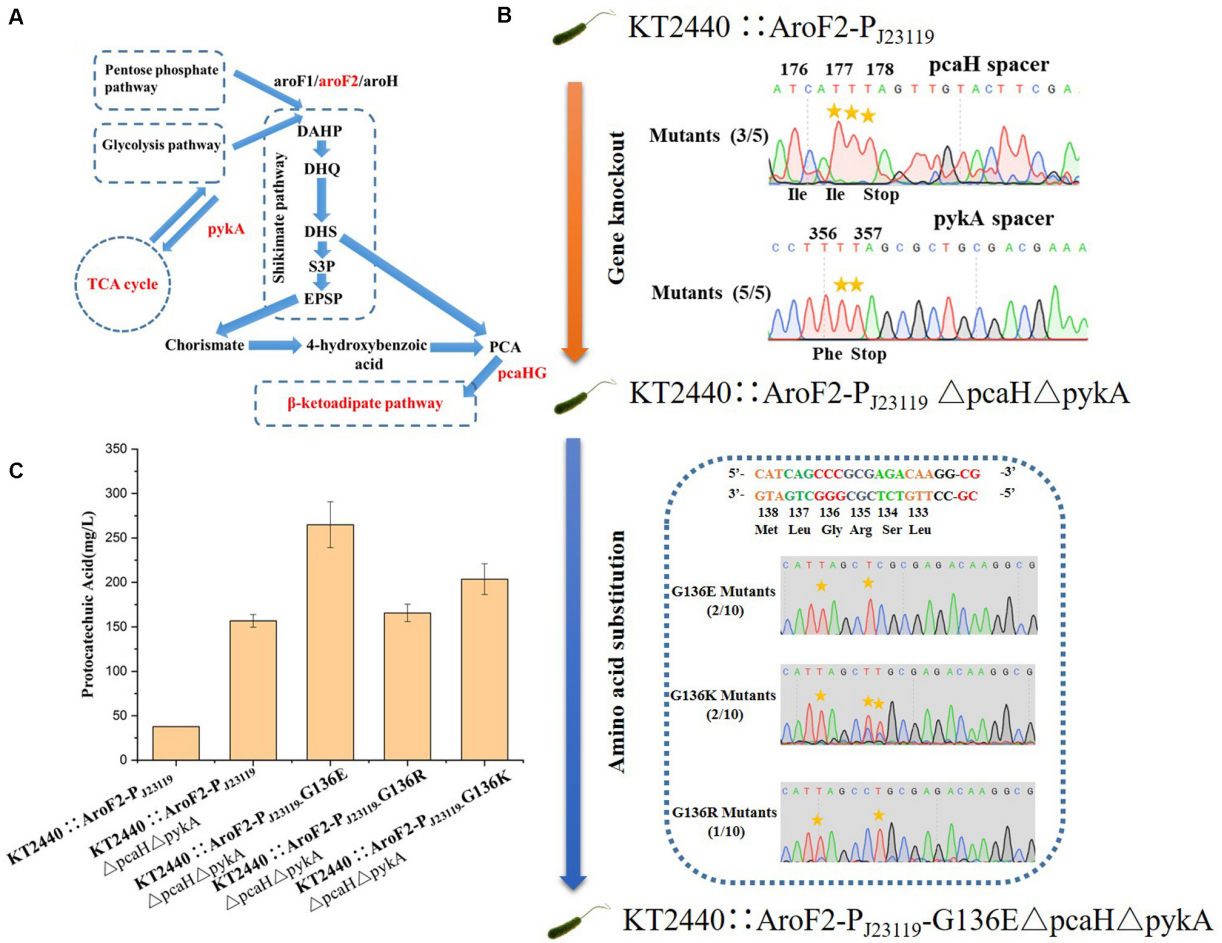


FIGURE 7 | Protocatechuic acid (PCA) production titer in KT2440::AroF2-P_{J23119}-derived strains after modification by the base editor. **(A)** A schematic representation of PCA synthesis in *P. putida* KT2440. **(B)** Genomic modification of KT2440::AroF2-P_{J23119} by gene knockout and amino acid substitution using the base editor. **(C)** The PCA titer in different KT2440::AroF2-P_{J23119}-derived strains after shaking-flask fermentation in King's medium. Abbreviations: DAHP, 3-deoxy-D-arabinoheptulosonate-7-phosphate. DHQ, 3-dehydroquinate. DHS, 3-dehydroshikimate. S3P, shikimate-3-phosphate. EPSP, 5-enolpyruvyl-shikimate-5-phosphate. PCA, protocatechuic acid.

In our study, we not only successfully developed the pSEVA6BE system as a knockout tool for *P. putida* KT2440, but also expanded the base-editing scope into *P. aeruginosa* PAO1, *P. fluorescens* Pf-5 and *P. entomophila* L48. *P. fluorescens* Pf-5 shows a great potential for application in biocontrol because of this strain possesses a wealth of antibacterial secondary metabolites (Paulsen et al., 2005). *P. entomophila* L48 is a model strain used for research of insect pathogenesis, which is hypothesized to produce hydrogen cyanide and novel secondary metabolites (Vodovar et al., 2006). Exploring the CRISPR-assisted cytosine base editing system scope in *P. fluorescens* Pf-5 and *P. entomophila* L48 will simplify genetic engineering of these strains and therefore advance research progress involving these strains.

During our study, Chen et al. (2018) reported a CRISPR-assisted cytosine base editor in multiple *Pseudomonas* strains, including *P. putida* KT2440 and *P. aeruginosa* PAO1. In this base editor, the expression of sgRNA and APOBEC1-nCas9 are under the control of constitutive promoters. The substitution of C to T can be identified from initial transformants after electroporation, which eliminates the required subculture step for colony generation. However, two limiting factors restrict its application in *Pseudomonas*: (i) High efficiency base editing was confined mainly to TC and CC motifs, with moderate efficiency toward the AC motif and no editing of the GC motif. (ii) The requirement of a NGG PAM reduces the candidate sites in target loci.

In our base editing system pSEVA6BE, cytosine base editing was achieved on all four NC motifs. The editing differences of Chen's base editing system (Chen et al., 2018) and ours on the four motifs are possibly because of the different expression types of the APOBEC1-nCas9 complex. In our study, sgRNA and APOBEC1-nCas9-UGI were under the control of the PJ23119 constitutive promoter and arabinose-inducible promoter, respectively. The inducible expression of APOBEC1-nCas9-UGI generally yields higher expression levels when

compared with that of the constitutive expression type used in the system reported by Chen et al. (2018), thereby achieving substitution of C to T in GC motif, even though this motif had the lowest catalytic order for APOBEC1. To expand the editing scope of the cytosine base editor, we constructed a NG-recognition base editor pSEVA6BE-NG by altering the PAM specificity of eSpCas9pp^{D10A}. By utilizing the pSEVA6BE-NG system, we introduced a premature stop codon into the *HexR* gene with an efficiency of 100% (**Figure 5A**). Multiplex base editing of double and triple loci was validated by one-plasmid and two-plasmid systems, which proved to be a convenient method for double-deletion of the *pykA* and *pcaH* genes (**Figure 7B**). Taken together, our base editing system is a significant extension of the previous base editor (Chen et al., 2018) in *Pseudomonas*.

Although the base editor possesses C→T mutations on all four NC motifs, we observed that base editing could not be achieved or inefficient to realize in randomly selected spacers PA2018-2 (targeting the *PA2018* gene), HexR-5 (targeting the *PP_1021* gene) and HexR-4 (targeting the *PP_1021* gene), which contains a GC motif in the editable window (**Supplementary Figures S12, S13**). The catalytic preference of APOBEC1 toward different motifs (TC ≥ CC ≥ AC > GC) or off-target effects, or both of them were probably the major causes. To overcome these potential limitations, spacers PA2018, HexR, HexR-2, and HexR-3 were screened from these target genes by gBIG. By utilizing gBIG, these newly screened spacers PA2018, HexR, HexR-2, and HexR-3 do not contain a GC motif in target cytidines and the chance of off-target at these spacers was minimized. After the base editing procedure, DNA sequencing showed that spacers HexR (**Supplementary Figure S5**), PA2018 (**Figure 4A**), HexR-2 and HexR-3 (**Figure 5B**) exhibited significantly higher mutation efficiencies than the original spacers HexR-5, PA2018-2, and HexR-4. To minimize the motif preference of APOBEC1 or the disruption of off-target effects, the design of the 20-bp spacer with a NGG or NG PAM for gene inactivation should be assessed by base editing designing tools (Xiao et al., 2014; Hwang et al., 2018; Wang et al., 2019) in an effort to overcome this potential limitation (Wang et al., 2019).

Multiplex gene editing provides a fast and convenient tool for generating mutant strain libraries of *P. putida* KT2440. However, there are few studies reporting the use of CRISPR/Cas9-mediated multiplex genome editing in *Pseudomonas*, probably because of the lethality caused by DSBs or the relatively low efficiency of recombination afforded by recombination proteins. In our study, multiplex base editing in double and triple loci was achieved by using a one-plasmid or two-plasmid cytosine base editing system, which was simple and effective. By utilizing multiplex base editing system, double-deletion of the *pykA* and *pcaH* genes was achieved in one step experiment to generate the PCA-accumulating strain KT2440::AroF2-P_{J23119}ΔpcaHΔpykA. On the basis of the resulting strain, amino acid substitutions were successfully introduced to relieve feedback resistance of AroF-2, which increased the PCA titer to 264.87 mg/L.

In conclusion, this study demonstrated that a high-efficient CRISPR-assisted cytosine base editing system can be used to achieve multiplex gene editing in *P. putida* KT2440. Additionally, the accessibility and universality of the cytosine base editor was successfully extended to *P. aeruginosa* PAO1, *P. fluorescens* Pf-5 and *P. entomophila* L48. The established cytosine base editing system is an efficient tool to facilitate future research of *Pseudomonas* species.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher. Plasmids pSEVA6BE-S, pSEVA6BE-NG-S, pSEVA2BE-S, and pSEVA6BE-YE1-S are available on Addgene (Addgene IDs: 155266–155269).

AUTHOR CONTRIBUTIONS

JS, L-RY, and J-PW designed the experiments. JS, L-BL, and T-XL performed the experiments. JS analyzed the results and wrote the manuscript. All the authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 21676240) and the National Key R&D Program of China (No. 2019YFA09005000).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00905/full#supplementary-material

REFERENCES

Aparicio, T., De Lorenzo, V., and Martínez-García, E. (2018). CRISPR/Cas9-based counterselection boosts recombineering efficiency in *Pseudomonas* putida. *Biotechnol. J.* 13, 1–10. doi: 10.1002/biot.201700161

Arazoe, T., Kondo, A., and Nishida, K. (2018). Targeted nucleotide editing technologies for microbial metabolic engineering. *Biotechnol. J.* 13, 1–36. doi: 10.1002/biot.201700596

Arias-Barrau, E., Olivera, E. R., Luengo, J. M., Fernández, C., Galán, B., García, J. L., et al. (2004). The homogentisate pathway: a central catabolic pathway involved in the degradation of L-phenylalanine, L-tyrosine, and 3-hydroxyphenylacetate in *Pseudomonas* putida. *J. Bacteriol.* 186, 5062–5077. doi: 10.1128/JB.186.15.5062-5077.2004

Chen, W., Zhang, Y., Zhang, Y., Pi, Y., Gu, T., Song, L., et al. (2018). CRISPR/Cas9-based genome editing in *Pseudomonas aeruginosa* and cytidine deaminase-mediated base editing in *Pseudomonas* species. *iScience* 6, 222–231. doi: 10.1016/j.isci.2018.07.024

Endo, M., Mikami, M., Endo, A., Kaya, H., Itoh, T., Nishimasu, H., et al. (2019). Genome editing in plants by engineered CRISPR–Cas9 recognizing NG PAM. *Nat. Plants* 5, 14–17. doi: 10.1038/s41477-018-0321-8

Graf, N., and Altenbuchner, J. (2011). Development of a method for markerless gene deletion in *Pseudomonas* putida. *Appl. Environ. Microbiol.* 77, 5549–5552. doi: 10.1128/AEM.05055-11

Hoang, T. T., Karkhoff-Schweizer, R. R., Kutchma, A. J., and Schweizer, H. P. (1998). A broad-host-range F1p-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. *Gene* 212, 77–86. doi: 10.1016/S0378-1119(98)00130-9

Hwang, G. H., Park, J., Lim, K., Kim, S., Yu, J., Yu, E., et al. (2018). Web-based design and analysis tools for CRISPR base editing. *BMC Bioinformatics* 19:4. doi: 10.1186/s12859-018-2585-4

Jakoèiunas, T., Jensen, M. K., and Keasling, J. D. (2016). CRISPR/Cas9 advances engineering of microbial cell factories. *Metab. Eng.* 34, 44–59. doi: 10.1016/j.ymben.2015.12.003

Kim, Y. B., Komor, A. C., Levy, J. M., Packer, M. S., Zhao, K. T., and Liu, D. R. (2017). Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat. Biotechnol.* 35, 371–376. doi: 10.1038/nbt.3803

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi: 10.1038/nature17946

Li, Q., Seys, F. M., Minton, N. P., Yang, J., Jiang, Y., Jiang, W., et al. (2019). CRISPR–Cas9 D10A nickase-assisted base editing in the solvent producer *Clostridium beijerinckii*. *Biotechnol. Bioeng.* 116, 1475–1483. doi: 10.1002/bit.26949

Liang, R., and Liu, J. (2010). Scarless and sequential gene modification in *Pseudomonas* using PCR product flanked by short homology regions. *BMC Microbiol.* 10:209. doi: 10.1186/1471-2180-10-209

Luo, X., Yang, Y., Ling, W., Zhuang, H., Li, Q., and Shang, G. (2016). *Pseudomonas* putida KT2440 markerless gene deletion using a combination of λ Red recombineering and Cre/loxP site-specific recombination. *FEMS Microbiol. Lett.* 363:fnw014. doi: 10.1093/femsle/fnw014

Martínez-García, E., and de Lorenzo, V. (2011). Engineering multiple genomic deletions in Gram-negative bacteria: Analysis of the multi-resistant antibiotic profile of *Pseudomonas* putida KT2440. *Environ. Microbiol.* 13, 2702–2716. doi: 10.1111/j.1462-2920.2011.02538.x

Moreno, R., and Rojo, F. (2014). Features of pseudomonads growing at low temperatures: another facet of their versatility. *Environ. Microbiol. Rep.* 6, 417–426. doi: 10.1111/1758-2229.12150

Nikel, P. I., Chavarría, M., Danchin, A., and de Lorenzo, V. (2016). From dirt to industrial applications: *Pseudomonas* putida as a Synthetic Biology chassis for hosting harsh biochemical reactions. *Curr. Opin. Chem. Biol.* 34, 20–29. doi: 10.1016/j.cbpa.2016.05.011

Nikel, P. I., Martínez-García, E., and De Lorenzo, V. (2014). Biotechnological domestication of pseudomonads using synthetic biology. *Nat. Rev. Microbiol.* 12, 368–379. doi: 10.1038/nrmicro3253

Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., et al. (2016). Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* 353:aaf8729. doi: 10.1126/science.aaf8729

Paulsen, I. T., Press, C. M., Ravel, J., Kobayashi, D. Y., Myers, G. S. A., Mavrodi, D. V., et al. (2005). Complete genome sequence of the plant commensal *Pseudomonas* fluorescens Pf-5. *Nat. Biotechnol.* 23, 873–878. doi: 10.1038/nbt1110

Poblete-Castro, I., Becker, J., Dohnt, K., dos Santos, V. M., and Wittmann, C. (2012). Industrial biotechnology of *Pseudomonas* putida and related species. *Appl. Microbiol. Biotechnol.* 93, 2279–2290. doi: 10.1007/s00253-012-3928-0

Quénée, L., Lamotte, D., and Polack, B. (2005). Combined sacB-based negative selection and cre-lox antibiotic marker recycling for efficient gene deletion in *Pseudomonas aeruginosa*. *Biotechniques* 38, 63–67. doi: 10.2144/05381ST01

Rees, H. A., and Liu, D. R. (2018). Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* 19, 770–788. doi: 10.1038/s41576-018-0059-1

Silby, M. W., Winstanley, C., Godfrey, S. A. C., Levy, S. B., and Jackson, R. W. (2011). *Pseudomonas* genomes: diverse and adaptable. *FEMS Microbiol. Rev.* 35, 652–680. doi: 10.1111/j.1574-6976.2011.00269.x

Slaymaker, I. M., Gao, L., Zetsche, B., Scott, D. A., Yan, W. X., and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science* 351, 84–88. doi: 10.1126/science.aad5227

Stover, C. K., Pham, X. Q., Erwin, A. L., Mizoguchi, S. D., Warrener, P., Hickey, M. J., et al. (2000). Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* 406, 959–964. doi: 10.1038/35023079

Sun, J., Wang, Q., Jiang, Y., Wen, Z., Yang, L., Wu, J., et al. (2018). Genome editing and transcriptional repression in *Pseudomonas* putida KT2440 via the type II CRISPR system. *Microb. Cell Fact.* 17, 1–17. doi: 10.1186/s12934-018-0887-x

Vodovar, N., Vallenet, D., Cruveiller, S., Rouy, Z., Barbe, V., Acosta, C., et al. (2006). Complete genome sequence of the entomopathogenic and metabolically versatile soil bacterium *Pseudomonas* entomophila. *Nat. Biotechnol.* 24, 673–679. doi: 10.1038/nbt1212

Wang, W., Wang, S., Bilal, M., Zhang, X., Zong, Y., and Hu, H. (2018). Development of a plasmid-free biosynthetic pathway for enhanced muconic acid production in *Pseudomonas* chlororaphis HT66. *ACS Synth. Biol.* 7, 1131–1142. doi: 10.1021/acssynbio.8b00047

Wang, Y., Liu, Y., Liu, J., Guo, Y., Fan, L., Ni, X., et al. (2018). MACBETH: multiplex automated *Corynebacterium glutamicum* base editing method. *Metab. Eng.* 47, 200–210. doi: 10.1016/j.ymben.2018.02.016

Wang, Y., Liu, Y., Li, J., Yang, Y., Ni, X., Cheng, H., et al. (2019). Expanding targeting scope, editing window, and base transition capability of base editing in *Corynebacterium glutamicum*. *Biotechnol. Bioeng.* 116, 3016–3029. doi: 10.1002/bit.27121

Wenzel, S. C., Gross, F., Zhang, Y., Fu, J., Stewart, A. F., and Müller, R. (2005). Heterologous expression of a myxobacterial natural products assembly line in pseudomonads via Red/ET recombineering. *Chem. Biol.* 12, 349–356. doi: 10.1016/j.chembiol.2004.12.012

Wu, Z., Chen, Z., Gao, X., Li, J., and Shang, G. (2019a). Combination of ssDNA recombineering and CRISPR-Cas9 for *Pseudomonas* putida KT2440 genome editing. *Appl. Microbiol. Biotechnol.* 103, 2783–2795. doi: 10.1007/s00253-019-09654-w

Wu, Z., Wang, Y., Zhang, Y., Chen, W., Wang, Y., and Ji, Q. (2019b). Strategies for developing CRISPR-based gene editing methods in bacteria. *Small Methods* 560, 1–21. doi: 10.1002/smtd.201900560

Wynands, B., Lenzen, C., Otto, M., Koch, F., Blank, L. M., and Wierckx, N. (2018). Metabolic engineering of *Pseudomonas* taiwanensis VLB120 with minimal genomic modifications for high-yield phenol production. *Metab. Eng.* 47, 121–133. doi: 10.1016/j.ymben.2018.03.011

Xiao, A., Cheng, Z., Kong, L., Zhu, Z., Lin, S., Gao, G., et al. (2014). CasOT: a genome-wide Cas9/gRNA off-target searching tool. *Bioinformatics* 30, 1180–1182. doi: 10.1093/bioinformatics/btt764

Yang, S., Liu, Q., Zhang, Y., Du, G., Chen, J., and Kang, Z. (2018). Construction and characterization of broad-spectrum promoters for synthetic biology. *ACS Synth. Biol.* 7, 287–291. doi: 10.1021/acssynbio.7b00258

Yu, W. S. W., Chen, W., Song, L., Zhang, L., Fangyou, L. S. (2018). CRISPR-Cas9 and CRISPR-assisted cytidine deaminase enable precise and efficient genome editing in *Klebsiella pneumoniae*. *Appl. Environ. Microbiol.* 84, 1–15.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

MINI REVIEW
published: 14 August 2020
doi: 10.3389/fbioe.2020.00969

Microbial Cell Factory for Efficiently Synthesizing Plant Natural Products via Optimizing the Location and Adaptation of Pathway on Genome Scale

*Bo Yang[1], Xudong Feng[2] and Chun Li[1,2,3]**

[1] SynBio Research Platform, Collaborative Innovation Center of Chemical Science and Engineering, Key Laboratory of Systems Bioengineering, Ministry of Education, School of Chemical Engineering and Technology, Tianjin University, Tianjin, China, [2] Institute for Synthetic Biosystem/Department of Biochemical Engineering, School of Chemistry and Chemical Engineering, Beijing Institute of Technology, Beijing, China, [3] Key Laboratory for Industrial Biocatalysis, Ministry of Education, Department of Chemical Engineering, Tsinghua University, Beijing, China

Plant natural products (PNPs) possess important pharmacological activities and are widely used in cosmetics, health care products, and as food additives. Currently, most PNPs are mainly extracted from cultivated plants, and the yield is limited by the long growth cycle, climate change and complex processing steps, which makes the process unsustainable. However, the complex structure of PNPs significantly reduces the efficiency of chemical synthesis. With the development of metabolic engineering and synthetic biology, heterologous biosynthesis of PNPs in microbial cell factories offers an attractive alternative. Based on the in-depth mining and analysis of genome and transcriptome data, the biosynthetic pathways of a number of natural products have been successfully elucidated, which lays the crucial foundation for heterologous production. However, there are several problems in the microbial synthesis of PNPs, including toxicity of intermediates, low enzyme activity, multiple auxotrophic dependence, and uncontrollable metabolic network. Although various metabolic engineering strategies have been developed to solve these problems, optimizing the location and adaptation of pathways on the whole-genome scale is an important strategy in microorganisms. From this perspective, this review introduces the application of CRISPR/Cas9 in editing PNPs biosynthesis pathways in model microorganisms, the influences of pathway location, and the approaches for optimizing the adaptation between metabolic pathways and chassis hosts for facilitating PNPs biosynthesis.

Keywords: plant natural products, microbial cell factory, pathway location, adaptation, heterologous pathway

OPEN ACCESS

Edited by:
Dawei Zhang,
Tianjing Institute of Industrial Biotechnology (CAS), China

Reviewed by:
Han Xiao,
Shanghai Jiao Tong University, China
Yanfeng Liu,
Jiangnan University, China

***Correspondence:**
Chun Li
lichun@tsinghua.edu.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *17 April 2020*
Accepted: *27 July 2020*
Published: *14 August 2020*

Citation:
Yang B, Feng X and Li C (2020) Microbial Cell Factory for Efficiently Synthesizing Plant Natural Products via Optimizing the Location and Adaptation of Pathway on Genome Scale. Front. Bioeng. Biotechnol. 8:969. doi: 10.3389/fbioe.2020.00969

INTRODUCTION

Plant natural products (PNPs) are secondary metabolites that are mainly used for defense and signal transduction in plants (Marienhagen and Bott, 2012). They have complex structures and various physiological as well as pharmacological activities. For instance, glycyrrhizin, betulinic acid, paclitaxel and resveratrol have antitumor activities, while lycopene, β-carotene, and astaxanthin possess antioxidant properties (Lü et al., 2016; Zhao et al., 2017; Zhang L. et al., 2019; Chen et al., 2020; Sun et al., 2020). Obtaining natural products through extraction from cultivated plants faces the problem of low natural content, uncertain climate

factors and destruction of the ecological environment (Zhao and Li, 2018; Sun et al., 2019). With the rapid advances of synthetic biology, the production of PNPs in microbial cell factories has paved the way for large-scale industrial production by shortening the synthesis cycle and reducing the difficulty of product separation (Liu et al., 2017).

There three main genome editing tools for pathway modification and regulation *in vivo*, including zinc-finger nucleases (ZFNs), transcription activation-like effector nucleases (TALENs), and the clustered regularly interspaced short palindromic repeats (CRISPR) system. In microorganisms, CRISPR can accurately edit and regulate the metabolic pathways of PNPs in an efficient manner without selection markers (Hou et al., 2018). However, the introduction of heterologous pathways usually disrupts the intracellular metabolic balance (Liu et al., 2019), while the integration loci of heterologous expression cassettes affect enzyme expression and product accumulation (Englaender et al., 2017). Thus, appropriately chosen loci can enhance the stable expression of heterologous genes. Additionally, organelles with a high local concentration of the substrates are a suitable location for PNPs synthesis. Numerous studies showed that the balance between endogenous and exogenous pathways contributed to the efficient synthesis of products (Zhang et al., 2015; Park et al., 2018), and the highest expression level of enzymes did not necessarily maximize the yield (Kang et al., 2018). At the same time, the adaptation of the chassis cells to the heterologous pathway is equally important. Traditional metabolic engineering is employed to regulate potential pathways and restore intracellular metabolic balance, which relies on targets obtained by analyzing metabolic flux distribution, reaction mechanisms or metabolic network models. However, a finite number of target modifications cannot produce the full possible variety of genotypes and phenotypes to select the optimal phenotypes. Reprogramming the expression of multiple genes is a complementary strategy, relying on approaches such as multiplex automated genome engineering (MAGE; Wang et al., 2009), or synthetic chromosome rearrangement and modification by loxP-mediated evolution (SCRaMbLE; Standage-Beier and Wang, 2017), which are tools for rewriting genomes. Moreover, heterologous synthesis of PNPs often involves multi-step enzymatic reactions, and synthetic scaffolds can improve the catalytic activity of the system to some extent (Zhang, 2011). To address the demand for multiple auxotrophic markers, the controllable decentralized assembly strategy was established and recyclable markers were employed for iterative integration (Xie et al., 2014). This review covers recent studies on pathway location and strategies for optimizing the fitness and intracellular metabolism of the chassis cells to promote the production of PNPs.

APPLICATION OF CRISPR/CAS9 IN PATHWAY INTEGRATION AND REGULATION

In 2013, the gene editing function of the CRISPR/Cas9 system was first verified in mammals (Cong et al., 2013). Soon afterward, CRISPR/Cas9 was universally applied in microorganisms. The CRISPR-mediated multi-locus gene integration strategy was developed to efficiently edit the β-carotene synthesis pathway in *Saccharomyces cerevisiae* (Ronda et al., 2015). Another significant research was that GTR-CPISPR could edit eight genes simultaneously with an efficiency up to 84% (Zhang Y. et al., 2019). The Cas9-based toolkit with high efficiency for integrating cassettes increased taxadiene production by 25-fold (Apel et al., 2017). In addition to genome editing functions, some derivatives have been exploited for precise and rapid regulation of gene expression. CRISPR interference (CRISPRi) can be used to regulate the transcription of multiple sites in the genome by fusing activators or repressors with catalytically-inactive Cas9 (dCas9) protein (Gilbert et al., 2013). Harnessing CRISPRi to synchronously down-regulate seven genes in the competing pathway boosted β-amyrin biosynthesis (Ni et al., 2019). The production of other PNPs such as pinosylvin (Wu et al., 2017a), resveratrol (Wu et al., 2017b), O-methylated anthocyanin (Cress et al., 2017), and α-amyrin (Yu et al., 2018) in microbes was also successfully enhanced using CRISPRi. Zalatan et al. (2015) extended sgRNA with modular RNA domains to form the scaffold RNA (scRNA), which can not only recognize target sequences specifically but also recruit RNA-binding proteins fused with effectors. The expression of multiple scRNAs made up for the defect of CRISPRi and achieved diverse types of regulation at respective loci in parallel. Employing the scRNA system, the best gRNA screened from 101 candidates was chosen to regulate multiple genes in the carotenoid biosynthesis pathway (Jensen et al., 2017). It is worth noting that disparate gRNAs competing for the same Cas9 is a likely limiting factor in scRNA systems. CRISPR-AID, an orthogonal tri-functional system, enabled the concurrent upregulation, downregulation and deletion of genes in *S. cerevisiae*, achieving the modular regulation of metabolic networks (Lian et al., 2017). Successful examples of the application of CRISPR/Cas9 and its derivatives in PNPs biosynthesis are summarized in **Table 1**.

OPTIMIZATION OF THE PATHWAY LOCATION

Microbial cell factories usually have to efficiently express multiple heterologous genes, and the introduction of episomal plasmids can impose a heavy metabolic burden. At the same time, it was found that the genetic stability of the host decreases when multiple plasmids are used in the same cell (Yan et al., 2005). By contrast, chromosomal integration of target genes is more stable and does not require a selective pressure (Zhu et al., 2018). For example, integrating target genes into the native GAL1 locus of engineered *S. cerevisiae* increased the oleanolic acid titer by 3.6 times compared with the strain utilizing multiple plasmids (Zhao et al., 2018). Optimizing the chromosomal integration sites of metabolic pathways and compartmentalizing metabolic pathways in organelles are both viable strategies for improving the synthesis of PNPs.

TABLE 1 | The application of CRISPR/Cas9 and derived technologies in the biosynthesis of PNPs.

Species	Products	Method	Regulation strategy	Culture method	Titer	References
E. coli	β-carotene	CRISPR/Cas9	Genomic modification	Fed-batch (5 L)	2.0 g/L	Li et al., 2015
S. cerevisiae	Mevalonate	CRISPR/Cas9	Multi-gene disruption	–	1.5 mg/L (41-fold)	Jakounas et al., 2015
Yarrowia lipolytica	Lycopene	CRISPR/Cas9	Multi-gene integration	Shake flask (250 mL)	8.6-fold	Schwartz et al., 2017
S. cerevisiae	Valencene	CRISPR/Cas9	Knock-out or down-regulation	Fed-batch (3 L)	539.3 mg/L (160-fold)	Chen et al., 2019
S. cerevisiae	guaia-6,10 (14)-diene	CRISPR/Cas9	Gene integration	Fed-batch (5 L)	0.8 g/L	Siemon et al., 2020
Candida tropicalis	β-carotene	CRISPR/Cas9	Multi-gene deletion or mutation	–	0.23 mg/g DCW	Zhang L. et al., 2020
Yarrowia lipolytica	β-carotene	CRISPR/Cas9	Gene integration	Fed-batch (5 L)	4.5 g/L	Zhang X.K. et al., 2020
S. cerevisiae	β-carotene	CrEdit	Multi-loci gene integration	–	12.7 mg/L	Ronda et al., 2015
S. cerevisiae	Patchoulol	Cas-3P	Multiplexed and sequential editing	Shake flask (250 mL)	20 mg/L	Li et al., 2020
S. cerevisiae	Taxadiene	Cas9-based toolkit	Including Cas9-sgRNA plasmids, promoters, protein tags	Test tube (5 mL)	20 mg/L (25-fold)	Apel et al., 2017
E. coli	Naringenin	CRISPRi	Multi-gene repression	Shake flask	421.6 mg/L (7.4-fold)	Wu et al., 2015
E. coli	Lycopene (-)-α-bisabolol	CRISPRi	Regulatable repression	Shake flask	50.6 mg/L 8.8-fold	Kim et al., 2016
E. coli	Peonidin 3-O-glucoside	CRISPRi	Silence expression	Shake flask (125 mL)	56 mg/L (2-fold)	Cress et al., 2017
E. coli	Pinosylvin	CRISPRi	Multi-gene repression	Shake flask (500 mL)	281 mg/L	Wu et al., 2017a
E. coli	Resveratrol	CRISPRi	Multi-gene down-regulation	Shake flask	304.5 mg/L	Wu et al., 2017b
S. cerevisiae	α-amyrin	CRISPRi	Down-regulation	–	11.97 mg/L	Yu et al., 2018
S. cerevisiae	β-amyrin	CRISPRi	Multi-gene one-step down-regulation	Fed-batch (2.5 L)	156.7 mg/L(1.44-fold)	Ni et al., 2019
S. cerevisiae	β-carotene	CRISPR-AID	Synchronous up/down-regulation, deletion	–	3-fold	Lian et al., 2017
Ogataea polymorpha	Resveratrol	CMGE	Multi-gene knock-out, integration	Shake flask (500 mL)	97.23 mg/L (20.73-fold)	Wang L. et al., 2018

Chromosomal Integration Loci of Metabolic Pathways

The integration loci of heterologous genes in the genome influence enzyme expression levels in various hosts, such as *Escherichia coli* (Englaender et al., 2017), yeast (Guo et al., 2018), actinomycetes (Bilyk et al., 2017), and *Lactococcus lactis* (Thompson and Gasson, 2001), and may even affect the titer of products (Bilyk et al., 2017). Enzymes with easily detectable activities or fluorescent proteins are normally chosen to characterize the expression intensity of sites, providing important guidance for the construction of efficient microbial cell factories.

In one study, a total of 1044 loci within the whole genome of *S. cerevisiae* were analyzed and the largest difference in expression levels among them was 13 times (Wu X.-L. et al., 2017). Sites with low expression intensity were mainly located near telomeres and centromeres (Ottaviani et al., 2008), and the loci with high expression levels were adjacent to autonomously replicating sequences. The robustness of position effect was even stronger with different promoters (Bai Flagfeldt et al., 2009), reporter genes, and carbon sources (Wu X.-L. et al., 2017). After taking into account the impact of chromosomal location on gene expression and growth rate, 11 out of 14 genomic loci in *S. cerevisiae* were found to be suitable for pathway integration (Mikkelsen et al., 2012). The introduction of eight genes at four sites enabled high indolylglucosinolate production. In *Yarrowia lipolytica*, some loci that could enhance the stable expression of the β-carotene (Zhang X.K. et al., 2020) or lycopene (Schwartz et al., 2017) biosynthesis pathways were screened out. Similar to plasmid expression, the copy number of the integration site also has a considerable effect on product synthesis (Wang L. et al., 2018). A fused expression cassette containing three genes for resveratrol biosynthesis was inserted into the multi-copy rDNA cluster in *Ogataea polymorpha* and strains with copy numbers ranging from 1 to 10 were obtained. Within this range, the increase of copy numbers of the expression cassette increased the resveratrol titer 20 times in comparison to integrating the three genes into scattered single-copy loci, respectively.

However, the protein expression level of the biosynthetic enzymes was not always positively correlated with the yield of PNPs (Englaender et al., 2017). Integration loci in the genome affect the efficiency of pathways, and inappropriate selection of

sites may even hinder product synthesis. The effects of integration sites on strain growth, genetic stability, and product synthesis should be taken into account comprehensively.

Subcellular Compartmentalization of Metabolic Pathways

Subcellular organelles, such as mitochondria, peroxisomes, and endoplasmic reticulum, have complex structures and act as independent membrane-bound compartments, which leads to higher local concentrations of substrates and generates physical separation between products and competing pathways (Huttanus and Feng, 2017). Compartmentalizing metabolic pathways in organelles achieved by fusing proteins with targeting signal tags can enhance the reaction rate and product synthesis efficiency.

Farnesyl diphosphate, which is an important intermediate in carotenoid production, is abundant in peroxisomes (Kovacs et al., 2002). When heterologous enzymes of the lycopene synthesis pathway were targeted into peroxisomes, it remarkably boosted the titer of lycopene to 73.9 mg/L in *Komagataella phaffii* (Bhataya et al., 2009). Another organelle commonly used in compartmentalization studies is mitochondria. Employing mitochondrion-targeted enzymes resulted in 8 and 20-fold increase in valencene and amorphadiene production, respectively (Farhi et al., 2011). The concentration of acetyl-CoA in mitochondria is 20–30 times higher than that in the cytoplasm (Galdieri et al., 2014). Many studies focused on the mitochondrial acetyl-CoA pool to promote the production of PNPs such as amorpha-4,11-diene (Yuan and Ching, 2016), geraniol, 8-hydroxygeraniol, and nepetalactol (Yee et al., 2019). Recently, dual engineering of metabolic pathways in the cytoplasm and organelles was performed for high-level PNPs production. Assembling the complete MVA pathway in peroxisomes or mitochondria, together with the cytoplasmic pathway, boosted the output of α-humulene (Zhang C et al., 2020), and linalool (Zhang Y et al., 2020), respectively.

ADAPTATION OF CHASSIS CELLS TO HETEROLOGOUS PATHWAYS

The introduction of a heterologous pathway almost always negatively affects metabolic homeostasis. It is therefore necessary to improve the fitness on a genome-wide level for efficient PNPs production. Here, we discuss the four strategies MAGE, SCRaMbLE, synthetic scaffolds, and decentralized assembly (**Figure 1**), which can be used to adapt the chassis to the introduced pathway.

MAGE

In the phage λ-Red recombination system, single-stranded DNA (ssDNA) binding protein β promotes homologous complementarity between ssDNA and the lagging strand of the replication fork during DNA replication. Consequently, sequences are integrated into the genome of offspring cells, leading to allelic replacement (Ellis et al., 2001). Based on ssDNA recombination, MAGE was developed to precisely and simultaneously program multiple specific sites (insertions, mismatches, or deletions) on a genome-wide scale (Wang et al., 2009). A total of 24 endogenous genes in the metabolic pathway were modified using a synthetic oligos library, which optimized the production of lycopene in *E. coli* (Wang et al., 2009). Among them, degenerate RBS sequences that differentially regulate gene expression and nonsense mutations were inserted for inactivation. After 3 days of evolution, mutants with a 5-fold increase in yield were obtained.

However, the inefficiency of inserting large fragments by MAGE limited the modification of genetic elements. Applying switchable co-selection markers, multiple T7 promoters were introduced into 12 genomic loci simultaneously, and a combination that maximized indigo synthesis was obtained (Wang et al., 2012). Large combinatorial populations were generated in each round of MAGE, including unproductive cheaters. To ease the screening burden, toggled selection was devised to remove cheaters (Raman et al., 2014). Oligos were targeted to SD sequences of candidate genes and toggled selection after each round of evolution resulted in naringenin production of up to 61 mg/L. In some cases, the biosynthesis of PNPs depended on NADPH, and insufficient NADPH restricted PNPs production. An RBS mutant library of four genes in the Entner-Doudoroff pathway was constructed and 40 cycles of MAGE generated mutants with a high NADPH pool, which led to a 97% increase in the titer of neurosporene (Ng et al., 2015). A common feature of these studies was the use of *mutS*-deficient *E. coli*, which could improve recombination efficiency. Strain-independent MAGE was developed to reduce the accumulation of undesired mutations and to broaden the range of applicable hosts (Ryu et al., 2014). Ryu et al. integrated a suicide plasmid carrying the λ Red recombination system into the *mutS* locus to switch MutS activity in *E. coli*. By constructing a mutant library of 5′-untranslational region of genes, MAGE fine-tuned the expression level of enzymes and resulted in a 38.2-fold improvement of the curcumin titer (Kang et al., 2018).

Due to the complicated genetic regulatory system in eukaryotes, the effect of employing MAGE directly is not ideal (Si et al., 2017). Researchers developed an oligo-mediated recombination method in *S. cerevisiae*, referred to as yeast oligo-mediated genome engineering (YOGE; DiCarlo et al., 2013). After inactivating the mismatch-repair system, overexpressing DNA recombinase and optimizing the reaction conditions, the editing efficiency was increased to 1%, but it was still low. With the in-depth investigation of the mechanism, eukaryotic MAGE (eMAGE) was established for multiplex genome engineering (Barbieri et al., 2017). The efficiency of precisely editing a single nucleotide exceeded 40%, and no additional mutations were introduced at the target sites. Apart from slowing DNA replication, the complementarity of ssDNA oligodeoxynucleotides (ssODNs) with the lagging strand rather than leading strand was equally conducive to allelic replacement. Incorporation of ssODNs at the replication fork formed combinatorial genomic diversity and facilitated β-carotene production in *S. cerevisiae* (Barbieri et al., 2017).

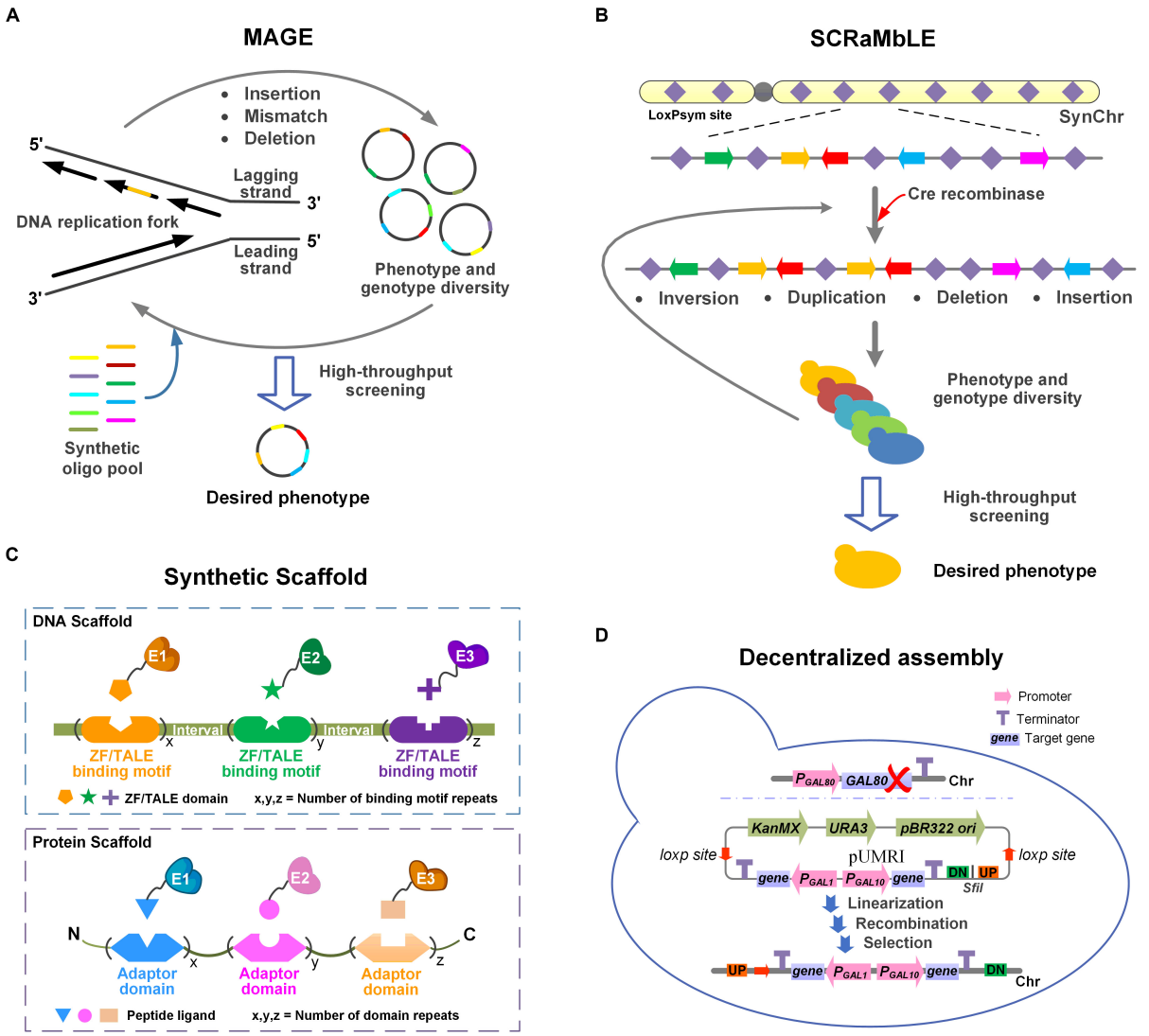


FIGURE 1 | Schematic illustration of four strategies for adapting the chassis cells to the heterologous pathway. **(A)** MAGE. Insertions, mismatches or deletions can be introduced at multiple specific sites in the genome by MAGE. Multiple iterations result in phenotypic and genotypic diversity. **(B)** SCRaMbLE. Genomic rearrangements (inversions, duplications, deletions, and insertions) occur at any two loxPsym sites stochastically. Phenotypic and genotypic diversity was obtained through iterative cycling. **(C)** Synthetic scaffold. In DNA scaffold, the target enzymes are fused with ZF/TALE domains through flexible linkers. In protein scaffold, the target enzymes are fused with peptide ligands via flexible linkers. **(D)** Decentralized assembly. GAL80 was deleted to obtain the GAL regulatory system. pUMRI, a marker-recyclable integrative toolbox with double selection was used to integrated genes into scattered loci. The controllable system and pUMRI constitute the decentralized assembly strategy.

To address the tediousness of oligonucleotide design, two web-based tools, MODEST (Bonde et al., 2014), and Merlin (Quintin et al., 2016), were developed. Iterative incorporation of ssDNA or ssODNs promoted sequence diversity. MAGE and derived technologies enable the introduction of genomic diversity, making it easier to evolve the desired phenotype compared with conventional methods and forming a large mutant library. This technology has been successfully applied to prokaryotes and eukaryotes to balance biosynthesis and cell growth, and improve the adaptation of hosts to heterologous pathways. However, the prerequisite for utilizing MAGE is to identify target genes that affect product synthesis and to possess suitable high-throughput screening methods, which has limited its wider application.

SCRaMbLE

In the international Synthetic Yeast Genome Project (Sc2.0), bidirectional loxPsym sites were inserted downstream of each non-essential gene (Kannan and Gibson, 2017). With Cre recombinase induction, rearrangements (deletion, insertion, duplication, and inversion events) occurred at any two loxPsym sites stochastically, resulting in the SCRaMbLE technique (Jin et al., 2018). The artificial rearrangement of synthetic chromosomes on the genome-wide scale accelerates strain

evolution to acquire the expected phenotypes (Wang J. et al., 2018). The genotypes were further analyzed by PCRTag and whole genome sequencing to identify the structural variations, providing an important reference for deciphering mechanism (Ma et al., 2019).

When the synthetic yeast strains containing a Cre recombinase plasmid were cultured, even without adding estradiol, a fraction of strains would switch on SCRaMbLE and their growth was slightly affected (Annaluru et al., 2014), indicating leaky expression of the plasmid. To resolve this, a genetic AND gate switch was devised to precisely control SCRaMbLE (Jia et al., 2018). A galactose-inducible promoter was used to regulate the expression of a fusion protein composed of Cre and an estrogen-binding domain. Therefore, only when galactose and estradiol were present simultaneously, genomic rearrangement can be triggered. Adopting an AND gate switch in synV haploid yeast raised the production of carotenoids by 50%. Heterozygous diploid yeasts with synthetic chromosomes and wild type counterparts were more likely to generate larger structural variations than haploid (Li et al., 2019). The synIII&V diploids formed by mating increased the likelihood of genome diversification and exhibited a 6.29- to 7.81-fold enhancement of the carotenoid yield (Jia et al., 2018). A large number of beneficial rearrangements was accumulated through multiplex SCRaMbLE iterative cycling, promoting the carotenoid synthesis up to 38.8 times.

In addition to the issue of leaky expression from plasmids, it is necessary to establish high-throughput screening methods to broaden the application of SCRaMbLE. In order for rapidly screening mutants after SCRaMbLE rearrangement, an ultra-fast LC-MS method using a guard column to substitute a standard analytical column was employed to boost betulinic acid production in *S. cerevisiae*, reducing the detection time per sample from 5 min to 84 s (Gowers et al., 2020). Ultimately, multiplex nanopore sequencing was utilized to identify the rearrangements in the high-yield strains and establish genotype-phenotype correlation.

To settle the problem that the exogenous synthesis pathway and the chassis cannot be optimized synchronously (Keasling, 2012), SCRaMbLE-in composed of an *in vitro* recombinase kit and *in vivo* chromosome rearrangement system was constructed to promote β-carotene biosynthesis (Liu et al., 2018). The recombinase toolkit inserted regulatory elements upstream of candidate genes to achieve different expression intensities. Apart from Cre, the recombinases Dre and VCre were used *in vitro* to construct the library. The *in vivo* rearrangement system induced by Cre randomly integrated the metabolic pathway into the engineered genome, leading to massive rearrangements of chromosomes. The β-carotene yield was increased by 2-fold through genomic rearrangement in comparison to integrating the pathway into the HO locus. The bottleneck of product biosynthesis can be solved through SCRaMbLE. In the process of astaxanthin production, the combination of *crtZ* and *crtW* from different sources was optimized stochastically. Among the darker-red colonies generated by SCRaMbLE, the highest yield of astaxanthin was boosted to 8.51 times (Qi et al., 2020).

Instead of using the inducer estradiol, which was toxic to humans, light-controlled SCRaMbLE (L-SCRaMbLE) was developed (Hochrein et al., 2018). The N-terminal and C-terminal of split Cre fused with chromophore-binding photoreceptor phytochrome B (PhyB) and phytochrome interacting factor PIF3, respectively, to constitute L-SCRaMbLE. Upon red light illumination, split Cre recombined and induced random rearrangements of genes among loxPsym sites. Compared with estradiol-inducible system, L-SCRaMbLE had lower recombination efficiency and mediated the recombination in a phycocyanobilin- and light-dependent manner.

Due to insufficient understanding of stress response mechanisms in microorganisms, some genes related to product synthesis were omitted during modification, which may sometimes be remedied via SCRaMbLE. This black-box approach provides a platform for rapid generation of phenotypic and genotypic diversity. Multiple rounds of SCRaMbLE were found to improve the production further (Liu et al., 2018). Meanwhile, an enormous mutant library was formed, and high-throughput screening method was a prerequisite for its wide application. Analyzing the differences in the expression of genes adjacent to rearrangement loci and the perturbation of the global metabolic network by the rearrangements is conductive to identifying possible regulatory mechanisms, which provides important reference for further rational modification to maximize PNPs synthesis in microbes.

Synthetic Scaffold

Some enzymes of heterologous pathways are located in different positions intracellularly. For multienzyme reaction systems, the large distance among different enzymes or between enzymes and substrates would restrict product synthesis (Morgado et al., 2018). The substrate channel formed by enzymes assembled on a scaffold shortens the distance between the active sites of different enzymes and alleviates the inhibitory effects of toxic intermediates (Horn and Sticht, 2015). DNA and protein scaffold have been adopted to facilitate PNPs synthesis in cell factories. The expression levels and unit numbers of scaffolds, the stoichiometric ratio and spatial orientation of enzymes, and the interval among enzymes are important parameters for scaffold optimization. To prevent the close proximity of multiple enzyme complexes from blocking access to the substrate binding site, flexible linkers are used in the scaffold. RNA scaffold has also been investigated, and was devised as discrete, one-dimensional, two-dimensional or triangular structure to promote the output of hydrogen (Delebecque et al., 2011), pentadecane and succinate (Sachdeva et al., 2014). Nonetheless, RNA scaffold is rarely used in the biosynthesis of PNPs. A brief comparison of three synthetic scaffolds was made in **Supplementary Table S1**.

DNA Scaffold

Zinc fingers (ZFs) can bind to specific DNA sequences and form a multienzyme complex with DNA scaffold (Greisman and Pabo, 1997). Fusing two key enzymes with Zif268 and PBSII ZF domains led to a nearly 5-fold increase in resveratrol production (Conrado et al., 2012). Similarly, the assembly of three enzymes in the lycopene synthesis pathway increased the production

up to 4.7 times in *E. coli* (Xu et al., 2020). Apart from ZFs, transcription activator-like effectors (TALEs) are also commonly used to construct DNA scaffold (Moscou and Bogdanove, 2009). The system composed of TALEs and corresponding DNA scaffold considerably increased the titer of indole-3-acetic acid by closely linking the two enzymes in the pathway (Zhu et al., 2016). The plasmid copy number of DNA scaffold was also found to affect the co-localization of heterologous enzymes (Xie et al., 2019).

Protein Scaffold

Adaptor domains SH3, SH2, PDZ, and GTPase binding domain (GBD) have strong affinity with peptide ligands and are widely used in protein scaffold (Horn and Sticht, 2015). In *E. coli*, the production of mevalonate was enhanced 77-fold by anchoring key enzymes fused with peptide ligands to the corresponding aptamer domain (Dueber et al., 2009). Based on this study, Zhao et al. (2015) adopted nine synthetic protein scaffolds from Dueber's research to capture and assemble enzymes for catechin biosynthesis in *E. coli*. They found that genes from diverse sources could also cause significant differences in scaffold function. With similar protein domains, the optimal scaffold ($GBD_1SH3_2PDZ_4$) increased the yield of resveratrol 2.7 times, compared with using the fusion protein in *S. cerevisiae* (Wang and Yu, 2012). Recently, SpyCatcher/SpyTag, and SnoopCatcher/SnoopTag were selected as protein tags to covalently conjugate enzymes in the mevalonate pathway, enhancing the titers of lycopene and astaxanthin to varying degrees (Qu et al., 2019).

Decentralized Assembly

Constructing controllable and genetically stable heterologous multi-gene metabolic pathways is an effective strategy to promote product biosynthesis. In order to avoid homologous recombination among multiple repeated sequences at the same location (Blount et al., 2012), genes were integrated into different loci on the chromosome for stable expression. To achieve this, a decentralized assembly strategy comprising integrative plasmids (pMRI) with recyclable markers and GAL regulatory system was developed (Xie et al., 2014). The pMRI plasmids contain ready to use *loxP-KanMX-pBR322ori-loxP* as the selection marker, which could be recombined and removed under Cre induction to realize the recyclable utilization of the marker. All candidate genes involved in β-carotene biosynthesis were controlled by *GAL1-GAL10* bidirectional promoters. The knockout of GAL80 enabled high-glucose inhibition and low-glucose induction, so that the glucose concentration could regulate the switching time of the integrated pathway. The dynamic regulation of the assembled pathway balanced the metabolic flux between regular cellular activities and product accumulation, shifting more intracellular resources to β-carotene synthesis.

To simplify marker excision and shorten the integration period, pUMRI with the marker *loxp-kanMX-URA-loxp* was designed to substitute pMRI (Lv et al., 2016). The marker was removed through low-frequency mitotic recombination and colony counterselection. The improved decentralized assembly strategy was applied to the synthesis of isoprenoids (Ye et al., 2017) and lycopene (Zhou et al., 2018).

CONCLUSION AND PERSPECTIVE

The location of metabolic pathways and the adaptation of the chassis cells to the heterologous pathways have a significant influence on PNPs biosynthesis. However, there are relatively few studies on the effects of different integration loci on PNPs production, which should be explored in the future. When targeting metabolic pathways to organelles, the possible negative effects of targeting signal tags on enzyme activity need to be considered. Although the genomic diversity generated by whole-genome engineering methods such as MAGE or SCRaMbLE can greatly accelerate strain evolution, high-throughput screening methods are required to identify beneficial mutants in the enormous mutant library, which is a pivotal step during laboratory evolution. There are few high-throughput screening methods. In most cases, these methods are only applicable to the specific reactions. The lack of efficient screening methods is a universal bottleneck in current research. For genome editing, an automated high-throughput screening platform is an important direction for future research. In addition to cytoplasmic protein scaffold, emerging membrane-bound scaffold is expected to improve PNPs biosynthesis. Examples include cohesion-based protein scaffold located on the membranes of lipid droplets (Lin et al., 2017), Tat-assisted scaffold in the thylakoid membrane (Henriques de Jesus et al., 2017), and membrane steroid-binding protein-mediated scaffold anchored to the ER membrane (Gou et al., 2018). To avert chromosomal rearrangement between repeated loxP segments, seamless recombination can be performed to optimize the decentralized assembly strategy. As synthetic biology tools and strategies move forward, these challenges will be addressed gradually. Microbial cell factories have the potential to achieve efficient synthesis and large-scale industrial production of PNPs.

AUTHOR CONTRIBUTIONS

All authors contributed to conception and design of the study. BY participated in searching and analyzing literature for this review and wrote the manuscript. XF and CL edited and corrected the manuscript. All authors approved the submitted version.

FUNDING

This work was supported by the National Key Research and Development Program of China (2018YFA0901800 and 2019YFA0905700) and the National Natural Science Foundation of China (21736002 and 21878021).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00969/full#supplementary-material

REFERENCES

Annaluru, N., Muller, H., Mitchell, L. A., Ramalingam, S., Stracquadanio, G., Richardson, S. M., et al. (2014). Total synthesis of a functional designer eukaryotic chromosome. *Science* 344, 55–58. doi: 10.1126/science.1249252

Apel, A. R., d'Espaux, L., Wehrs, M., Sachs, D., Li, R. A., Tong, G. J., et al. (2017). A Cas9-based toolkit to program gene expression in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 45, 496–508. doi: 10.1093/nar/gkw1023

Bai Flagfeldt, D., Siewers, V., Huang, L., and Nielsen, J. (2009). Characterization of chromosomal integration sites for heterologous gene expression in *Saccharomyces cerevisiae*. *Yeast* 26, 545–551. doi: 10.1002/yea.1705

Barbieri, E. M., Muir, P., Akhuetie-Oni, B. O., Yellman, C. M., and Isaacs, F. J. (2017). Precise editing at DNA replication forks enables multiplex genome engineering in Eukaryotes. *Cell* 171, 1453–1467. doi: 10.1016/j.cell.2017.10.034

Bhataya, A., Schmidt-Dannert, C., and Lee, P. C. (2009). Metabolic engineering of *Pichia pastoris* X-33 for lycopene production. *Process Biochem.* 44, 1095–1102. doi: 10.1016/j.procbio.2009.05.012

Bilyk, B., Horbal, L., and Luzhetskyy, A. (2017). Chromosomal position effect influences the heterologous expression of genes and biosynthetic gene clusters in *Streptomyces albus* J1074. *Microb. Cell Fact.* 16:5. doi: 10.1186/s12934-016-0619-z

Blount, B. A., Weenink, T., and Ellis, T. (2012). Construction of synthetic regulatory networks in yeast. *FEBS Lett.* 586, 2112–2121. doi: 10.1016/j.febslet.2012.01.053

Bonde, M. T., Klausen, M. S., Anderson, M. V., Wallin, A. I. N., Wang, H. H., and Sommer, M. O. A. (2014). MODEST: a web-based design tool for oligonucleotide-mediated genome engineering and recombineering. *Nucleic Acids Res.* 42, W408–W415. doi: 10.1093/nar/gku428

Chen, H., Zhu, C., Zhu, M., Xiong, J., Ma, H., Zhuo, M., et al. (2019). High production of valencene in *Saccharomyces cerevisiae* through metabolic engineering. *Microb. Cell Fact.* 18:195. doi: 10.1186/s12934-019-1246-2

Chen, R., Yang, S., Zhang, L., and Zhou, Y. J. (2020). Advanced strategies for production of natural products in yeast. *iScience* 23:100879. doi: 10.1016/j.isci.2020.100879

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823. doi: 10.1126/science.1231143

Conrado, R. J., Wu, G. C., Boock, J. T., Xu, H., Chen, S. Y., Lebar, T., et al. (2012). DNA-guided assembly of biosynthetic pathways promotes improved catalytic efficiency. *Nucleic Acids Res.* 40, 1879–1889. doi: 10.1093/nar/gkr888

Cress, B. F., Leitz, Q. D., Kim, D. C., Amore, T. D., Suzuki, J. Y., Linhardt, R. J., et al. (2017). CRISPRi-mediated metabolic engineering of *E. coli* for O-methylated anthocyanin production. *Microb. Cell Fact.* 16:10. doi: 10.1186/s12934-016-0623-3

Delebecque, C. J., Lindner, A. B., Silver, P. A., and Aldaye, F. A. (2011). Organization of intracellular reactions with rationally designed RNA assemblies. *Science* 333, 470–474. doi: 10.1126/science.1206938

DiCarlo, J. E., Conley, A. J., Penttilä, M., Jäntti, J., Wang, H. H., and Church, G. M. (2013). Yeast Oligo-Mediated Genome Engineering (YOGE). *ACS Synth. Biol.* 2, 741–749. doi: 10.1021/sb400117c

Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., Ullal, A. V., et al. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* 27, 753–759. doi: 10.1038/nbt.1557

Ellis, H. M., Yu, D. G., DiTizio, T., and Court, D. L. (2001). High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* 98, 6742–6746. doi: 10.1073/pnas.121164898

Englaender, J. A., Jones, J. A., Cress, B. F., Kuhlman, T. E., Linhardt, R. J., and Koffas, M. A. G. (2017). Effect of genomic integration location on heterologous protein expression and metabolic engineering in *E. coli*. *ACS Synth. Biol.* 6, 710–720. doi: 10.1021/acssynbio.6b00350

Farhi, M., Marhevka, E., Masci, T., Marcos, E., Eyal, Y., Ovadis, M., et al. (2011). Harnessing yeast subcellular compartments for the production of plant terpenoids. *Metab. Eng.* 13, 474–481. doi: 10.1016/j.ymben.2011.05.001

Galdieri, L., Zhang, T., Rogerson, D., Lleshi, R., and Vancura, A. (2014). Protein acetylation and acetyl coenzyme a metabolism in budding yeast. *Eukaryot. Cell* 13, 1472–1483. doi: 10.1128/ec.00189-14

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in Eukaryotes. *Cell* 154, 442–451. doi: 10.1016/j.cell.2013.06.044

Gou, M., Ran, X., Martin, D. W., and Liu, C.-J. (2018). The scaffold proteins of lignin biosynthetic cytochrome P450 enzymes. *Nat. Plants* 4, 299–310. doi: 10.1038/s41477-018-0142-9

Gowers, G. O. F., Chee, S. M., Bell, D., Suckling, L., Kern, M., Tew, D., et al. (2020). Improved betulinic acid biosynthesis using synthetic yeast chromosome recombination and semi-automated rapid LC-MS screening. *Nat. Commun.* 11:868. doi: 10.1038/s41467-020-14708-z

Greisman, H. A., and Pabo, C. O. (1997). A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. *Science* 275, 657–661. doi: 10.1126/science.275.5300.657

Guo, X.-J., Xiao, W.-H., Wang, Y., Yao, M.-D., Zeng, B.-X., Liu, H., et al. (2018). Metabolic engineering of *Saccharomyces cerevisiae* for 7-dehydrocholesterol overproduction. *Biotechnol. Biofuels* 11:192. doi: 10.1186/s13068-018-1194-9

Henriques de Jesus, M. P. R., Zygadlo Nielsen, A., Busck Mellor, S., Matthes, A., Burow, M., Robinson, C., et al. (2017). Tat proteins as novel thylakoid membrane anchors organize a biosynthetic pathway in chloroplasts and increase product yield 5-fold. *Metab. Eng.* 44, 108–116. doi: 10.1016/j.ymben.2017.09.014

Hochrein, L., Mitchell, L. A., Schulz, K., Messerschmidt, K., and Mueller-Roeber, B. (2018). L-SCRaMbLE as a tool for light-controlled Cre-mediated recombination in yeast. *Nat. Commun.* 9:1931. doi: 10.1038/s41467-017-02208-6

Horn, A. H. C., and Sticht, H. (2015). Synthetic protein scaffolds based on peptide motifs and cognate adaptor domains for improving metabolic productivity. *Front. Bioeng. Biotechnol.* 3:191. doi: 10.3389/fbioe.2015.00191

Hou, S., Qin, Q., and Dai, J. (2018). Wicket: a versatile tool for the integration and optimization of exogenous pathways in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* 7, 782–788. doi: 10.1021/acssynbio.7b00391

Huttanus, H. M., and Feng, X. (2017). Compartmentalized metabolic engineering for biochemical and biofuel production. *Biotechnol. J.* 12:1700052. doi: 10.1002/biot.201700052

Jakounas, T., Sonde, I., Herrgard, M., Harrison, S. J., Kristensen, M., Pedersen, L. E., et al. (2015). Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. *Metab. Eng.* 28, 213–222. doi: 10.1016/j.ymben.2015.01.008

Jensen, E. D., Ferreira, R., Jakociunas, T., Arsovska, D., Zhang, J., Ding, L., et al. (2017). Transcriptional reprogramming in yeast using dCas9 and combinatorial gRNA strategies. *Microb. Cell Fact.* 16:46. doi: 10.1186/s12934-017-0664-2

Jia, B., Wu, Y., Li, B. Z., Mitchell, L. A., Liu, H., Pan, S., et al. (2018). Precise control of SCRaMbLE in synthetic haploid and diploid yeast. *Nat. Commun.* 9:1933. doi: 10.1038/s41467-018-03084-4

Jin, J., Ma, Y., and Liu, D. (2018). SCRaMbLE drive application of synthetic yeast genome. *Front. Chem. Sci. Eng.* 12:832–834. doi: 10.1007/s11705-018-1749-0

Kang, S.-Y., Heo, K. T., and Hong, Y.-S. (2018). Optimization of artificial curcumin biosynthesis in *E. coli* by randomized 5'-UTR sequences to control the multienzyme pathway. *ACS Synth. Biol.* 7, 2054–2062. doi: 10.1021/acssynbio.8b00198

Kannan, K., and Gibson, D. G. (2017). Yeast genome, by design. *Science* 355, 1024–1025. doi: 10.1126/science.aam9739

Keasling, J. D. (2012). Synthetic biology and the development of tools for metabolic engineering. *Metab. Eng* 14, 189–195. doi: 10.1016/j.ymben.2012.01.004

Kim, S. K., Han, G. H., Seong, W., Kim, H., Kim, S. W., Lee, D. H., et al. (2016). CRISPR interference-guided balancing of a biosynthetic mevalonate pathway increases terpenoid production. *Metab. Eng.* 38, 228–240. doi: 10.1016/j.ymben.2016.08.006

Kovacs, W. J., Olivier, L. M., and Krisans, S. K. (2002). Central role of peroxisomes in isoprenoid biosynthesis. *Prog. Lipid Res.* 41, 369–391. doi: 10.1016/S0163-7827(02)00002-4

Li, Y., Lin, Z., Huang, C., Zhang, Y., Wang, Z., Tang, Y. J., et al. (2015). Metabolic engineering of *Escherichia coli* using CRISPR-Cas9 meditated genome editing. *Metab. Eng.* 31, 13–21. doi: 10.1016/j.ymben.2015.06.006

Li, Y., Wu, Y., Ma, L., Guo, Z., Xiao, W., and Yuan, Y. (2019). Loss of heterozygosity by SCRaMbLEing. *Sci. China Life Sci.* 62, 381–393. doi: 10.1007/s11427-019-9504-5

Li, Z. H., Meng, H., Ma, B., Tao, X., Liu, M., Wang, F. Q., et al. (2020). Immediate, multiplexed and sequential genome engineering facilitated by CRISPR/Cas9 in *Saccharomyces cerevisiae*. *J. Ind. Microbiol. Biotechnol.* 47, 83–96. doi: 10.1007/s10295-019-02251-w

Lian, J., HamediRad, M., Hu, S., and Zhao, H. (2017). Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. *Nat. Commun.* 8:1688. doi: 10.1038/s41467-017-01695-x

Lin, J. L., Zhu, J., and Wheeldon, I. (2017). Synthetic protein scaffolds for biosynthetic pathway co-localization on lipid droplet membranes. *ACS Synth. Biol.* 6, 1534–1544. doi: 10.1021/acssynbio.7b00041

Liu, H., Fan, J., Wang, C., Li, C., and Zhou, X. (2019). Enhanced β-amyrin synthesis in *Saccharomyces cerevisiae* by coupling an optimal acetyl-CoA supply pathway. *J. Agric. Food Chem.* 67, 3723–3732. doi: 10.1021/acs.jafc.9b00653

Liu, W., Luo, Z., Wang, Y., Pham, N. T., Tuck, L., Pérez-Pi, I., et al. (2018). Rapid pathway prototyping and engineering using in vitro and *in vivo* synthetic genome SCRaMbLE-in methods. *Nat. Commun.* 9:1936. doi: 10.1038/s41467-018-04254-0

Liu, X., Ding, W., and Jiang, H. (2017). Engineering microbial cell factories for the production of plant natural products: from design principles to industrial-scale production. *Microb. Cell Fact.* 16:125. doi: 10.1186/s12934-017-0732-7

Lü, B., Yang, X., Feng, X., and Li, C. (2016). Enhanced production of glycyrrhetic acid 3-O-mono-β-d-glucuronide by fed-batch fermentation using pH and dissolved oxygen as feedback parameters. *Chinese J. Chem. Eng.* 24, 506–512. doi: 10.1016/j.cjche.2015.12.003

Lv, X., Wang, F., Zhou, P., Ye, L., Xie, W., Xu, H., et al. (2016). Dual regulation of cytoplasmic and mitochondrial acetyl-CoA utilization for improved isoprene production in *Saccharomyces cerevisiae*. *Nat. Commun.* 7:12851. doi: 10.1038/ncomms12851

Ma, L., Li, Y., Chen, X., Ding, M., Wu, Y., and Yuan, Y.-J. (2019). SCRaMbLE generates evolved yeasts with increased alkali tolerance. *Microb. Cell Fact.* 18:52. doi: 10.1186/s12934-019-1102-4

Marienhagen, J., and Bott, M. (2012). Metabolic engineering of microorganisms for the synthesis of plant natural products. *J. Biotechnol.* 163, 166–178. doi: 10.1016/j.jbiotec.2012.06.001

Mikkelsen, M. D., Buron, L. D., Salomonsen, B., Olsen, C. E., Hansen, B. G., Mortensen, U. H., et al. (2012). Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metab. Eng.* 14, 104–111. doi: 10.1016/j.ymben.2012.01.006

Morgado, G., Gerngross, D., Roberts, T. M., and Panke, S. (2018). "Synthetic biology for cell-free biosynthesis: fundamentals of designing novel in vitro multi- enzyme reaction networks," in *Synthetic Biology - Metabolic Engineering*, eds H. Zhao and A. P. Zeng (Cham: Springer), 117–146. doi: 10.1007/10_2016_13

Moscou, M. J., and Bogdanove, A. J. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science* 326:1501. doi: 10.1126/science.1178817

Ng, C. Y., Farasat, I., Maranas, C. D., and Salis, H. M. (2015). Rational design of a synthetic Entner–Doudoroff pathway for improved and controllable NADPH regeneration. *Metab. Eng.* 29, 86–96. doi: 10.1016/j.ymben.2015.03.001

Ni, J., Zhang, G., Qin, L., Li, J., and Li, C. (2019). Simultaneously down-regulation of multiplex branch pathways using CRISPRi and fermentation optimization for enhancing β-amyrin production in *Saccharomyces cerevisiae*. *Synth. Syst. Biotechnol.* 4, 79–85. doi: 10.1016/j.synbio.2019.02.002

Ottaviani, A., Gilson, E., and Magdinier, F. (2008). Telomeric position effect: from the yeast paradigm to human pathologies? *Biochimie* 90, 93–107. doi: 10.1016/j.biochi.2007.07.022

Park, S. Y., Yang, D., Ha, S. H., and Lee, S. Y. (2018). Metabolic engineering of microorganisms for the production of natural compounds. *Adv. Biosyst.* 2:1700190. doi: 10.1002/adbi.201700190

Qi, D.-D., Jin, J., Liu, D., Jia, B., and Yuan, Y.-J. (2020). In vitro and *in vivo* recombination of heterologous modules for improving biosynthesis of astaxanthin in yeast. *Microb. Cell Fact.* 19:103. doi: 10.1186/s12934-020-01356-7

Qu, J., Cao, S., Wei, Q., Zhang, H., Wang, R., Kang, W., et al. (2019). Synthetic multienzyme complexes, catalytic nanomachineries for cascade biosynthesis *in vivo*. *ACS Nano* 13, 9895–9906. doi: 10.1021/acsnano.9b03631

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

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

Ronda, C., Maury, J., Jakociunas, T., Jacobsen, S. A. B., Germann, S. M., Harrison, S. J., et al. (2015). CrEdit: CRISPR mediated multi-loci gene integration in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* 14:97. doi: 10.1186/s12934-015-0288-3

Ryu, Y. S., Biswas, R. K., Shin, K., Parisutham, V., Kim, S. M., and Lee, S. K. (2014). A simple and effective method for construction of *Escherichia coli* strains proficient for genome engineering. *PLoS One* 9:e94266. doi: 10.1371/journal.pone.0094266

Sachdeva, G., Garg, A., Godding, D., Way, J. C., and Silver, P. A. (2014). *In vivo* co-localization of enzymes on RNA scaffolds increases metabolic production in a geometrically dependent manner. *Nucleic Acids Res.* 42, 9493–9503. doi: 10.1093/nar/gku617

Schwartz, C., Shabbir-Hussain, M., Frogue, K., Blenner, M., and Wheeldon, I. (2017). Standardized markerless gene integration for pathway engineering in *Yarrowia lipolytica*. *ACS Synth. Biol.* 6, 402–409. doi: 10.1021/acssynbio.6b00285

Si, T., Chao, R., Min, Y., Wu, Y., Ren, W., and Zhao, H. (2017). Automated multiplex genome-scale engineering in yeast. *Nat. Commun.* 8:15187. doi: 10.1038/ncomms15187

Siemon, T., Wang, Z., Bian, G., Seitz, T., Ye, Z., Lu, Y., et al. (2020). Semisynthesis of plant-derived Englerin A enabled by microbe engineering of guaia-6,10(14)-diene as building block. *J. Am. Chem. Soc.* 142, 2760–2765. doi: 10.1021/jacs.9b12940

Standage-Beier, K., and Wang, X. (2017). Genome reprogramming for synthetic biology. *Front. Chem. Sci. Eng.* 11, 37–45. doi: 10.1007/s11705-017-1618-2

Sun, W., Qin, L., Xue, H., Yu, Y., Ma, Y., Wang, Y., et al. (2019). Novel trends for producing plant triterpenoids in yeast. *Crit. Rev. Biotechnol.* 39, 618–632. doi: 10.1080/07388551.2019.1608503

Sun, W., Xue, H., Liu, H., Lv, B., Yu, Y., Wang, Y., et al. (2020). Controlling chemo- and regioselectivity of a plant P450 in yeast cell toward rare licorice triterpenoid biosynthesis. *ACS Catal.* 10, 4253–4260. doi: 10.1021/acscatal.0c00128

Thompson, A., and Gasson, M. J. (2001). Location effects of a reporter gene on expression levels and on native protein synthesis in *Lactococcus lactis* and *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* 67, 3434–3439. doi: 10.1128/aem.67.8.3434-3439.2001

Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., Xu, G., Forest, C. R., et al. (2009). Programming cells by multiplex genome engineering and accelerated evolution. *Nature* 460, 894–898. doi: 10.1038/nature08187

Wang, H. H., Kim, H., Cong, L., Jeong, J., Bang, D., and Church, G. M. (2012). Genome-scale promoter engineering by coselection MAGE. *Nat. Methods* 9, 591–596. doi: 10.1038/nmeth.1971

Wang, J., Jia, B., Xie, Z., Li, Y., and Yuan, Y. (2018). Improving prodeoxyviolacein production via multiplex SCRaMbLE iterative cycles. *Front. Chem. Sci. Eng.* 12, 806–814. doi: 10.1007/s11705-018-1739-2

Wang, L., Deng, A., Zhang, Y., Liu, S., Liang, Y., Bai, H., et al. (2018). Efficient CRISPR-Cas9 mediated multiplex genome editing in yeasts. *Biotechnol. Biofuels* 11:277. doi: 10.1186/s13068-018-1271-0

Wang, Y., and Yu, O. (2012). Synthetic scaffolds increased resveratrol biosynthesis in engineered yeast cells. *J. Biotechnol.* 157, 258–260. doi: 10.1016/j.jbiotec.2011.11.003

Wu, J., Du, G., Chen, J., and Zhou, J. (2015). Enhancing flavonoid production by systematically tuning the central metabolic pathways based on a CRISPR interference system in *Escherichia coli*. *Sci. Rep.* 5:13477. doi: 10.1038/srep13477

Wu, J., Zhang, X., Zhu, Y., Tan, Q., He, J., and Dong, M. (2017a). Rational modular design of metabolic network for efficient production of plant polyphenol pinosylvin. *Sci. Rep.* 7:1459. doi: 10.1038/s41598-017-01700-9

Wu, J., Zhou, P., Zhang, X., and Dong, M. (2017b). Efficient de novo synthesis of resveratrol by metabolically engineered *Escherichia coli*. *J. Ind. Microbiol. Biotehnol.* 44, 1083–1095. doi: 10.1007/s10295-017-1937-9

Wu, X.-L., Li, B.-Z., Zhang, W.-Z., Song, K., Qi, H., Dai, J. B., et al. (2017). Genome-wide landscape of position effects on heterogeneous gene expression

in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* 10:189. doi: 10.1186/s13068-017-0872-3

Xie, S. S., Qiu, X. Y., Zhu, L. Y., Zhu, C. S., Liu, C. Y., Wu, X. M., et al. (2019). Assembly of TALE-based DNA scaffold for the enhancement of exogenous multi-enzymatic pathway. *J. Biotechnol.* 296, 69–74. doi: 10.1016/j.jbiotec.2019.03.008

Xie, W., Liu, M., Lv, X., Lu, W., Gu, J., and Yu, H. (2014). Construction of a controllable beta-carotene biosynthetic pathway by decentralized assembly strategy in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* 111, 125–133. doi: 10.1002/bit.25002

Xu, X., Tian, L., Tang, S., Xie, C., Xu, J., and Jiang, L. (2020). Design and tailoring of an artificial DNA scaffolding system for efficient lycopene synthesis using zinc-finger-guided assembly. *J. Ind. Microbiol. Biotechnol.* 47, 209–222. doi: 10.1007/s10295-019-02255-6

Yan, Y. J., Kohli, A., and Koffas, M. A. G. (2005). Biosynthesis of natural flavanones in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* 71, 5610–5613. doi: 10.1128/aem.71.9.5610-5613.2005

Ye, L., Lv, X., and Yu, H. (2017). Assembly of biosynthetic pathways in *Saccharomyces cerevisiae* using a marker recyclable integrative plasmid toolbox. *Front. Chem. Sci. Eng.* 11, 126–132. doi: 10.1007/s11705-016-1597-8

Yee, D. A., DeNicola, A. B., Billingsley, J. M., Creso, J. G., Subrahmanyam, V., and Tang, Y. (2019). Engineered mitochondrial production of monoterpenes in *Saccharomyces cerevisiae*. *Metab. Eng.* 55, 76–84. doi: 10.1016/j.ymben.2019.06.004

Yu, Y., Chang, P., Yu, H., Ren, H., Hong, D., Li, Z., et al. (2018). Productive amyrin synthases for efficient alpha-amyrin synthesis in engineered *Saccharomyces cerevisiae*. *ACS Synth. Biol.* 7, 2391–2402. doi: 10.1021/acssynbio.8b00176

Yuan, J., and Ching, C. B. (2016). Mitochondrial acetyl-CoA utilization pathway for terpenoid productions. *Metab. Eng.* 38, 303–309. doi: 10.1016/j.ymben.2016.07.008

Zalatan, J. G., Lee, M. E., Almeida, R., Gilbert, L. A., Whitehead, E. H., La Russa, M., et al. (2015). Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* 160, 339–350. doi: 10.1016/j.cell.2014.11.052

Zhang, C., Li, M., Zhao, G.-R., and Lu, W. (2020). Harnessing yeast peroxisomes and cytosol acetyl-coA for sesquiterpene α-humulene production. *J. Agric. Food Chem.* 68, 1382–1389. doi: 10.1021/acs.jafc.9b07290

Zhang, G., Cao, Q., Liu, J., Liu, B., Li, J., and Li, C. (2015). Refactoring β-amyrin synthesis in *Saccharomyces cerevisiae*. *AIChE J.* 61, 3172–3179. doi: 10.1002/aic.14950

Zhang, L., Gao, Y., Liu, X., Guo, F., Ma, C., Liang, J., et al. (2019). Mining of sucrose synthases from Glycyrrhiza uralensis and their application in the construction of an efficient UDP-recycling system. *J. Agric. Food Chem.* 67, 11694–11702. doi: 10.1021/acs.jafc.9b05178

Zhang, L., Zhang, H., Liu, Y., Zhou, J., Shen, W., Liu, L., et al. (2020). A CRISPR–Cas9 system for multiple genome editing and pathway assembly in *Candida tropicalis*. *Biotechnol. Bioeng.* 117, 531–542. doi: 10.1002/bit.27207

Zhang, X. K., Wang, D. N., Chen, J., Liu, Z. J., Wei, L. J., and Hua, Q. (2020). Metabolic engineering of β-carotene biosynthesis in *Yarrowia lipolytica*. *Biotechnol. Lett.* 42, 945–956. doi: 10.1007/s10529-020-02844-x

Zhang, Y., Wang, J., Cao, X., Liu, W., Yu, H., and Ye, L. (2020). High-level production of linalool by engineered *Saccharomyces cerevisiae* harboring dual mevalonate pathways in mitochondria and cytoplasm. *Enzyme Microb. Technol.* 134:109462. doi: 10.1016/j.enzmictec.2019.109462

Zhang, Y., Wang, J., Wang, Z., Zhang, Y., Shi, S., Nielsen, J., et al. (2019). A gRNA-tRNA array for CRISPR-Cas9 based rapid multiplexed genome editing in *Saccharomyces cerevisiae*. *Nat. Commun.* 10:1053. doi: 10.1038/s41467-019-09005-3

Zhang, Y. H. P. (2011). Substrate channeling and enzyme complexes for biotechnological applications. *Biotechnol. Adv.* 29, 715–725. doi: 10.1016/j.biotechadv.2011.05.020

Zhao, S., Jones, J. A., Lachance, D. M., Bhan, N., Khalidi, O., Venkataraman, S., et al. (2015). Improvement of catechin production in *Escherichia coli* through combinatorial metabolic engineering. *Metab. Eng.* 28, 43–53. doi: 10.1016/j.ymben.2014.12.002

Zhao, Y., Fan, J., Wang, C., Feng, X., and Li, C. (2018). Enhancing oleanolic acid production in engineered *Saccharomyces cerevisiae*. *Bioresour. Technol.* 257, 339–343. doi: 10.1016/j.biortech.2018.02.096

Zhao, Y., Lv, B., Feng, X., and Li, C. (2017). Perspective on biotransformation and de novo biosynthesis of licorice constituents. *J. Agric. Food Chem.* 65, 11147–11156. doi: 10.1021/acs.jafc.7b04470

Zhao, Y.-J., and Li, C. (2018). Biosynthesis of plant triterpenoid saponins in microbial cell factories. *J. Agric. Food Chem.* 66, 12155–12165. doi: 10.1021/acs.jafc.8b04657

Zhou, P., Xie, W., Yao, Z., Zhu, Y., Ye, L., and Yu, H. (2018). Development of a temperature-responsive yeast cell factory using engineered Gal4 as a protein switch. *Biotechnol. Bioeng.* 115, 1321–1330. doi: 10.1002/bit.26544

Zhu, L. Y., Qiu, X. Y., Zhu, L. Y., Wu, X. M., Zhang, Y., Zhu, Q. H., et al. (2016). Spatial organization of heterologous metabolic system *in vivo* based on TALE. *Sci. Rep.* 6:26065. doi: 10.1038/srep26065

Zhu, M., Wang, C., Sun, W., Zhou, A., Wang, Y., Zhang, G., et al. (2018). Boosting 11-oxo-β-amyrin and glycyrrhetinic acid synthesis in *Saccharomyces cerevisiae* via pairing novel oxidation and reduction system from legume plants. *Metab. Eng.* 45, 43–50. doi: 10.1016/j.ymben.2017.11.009

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 02 September 2020
doi: 10.3389/fbioe.2020.524676

Design and Construction of Portable CRISPR-Cpf1-Mediated Genome Editing in *Bacillus subtilis* 168 Oriented Toward Multiple Utilities

Wenliang Hao[1], Feiya Suo[1], Qiao Lin[1], Qiaoqing Chen[1], Li Zhou[1], Zhongmei Liu[1], Wenjing Cui[1] and Zhemin Zhou[1,2]**

[1] *The Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi, China,* [2] *Jiangnan University (Rugao) Food Biotechnology Research Institute, Jiangsu, China*

OPEN ACCESS

Edited by:
Yi Wang,
Auburn University, United States

Reviewed by:
Zengyi Shao,
Iowa State University, United States
Junjie Yang,
Center for Excellence in Molecular Plant Sciences (CAS), China

***Correspondence:**
Wenjing Cui
wjcui@jiangnan.edu.cn
Zhemin Zhou
zhmzhou@jiangnan.edu.cn

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *06 January 2020*
Accepted: *12 August 2020*
Published: *02 September 2020*

Citation:
Hao W, Suo F, Lin Q, Chen Q, Zhou L, Liu Z, Cui W and Zhou Z (2020) Design and Construction of Portable CRISPR-Cpf1-Mediated Genome Editing in Bacillus subtilis 168 Oriented Toward Multiple Utilities. Front. Bioeng. Biotechnol. 8:524676. doi: 10.3389/fbioe.2020.524676

Bacillus subtilis is an important Gram-positive bacterium for industrial biotechnology, which has been widely used to produce diverse high-value added chemicals and industrially and pharmaceutically relevant proteins. Robust and versatile toolkits for genome editing in *B. subtilis* are highly demanding to design higher version chassis. Although the *Streptococcus pyogenes* (*Sp*) CRISPR-Cas9 has been extensively adapted for genome engineering of multiple bacteria, it has many defects, such as higher molecular weight which leads to higher carrier load, low deletion efficiency and complexity of sgRNA construction for multiplex genome editing. Here, we designed a CRISPR-Cpf1-based toolkit employing a type V Cas protein, Cpf1 from *Francisella novicida.* Using this platform, we precisely deleted single gene and gene cluster in *B. subtilis* with high editing efficiency, such as *sacA*, *ganA, ligD* & *ligV,* and *bac* operon. Especially, an extremely large gene cluster of 38 kb in *B. subtilis* genome was accurately deleted from the genome without introducing any unexpected mutations. Meanwhile, the synthetic platform was further upgraded to a version for multiplex genome editing, upon which two genes *sacA* and *aprE* were precisely and efficiently deleted using only one plasmid harboring two targeting sequences. In addition, we successfully inserted foreign genes into the genome of the chassis using the CRISPR-Cpf1 platform. Our work highlighted the availability of CRISPR-Cpf1 to gene manipulation in *B. subtilis*, including the flexible deletion of a single gene and multiple genes or a gene cluster, and gene knock-in. The designed genome-editing platform was easily and broadly applicable to other microorganisms. The novel platforms we constructed in this study provide a promising tool for efficient genome editing in diverse bacteria.

Keywords: CRISPR-Cpf1, *Bacillus subtilis*, multiplex genome editing, large fragment deletion, gene insertion, chassis microorganisms

INTRODUCTION

Bacillus subtilis, a well-characterized Gram-positive bacterium, has been regarded to be a "generally recognized as safe" (GRAS) microbe that can naturally secrete numerous extracellular proteins (Correa et al., 2020). *B. subtilis* is an ideal organism for industrial application, however, the available genetic tools are insufficient compared to other widely used microbial chassis, such as *Escherichia coli* and *Saccharomyces cerevisiae* (Keasling, 2010; Dong and Zhang, 2014). Gene-editing is of great utilization to reprogramming and reshaping the genome of synthetic chassis (Bai et al., 2018). In previous studies, several *B. subtilis* genome editing tools have been developed. Common gene knockout systems in *B. subtilis* include Cre/*loxP* recombination (Suzuki et al., 2005; Choi et al., 2015), MazF counter-selectable markers (Westbrook et al., 2016), and synthetic gene circuits (Jeong et al., 2015). Cre/*loxP* is a recombination system based on resistance selection markers, which can knock out or insert genes by recombining homologous fragments with Cre recombinase. However, this method needs to introduce a foreign resistance gene, which is not in line with the needs of eco-friendly hosts. The counter-selectable method based on MazF knocks out genes by introducing a toxin-antitoxin (TA) system from *E. coli* (Zhang et al., 2006). Although this method does not introduce resistance markers on chromosomes, its efficiency is very low. Recently, gene knockout methods based on synthetic gene circuits have been constructed in *B. subtilis* (Jeong et al., 2015). Although it was unnecessary to introduce foreign resistance marker genes, the efficiency of knocking out large gene clusters is very low. Thus, the genome engineering of *B. subtilis* needs an effective method without antibiotic resistance markers (Suzuki et al., 2005).

Recently, the Class 2 clustered regularly interspaced short palindromic repeat (CRISPR) system has been employed as a powerful tool for genome editing and transcription regulation in many organisms, including bacteria (Jiang et al., 2015; Westbrook et al., 2016; Wu et al., 2018), yeast (Bao et al., 2015), plant (Gao et al., 2017), and mammals (Hwang et al., 2013). CRISPR systems are divided into two categories on the basis of the configuration of their effector molecules (Zetsche et al., 2015). Different from the class 1 CRISPR system, which requires various Cas proteins to coordinate with each other and bind to the crRNA to form a ribonucleoprotein (RNP) complex, class 2 CRISPR system employ a large single-component Cas protein in conjunction with crRNA to mediate genome editing (Zetsche et al., 2015). In type II CRISPR system, Cas9 from *Streptococcus pyogenes* is widely used because it has been studied very clearly. Zhang et al. (2016) constructed the AIO system in *B. subtilis* ATCC 6051a by using Cas9 from *Streptococcus pyogenes.* Using this system, the authors disrupted specific genes (including *srfC*, *spoIIAC*, *nprE*, *aprE,* and *amyE*) in *B. subtilis* ATCC 6051a with 33–53% efficiency (Zhang et al., 2016). The production of β-cyclodextrin glycosyltransferase by modified *B. subtilis* ATCC 6051a (ΔsrfC, ΔspoIIAC, ΔnprE, ΔaprE, and ΔamyE) is more than 2.5 times that of wild-type *B. subtilis* ATCC 6051a. Similarly, Westbrook et al. (2016) developed a CRISPR-Cas9-based toolkit that can knockout, knock-in, knockdown and point mutations of target genes in *B. subtilis*. They employed a strategy of expressing Cas9 and transcribing gRNA on chromosomes, and the authors believe that this method obviates the instability of multicopy plasmid in the host and the pressure of plasmid on the host. CRISPR-Cas9 system requires three essential factors to cleave the genomic DNA: the CRISPR RNA (crRNA), the *trans*-activating CRISPR RNA (tracrRNA), and the Cas9 nuclease (Chen et al., 2017). In this system, crRNA binds partially to the complementary tracrRNA prior to association with Cas9, allowing to form a gRNA-Cas9 complex. The complex identifies specific target site of genomic DNA based on the PAM sequence and generates blunt-ended double strand breakage (DSB) (Hong et al., 2018). Although CRISPR-Cas9 system has achieved huge success in genome editing in different organisms, it has prominent drawbacks, including severe off-target effects and certain unknown toxicity, resulting in low efficiency in a specific case (Fu et al., 2013; Hsu et al., 2013; Jiang et al., 2017; Hong et al., 2018). Recently, an array of novel Cas proteins has been increasingly developed, such as Cpf1 (Zetsche et al., 2015; Hong et al., 2018), Cas12b (Teng et al., 2018; Strecker et al., 2019), and CasX (Liu et al., 2019) from diverse bacteria. Among these Cas proteins, Cpf1 is a protein from bacterial immune system, which has recently been engineered as a genome editing tool in *Clostridium difficile* (Hong et al., 2018), *Corynebacterium glutamicum* (Jiang et al., 2017), and rice (Wang et al., 2017).

Cpf1, derived from a type V CRISPR system, is an effector Cas protein distinct from Cas9 in structure and function (Zetsche et al., 2015). The prominent advantage of Cpf1 over Cas9 is that the maturity of CRISPR arrays does not require additional tracrRNA, so that Cpf1 is able to process pre-crRNA to mature crRNA (Zetsche et al., 2015, 2017). This feature resolves the drawback in the construction of multiple or large expression constructs using Cas9, upon which the procedure of multiplexed-gene editing is possible to be simplified (Zetsche et al., 2015, 2017; Fonfara et al., 2016). Moreover, CRISPR-Cpf1 complexes efficiently cleave the target DNA utilizing a T-rich PAM sequence rather than the G-rich PAM sequence in CRISPR-Cas9 system (Zetsche et al., 2015), which is probably more efficient in *B. subtilis*. In addition, Cpf1 leaves a staggered end with a 5′ overhang after cleaving DNA, which facilitates repairing the nicked DNA by non-homologous end joining (NHEJ) or homology-directed repair (HDR) after cutting (Zetsche et al., 2015; Jiang et al., 2017).

Previous studies have shown that Cpf1 is greatly superior to Cas9 in genome editing (Kim et al., 2016). Compared with Cas9, Cpf1 is a small protein that contains a single well-identified nuclease domain rather than two nuclease domains for Cas9 (Zetsche et al., 2015). For instance, Cpf1 has only 1300 amino acids (Zetsche et al., 2015), which is more suitable to deliver Cas-sgRNA complex. Importantly, Cpf1 is lower in potential toxicity to the host compared to Cas9 (such as *Sp*Cas9) (Jiang et al., 2017). Therefore, it is a more suitable candidate Cas protein for genome editing (Zetsche et al., 2017). In previous studies, CRISPR-Cpf1 system has been engineered as a powerful genome-editing tool and applied to different organisms, including rice (Wang et al., 2017), soybean (Kim et al., 2017), mouse (Hur et al., 2016), zebrafish (Hwang et al., 2013), human cell (Kim et al., 2016), *Mycobacterium smegmatis* (Yan et al., 2017), *yeast*

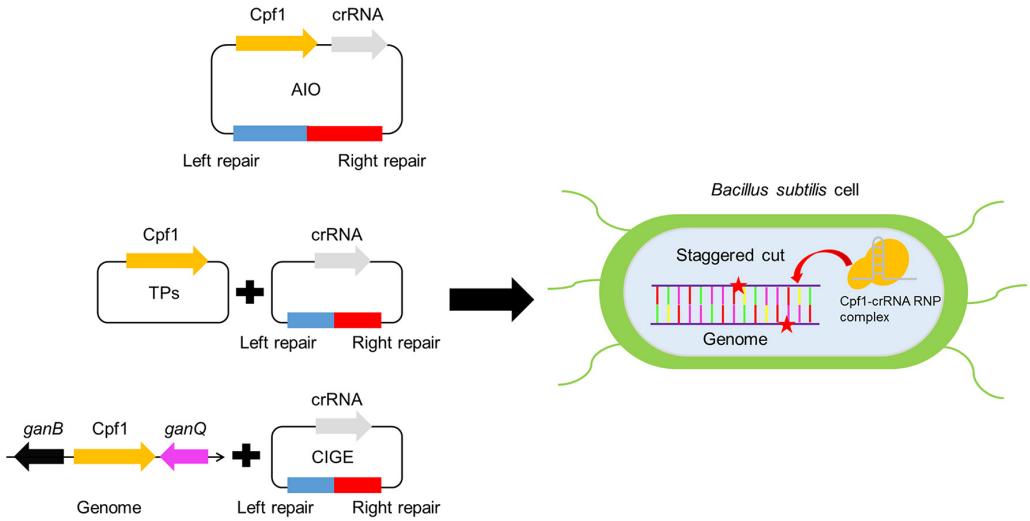


FIGURE 1 | Development of genome editing toolkit based on CRISPR-Cpf1 in *Bacillus subtilis*.

(Buchmuller et al., 2019), and *C. difficile* (Hong et al., 2018). Recently, Wu et al. (2020) constructed a gene editing system based on CRISPR-Cpf1 in *B. subtilis*, and used the system to perform gene knock-out, knock-in, and regulation of target gene expression. The authors employed a two-plasmid model and improved the efficiency of gene editing (including double gene knockout, knock-in, and point mutation) by overexpressing mutated NgAgo* on plasmids. However, it is unclear whether the efficiency of gene knock-out by CRISPR-Cpf1 system in *B. subtilis* will be varied with the size of the target gene fragment. And it is still to be determined whether CRISPR-Cpf1 system can mediate deletion of large gene cluster in *B. subtilis*.

In this study, we broadened genome editing toolkit based on CRISPR-Cpf1 system employing different strategies in *B. subtilis* and successfully applied the CRISPR-Cpf1 system to the deletion of single gene of various sizes as well as multiplex-gene editing in high efficiency (**Figure 1**). Importantly, we also established an efficient chromosome-integration genome editing (CIGE) platform to achieve precise insertion of gene of interest. These results exhibit that the CRISPR-Cpf1-based tools we have designed and built in this study have highly flexible property that is not only used to deletion of gene of diverse sizes but also serve as a proficient platform to precisely insert heterologous genes into chromosome. This toolbox is of great importance to develop high version of chassis by precisely editing the genome *B. subtilis*, which is great potential to extend the synthetic biology of *B. subtilis*.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

All the *E. coli* and *B. subtilis* strains used in this study are listed in **Supplementary Table S1**. The JM109 clone *E. coli* strains (General Biosystems, China) was used as the general host for plasmid construction and gene cloning. The transformation of the JM109 *E. coli* strains was conducted through chemical transformation using Competent cells from General Biosystems. *E. coli* strains were grown in Luria-Bertani (LB) medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, pH 7.0) supplement with ampicillin (100 μg/mL) or spectinomycin (100 μg/mL) when necessary. Transformation of *B. subtilis* cells was carried out by the two-step transformation procedure (Anagnostopoulos and Spizizen, 1961). *B. subtilis* strains were cultivated in Luria-Bertani (LB) medium supplement with spectinomycin (100 μg/mL), chloramphenicol (6 μg/mL) or kanamycin (50 μg/mL) and LB solid medium supplemented with 1% glucose.

Plasmids Construction

All the plasmids used in this study were listed in **Supplementary Table S1**. All the DNA primers used in this study were listed in **Supplementary Table S2**. All the crRNA used in this study were listed in **Supplementary Table S3**. The nucleotide and amino acid sequences of *Fn*Cpf1 and *Sp*Cas9 nucleases were shown in the complementary sequences. The crRNA and sgRNA sequences used in this paper were also listed in the Supplementary Sequences.

Construction of All-in-One System

The pHTsacA plasmid was derived from pHT01, an expression plasmid for *B. subtilis* (MoBiTec, Göttingen, Germany). P43 promoter and RBS were amplified from pBSG03 using the primers pHT-P43-F and pHT-P43-R while the backbone of the plasmid, harboring *lacI* gene, was amplified using the primer pHT-P43-b-F and pHT-P43-b-R. Then, P43 promoter and RBS were inserted into the shuttle vectors using Gibson assembly (Gibson et al., 2009), yielding pHT-P43-RBS. Accordingly, *Cpf*1

gene was cloned into pHT-P43-RBS using the primers P43-FnCpf1-F/R and P43-FnCpf1-b-F/R, yielding pHT-P43-RBS-FnCpf1. The *sacA*-targeting crRNA expression cassette under the control of a strong promoter P_{veg} was synthesized by GENEWIZ Inc., Ltd. (Wuxi, China) and cloned into pHT-P43-RBS-FnCpf1 using the primers pHT-pVeg-sacAcrRNA-F/R and pHT-sacAcr-b-F/R, producing pHT-FnCpf1-*sacA*crRNA. Primer pair sacA-HA-F/R and pHT-HA-b-F/R was used to amplify the 1.2-kb donor DNA template and its bone, respectively. Finally, *sacA* homologous arm was cloned into pHT-FnCpf1-sacAcrRNA, yielding pHTsacA (for the detailed construction method of pHTganA-Cpf1,pHTganA-Cas9, and pHTDV, refer to the Supplementary Method in the **Supplementary Material**).

The plasmid pHTsfGFPKiT was used to delete *sacA* from chromosome of *B. subtilis*. To construct this plasmid, the sfGFP DNA fragment was amplified from pBPylbp-sfGFP-Ter plasmid using primers sfGFPKi-F/R. The backbone of the plasmid pHTsacA was amplified at the same time using primers pHT-all-sfGFPKi-b-F/R. Then, sfGFP was cloned between the homologous arm of the plasmid pHTsacA, yielding pHTsfGFPKi.

Construction of Two Plasmids System

To construct gene deletion tool of two-plasmid format, gene *Cpf1* was firstly cloned into pHT01 vector using the primer pHT-Pgrac-Cpf1-F/R and pHT-Pgrac-Cpf1-b-F/R by the Gibson assembly, generating pHT01-Cpf1 activated by Isopropyl-β-D-thiogalactopyranoside (IPTG). To generate the vector for expression of the *sacA*-targeting crRNA, we cloned *sacA*-targeting crRNA expression cassette to pAD123 vector, harboring coding sequences (CDSs) of gfpmut3a and rep60 from pAT1060 origin using the primers pAD123-pVeg-sacAcrT-F/R and pAD123-sacAcrT-b-F/R, generating pAD-pVeg-*sacA*crRNA. The *sacA* homologous arm was cloned into pAD-pVeg-*sacA*crRNA using the primer pAD-sacAH-F/R and pAD-sacAH-b-F/R, yielding pADsacA.

Construction of Chromosomally Integrated Genome Editing System (CIGE)

To construct a CIGE system that had *Cpf1* integrated in the chromosome in *B. subtilis*, pAX01-Cpf1 under the control of P_{xylA} was constructed to this end using the primer pAX01-Cpf1-F/R and pAX01-Cpf1-b-F/R. Then, rrnBT1 terminator and rrnBT2 terminator was fused downstream of Cpf1 expression cassette. Chloramphenicol-resistant gene (*cat*), *lox*66-71 site and the homologous arm of *ganA* gene was amplified using the primer lacA-Cpf1-F/R, the PCR fragment was integrated specifically into *ganA* site of genome, yielding the strain BS-*ganA*'-Cpf1. To construct plasmid harboring *sacA*-targeting crRNA, we cloned sacAcrRNA expression cassette and *sacA* homologous arm from pHTsacA using the primer pB-pVeg-sacAHA-F/R and pB-pVeg-sacAHA-b-F/R into pBSG03 vector, producing the plasmid pBsacA (for the detailed construction method of pBbac and pBpps, refer to the Supplementary Method in the **Supplementary Material**).

pBsfGFPKi was constructed to insert super folder green fluorescence protein (sfGFP) into the *sacA* locus. The sacAcrRNA expression cassette, the homologous arms of *sacA*, and the sequence of sfGFP were amplified from pHTsfGFPKi using the primers pB-sacAHA-sfGFP-F/R. Meanwhile, the backbone of the plasmid pHTsfGFPKi was amplified using the primers pB-sacAHA-sfGFP-b-F/R. Accordingly, the two PCR products were fused by Gibson assembly, generating pBsfGFPKi.

The plasmid pB-sacAKo-aprEKi was constructed by one-pot Golden Gate assembly reaction (Engler et al., 2009). Specifically, pB-sacAcrRNA-HA used primers AmpM-*Bsa*I-F/R and RepM-*Bsa*I-F/R to mutate ampicillin-resistance gene and repB replication gene to eliminate *Bsa*I recognition sites. The purpose to do this was to prevent the *Bsa*I restriction enzyme from non-specifically cutting other positions in the reaction of Golden Gate assembly. Then, we amplified the 500-bp homologous arms upstream and downstream of *aprE* from the plasmid pB-aprEHA using the primers pB-aprEHA-F/R. The amplicon was further fused to the plasmid pBsacA using primers pB-aprEHA-b-F/R, producing pBsacA-aprEHA. The gene of mCherry was amplified from the plasmid pBP43-GFP-mCherry using primers pB-mCherry-F/R prior to fusing with the linearized plasmid pBsacA-aprEHA (amplified by primers pB-mCherry-b-F/R) by Gibson assembly, generating pBsacA-aprEHA-mCh. As Golden gate assembly requires two *Bsa*I restriction site and orthogonal overhangs to ensure the target fragment correctly being inserted into the recipient vector, we used pBsacA-aprEHA-mCh as a template and perform two rounds of rPCR to obtain two *Bsa*I restriction sites using primers *Bsa*I-1-F/R and *Bsa*I-2-F/R, so as to enable the resulting plasmid being digested by *Bsa*I. The crRNA sequence targeting *aprE* was amplified from pHTaprE using primers aprEcrRNA-*Bsa*I-F/R. To construct pB-sacAKo-aprEKi, a restriction-ligation was performed in a mixture containing crRNA expression cassette targeting *aprE*, the recipient vector pBsacA-aprEHA-mCh, *Bsa*I enzyme and T4 DNA ligase, generating pB-sacAKo-aprEKi. For double deletion of *sacA* and *aprE*, a single crRNA array was synthesized from GENEWIZ, Inc. (Suzhou, China), which contained P_{vegM} promoter, crRNA targeting *sacA* and *aprE*, as well as BT5 terminator (screened by our laboratory). The array was amplified from pUC57 plasmid using primers SA-BT5-F/R, the resulting PCR product was inserted into pB-sacAHA-aprEHA plasmid (containing the homologous arms of *sacA* and *aprE*) with primers SA-BT5-b-F/R by Gibson Assembly, yielding the recombinant plasmid pB-PvegM-SAKo.

Plasmid Curing

To cure mutant strains of the pBSG so as to enable their use in a second round of genome editing, the mutant strains were inoculated into LB medium with a final concentration of 0.0005% SDS without antibiotics (Trevors, 1986), and then incubated at 37°C, 200 rpm for 20 h. Accordingly, the culture was diluted and spread on LB plates without antibiotics. Colonies were carefully picked up and dotted at the same positions on two LB plates with and without antibiotics, respectively. Antibiotics-sensitive colonies were

picked and propagated in 5-mL LB medium. Then, plasmid-free mutants were further confirmed through PCR. For elimination of pHT01 and pAD123 plasmid, we referred to previous research and achieved it (Yamashiro et al., 2011; So et al., 2017).

RESULTS AND DISCUSSION

Design of a CRISPR-Cpf1 in All-in-One (AIO) System for Deletion of Large Genes in *B. subtilis*

To achieve the goal for highly efficient editing the genome in *B. subtilis*, we employed a type V Cas protein, Cpf1 from *Francisella novicida* to design an all-in-one system (AIO). In this system, crRNA and Cpf1 were expressed constitutively by P_{veg} and P43 promoters, respectively. The expressed crRNA carried Cpf1 to a specific site of the target gene, where the cleavage of Cpf1 to the target gene would cause the precise recombination of homologous segments, leading to the deletion of the target gene (**Figure 2A**). To compare the performance of gene editing between Cpf1 and Cas9, we constructed a CRISPR-Cas9-based AIO system. To verify the functionality of two AIO systems, we selected the *ganA* gene (gene of medium size) to identify the efficiency of deletion mediated by the two AIO systems. The *ganA*, which was of 2064 bp in length, encodes β-galactosidase, which participates in the degradation of galactose. Disruption or complete deletion of *ganA* does not affect the growth of host. Therefore, we chose *ganA* gene as the target to explore the functionality of the two designed CRISPR systems (CRISPR-Cpf1 and CRISPR-Cas9) for deletion of large genes. The two editing plasmids, pHTganA-Cpf1 (refer to addgene #158647) and pHTganA-Cas9, were separately introduced into *B. subtilis* 168. The transformants were separately picked up and then inoculated into newly prepared LB medium to grow. Then the culture was diluted to appropriate density and spread onto LB agar plates. Twenty clones were randomly picked for cPCR to screen *ganA*-disrupted mutants from pHTganA-Cpf1 and pHTganA-Cas9 plates, respectively. The results showed that 20 clones of *ganA*-disrupted mutants had smaller cPCR product than that of wild-type strain (ck), indicating that *ganA* has been successfully deleted on pHTganA-Cpf1 plate (efficiency was 100%, **Figure 2B**). However, we only screened 15 *ganA*-deletion mutants on the pHTganA-Cas9 plate (efficiency was 75%, **Figure 2B**). Furthermore, the precision of deletion sites for these engineered strains were validated by sequencing the cPCR products (**Figure 2B**). These data displayed that the efficiency of CRISPR-Cpf1 and CRISPR-Cas9 systems was sufficient for general gene editing. However, CRISPR-Cpf1 system would be more suitable for larger fragments to be knocked out. Therefore, our following experiments all aimed at the engineering design of CRISPR-Cpf1 system. To further verify the functionality of the AIO system based on CRISPR-Cpf1, we selected *sacA* to identify the efficiency of deletion mediated by the system. The *sacA* (NC_000964:3902858), which was of 2400 bp in length, encodes sucrose-6-phosphate hydrolase, which is not an essential gene for *B. subtilis* (Zhu and Stulke, 2018). In most integrated systems for *B. subtilis*, *sacA* site was widely used to an integration site (Radeck et al., 2013). Here, we selected PAM sequence 5′-TTTG-3′ (Zetsche et al., 2015) in *sacA* gene, and then constructed PHTsacA plasmid harboring *Cpf1* gene and target-*sacA* crRNA that under the control of constitutive promoters P43 and P_{veg}, respectively (**Figure 2A**). The 1200-bp homologous arm was inserted downstream of the CRISPR-Cpf1 expression cassette. Based on the deletion design, the fragment should be 240 bp after deletion. Plasmid pHTsacA was then introduced into *B. subtilis* 168. The transformations were picked, inoculated into 5 mL of LB medium, and incubated aerobically at 37°C overnight. Then the culture was spread onto LB plates. Twelve colonies were randomly collected to perform colony PCR (cPCR) to verify the deletion efficiency. The PCR product should be of 1655 bp if *sacA*-deletion was unsuccessful using designed primers, in contrast, it should be of 240 bp smaller than that of the wild type *B. subtilis* 168. The data showed that introduction of pHTsacA into *B. subtilis* 168 resulted in 100% deletion (**Figure 2C**). These results were further confirmed by sequencing of one successfully deleted mutant (**Figure 2C**), indicating that the designed synthetic CRISPR-Cpf1 tool has relatively high editing efficiency targeting genes with short lengths.

Furthermore, to further identify the available range of AIO-based CRISPR-Cpf1 system, we employed the system to delete larger gene cluster. The *ligD* & *ligV* gene cluster, which is of 2775 bp in length, was selected as the target. The gene cluster *ligD* & *ligV* is involving in the non-homologous end-joining (NHEJ) process (Pitcher et al., 2007). Although the gene cluster is important in maintaining chromosomal stability in bacteria, it is non-fatal to *B. subtilis* when *ligD* & *ligV* cluster is deficient. Therefore, we chose *ligD* & *ligV* gene cluster as the target to test the functionality of deletion of large gene cluster using designed CRISPR-Cpf1 system. The AIO system was selected to construct the editing system (**Figure 2A**). The pHTDV plasmid harboring *ligD & ligV*-targeting crRNA and the homologous arms was constructed in the similar procedure as that of pHTganA but had longer homologous arms (600 bp). The editing plasmid, pHTDV, was introduced into *B. subtilis* 168. The transformants were picked and then inoculated into newly prepared LB medium to grow. Then the culture was diluted to appropriate density and spread onto LB agar plates. Nineteen colonies were randomly picked up to screen *ligD* & *ligV*-disrupted mutants by cPCR. The results showed that 18 of the 19 clones were confirmed to be the *ligD* & *ligV*-deficient strains, suggesting an editing efficiency of 94.7% (**Figure 2D**). Furthermore, the precision of deletion sites for these engineered strains were validated by sequencing the cPCR products. The data displayed that all these *ganA*, *sacA* and *ligD* & *ligV* mutants had accurate deletion sites as we had designed (**Figures 2B–D**).

In this work, we authenticated that the AIO system was capable of being applied to delete single gene with high efficiency in *B. subtilis*. AIO-based gene editing employing CRISPR-Cas9 was also constructed in *B. subtilis* ATCC 6051a (Zhang et al., 2016). The prominent feature of AIO was that Cas protein, sgRNA and homologous arms were all concentrated on one plasmid. When the gene editing was completed, the

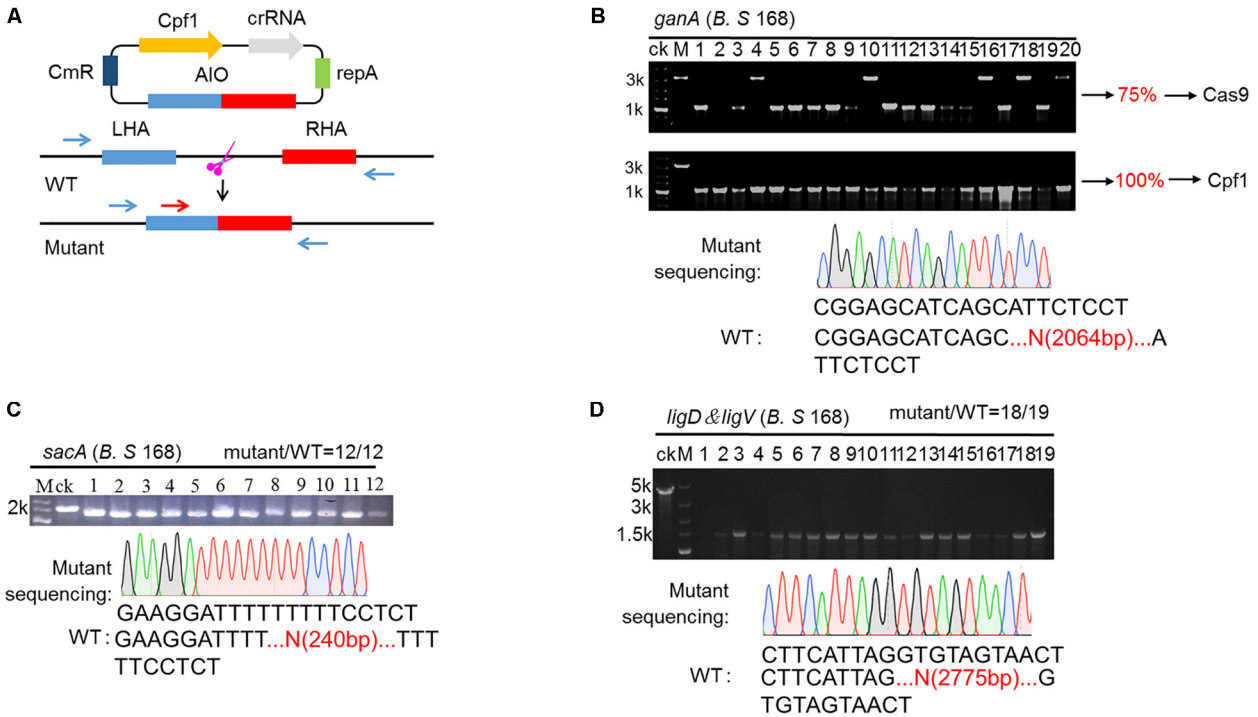


FIGURE 2 | CRISPR-Cpf1-mediated genome editing in the *Bacillus subtilis* via using AIO system. **(A)** Schematic illustration of the editing procedures. The blue arrows are the primers utilized for PCR validation of the editing efficiency. The red arrow is the primer used for sequencing. **(B)** Deletion of *ganA* gene mediated by AIO system based on CRISPR-Cpf1 and CRISPR-Cas9 in *B. subtilis* 168. The efficiency of deletion of *ganA* by AIO system based on CRISPR-Cpf1 was 20/20 (100%). The efficiency of deletion of *ganA* gene by AIO system based on CRISPR-Cas9 was 15/20 (75%). Lane M, the 5-k DNA marker from Takara with number on the left representing the band size in kb. The lane labeled "ck" is the PCR product from the wild-type strain as a control. **(C)** AIO-mediated deletion of the *sacA* gene in the *B. subtilis* 168 strain. The editing efficiency was 12/12. **(D)** AIO-mediated deletion of the *ligD* & *ligV* genes in the *B. subtilis* 168 strain. The editing efficiency was 18/19.

plasmid will be eliminated and no foreign genes will be introduced into the genome. However, the plasmids of AIO were often large, which led to low transformation efficiency and instability of plasmids (**Supplementary Table S5**). To improve the stability of plasmid replication, we use pHT01 vector as the skeleton of AIO, which carries a repA replicon. The vector pHT01 belongs to a medium copy plasmid. Because repA protein in the way of θ replication to multiply, the pHT01 vector is relatively stable. If we want to further improve the stability of the plasmid, we can integrate Cpf1 into the host genome, which can reduce the replication pressure of plasmid. Therefore, to break through the limitations of AIO system and provide flexible tools to facilitate gene editing in complex genetic context, more robustness, and efficiency of CRISPR-Cpf1 system should be built to allow reliable genome editing.

Construction and Validation of Two-Plasmid (TP)-Based, Cpf1-Mediated Gene-Deleting System in *B. subtilis*

To broaden an alternative form of CRISPR-Cpf1 system, the two plasmids (TP) system, and verify whether the system was able to efficiently delete genes in *B. subtilis*. In this system, we also chose *sacA* as the target. First, we constructed plasmid pHT01-Cpf1 (refer to addgene #158648), harboring Cpf1 under the control of the inducible promoter P_{grac} (**Figure 3A**). Because pHT01 and pAD123 contained the same chloramphenicol resistance gene, we substituted the chloramphenicol-resistant gene from pHT01 vector with spectinomycin-resistant gene, allowing to screen the transformants with different antibiotics (**Figure 3A**). We constructed plasmid pADsacA (refer to addgene #158649), derived from the backbone of pAD123, to constitutively express the crRNA targeting *sacA* mediated by P_{veg} promoter (**Figure 3A**). Then, we sought to identify the efficiency of disruption of *sacA* by TPs system. We sequentially transformed plasmids pHT01-Cpf1 and pAD-sacA into *B. subtilis*. The two types of transformants were picked up, and inoculated into LB medium to propagation. When OD_{600} reached approximately 0.5, Isopropyl-β-D-thiogalactopyranoside (IPTG) with final concentration of 1 mM was added to induce the expression of Cpf1. Then the culture was spread onto LB agar plates. After incubated overnight, 23 colonies were randomly picked to perform cPCR to verify the deletion efficiency. The results exhibit that all the colonies (23/23) were the successfully deleted mutant harboring a 240 bp-deleted sequence of *sacA* (**Figure 3B**). To further validate the accuracy of gene deletion, we confirmed the sequence of disrupted *sacA* by sequencing.

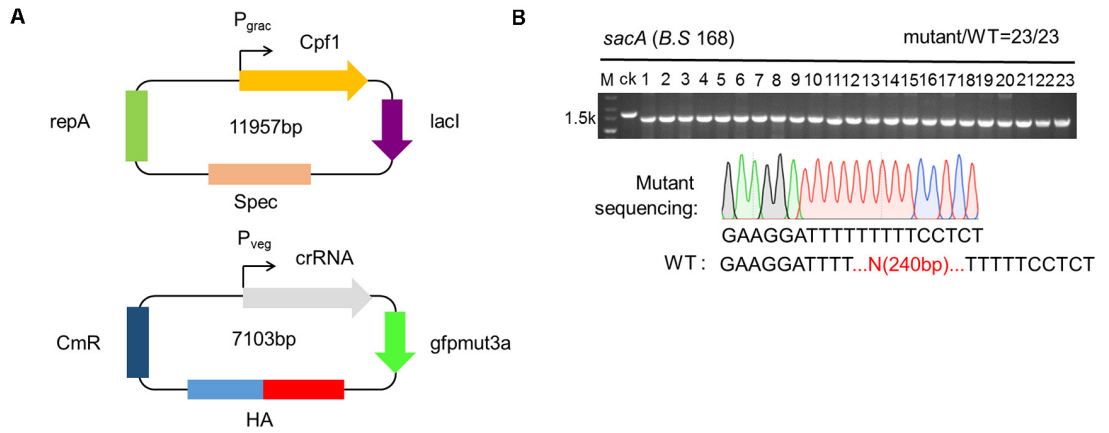


FIGURE 3 | Construction of the two plasmids (TPs) for genome engineering system. **(A)** Schematic illustration of pHTCpf1 plasmid and pADsacA plasmid containing *sacA* crRNA transcription module as well as donor DNA template. **(B)** TPs-mediated disruption of the *sacA* gene in the *Bacillus subtilis 168* strain. The editing efficiency was 20/20. The lane labeled "ck" is the PCR product from the wild-type strain as a control.

The data validated the successful deletion of 240 bp of *sacA*, suggesting that the deletion efficiency mediated by TPs strategy is as high as 100% (**Figure 3B**). To further compare the performance CRISPR-Cpf1 and CRISPR-Cas9 using two-plasmid strategy to knock out the same gene in *B. subtilis* 168, we selected *ganA* (2064 bp) gene as our target. The results showed that the knockout efficiency of CRISPR-Cpf1 system was much higher than that of CRISPR-Cas9 system (10/11 vs. 6/11, **Supplementary Figure S2**). These data suggest that the function of CRISPR-Cpf1 in knockout of single gene is higher than that of CRISPR-Cas9 with the same strategy (two-plasmid system).

CRISPR-Cpf1-Based CIGE (CCB-CIGE) Platform Is Superior to Highly Efficient Gene Insertion in *B. subtilis*

Gene insertion is another critical issue to genome editing. Therefore, robust and efficient insertion system is valuable tool for precisely editing the genome of *B. subtilis*. To evaluate the availability of the CRISPR-Cpf1-based AIO system for insertion of heterologous gene into the genome of *B. subtilis*, we firstly constructed an AIO plasmid pHTsfGFPKi, of which the expression of Cpf1 and *sacA*-targeting crRNA was controlled by P43 and P_{veg}, respectively. The coding region of *sfGFP* was flanked by upstream and downstream homologous arms of *sacA* (**Figure 4A**). We introduced pHTsfGFPKi into *B. subtilis* 168 to implement insertion function. We randomly chose 23 colonies from the plate and performed cPCR. By evaluating the size of cPCR product for each selected mutant, we found that two clones contained larger size of cPCR product than that of the "ck," whereas the others had the same size as that of the ck, indicating that *sfGFP* was successfully inserted into the genome of the two clones among the 23 selected colonies. The efficiency of insertion was calculated to be 9% (2/23) (**Figure 4B**). We infer that because the expression of Cpf1 or crRNA might be unstable from the plasmid that is too large in size.

To resolve the problem of low insertion efficiency by AIO system, we designed CRISPR-Cpf1-based CIGE (CCB-CIGE) platform to improve the expression of the components to elevate efficiency of gene insertion (**Figure 4C**). In this strategy, we chose pBSG as the skeleton and constructed a plasmid, pBsfGFPKi, harboring a *sacA*-targeted crRNA sequence controlled by P_{veg}. The coding region of *sfGFP* was flanked by upstream and downstream homologous arms. Because pBSG vector belongs to high copy plasmid, we can improve the expression of each component by using it. However, the replication protein of repB carried by the plasmid will multiply in the form of rolling circle replication, which may lead to instability of the plasmid. The plasmid was then introduced into the host BS-*ganA*'-Cpf1 (the Cpf1 gene under the control of P_{xylA} was integrated in the chromosome of *B. subtilis* via substitution for *ganA*). After 16 h induction by 1% xylose, 23 colonies were randomly chosen to identify the insertion results by cPCR. The data showed that the cPCR products from 17 out of 23 clones were larger compared to that of ck, indicating that the *sfGFP* had been successfully inserted into the chromosome of these clones (**Figure 4D**). The insertion efficiency was confirmed to be 74% (**Figure 4D**). According to the previous research, efficiency of gene knock-in in *B. subtilis* can be significantly improved by iterative genome engineering (So et al., 2017). Therefore, we prolonged the incubation time to 24 h, and the insertion efficiency increased to 82% (**Figure 4D**). To further validate that insertion occurred precisely at the *sacA* site, we sequenced the region from the upstream to the downstream homologous arm of No. 1 mutant among the successfully-inserted mutants. The map of sequencing revealed that *sfGFP* has been accurately inserted into specific site of the *sacA* gene using an optimized CRISPR-Cpf1-mediated gene knock-in strategy, without any unintended mutations (**Figure 4D**). These results manifest that our optimized CRISPR-Cpf1 system accurately insert target genes into preset positions in the chromosome.

Gene knock-in based on CRISPR system is an important technology for metabolic engineering and genomic function research. A potential application is that when we construct

metabolic engineering strains, we often need to overexpress some genes to improve the titer of metabolites. However, overexpression of some genes on plasmids often brings great pressure to the host. Therefore, integrating the target gene into the genome for overexpression has become a preferred strategy. Another advantage is to study the function of some genes in the genome. When we study a gene in the genome, we often take the surrounding genetic environment into account. At this time, *in situ* expression of gene is often much more important than overexpression on plasmid. Although we have constructed a gene knock-in system based on CRISPR-Cpf1 in this paper, we only verified it using *sf* GFP. In the current research, large gene cluster knock-in is an indispensable technology for the study of synthetic biology. Therefore, the strategy of gene knock in based on CRISPR-Cpf1 constructed in this work needs to be further optimized and the ability of large gene cluster knock-in needs to be improved. As the CCB-CIGE strategy was more efficient than AIO, we used CCB-CIGE in the following experiments.

Design of CRISPR-Cpf1 for Precise Deletion of Large Gene Cluster

Previous studies have shown that the deletion of large genome in *Bacillus* sp. plays a crucial role in heterologous expression of proteins, genome reduction (Westers et al., 2003), strain improvement (Thwaite et al., 2002) and overproduction of antibiotics (Zobel et al., 2015). There are three large gene clusters in *B. subtilis* encoding polyketide synthase (*pks*), plipastatin synthetase (*pps*), and surfactin (*srf*), which account for 7.7% of the total genome (So et al., 2017). There are also gene clusters in *B. subtilis* that synthesize peptide antibiotics, such as *bac* operon (Ozcengiz and Ogulur, 2015). The deletion of these gene clusters that synthesize secondary metabolites has minor effect on the growth of *B. subtilis*. Moreover, deletion of these non-essential regions in *B. subtilis* is essential to construction of the minimal genome in microbial chassis. Although a counter-selectable marker system based on synthetic gene circuits has been developed to delete *pps* operon, the knockout efficiency was only 6.4% (Jeong et al., 2015). Recently, a new editing system based on CRISPR-Cas9 employing a single sgRNA was developed to delete *pps* operon (So et al., 2017). However, the *pps* operon was not completely deleted even though the efficiency of deletion was further improved by optimizing. We infer that the DSB site and the repair site of homologous arm are too far to initiate an efficient double cross-over.

In view of those results that CIGE-based CRISPR-Cpf1 system was highly functioned to insertion of gene in chromosome, we considered that whether the system was also suitable to deletion of large gene clusters in *B. subtilis*. To verify the deduction, we first evaluated the function of deletion of *sacA* (**Figure 5A**). Plasmid pBsacA (refer to addgene #158650), containing *sacA*-targeted crRNA and the corresponding homologous arms identical to that of AIO and TPs (refer to pHTsacA), was constructed to implement the function. Then, pBsacA was introduced into BS-*ganA*'-Cpf1. We randomly selected 20 colonies from the agar plate and perform cPCR to identify the deletion efficiency. The results demonstrated that all the colonies harbored the disrupted *sacA* smaller than that of the wild-type *sacA* (ck), suggesting that the genome editing efficiencies mediated by CIGE system was 100% (20/20) (**Figure 5B**, upper panel). Sequencing results also authenticated that *sacA* has been disrupted by 240 bp in the CIGE edited hosts (**Figure 5B**, lower panel).

After validating that *sacA* was deleted by the engineering CRISPR-Cpf1 system, we sought to further explore whether the system can be used to delete larger gene clusters. The *bac* operon was selected as the deletion target. Firstly, we constructed the targeting plasmid, pBbac, containing expression cassette composed of a pair of 500-bp homologous arms and a *bacD*-targeted crRNA under the control of P_{veg}. After transformation in BS-*ganA*'-Cpf1, the transformants harboring pBbac was cultured and screened in line with the procedure above. Since the designed primers for verification flanked the homologous arms, the product of cPCR should be approximate of 1320 bp rather than 8008 bp (the full-length of *bac* operon) if *bac* operon was successfully deleted. We randomly selected 10 colonies from the agar plate and performed cPCR to identify the deletion efficiency. Interestingly, the cPCR product from the 10 colonies had expected gene size while the control (ck) hadn't, indicating that *bacD* was successfully deleted from the operon (**Figure 5C**). In parallel, we sequenced the cPCR product from one of the successfully deleted mutants and confirmed the deletion (**Figure 5C**).

To testify whether the tool functioned to a broad range of targets, we selected another target, the *pps* operon, to evaluate the deletion efficiency. We firstly verified that the *pps* operon exists in the *B. subtilis* 168 prior to performing the deletion of the *pps* operon (**Supplementary Figure S1**). Then, we replaced *bacD*-targeting crRNA with *ppsC*-targeting crRNA on plasmid pBbac, and extended the length of homologous arms to 800 bp to ensure the deletion efficiency. Plasmid used to delete *pps* was termed pBpps. Then, pBpps was transformed into BS-*ganA*'-Cpf1. The screening and verification procedure were identical to that of *bacD* deletion. The cPCR product from *ppsC*-deleted mutant should be of 1772 bp by the designed primers. The results of cPCR showed that 8 out of 10 colonies had the products of expected size, suggesting that *ppsC* is deleted in these 8 mutants (**Figure 5D**). These results were further confirmed by sequencing (**Figure 5D**). The deletion efficiency for *bac* and *pps* was of 100% and 80%, respectively, based on the CRISPR-Cpf1-dependent CIGE platform (**Figures 5C,D**).

These results manifest that our designed CRISPR-Cpf1 system (CIGE) is more reliable and portable than the CRISPR-Cas9 system, especially for deletion of large gene clusters. To our best knowledge, this is the most efficient system to delete a *pps* operon with a single crRNA (**Figure 5D**). Deletion of large gene fragments is very important in synthetic biology, especially in deleting non-essential genes to construct minimal genomic microorganisms. Therefore, our customized CRISPR-Cpf1 system has great potential to implement this.

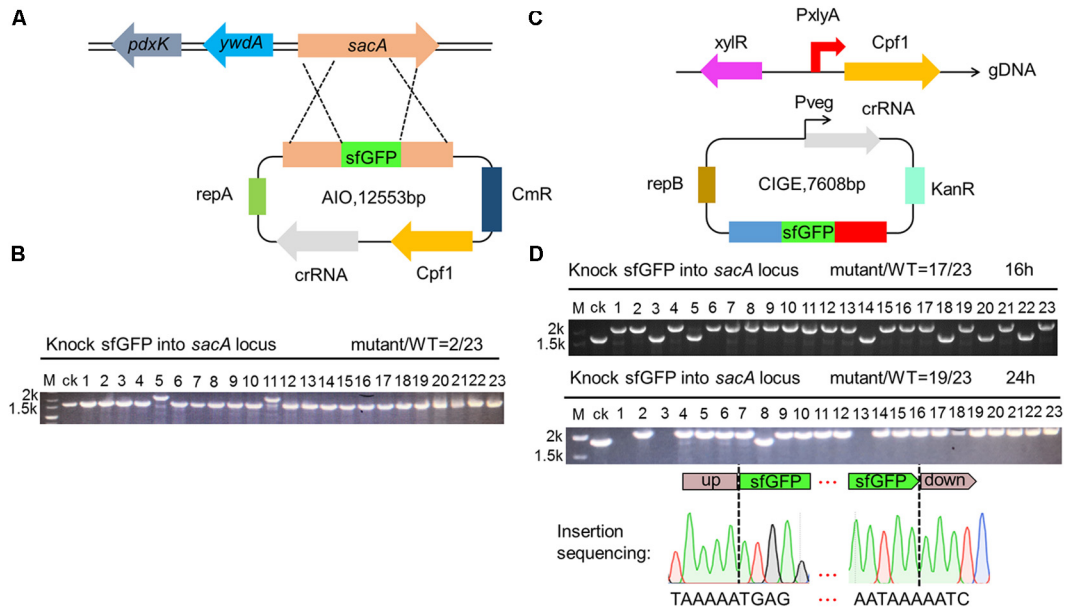


FIGURE 4 | Construction of superfolder green fluorescence protein (sfGFP) insertion in *sacA*. **(A)** Scheme showing the procedures for gene insertion into *Bacillus subtilis* 168 by AIO system. **(B)** cPCR results shown that 9% (lane 5 and lane 11) of colonies had the sfGFP insertion mutant. **(C)** Scheme showing the procedures for gene insertion into *B. subtilis* 168 by CCB-CIGE system. **(D)** CIGE enables highly efficient sfGFP insertion mutation in the *B. subtilis* 168 strain. The *sacA* gene was replaced by the *sfgfp* gene. The efficiency for *sfgfp* gene insertion was 17/23 in the *B. subtilis* 168 strain. Prolonged incubation time under selective pressure increased the mutation efficiency to 82%.

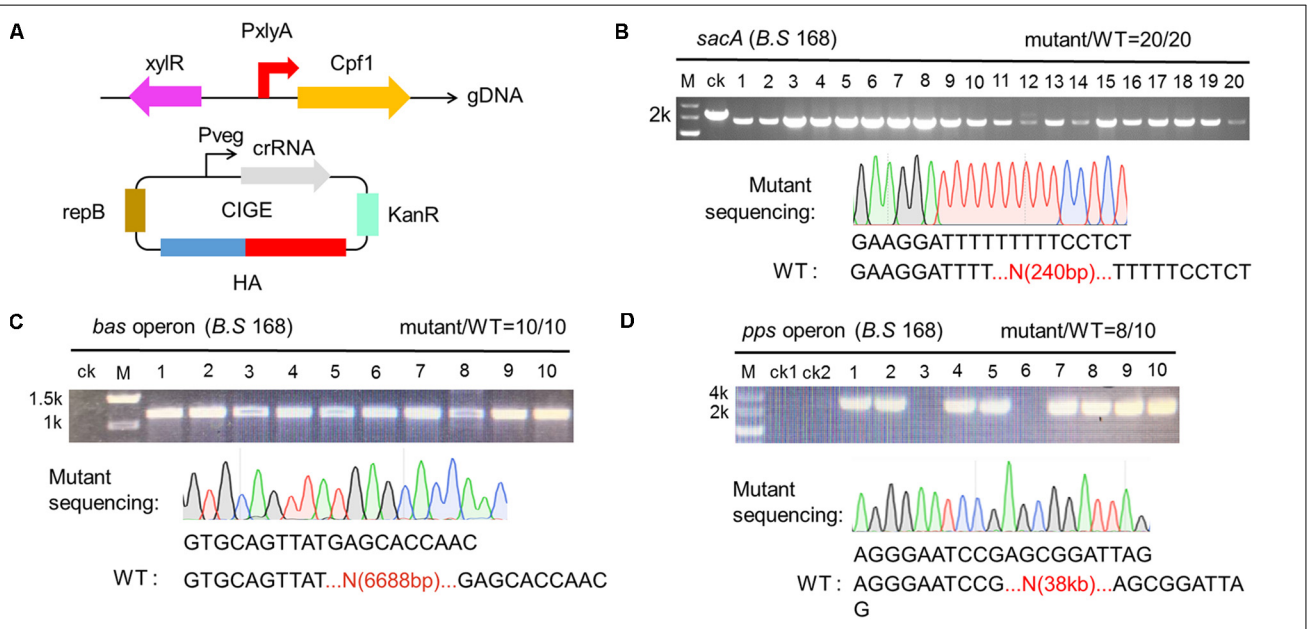


FIGURE 5 | Deletion of larger gene clusters mediated by CCB-CIGE system derived from CRISPR-Cpf1. **(A)** Scheme showing the composition of the CIGE system for the deletion of larger gene clusters. **(B)** CIGE-mediated disruption of the *sacA* gene in the *Bacillus subtilis* 168 strain. The editing efficiency was 20/20. **(C)** Identification of *bac* operon deletion mutants. *bac* operon is composed of seven genes (*bac*ABCDEFG), about 6688 bp. Ten transformants were selected and verified by colony PCR. M and ck represent the 2-kb DNA ladder and wild-type *B. subtilis* 168, respectively. **(D)** Confirmation of *pps* operon deletion using colony PCR. Lane M, the 5-kb DNA marker from Takara. Lane ck1 and ck2, PCR amplification with the *B. subtilis* 168 and *B. subtilis* comK genomic DNA (gDNA) using external primers, respectively. Lane 1-10, PCR amplification with the mutant gDNA using external primer. The editing efficiency was 8/10.

CCB-CIGE Platform Is Upgraded to a Higher Version to Exert Multiplex-Gene Editing

Wild-type bacteria usually are unable to exert programmed functions preset in cell factories. Therefore, engineering natural bacteria is a practical strategy to design a higher version of microbial chassis for synthetic biology. Precise and portable multiplex gene editing is an absolutely indispensable approach to do this. However, some commonly used methods, Cre/*lox*P (Suzuki et al., 2005; Choi et al., 2015) and *Red* recombination system (Datsenko and Wanner, 2000), were difficult to efficiently edit at genome-scale because of some technical hurdles. First of all, Cre/*lox*P system is not a scarless knockout technology, which leaves loxPLR sites at the edited position (Suzuki et al., 2007). These scars may affect the growth of host cells, and the cycle of this technique is longer. Secondly, the efficiency of this gene deletion method (Cre/*lox*P) is not very efficient because two rounds of homologous recombination are required and mutant selection after the second recombination is time-consuming (Cleto et al., 2016). Thirdly, it is difficult to achieve simultaneous editing of multiple targets in the genome editing system based on Cre/*lox*P. For genome editing in prokaryotes, phage-derived lambda red recombinases have been employed in recombineering, which facilitates homology-dependent integration/replacement of a donor DNA or oligonucleotide. However, these systems require the genetic background of the target strains such as deficiency of methyl-directed mismatch repair or *RecA* that involves in the recombinational DNA repair system (Wang et al., 2009). Multiplex genome editing systems based on CRISPR-Cas9 have been applied to different organisms, including *E. coli* (Zerbini et al., 2017), *B. subtilis* (Westbrook et al., 2016), *Streptomyces* species (Cobb et al., 2015), *Rhodosporidium toruloides* (Otoupal et al., 2019), and mammalian (Cong et al., 2013).

To fully exploit the function of CRISPR-Cpf1 in multiplex genome editing, we employed two genes used above, *sacA* and *aprE*, to design and build a multiple gene editing system. Firstly, we incorporated *sacA*- and *aprE*- crRNA cassettes into pBSG, by which the two crRNAs can simultaneously target *sacA* and *aprE*. In this form, we found when the crRNAs were transformed into BS-*ganA*'-Cpf1, only one gene was deleted in the same transformant (data not shown). Previous study reported that *B. subtilis* has more complex recombination systems and diverse plasmid replication modes. Therefore, we inferred that the failure might be due to the exchange of some fragments of the plasmid in the process of replication. After sequencing the transformants with single gene deletion, we confirmed that the sacA crRNA expression cassette was lost from plasmid, resulting in failure of deleting *sacA* (data not shown).

According to previous studies, Cpf1 intrinsically processes pre-crRNA into mature crRNA by cleaving specific site (Fonfara et al., 2016; Zetsche et al., 2017). Thus, we inferred that it might be feasible to implement multiplex genome editing by integrating multiple crRNAs to single expression cassette. Two crRNAs targeting *sacA* and *aprE* were genetically fused and insulated them by synthetic two repeats of "Direct Repeat (DR)-Spacer" units. The homologous arms of *sacA* and *aprE* tandem were sequentially cloned downstream of these crRNAs, generating a complete targeting sequence. P_{veg} was equipped to trigger the expression cassette. The expression cassette was then cloned into pBSG, yielding pB-PvegW-SAKo (**Figure 6A**). pB-PvegW-SAKo was transformed into BS-*ganA*'-Cpf1. Through the cultivation and screening as described above, cPCR results showed that 27.2% (6/22) of colonies were *sacA*- and *aprE*-deficient (**Figure 6B**). Although we successfully achieved double-gene deletion by CCB-CIGE platform, the efficiency still needs to be improved so as to elevate the work performance.

Accordingly, the next thrust is to engineer the components on this platform. Because the wild type P_{veg} promoter is a super strong promoter, we deduced that excessive activity of P_{veg} in the cell might induce growth burden in *B. subtilis*, rendering low editing efficiency through disturbing expression of crRNAs. Therefore, to rebalance the expression, we mutated the -10 region of the wild-type P_{veg} promoter to decrease its activity, variant was termed P_{vegM} (**Figure 6A**). By substitution for the wild-type P_{veg}, pB-PvegM-SAKo was constructed, and then transformed into BS-*ganA*'-Cpf1. We randomly chose 12 colonies to verify the deletion by cPCR. The results showed that 7 selected clones among the 12 clones harbored the deleted *sacA* and *aprE* genome as the bands were smaller than that of the control "ck" (**Figure 6C**). The double-deficient mutants accounted for 58.3% of total tested mutants (**Figure 6C**), which was higher than that of previously reported by two folds (**Figures 6B,C**). These results suggest that the optimized CRISPR-Cpf1 system has great potential to achieve multiplex-gene editing at genome-scale.

Up to date, CRISPR-Cas9 system has been widely used in different organisms for genome editing. However, the function relies on a complex composed of crRNA and tracrRNA (or a chimera gRNA), which guides Cas9 to the target in genome. In contrast, one single crRNA is sufficient to guide CRISPR-Cpf1 RNP to a gene target. Practically, CRISPR-Cas9 system mediates multiplex genome editing by expression of multiple gRNAs. Nevertheless, the functional construct is often relatively large and complex, which would be rather difficult to construct and transform plasmid. Compared with Cas9 nucleases, the most significant feature of Cpf1 nucleases is that they not only have DNase activity but also RNase activity, which gives them the ability to process their own crRNA from a long precursor. This feature greatly facilitates their application for multiplex genome editing, transcriptional regulation and imaging, which tasks typically need to locate multiple loci in the genome for efficient operation. By taking advantage of this feature, we successfully constructed the multiplex genome editing system by using a single crRNA array in one vector (**Figure 6A**). The potential advantage of this system is that it can delete those genes that are difficult to delete individually. Another consideration is that the transformation efficiency of multiplex-gene-targeting plasmid is significantly lower than that of single-gene-targeting plasmid. It might also be due to the strong cleavage effect of crRNA by the constitutive transcription of the P_{veg} promoter, which makes it difficult for bacteria to repair in time. Therefore, inducible promoters can be

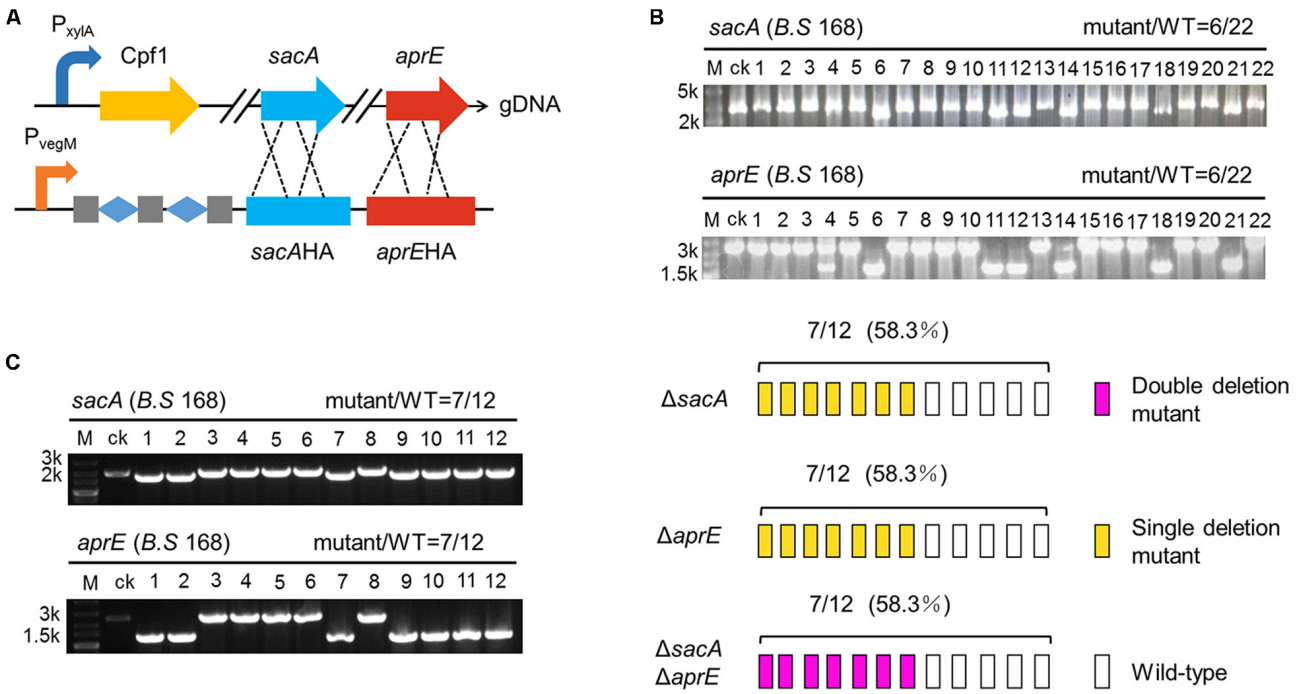


FIGURE 6 | CRISPR-Cpf1-assisted simultaneous deletion of double genes (ΔsacAΔaprE). **(A)** The construction of the Cpf1 multiplex gene editing system contains a Cpf1 expression cassette and a multi-crRNA expression cassette. Cpf1 was inserted *ganA* site of genome DNA (gDNA) of *Bacillus subtilis* 168. The multi-crRNA expression cassette contains three DR guide units, and each unit includes one mature DR and 23 bp of guide sequence. **(B)** Multi-crRNA expression cassette is controlled by the P_{veg} promoter. Twenty-two colonies were picked and screened for mutations. The efficiency of simultaneous deletion of double genes is 6/22. **(C)** Multi-crRNA expression cassette is controlled by the P_{veg} promoter variant P_{vegM}. The colonies of 7/12 is identified as a double deletion mutant strain (ΔsacAΔaprE). The open rectangle, orange rectangle, and dark pink rectangle represent the WT strain, the single deletion mutant of ΔsacA or ΔaprE, and the ΔsacAΔaprE double deletion mutant, respectively.

used to regulate the expression of crRNA to enrich biomass in later studies.

Up to now, we have constructed a set of CRISPR-Cpf1-based toolkit (including AIO, TP, and CCB-CIGE) in *B. subtilis*. Zhang et al., Westbrook et al., and Wu et al., have developed genome editing tools based on CRISPR-Cas9 and CRISPR-Cpf1 in previous studies. However, compared with their systems, our newly constructed system has some advantages.

Firstly, although Zhang et al., constructed a genome editing system based on CRISPR-Cas9 in *B. subtilis* ATCC 6051a, they only verified that CRISPR-Cas9 system can disrupt gene in *B. subtilis* ATCC 6051a, and did not testify that the system can accurately delete and insert gene in the frame. In this work, we constructed a genome editing system based on CRISPR-Cpf1 with the same strategy (AIO). In this system, we used strong promoters P43 and P_{veg} from *B. subtilis* to express Cpf1 and crRNA, respectively. And we also verified that our engineered AIO system is capable of knocking out fragments of different sizes (240 bp, 2064 bp and 2775 bp) with efficiency over 95% (**Figures 2B–D**). Simultaneously, we construct two AIO systems (based on CRISPR-Cpf1 and CRISPR-Cas9) to prove that CRISPR-Cpf1 system is superior to CRISPR-Cas9 system in *B. subtilis* (take gene *ganA* as an example, 100% of Cpf1 vs. 75% of Cas9, **Figure 2B**). However, the AIO system we constructed in this study is relatively larger, which can make the construction and transformation of plasmid a little more difficult.

Secondly, in the study of Westbrook et al. (2016) they constructed a tool kit based on CRISPR-Cas9 in *B. subtilis*. In their whole research, they all adopt chromosome integration strategy to achieve different functions (including knockout, knock-in and transcription interference) based on CRISPR-Cas9. In this way, foreign genes (for example, Cas9, gRNA, and resistance marker) will inevitably be introduced into the genome of *B. subtilis*. However, as an important platform for metabolic engineering and synthetic biology research, the effect of exogenous gene introduction on the expected function of *B. subtilis* is unknown. Concurrently, insertion of antibiotic resistance marker will disqualify the use of engineered *Bacillus* strains as an eco-friendly host for food-grade applications, industrial fermentation and bioremediation. Westbrook et al. (2016) have also constructed a multi-gRNA delivery vector for multiplex genome editing. However, it is relatively difficult and time-consuming to insert multiple repeat sequences into one vector at the same time. In this study, we constructed a toolkit based on CRISPR-Cpf1, which can be used in different scenarios. AIO and TP systems can be used for traceless knockout, and are suitable for use without introducing foreign gene (for example, fermentation of food-grade enzymes). CCB-CIGE system can be used for engineering modification of *B. subtilis* to produce

industrial enzyme. Recently, Wu et al., developed a toolkit (including knockout, knock-in, point mutation and transcription interference and activation) based on CRISPR-Cpf1 in *B. subtilis*. However, they did not explore the editing efficiency of a single gene of CRISPR-Cpf1 system in *B. subtilis*. It is unclear whether the editing efficiency varied in deletion of different size of the fragment. In our study, the AIO system we developed can knock-out small and medium-sized fragment, and the knock-out efficiency is almost not affected (efficiency is between 95% and 100%, **Figures 2B–D**). And in previous studies, it has been shown that the knockout of non-essential large gene clusters is indispensable for the construction of the minimal genomic chassis microorganism. However, in the study of Wu et al., deletion of large gene clusters was not validated. Here, we developed an upgraded CRISPR-Cpf1 system, which can knock-out *bac* (efficiency is 100%) and *pps* operons (efficiency is 80%, **Figures 5C,D**). And as far as we know, this is the first example to knock out 38-kb gene cluster with CRISPR-Cpf1 system in *B. subtilis*. Next, we also try to use CCB-CIGE platform to delete extremely large cluster *pks* operon (about 78 kb). However, we only obtain lower deletion efficiency (about 16.7%, data not show). Although, we still believe that the optimized CCB-CIGE platform can continue to improve the knock-out efficiency of *pks* operon. For double gene editing, Wu et al., used a double plasmid expression system, and enhanced the efficiency of homologous recombination by overexpression of mutated NgAgo*. However, they need to add two inducers to achieve the goal of knocking out two genes, which is more laborious. Moreover, the transformation efficiency of the double plasmid system is significantly lower than that of the CCB-CIGE system in our study. Similarly, in the research of Jiang et al., they also overexpressed the recombinant factor *recT* to improve the efficiency of recombination. However, in our study, Cpf1 is integrated into the chromosome, the decrease of Cpf1 expression may weaken the knockout efficiency. And we did not introduce the factor NgAgo* and recT, to enhance homologous recombination, which will also reduce the editing efficiency. It is necessary to continue to improve the efficiency of homologous recombination and multi-target editing of our system.

Overall, we identified that engineering CRISPR-Cpf1 system significantly facilitates gene manipulation in *B. subtilis*. The development of this system has accelerated the construction of high-version microbial chassis. We believe that the optimized CRISPR-Cpf1 system can also be applied to other Gram-positive bacteria, such as *Bacillus thuringiensis* and *Bacillus licheniformis*.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

WH and WC conceived the project and designed the experiments. FS performed the molecular cloning experiments. QL and QC performed the colony PCR. WH, WC, and ZZ analyzed the data and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was financially supported by a project funded by the International S&T Innovation Cooperation Key Project (2017YFE0129600), the National Natural Science Foundation of China (21878125), the Natural Science Foundation of Jiangsu Province (BK20181206), the Priority Academic Program Development of Jiangsu Higher Education Institutions, the 111 Project (No. 111-2-06), the Jiangsu Province "Collaborative Innovation Center for Advanced Industrial Fermentation" Industry Development Program, and First-Class Discipline Program of Light Industry Technology and Engineering (LITE2018-04).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.524676/full#supplementary-material

REFERENCES

Anagnostopoulos, C., and Spizizen, J. (1961). Requirements for transformation in bacillus subtilis. *J. Bacteriol.* 81, 741–746.

Bai, H., Deng, A., Liu, S., Cui, D., Qiu, Q., Wang, L., et al. (2018). A novel tool for microbial genome editing using the restriction-modification system. *ACS Synth. Biol.* 7, 98–106. doi: 10.1021/acssynbio.7b00254

Bao, Z., Xiao, H., Liang, J., Zhang, L., Xiong, X., Sun, N., et al. (2015). Homology-integrated CRISPR-Cas (HI-CRISPR) system for one-step multigene disruption in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* 4, 585–594. doi: 10.1021/sb500255k

Buchmuller, B. C., Herbst, K., Meurer, M., Kirrmaier, D., Sass, E., Levy, E. D., et al. (2019). Pooled clone collections by multiplexed CRISPR-Cas12a-assisted gene tagging in yeast. *Nat. Commun.* 10:2960. doi: 10.1038/s41467-019-10816-7

Chen, W., Zhang, Y., Yeo, W. S., Bae, T., and Ji, Q. (2017). Rapid and efficient genome editing in *Staphylococcus aureus* by using an engineered CRISPR/Cas9 system. *J. Am. Chem. Soc.* 139, 3790–3795. doi: 10.1021/jacs.6b13317

Choi, J. W., Yim, S. S., Kim, M. J., and Jeong, K. J. (2015). Enhanced production of recombinant proteins with *Corynebacterium glutamicum* by deletion of insertion sequences (IS elements). *Microb. Cell Fact.* 14:207. doi: 10.1186/s12934-015-0401-7

Cleto, S., Jensen, J. V., Wendisch, V. F., and Lu, T. (2016). *Corynebacterium glutamicum* metabolic engineering with CRISPR Interference (CRISPRi). *ACS Synth. Biol.* 5, 375–385. doi: 10.1021/acssynbio.5b00216

Cobb, R. E., Wang, Y., and Zhao, H. (2015). High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS Synth. Biol.* 4, 723–728. doi: 10.1021/sb500351f

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823. doi: 10.1126/science.1231143

Correa, G. G., Rachel da Costa Ribeiro Lins, M., Silva, B. F., Barbosa, de Paiva, G., Bertolazzi Zocca, V. F., et al. (2020). A modular autoinduction device for control of gene expression in *Bacillus subtilis*. *Metab. Eng.* 61, 326–334. doi: 10.1016/j.ymben.2020.03.012

Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6640–6645. doi: 10.1073/pnas.120163297

Dong, H., and Zhang, D. (2014). Current development in genetic engineering strategies of *Bacillus* species. *Microb. Cell Fact.* 13, 63. doi: 10.1186/1475-2859-13-63

Engler, C., Gruetzner, R., Kandzia, R., and Marillonnet, S. (2009). Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS One* 4:e5553. doi: 10.1371/journal.pone.0005553

Fonfara, I., Richter, H., Bratovic, M., Le Rhun, A., and Charpentier, E. (2016). The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* 532, 517–521. doi: 10.1038/nature17945

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., et al. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 31, 822–826. doi: 10.1038/nbt.2623

Gao, W., Long, L., Tian, X., Xu, F., Liu, J., Singh, P. K., et al. (2017). Genome editing in cotton with the CRISPR/Cas9 system. *Front. Plant Sci.* 8:1364. doi: 10.3389/fpls.2017.01364

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A. III, and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. doi: 10.1038/nmeth.1318

Hong, W., Zhang, J., Cui, G., Wang, L., and Wang, Y. (2018). Multiplexed CRISPR-Cpf1-Mediated Genome Editing in *Clostridium difficile* toward the Understanding of Pathogenesis of *C. difficile* Infection. *ACS Synth. Biol.* 7, 1588–1600. doi: 10.1021/acssynbio.8b00087

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832. doi: 10.1038/nbt.2647

Hur, J. K., Kim, K., Been, K. W., Baek, G., Ye, S., Hur, J. W., et al. (2016). Targeted mutagenesis in mice by electroporation of Cpf1 ribonucleoproteins. *Nat. Biotechnol.* 34:807. doi: 10.1038/nbt.3596

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., et al. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 227–229. doi: 10.1038/nbt.2501

Jeong, D. E., Park, S. H., Pan, J. G., Kim, E. J., and Choi, S. K. (2015). Genome engineering using a synthetic gene circuit in *Bacillus subtilis*. *Nucleic Acids Res.* 43:e42. doi: 10.1093/nar/gku1380

Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., and Yang, S. (2015). Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl. Environ. Microbiol.* 81, 2506–2514. doi: 10.1128/aem.04023-14

Jiang, Y., Qian, F., Yang, J., Liu, Y., Dong, F., Xu, C., et al. (2017). CRISPR-Cpf1 assisted genome editing of *Corynebacterium glutamicum*. *Nat. Commun.* 8:15179. doi: 10.1038/ncomms15179

Keasling, J. D. (2010). Manufacturing molecules through metabolic engineering. *Science* 330, 1355–1358. doi: 10.1126/science.1193990

Kim, D., Kim, J., Hur, J. K., Been, K. W., Yoon, S. H., and Kim, J. S. (2016). Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat. Biotechnol.* 34, 863–868. doi: 10.1038/nbt.3609

Kim, H., Kim, S. T., Ryu, J., Kang, B. C., Kim, J. S., and Kim, S. G. (2017). CRISPR/Cpf1-mediated DNA-free plant genome editing. *Nat. Commun.* 8:4406. doi: 10.1038/ncomms14406

Liu, J. J., Orlova, N., Oakes, B. L., Ma, E., Spinner, H. B., Baney, K. L. M., et al. (2019). CasX enzymes comprise a distinct family of RNA-guided genome editors. *Nature* 566, 218–223. doi: 10.1038/s41586-019-0908-x

Otoupal, P. B., Ito, M., Arkin, A. P., Magnuson, J. K., Gladden, J. M., and Skerker, J. M. (2019). Multiplexed CRISPR-Cas9-Based genome editing of *Rhodosporidium toruloides*. *mSphere* 4:e00099-19. doi: 10.1128/mSphere.00099-19

Ozcengiz, G., and Ogulur, I. (2015). Biochemistry, genetics and regulation of bacilysin biosynthesis and its significance more than an antibiotic. *N. Biotechnol.* 32, 612–619. doi: 10.1016/j.nbt.2015.01.006

Pitcher, R. S., Brissett, N. C., and Doherty, A. J. (2007). Nonhomologous end-joining in bacteria: a microbial perspective. *Annu. Rev. Microbiol.* 61, 259–282. doi: 10.1146/annurev.micro.61.080706.093354

Radeck, J., Kraft, K., Bartels, J., Cikovic, T., Durr, F., Emenegger, J., et al. (2013). The *Bacillus* BioBrick Box: generation and evaluation of essential genetic building blocks for standardized work with *Bacillus subtilis*. *J. Biol. Eng.* 7:29. doi: 10.1186/1754-1611-7-29

So, Y., Park, S. Y., Park, E. H., Park, S. H., Kim, E. J., Pan, J. G., et al. (2017). A highly efficient CRISPR-Cas9-mediated large genomic deletion in *Bacillus subtilis*. *Front. Microbiol.* 8:1167. doi: 10.3389/fmicb.2017.01167

Strecker, J., Jones, S., Koopal, B., Schmid-Burgk, J., Zetsche, B., Gao, L., et al. (2019). Engineering of CRISPR-Cas12b for human genome editing. *Nat. Commun.* 10:212. doi: 10.1038/s41467-018-08224-4

Suzuki, N., Inui, M., and Yukawa, H. (2007). Site-directed integration system using a combination of mutant lox sites for *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 77, 871–878. doi: 10.1007/s00253-007-1215-2

Suzuki, N., Tsuge, Y., Inui, M., and Yukawa, H. (2005). Cre/loxP-mediated deletion system for large genome rearrangements in *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 67, 225–233. doi: 10.1007/s00253-004-1772-6

Teng, F., Cui, T., Feng, G., Guo, L., Xu, K., Gao, Q., et al. (2018). Repurposing CRISPR-Cas12b for mammalian genome engineering. *Cell Discov.* 4:63. doi: 10.1038/s41421-018-0069-3

Thwaite, J. E., Baillie, L. W., Carter, N. M., Stephenson, K., Rees, M., Harwood, C. R., et al. (2002). Optimization of the cell wall microenvironment allows increased production of recombinant *Bacillus anthracis* protective antigen from *B. subtilis*. *Appl. Environ. Microbiol.* 68, 227–234. doi: 10.1128/aem.68.1.227-234.2002

Trevors, J. T. (1986). Plasmid curing in bacteria. *FEMS Microbiol. Rev.* 32, 149–157. doi: 10.1111/j.1574-6968.1986.tb01189.x

Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., and Xu, G. (2009). Programming cells by multiplex genome engineering and accelerated evolution. *Nature* 460, 894–898. doi: 10.1038/nature08187

Wang, M., Mao, Y., Lu, Y., Tao, X., and Zhu, J. K. (2017). Multiplex gene editing in rice using the CRISPR-Cpf1 system. *Mol. Plant* 10, 1011–1013. doi: 10.1016/j.molp.2017.03.001

Westbrook, A. W., Moo-Young, M., and Chou, C. P. (2016). Development of a CRISPR-Cas9 tool kit for comprehensive engineering of *Bacillus subtilis*. *Appl. Environ. Microbiol.* 82, 4876–4895. doi: 10.1128/aem.01159-16

Westers, H., Dorenbos, R., van Dijl, J. M., Kabel, J., Flanagan, T., Devine, K. M., et al. (2003). Genome engineering reveals large dispensable regions in *Bacillus subtilis*. *Mol. Biol. Evol.* 20, 2076–2090. doi: 10.1093/molbev/msg219

Wu, Y., Chen, T., Liu, Y., Lv, X., Li, J., Du, G., et al. (2018). CRISPRi allows optimal temporal control of N-acetylglucosamine bioproduction by a dynamic coordination of glucose and xylose metabolism in *Bacillus subtilis*. *Metab. Eng.* 49, 232–241. doi: 10.1016/j.ymben.2018.08.012

Wu, Y., Liu, Y., Lv, X., Li, J., Du, G., and Liu, L. (2020). CAMERS-B: CRISPR/Cpf1 assisted multiple-genes editing and regulation system for *Bacillus subtilis*. *Biotechnol. Bioeng.* 117, 1817–1825. doi: 10.1002/bit.27322

Yamashiro, D., Minouchi, Y., and Ashiuchi, M. (2011). Moonlighting role of a poly-gamma-glutamate synthetase component from *Bacillus subtilis*: insight into novel extrachromosomal DNA maintenance. *Appl. Environ. Microbiol.* 77, 2796–2798. doi: 10.1128/aem.02649-10

Yan, M. Y., Yan, H. Q., Ren, G. X., Zhao, J. P., Guo, X. P., and Sun, Y. C. (2017). CRISPR-Cas12a-assisted recombineering in bacteria. *Appl. Environ. Microbiol.* 83:e00947-17. doi: 10.1128/aem.00947-17

Zerbini, F., Zanella, I., Fraccascia, D., Konig, E., Irene, C., Frattini, L. F., et al. (2017). Large scale validation of an efficient CRISPR/Cas-based multi gene editing protocol in *Escherichia coli*. *Microb. Cell Fact.* 16:68. doi: 10.1186/s12934-017-0681-1

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163, 759–771. doi: 10.1016/j.cell.2015.09.038

Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E. M., et al. (2017). Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat. Biotechnol.* 35, 31–34. doi: 10.1038/nbt.3737

Zhang, K., Duan, X., and Wu, J. (2016). Multigene disruption in undomesticated *Bacillus subtilis* ATCC 6051a using the CRISPR/Cas9 system. *Sci. Rep.* 6:27943. doi: 10.1038/srep27943

Zhang, X. Z., Yan, X., Cui, Z. L., Hong, Q., and Li, S. P. (2006). mazF, a novel counter-selectable marker for unmarked chromosomal manipulation in *Bacillus subtilis*. *Nucleic Acids Res.* 34:e71. doi: 10.1093/nar/gkl358

Zhu, B., and Stulke, J. (2018). SubtiWiki in 2018: from genes and proteins to functional network annotation of the model organism *Bacillus subtilis*. *Nucleic Acids Res.* 46, D743–D748. doi: 10.1093/nar/gkx908

Zobel, S., Kumpfmuller, J., Sussmuth, R. D., and Schweder, T. (2015). *Bacillus subtilis* as heterologous host for the secretory production of the non-ribosomal cyclodepsipeptide enniatin. *Appl. Microbiol. Biotechnol.* 99, 681–691. doi: 10.1007/s00253-014-6199-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 27 October 2020
doi: 10.3389/fbioe.2020.560726

The Metabolism of *Clostridium ljungdahlii* in Phosphotransacetylase Negative Strains and Development of an Ethanologenic Strain

Jonathan Lo[1], Jonathan R. Humphreys[1], Joshua Jack[2], Chris Urban[1], Lauren Magnusson[1], Wei Xiong[1], Yang Gu[3], Zhiyong Jason Ren[2] and Pin-Ching Maness[1]*

[1] National Renewable Energy Laboratory, Golden, CO, United States, [2] Andlinger Center for Energy and Environment, Princeton University, Princeton, NJ, United States, [3] Key Laboratory of Synthetic Biology, CAS Center for Excellence in Molecular Plant Sciences, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China

OPEN ACCESS

Edited by:
Yi Wang,
Auburn University, United States

Reviewed by:
Volker Müller,
Goethe University Frankfurt, Germany
Largus T. Angenent,
University of Tübingen, Germany
Michael Köpke,
LanzaTech, United States
Mirko Basen,
Universität Rostock, Germany

***Correspondence:**
Jonathan Lo
Jonathan.Lo@NREL.gov

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *10 May 2020*
Accepted: *05 October 2020*
Published: *27 October 2020*

Citation:
Lo J, Humphreys JR, Jack J, Urban C, Magnusson L, Xiong W, Gu Y, Ren ZJ and Maness P-C (2020) The Metabolism of Clostridium ljungdahlii in Phosphotransacetylase Negative Strains and Development of an Ethanologenic Strain. Front. Bioeng. Biotechnol. 8:560726. doi: 10.3389/fbioe.2020.560726

The sustainable production of chemicals from non-petrochemical sources is one of the greatest challenges of our time. CO_2 release from industrial activity is not environmentally friendly yet provides an inexpensive feedstock for chemical production. One means of addressing this problem is using acetogenic bacteria to produce chemicals from CO_2, waste streams, or renewable resources. Acetogens are attractive hosts for chemical production for many reasons: they can utilize a variety of feedstocks that are renewable or currently waste streams, can capture waste carbon sources and covert them to products, and can produce a variety of chemicals with greater carbon efficiency over traditional fermentation technologies. Here we investigated the metabolism of *Clostridium ljungdahlii*, a model acetogen, to probe carbon and electron partitioning and understand what mechanisms drive product formation in this organism. We utilized CRISPR/Cas9 and an inducible riboswitch to target enzymes involved in fermentation product formation. We focused on the genes encoding phosphotransacetylase (*pta*), aldehyde ferredoxin oxidoreductases (*aor1* and *aor2*), and bifunctional alcohol/aldehyde dehydrogenases (*adhE1* and *adhE2*) and performed growth studies under a variety of conditions to probe the role of those enzymes in the metabolism. Finally, we demonstrated a switch from acetogenic to ethanologenic metabolism by these manipulations, providing an engineered bacterium with greater application potential in biorefinery industry.

Keywords: syngas, acetogen, metabolic engineering, CO_2 fixation, *Clostridium ljungdahlii*, autotrophic

INTRODUCTION

Acetogenic bacteria utilize the Wood-Ljungdahl pathway (WLP) to non-photosynthetically fix inorganic carbon into acetyl-CoA, which is then converted into products, normally acetate and to a lesser extent, ethanol. Acetate production is an important feature of acetogenic metabolism. In *Clostridium ljungdahlii*, a model acetogen, acetate formation primarily occurs sequentially with the conversion of acetyl-CoA to acetyl-P via phosphotransacetylase (PTA) and then from

acetyl-P to acetate via acetate kinase (ACK), with the ACK step generating ATP (Köpke et al., 2010). Acetate generation is important for ATP formation, as acetogens have low ATP budgets because they survive on the "thermodynamic limit of life", on a small free energy change (Schuchmann and Müller, 2014).

Based on genome information and published data, *C. ljungdahlii* is believed to make ethanol from acetyl-CoA via two established mechanisms: aldehyde ferredoxin oxidoreductase (AOR) and bifunctional aldehyde/alcohol dehydrogenase (AdhE) (Köpke et al., 2010; Leang et al., 2013). AOR catalyzes the reversible conversion of acetate to acetaldehyde, with ferredoxin as the electron acceptor/donor. AdhE is a bifunctional enzyme, catalyzing the reversible conversion of acetyl-CoA to acetaldehyde, then to ethanol, using NAD(P)H as the electron donors.

Acetogens have been studied for their ability to generate value-added products from either syngas alone or in conjunction with sugar for increased carbon yield (Köpke et al., 2010; Jones et al., 2016). Many of these value-added products are derived from acetyl-CoA, which is generated from the WLP and glycolysis and serves as a primary intermediate in central carbon metabolism. Thus, it would be useful to understand the metabolism around this acetyl-CoA node to help guide metabolic engineering efforts.

Pta is the most obvious target, as it is the first and direct step to making acetate from acetyl-CoA. *C. ljungdahlii pta* (CLJU_c12770) has been targeted in three studies via gene disruption, CRISPR interference, and CRISPR/Cas9, but in all cases it was ambiguous what was happening. In the gene disruption study of *pta*, *pta* was replaced with a butyrate pathway that could have side reactions to generate acetate (Ueki et al., 2014). In the CRISPR interference study, knockdown of *pta* was incomplete and did not show an obvious growth phenotype leaving open the possibility that leaky expression was responsible for acetate production (Woolston et al., 2018). In the CRISPR/Cas9 study, the *pta* deletion strain growth data was only characterized vs. the wild type on syngas, showing no significant growth after 96 h or characterization of heterotrophic growth (Huang et al., 2016). High yields of acetate remained in strains targeting *pta*, suggesting that alternative pathways may be important for acetate production. Studies of acetyl-CoA metabolism suggest that there may be multiple important enzymes and pathways utilizing acetyl-CoA (Ueki et al., 2014; Whitham et al., 2015; Huang et al., 2016). Thus, we set out to delete *pta* and other key genes to understand their roles in metabolism and product formation.

MATERIALS AND METHODS

Microbial Strains and Media Composition

Clostridium ljungdahlii DSM 13528 was acquired from The Leibniz Institute DSMZ (Braunschweig, Germany). Routine growth was performed in YTF media (10 g L^{-1} yeast extract, 16 g L^{-1} Bacto tryptone, 4 g L^{-1} sodium chloride, 5 g L^{-1} fructose, 0.5 g L^{-1} cysteine, pH 6) at 37°C with a N_2 headspace. Routine manipulations and growth were performed in a COY (Grass Lake, MI) anaerobic growth chamber, flushed with 95% N_2 and 5% H_2 and maintained anaerobic by palladium catalyst.

For growth and fermentation characterization, cells were grown anaerobically in PETC 1754 media (ATCC) with sodium sulfide omitted. The gas pressure was at 1 atm, and the headspace was flushed with N_2 unless indicated otherwise. Cells were grown in 18 × 150 mm Balch tubes (approximate total volume of 29 mL) crimped shut with 5 mL of PETC, with fructose, CO, or H_2 added as described, standing upright in a 37°C incubator, shaking at 250 RPM. All fermentation data shown was the result of three replicates. Media components were supplied by Sigma-Aldrich (St. Louis, MO), while gases were > 99% purity, supplied by Airgas (Radnor Township, PA).

Escherichia coli strains were acquired from New England Biolabs (NEB) (Ipswich, MA). NEB 10-beta was utilized for general molecular purposes and plasmid propagation, while *E. coli* strain NEB Express was used for plasmid preparation for transformation into *C. ljungdahlii* due to methylation compatibility (Leang et al., 2013).

Molecular Techniques

Standard molecular techniques were used with enzymes from NEB. For routine PCR and cloning, Phusion polymerase was used to create and amplify fragments. For PCR screening the genetic loci of target genes, LongAmp polymerase was used. The ladder used was 1 kb Opti-DNA Marker from Applied Biological Materials (Vancouver, Canada). To generate new constructs, DNA was ligated together using Gibson assembly (NEB).

Plasmids from the pMTL80000 modular system from Chain Biotech (Nottingham, United Kingdom) were used to generate the constructs transformed into *C. ljungdahlii*. The original construct with Cas9 targeting *pta* was acquired from the original paper's author describing the Cas9 system in *C. ljungdahlii* and ligated into pMTL83151 to generate the *pta* deletion construct (Huang et al., 2016). The plasmids were retargeting accordingly by ligating the new guide RNA and homology arms via PCR overlap extension (plasmids and primers are listed in **Supplementary Table S1**).

Complement plasmid for the *aor2* and *pta/ack* were created using HiFi assembly (NEB) as follows. Gene fragments were gel purified from PCR amplification using Phusion DNA polymerase and the primers stated in **Supplementary Table S1**. pMTL83151 was cut using *Not*I and gel purified. Purified gene fragments and *Not*I cut pMTL83151 were assembled using HiFi assembly following the manufacturer's instructions. Resulting colonies were screened with the primers used to make the gene fragments. Positive colonies were grown overnight, and plasmids purified using Monarch plasmid miniprep kit (NEB). Confirmation of complete plasmids was confirmed by whole plasmid sequencing by MBH DNA Core (Cambridge, MA).

Electrocompetent Cell Preparation and Transformation Protocol

Transformation protocols were similar to previously reported conditions (Leang et al., 2013). Briefly, *C. ljungdahlii* was

grown on YTF with 40 mM DL-Threonine to early log phase (0.2–0.7), harvested and washed with ice cold SMP buffer (270 mM sucrose, 1 mM $MgCl_2$, 7 mM sodium phosphate, pH 6), then resuspended in SMP buffer with 10% (dimethyl sulfoxide) DMSO and frozen at −80°C until use for transformation.

Electroporation was performed in a COY chamber with a Bio-Rad (Hercules, California) Gene Pulser Xcell electroporator under the following conditions: 25 μL cells were mixed with 2–10 μg of DNA in a 1 mm cuvette, pulsed 625 kV with resistance at 600 Ω and a capacitance of 25 μF. They were then resuspended in 5 mL of YTF and recovered overnight at 37°C. Cells were then plated in 1.5% agar YTF with thiamphenicol (Tm) at a final concentration of 10 μg/mL, which was Tm concentration used for general propagation. Typically, colonies appeared between 3 and 7 days after plating.

Generating and Validating Deletion Strains

To generate the plasmids targeting *pta* (CLJU_c12770) NC_014328.1 (1376316.1377317), we received the *pta* targeting Cas9 fragment described from Huang et al. (2016) and recreated the *pta* deletion vector by cloning in the fragment via restriction cloning into pMTL83151. Subsequent *aor2* (CLJU_c20210) (2192715.2194538), and *adhE1*/*adhE2* (CLJU_c16510/CLJU_c16520) (1791269.1793881/1794033.1796666), Cas9 deletion constructs were generated with the *pta* targeting vector as a backbone, using PCR to change gene targeting, similar to previously reported protocols (Huang et al., 2016). For the *aor1* (CLJU_c20110) (2179645.2181468) deletion, the original promoter driving *cas9* was replaced with a riboswitch-linked promoter (pbuE, P = 8 bp) activated by 2-aminopurine (2-AP) at 2 mM (Marcano-Velazquez et al., 2019). These strains were dilution plated on YTF agar with Tm and 2-AP and allowed for grow until colonies appeared (around 1 week). Once colonies appeared, they were individually picked, and PCR screened to isolate bacteria with successful genetic editing. To cure the plasmid and restore sensitivity to Tm, these strains were serially transferred 2 times in non-selective media and dilution plated on YTF agar for single colony isolation, then tested for sensitivity to Tm and PCR screened for loss of the Tm resistance gene.

Analytical Techniques

Fermentation liquid samples of 150 μL were extracted by syringe, filtered using Costar Spin-X 0.45 μm filters (Corning, Corning, NY), and stored at −20°C until the experiments completed. Fermentation products in the liquid phase (acetate, ethanol, and 2,3 butanediol) were measured by high-performance liquid chromatography on a 1,200 series Agilent (Santa Clara, CA) with an Aminex HPX-87H column using a Micro Guard Cation H Cartridge, at 55°C with 4 mM H_2SO_4 mobile phase, as previously described (Marcano-Velazquez et al., 2019). Optical density was measured at 600 nm using a Milton Roy (Ivyland, PA) Spectronic 21D.

Biochemical Assays

Wild type and derived cultures were grown in 50 mL PETC medium with 5 g/L fructose and harvested at mid log phase (0.3–0.7 OD_{600}) by centrifugation at 4,000 g for 10 min. The supernatants were discarded, and pellets were frozen at −80°C for future use. Cell pellets were resuspended in 200 μL phosphate buffer (pH 7.6) which contained 25 mg/ml lysozyme and 10 mM dithiothreitol. Resuspended cells were added to a screwcap tube contained acid washed glass beads (Sigma-Aldrich, United States) and lysed using a Tissuelyser II (Qiagen, Germany) at a frequency of 30 1/s for 2 min. Lysed cells were subsequently centrifuged at 17,000 g for 2 min. The resulting supernatant was extracted and stored in anaerobic vials at –80°C for future use. Protein concentration was determined using Bradford reagent (Sigma-Aldrich) following the manufacturer's instructions using a Tecan infinite M200 pro plate reader (Tecan Life Sciences, Switzerland). All assays are reported as specific activities (μmol min^{-1} mg^{-1}) and were performed in triplicate in the anaerobic COY chamber at 27°C. Enzyme assay mixtures without cell free extract were used to track baseline changes in absorbance.

To test Pta activity in the wild type and derived mutants, the procedure employed was based on formation of thioester bond formation of acetyl-CoA, following the change in absorbance at 233 nm using a molar extinction coefficient of 4,360 M^{-1} cm^{-1} (Klotzsch, 1969). In short, 200 μL of a 100 mM Tris-HCl (pH 7.6) containing 1.6 mM glutathione, 0.43 mM coenzyme-A (CoA), 7.23 mM acetyl-phosphate, and 13.3 mM ammonium sulfate was added to 1 μL of cell free extract, and monitored for 233 nm change in a NanoDrop One from Thermo Fisher Scientific (Waltham, MA) at 15 s intervals for 4 min.

To test Aor activity in the wild type and derived mutants, the procedure employed was based on benzyl viologen reduction (Heider et al., 1995). In short, 200 μL of a 100 mM phosphate buffered mix (pH 7.6) containing 1.6 mM benzyl viologen, 1 mM acetaldehyde and 1 mM dithiothreitol was added 10 μL of cell free extract in a flat bottom 96 well plate. Reduction of benzyl viologen was measured at 600 nm using a molar extinction coefficient of 7,400 M^{-1} cm^{-1} using BioTek Synergy Neo2 plate reader (BioTek Instruments, United States) at 10 s intervals for 5 min.

To test both Adh and Aldh activity in the wild type and derived mutants, enzymatic reduction of NAD^+/$NADP^+$ was employed based on previous published methods (Lo et al., 2015). For the Adh assay, 200 μL of a 100 mM Tris-HCL (pH 7.8) buffer mixture containing 10 mM ethanol, 1 mM NAD^+ or 1 mM $NADP^+$ and 1 mM dithiothreitol was added to 10 μL of cell free extract in a flat bottom 96 well plate. Reduction of NAD^+ or $NADP^+$ was measured at 340 nm using BioTek Synergy Neo2 plate reader at 10 s intervals for 5 min. For the Aldh assay, 200 μL of 100 mM Tris-HCL (pH 7.8) buffer mixture containing 0.43 mM CoA, 1 mM NAD^+ or 1 mM $NADP^+$, 10 mM acetaldehyde and 1 mM dithiothreitol was added to 10 μL of cell free extract in a flat bottom 96 well plate. Reduction

$NAD(P)^+$ was calculated using a molar extinction coefficient of 6,220 M^{-1} cm^{-1}.

RESULTS

Generation and Characterization of Phosphotransacetylase Gene(*pta*) Negative Strain

To determine the role of *pta* in acetate formation, we utilized Cas9 to delete *pta*. Cas9 can be utilized to generate markerless deletion strains, which is useful for generating strains with multiple deletions when selective markers are limited. Because of the importance of *pta* to acetogenic metabolism and as a target for genetic engineering for increased product formation, we acquired the Cas9 *pta* deletion cassette and with which created the *pta* deletion strain (**Figure 1**).

Surprisingly, in our initial characterization of heterotrophic growth, there was little initial difference between the Δ*pta* and wild-type strain (**Figure 2**). The wild type and Δ*pta* strain grew at similar rates and produced a similar amount of products. Both the wild type and Δ*pta* strain produced high levels of acetate during heterotrophic growth (51 vs. 47 mM) and similar amounts of ethanol (∼6 mM). The only notable difference is that the Δ*pta* strain produced markedly more 2,3 butanediol (2,3 BDO) than the wild-type strain (8 vs. 2 mM). The carbon/electron recovery from fermentation products was 0.79/0.85 for the wild type and 0.86/0.97 for the Δ*pta* strain.

Predictions from the metabolic model developed by Nagarajan et al. (2013) suggested loss of *pta* would redirect flux toward ethanol formation through AdhE, yet we observed no increase in ethanol. We reasoned that if acetyl-CoA was no longer being directed toward acetate via acetyl-P, perhaps it was through an acetaldehyde intermediate via Aldh and AOR. If this was the case, NADH would be consumed, ferredoxin would be reduced, and acetaldehyde would be oxidized to acetate. The increased reduced ferredoxin (Fd_{red}) could then be utilized by the WLP to fix the CO_2 evolved by glycolysis. Redox consumption via the WLP may be preventing ethanol formation.

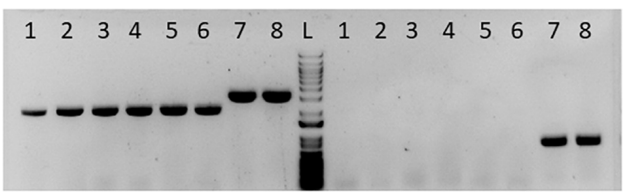


FIGURE 1 | Gel picture showing successful deletion of *pta* in *C. ljungdahlii*. Columns labeled 1–6 to the left of the ladder "L" were presumptive *pta* deletion colonies for single colony purification. Column 7 was wild type DNA mixed with the *pta* deletion plasmid. Column 8 is the wild type DNA alone. The set of lanes to the left of the ladder "L" was PCR amplifying the *pta* region. Columns 1–6 show a smaller *pta* region, indicating loss of *pta* at that region in the genome. Columns 7 and 8 show the wild type band size. The set of lanes to the right of the ladder (L) was PCR amplifying a region inside *pta*. Columns 1–6 show no fragment amplified, indicating no detectable *pta* in those colonies. Columns 7 and 8 show the fragment of *pta* in unedited DNA.

Creation and Characterization of Δ*pta* Derivative Strains

If *pta* is not essential for acetate formation, how is acetate enzymatically produced in the Δ*pta* strain? There are a number of enzymes identified that can influence acetyl-CoA that may change product formation, in particular AOR and the AdhE. We therefore targeted these genes using CRISPR/Cas9 to understand their metabolic function (**Figure 3**). We show deletion of these genes using PCR to amplify external target gene regions (**Figure 3C**) and confirmed loss of the genes by PCR and sequencing (**Supplementary Figures S1, S2**).

AOR catalyzes the fully reversible interconversion of acetate and acetaldehyde with ferredoxin as the electron donor or acceptor. Although reversible, AOR's physiological directionality is of question but could be the source of acetate, if acetaldehyde is derived from acetyl-CoA (Ueki et al., 2014; Richter et al., 2016; Nissen and Basen, 2019). Transcriptomics studies under heterotrophic and autotrophic conditions have shown that *aor2* (CLJU_c20210) is the predominantly expressed AOR (2085 heterotrophic/612 autotrophic FPKM), vs. *aor1* (CLJU_c20110) (10 heterotrophic/7 autotrophic FPKM) (Nagarajan et al., 2013). CRISPRi targeting *aor2* showed an effect on fermentation products, although only while targeting *pta* as well (Woolston et al., 2018). Thus, we decided to first target *aor2* using CRISPR/Cas9 in the Δ*pta* strain background.

The double Δ*pta* Δ*aor2* strain under heterotrophic conditions produced much more ethanol than the wild type or parent Δ*pta* strains. Whereas ethanol was produced in trace amounts in both those strains with an acetate to ethanol ratios of 9:1 (**Figure 2**), the Δ*pta* Δ*aor2* strain consistently produced acetate to ethanol in a 2:1 ratio (**Figure 4A**). This suggests that *aor2* directionality under heterotrophic growth at pH 6.0 is in the acetate forming direction. It also suggests *aor2* was important for reduced ferredoxin production and/or acetaldehyde oxidation and prevented ethanol formation under heterotrophic conditions. In contrast to the Δ*pta* strain, the Δ*pta* Δ*aor2* strain under these conditions produced similar 2,3 BDO (2 mM) to the wild type, despite an increase in the production of ethanol. The carbon/electron recovery from fermentation products was 0.82/0.97 for the Δ*pta* Δ*aor2* strain.

Next, we decided to target the bifunctional alcohol/aldehyde dehydrogenases (*adhE*) genes in the Δ*pta* strain. Bifunctional alcohol/aldehyde dehydrogenases catalyze the conversion of acetyl-CoA to acetaldehyde to ethanol using 2 NAD(P)H as electron donors, and could be the source of acetaldehyde substrate for AOR. These genes are also considered important for ethanol production in other *Clostridia*, so the production of both acetate and ethanol could be affected (Cooksley et al., 2012; Lo et al., 2015). There are two annotated *adhE* genes in the genome, *adhE1* and *adhE2* (CLJU_c16510 and CLJU_c16520), which are colocalized in the genome. It was previously shown that disruption of *adhE1* but not *adhE2* resulted in a decrease in ethanol when grown on fructose (Leang et al., 2013). Omics

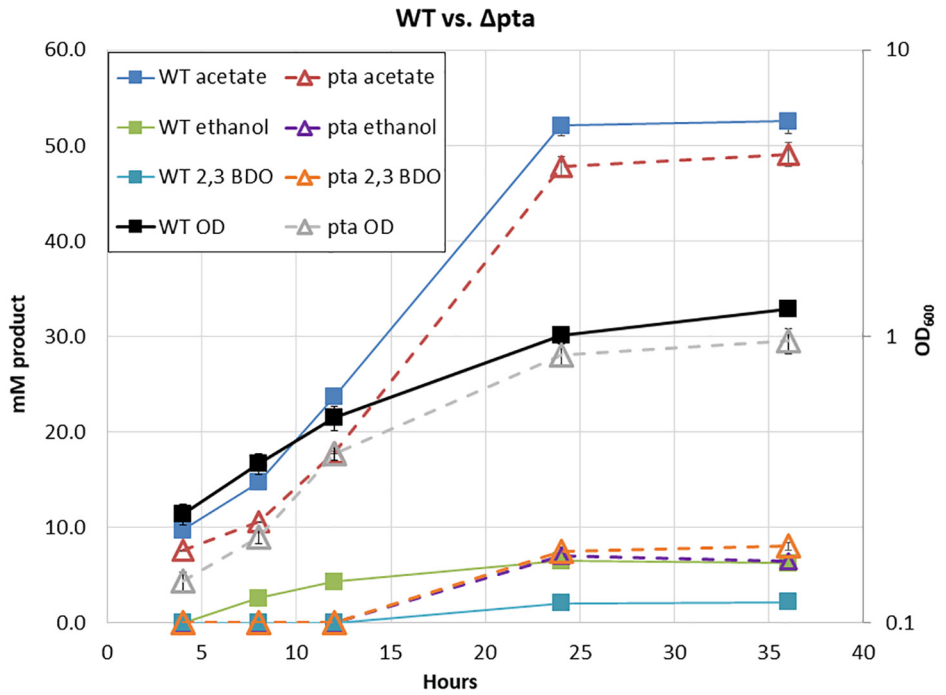


FIGURE 2 | Heterotrophic growth and fermentation products of *C. ljungdahlii* wild type and the *pta* deletion strain on PETC with 5 g/L fructose. Wild type is marked with solid lines and squares, while Δ*pta* is marked dashed lines with hollow triangles. Growth rate, OD_{600}, and fermentation products are similar across the duration of the experiment.

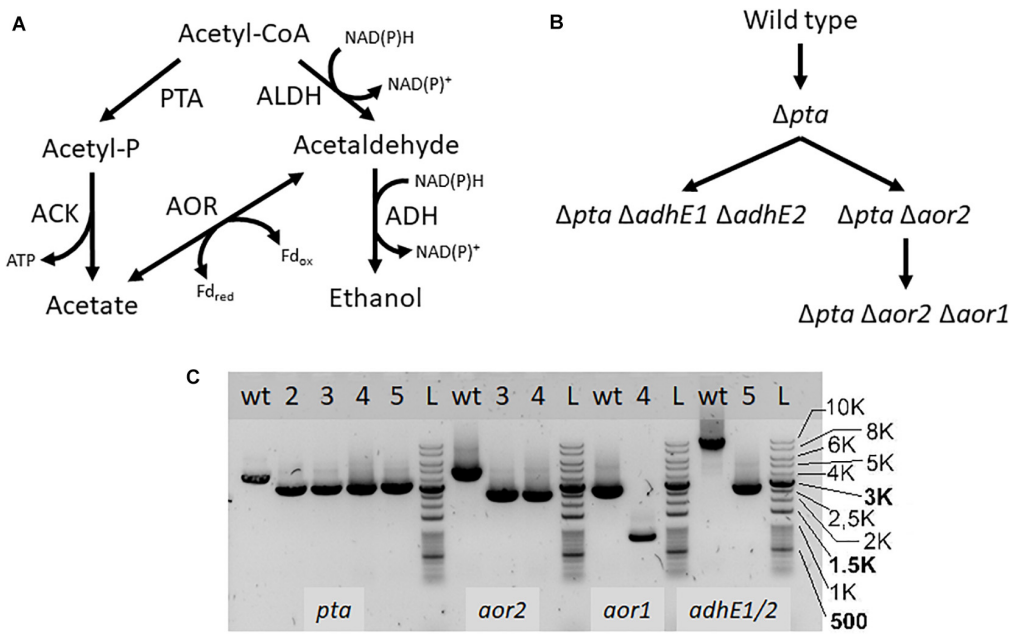


FIGURE 3 | Acetyl-CoA metabolism and creation of deletion mutants. PTA, phosphotransacetylase; ACK, acetate kinase; ALDH/ADH, aldehyde/alcohol dehydrogenase; AOR, aldehyde ferredoxin oxidoreductase **(A)** Map of acetyl-CoA metabolism with metabolites, cofactors, and enzyme reactions. **(B)** Strain map showing lineage of deletions made. **(C)** PCR confirmation of deletion strains. Primers were used to amplify the genomic regions of *pta*, *aor2*, *aor1*, *adhE1,* and *adhE2*, to confirm genomic deletions. Lanes labeled with "wt" is wild type DNA control, while lanes with "L" are the DNA ladder. Lanes with numbers correspond to the following strains: 2- Δ*pta*, 3- Δ*pta* Δ*aor2*, 4- Δ*pta* Δ*aor2* Δ*aor1*, 5- Δ*pta* Δ*adhE1* Δ*adhE2*. The lower bands in the lanes correspond to expected gene deletions targeted by the homology arms.

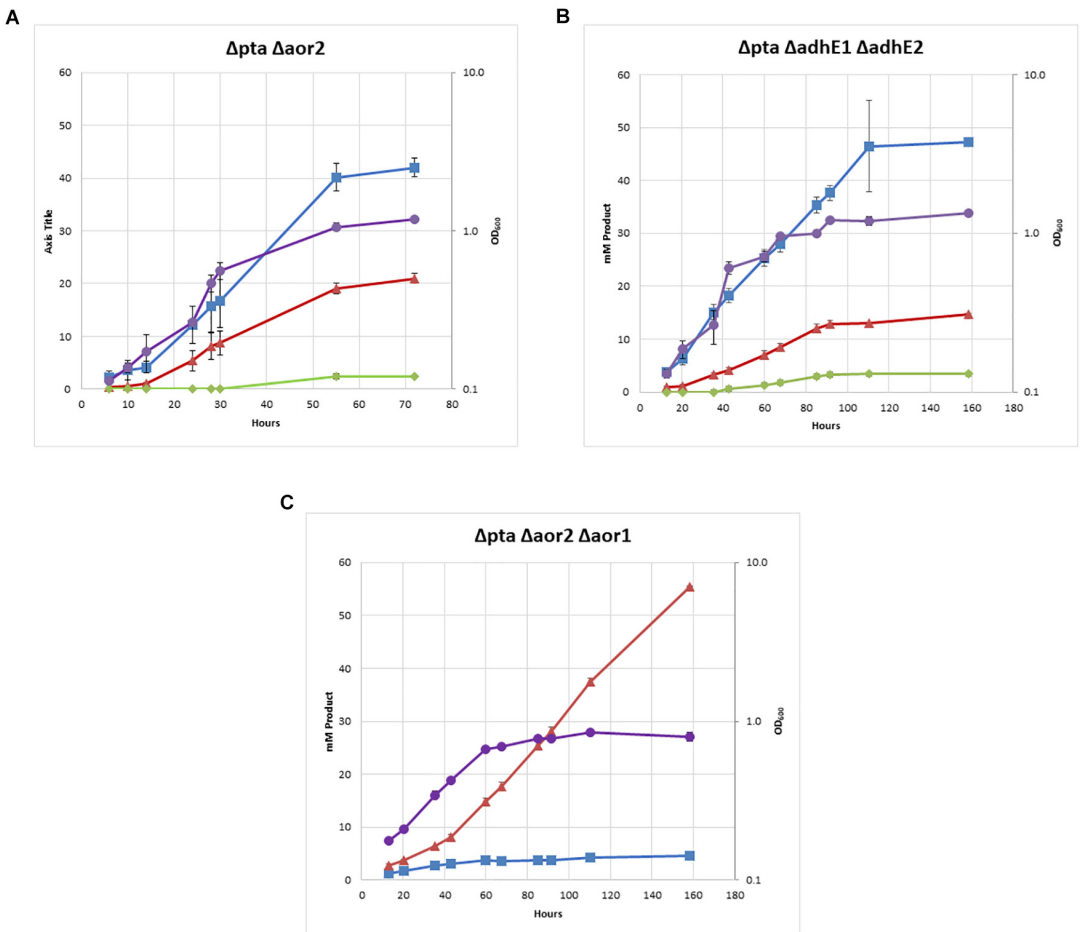


FIGURE 4 | Heterotrophic growth and fermentation products of strains on PETC with 5 g/L fructose. Acetate in blue squares, ethanol in red triangles, 2,3-butanediol in green diamonds, OD_{600} in purple circles. **(A)** Δ*pta* Δ*aor2,* **(B)** Δ*pta* Δ*adhE1* Δ*adhE2,* and **(C)** Δ*pta* Δ*aor2* Δ*aor1.*

studies have also shown that this is probably because *adhE1* is the predominantly expressed AdhE (166 heterotrophic/0.1 autotrophic FPKM) vs. *adhE2* (5.3 heterotrophic/0 autotrophic FPKM) (Nagarajan et al., 2013). Due to colocalization of *adhE1* and *adhE2*, we decided to simultaneously target both to create the strain Δ*pta* Δ*adhE1* Δ*adhE2* (**Figure 3C**). We targeted *adhE1* using the guide RNA and supplied recombination regions downstream of the *adhE2* for repair. The Δ*pta* Δ*adhE1* Δ*adhE2* strain showed increased ethanol formation (15 mM) over the parent Δ*pta* strain (∼5 mM), suggesting that there are additional proteins involved in ethanol formation, perhaps from other annotated acetaldehyde/alcohol dehydrogenases (**Figure 4B**; Whitham et al., 2015). The acetate amount is comparable to that of the Δ*pta* parental strain. The carbon/electron recovery from fermentation products was 0.83/0.94 for the Δ*pta* Δ*adhE1* Δ*adhE2* strain.

Finally, we targeted *aor1* in the Δ*pta* Δ*aor2* strain. Whether *aor1* had an influence on metabolism in *C. ljungdahlii* was unclear in literature. Low detection of *aor1* was shown in a previous study (10.4 heterotrophic/6.5 autotrophic FPKM) (Nagarajan et al., 2013), although there were reports of increased *aor1* under some conditions and variability in some reports (Whitham et al., 2015; Richter et al., 2016; Aklujkar et al., 2017). To target *aor1* we decided to test the original Cas9 construct vs. a riboswitch inducible Cas9 recently developed (Marcano-Velazquez et al., 2019). Working with Cas9 construct for the previous deletions was difficult due to several reasons. In *E. coli*, the plasmid was unstable and often produced truncated plasmids, and did not readily produce concentrated plasmid. In *C. ljungdahlii* transformation efficiencies were low due to constitutive expression of Cas9, which would only allow growth of colonies that had both taken up the plasmid and undergone successful genome editing (Huang et al., 2016). Coupling of an inducible riboswitch with Cas9 has shown improved genomic editing in other *Clostridia* (Cañadas et al., 2019). Switching the original thiolase promoter with the riboswitch inducible promoter (the 2-AP riboswitch linked with the Cthe_2638 promoter, pbuE, P = 8 bp) resulted in improved performance in *E. coli* and allowed comparable transformation efficiency of non-Cas9 plasmids, while we were unable to get successful transformants of the constitutive Cas9 construct targeting *aor1*. After getting successful colonies, we induced Cas9 using 2-AP

and plated cells on 2-AP plates. Five colonies were isolated and tested, with two showing the edited *aor1* genotype.

The Δ*pta* Δ*aor2* Δ*aor1* strain showed several distinct features compared to Δ*pta* Δ*aor2* (**Figure 4C**). There was a dramatic decrease in acetate and increase in ethanol, with a ∼85% selectivity for ethanol compared to acetate. No 2,3 BDO was detected. The strain also grew much slower than other strains, including Δ*pta* Δ*aor2*, but was eventually able to consume all the fructose. The carbon/electron recovery from fermentation products was 0.72/1.05 for the Δ*pta* Δ*aor2* Δ*aor1* strain. To confirm the acetate phenotype, the Δ*pta* Δ*aor2* Δ*aor1* strain was retransformed with the control plasmid or a plasmid expressing *pta*/*ack* or *aor2*, which resulted in increased acetate vs. Δ*pta* Δ*aor2* Δ*aor1* with the control plasmid (**Supplementary Figure S3**).The plasmid complementation did not completely restore acetate production, but that may be due to inferior expression from plasmids (Leang et al., 2013).

Enzyme Assays on Strains

To examine changes brought by targeted gene deletions, we assayed strains for enzymatic activity based on the targeted genes. First, we measured phosphotransacetylase activity of the wild type and Δ*pta* strain and confirmed loss of detectable phosphotransacetylase activity, measuring wild type specific activity of 1.47 vs. < 0.01 for the Δ*pta* strain. Next, we measured the activity of AOR, Aldh, and Adh (**Table 1**). For AOR activity, we measured relatively low activity in the wild type strain, which increased fivefold in the Δ*pta* strain. Interestingly, the Δ*pta* Δ*aor2* strain had the highest detectable AOR activity, while the Δ*pta* Δ*aor2* Δ*aor1* strain had the lowest activity of the Δ*pta* derivative strains, but still significantly higher activity than the wild type. For Aldh NAD^+ activity, we found that a large increase in the Δ*pta* strain vs. the wild type, and a subsequent reduction in Aldh activity in the Δ*pta* Δ*adhE1* Δ*adhE2*, although still above wild type levels. For ADH activity, activities were low for wild type, Δ*pta*, and Δ*pta* Δ*adhE1* Δ*adhE2*, although the highest activity was found in the Δ*pta* Δ*adhE1* Δ*adhE2*.

Autotrophic Growth of Strains

One of the most important predicted functions of Pta in *C. ljungdahlii* is to generate acetyl-P for ATP synthesis. The previous CRISPR/Cas9 study on the Δ*pta* strain showed no significant growth on syngas (Huang et al., 2016). In contrast with previously data, we demonstrated robust growth of most strains on 100% CO in PETC media without sugar over a 7-day period (**Figure 5**), except for Δ*pta* Δ*aor2* Δ*aor1* which did not grow (OD did not increase by more than 0.1 over 7 days). Acetate was the primary product formed in all strains, with Δ*pta* Δ*aor2* having the least (16 mM) and wild type and Δ*pta* the most (∼23 mM). Δ*pta* Δ*aor2* had the most (7 mM) of ethanol formed, while the others had similar levels (∼5 mM). 2,3 BDO was not a significant product (<2 mM formed). Final OD of strains ranged between 0.35 and 0.48, with strains that generated more acetate having a higher OD.

H_2-Enhanced Mixotrophic Growth for Carbon Conversion

Traditional fermentation ethanol yields are limited to a theoretical 66% carbon yield due to loss of CO_2 via decarboxylation when converting pyruvate to acetyl-CoA. Acetogens can overcome these limitations by simultaneously utilizing fructose and syngas in mixotrophic growth, allowing for capture of CO_2 and conversion to acetyl-CoA (Jones et al., 2016). *C. ljungdahlii* was shown to be one of the best acetogens at syngas-enhanced mixotrophic growth as well as ethanol tolerance.

Since the Δ*pta* Δ*aor2* Δ*aor1* was able to produce ethanol at a high yield heterotrophically, we were interested in its performance under H_2-enhanced mixotrophic growth compared to the wild type. Theoretically H_2 could be used to capture lost CO_2 from glycolysis and supply reducing power to reduce acetyl-CoA to ethanol, resulting in a carbon yield increase on a per sugar basis. In a wild type strain, addition of H_2 with growth on sugar was shown to shift fermentation products toward more ethanol (Jones et al., 2016).

While the Δ*pta* Δ*aor2* Δ*aor1* grew heterotrophically but not on CO alone, we wondered whether the strain could utilize fructose and H_2 simultaneously to get a larger ethanol yield than on fructose alone. To test this, we grew the wild type strain and the Δ*pta* Δ*aor2* Δ*aor1* strain on 5 g/L fructose with either 100% H_2 or 100% CO_2 (as a control) in the headspace (**Figure 6**). After a week of growth, the wild type with CO_2 or H_2 and Δ*pta* Δ*aor2* Δ*aor1* with CO_2 consumed most of the sugar, while the Δ*pta* Δ*aor2* Δ*aor1* with H_2 failed to grow at all. Based on the fructose added, the carbon recovery of the wild type strain was ∼90% in CO_2 and ∼105% in H_2, while only 75% Δ*pta* Δ*aor2* Δ*aor1* in CO_2.

TABLE 1 | AOR, Adh, and Aldh cell free extract specific activity.

	BV AOR	NAD^+ ADH	$NADP^+$ ADH	NAD^+ ALDH	$NADP^+$ ALDH
WT	0.09 (0.00)	0.01 (0.01)	0.03 (0.01)	0.01 (0.01)	0.01 (0.01)
pta	0.38 (0.02)	0.01 (0.01)	0.00 (0.00)	0.46 (0.01)	0.01 (0.01)
pta aor2	0.47 (0.02)	ND	ND	ND	ND
pta aor2 aor1	0.31 (0.02)	ND	ND	ND	ND
pta adhE1 adhE2	ND	0.03 (0.01)	0.01 (0.01)	0.08 (0.01)	0.01 (0.01)

Cell free extract activities of acetaldehyde ferredoxin oxidoreductase, alcohol dehydrogenase, and aldehyde dehydrogenase, reported in μmol min^{-1} mg^{-1}, measured as reduction of the electron acceptor. The top line lists the electron acceptor, either $NAD(P)^+$ or benzyl viologen (BV). AOR, aldehyde ferredoxin oxidoreductase; ADH, alcohol dehydrogenase; ALDH, aldehyde dehydrogenase; ND, not determined. Standard deviation is in parenthesis. For all assays, n = 3.

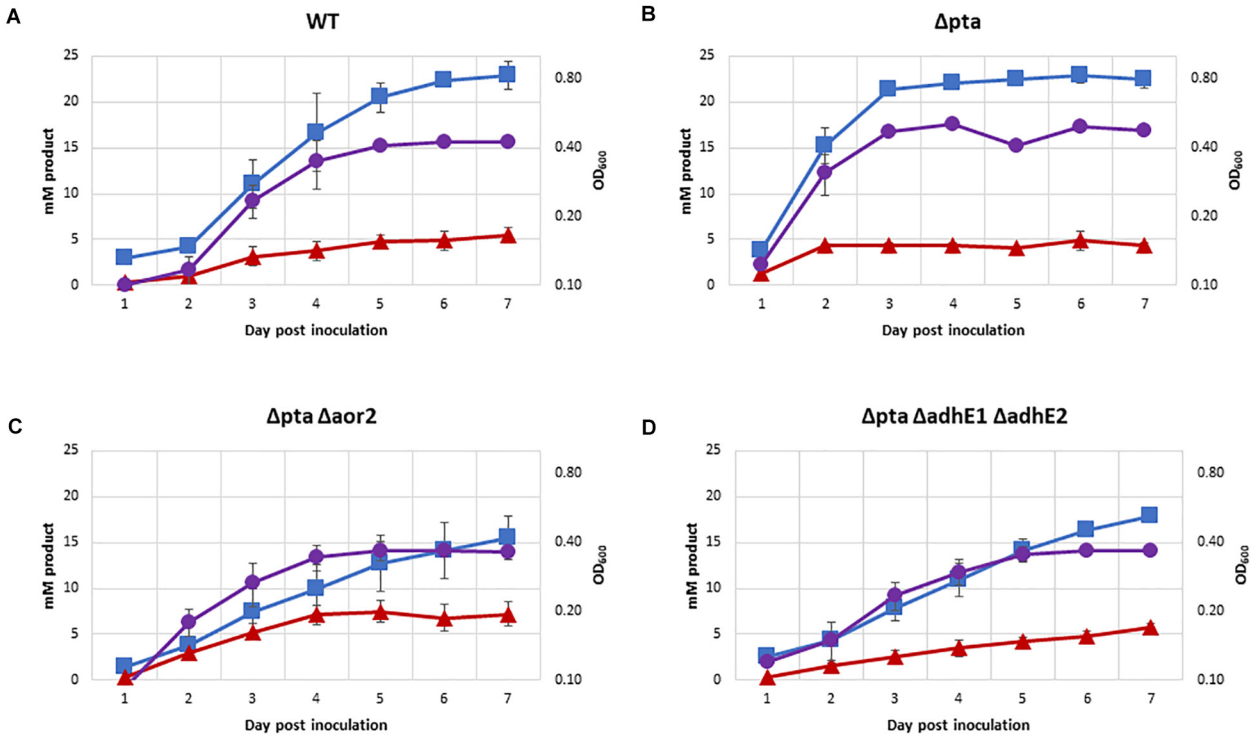


FIGURE 5 | Growth and fermentation products of strains on PETC with 100% CO. Strains were grown over 7 days with fermentation products and growth monitored, with approximately a 5:1 headspace to liquid ratio. Acetate as blue squares, ethanol as red triangles, OD_{600} as purple circles. **(A)** Wild type, **(B)** Δ*pta*, **(C)** Δ*pta* Δ*aor2*, and **(D)** Δ*pta* Δ*adhE1* Δ*adhE2*.

DISCUSSION

Acetate formation is a key characteristic of acetogens, which utilize the WLP to fix CO_2 and produce primarily acetate. Surprisingly under a wide variety of conditions in *C. ljungdahlii*, Pta is unnecessary for acetate formation, despite being normally understood as a critical enzyme for ATP production and acetate formation. What are the underlying mechanisms and enzymes that drive *C. ljungdahlii* metabolism? There are a few models that attempt to predict *C. ljungdahlii* metabolism. Nagarajan's genome-scale model suggested that deleting *pta* and *ack* would force ethanol production rather than acetate, but this model was missing the AOR reaction functioning in the acetate forming direction (Nagarajan et al., 2013). Richter et al. (2016) suggested that thermodynamics, rather than enzyme expression, is a driving force behind the product formation in *C. ljungdahlii*, and that AOR is an important arbiter in determining acetate:ethanol ratios. We provide evidence that model has value. Despite *pta* as one of the most highly expressed genes and an important contributor to ATP/acetate formation, the Δ*pta* and wild-type strain had very similar product profiles. Richter et al. (2016), however, suggested that a *pta/ack* knockout strain would not produce acetate, but we have shown high acetate yield without *pta*. We believe this is through the activity of AOR. AOR has been proposed to be important for metabolism in *C. ljungdahlii*, and *aor2* has been found to be expressed under both heterotrophic and autotrophic growth conditions (Nagarajan et al., 2013; Ueki et al., 2014; Whitham et al., 2015; Richter et al., 2016; Aklujkar et al., 2017; Woolston et al., 2018; Zhu et al., 2020). What is the purpose of AOR in acetogens? There are several potential functions: ferredoxin reoxidation under growth on CO, detoxifying short-chain fatty acids when pH is low, and maximizing ATP yield during autotrophic ethanol production (Mock et al., 2015; Richter et al., 2016; Liew et al., 2017).

We provide evidence that a potential function of AOR may be to supply ferredoxin via an acetaldehyde intermediate from acetyl-CoA. Although in acetogens, AOR was thought to primarily function in the reduction of carboxylic acids, oxidation of aldehydes may be an important feature as well. Balancing redox cofactors is likely to be an important part of *C. ljungdahlii* metabolism and AOR could be a method of indirectly "upgrading" electrons from NAD(P)H (E_0' = −320 mV) to Fd_{red} (E_0' = −400 mV). The directionality of AOR is partially determined by pH, as AOR is only active on protonated short-chain fatty acids (i.e., acetic acid, pKa = 4.7) (Huber et al., 1995; Napora-Wijata et al., 2014). A higher pH would drive the reaction toward reducing ferredoxin as acetic acid would be deprotonated to acetate. High expression of AORs has been shown in both heterotrophic and autotrophic conditions, despite high yields of acetate, so even in Pta and Ack competent strains AOR may be converting acetaldehyde to acetate. This process would provide

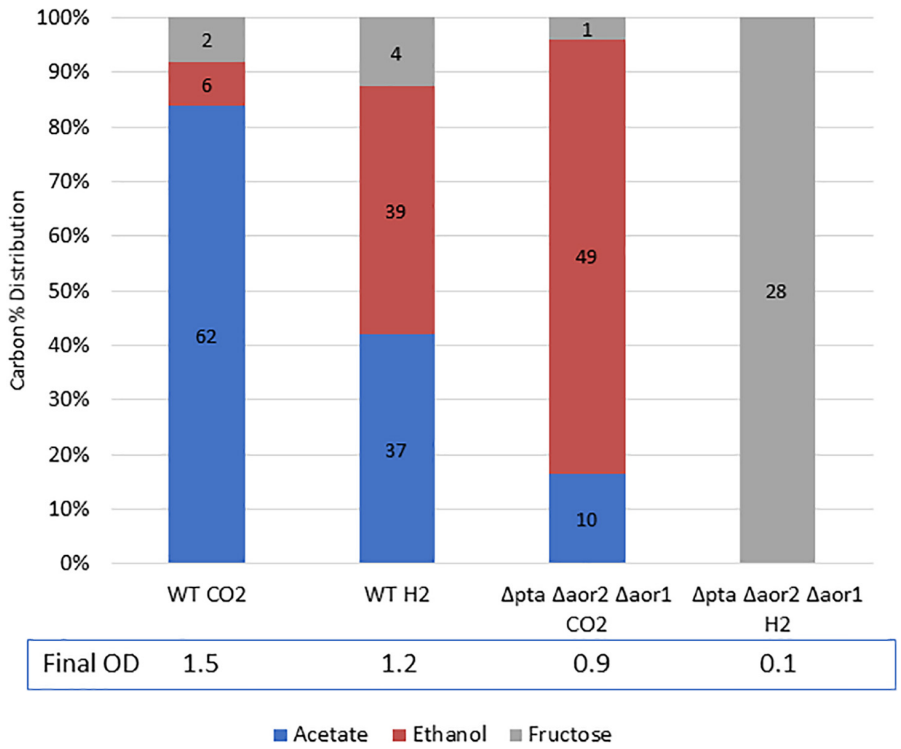


FIGURE 6 | Mixotrophic growth and fermentation products of wild type vs. Δ*pta* Δ*aor2* Δ*aor1* on PETC with 5 g/L fructose with either 100% CO_2 or 100% H_2. Strains were grown on 5 g/L (28 mM) fructose with the indicated headspace gas (CO_2 or H_2). Carbon distribution of fermentation products/residual fructose is indicated by bar chart%, while the number label is mM amount after 5 days. The OD_{600} after 5 days is indicated at the bottom.

redox flexibility and ATP through the conversion of Fd_{red} to NAD(P)H via the NADH ferredoxin:$NADP^+$ oxidoreductase and proton-translocating ferredoxin:NAD^+ oxidoreductase, which would then be consumed by the electron demanding WLP. Indeed, in many of our Δ*pta* strains, growth and product formation was like the wild type and only appeared to be severely compromised when both *aor1* and *aor2* were also deleted (**Figure 4C**). Interestingly, enzyme assays in the Δ*pta* strain of both NAD^+-linked Aldh and AOR were significantly higher than the wild type strain, suggesting that this pathway is the main route of acetate formation in this strain. Even with loss of *aor2*, we detected the highest AOR activity in the Δ*pta* Δ*aor2* strain (**Table 1**). It should be noted that the AOR assay used benzyl viologen rather than ferredoxin, and thus these activities may not be true indicators of AOR activity. AOR genes are widely distributed among acetogens and could serve a similar role in those organisms (Nissen and Basen, 2019). In the closely related *Clostridium autoethanogenum*, *aor1,* and *aor2* were deleted, and it was shown that loss of *aor2*, but not *aor1*, improves autotrophic ethanol production. In the same study, it was also shown that loss of these genes prevents reduction of carboxylic acids (Liew et al., 2017). It is unknown why there is a difference in the response between *C. ljungdahlii* and *C. autoethanogenum*, but potentially it was because the *pta/ack* pathway was functional in those strains. Pta and Ack are usually some of the highest detected enzymes, which points to their central role in acetate formation. In our Δ*pta* derivative strains, AOR may be the only functional route to acetate, which may be a preferred product due to electron consumption by the WLP. It is still unknown why there are two genes and what factors drive their expression. Additionally, while these organisms are over 99% identical, there are significant differences in metabolism, substrate consumption, and product formation (Bruno-Barcena et al., 2013; Jones et al., 2016). Specific differences in the WLP enzymes include a uniquely truncated *C. autoethanogenum acsA* gene which could have a major effect on metabolism (Liew et al., 2016).

The Δ*pta* Δ*aor2* Δ*aor1* strain shows slower heterotrophic growth. Why is this the case? The WLP is an electron demanding process and the formation of ethanol would potentially cause redox imbalance. Alternatively, if AOR "upgrading" of electrons is disabled, there may be a cofactor mismatch between electrons produced by glycolysis (Fd_{red} and NADH) and demanded by the WLP. Electron requiring enzymes of the WLP require redox cofactors (i.e., NADPH, Fd_{red}, and possibly NADH) in specific amounts. Fd_{red} can be used to reduce $NAD(P)^+$ through the activity of Rnf and Nfn, which has been detected in high activities, suggesting these interconversions are actively taking place.

The Δ*pta* Δ*aor2* Δ*aor1* strain also showed no detectable autotrophic growth, even though ethanol formation through *adhE* is a net ATP positive process (Bertsch and Müller, 2015; Mock et al., 2015). We suspect that this is due to the regulation of *adhE1*, as *adhE1* appears to be highly expressed

under heterotrophic conditions but virtually undetectable in autotrophic conditions (Nagarajan et al., 2013). This makes sense as ethanol production through ADHE is much less ATP efficient than through AOR, but the underlying regulatory mechanism of *adhE* control is still unknown. Mixotrophic conditions of Δ*pta* Δ*aor2* Δ*aor1* also showed no detectable growth. We think this may be due to CO/H_2 inhibiting strong expression of *adhE1*, like on autotrophic growth. If this is the case, *adhE1* regulation may not be an autotrophic mechanism *per se*, but rather in response to reducing gases or redox state, regardless of the presence of sugar. Alternatively, a recent study looking at *C. autoethanogenum* autotrophic growth suggests that ethanol formation through ADHE may be thermodynamically unfeasible under certain conditions and that AOR is important for energy generation and redox balance (Mahamkali et al., 2020). These factors could explain the lack of growth in the presence of syngas. Although there have been studies looking at heterotrophic, autotrophic, and mixotrophic growth, little is known about the mechanisms that govern the metabolism of these organisms. Since acetogenic metabolism is ATP and thermodynamically limited, there may be multiple mechanisms at work (Schuchmann and Müller, 2014; Molitor et al., 2017).

The Δ*pta* Δ*adhE1* Δ*adhE2* strain raises several questions around ethanol and acetate formation in this organism. Despite loss of these highly expressed enzymes under heterotrophic conditions, loss of these enzymes increased ethanol formation and significant amounts of ethanol and acetate were formed. Acetate was still the primary product, which is probably from the active AORs in this strain, but ethanol was about 33% of the product formed. AdhE's main function is probably the reversible conversion of acetyl-CoA to ethanol (and vice versa) and not acetaldehyde generation for AOR activity. AdhEs are known to form spirosomes, which are large structures believed to sequester the toxic acetaldehyde from the rest of the cell (Kim et al., 2019). However, the Δ*pta* strain had higher Aldh activity than the Δ*pta* Δ*adhE1* Δ*adhE2* strain, suggesting that these *adhE* genes may be important for Aldh activity, which has been seen in other *Clostridia* (Yao and Mikkelsen, 2010). Even with the loss of both *adhE* genes, the Δ*pta* Δ*adhE1* Δ*adhE2* strain had some Aldh activity, and slightly higher Adh activity, which could explain the increased ethanol formation. This suggests the presence of other important aldehyde and alcohol dehydrogenases. There are many others annotated and expressed, but their function and importance are not well understood (Whitham et al., 2015). Other studies have pointed out other potential candidates that are highly expressed, but as we have shown for *aor2* vs. *aor1*, expression levels under those studies may be deceptive for predicted metabolic effect. The metabolic effect of these enzymes change depending on the state of cells. For instance, during growth on syngas, ethanol can be both produced and consumed, which is dependent on both AdhE and AOR (Liu et al., 2020). Additionally, there are many potential redundant enzymes and their regulation is unknown, and while many of these genes are annotated as alcohol or aldehyde dehydrogenases, they may not be specific for acetaldehyde/ethanol (Tan et al., 2015). In general, we did not detect high Adh activity, which suggests a lack of highly expressed/active alcohol dehydrogenases. The leftover acetate in the Δ*pta* Δ*aor2* Δ*aor1* fermentations could perhaps be formed from other annotated AORs or from *ctf* (Ueki et al., 2014; Whitham et al., 2015). While we noted strain differences in enzymatic activity, these assays were extremely limited. We did not test cells under different growth conditions or attempt to optimize enzymatic activity, and we used benzyl viologen as a proxy for ferredoxin. A more careful enzymatic study may be expected to find different results (Lo et al., 2015; Mock et al., 2015). It will be important to rigorously test enzymatic activity and identify responsible enzymes, for both metabolic engineering and fundamental understanding.

Understanding the underlying mechanisms of acetogen metabolism is important for engineering them for targeted chemical formation. Acetogens are being investigated for their autotrophic, heterotrophic, and mixotrophic growth capabilities to produce a wide variety of products with high carbon-conversion efficiency. We show that substrate level phosphorylation of ATP via *pta* is not required for robust growth on CO. This may be helpful when designing pathways for product formation, as ATP production via substrate level phosphorylation is not an absolute requirement and ATP production through proton motive force generated by Rnf is sufficient (and necessary) for autotrophic growth (Tremblay et al., 2013). We did not test autotrophic growth differences of the strains on H_2/CO_2, which is expected to have lower ATP yield than CO growth, but H_2/CO_2 growth is still expected to be ATP positive assuming acetate is formed (Mock et al., 2015).

Indeed, *C. ljungdahlii* has already been engineered to produce acetone, butanol, and butyrate (among many others), and acetate has been an ever-present product in those studies. We have successfully reduced acetate production by $>$ 80% via targeted gene knockout, which should help inform further work in engineering new products in acetogens and increasing yield of desired products. In this work, we engineered *C. ljungdahlii* and improved yield of ethanol from ∼10% to over 80%. In the process, we clarified the role of several genes involved in acetyl-CoA metabolism and showed that substrate phosphorylation via *pta* is unnecessary for autotrophic growth on CO (**Figure 5**). We also identify gaps in our knowledge of acetogens and highlight important areas of study in acetogens for further research in both fundamental understanding and applied metabolic engineering.

AUTHOR'S NOTE

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by DOE and Bioenergy Technology Office (BETO). The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

JL, P-CM, YG, and ZR led the research. JL and JH designed the experiments. JL, JH, JJ, CU, LM, and WX performed the experiments. JL, WX, and P-CM wrote the article. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the US Department of Energy (DOE) Bioenergy Technology Office (BETO) and National Renewable Energy Laboratory (NREL) under Contract DE-AC36-08-GO28308.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.560726/full#supplementary-material

REFERENCES

Aklujkar, M., Leang, C., Shrestha, P. M., Shrestha, M., and Lovley, D. R. (2017). Transcriptomic profiles of clostridium ljungdahlii during lithotrophic growth with syngas or H 2 and CO 2 compared to organotrophic growth with fructose. *Sci. Rep.* 7, 1–14. doi: 10.1038/s41598-017-12712-w

Bertsch, J., and Müller, V. (2015). Bioenergetic constraints for conversion of syngas to biofuels in acetogenic bacteria. *Biotechnol. Biofuels* 8:210. doi: 10.1186/s13068-015-0393-x

Bruno-Barcena, J. M., Chinn, M. S., and Grunden, A. M. (2013). Genome sequence of the autotrophic acetogen clostridium autoethanogenum JA1-1 Strain DSM 10061, a producer of ethanol from carbon monoxide. *Genome Announc.* 1:e00628-13. doi: 10.1128/genomeA.00628-13

Cañadas, I. C., Groothuis, D., Zygouropoulou, M., Rodrigues, R., and Minton, N. P. (2019). RiboCas: a universal CRISPR-based editing tool for clostridium. *ACS Synth. Biol.* 8, 1379–1390. doi: 10.1021/acssynbio.9b00075

Cooksley, C. M., Zhang, Y., Wang, H., Redl, S., Winzer, K., and Minton, N. P. (2012). Targeted mutagenesis of the clostridium acetobutylicum acetone–butanol–ethanol fermentation pathway. *Metab. Eng.* 14, 630–641. doi: 10.1016/j.ymben.2012.09.001

Heider, J., Ma, K., and Adams, M. W. (1995). Purification, characterization, and metabolic function of tungsten-containing aldehyde ferredoxin oxidoreductase from the hyperthermophilic and proteolytic archaeon thermococcus strain ES-1. *J. Bacteriol.* 177, 4757–4764.

Huang, H., Chai, C., Li, N., Rowe, P., Minton, N. P., Yang, S., et al. (2016). CRISPR/Cas9-based efficient genome editing in clostridium ljungdahlii, an autotrophic gas-fermenting bacterium. *ACS Synth. Biolo.* 5, 1355–1361. doi: 10.1021/acssynbio.6b00044

Huber, C., Skopan, H., Feicht, R., White, H., and Simon, H. (1995). Pterin cofactor, substrate specificity, and observations on the kinetics of the reversible tungsten-containing aldehyde oxidoreductase fromclostridium thermoaceticum. *Arch. Microbiol.* 164, 110–118. doi: 10.1007/BF02525316

Jones, S. W., Fast, A. G., Carlson, E. D., Wiedel, C. A., Au, J., Antoniewicz, M. R., et al. (2016). CO2 fixation by anaerobic non-photosynthetic mixotrophy for improved carbon conversion. *Nat. Commun.* 7:12800.

Kim, G., Azmi, L., Jang, S., Jung, T., Hebert, H., Roe, A. J., et al. (2019). Aldehyde-alcohol dehydrogenase forms a high-order spirosome architecture critical for its activity. *Nat. Commun.* 10, 1–11. doi: 10.1038/s41467-019-12427-8

Klotzsch, H. R. (1969). "[59] phosphotransacetylase from *Clostridium kluyveri*," in *Methods in Enzymology*, Vol. 13, eds N. O. Kaplan and S. P. Colowick (Cambridge, MA: Academic Press.), 381–386. doi: 10.1016/0076-6879(69)13065-7

Köpke, M., Held, C., Hujer, S., Liesegang, H., Wiezer, A., Wollherr, A., et al. (2010). Clostridium ljungdahlii represents a microbial production platform based on syngas. *Proc. Natl. Acad. Sci. U.S.A.* 107, 13087–13092. doi: 10.1073/pnas.1004716107

Leang, C., Ueki, T., Nevin, K. P., and Lovley, D. R. (2013). A genetic system for clostridium ljungdahlii: a chassis for autotrophic production of biocommodities and a model homoacetogen. *Appl. Environ. Microbiol.* 79, 1102–1109. doi: 10.1128/AEM.02891-12

Liew, F., Henstra, A. M., Köpke, M., Winzer, K., Simpson, S. D., and Minton, N. P. (2017). Metabolic engineering of clostridium autoethanogenum for selective alcohol production. *Metab. Eng.* 40, 104–114. doi: 10.1016/j.ymben.2017.01.007

Liew, F., Henstra, A. M., Winzer, K., Köpke, M., Simpson, S. D., and Minton, N. P. (2016). Insights into CO2 fixation pathway of clostridium autoethanogenum by targeted mutagenesis. *mBio* 7:e00427-16. doi: 10.1128/mBio.00427-16

Liu, Z.-Y., Jia, D. C., Zhang, K. D., Zhu, H. F., Zhang, Q., Jiang, W. H., et al. (2020). Ethanol metabolism dynamics in clostridium ljungdahlii grown on carbon monoxide. *Appl. Environ. Microbiol.* 86:e00730-20. doi: 10.1128/AEM.00730-20

Lo, J., Zheng, T., Hon, S., Olson, D. G., and Lynd, L. R. (2015). The bifunctional alcohol and aldehyde dehydrogenase gene, AdhE, is necessary for ethanol production in clostridium thermocellum and thermoanaerobacterium saccharolyticum. *J. Bacteriol.* 197, 1386–1393. doi: 10.1128/JB.02450-14

Mahamkali, V., Valgepea, K., de Souza, R., Lemgruber, P., Plan, M., Tappel, R., et al. (2020). Redox controls metabolic robustness in the gas-fermenting acetogen clostridium autoethanogenum. *Proc. Natl. Acad. Sci. U.S.A.* 117, 13168–13175. doi: 10.1073/pnas.1919531117

Marcano-Velazquez, J. G., Lo, J., Nag, A., Maness, P. C., and Chou, K. J. (2019). Developing riboswitch-mediated gene regulatory controls in thermophilic bacteria. *ACS Synth. Biol.* 8, 633–640. doi: 10.1021/acssynbio.8b00487

Mock, J., Zheng, Y., Mueller, A. P., Ly, S., Tran, L., Segovia, S., et al. (2015). Energy conservation associated with ethanol formation from H $_2$ and CO $_2$ in clostridium autoethanogenum involving electron bifurcation. *J. Bacteriol.* 197, 2965–2980. doi: 10.1128/JB.00399-15

Molitor, B., Marcellin, E., and Angenent, L. T. (2017). Overcoming the energetic limitations of syngas fermentation. *Curr. Opin. Chem. Biol. Mech. Biol. Energy* 41, 84–92. doi: 10.1016/j.cbpa.2017.10.003

Nagarajan, H., Sahin, M., Nogales, J., Latif, H., Lovley, D. R., Ebrahim, A., et al. (2013). Characterizing acetogenic metabolism using a genome-scale metabolic reconstruction of clostridium ljungdahlii. *Microb. Cell Factor.* 12:118. doi: 10.1186/1475-2859-12-118

Napora-Wijata, K., Strohmeier, G. A., and Winkler, M. (2014). Biocatalytic reduction of carboxylic acids. *Biotechnol. J.* 9, 822–843. doi: 10.1002/biot.201400012

Nissen, L. S., and Basen, M. (2019). The emerging role of aldehyde:ferredoxin oxidoreductases in microbially-catalyzed alcohol production. *J. Biotechnol.* 306, 105–117. doi: 10.1016/j.jbiotec.2019.09.005

Richter, H., Molitor, B., Wei, H., Chen, W., Aristilde, L., and Angenent, L. T. (2016). Ethanol production in syngas-fermenting clostridium ljungdahlii is controlled by thermodynamics rather than by enzyme expression. *Energy Environ. Sci.* 9, 2392–2399. doi: 10.1039/C6EE01108J

Schuchmann, K., and Müller, V. (2014). Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria. *Nat. Rev. Microbiol.* 12, 809–821. doi: 10.1038/nrmicro3365

Tan, Y., Liu, Z.-Y., Liu, Z., and Li, F.-L. (2015). Characterization of an acetoin Reductase/2,3-butanediol dehydrogenase from clostridium ljungdahlii DSM 13528. *Enzyme Microb. Technol.* 79–80, 1–7. doi: 10.1016/j.enzmictec.2015.06.011

Tremblay, P. L., Zhang, T., Dar, S. A., Leang, C., and Lovley, D. R. (2013). The Rnf complex of clostridium ljungdahlii is a proton-translocating ferredoxin:NAD+ oxidoreductase essential for autotrophic growth. *mBio* 4:e00406-12. doi: 10.1128/mBio.00406-12

Ueki, T., Nevin, K. P., Woodard, T. L., and Lovley, D. R. (2014). Converting carbon dioxide to butyrate with an engineered strain of clostridium ljungdahlii. *mBio* 5:e01636-14. doi: 10.1128/mBio.01636-14

Whitham, J. M., Tirado-Acevedo, O., Chinn, M. S., Pawlak, J. J., and Grunden, A. M. (2015). Metabolic response of clostridium ljungdahlii to oxygen exposure. *Appl. Environ. Microbiol.* 81, 8379–8391. doi: 10.1128/AEM.02491-15

Woolston, B. M., Emerson, D. F., Currie, D. H., and Stephanopoulos, G. (2018). Rediverting carbon flux in clostridium ljungdahlii using CRISPR interference (CRISPRi). *Metab. Eng.* 48, 243–253. doi: 10.1016/j.ymben.2018.06.006

Yao, S., and Mikkelsen, M. J. (2010). Identification and overexpression of a bifunctional aldehyde/alcohol dehydrogenase responsible for ethanol production in thermoanaerobacter mathranii. *J. Mol. Microbiol. Biotechnol.* 19, 123–133. doi: 10.1159/000321498

Zhu, H.-F., Liu, Z.-Y., Zhou, X., Yi, J.-H., Lun, Z.-M., Wang, S.-N., et al. (2020). . Energy conservation and carbon flux distribution during fermentation of CO or H2/CO2 by clostridium ljungdahlii. *Front. Microbiol.* 11:416. doi: 10.3389/fmicb.2020.00416

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: info@frontiersin.org | +41 21 510 17 00

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership