BIOSENSOR-ASSISTED METABOLIC ENGINEERING AND BIOSENSOR DEVELOPMENT

by

CHENYI LI

(Under the Direction of Yajun Yan)

ABSTRACT

Establishing microbial cell factories using metabolic engineering strategies or synthetic biology tools has enabled fast and cost-effective production of a series of valuable natural and synthetic compounds. To further improve the productivities of microbial cell factories, fine-tuning the gene expression, through either intelligently enhancing the target production or effectively inhibiting the competitive pathways with the assistance of genetically-encoded biosensors, has been demonstrated to be efficient in constructing highly efficient microbial cell factories. Herein, we explored the biosensor-assisted fine-tuning of gene expression in microbial cell factories to improve the biosynthesis of target compounds, and the further development and optimization of biosensors with desired dynamic performance and expanded substrate scopes. With the understanding of the replication mechanism for ColE1-derived plasmid origins, we established the dynamic pathway regulation at gene copy level to fine-tune the gene expression through dynamically controlling the plasmid replication with the *p*-coumaric acid responsive transcriptional repressor PadR derived from *Bacillus subtilis 168*, which enabled a significantly improved *p*-coumaric acid production in *Escherichia coli*. Then, we targeted the central

metabolism of *E. coli* using CRISPRi with the mismatched sgRNA arrays to finely rewire the carbon flux to the production of *p*-coumaric acid or butyric acid. The PadR and the butyric acid responsive regulator HpdR were applied to screen the high-producers of *p*-coumaric acid and butyric acid in a high-throughput manner, respectively. To further improve the applicability of the transcriptional factors-based biosensors, we broadened the substrate scope of the phenolic acid-responsive regulator PadR and enabled the recognition of seven aromatic compounds with different structures. Our work demonstrated the enormous potential of utilizing transcriptional factors-based biosensors in constructing highly efficient microbial cell factories, and the newly developed regulator variants in our study that harbor altered substrate scopes would also be particularly beneficial for future applications in metabolic engineering and synthetic biology.

INDEX WORDS: Metabolic engineering, synthetic biology, dynamic regulation, genetic circuits, high-throughput screening, CRISPRi, *p*-coumaric acid, PadR, protein engineering, transcriptional regulators, biosensors, *Escherichia coli*

BIOSENSOR-ASSISTED METABOLIC ENGINEERING AND BIOSENSOR DEVELOPMENT

by

CHENYI LI

BS, Beijing Institute of Technology, China, 2017

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2022

BIOSENSOR-ASSISTED METABOLIC ENGINEERING AND BIOSENSOR DEVELOPMENT

by

CHENYI LI

Major Professor:	Yajun Yan
Committee:	K.C.Das
	Y. George Zheng
	Cheryl Gomillion
	William Kisaalita

Electronic Version Approved:

Ron Walcott

Vice Provost for Graduate Education and Dean of the Graduate School

The University of Georgia

May 2022

DEDICATION

This Dissertation is dedicated to my supportive parents who never stop giving of themselves in countless ways; To my beloved wife Tian Jiang who encouraged me when I was in the valley of darkness with the light of hope and her love; To my respectful grandmother who cares me so much with her kindness and love.

ACKNOWLEDGEMENTS

I would like to express my gratitude for my major advisor Dr. Yajun Yan for his unreserved support and assistance during the past five years. Starting the Ph.D. journey directly after obtaining my Bachelor's degree is challenging for me, but Dr. Yan provided me with professional advices not only in my scientific research and but also in developing my career. I really enjoy the discussion with Dr. Yan and I am excited every time we came up with a new research idea. The trainings I received from Dr. Yan made me become a rigorous scientific researcher and paved my way for my future career. I also would like to say thank you to my committee members, Dr. Y. George Zheng, Dr. Cheryl Gomillion, Dr. William Kisaalita, and Dr. K.C.Das, for their insightful suggestions and advices in revising my dissertation. I also want to thank all the members in the lab. It has been a wonderful journey working with you. In particular, I would like to thank Dr. Wang Jian for mentoring me when I was in my early stage of research. Last but not least, a special thankfulness goes to my wife Tian Jiang, who is also my colleague in the lab, for her love and support which inspired me during challenging times.

Table of Contents

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW [1,2]

---

1 Li, C., Jiang, T., Li, M., Zou, Y., Yan, Y., 2021. Fine-tuning gene expression for improved biosynthesis of natural products: From transcriptional to post-translational regulation. Biotechnology Advances, 107853. Reproduced here with permission of Elsevier.

2 Li, C., Zhang, R., Wang, J., Wilson, L.M., Yan, Y., 2020. Protein engineering for improving and diversifying natural product biosynthesis. Trends in biotechnology 38(7), 729-744. Reproduced here with permission of Elsevier.

1.1 Microbial biosynthesis of value-added compounds

Value-added chemicals, ranging from small molecules like biofuels, biopolymers, terpenoids, flavonoids, alkaloids, or coumarins to large chemical moieties such as polyketides, nonribosomal peptides, or ribosomally synthesized and post-translationally modified peptides (RiPPs), are widely used as nutritionally, pharmaceutically, or industrially important components in biomedicine applications, dietary supplements, cosmetic industries, and polymer synthesis (Chen, J. et al., 2015; Kusumawati and Indrayanto, 2013; Liu and Nielsen, 2019; Rodrigues et al., 2016; Ververidis et al., 2007). The rapidly growing demand for such chemicals (Wang, R. et al., 2020; Yang et al., 2020) motivates the exploration into cost-effective manufacturing of these high-value compounds. The past decades have witnessed a rapid development of establishing microbial cell factories for synthesizing such compounds (Cravens et al., 2019; Zha et al., 2020; Zhang et al., 2018). With the expanding genetic toolsets, some amenable microbial hosts with clear genetic backgrounds and strong production capacities were explored for natural product synthesis, including *Escherichia coli*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, and *Corynebacterium glutamicum* (Calero and Nikel, 2019; Fletcher et al., 2016). Along with these traditional hosts, a series of unconventional hosts with special production capability, such as *Cyanobacteria*, *Streptomyces*, *Yarrowia Lipolytica*, and *Klebsiella pneumoniae*, have also been exploited (Bekker et al., 2014; Carroll et al., 2018; Rhie et al., 2019; Zhu and Jackson, 2015). Fast dividing microorganisms (0.5–3 h doubling time) can enable shortened production cycles. The purification workflow can also be potentially simplified since only a limited range of similar metabolites are endogenously produced in commonly used industrial microbial hosts (Ehrenworth and Peralta-Yahya, 2017). In addition to the aforementioned benefits, the ease of cultivation and manipulation, as well as potentially lower production costs, have attracted more attention in developing efficient

microbial cell factories for biosynthesis of natural products (Ehrenworth and Peralta-Yahya, 2017; Wang, J. et al., 2018). Thus, exciting achievements have been made in the microbial production of biofuels (Adegboye et al., 2021; Choi et al., 2020), biopolymers (Moradali and Rehm, 2020; Sohn et al., 2020), small aromatic compounds (Chouhan et al., 2017; Liu, Q. et al., 2019; Wang, J. et al., 2018), coumarins (Lin et al., 2013a; Lin et al., 2013b; Yang et al., 2015), terpenoids (Ajikumar et al., 2010; Keasling, 2010), flavonoids (Biggs et al., 2016; Liu, X. et al., 2018; Palmer et al., 2019), and polyketides (Wang, W. et al., 2020; Weissman, 2016), in the past decades.

Despite the advantages of utilizing microbial hosts to synthesize natural products, there are many challenges in constructing efficient microbial cell factories for producing these value-added compounds. First, the production of many natural compounds requires the consumption of precursors from microbial central metabolism. The synthesis of these compounds can cause competition between cell growth and production, which may negatively affect the productivities of microbial chassis (Wu et al., 2016). Additionally, the biosynthetic pathways for natural products usually involve metabolites that do not naturally exist in the host, and the accumulation of these chemicals can be toxic to the chassis and thus hinder the production efficiencies (Jiang et al., 2020). Moreover, some reactions in the synthesis of natural compounds require the input of co-factors or redox molecules. The consumption of these biomolecules may cause imbalances of co-factors, which can also hamper cell growth and production.

To address these problems, a series of metabolic engineering strategies, including adaptive laboratory evolution (Sandberg et al., 2019), protein/enzyme engineering (Li et al., 2020), compartmentalization (Abernathy et al., 2017; Hammer and Avalos, 2017), scaffold engineering

(Park et al., 2018), co-factor engineering (Black et al., 2020; Zhao, X. et al., 2015), dynamic pathway regulations (Chen and Liu, 2018; Jones, J. A. et al., 2015; Shen et al., 2019), and combinatorial optimizations (Galanie et al., 2015; Jeschek et al., 2016; Zhao, S. et al., 2015), have been widely applied to improve the biosynthesis of natural products. Besides these strategies, fine-tuning the gene expression, either via controlling the biosynthetic pathways of natural products or regulating the central metabolisms of the microbial chassis, has been proven to be a promising and efficient way to address these challenges. When the microbial cell factories are reasonably regulated, the engineered strains can achieve better cell growth and higher titers. Gene expression can be typically controlled at three levels: the DNA level, (transcriptional level), RNA level (post-transcriptional and translational level), protein level (post-translational level), and multilevel regulation. In this review, we focus on recent advances in fine-tuning gene expression at these three levels to improve the biosynthesis of natural products. Commonly used toolsets in each level are reviewed and outlooks for future direction in this area are provided.

1.2 Fine-tuning the gene expression for improved biosynthesis

Due to the stochasticity and complex regulations lying in microbial systems, sometimes it is challenging to achieve high productivities of natural products in microbial hosts (Kaern et al., 2005; Rugbjerg et al., 2018). The synthetic pathways for natural products usually involve multiple steps, and some metabolites may be harmful to microbial hosts (Dahl et al., 2013; Yang et al., 2020). Therefore, controlling the gene expression to reduce the competition between cell growth and production, avoid or delay the accumulation of toxic metabolites, and alleviate the feedback inhibition, is necessary for achieving high titers of value-added natural compounds. Plenty of efficient regulation toolboxes or novel regulation strategies have been established and well-

characterized in recent years, which greatly promoted the fine-tuning of gene expression in natural product biosynthesis (**Table 1.1**). These strategies were divided into three levels and discussed in this paper based on which process they regulate.

1.2.1 Regulation at DNA level

The first step in protein expression is the transcription of DNA to RNA, where the promoters, RNA polymerase, sigma factors, and allosteric transcriptional factors (aTF) are typically involved (Clancy, 2008). Plenty of transcriptional regulation strategies and toolsets have been developed in the past decades, such as promoter engineering (Jones, J. Andrew et al., 2015; Lim Chin et al., 2011; Zhou et al., 2019), allosteric transcriptional factors (Mahr and Frunzke, 2016), CRISPR-Cas based gene regulation (Gilbert et al., 2014), TALE-mediated (transcriptional activator-like effector) transcriptional control (Crocker and Stern, 2013), zinc finger-mediated transcriptional control (Laity et al., 2001), and DNA aptamers (Jang et al., 2017; Wang, Jing et al., 2017). Regulation at the DNA level is to control the transcription of target genes, which is the most common strategy in pathway control due to the ease of manipulation. As transcription is the start of protein synthesis, inhibition on transcription would also negatively affect the translation and post-translation modifications. Hence, when applying DNA level regulation, the whole protein synthesis process can be regulated with just a simple control.

Promoters directly affect the transcription rate, and thus are common targets used for engineering biosynthesis of natural products (Alper et al., 2005; Kang et al., 2016). Constructing promoter variants with different transcriptional activities and applying them to fine tune the biosynthetic pathways are becoming increasingly feasible with the development of biosensors and high-

throughput screening techniques (Lee et al., 2013; Xiu et al., 2017; Zhou et al., 2019) . Naringenin is a valuable natural product with great therapeutic potential (Salehi et al., 2019). The microbial synthesis of naringenin is achieved via converting the tyrosine to *p*-coumaric acid and then *p*-coumaroyl-CoA by TAL and 4CL, followed by condensation with three molecules of malonyl-CoA and isomerization to form naringenin by CHS and CHI (Miyahisa et al., 2006). These four genes need to be fine-tuned to balance the carbon flux so that the naringenin production can be enhanced. Therefore, an iterative high-throughput balancing (IHTB) strategy was established to thoroughly fine-tune the naringenin biosynthetic pathway (Zhou et al., 2019). Constitutive promoters with gradient strengths were randomly picked to control the expression of pathway genes, resulting a library with differed naringenin production capacities. After screening of over 1,200 candidates, the metabolic flux of the naringenin synthetic pathway was appropriately balanced, with the final naringenin titer reaching 191.9 mg/L (Zhou et al., 2019). Similar strategy has also been demonstrated in engineering the naringenin production in *S.cerevisiae* (Wang, R. et al., 2019). Promoter engineering is a common strategy to fine-tune the gene expression at DNA level, but the large library size usually make the characterization processes rely on the high-throughput screening techniques (Lee et al., 2013). Thus, future engineering and development of efficient high-throughput detection techniques would benefit and accelerate the processes of developing high-producers using promoter engineering.

Allosteric transcriptional factor (aTF) is the most commonly used tool in controlling gene expression at the DNA level (Latchman, 1993). aTF is a regulatory protein that can control, either by enhancing (transcriptional activator) or inhibiting (transcriptional repressor), the transcription by binding to a specific DNA sequence in the promoter region (**Fig. 1.1a**). This region often

overlaps with the -35 or -10 box that is for RNA polymerase binding (Karin, 1990; Shen et al., 2019). Due to the ligand-sensing ability of aTFs, they were usually used to enable dynamic and real-time fine-tuning of the gene expression for natural product synthesis. The first dynamic control was achieved in lycopene biosynthesis by Liao's group (Farmer and Liao, 2000). After that, this strategy was wildly applied in the synthesis of isoprenoid, fatty acids, glucaric acid, flavonoids, and aromatic compounds (Dahl et al., 2013; Dinh and Prather, 2019; Doong et al., 2018; Gupta et al., 2017; Liang et al., 2020; Xu et al., 2014; Yang et al., 2018). The regulatory circuits in these studies are powered by the inducible aTFs to improve the production of the target compounds and reduce the competition between cell growth and production. For example, vanillin is a valuable natural product with a projected market size of USD 734.5 million in 2025 (Global Vanillin Market Size). Despite the high value of this compound, the microbial synthesis of vanillin can compete with cell growth, and high pathway expression can lead to growth defects and plummeted productivity as accumulated vanillin displays antimicrobial activity (Liang et al., 2020). To solve these problems, a dynamic pathway control circuit powered by the aTF HucR variants was designed and applied to vanillin biosynthesis. The delayed pathway expression significantly improved the cell growth, which in turn resulted in an increased vanillin production to a titer of over 1.5 g/L (Liang et al., 2020). Plant-derived synthetic aTFs were implemented in improving the production of β-ionone and naringenin in *Saccharomyces cerevisiae* (Naseri et al., 2019). Through balancing the pathway expression, a 4.2-fold increase in β-ionone production was achieved. The improved production of naringenin was also detected by a biosensor (Naseri et al., 2019). While the aTF is a powerful tool to regulate gene expression, the limited substrate scopes and narrow dynamic ranges of natural aTFs may constrain their application potentials in dynamic gene regulations. Mining new aTFs or engineering existing aTFs to respond to new compounds and

improving the dynamic range of existing aTFs (Jiang et al., 2021) are becoming increasingly important to expand the applicability of aTF in fine-tuning natural product biosynthesis (Li et al., 2020).

In addition to the aTFs and promoter engineering, CRISPR (clusters of regularly interspaced short palindromic repeats) based gene regulation is also a trending tool for improving the natural product biosynthesis (**Fig. 1.1b**). By using the nuclease-deficient CRISPR protein (dCas9), transcriptional activation (CRISPRa) (Bester et al., 2018; Fontana et al., 2020; Konermann et al., 2015) and interference (CRISPRi) (Gilbert et al., 2014; Wu et al., 2020; Zhao et al., 2021) have been demonstrated in various hosts for efficient transcriptional control (Lian et al., 2019). An important advantage of using CRISPR-based gene expression control is the ease of achieving simultaneous multiplex regulation. For example, the biosynthesis of naringenin requires the input of malonyl-CoA and *p*-coumaric acid derived from tyrosine. Thus, CRISPRi was applied in fine-tuning multiple target genes involved in the central metabolism of *E. coli* to rewire the carbon flux to the precursor of naringenin (Wu et al., 2015). The simultaneous inhibition of *fabF*, *fumC*, *fabB*, *sucC*, and *adhE* resulted in a 7.6-fold increase in naringenin titer to 421.6 mg/L without significantly altering the final biomass accumulation (Wu et al., 2015). In another study, the simultaneous down-regulation of multiple competing pathways using CRISPRi combined with fermentation optimization enhanced the production of β-amyrin in *S. cerevisiae* to 156.7 mg/L, a 44.3% increase in the titer compared with the unregulated strain (Ni et al., 2019). CRISPRi based control was also employed to alleviate the feedback inhibition lying in the biosynthetic pathway (Sander et al., 2019). Amino acid biosynthesis is known to be tightly regulated in microbial cells. Sander and colleagues applied CRISPRi in downregulating the feedback inhibition mediated by ArgR in *E.*

coli and improved the cell growth and production of the arginine-overproducing strain (Sander et al., 2019). As for CRISPRa, this system has been well-established in engineering the biosynthesis of natural products such as violacein (Liu, Y. et al., 2019) and epothilone (Fontana et al., 2020; Peng, R. et al., 2018). In 2021, CRISPRa was ported and engineered to function in *Pseudomonas putida* (Kiattisewee et al., 2021). With systematic analysis and optimizations, CRISPRa was applied to regulate the biosynthesis of biopterin and mevalonate, resulting in over 5-fold and 40-fold increase in titers, respectively (Kiattisewee et al., 2021). Besides the application of simple inhibition or activation, CRISPR-based multiplex regulation for the production of β-carotene has also been demonstrated in yeast (Lian et al., 2017). By u*p*-regulating the rate-limiting step catalyzed by *HMG1*, inhibiting the *ERG9* that competes the β-carotene biosynthesis for endogenous sterol generation, and deleting a stress responsive transcriptional regulator *ROX1*, the titer of β-carotene increased by up to 2.8-fold compared to the control strain (Lian et al., 2017). As a trending regulation method, CRISPR-based gene expression control demonstrated great potentials in improving the biosynthesis of natural products. Future optimization of the CRISPR system such as minimizing the off-target effect and engineering relaxed PAM specificity may continue to drive the application of this regulation method in metabolic engineering (Chatterjee et al., 2020; Collias and Beisel, 2021).

DNA aptamers that can bind to specific ligands are also used to construct genetic circuits which control the gene expression at the transcriptional level (**Fig. 1.1c**) (Pfeiffer and Mayer, 2016; Ruscito and DeRosa, 2016). The binding of ligand would cause unwinding of double stranded DNA, which would allow easier recognition and binding by RNA polymerases if the aptamer is placed near the promoter region, leading to the enhanced expression of downstream genes at the

transcriptional level (Wang, Jing et al., 2017) (Fig. 1c). The well-characterized thrombin-bound aptamer was engineered and applied into 2′-fucosyllactose synthesis in *B. subilits* (Deng et al., 2019). The introduction of DNA aptamer-mediated control enabled a titer of 511 mg/L, a 22.3-fold increase over the parental strain (Deng et al., 2019). While the regulation mediated by DNA aptamers is easy to achieve, the limited set of available DNA aptamers significantly hindered the further applications of this regulatory component. Therefore, larger libraries of signal molecules and responsive components are required in the future for multiplex, fine-tuning natural product biosynthesis via DNA aptamers.

While not as common as other regulation tools at DNA level, the TALE-mediated regulation was also applied in engineering biosynthesis of natural product. In 2016, TALE-fused enzymes were gathered around the TALE DNA scaffolds and enriched the local enzyme concentrations, which led to a 9.6-fold increase in inole-3-acetic acid production (Zhu et al., 2016).

Applying transcriptional regulation strategies for tunable gene expression at DNA level, which is the start of protein expression, can control the synthesis of target proteins with minimal effort. Thus, it is widely used in metabolic pathway control, especially for natural products. While this type of regulation is efficient and easy to achieve, the stop of transcription did not necessarily represent a decreased availability of proteins, because the previously transcribed mRNA can still be translated to proteins by the ribosomes, and there are also residual enzymes/proteins in the cells that can still function as normal (Sekar et al., 2016; Yuan and Ching, 2015). Therefore, if only the transcriptional control is applied, there will often be a time lag for the cells to exhibit the regulation effects due to the above reasons, and this type of control sometimes can be less effective in some

microorganisms if the vast majority of intracellular proteins are long-lived, such as *S. cerevisiae* and *S. pombe* (Christiano et al., 2014; Yuan and Ching, 2015).

1.2.2 Regulation at RNA level

Messenger RNA (mRNA) is the bridge connecting DNA and proteins. RNA-level regulations are usually achieved by blocking the access of ribosome to the ribozyme binding site (RBS) in mRNA or decreasing the mRNA availabilities to hinder the translation and thus decrease the protein expression level. The commonly used tools include small regulatory RNAs, antisense RNAs, and riboswitches (Wang and Cirino, 2016).

Small regulatory RNA (sRNA) is a type of short non-coding RNA that natively exists in many microorganisms (**Fig. 1.2a**). It was first discovered in *E. coli* in 1984 (Mizuno et al., 1984; Svensson and Sharma, 2016; Vogel and Wagner, 2007). sRNA consists of two major components: one is the scaffold that is used to stabilize the RNA structure, the other is a short binding sequence (usually 20-200 bp) that can bind with target mRNA by complementary pairing. When the target region includes the RBS, the binding of sRNA can both block the access of RBS and also accelerate the degradation of mRNA to decrease the mRNA concentration. When the RBS is not in the sRNA target region, the binding of sRNA would only cause a decreased availability of mRNA (Aiba, 2007; Culver, 2001). Notably, the degradation of the RNA complex upon binding with target mRNA requires the recruitment of hfq protein, an abundant bacterial RNA binding protein that is physiologically important in interacting with sRNAs and facilitating their antisense interactions with their targets (Holmqvist et al., 2016; Tsui et al., 1994). In 2013, Sang Yup Lee and coworkers established a standard workflow to engineer the small regulatory RNA (sRNA) for regulating

biosynthesis pathways (Na et al., 2013; Yoo et al., 2013). By screening over 100 candidate sRNAs, they select the micC as the optimal scaffold to design and engineer synthetic sRNAs. The designed sRNAs were applied to inhibit the competing pathways of lysine biosynthesis and rewire the carbon flux from central metabolism to cadaverine production. They screened a library containing 130 synthetic sRNAs, and a 55% increase in cadaverine titer was achieved when repressing *murE* (Na et al., 2013). In 2019, the same group developed an expanded synthetic sRNA expression platform to enable rapid, multiplexed, and genome-scale target gene knockdown by improving its compatibility for commonly used antibiotic markers and origin of replications (ORI, e.g. p15A, ColE1, and pSC101) (Yang et al., 2019). Using this strategy, they demonstrated increased productions of L-threonine (22.9 g/L), L-proline (54.1 g/L), crude violacein (5.19 g/L), and indigo (135 mg/L) with corresponding engineered strains that were developed from other studies (Yang et al., 2019).

Antisense RNA (asRNA) is also a common tool used for RNA-level regulation (**Fig. 1.2b**). The function of asRNA is very similar to sRNA: the RNA-level regulation mediated by asRNA is also achieved by complementary pairing with target mRNA, but unlike sRNA, the degradation of the asRNA-mRNA complex does not require the recruitment of hfq protein (Pelechano and Steinmetz, 2013). Instead, it is usually achieved by the native RNA degradation enzymes, such as RNAse III in Gram-positive bacteria (Lasa et al., 2011). Most notable applications of asRNAs are in improving the biosynthesis of flavonoids. In 2015, the antisense RNA was engineered to rewire carbon flux from central metabolisms to flavonoids biosynthesis (Wu et al., 2014; Yang et al., 2015). The asRNA was equipped with a 100 bp artificial stem loop in these two studies to improve its stability. The fatty acid synthesis pathway was inhibited by the engineered asRNA to enhance

the availability of malonyl-CoA, an important precursor for the synthesis of a series of flavonoids. In Yang's paper, the asRNA-mediated down-regulation was applied towards *fabD*, which converts the malonyl-CoA to fatty acids. This regulation resulted in a 4.5-fold increase of the malonyl-CoA availability, which in turn enabled a 1.53-fold, 1.70-fold, and 2.53-fold increase in the production of naringenin, resveratrol, and 4-hydroxycoumarin, respectively (Yang et al., 2015). In Wu's paper, the inhibition of *fabB*/*fabF* using asRNA improved the production titer of naringenin by 4.31-fold (391 mg/liter) (Wu et al., 2014). These examples demonstrated that asRNA is a powerful tool for fine-tuning the central metabolism of bacteria to improve the biosynthesis of natural products.

Unlike the sRNA and asRNA, regulation mediated by riboswitch does not affect the intracellular mRNA availability. Riboswitch is a type of regulator that controls the gene expression via conformational change of RNA secondary structures (**Fig. 1.2c**) (Mandal et al., 2003). There are two major components in a typical riboswitch, one is the ligand binding domain, where the binding of target compounds can cause a conformational change to the secondary structure. The other is the regulation domain, where the RNA secondary structure can block the access of ribozyme to RBS and thus hinder the translation. While this type of regulator is wildly distributed in microorganisms, they were often used as biosensors to screen high-producers (Lee and Oh, 2015; Wang et al., 2015; Yang et al., 2013), and only a handful of examples have achieved gene regulation via riboswitches. Deng and co-workers achieved simultaneous u*p*-regulation and down-regulation of *in vivo* gene expression by engineering the ligand thrombin-bound aptamer (Deng et al., 2019). Using this, they enhanced the expression of 2'-fucosyllactose pathway genes but reduced the unnecessary induce of biofilm formation in *Bacillus subtilis*. The final strain reached a titer of 674 mg/L when the bi-functional regulation was applied, which was 27.3-fold that of the

un-controlled original strain (Deng et al., 2019). This type of dual-control has also been demonstrated in another study. The lysine-responsive riboswitches were applied in *Corynebacterium glutamicum* for metabolic control (Zhou and Zeng, 2015). The screened lysine-ON riboswitches were used to control the expression of *lysE* gene (encoding a lysine transport protein), achieving dynamic u*p*-regulation of lysine transport in the *C. glutamicum*. On the other hand, a lysine-OFF riboswitch was applied to dynamically down-regulate the expression of citrate synthase. The combined dual-control enabled an 89% increase in lysine yield, and also demonstrated the synergetic effect of lysine-ON and -OFF riboswitches for improving lysine production (Zhou and Zeng, 2015).

Instead of using regulatory proteins compared with transcriptional regulation, the RNA-level regulations are usually achieved via non-coding RNAs. Introducing such non-coding RNAs can often bring less cell burden to the host than using regulatory proteins. First, non-coding RNAs (usually less than 200 nt) are often shorter than normal mRNA (usually over 200 nt). Moreover, compared with proteins that need to be synthesized from amino acids, RNAs do not consume amino acids for synthesis, which can effectively save some resources for cell growth or production. However, similar to DNA-level control, even though the RNA-level regulation can decrease the availability of mRNA or hinder the translation, this cannot quickly transition to a decreased protein level, because the residual proteins in cells usually would not be quickly degraded so they can still function as normal (Sekar et al., 2016; Yuan and Ching, 2015). Also, when applying RNA-level regulation, the synthesis of mRNA (transcription) is not controlled, which means the mRNA can still be synthesized continuously, which can cause a waste of cell resources for unnecessary synthesis of mRNAs.

1.2.3 Regulation at protein level

One of the essential goals for gene regulation is to adjust the intracellular protein concentrations. Therefore, regulation at the protein level provides a more direct way to achieve fast and accurate control of the protein availabilities and thus has attracted increasing attention in recent years (Cameron and Collins, 2014; Chung et al., 2015). Unlike the regulations at DNA or RNA level that can achieve both u*p*- and down-regulation (Lian et al., 2017; Zhou and Zeng, 2015), most studies with protein level control usually focused on controllable degradation (down-regulation). Down-regulation by tunable protein degradation reduces the intracellular protein availabilities to a low level, which would be helpful in shutting down competing pathways or rewiring carbon fluxes. The major ways to achieve tunable protein degradation are through engineering protein degradation tags or degrons, as well as controllable expression of proteases (**Fig. 1.3**) (Cameron and Collins, 2014; Gao et al., 2019; Lu et al., 2021).

Regulated protein depletion via degradation tags or degrons was commonly used in terpenoids biosynthesis (Lu et al., 2021; Peng, B. et al., 2018; Peng et al., 2017). Terpenoids biosynthesis requires the consumption of building blocks like isopentenyl pyrophosphate (IPP), geranyl pyrophosphate (GPP), and farnesyl pyrophosphate (FPP), but these precursors are also essential for biomass generation and cell growth. Therefore, the carbon flux needs to be appropriately controlled to minimize the competition of these precursors between cell growth and production. To this end, an autonomous controllable degradation was established in *S. cerevisiae* via a PEST (rich in Pro, Glu/Asp, Ser, and Thr) sequence-dependent endoplasmic reticulum-associated protein degradation (ERAD) mechanism (Peng et al., 2017). The existence of Erg9p would direct excessive FPP to sterol synthesis and decrease the production of the target product nerolidol, but

transcriptional inhibition of Erg9p was insufficient to regulate its availability due to the high stability of the enzyme (Christiano et al., 2014; Yuan and Ching, 2015). As a result, the protein-level regulation via the degradation tags was carried out to control the intracellular Erg9p availability. The Erg9p protein can be normally expressed and function during the growth phase, but it will be degraded during the production phase. Through this strategy, they successfully minimized the competition between cell growth and nerolidol production and enabled an 86% increase in the final titer (Peng et al., 2017). Inspired by this achievement, the same group further applied a similar strategy towards controlling the endogenous farnesyl pyrophosphate synthase (FPPS, Erg20p) to minimize the competition between biomass generation and linalool production, with a 27-fold increase in the final titer (Peng, B. et al., 2018). Recently, an engineered auxin-inducible protein degradation strategy for several terpenoids biosynthesis was developed (Lu et al., 2021). The natural auxin-inducible protein degradation system suffers from auxin-independent basal degradation, making it troublesome for controlling specific responses (Li et al., 2019; Nishimura et al., 2009). Hence, in the referred study, the auxin-inducible degradation system was optimized and improved with reduced non-auxin-dependent degradation. The engineered system was applied to improve the production of limonene and nerolidol. The titer of the monoterpene limonene and sesquiterpene (C15) nerolidol was increased to 76 mg/L and 3.5 g/L in flask cultivation, respectively (Lu et al., 2021).

While most studies focus on the employment of degradation tags and degrons, Cong and colleagues established a dynamic protein degradation genetic circuit via biosensor-controlled protease expression (Gao et al., 2019). Three orthogonal proteases, tobacco etch virus protease (TEVp), tobacco vein mottling virus protease (TVMVp), and sunflower mild mosaic virus protease

(SuMMVp) (Fernandez-Rodriguez and Voigt, 2016), were placed under the control of inducible promoters. The expression of these proteases will activate the degradation of target proteins harboring the corresponding tags. By manipulating these components and adjusting the control logic, they built several classic genetic circuits, including inverter and oscillator, and demonstrated the desired dynamic performance. The proteases-mediated dynamic protein degradation was then applied in shikimate and D-xylonate biosynthesis, with titers reaching 12.63 g/L and 199.44 g/L, respectively (Gao et al., 2019).

Regulation at the protein level is a direct control of protein availability, which is usually faster than regulations at DNA and RNA levels. This type of control is especially efficient for proteins with very long half-lives in cells (Peng et al., 2017). However, control at the protein level does not affect transcription and translation, which means the continued protein synthesis is also a waste of cell resources for unnecessary protein expression, and this may also bring extra burden for the degradation and decrease the regulation efficiencies. Additionally, the limited tunability, relaxed orthogonality, and the lack of protein-protein interaction regulators, may restrain the application of tag-based protein degradation in engineering the biosynthesis of natural product. Therefore, future engineering effort may be taken towards engineering toolsets with expanded tunability and mining new protein degradation mechanisms with improved orthogonality.

1.2.4 Multilevel regulation

Despite the encouraging advances in using single level regulation for the synthesis of natural compounds, there are some potential pitfalls in each type of regulation as discussed in the previous sections. Therefore, combination of multi-level regulations becomes a growing trend to improve

the regulation efficiencies in engineering biosynthesis of natural products, as this can possibly make up for deficiencies in single-level regulation (Calles et al., 2019; Greco et al., 2021; Martínez et al., 2017; Westbrook and Lucks, 2017).

For example, the DNA level regulations can be combined with protein degradation to minimize the regulation time lag by quickly degrading the proteins, and meanwhile, reduce the waste of cell resources for unnecessary protein synthesis by stopping the transcription of target genes. A notable example was seen in the engineering of glucaric acid and myo-inositol production in *E. coli* by combining quorum sensing system and protein degradation to control the transcription and protein availability of selected targets (Gupta et al., 2017). The quorum sensing was used to dynamically down-regulate the transcription of *pfkA* and reduce the glycolytic flux with the cell growth. Meanwhile, a standard ssrA degradation tag was fused to the enzyme Pfk-1 (protein encoded by *pfkA*) to accelerate the depletion of its intracellular availability and quickly shutdown the carbon flux towards glycolysis. This dynamic multilevel regulation enabled a 5.5-fold increase in titers of myo-inositol and pushed the titer of glucaric acid from undetectable to over 0.8 g/L (Gupta et al., 2017). Similar strategy has been applied in the biosynthesis of α-farnesene in *S. cerevisiae* (Yang et al., 2021). In the referred study, the quorum sensing regulators were combined with the auxin-mediated protein degradation system in *S. cerevisiae* to dynamically control the transcription and stability of Erg9, resulting in an 80% increase in the α-farnesene titer (Yang et al., 2021). Multilevel regulation can efficiently make up the deficiencies in single-level regulation, and thus has seen growing applications in engineering biosynthesis of natural products.

While most reported applications of multilevel regulation are combining regulations at DNA- and protein-level, combining DNA level and RNA level regulations, or RNA level and protein level

regulations, though were not applied in improving biosynthesis of natural products, have also been reported. RNA level regulation was combined with protein degradation to modify the protein synthesis rate and stability (Martínez et al., 2017). For multilevel regulations combining DNA and RNA level control, a double feedback circuit with bi-functional control of gene expression was engineered. The engineered circuit involves a translation-inhibitory sRNA with the translational coupling of the target gene to a repressor, and the target gene was under the control of the repressor-regulated promoter (Calles et al., 2019). In another study, Westbrook and co-workers configured and optimized the naturally-derived antisense RNA-mediated transcriptional regulators to control both transcription and translation in *E. coli* (Westbrook and Lucks, 2017). These approaches provide valuable insights for applying the multilevel regulations to achieve tunable gene expression. Thus, we will see more applications of multilevel regulations in the near future with the continuous development and optimizations of these toolsets.

1.2.5 Concluding remarks and future perspectives

The biosynthesis of natural products remains a hot topic in metabolic engineering and synthetic biology due to their high value in food, cosmetic, and pharmaceutical industries. As summarized in this review, tuning the gene expression at DNA, RNA, and protein levels via versatile regulation tools has been carried out to effectively improve the synthesis of many natural products. Each level of regulation has its advantages and disadvantages. Thus, multilevel regulation, as a trending regulatory strategy, has seen more and more applications in improving the biosynthesis of natural products, because multilevel regulations can potentially avoid the problems and challenges in using single-level regulation (Calles et al., 2019; Greco et al., 2021; Martínez et al., 2017; Westbrook and Lucks, 2017). Besides the application of improving natural products biosynthesis,

the advanced regulatory tools and strategies discussed in this paper can also be employed in engineering production of large protein complexes or non-natural chemicals (Fletcher et al., 2016; Nielsen, 2013), as well as for studying gene regulatory networks (Soma et al., 2017; Wang, T. et al., 2019).

With the expanding availability of biosensors and improvement on gene regulatory toolsets, a rapid increase in applications of dynamic pathway regulation has been observed in engineering biosynthesis of natural products for the last decade. Conventionally, static regulations were carried out to engineer the biosynthesis of natural products due to the ease of manipulation. However, static regulations are one-time, irreversible, and sometimes are manually carried out regardless of the cell status, which would cause undesired performance of microbial cell factories and reduce the productivities (Liu, D. et al., 2018). Dynamic pathway control mediated by genetically encoded biosensors can tune the gene expression and rewire carbon fluxes based on intracellular chemical concentration or environmental signals and thus it can be more efficient when encountering fluctuated chemical concentrations or complex environmental conditions. Thus, it may offer a more flexible and intelligent way to control the biosynthesis of natural products.

While fine-tuning the gene expression sees encouraging progress in engineering natural products, there are some challenges that cannot be fully addressed by these regulatory toolsets and strategies, such as low enzymatic activities or vulnerability to toxic end-products. Therefore, metabolic engineers will see combinatory optimizations that combining other engineering strategies, such as protein engineering and adaptive laboratory evolution, with the gene expression fine-tuning to

improve the biosynthesis of natural products to achieve better titers, yields and productivities in future engineering efforts.

1.3 Biosensor optimization for improved dynamic fine-tuning of gene expression

While the fine-tuning of gene expression has become a powerful strategy in improving the biosynthesis of microbial cell factories, the implementation of this strategy often requires the involvement of genetically-encoded biosensors, which can monitor the intracellular chemical concentrations or environmental signals (pH, temperature, dissolved oxygen, or osmatic pressure) in a real-time manner (Shen et al., 2019; Zhang and Keasling, 2011). Thus, to fulfill different regulation desires and expand the applicability of fine-tuning gene expression in various application scenarios, the optimizations and engineering of biosensor systems become increasingly essential. Among the optimizations, broadening the ligand scope and improving the dynamic ranges of biosensors are two major goals for biosensor engineering.

1.3.1 Engineering biosensors for expanded ligand spectra

Most biosensors are allosteric transcription factors (aTFs) with specific inducers and can recognize certain promoter sequences to repress or activate the downstream expression (Mahr and Frunzke, 2016). Besides real-time monitoring of intermediates or products, aTFs are now enabling synthetic biology applications in dynamic pathway regulating to enhance titer of natural products production and biosensor-based high-throughput screening of efficient enzymes or high-performance strains (Dahl et al., 2013; Eggeling et al., 2015; Liu et al., 2015; Rogers et al., 2016; Schallmey et al., 2014; Shen et al., 2019; Zhang and Keasling, 2011). However, the number of natural aTFs are limited, and they are sometimes not optimal for dynamic controls or biosensor-based screening.

Although identifying and characterizing novel sensors has been fruitful (Ganesh et al., 2013; Ganesh et al., 2015; Liu, C. et al., 2018; Liu, C.L. et al., 2018; Siedler et al., 2014), mining natural aTFs is difficult to catch up with the discovery of new compounds and meet complex regulation or screening requirements. Thus, it is important to engineer existing sensor-regulators with enhanced dynamic range or expanded scope of responsive ligands to improve the efficiency of biosensor-based screening and engineering.

Engineering existing aTFs to broaden the ligand scope of biosensors to sense new chemicals is of considerable interest in the field of regulator engineering (Figure 3a) (Libis et al., 2016). For example, the traditional L-arabinose biosensor AraC has been modified to sense a series of ligands, including D-arabinose (Tang et al., 2008), MVA (Tang and Cirino, 2011), triacetic acid lactone (Tang et al., 2013), and ectoine (Chen, W. et al., 2015). Site-saturated mutagenesis at certain amino acid positions enabled these AraC variants to respond to ectoine and the best ectoine dose responsive variants were selected. Through biosensor-based high-throughput screening, the specific activity of rate-limiting enzyme L-2,4-diaminobutyric acid (DABA) aminotransferase (EctB) was enhanced by up to 4.1-fold (Chen, W. et al., 2015). With a similar approach, the well-known *lac* repressor LacI was engineered to respond to four new inducers: fucose (Q291T), gentiobiose (H70D, H74S), lactitol (I79T, R101H) and sucralose (D149G, I160V, H163Y, S193E), through computational-guided protein design and site-saturated mutagenesis (a typical variant for each inducer is in the corresponding parentheses) (Taylor et al., 2016). Besides these well-characterized aTFs, many unconventional biosensors have also been engineered to sense natural products and thus can be used to regulate and improve their biosynthesis. TtgR from *Pseudomonas putida* (*P.p*), which is normally responsible for resistance to multiple antibiotics and plant

secondary metabolites, was engineered to specifically respond towards resveratrol. A variant with only a single amino acid substitution A38T enabled TtgR to respond to resveratrol and was applied to screen for *p*-coumarate:CoA ligase with improved activity (Xiong et al., 2017). One caveat is that crosstalk effects might exist when modifying the inducer scopes of biosensors. If the mechanism for ligand binding of an engineered biosensor is not fully understood, the regulator could possibly be induced by effectors with similar structures as the ligand. The crosstalk effect might hinder the application of the modified biosensor. Careful examination of crosstalk effects should be conducted after ligand scope engineering.

1.3.2 Engineering biosensors for versatile dynamic performance

Expanding the dynamic range of biosensors could enhance efficiency in biosensor-aided screening or dynamic regulation (Figure 3b). BenM, a prokaryotic transcription factor for *cis, cis*-muconic acid (CCM), was engineered to function in *Saccharomyces cerevisiae* through PCR-based random mutagenesis on the previously identified effector binding domains. The substitutions of H110R, F211V, and Y286N together showed doubled GFP output with the CCM induction. This variant was further applied in screening the best enzyme combinations for the rate-limiting protocatechuic acid decarboxylase in synthesizing CCM (Skjoedt et al., 2016). However, the dynamic range of a biosensor is typically affected by both the DNA binding affinity and the activity of the corresponding promoter. A repressor can bind to the specific promoter to inhibit the transcription unless alleviated by its inducer(s). When the binding between the repressor and the promoter is completely eliminated, the promoter will reach its maximum activity, which is independent of the regulator. Dynamic range expansion should cooperate with promoter engineering or RBS engineering as well. Combined with promoter engineering and directed evolution of a vanillic acid

sensor VanR from *Caulobacter crescentus*, the dynamic range of the regulator was improved 14-fold. The evolved biosensor was applied in screening natural catechol *O*-methyltransferases, and three active O-methyltransferases were successfully identified (Kunjapur and Prather, 2019). In another example, the dynamic range of the MarR-P_{marO} sensor system was positively shifted through tuning the expression level of MarR and promoter engineering of P_{marO} (Zou et al., 2021). A more comprehensive engineering, combining the fine-tuning of regulator expression level, protein engineering, and promoter engineering, of the phenolic acid responsive sensor system PadR-P_{padC} was recently reported (Jiang et al., 2021). The DNA binding region of PadR was mutated and two notable variants, H38A and K64, exhibiting improved sensitivity and expanded dynamic range, respectively, were selected. On establishing the enhanced dynamic ranges, Jiang and coworkers further investigated whether the two PadR binding boxes (located in the original P_{padC} promoter) can function in a "plug-and-play" manner by placing the binding boxes in a strong constitutive promoter pL. Through the hybrid promoter construction, a series of highly sensitive promoters with expanded dynamic ranges were obtained (Jiang et al., 2021). This work represents a comprehensive engineering of the biosensor dynamic performance.

1.3.3 Concluding remarks and future perspectives

Engineering biosensors rapidly expanded the number of detectable compounds and optimized the responsiveness of these regulators. Microbial biosynthesis often encounters complex conditions and thus the applicable biosensors need to fulfill high standards, such as little or no crosstalk effect, high sensitivity to inducers (some complex products can only be produced at milligram level) and an obvious on-off switch (appropriate dynamic range). Future research for engineering sensor-regulators would prioritize these directions. While significant advances have been made in

engineering existing biosensor systems for broader ligand spectra or enhanced dynamic performance, identifying novel biosensors by genome mining has also become increasingly important in developing new biosensor systems with the help of new computational algorithms and user-friendly databases. The rapid development of advanced bioinformatic tools and fast-sequencing techniques has led to enormous amounts of genomic sequence data, which revealed a tremendous reservoir of likely transcriptional factors awaiting to be characterized (Novichkov et al., 2013; Shlomi et al., 2007; Sun et al., 2020). Thus, in future studies, progresses will be seen in both engineering existing biosensors and mining novel regulators.

1.4 Research Objective

The primary goals of this work are the exploration of how fine-tuning the gene expression can improve the production of value-added compounds, and the development and optimization of biosensor variants with expanded substrate scope and versatile dynamic performance. Chapter 2 describes the development of a biosensor-enabled dynamic gene expression regulation at the gene copy level through controllable plasmid replication. This strategy expanded the dynamic regulation to gene replication level in order to fine-tune gene expression, and its real-world applicability has been validated in improving the biosynthesis of *p*-coumaric acid, the important precursor for a series of valuable flavonoids and coumarins. In chapter 3, we employed the mismatched sgRNA to enable titratable inhibition on cellular metabolism so that the carbon flux can be finely rewired to enhance the target productions. During this process, a biosensor-assisted high-throughput screening was established to accelerate the selection of high-producing phenotypes. These two chapters demonstrated the fine-tuning of gene expression can enhance the

production of value-added compounds, and also validated how the utilization of biosensors can assist the establishment of highly efficient microbial cell factories. However, the applications of biosensors in metabolic engineering may be limited by their ligand scopes and dynamic ranges. Thus, to improve the applicability and usability of the transcriptional factors-based biosensors, we broadened the ligand spectra of the phenolic acid-responsive regulator PadR from *Bacillus subtilis* through both rational and semi-rational strategies in Chapter 4. Our work highlighted the importance of transcriptional factors-based biosensor in engineering microbial biosynthesis, and demonstrated how the applicability of such biosensors can be improved by engineering existing transcriptional regulators.

1.5 Tables and Figures

Table 1.1 Summary of notable examples in fine-tuning gene expression to improve the microbial biosynthesis

Compounds (hosts[a])	Tool/strategy used (regulation level)	Achievements[b]	Reference[c]
Naringenin (*S.c*)	promoter engineering (DNA)	Increased the final titer to 52.0 mg/L	(Lee et al., 2013)
Naringenin (*E.c*)	promoter engineering (DNA)	A 2.1-fold increase in the final titer (191 mg/L)	(Zhou et al., 2019)
Lycopene (*E.c*)	aTFs (DNA)	Increased the final titer to over 150 mg/L	(Farmer and Liao, 2000)
Amorphadiene (*E.c*)	aTFs (DNA)	A 2-fold increase in the final titer (1.6 g/L)	(Dahl et al., 2013)
Fatty acids (*E.c*)	aTFs (DNA)	A 15.7-fold increase in the final titer (3.86 g/L) compared to wild-type strain	(Xu et al., 2014)
β-Ionone (*S.c*)	aTFs (DNA)	A 4.2-fold increase in the final titer compared to wild-type strain	(Naseri et al., 2019)
Salicylic acid (*E.c*)	aTFs (DNA)	A 1.8-fold improvement in the final titer (520 mg/L)	(Dinh and Prather, 2019)

Vanillin (*E.c*)	aTFs (DNA)	Increased the final titer to1.5 g/L	(Liang et al., 2020)
Naringenin (*E.c*)	CRISPRi (DNA)	A 7.6-fold increase in the final titer (421.6 mg/L)	(Wu et al., 2015)
β-Amyrin (*S.c*)	CRISPRi (DNA)	A 44.3% increase in the final titer (156.7 mg/L)	(Ni et al., 2019)
Arginine (*E.c*)	CRISPRi (DNA)	A 2-fold increase in growth rate with similar arginine production	(Sander et al., 2019)
Epothilone (*M.x*)	CRISPRa (DNA)	Increased the final titer of to over 25 mg/L	(Peng, R. et al., 2018)
Biopterin (*P.p*)	CRISPRa (DNA)	A more than 5-fold increase in the final titer	(Kiattisewee et al., 2021)
Mevalonate (*P.p*)	CRISPRa (DNA)	A 40-fold increases in the final titer (402 mg/L)	
β-Carotene (*S.c*)	Multiplex-CRISPR (DNA)	A 2.8-fold increase in the final titer	(Lian et al., 2017)
2′-Fucosyllactose (*B.s*)	DNA aptamer (DNA)	A 22.3-fold increase in the final titer (511 mg/L)	(Deng et al., 2019)
Inole-3-acetic acid (*E.c*)	TALE (DNA)	A 9.6-fold increase in the final titer	(Zhu et al., 2016)

Cadaverine (*E.c*)	Small regulatory RNA (RNA)	A 55% increase in the final titer	(Na et al., 2013)
L-Threonine (*E.c*)	Small regulatory RNA (RNA)	Increased the final titer to 22.9 g/L	(Yang et al., 2019)
L-Proline (*E.c*)	Small regulatory RNA (RNA)	Increased the final titer to 54.1 g/L	
Crude violacein (*E.c*)	Small regulatory RNA (RNA)	Increased the final titer to 5.19 g/L	
Indigo (*E.c*)	Small regulatory RNA (RNA)	Increased the final titer to 135 mg/L	
Naringenin (*E.c*)	Antisense RNA (RNA)	A 4.31-fold increase in the final titer (391 mg/L)	(Wu et al., 2014)
Naringenin (*E.c*)	Antisense RNA (RNA)	A 1.53-fold increase in the final titer (91.31 mg/L)	(Yang et al., 2015)
4-Hydroxycoumarin (*E.c*)	Antisense RNA (RNA)	A 2.53-fold increase in the final titer (270.85 mg/L)	
Resveratrol (*E.c*)	Antisense RNA (RNA)	A 1.70-fold increase in the final titer (268.20 mg/L)	
L-Lysine (*E.c*)	Riboswitch (RNA)	An 89% increase in the final yield	(Zhou and Zeng, 2015)

2′-Fucosyllactose (*B.s*)	Riboswitch (RNA)	A 27.3-fold increase in the final titer (674 mg/L)	(Deng et al., 2019)
Nerolidol (*S.c*)	PEST Degron (Protein)	An 86% increase in the final titer	(Peng et al., 2017)
Linalool (*S.c*)	Degron (Protein)	A 27-fold increase in the final titer (11 mg/L)	(Peng, B. et al., 2018)
Shikimate (*E.c*)	Dynamic heterologous proteases expression (Protein)	Increased the titer to 12.63 g/L	(Gao et al., 2019)
D-Xylonate (*E.c*)	Dynamic heterologous proteases expression (Protein)	Increased the titer to 199.44 g/L	
Limonene (*S.c*)	Auxin-inducible degradation (Protein)	Increased the final titer to 76 mg/L	(Lu et al., 2021)
Nerolidol (*S.c*)	Auxin-inducible degradation (Protein)	Increased the final titer to 3.5 g/L	
Glucaric acid (*E.c*)	Multilevel regulations: quorum sensing (DNA) with protein degradation (Protein)	Increased the final titer from undetectable to over 0.8 g/L	(Gupta et al., 2017)
Myo-Inositol (*E.c*)	Multilevel regulations: quorum sensing (DNA)	A 5.5-fold increase in the final titer	

	with protein degradation (Protein)		
α-Farnesene (*S.c*)	Multilevel regulations: quorum sensing (DNA) with protein degradation (Protein)	An 80% increase in the final titer	(Yang et al., 2021)

[a] The abbreviations for hosts listed in the table: E.c, *Escherichia coli*; S.c, *Saccharomyces cerevisiae*; B.s, *Bacillus subtilis*; P.p, *Pseudomonas putida*; M.x, *Myxococcus xanthus*.

[b] All fold-change were calculated based on the titer of the unregulated control unless specified in the table. The unregulated control means the strain with the same genotype and pathway but was not regulated by the toolset/strategy.

[c] When there were multiple examples demonstrated in one reference, the reference cells were merged to indicate that these examples/chemicals were using the same reference.

Table 1.2 Biosensor development for improved fine-tuning of gene expression

Enzyme	Chassis	Ligand	Mutation(s)*	Effects	Reference
AraC	*E.c*	L-arabinose	P8R/T24D/ H80L/ Y82Q; P8G/ F15W/ T24P/ H80A	Respond to D-arabinose	(Tang et al., 2008)
			P8P/ T24L/H80L/ Y82L/ H93R	Respond to mevalonate	(Tang and Cirino, 2011)
			P8V/ T24I/ H80G/ Y82L/ H93R	Respond to triacetic acid lactone	(Tang et al., 2013)
			P8Y/ T24M/ H80T/ Y82V/ H93V	Respond to ectoine	(Chen, W. et al., 2015)
LacI	*E.c*	Lactose/ Isopropyl β-d-1-thiogalactopyranoside (IPTG)	Q291T	Respond to fucose	(Taylor et al., 2016)
			H70D/ H74S	Respond to gentiobiose	
			I79T/ R101H	Respond to lactitol	
			D149G/ I160V/ H163Y/ S193E	Respond to sucralose	

TtgR	*P.p*	antibiotics	A38T	Respond to resveratrol	(Xiong et al., 2017)
BenM	*S.c*	*cis, cis*-muconic acid	H110R/ F211V/ Y286N	Doubled dynamic range	(Skjoedt et al., 2016)
VanR	*E.c*	Vanillic acid	T49I/ A111V/ P179S	14-fold increase in dynamic range	(Kunjapur and Prather, 2019)
MarR	*E.c*	Salicylic acid	N/A	Positively shifted the dynamic ranges	(Zou et al., 2021)
PadR	*E.c*	*p*-coumaric acid	H38A; K64A.		(Jiang et al., 2021)

* The symbol "/" in this column means these mutations are present at the same time. Symbol ";" was used to separate multiple functional variants in a study. "N/A" means no mutation was made in the experiment or the mutation was not specified in the reference.

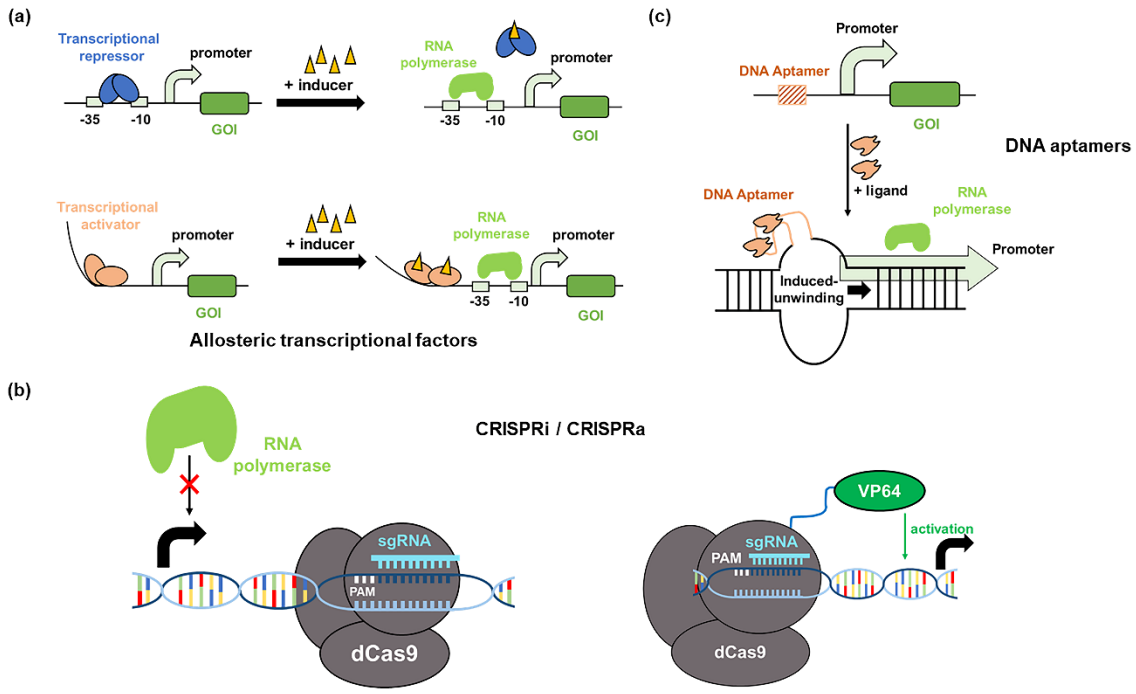


Figure 1.1. Schemes of DNA-level regulation toolsets. (a) Regulations mediated by transcriptional repressors (up) and transcriptional activators (down). Transcriptional repressors will occupy the -35 and -10 regions, thus blocking the access of RNA polymerase. When the inducer is present, the repressor can bind to the inducer. This process will result in a conformational change of the repressor, making it unable to bind the DNA sequence and thus release the transcriptional inhibition. For the transcriptional activator, the regulatory protein can bind to the DNA and force the DNA to bend. When the inducer is present, the binding of the inducer can cause a conformational change of the activator, which consequently leads to interaction with the RNA polymerase, and this will activate the expression of the target gene. GOI, gene of interest. **(b)** CRISPR based gene interference (CRISPRi, left) and activation (CRISPRa, right). For CRISPRi, the sgRNA (single guide RNA) will guide the dCas9 to the promoter region, blocking the access of RNA polymerase to the promoter, and thus inhibiting the transcription. For CRISPRa,

an activation domain (e.g., VP64, derived from the herpes simplex virus) is fused to dCas9 protein, and the sgRNA can guide the dCas9 to the appropriate location and bring the activator into contact with RNA polymerase. Upon the interaction, the promoter can be activated. PAM, protospacer adjacent motif, is a short DNA sequence that follows the DNA region targeted for cleavage by the CRISPR system. **(c)** Regulations mediated by DNA aptamers. The binding of ligand would cause unwinding of double stranded DNA. This would allow easier recognition and binding by RNA polymerases if the aptamer is placed near the promoter region, leading to the enhanced expression of downstream genes at the transcriptional level.

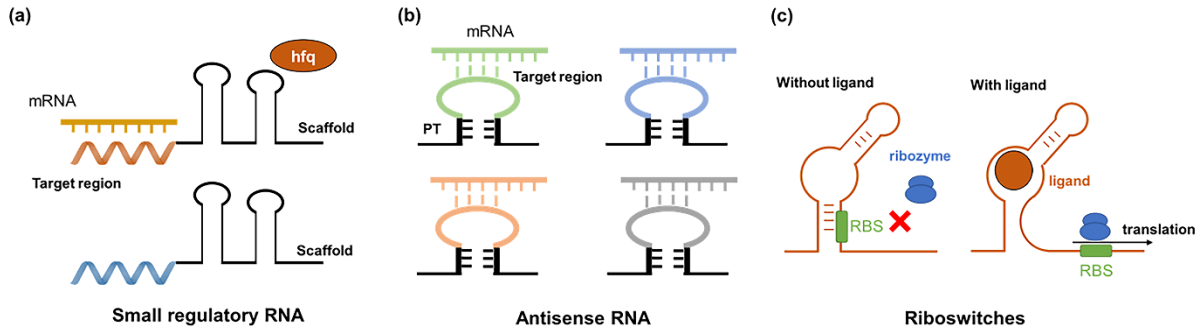


Figure 1.2. Schemes of RNA-level regulation toolsets. (a) Regulations mediated by small regulatory RNAs. A scaffold is used to enhance the stability of sRNA. A short binding sequence (target region) that can bind with target mRNA by complementary pairing and hinder the translation. Two small regulatory RNAs indicate that the target region is programmable/designable based on different regulation sequence but the scaffold can be similar. The scaffolds presented here for the two sRNAs (in black) are identical. Hfq, an abundant bacterial RNA binding protein that is physiologically important in interacting with sRNA. **(b)** Regulations mediated by antisense RNAs. Similar to sRNA, the asRNA achieves the regulation via complementary pairing with target mRNA. A 100 bp pair termini (PT) structure (in black) is used to stabilize the asRNA. The different colors in the asRNA represent varied target regions that can bind with different sequences. **(c)** Regulations mediated by riboswitches. The RNA secondary structure of riboswitch can block the access of ribozyme to the ribozyme binding site (RBS) and thus hinder the translation. This inhibition will be released upon ligand binding, and the ribozyme can then recognize the exposed RBS and start the translation.

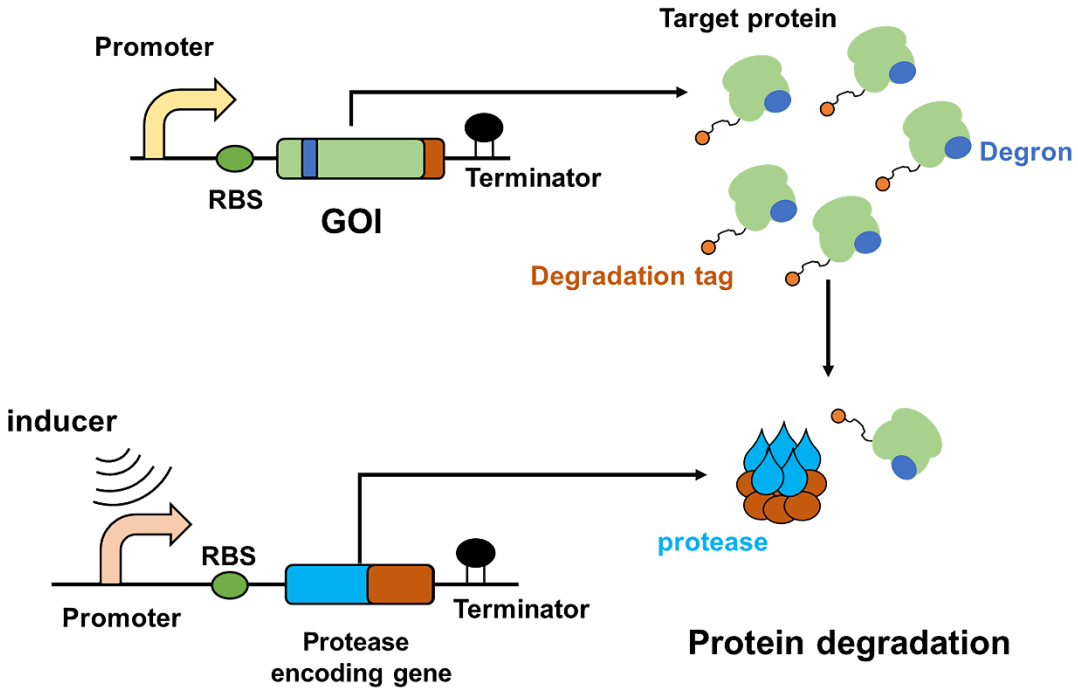


Figure 1.3. Schemes of protein-level regulation. Regulations at protein level are usually achieved by introducing the degradation tags into target proteins. The blue rectangle in the GOI represents the DNA sequence of degron, and the brown rectangle near the GOI represents the coding sequence of the degradation tag. The tagged protein will be recognized by corresponding protease and then degraded. The protease usually contains two functional domains, one for tag recognition (in blue) and one for degrading the tagged protein (in brown). For dynamic protein degradation, the protease can also be controlled by an inducible promoter (in pink) to enable inducible protein degradation. GOI, gene of interest.

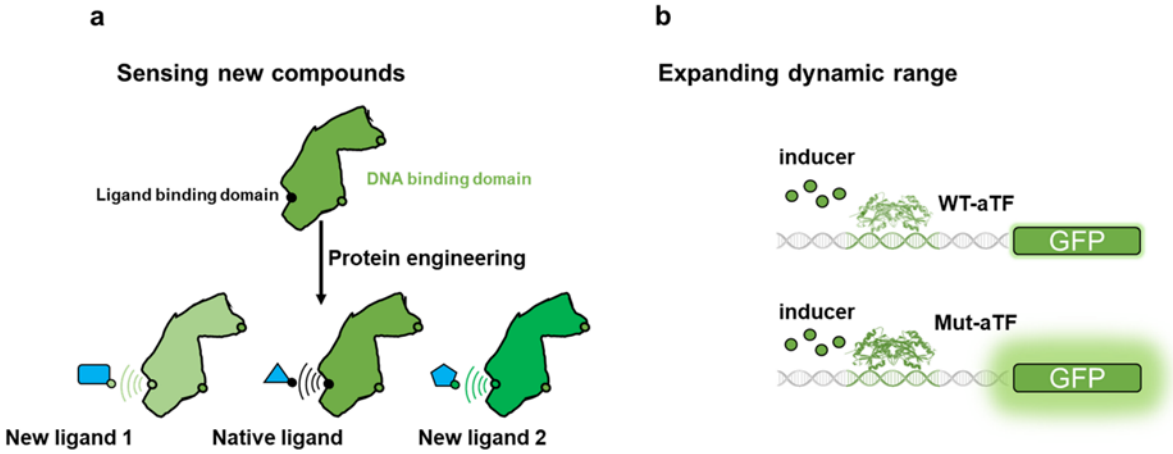


Figure 1.4. Protein engineering aid the optimization of genetically encoded biosensors. a) Engineering biosensors for responding to novel inducers. Through protein engineering, variant regulators with novel binding functions can be generated and applied in metabolic engineering of natural products biosynthesis. **b)** Engineering biosensors for expanded dynamic range. The enhanced regulators usually have stronger binding affinities and thus can result in more obvious output.

CHAPTER 2

HARNESSING PLASMID REPLICATION MECHANISM TO ENABLE DYNAMIC CONTROL OF GENE COPY IN BACTERIA[1]

---

1 Li, C., Zou, Y., Jiang, T., Zhang, J., Yan, Y., 2022. Harnessing plasmid replication mechanism to enable dynamic control of gene copy in bacteria. Metabolic Engineering.

2.1 Abstract:

Dynamic regulation has been proved efficient in controlling gene expression at transcriptional, translational, and post-translational level. However, the dynamic regulation at gene replication level has been rarely explored so far. In this study, we established dynamic regulation at gene copy level through engineering controllable plasmid replication to dynamically control the gene expression. Prototypic genetic circuits with different control logic were applied to enable diversified dynamic behaviors of gene copy. To explore the applicability of this strategy, the dynamic gene copy control was employed in regulating the biosynthesis of *p*-coumaric acid, which resulted in a up to 78% increase in *p*-coumaric acid titer to 1.69 g/L in shake flasks. These results indicated the great potential of applying dynamic gene copy control for engineering biosynthesis of valuable compounds in metabolic engineering.

2.2 Introduction

Dynamic pathway regulations have seen growing applications in engineering microbial production of natural products, biofuels, biopharmaceuticals, and bulk chemicals (Dahl et al., 2013; Doong et al., 2018; Farmer and Liao, 2000; Gupta et al., 2017; Xu et al., 2014; Yang et al., 2018; Zhang et al., 2012). Unlike traditional static regulations, dynamic regulation can adjust gene expression and direct carbon flux based on intracellular chemical concentrations or environmental signals (Liu, D. et al., 2018; Shen et al., 2019; Tan and Prather, 2017). To enable efficient dynamic pathway regulations, engineering efforts have been devoted to developing regulation toolsets and strategies at multiple levels. The transcriptional factors (Chen et al., 2018; Dinh and Prather, 2019; Gupta et al., 2017; Kim et al., 2017; Mahr and Frunzke, 2016), CRISPR based gene activation or inhibition

(Konermann et al., 2015; Larson et al., 2013; Lian et al., 2017; Perez-Pinera et al., 2013), and ligand-binding DNA aptamers (Deng et al., 2019; Wang, Jing et al., 2017), were used to target the transcription of target genes and thus achieving dynamic regulations at transcriptional level. At translational level, small regulatory RNA (Na et al., 2013; Yoo et al., 2013), antisense RNA (Yang et al., 2015), and riboswitches (Barrick and Breaker, 2007; Muranaka et al., 2009), were the most commonly used tools and strategies due to the short RNA lengths and the ease of manipulation. Besides, post-translational level regulation of protein availabilities by timely regulating the expression of heterologous proteases or assisting factors in protein degradation were also reported (Brockman and Prather, 2015; Dinh et al., 2020; Gao et al., 2019). All these regulation strategies were proven to be efficient in dynamic pathway regulation for engineering microbial production.

Beyond the regulations on transcription, translation and post-translation, controlling the copy numbers of genes, however, was less investigated in dynamic pathway regulations. While fine-tuning gene copies by using various plasmids has been commonly applied in metabolic engineering (Ajikumar et al., 2010; Jones et al., 2000), this strategy was usually used to statically or permanently adjust the gene copy during the production. Dynamically modulating the gene copy using biosensors in microbial production has not been explored so far. Dynamic gene copy regulations, in addition to the transcriptional, translational, or post-translational control, provides a new level of control to address the metabolic burdens arising from inappropriate high copies of genes or low-productivities from insufficient pathway copies by controlling the process of gene/plasmid replication based on cellular status or environmental conditions. Also, dynamic regulation of the gene copy through controllable plasmid replication, with just a simple regulatory device, can enable simultaneous u*p*- or down-regulation of the expression levels for all the genes

in the same plasmid. Therefore, this strategy may also be useful for efficient regulation of complex pathways with multiple gene targets.

With the goal of developing dynamic control at the gene copy level, we harnessed the replication mechanism of ColE1 derived origins. Plasmids with ColE1 derived origins, such as pUC18/19 (pMB1 variant) (Yanisch-Perron et al., 1985), pZE12luc (ColE1) (Lutz and Bujard, 1997), pCS27 (p15A) (Shen and Liao, 2008), and pET and its derivatives (pBR322) (Balbás and Bolívar, 2004), are commonly used in synthetic biology and metabolic engineering. Upon understanding the mechanism, we successfully established controllable replication of ColE1 derived origins. Based on this, three prototypic genetic circuits with different dynamic behaviors were constructed to enable diversified dynamic regulation logics on gene copy. Finally, the dynamic gene copy control strategy was applied in *p*-coumaric acid biosynthesis and enhanced its microbial production, which demonstrated the effectiveness of dynamic gene copy regulation in real-world application. Overall, our research provides a broadly applicable approach to dynamically and efficiently control microbial synthesis at DNA level.

2.3 Results

2.3.1 Establishing controllable plasmid replication to inhibit gene copy

The copy number of ColE1-derived plasmids is mainly regulated by a pair of plasmid-encoded RNAs, namely the RNAI and RNAII (Cesareni et al., 1991; Del Solar et al., 1998; Selzer et al., 1983). The RNAII serves as a pre-primer and folds into a secondary structure which can stabilize the interaction between the immature RNA and the DNA of origin. This hybrid is then processed by RNase H, which cleaves the RNA strand and exposes the 3' hydroxyl group. The DNA Polymerase I will then begin to synthesize the leading strand and initiate the replication **(Fig. 2.1a)**.

RNAI is a section of antisense RNAII which can bind to the 5' end of RNAII, forming kissing complex with the help of Rop protein. This will prevent the folding of RNAII and destroy the formation of the DNA-RNA hybrid, abolishing the aforementioned cleavage and slowing down the plasmid replication (Cesareni et al., 1991; Del Solar et al., 1998) **(Fig. 2.1a)**. All regulatory components (RNAI, RNAII, and Rop protein) are encoded in the plasmid (Camps, 2010; Cesareni et al., 1991; Cesareni et al., 1982).

Based on this mechanism, we hypothesized that the copy number of the plasmid harboring ColE1 derived origins can be controlled by manipulating the relative RNAI/RNAII availability. To test this hypothesis, we first selected the p15A origin as the target. The p15A origin is a ColE1-derived origin with a medium copy number (15-20 copies per cell). A reporter gene encoding the enhanced green fluorescence protein (egfp) was placed in the pCS27 plasmid (harboring the p15A origin) under the control of a constitutive promoter lpp0.5 (Wang, J. et al., 2017a). The fluorescence level normalized with the cell density (RFU/OD_{600}) was employed as an observable evaluation of plasmid copy number (PCN). The RNAI (p15A) was constructed into a plasmid pZE12-luc under the control of an IPTG (isopropylthio-β-galactoside)-inducible promoter pL_{lacO1}, resulting in pZE-RNAI (p15A). As expected, the over-transcribed RNAI (p15A) induced by IPTG resulted in a 68.87% decrease of the fluorescence level compared with the uninduced control **(Fig. 2.1b)**. To further validate that the lowered fluorescence intensity was caused by a decreased copy number, the relative PCNs of pCS-lpp0.5-egfp with and without the over-transcription of RNAI (p15A) were measured by quantitative PCR (qPCR) (Lee et al., 2006). The result showed that a 71.63% reduction in relative PCN was observed when RNAI was over-transcribed **(Table S2.1)**. We noticed that the inhibition efficiency obtained by qPCR was slightly higher than what we observed using fluorescence assay. This was likely due to the lag of fluorescent protein degradation inside

the cells after the decrease of PCN. These results demonstrated that the replication of p15A origin can be inhibited by increasing the RNAI availability.

With the success in regulating the replication of p15A origin, we then proceeded to test whether this strategy can be applied towards other ColE1 derived origins. Given that the p15A origin often provides a medium copy number in cells, we selected the ColE1 origin as our next target which gives a relatively higher copy number (70-100 copies per cell). The RNAI of the ColE1 origin was placed under the control of pL_{lacO1} in the pCS27 plasmid, resulting in pCS-RNAI (ColE1). The pZE-lpp0.5-egfp with the constitutively expressed egfp was used as the reporter plasmid. These two plasmids were co-transferred into *E. coli* BW25113 F' to test the inhibition efficiency of RNAI on the replication of ColE1 origin. Surprisingly, the over-transcription of RNAI (ColE1) only resulted in a very slight decrease (less than 8%) in the fluorescence intensity (**Fig. 2.1c**), indicating an insufficient repression on the replication of the ColE1 origin. We hypothesized that this was due to an inadequate amount of RNAI (ColE1), as the plasmid replication was controlled by the relative balance of RNAI and RNAII. The successful manipulation of the replication of p15A origin was likely because the over-transcribed RNAI (p15A) posed evident alteration of the intracellular RNAI/RNAII ratio, while for the ColE1 origin, the over-transcribed RNAI (ColE1) did not result in significant increase in the ratio of RNAI/RNAII. To examine this hypothesis, two parallel experiments were designed: we increased the RNAI (ColE1) availability by moving it from the medium-copy plasmid pCS27 to a high-copy plasmid pSC74 (with a CloDF13 origin that gives 20-40 copies per cell) (Camps, 2010; Stuitje et al., 1979), resulting in pSC-RNAI (ColE1). Meanwhile, the RNAI (p15A) was also moved to the plasmid pSA74 (Huo et al., 2011) (with an origin of pSC101* that gives 3-5 copies in cells) to reduce its intracellular availability. The RNAI

(p15A) transcribed from pSA only resulted in negligible inhibition on the replication of origin p15A (**Fig. 2.1d**), which was consistent with our expectation due to a decreased RNAI availability. However, the transcription of RNAI (ColE1) from the plasmid pSC-RNAI (ColE1) still did not result in any inhibition on the replication of ColE1 (**Fig. 2.1e**), indicating that the replication of pZE12 was insensitive to the increase of intracellular RNAI concentration, which was likely because the increased availability of RNAI (ColE1) did not result in significant change in the intracellular ratio of RNAI/RNAII. This suggested that further optimization of the copy number control strategy is required to manipulate the replication of the ColE1 origin. Our results indicated that the relative ratio of RNAI/RNAII was an important parameter for tuning the regulation efficiency on plasmid replication.

2.3.2 Modifying RNAI secondary structures for varied regulation efficiencies

The feasibility to inhibit the replication of p15A origin by over-transcribing RNAI was successfully demonstrated. However, this strategy was less successful when regulating the replication of the ColE1 origin which gives a relatively higher copy number. Therefore, we sought to investigate the factors related to the RNAI-RNAII interaction to enhance the inhibition efficiency. It is commonly recognized that the stability of RNA secondary structure plays an important role in RNA-RNA interactions (Cesareni et al., 1991; Guil and Esteller, 2015). Thus, we hypothesized that by optimizing the secondary structure of RNAI, the interaction between RNAI and RNAII can be enhanced and this will be beneficial for regulating the plasmid replication.

In our previous research, a double-stranded stem structure consisting of two inverted repeat DNA sequences (38 bp), namely the paired termini (PT) component, was developed to assist the inhibition of antisense RNA (Yang et al., 2015). We believe the long antisense stem in the PT

component will stabilize the secondary structure of the RNAI and thus provide an enhanced inhibition efficiency on RNAII. Using the RNAI (p15A) as a proof-of-concept target, we simulated the secondary structures for both original RNAI (p15A) and RNAI (p15A) with the addition of PT component (PT_RNAI(p15A)) through a web-based modeling toolset named RNAfold (http://rna.tbi.univie.ac.at//cgi-bin/RNAWebSuite/RNAfold) (**Fig. 2.2a**). Based on the predicted outputs, the addition of PT component led to a drastically decreased Gibbs free energy ΔG_0 of RNAI (p15A) from -42.29 to -129.21 kcal/mol (**Fig. 2.2a**). To experimentally test whether the improvement in RNAI stability by the addition of the PT component can enhance the regulation efficiency, the DNA sequence of PT_RNAI (p15A) was placed in the plasmid pZE12luc, forming pZE-PT_RNAI (p15A). The reporter plasmid pCS-lpp0.5-egfp and the pZE-PT_RNAI (p15A) were co-transferred into *E. coli*. The strain with the original RNAI (p15A) was used as a control. The normalized fluorescence intensities of induced and uninduced strains were tested and used to calculate the corresponding inhibition efficiencies. The transcription of PT_RNAI resulted in a 73.43% reduction in egfp expression, while the original RNAI (p15A) led to an inhibition efficiency of 68.58% (**Fig. 2.2b**). These results showed that the addition of the PT component on the RNAI slightly enhanced the regulation efficiency on plasmid replication. Observing this, we assumed that the addition of PT structure in RNAI (ColE1) might also improve the regulation effectiveness towards the ColE1 origin. Thus, the PT_RNAI (ColE1) was constructed and used to regulate the replication of ColE1 replication. This time, the introduction of PT_RNAI (ColE1) only enabled a 5% decrease in egfp expression level, indicating that the regulation efficiency of PT_RNAI (ColE1) was still insufficient to control the replication of ColE1 origin **(Fig. S2.1)**.

After achieving enhanced stability of RNAI by the addition of an external structure, we next sought to remove the internal unstable factors in the RNAI to further enhance its stability. In the simulated structure of RNAI (p15A), there are three stem loops, SL1, SL2, and SL3 (Cesareni et al., 1991). There is an imperfect match in each of the stem loop located in the antisense paired stem (**Fig. 2.2a**). Specifically, an extra U is located in the stem of SL1, and an extra G is located in the stem of SL3. For SL2, the sequence C-AAA-G cannot pair perfectly with G-AGA-C, because "AAA" is not the reverse complementary counterpart of "AGA". Thus, there is a bulgy ring in the stem of SL2 (**Fig. 2.2a**). We believe these mismatches may negatively affect the stability of RNAI, and the elimination of these unstable mismatches may improve the inhibition efficiency of RNAI due to an elevated stability. Therefore, we removed the extra U and G in SL1 and SL3, respectively, and the "C-AAA-G" sequence in the SL 2 was replaced by "C-TCT-G" to form a perfect antisense pair. The secondary structure of optimized RNAI was simulated in RNAfold and showed a decreased ΔG_0 of -59.39 kcal/mol compared to the original RNAI (p15A) with -42.29 kcal/mol (**Fig. 2.2a**). To experimentally test the inhibition efficiency of the optimized RNAI (p15A), the plasmid pZE-RNAI-opt (p15A) using the optimized DNA sequence of RNAI (p15A) was constructed. This plasmid was co-transferred with the reporter plasmid pCS-lpp0.5-egfp. The strain with the original RNAI (p15A) was used as the control. Unexpectedly, while the unmodified RNAI (p15A) still maintained an inhibition efficiency of 68.81%, the transcription of optimized RNAI (p15A) showed a lower inhibition efficiency, with only 48.31% reduction in the egfp expression level (**Fig. 2.2c**). This decrease in inhibition efficiency was presumably because the optimized sequence of RNAI was no longer a perfect antisense counterpart of RNAII. The sequence modification introduced mismatches between RNAI and RNAII, which caused a decreased binding affinity between the RNAI and RNAII. These results demonstrated that

although the stabilized secondary structure can enhance the inhibition efficiency, the perfect antisense match also plays a vital role in the RNAI-RNAII interaction. While the optimized RNAI (p15A) did not improve the regulation efficiency as we would expect, it can serve as a more relaxed control on the replication of p15A origin. The PT_RNAI, original RNAI, and RNAI-opt developed here provide varied inhibition efficiencies and may be useful in different application scenarios.

2.3.3 CRISPRi-mediated replication control on high-copy origin

While controllable replication inhibition can be established on the origins with medium or low copy number (e.g., p15A), efficient regulation of the high-copy origins, such as ColE1, remains challenging. As we shown before, even with the optimization of RNAI (i.e., addition of PT structure to RNAI(ColE1)), effective regulation on ColE1 replication was still difficult to achieve. We suspected that this was because the RNAI cannot efficiently neutralize RNAII due to a low ratio of RNAI/RNAII. Thus, a highly efficient regulatory tool that can control the RNAII transcription may be required to regulate the replication of high-copy origins. CRISPRi was demonstrated to be a very effective inhibition tool at the transcriptional level (Larson et al., 2013). We hypothesized that the replication of ColE1 origin can be efficiently controlled by introducing the CRISPRi to repress the transcription of RNAII, which can lead to decreased intracellular RNAII availability. Therefore, we adopted the CRISPRi into our strategy to test the feasibility of controlling the ColE1 replication. Two small guide RNAs (sgRNA) were designed to target different regions of the RNAII transcription cassette in the ColE1 origin. One is sgRNA-TIS with the target position located at 6-26 bp upstream of the transcription initiation site (TIS), namely the P2 promoter region. The other one is sgRNA-CDS with the target position located at 13-33 bp downstream of the TIS (**Fig. 2.3a**). These two sgRNAs were placed under the control of the IPTG-inducible promoter pL_{lacO1} and constructed into the plasmid pSA-dCas9 containing the deactivated

Cas9 protein (D10A/H841A) from *Streptococcus pyogenes* (Wang, J. et al., 2017b) , forming pSA-dCas9-sgRNA (ColE1-TIS) and pSA-dCas9-sgRNA (ColE1-CDS), respectively.

A 40.87% decrease of egfp expression level was achieved when targeting the promoter region by sgRNA-TIS, while only a less than 3% inhibition can be observed when using sgRNA-CDS **(Fig. 2.3b)**. The difference in the inhibition efficiency was likely because the binding of dCas9 protein guided by sgRNA-TIS caused steric hindrance and blocked the access of RNA polymerase to the promoter P2, but this hindrance was less prominent when the binding of dCas9 moved to the downstream sequence, and such decrease in regulation efficiency would no longer be able to cause obvious inhibition on plasmid replication. qPCR results also confirmed that the reduction in the egfp expression level was due to a decreased plasmid copy number, with a 53.98% decrease in relative plasmid concentration can be detected **(Table. S2.2)**. These results proved our hypothesis that inhibiting the transcription of RNAII can efficiently control the replication of ColE1 origin. We then proceed to see how well CRISPRi can perform in regulating the p15A origin, and a similar trend was observed when using CRISPRi to target the transcription of RNAII in p15A origin **(Fig. 2.3c)**. The sgRNA-TIS for RNAII (p15A) caused a 55.56% reduction in egfp expression, while the sgRNA-CDS for RNAII (p15A) did not result in any obvious decrease in the normalized fluorescence intensity. We noticed that the regulation efficiency on plasmid pCS is slightly better than on pZE. This was likely because the copy number of pZE is higher than that of pCS, and this means a lower concentration of sgRNA is required to saturate the inhibition on the replication of plasmid pCS. However, the inhibition efficiency on p15A replication achieved by CRISPRi was lower than that by over-transcribing RNAI (p15A), which can maintain an inhibition efficiency of 68% (**Fig. 2.1b**). We suspect this was likely because the RNAI (p15A) was over-transcribed from

a high-copy plasmid pZE12 (70-100 copies per cell), while the dCas9 and sgRNA were expressed from a low-copy plasmid pSA74 with pSC101* origin (3-5 copies per cell). Thus, the introduced dCas9 and sgRNA may not be adequate to efficiently repress the RNAII transcription, and this can lead to a decreased inhibition efficiency in comparison to RNAI (p15A).

2.3.4 Controllable gene copy enhancement by reducing RNAI availability

On validating the controllable inhibition on replication of ColE1 derived origins, we then sought to test whether the controllable activation is achievable by reducing the intracellular RNAI/RNAII ratio. We hypothesized that, if the intracellular RNAI/RNAII ratio can be reduced, either through lowering the RNAI concentration or repressing the transcription of RNAI, the plasmid replication would be enhanced, pushing the copy number to a higher level. To test this hypothesis, we designed the antisense RNAI (asRNAI, reverse complementary to RNAI) of p15A origin that would bind with RNAI and decrease the intracellular concentration of free RNAI. The coding sequence of asRNAI was placed under the control of pL_{lacO1} promoter, forming pZE-pL_{lacO1}-asRNAI (p15A). This plasmid was used to target the replication of p15A origin using plasmid pCS-lpp0.5-egfp as the reporter (**Fig. 2.4a**). As expected, the introduction of asRNAI resulted in a 5.39-fold increase in the egfp expression level (**Fig. 2.4b**), which demonstrated that it was feasible to increase the copy number of p15A origin by decreasing the RNAI availability. Since the PT structure was demonstrated to be efficient in improving the regulation efficiency of RNAI, we also tested whether the introduction of the PT structure into the asRNAI would cause increased regulation effectiveness. While the PT_asRNAI still enabled a 5.08-fold increase in egfp expression level (**Fig. 2.4b**), The results suggested that both asRNAI and PT_asRNAI can achieve high activation efficiencies.

After achieving the controllable activation on p15A replication, we next proceeded to investigate whether the replication of the high-copy ColE1 origin can be enhanced by repressing the transcription of RNAI. As the replication of ColE1 origin is insensitive to RNA-level control, the CRISPRi, which was demonstrated successful in inhibiting the ColE1 replication, was applied to target the transcription of RNAI (ColE1). Since the p15A origin can also be regulated by CRISPRi **(Fig. 2.3c)**, we also include the p15A origin as a target in the test. Thus, two sgRNAs targeting the RNAI promoter region of ColE1 and p15A origin **(Fig. 2.4c)**, namely the sgColE1 and sgp15A, respectively, were designed and constructed in the pSA-dCas9 plasmid containing the dCas9 protein expression cassette, forming pSA-dCas9-sgColE1 and pSA-dCas9-sgp15A. The strain harboring only the dCas9 protein expression cassette but without any sgRNA was used as a control. The introduction of CRISPRi on the RNAI (ColE1) transcription enabled a 2.14-fold increase in egfp expression over the control **(Fig. 2.4d)**. Similarly, the inhibition on RNAI (p15A) transcription by CRISPRi also resulted in a 3.22-fold increase in the egfp expression level **(Fig. 2.4d)**. The regulation on RNAI transcription led to a more prominent effect on plasmid pCS than on pZE, which was similar to what we observed when using CRISPRi to inhibit the transcription of RNAII. Also, the CRISPRi-mediated enhancement on p15A replication was less effective compared with what can be achieved by RNA-level control (3.22-fold vs 5.39-fold), which was consistent with what we observed in inhibition of the p15A replication (**Fig. 2.1b, 2.3c**). Our results demonstrated that the decrease of RNAI/RNAII ratio, either by lowering the RNAI concentration or repressing the transcription of RNAI, would cause enhanced replication of ColE1 derived origins. The controllable activation, along with the controllable inhibition of the plasmid replication demonstrated in previous sections, can be used to enable diversified dynamic control of gene copy via regulating the plasmid copy number.

2.3.5 Exploration of simultaneous control of two plasmids

Applying the CRISPRi in dynamic gene copy control enabled the regulation on both the p15A and ColE1 origin. One advantage of CRISPRi is the ease of achieving multiplex regulation by designing different sgRNAs. As the two-plasmid configuration is usually used in metabolic engineering and synthetic biology research, we explored the potential of using sgRNAs to target the ColE1 and p15A origin at the same time for simultaneous regulation of both origins. As a demonstration, we selected the sgRNA-TIS (ColE1) and sgRNA-TIS (p15A), which target the RNAII transcription of ColE1 and p15A origin, respectively. The two sgRNA transcription cassettes were placed on plasmid pSA-dCas9, resulting in plasmid pSA-dCas9-Co_sgRNA. Both the pZE12 plasmid carrying the red fluorescence protein (RFP) and pCS27 plasmid carrying the egfp were used as the reporter plasmids. The induction of sgRNA transcriptions resulted in a 34.63% reduction of RFP expression level and a 47.49% in egfp expression level **(Fig. S2.2)**, demonstrating the capability of co-inhibiting two plasmids at the same time. To see whether the regulation on two plasmids is orthogonal, we individually targeted only one plasmid when both reporter plasmids were present. The inhibition on p15A origin successfully led to a 48.11% decrease in egfp expression level, but only slightly affected the red fluorescence level. The inhibition on ColE1 origin resulted in a 31.75% decrease in RFP expression level, and did not affect the egfp expression level **(Fig. S2.2)**. These results proved that the co-regulation of two origins is feasible, and the regulations on two plasmids are orthogonal.

2.3.6 Designing genetic circuits to enable diverse dynamic control of gene copy

With the validation of both controllable reduction and increase of gene copy through regulating the replication of origins, our next goal is to achieve diversified dynamic regulation of the gene

copy using biosensor-enabled genetic circuits with different control logics. To this end, three different control logic were designed. The first control logic enables a dynamic behavior of direct "normal-to-low" control: the gene copy would be maintained at a normal level in the beginning but reduced when there is a signal input. As a result, the gene copy would be reduced to a lower level. The second control logic on the other hand, is direct enhancement of gene copy, forming a "normal-to-high" dynamic behavior: the gene copy would be maintained at the normal level when there is no signal input, but the replication would be enhanced to improve the gene copy when the input is present. The third control logic enables a more complex regulation: the inhibition on gene copy will be executed at the beginning, forcing a low-level of gene copy. When there is a signal input, the inhibition on gene copy would be released, and the gene copy would be restored to the normal level.

To apply the designed control logic, biosensor-enabled genetic circuits with the desired dynamic behaviors are employed. As a proof-of-concept demonstration, the optimized *p*-coumaric acid-responsive sensor system PadR-P_{padC} and its variants developed in our previous study (Jiang et al., 2021) were selected. PadR is a transcriptional repressor that can bind with P_{padC} (and its variant/hybrid promoters) to inhibit the promoter activity, and this inhibition would be released when *p*-coumaric acid is present (Jiang et al., 2021; Nguyen Thi Kim et al., 2011). To enable varied dynamic performance of the genetic circuits, we employed both the wild type (WT) PadR and a PadR (K64A) variant with increased responsiveness but less leaky activity from our previous study (Jiang et al., 2021) in demonstrating all three control logics. For each control logic, we designed two versions of genetic circuits, one for regulating the medium-copy p15A origin and the other is for high-copy ColE1 origin. When regulating the replication of medium-copy origin p15A, the PT_RNAI/asRNAI were used, while the CRISPRi system (dCas9 with sgRNAs) was applied when

regulating the replication of high-copy origin ColE1. As we demonstrated in previous section, the control of p15A replication required a high concentration of RNA (PT_RNAI or asRNAI), while a small amount of sgRNA was sufficient to regulate the ColE1 replication when using CRISPRi. Thus, the strong promoter P9 that responds to *p*-coumaric acid (Jiang et al., 2021) was used to control the transcription of PT_RNAI or asRNAI, while the weaker promoter $P_{padC\text{-}FG}$ (Jiang et al., 2021) was used for sgRNAs targeting the ColE1 replication.

For constructing the circuits with the first control logic (circuit 1), the PT_RNAI (p15A) and sgRNA-TIS (ColE1) were placed under the control of the *p*-coumaric acid responsive promoters P9 and $P_{padC\text{-}FG}$, respectively. The regulator gene *padR* was under the control of the constitutive promoter lpp0.2 (Wang, J. et al., 2017a) for optimal sensor performance (Jiang et al., 2021) **(Fig. 2.5a,c)**. As we designed, the promoter P9 and $P_{padC\text{-}FG}$ can be repressed by PadR when no *p*-coumaric acid is present, but such repression will be released when *p*-coumaric acid is added. The P9 or $P_{padC\text{-}FG}$ would then initiate the transcription of PT_RNAI (p15A) or sgRNA-TIS (ColE1), respectively, which would lead to inhibition on plasmid replication and reduce the gene copy. To test the circuit performance, gradient concentrations of *p*-coumaric acid (0, 100 mg/L, 500 mg/L) were fed into the media after 2h of inoculation. The dynamic performance of an unregulated control (without the PT_RNAI or sgRNA transcription operon) was also tested. As expected, with similar cell growth (**Fig. S2.3a,b**), the addition of *p*-coumaric acid resulted in obvious reductions in egfp expression level when targeting either the p15A **(Fig. 2.5b)** or the ColE1 origin **(Fig. 2.5d)** (up to an inhibition efficiency of 37.45% for regulating p15A origin and 55.76% for ColE1 origin). Interestingly, we noticed that the dynamic performance of genetic circuits was largely dependent on the PadR regulator. For example, since the PadR (K64A) exhibits less leaking activity (Jiang et al., 2021), when no *p*-coumaric acid was added, the reduction in egfp expression level of the

circuit harboring the PadR (K64A) variant (K64A-0, red line, 2.55% reduction at 12h) was less prominent compared with the circuit harboring the WT PadR (WT-0, orange line, 6.53% reduction at 12h) when regulating the p15A origin **(Fig. 2.5b)**. This was especially obvious in regulating the ColE1 origin (17.79% reduction in WT-0 vs 3.08% reduction in K64A-0) **(Fig. 2.5d)**. Also, due to a higher responsiveness of PadR (K64A) towards *p*-coumaric acid, the inhibition on gene copy was more prominent when using PadR (K64A), as both the highest inhibition efficiencies were achieved when the circuit was equipped with K64A and fed with 500 mg/L *p*-coumaric acid. (**Fig. 2.5b,d**). Prominent dose-dependent effect was observed in regulating ColE1 origin **(Fig. 2.5d)**. The reduction in egfp expression level induced by 500 mg/L *p*-coumaric acid was higher than that of induced by 100 mg/L *p*-coumaric acid in both circuits harboring the WT PadR- (43.44% vs 54.97%) and PadR (K64A) (43.79% vs 55.76%). All these results showed that the performance of the biosensor was well-transited to the dynamic performance of the genetic circuits.

The genetic circuits with the second regulation logic (circuit 2) were constructed with a similar configuration to the circuit 1, but the regulation tools that were under the control of *p*-coumaric acid-responsive promoters were changed to asRNAI (p15A) or sgColE1 **(Fig. 2.5e,g)**. We also tested the dynamic performance of circuit 2 with both WT PadR and PadR (K64A), aiming to see whether the sensor dynamic properties would again, like what we observed in testing the first control logic, be transited to the dynamic performance of the circuit 2. Gradient concentrations of *p*-coumaric acid (0, 100 mg/L, 500 mg/L) were added to the media after 2h of inoculation. An unregulated strain (without the asRNAI or sgColE1 transcription cassettes) was also included in the test as a control. Both versions of the circuit 2 performed well as we expected. The addition of *p*-coumaric acid enhanced the plasmid replication and improved the egfp expression level **(Fig. 2.5f,h)**, with an up to 1.78-fold and 2.37-fold increase can be observed at 14h for regulating p15A

and ColE1 origin, respectively. Slight growth defects can be observed when applying circuit 2 in regulating the high-copy ColE1 origin (**Fig. S2.3c,d**), which was likely because the further increased gene copy led to elevated burdens on cells. Dose-dependent effect was also observed in both regulations of p15A and ColE1 replication **(Fig. 2.5f,h)**. Notably, the dynamic properties of the biosensor was again well transited to the dynamic performance of the circuit, as the circuit 2 harboring the PadR (K64A), compared to the circuit equipped with WT PadR, exhibited less leaky enhancement but higher activation efficiencies **(Fig. 2.5f,h)**.

The control logic for the third type of genetic circuits (circuit 3) resembles the function of a genetic inverter, which adds an extra layer of regulation compared to the previous two strategies. To achieve this designed function, this circuit requires an additional regulator, for which we chose the well-studied and commonly used "repressor-promoter" combo, TetR-pL_{tetO1}. The pL_{tetO1} is a constitutive promoter when the repressor TetR is not present, but its activity would be inhibited when TetR is expressed (Lutz and Bujard, 1997). Thus, the pL_{tetO1} promoter was used to control the transcription of PT_RNAI (p15A) or sgRNA-TIS (ColE1). The *tetR* gene was under the control of a weaker promoter $P_{padC\text{-}FG}$ instead of hybrid promoter P9 due to the high sensitivity of the TetR-pL_{tetO1} system (a small amount of TetR protein is sufficient to fully repress the pL_{tetO1} promoter) (Lutz and Bujard, 1997). We first tested how this circuit can manipulate the gene copy via regulating the medium-copy p15A origin. An unregulated strain with only the sensor and the reporter plasmid (no PT_RNAI transcription cassette) was used as the positive control (Control). Gradient concentrations of *p*-coumaric acid (0, 100 mg/L, 500 mg/L) were added to the media after 2h of inoculation. The induction of *tetR* expression by *p*-coumaric acid enabled increased fluorescence intensities compared with the strain without any *p*-coumaric acid feeding (WT-0 and K64A-0), indicating the inverter function of circuit 3 was successfully established on regulating

p15A origin **(Fig. 2.5j)**. Dose-dependent effect was again observed, as the increased *p*-coumaric acid concentration resulted in a better recovery of egfp expression, with up to 77.17% activity can be recovered by the circuit harboring the PadR (K64A) induced by 500 mg/L *p*-coumaric acid **(Fig. 2.5j)**. Similar to the previous two circuits, the sensor performance was also well-transited to the circuit, as the K64A variant can enable a better recovery of egfp expression (up to 77.17% egfp expression can be recovered) than that of WT PadR (up to 64.03%). We then tested whether this control logic could also work on regulating the high-copy ColE1 origin. For the circuit 3 controlled by WT PadR, the introduction of *p*-coumaric acid released the inhibition on ColE1 replication and resulted in increased egfp expression with dose-dependent effect **(Fig. 2.5l)**. Up to 62.54% egfp expression level can be resumed by the genetic inverter. Notably, the strains with induced circuit 3 (WT-100 and WT-500) showed better cell growth (**Fig. S2.3f**), indicating that the circuit 3 alleviated the cell burdens from excessive copy of egfp. However, unlike what we observed in regulating p15A replication **(Fig. 2.5j)**, the K64A variant did not result in better recovery of egfp expression this time. Instead, the circuit 3 harboring the PadR (K64A) can be barely activated by *p*-coumaric acid **(Fig. 2.5l)**. We suspected that this was due to a stringent control of $P_{padC\text{-}FG}$ by PadR (K64A), and this led to a stricter inhibition on TetR expression, which subsequently resulted in a higher activity of pL_{tetO1} and increased sgRNA concentration. As the regulation on ColE1 replication was sensitive to the sgRNA availability, the increased sgRNA concentration make it harder for the circuit 3 to eliminate the inhibition and restore the gene copy of egfp. Overall, the function of all three designed control logics have been demonstrated, and the developed genetic circuits with these dynamic behaviors would be beneficial for diversifying the dynamic regulation of gene copy in metabolic engineering.

2.3.7 Application of dynamic control of gene copy in *p*-coumaric acid biosynthesis

While we have validated and diversified the dynamic control of gene copy, the potential of this strategy in real-world applications remains unexploited. To test the applicability of dynamic control of gene copy, the biosynthesis of *p*-coumaric acid was selected as a proof-of-concept demonstration. *p*-coumaric acid is a high-value aromatic compound that serves as an important precursor for synthesizing many valuable natural products, such as naringenin, resveratrol, and apigenin (Siedler et al., 2014; Yang et al., 2015). Microbial production of *p*-coumaric acid is usually catalyzed by tyrosine ammonia lyase (TAL), which employs the L-tyrosine as the precursor. In our previous studies, we have established an efficient *de novo p*-coumaric acid biosynthesis pathway by employing four key enzymes from *E. coli*, TyrA*, PpsA, TktA, and AroG* (hereafter referred as TPTA, * indicated that the enzyme is feedback-resistant) along with the TAL from *Rhodotorula glutinis* (RgTAL) (Huang et al., 2013; Wang, J. et al., 2017a) **(Fig. 2.6a)**. Thus, the *p*-coumaric acid biosynthesis pathway containing the genes encoding the TPTA and RgTAL was assembled to the pCS27 plasmid, resulting in pCS-TPTA-RgTAL. As demonstrated in our previous study (Yang et al., 2018), the inhibition on *ppc* gene via antisense RNA would enhance the carbon flux towards shikimate pathway, which would be beneficial in improving the production of *p*-coumaric acid. Thus, the asppc (antisense RNA of ppc that inhibit the expression of genomic *ppc*) transcription cassette was also integrated with the *p*-coumaric acid pathway, forming pCS-TPTA-RgTAL-asppc. To direct more carbon flux to the *p*-coumaric acid biosynthesis and enhance the tyrosine accumulation, the competing pathways in *E. coli* BW25113 F'(Atsumi et al., 2008) were eliminated by knocking out *pheA* (for phenylalanine biosynthesis), *pykAF* (consuming the precursor phosphoenolpyruvate, PEP), and *tyrR* (encoding the regulator TyrR that conducts tyrosine-responsive feedback inhibition on tyrosine biosynthesis) **(Fig. 2.6a)**.

The final strain *E. coli* BW-PCA (BW25113 F' Δ*pheA* Δ*tyrR* Δ*pykA* Δ*pykF*) was used for applying dynamic gene copy regulation in *p*-coumaric acid biosynthesis.

The biosynthesis of *p*-coumaric acid requires directing carbon flux from glycolysis to shikimate pathway **(Fig. 2.6a)**. It will compete with central metabolism for PEP (phosphoenolpyruvate) and E4P (erythrose 4-phosphate), which may reduce the biomass generation and negatively affect the cell performance. To solve this problem, we first employed the dynamic "low-to-normal" control logic **(Fig. 2.5i)** to regulate the replication of all genes in the plasmid pCS-TPTA-RgTAL-asppc (with a p15A origin) **(Fig. 2.6b)**. The copy numbers of the pathway genes would be kept at a low level when no *p*-coumaric acid is present. With cell growing and the accumulation of *p*-coumaric acid, the inhibition on gene replication would be released, and the pathway expression will be restored to a normal level, directing carbon flux towards the final product. We aimed to reduce the competition between cell growth and *p*-coumaric acid production using this design. To achieve the function, the circuit with the third control logic **(Fig. 2.5i)** was used to control the replication of all pathway genes in the plasmid. The circuit harboring the PadR (K64A) variant was used as this circuit exhibited a stricter inhibition and higher recovery efficiency of gene copy than the circuit with WT PadR **(Fig. 2.5j)**. The strain without the dynamic gene copy regulation was used as the control. As expected, the cell growth was boosted by the dynamic gene copy regulation, with up to 24.52% increase in cell density at 36h and an 11.33% increase at 72h can be observed **(Fig. 2.6c)**. However, the *p*-coumaric acid production did not further increased with the enhanced cell growth **(Fig. 2.6c)**. A titer of 1.29 g/L *p*-coumaric acid can be accumulated using the dynamic gene copy control, while the unregulated control can generate a slightly higher titer of 1.33 g/L **(Fig. 2.6c)**. We suspected this was due to the trade-offs between cell growth and production when using dynamic pathway control: when the pathway expression was at a low level, more carbon fluxes

would be directed to biomass generation to improve the cell growth. However, the increased biomass generation, at the price of reduced pathway expression, reduced the carbon fluxes that can be used for microbial synthesis and led to a declined *p*-coumaric acid production.

As the better cell growth did not lead to increased production of *p*-coumaric acid, we hypothesized that the metabolic burdens arising from the pathway expression was still in an acceptable range. Next, the circuit with the "normal-to-high" control logic (circuit 2) was employed to dynamically improve the pathway gene copies. In this design, the pathway genes would maintain a normal copy number at the beginning of fermentation, but the gene replications would be dynamically enhanced with the accumulation of *p*-coumaric acid. Again, due to a better dynamic performance, the PadR (K64A)-enabled circuit was used to u*p*-regulate the replication of p15A origin. The strain without dynamic gene copy regulation and the strain with static gene copy regulation (constitutively enhanced gene copy) were used as the control. As expected, the dynamic u*p*-regulation of gene copies improved the production of *p*-coumaric acid, with the highest titer reaching 1.69 g/L in shake flasks **(Fig. 2.6d)**. The unregulated strain accumulated 1.29 g/L *p*-coumaric acid, while the strain with static regulation exhibited growth defects and only produce a titer of 0.95 g/L *p*-coumaric acid **(Fig. 2.6d)**. The dynamic regulation of gene copy resulted in 31.01% and 77.89% increase in *p*-coumaric acid titers compared to the unregulated control and static control, respectively **(Fig. 2.6d)**. The plummeted titer by the static regulation confirmed that the excessive copies of pathway genes would negatively affect the cell growth and led to decreased production performance. The dynamic u*p*-regulation, on the other hand, avoided the overwhelmed metabolic burden at the beginning compared to the static regulation, but gradually directed more carbon fluxes to *p*-coumaric acid synthesis when cells enter to post-log phase and stationary phase compared to the unregulated strain. Our results demonstrated the great potential of employing the

dynamic gene copy control in real-world applications. Notably, our results also highlighted the importance of appropriately applying dynamic pathway control based on different regulation scenarios. When the metabolic burdens or competition between biomass generation and products accumulation did not cause significant defects on cell growth, applying dynamic down-regulation would reduce the carbon fluxes towards target compounds. This loss in compounds accumulation is unlikely to be made up by the better cell growth in this case. At this time, the "dynamic-up" regulation with gradually increased pathway expression would be more helpful in improving the biosynthesis, as demonstrated in our experiments. Overall, these results highlighted the great potential of applying the dynamic pathway control at gene replication level for engineering microbial cell factories.

2.4 Discussion

Dynamic regulation has attracted growing attentions in recent years as it provides a new solution to intelligently direct carbon flux, inhibit competing pathways that are essential to cell growth, or minimize accumulation of toxic intermediates, based on intracellular chemical concentrations and environmental signals. To enable efficient dynamic regulation, plenty of toolsets and strategies, as discussed in the introduction section, have been developed and engineered to improve the performance of microbial cell factories. While these strategies have been demonstrated efficient in regulating gene expression at transcriptional, translational, and post-translational level, the dynamic regulation at gene replication level, or gene copy level, to address the inherent burdens coming from unnecessary high copies of genes or low pathway efficiencies from insufficient gene replications, has been rarely explored.

Here, we successfully established dynamic control at gene replication level on plasmids with ColE1 derived origins using tunable biosensor and genetic circuits. As demonstrated in this study, by manipulating the transcription of RNAI and RNAII, the gene copy can either be dynamically increased or decreased in a tunable way. To diversify the control logic of regulating gene replication, genetic circuits equipped with biosensor were constructed and enabled different dynamic behaviors of gene replication. To investigate the industrial potential of this strategy, we applied dynamic gene copy control in *p*-coumaric acid biosynthesis, and enhanced the production of *p*-coumaric acid by up to 78% to a titer of 1.69 g/L in shake flasks. More importantly, the dynamic gene copy regulation developed in this study is tunable at multiple levels. By engineering the secondary structure of RNAI, using different regulatory tools (e.g., CRISPRi), and employing biosensor variants with different dynamic performance, the performance of the genetic circuits can be fine-tuned to meet different regulation requirements in various application scenarios. The *p*-coumaric acid titer was moderately enhanced by dynamic gene copy regulation. However, the production improvement by dynamic pathway regulation can be case-dependent, and it is also closely related to factors such as the dynamic properties of the biosensors, the pathway kinetics, and the cultivation conditions. Thus, the case of *p*-coumaric acid biosynthesis in this study is sufficient to demonstrate the applicability of the dynamic gene copy control in regulating chemical production.

Dynamic control of the gene copy expands the dynamic pathway regulation to gene replication level, which brings several advantages compared with current strategies of gene expression control at transcriptional, translational, and post-translational level. Taking the dynamic down-regulation as an example, the dynamic down-regulation of gene copy would repress the gene expression

before it is replicated and save the cellular resources for replication, transcription, and translation. However, other level of regulations like transcriptional, translational, or post-translational control, repress the gene expression after the gene is replicated, transcribed, or even translated, which would cause a waste of cellular resources for the unnecessary gene expression. Besides, regulation at the replication level is generally more responsive compared with other level of controls that take place after the gene is replicated. An additional advantage of the dynamic gene copy control via controllable plasmid replication is the ease of achieving multiplex regulation. As demonstrated in the example of *p*-coumaric acid production, our approach enabled a simultaneous dynamic regulation of six targets in a plasmid using just one regulatory circuit. This feature can be especially useful when regulating complex pathways that involves multiple genes. Moreover, dynamic gene copy regulation can be combined with transcriptional, translational, or post-translational control to establish multi-layered dynamic regulation on multiple targets. Overall, the dynamic regulation of gene copy enabled a new level of dynamic control on gene expression and offered an additional layer of regulation on dynamic pathway control, which would be beneficial for engineering microbial production of valuable compounds in metabolic engineering.

However, the excessive use of plasmids is often not preferred by industrial biotechnology, mainly due to the low stability of the plasmid (Rugbjerg et al., 2018; Silva et al., 2012). Thus, developing stable plasmid regulation system will be necessary to improve the applicability of the dynamic gene copy control. There have been many recent engineering efforts at establishing strategies to stably maintain the plasmids (Kang et al., 2018; Zhang et al., 2019). Thus, future research can be conducted to combine such strategies and develop stable plasmid regulation systems for dynamic regulation of gene expression in metabolic engineering and synthetic biology.

2.5 Materials and Methods

2.5.1 Strains, plasmids and reagents

All plasmids and strains used in this study are listed in **Table S2.3**. The *E. coli* strain XL1-Blue (Stratagene) was used for gene cloning and plasmid constructions. The *E. coli* strain BW25113 (F') (Atsumi et al., 2008) was used for validation of controllable replication control of origins and dynamic performance test of genetic circuits. P1 transduction was used to sequentially knock out the *pykA*, *pykF*, *tyrR*, and *pheA* in *E. coli* strain BW25113 (F'), forming the strain BW-PCA that was used for *p*-coumaric acid biosynthesis. The high-fidelity Phusion DNA polymerase, restriction enzymes and Quick Ligase kit were purchased from NEB (New England Biolabs). Plasmid miniprep and DNA gel purification kits were purchased from Zymo Research (Irvine, CA, USA). Luria-Bertani (LB) medium was made from the LB powder purchased from Fisher Scientific and was used for plasmid propagation and cell inoculation. The M9Y medium containing 20 g/L glycerol, 5 g/L yeast extract, 6 g/L Na_2HPO_4, 0.5 g/L NaCl, 3 g/L KH_2PO_4, 1 g/L NH_4Cl, 246.5 mg/L $MgSO_4 \cdot 7H_2O$ and 14.7 mg/L $CaCl_2 \cdot 2H_2O$ was used for production of *p*-coumaric acid. Antibiotics including ampicillin (100 mg/L), kanamycin (50 mg/L), or chloramphenicol (34 mg/L) were added to cultures as needed. The *p*-coumaric acid standard was purchased from MP Biomedical LLC.

2.5.2 DNA manipulation and genetic circuits construction

The plasmids constructed in this study were listed in **Table S2.3**. The DNA sequences of the components used in this study were listed in **Table S2.4**. The reporter plasmid pCS-lpp0.5-egfp was constructed in our previous study (Wang, J. et al., 2017a), and the reporter plasmid pZE-lpp0.5-egfp was constructed by inserting the expression cassette of lpp0.5-egfp (containing lpp0.5 constitutive promoter, RBS, and egfp encoding gene) from pCS-lpp0.5-egfp into the plasmid

pZE12luc using XhoI and XbaI. The pZE-lpp0.5-RFP was constructed by replacing the coding sequence of egfp in pZE-lpp0.5-egfp with the coding sequence of red fluorescence protein (RFP) using Acc65I and XbaI. The plasmid pZE12luc harboring the ColE1 origin was used as the template for amplifying the DNA sequence of RNAI (ColE1) by polymerase chain reaction (PCR). The fragment of RNAI (ColE1) was inserted into the multi-cloning site (MCS) of pCS27 using EcoRI and BamHI to form plasmid pCS-RNAI (ColE1). The plasmid pCS27 harboring the p15A origin was used as the template for amplifying the DNA sequence of RNAI (p15A) and asRNAI (p15A) by PCR. The fragments of RNAI (p15A) and asRNAI (p15A) were inserted into the multi-cloning site (MCS) of pZE12luc and pZE-PT (Yang et al., 2015) using EcoRI and XbaI, or Acc65I and BamHI to form the plasmid pZE-RNAI (p15A), pZE-PT_RNAI (p15A), pZE-asRNAI (p15A), pZE-PT_asRNAI (p15A) respectively. The RNAIopt (p15A) were generated by overlap extension PCR using RNAI (p15A) as the template. pZE-PT_RNAI (ColE1) was constructed by replacing the RNAI (p15A) in the pZE-PT_RNAI (p15A) using Acc65I and BamHI. pSC-PT_RNAI (ColE1) was constructed by inserting the PT_RNAI (ColE1) transcriptional cassette from pZE-PT_RNAI (ColE1) to plasmid pSC74 (Wang, J. et al., 2020) using XhoI and XbaI. For the CRISPRi-based replication control of ColE1 origin, the sgRNA-TIS (ColE1) was designed to target the promoter region (from -7 to -27 bp) of RNAII, and the sgRNA-CDS (ColE1) was designed to target the transcript of RNAII (from +43 to +63 bp). Similarly, the sgRNA-TIS (p15A) was designed to target the promoter region (from -25 to -45 bp) of RNAII, and the sgRNA-CDS (p15A) was designed to target the transcript of RNAII (from +16 to +36 bp). The sgColE1 was designed to target the promoter region of RNAI in the ColE1 origin (from -3 to -23 bp) and the sgp15A was designed to target the promoter region of RNAI in the p15A origin (from +1to -20 bp). All DNA sequences of sgRNAs were listed in **Table S2.4**. The DNA sequences of all designed sgRNAs

were then constructed in pCS27 plasmid to form the pCS-sgRNA-TIS (ColE1), pCS-sgRNA-CDS (ColE1), pCS-sgRNA-TIS (p15A), pCS-sgRNA-CDS (p15A), pCS-sgColE1, and pCS-sgp15A. The transcriptional operons of the sgRNAs were constructed to plasmid pSA-dCas9 (Wang, J. et al., 2017b) using AatII and XhoI, resulting in pSA-dCas9-sgRNA-TIS (ColE1), pSA-dCas9-sgRNA-CDS (ColE1), pSA-dCas9-sgRNA-TIS (p15A), pSA-dCas9-sgRNA-CDS (p15A), pSA-dCas9-sgColE1, and pSA-dCas9-sgp15A.

For genetic circuits construction, the plasmids pZE-lpp0.2-PadRwt and pZE-lpp0.2-PadR (K64A) were constructed by inserting the PadR expression cassettes from pCS-lpp0.2-PadRwt and pCS-lpp0.2-PadR (K64A) (Jiang et al., 2021) to pZE12luc using XhoI and AvrII. The PT_RNAI (p15A) and asRNAI (p15A) were constructed to pZE-P9-egfp using EcoRI and XbaI. The operons of P9-PT_RNAI (p15A) and P9-asRNAI (p15A) were constructed into pZE-lpp0.2-PadRwt and pZE-lpp0.2-PadR (K64A) to form pZE-lpp0.2-PadRwt-P9-PT_RNAI (p15A), pZE-lpp0.2-PadR (K64A)-P9-PT_RNAI (p15A), pZE-lpp0.2-PadRwt-P9-asRNAI (p15A), and pZE-lpp0.2-PadR (K64A)-P9-asRNAI (p15A). The tetR gene was placed under the control of promoter $P_{padC\text{-}FG}$. The expression operon of $P_{padC\text{-}FG}$-tetR was inserted to pZE-lpp0.2-PadRwt, pZE-lpp0.2-PadR (K64A), pCS-lpp0.2-PadRwt, and pCS-lpp0.2-PadR (K64A), resulting in plasmids pZE-lpp0.2-PadRwt-$P_{padC\text{-}FG}$-tetR, pZE-lpp0.2-PadR (K64A)-$P_{padC\text{-}FG}$-tetR, pCS-lpp0.2-PadRwt-$P_{padC\text{-}FG}$-tetR, and pCS-lpp0.2-PadR (K64A)-$P_{padC\text{-}FG}$-tetR. The pL_{tetO1}-controlled transcriptional cassette of PT_RNAI (p15A) was then inserted to plasmids pZE-lpp0.2-PadRwt-$P_{padC\text{-}FG}$-tetR and pZE-lpp0.2-PadR (K64A)-$P_{padC\text{-}FG}$-tetR, forming pZE-lpp0.2-PadRwt-$P_{padC\text{-}FG}$-tetR-pL_{tetO1}-PT_RNAI (p15A) and pZE-lpp0.2-PadR (K64A)-$P_{padC\text{-}FG}$-tetR- pL_{tetO1}-PT_RNAI (p15A). For controlling ColE1 origin, the $P_{padC\text{-}FG}$-controlled transcriptional cassettes of sgRNA-TIS (ColE1) or sgColE1,

and the pL_{tetO1}-controlled transcriptional cassette of sgRNA-TIS (ColE1), were constructed to pSA-dCas9, resulting in plasmids pSA-dCas9-$P_{padC\text{-}FG}$-sgRNA-TIS (ColE1), pSA-dCas9-$P_{padC\text{-}FG}$-sgColE1, and pSA-dCas9-pL_{tetO1}-sgRNA-TIS (ColE1). The plasmid for *p*-coumaric acid biosynthesis was constructed by inserting the RgTAL expression operon from pZE-RgTAL into plasmid pCS-TPTA, resulting in pCS-TPTA-RgTAL. The pL_{lacO1}-controlled transcriptional cassette of asppc was constructed to pCS-TPTA-RgTAL to form the final plasmid pCS-TPTA-RgTAL-asppc.

2.5.3 Validation of controllable replication of ColE1 derived origins

The *E. coli* transformants were inoculated and cultivated for 12 h in 3.5 mL LB media at 37℃ as seeds. 150 μl culture was transferred into test tubes with 3.5 mL fresh LB medium. IPTG (0.5 mM, final concentration) was used to induce the pL_{lacO1} promoter after 2 h of inoculation. The samples were collected at 14 h of inoculation. The cell densities (OD_{600}, optical density at 600 nm) and the expression level of egfp (green fluorescence intensity) were measured using the plate reader *Synergy HT* purchased from BioTek. All values obtained from the plate reader misused the value of a blank control (pure water). For detecting the green fluorescence intensity, the emission and excitation parameter was set to 528/20 and 485/20, respectively.

2.5.4 Test dynamic performance of genetic circuits

The *E. coli* transformants were inoculated and cultivated for 12 h in 3.5mL LB media at 37℃ as seeds. 150 μl culture was transferred into test tubes with 3.5 mL fresh LB medium. IPTG (0.5 mM, final concentration) was used to induce the pL_{lacO1} promoter and gradient concentrations of *p*-coumaric acid (0, 100 mg/L, 500 mg/L) were added into media after 2 h of inoculation. The

samples were collected every 2 h for 14 h. The cell densities (OD_{600}) and the expression level of egfp (green fluorescence intensity) were measured using the plate reader *Synergy HT* purchased from BioTek. All values obtained from the plate reader misused the value of a blank control (pure water). For detecting the green fluorescence intensity, the emission and excitation parameter was set to 528/20 and 485/20, respectively.

2.5.5 Relative plasmid copy number determination by quantitative PCR

The quantification of relative plasmid copy number was carried out by quantitative polymerase chain reaction (qPCR) (Skulj et al., 2008). Cultured samples were diluted to OD_{600}=1.5 (corresponding to $1.5*10^6$ cells/μL) and were centrifuged at 12,000 rpm for 2 min. The plasmid mixtures from every sample were purified using Zymo plasmid miniprep kit with 40 ul water for dissolving the final products. The plasmid mixtures were diluted 100 times (1:100) by water before used as templates for qPCR. The primer pairs (**Table S2.5**) targeting the antibiotic markers (Amp^R, Kan^R, and Chl^R) were designed by inputting the corresponding sequences into Primer 3 (https://bioinfo.ut.ee/primer3-0.4.0/). The designed primers were then employed to specifically amplify the plasmids derived from pZE12 (Amp^R), pCS27 (Kan^R), and pSA74 (Chl^R), respectively. The sybr green supermix purchased from BIO-RAD was used for qPCR. The QuantStudio 3 Real-time PCR system (Thermo Fisher Scientific) was used for amplification and signal detection. The $2^{-\Delta\Delta Ct}$ method was used to analyze the relative plasmid concentration after normalization with the corresponding controls.

2.5.6 *p*-coumaric acid biosynthesis and product analysis

The plasmid pCS-TPTA-RgTAL-asppc were used for *p*-coumaric acid biosynthesis. The plasmid harboring the circuit with the third control logic (pZE-lpp0.2-PadR (K64A)-$P_{padC\text{-}FG}$-tetR-P9-PT_RNAI (p15A)) or the plasmid harboring the circuit with the second control logic (pZE-lpp0.2-PadR (K64A)-P9-asRNAI (p15A)) were used to control the plasmid replication of pCS-TPTA-RgRAL-asppc. The transformants were inoculated in LB media for 12 h as seeds for fermentation. 1 mL seeds were then transferred into shake flasks with 20 mL fresh M9Y media and cultivated at 37℃. After 2.5 h of cultivation, IPTG was added to every culture with a final concentration of 0.5 mM to induce the *p*-coumaric acid pathway. The samples were collected every 12 h for 72 h and then subject to cell density measurement and HPLC analysis. The cell density (OD_{600}) was measured using the V-1200 spectrophotometer (VWR) with the pure water as a blank. The concentrations of *p*-coumaric acid were quantitatively analyzed by Agilent HPLC 1260 Infinity II (1260 Infinity II Diode Array Detector WR) with a reverse-phase ZORBAX SB-C18 column. A methanol-water (containing 0.1% trifluoroacetic acid) gradient system at a flow rate of 1 mL/min was used. The analyzing method was set as follows: 5% methanol from 0-2 min, 5 to 80% methanol from 2-10 min, maintain 80% methanol from 10-14 min, 80 to 5% methanol from 14-18 min, and 5% methanol from 18-20 min. Standard curve was drawn using gradient concentrations of *p*-coumaric acid solution prepared from commercially available *p*-coumaric acid powder (MP Biomedical LLC). The retention time of *p*-coumaric acid is at around 8.89 min.

2.6 Tables and Figures

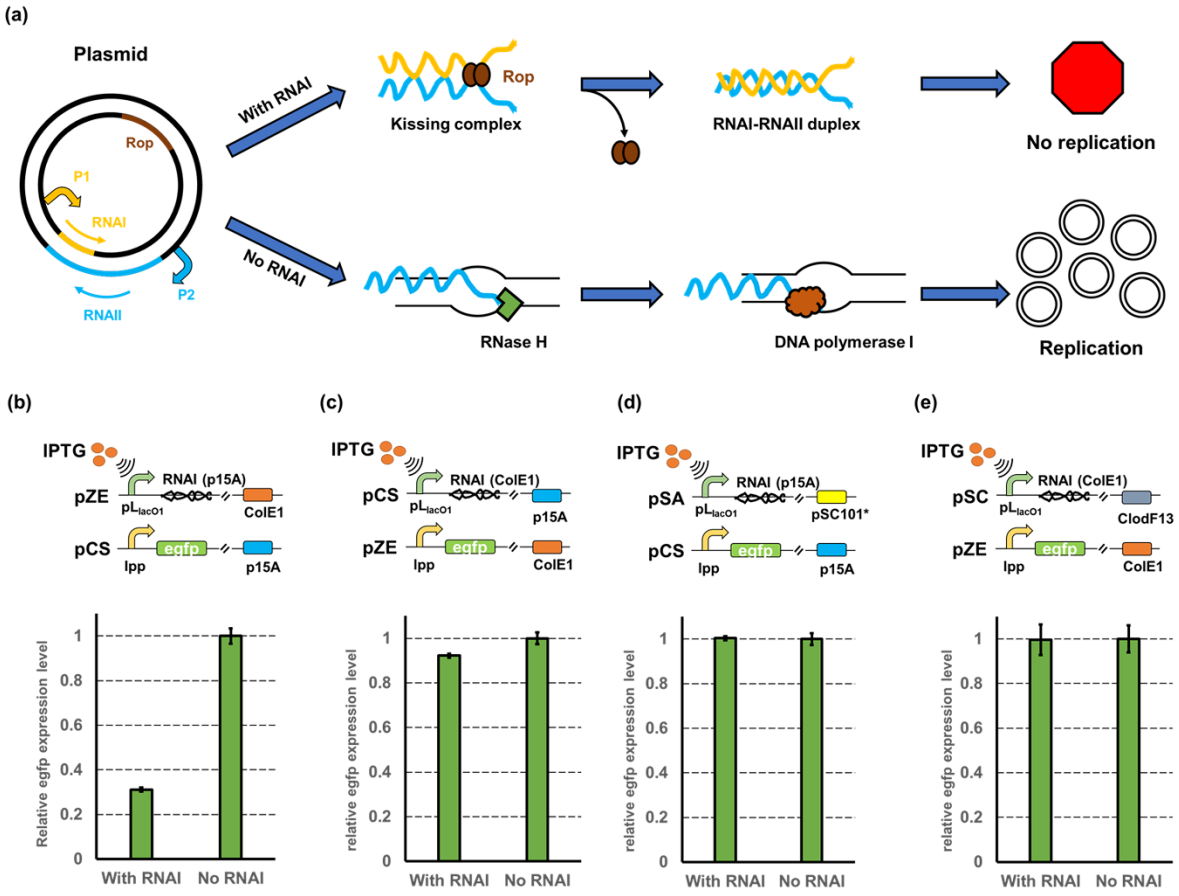


Figure 2.1. **Validation of the controllable inhibition on plasmid replication by overexpressing RNAI. (a)** Mechanism of how antisense RNAs regulate the replication of ColE1/p15A origin. **(b)** The inhibition efficiency on medium-copy plasmid pCS harboring the p15A origin. The RNAI of p15A was overexpressed in a high-copy plasmid pZE and the green fluorescence intensity was used for measuring the expression level of the reporter gene. **(c)** The inhibition efficiency on high-copy plasmid pZE harboring the ColE1 origin. The RNAI of ColE1 was overexpressed in a medium-copy plasmid pCS and the green fluorescence intensity was used for measuring the expression level of the reporter gene. **(d)** The inhibition efficiency on plasmid pCS when using a low-copy plasmid pSA to over-transcribe the RNAI (p15A). **(e)** The inhibition efficiency on

plasmid pZE when using a high-copy plasmid pSC to over-transcribe the RNAI (ColE1). All data represent the mean of 3 biologically independent samples and error bars show standard deviation.

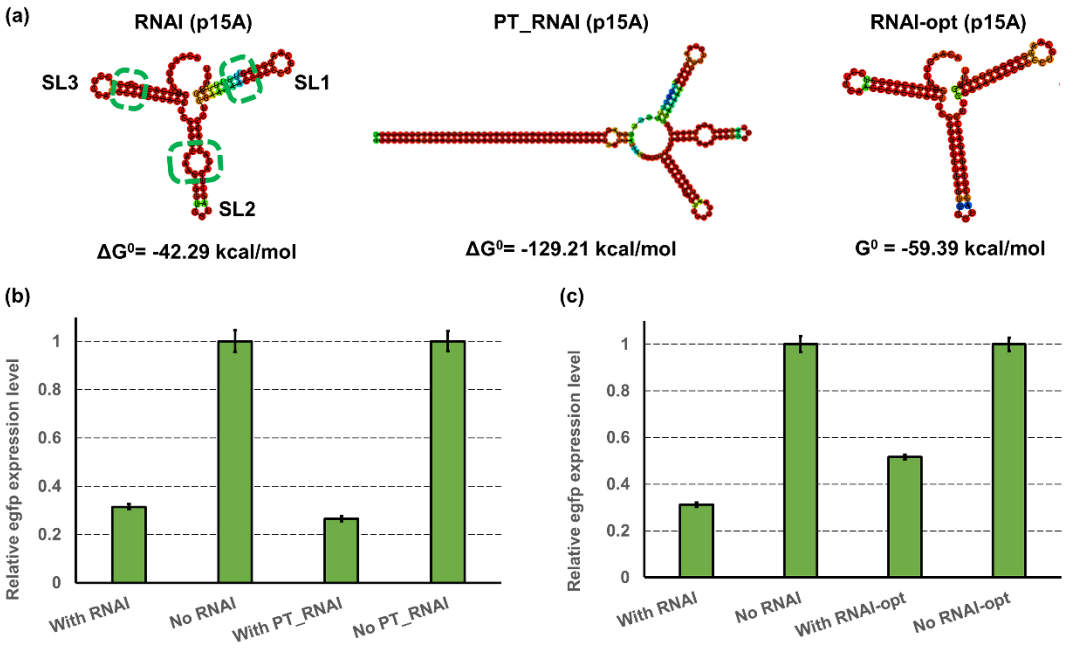


Figure 2.2. Optimizing the secondary structure of RNAI for enhanced inhibition efficiencies. **(a)** The simulated secondary structures of wild-type RNAI (left), PT_RNAI (middle), and RNAI-opt (right). The structure predictions and their corresponding Gibbs free energy calculations were performed using a web-based prediction tool, RNAfold (http://rna.tbi.univie.ac.at//cgi-bin/RNAWebSuite/RNAfold). The green dashed circles label the imperfect match in each stem. **(b)** the inhibition efficiency on medium-copy plasmid pCS by RNAI (p15A) and PT_RNAI (p15A). The RNAI or PT_RNAI of p15A was overexpressed in a high-copy plasmid pZE and the normalized green fluorescence intensity was used for measuring the expression level of egfp. **(c)** the inhibition efficiency on medium-copy plasmid pCS by optimized RNAI (p15A) and RNAI-opt. The RNAI and RNAI-opt were overexpressed in a high-copy plasmid pZE and the normalized green fluorescence intensity was used for measuring the expression level of the egfp. All data represent the mean of 3 biologically independent samples and error bars show standard deviation.

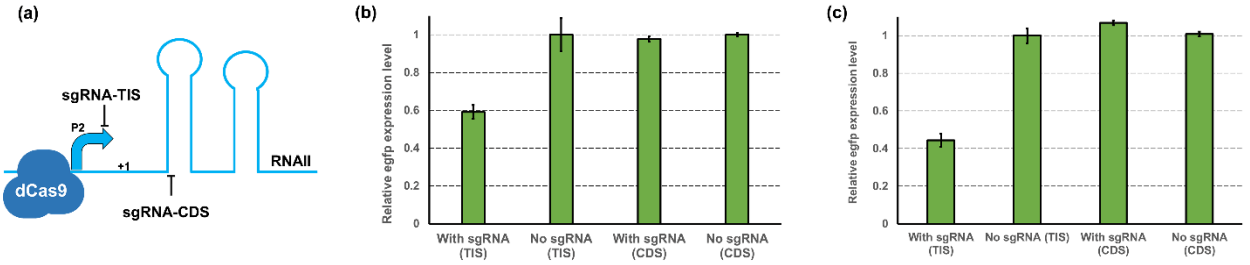


Figure 2.3. CRISPRi-mediated controllable inhibition of replication (a) Scheme of CRISPRi-mediated transcriptional repression on RNAII of ColE1. The sgRNA-TIS is targeting the promoter region (P2) of RNAII and the sgRNA-CDS is targeting the transcript sequence of RNAII. **(b)** The inhibition efficiency on high-copy plasmid pZE by sgRNA-TIS and sgRNA-CDS. These two sgRNA and the dCas9 protein were overexpressed in a low-copy plasmid pSA74. The normalized green fluorescence intensity was used for measuring the expression level of the reporter gene. **(c)** The inhibition efficiency on medium-copy plasmid pCS by sgRNA-TIS and sgRNA-CDS. These two sgRNA and the dCas9 protein were overexpressed in a low-copy plasmid pSA74. The normalized green fluorescence intensity was used for measuring the expression level of the reporter gene. All data represent the mean of 3 biologically independent samples and error bars show standard deviation.

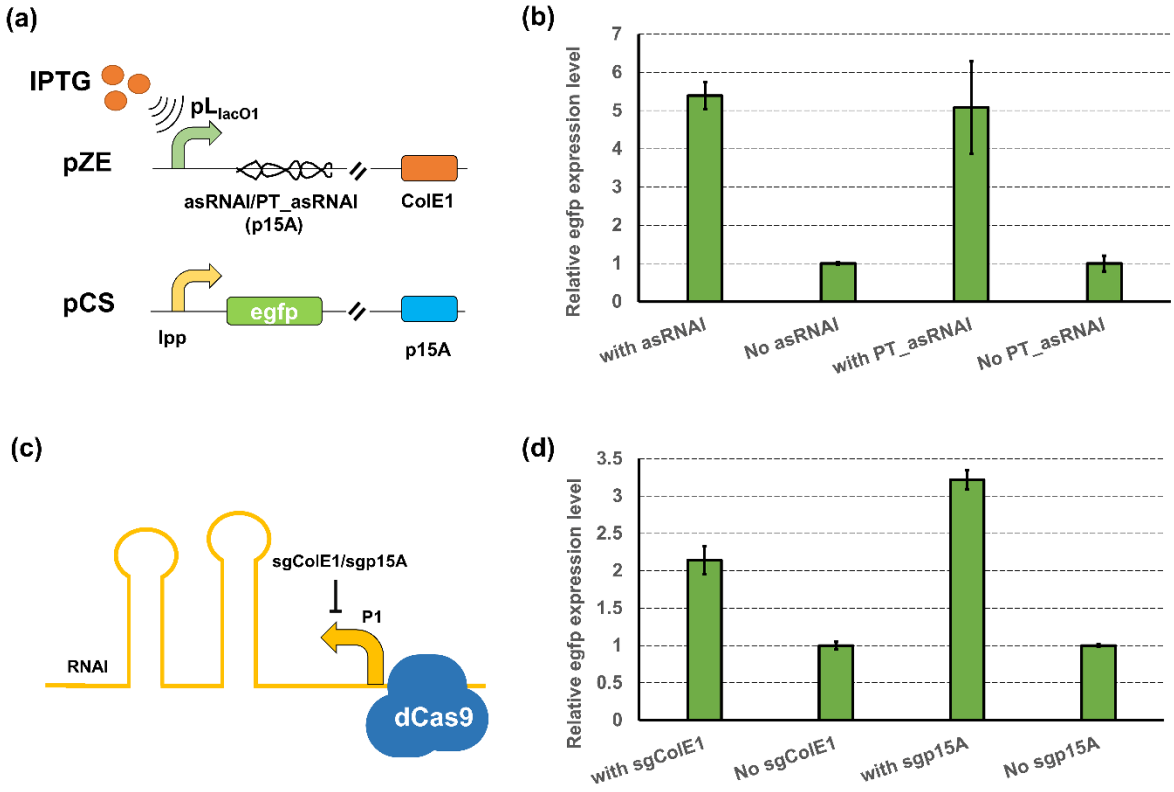


Figure 2.4. Controllable gene copy enhancement by reducing the RNAI availability. (a) Plasmid configuration of asRNAI- or PT_asRNAI-mediated repression on RNAI (p15A) availability **(b)** The activation efficiencies of asRNAI- or PT_asRNAI-mediated replication enhancement on p15A replication. **(c)** Scheme of CRISPRi-mediated transcriptional repression on RNAII of ColE1 **(d)**. The activation efficiency when using sgColE1 or sgp15A to target RNAI transcription in ColE1 or p15A origin, respectively. All data represent the mean of 3 biologically independent samples and error bars show standard deviation.

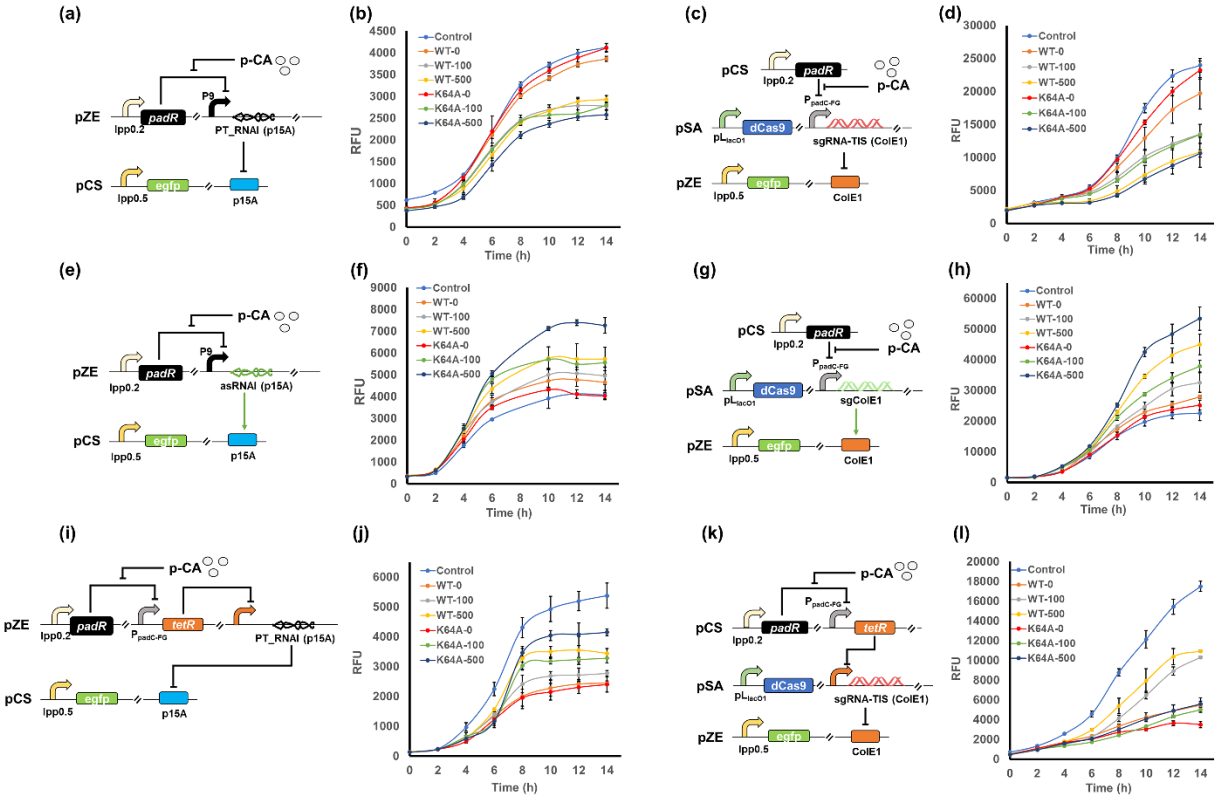


Figure 2.5. Constructing genetic circuits with different control logic to diversify the dynamic gene copy control strategy. **(a)** Scheme of the dynamic genetic circuit 1 targeting p15A origin using PT_RNAI (p15A). **(b)** The time-profiled egfp expression level when applying circuit 1 to regulate the replication of p15A origin. WT, circuit harboring the wild type PadR. K64A, circuit harboring PadR (K64A) variant. The number behind the dash indicated the concentration (mg/L) of *p*-coumaric acid. **(c)** Scheme of the dynamic genetic circuit 1 targeting ColE1 origin using sgRNA-TIS (ColE1). **(d)** The time-profiled egfp expression level when applying circuit 1 to regulate the replication of ColE1 origin. WT, circuit harboring the wild type PadR. K64A, circuit harboring the PadR (K64A) variant. The number behind the dash indicated the concentration (mg/L) of *p*-coumaric acid. **(e)** Scheme of the dynamic genetic circuit 2 targeting p15A origin using asRNAI (p15A). **(f)** The time-profiled egfp expression level when applying circuit 2 to regulate the replication of p15A origin. **(g)** Scheme of the dynamic genetic circuit 2 targeting

ColE1 origin using sgColE1. **(h)** The time-profiled egfp expression level when applying circuit 2 to regulate the replication of ColE1 origin. **(i)** Scheme of the dynamic genetic circuit 3 targeting p15A origin using PT_RNAI (p15A). **(j)** The time-profiled egfp expression level when applying circuit 3 to regulate the replication of p15A origin. **(k)** Scheme of the dynamic genetic circuit 3 targeting ColE1 origin using sgRNA-TIS (ColE1). **(l)**. The time-profiled egfp expression level when applying circuit 3 to regulate the replication of ColE1 origin. RFU, relative fluorescence unit. All data represent the mean of 3 biologically independent samples and error bars show standard deviation.

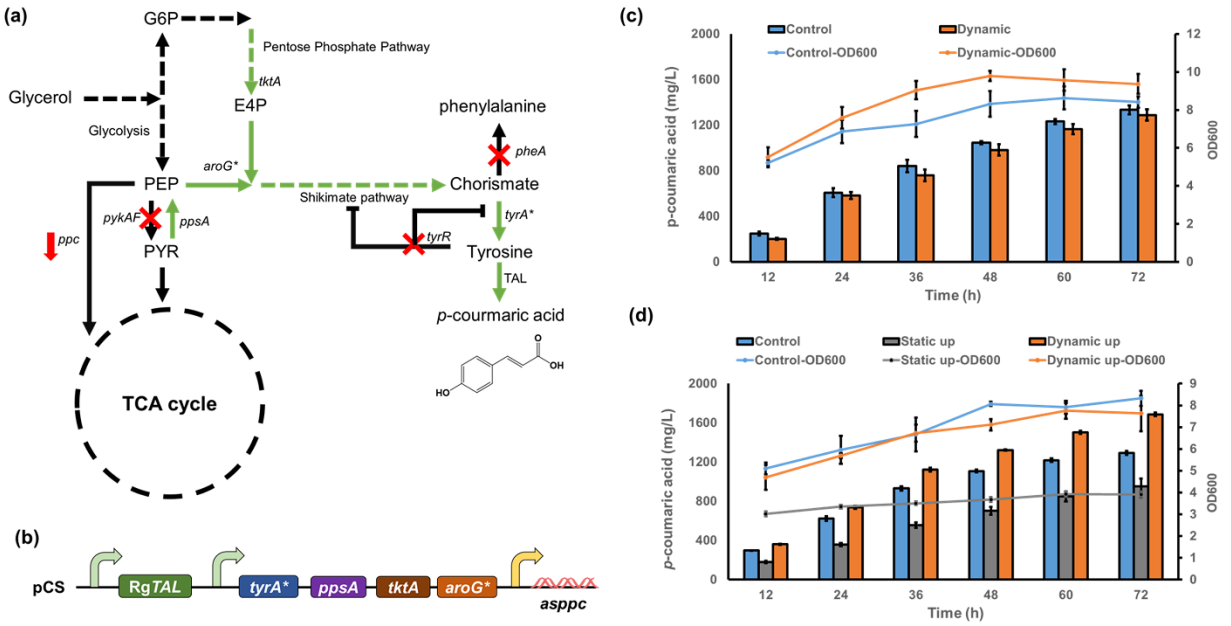


Figure 2.6. Implementing dynamic gene copy control in *p*-coumaric acid biosynthesis. (a) The *de novo* biosynthetic pathway for *p*-coumaric acid in *E. coli*. **(b)** The plasmid harboring the *p*-coumaric acid biosynthesis pathway, which consist of three operons and a total of six target genes. **(c)** Growth curves and *p*-coumaric acid production profile by dynamic "low-to-normal" strategy **(d)** Growth curves and *p*-coumaric acid production profile by dynamic "normal-to-high" strategy. All data represent the mean of 3 biologically independent samples and error bars show standard deviation.

2.7 Supplementary Information

Table S2.1. Quantitative PCR (qPCR) data for measuring the relative plasmid concentration

	Sample 1		Sample 2		Sample 3	
Target*	**pZE**	**pCS**	**pZE**	**pCS**	**pZE**	**pCS**
CTC value in Control condition (no IPTG added)**	17.481	19.502	18.016	19.405	18.764	19.708
CTE value in Experimental condition (with IPTG)**	16.992	20.965	17.422	20.753	18.522	21.055
ΔΔCT (ΔCTE-ΔCTC)	1.952		1.942		1.589	
Concentration fold change ($2^{(-\Delta\Delta CT)}$)	0.2585		0.2603		0.3323	
Inhibition efficiency ± standard deviation ($1-2^{(-\Delta\Delta CT)}$)	0.7163 ± 0.0421					

*The "pZE" in the target represents the plasmid pZE-RNAI (p15A), which was used as a control for measuring qPCR efficiency. The "pCS" represents the plasmid pCS-lpp0.5-egfp, which was the reporter plasmid.

**CTC: CT value in control condition; CTE: CT value in experimental condition. The relative PCN reduction on medium-copy plasmid pCS harboring the p15A origin when RNAI of p15A was

overexpressed in a high-copy plasmid pZE was calculated based on the ΔΔCT value. Three biological samples for each group were measured in this experiment.

Table S2.2. Quantitative PCR (qPCR) data for measuring the relative plasmid concentration

	Sample 1		Sample 2		Sample 3	
Target*	**pSA**	**pZE**	**pSA**	**pZE**	**pSA**	**pZE**
CTC value in Control condition (no IPTG added)**	24.365	17.919	25.075	18.330	24.784	17.796
CTE value in Experimental condition (with IPTG)**	24.752	19.211	24.477	19.014	24.356	18.568
ΔΔCT (ΔCTE-ΔCTC)	0.905		1.282		1.200	
Concentration fold change ($2^{(-\Delta\Delta CT)}$)	0.5339		0.4112		0.4354	
Inhibition efficiency ± standard deviation ($1-2^{(-\Delta\Delta CT)}$)	0.5398 ± 0.0650					

*The "pSA" in the target represents the plasmid pSA-dCas9-sgRNA-TIS (ColE1), which was used as a control for measuring qPCR efficiency. The "pZE" represents the plasmid pZE-lpp0.5-egfp, which was the reporter plasmid.

**CTC: CT value in control condition; CTE: CT value in experimental condition. The relative PCN reduction on medium-copy plasmid pCS harboring the p15A origin when RNAI of p15A was

overexpressed in a high-copy plasmid pZE was calculated based on the ΔΔCT value. Three biological samples for each group were measured in this experiment.

Table S2.3. Strains and plasmids used in this study

Strains	Genotype	References
E. coli BW25113 (F')	*rrnBT*14 Δ*lacZWJ*16 *hsdR*514 Δ*araBADAH*33 Δ*rhaBADLD*78 F′ [*traD*36 *proAB lacIqZ*Δ*M*15 *Tn*10(*Tet*^{r})]	(Atsumi et al., 2008)
E. coli XL1-Blue	*recA1 endA1gyrA*96*thi*-1*hsdR*17*supE*44*relA*1*lac*	Stratagene
E. coli BW-PCA	*E. coli* BW25113 (F') but Δ*pheA* Δ*tyrR* Δ*pykA* Δ*pykF*	This study
Plasmids		
pCS27	pL_{lacO1}; *p15A ori*; Kan^R	(Shen and Liao, 2008)
pZE12luc	pL_{lacO1}; *luc*; *ColE1 ori*; Amp^R	(Lutz and Bujard, 1997)
pSA74	pL_{lacO1}; *glk*; *pSC101* ori*, Cl^R	(Huo et al., 2011)
pCS-lpp0.5-egfp	pCS27 harboring the expression cassette of lpp0.5-egfp, which contains lpp0.5 constitutive promoter, RBS, egfp encoding gene and T1 terminator.	(Wang, J. et al., 2017a)
pZE-lpp0.5-egfp	pZE12luc with the luc expression cassette replaced by the expression cassette of lpp0.5-egfp, which contains lpp0.5 constitutive promoter, RBS, egfp encoding gene and T1 terminator	This study

pZE-lpp0.5-RFP	pZE-lpp0.5-egfp with the egfp gene replaced by the RFP gene	This study
pCS-RNAI (ColE1)	pCS27 harboring the coding sequence of RNAI for ColE1 origin under the control of pL_{lacO1} promoter	This study
pZE-RNAI (p15A)	pZE12luc with the luc gene and RBS replaced by the coding sequence of RNAI for p15A origin	This study
pSA-RNAI (p15A)	pSA74 with the glk gene and RBS replaced by the coding sequence of RNAI for p15A origin	This study
pSC74	pL_{lacO1}; *glk*; *ClodF13 ori*, Cl^R. derived from pSA74 by replacing its replication origin with CloDF13 origin from pCDFDuet-1	(Wang, J. et al., 2020)
pSC-RNAI (ColE1)	pSC74 with the glk gene and RBS replaced by the coding sequence of RNAI for ColE1 origin	This study
pZE-PT	pZE12-luc with the luc gene replaced by the PT template	(Yang et al., 2015)
pZE-PT_RNAI (p15A)	pZE-RNAI (p15A) with the addition of two flanked PT arms to the RNAI (p15A)	This study
pZE-PT_RNAI (ColE1)	pZE-RNAI (p15A) with the addition of two flanked PT arms to the RNAI (ColE1)	This study
pZE-RNAIopt (p15A)	pZE12luc with the luc gene and RBS replaced by the optimized coding sequence of RNAI for p15A origin	This study
pSA-dCas9	pSA74 harboring cas9 from *Streptococcus pyogene*s with mutation D10A/H841A	(Wang, J. et al., 2017b)

pCS-sgRNA-TIS (ColE1)	pCS27 harboring the pL_{lacO1}-controlled transcriptional cassette of sgRNA targeting the promoter region (from -7 to -27 bp) of RNAII (ColE1)	This study
pCS-sgRNA-CDS (ColE1)	pCS27 harboring the pL_{lacO1}-controlled transcriptional cassette of sgRNA targeting the RNAII (ColE1) transcripts (+43 to +63 bp)	This study
pCS-sgRNA-TIS (p15A)	pCS27 harboring the pL_{lacO1}-controlled transcriptional cassette of sgRNA targeting the promoter region (-25 to -45 bp) of RNAII (p15A)	This study
pCS-sgRNA-CDS (p15A)	pCS27 harboring the pL_{lacO1}-controlled transcriptional cassette of sgRNA targeting the RNAII (p15A) transcripts (+16 to +36 bp)	This study
pCS-sgColE1	pCS27 harboring the pL_{lacO1}-controlled transcriptional cassette of sgRNA targeting the promoter region (-3 to -23 bp) of RNAI (ColE1)	This study
pCS-sgp15A	pCS27 harboring the pL_{lacO1}-controlled transcriptional cassette of sgRNA targeting the promoter region (+1to -20 bp) of RNAI (p15A)	This study
pSA-dCas9-sgRNA-TIS (ColE1)	pSA-dCas9 harboring the pL_{lacO1}-controlled transcriptional cassette of sgRNA-TIS (ColE1)	This study

pSA-dCas9-sgRNA-CDS (ColE1)	pSA-dCas9 harboring the pL_{lacO1}-controlled transcriptional cassette of sgRNA-CDS (ColE1)	This study
pSA-dCas9-sgRNA-TIS (p15A)	pSA-dCas9 harboring the pL_{lacO1}-controlled transcriptional cassette of sgRNA-TIS (p15A)	This study
pSA-dCas9-sgRNA-CDS (p15A)	pSA-dCas9 harboring the pL_{lacO1}-controlled transcriptional cassette of sgRNA-CDS (p15A)	This study
pZE-asRNAI (p15A)	pZE12luc with the luc gene and RBS replaced by the complementary reversed (antisense) coding sequence of RNAI for p15A origin	This study
pSA-dCas9-sgColE1	pSA-dCas9 harboring the pL_{lacO1}-controlled transcriptional cassette of sgColE1	This study
pSA-dCas9-sgp15A	pSA-dCas9 harboring the pL_{lacO1}-controlled transcriptional cassette of sgp15A	This study
pSA-dCas9-Co_sgRNA	pSA-dCas9-sgRNA-TIS (ColE1) harboring the pL_{lacO1}-controlled transcriptional cassette of sgRNA targeting the promoter region (-25 to -45 bp) of RNAII (p15A)	This study
pCS-lpp0.2-PadRwt	pCS27 harboring the wild type PadR under the control of lpp0.2 promoter	(Jiang et al., 2021)
pCS-lpp0.2-PadR (K64A)	pCS27 harboring the PadR (K64A) variant under the control of lpp0.2 promoter	(Jiang et al., 2021)

pZE-lpp0.2-PadRwt	pZE12luc harboring the wild type PadR under the control of lpp0.2 promoter	This study
pZE-lpp0.2-PadR (K64A)	pZE12luc harboring the PadR (K64A) variant under the control of lpp0.2 promoter	This study
pZE-P9-egfp	pZE12luc with the pL_{lacO1} promoter replaced by P9 promoter	(Jiang et al., 2021)
pZE-P9-PT_RNAI (p15A)	pZE-P9-egfp with the egfp and RBS replaced by PT_RNAI (p15A)	This study
pZE-P9-asRNAI (p15A)	pZE-P9-egfp with the egfp and RBS replaced by asRNAI (p15A)	This study
pZE-lpp0.2-PadRwt-P9-PT_RNAI (p15A)	pZE-lpp0.2-PadRwt harboring the transcription cassette consist of P9 promoter, PT_RNAI (p15A), and T1 terminator	This study
pZE-lpp0.2-PadR (K64A)-P9-PT_RNAI (p15A)	pZE-lpp0.2-PadR (K64A) harboring the transcription cassette consist of P9 promoter, PT_RNAI (p15A), and T1 terminator	This study
pZE-lpp0.2-PadRwt-P9-asRNAI (p15A)	pZE-lpp0.2-PadRwt harboring the transcription cassette consist of P9 promoter, asRNAI (p15A), and T1 terminator	This study
pZE-lpp0.2-PadR (K64A)-P9-asRNAI (p15A)	pZE-lpp0.2-PadR (K64A) harboring the transcription cassette consist of P9 promoter, asRNAI (p15A), and T1 terminator	This study

pZE-lpp0.2-PadRwt-$P_{padC-FG}$-tetR	pZE-lpp0.2-PadRwt harboring tetR expression cassette under the control of $P_{padC-FG}$ promoter	This study
pZE-lpp0.2-PadR (K64A)-$P_{padC-FG}$-tetR	pZE-lpp0.2-PadR (K64A) harboring tetR expression cassette under the control of $P_{padC-FG}$ promoter	This study
pZE-lpp0.2-PadRwt-$P_{padC-FG}$-tetR-pL_{tetO1}-PT_RNAI (p15A)	pZE-lpp0.2-PadRwt-$P_{padC-FG}$-tetR harboring the PT_RNAI (p15A) transcriptional cassette under the control of pL_{tetO1} promoter	This study
pZE-lpp0.2-PadR (K64A)-$P_{padC-FG}$-tetR- pL_{tetO1}-PT_RNAI (p15A)	pZE-lpp0.2-PadR (K64A)-$P_{padC-FG}$-tetR harborin the PT_RNAI (p15A) transcriptional cassette under the control of pL_{tetO1} promoter	This study
pSA-dCas9-$P_{padC-FG}$-sgRNA-TIS (ColE1)	pSA-dCas9 harboring the $P_{padC-FG}$-controlled transcriptional cassette of sgRNA targeting the promoter region (from -7 to -27 bp) of RNAII (ColE1)	This study
pSA-dCas9-$P_{padC-FG}$-sgColE1	pSA-dCas9 harboring the $P_{padC-FG}$-controlled transcriptional cassette of sgRNA targeting the promoter region (-3 to -23 bp) of RNAI (ColE1)	This study
pSA-dCas9-pL_{tetO1}-sgRNA-TIS (ColE1)	pSA-dCas9 harboring the pL_{tetO1}-controlled transcriptional cassette of sgRNA targeting the promoter region (from -7 to -27 bp) of RNAII (ColE1)	This study

pCS-lpp0.2-PadRwt-$P_{padC-FG}$-tetR	pCS-lpp0.2-PadRwt harboring tetR expression cassette under the control of $P_{padC-FG}$ promoter	This study
pCS-lpp0.2-PadR (K64A)-$P_{padC-FG}$-tetR	pCS-lpp0.2-PadR (K64A) harboring tetR expression cassette under the control of $P_{padC-FG}$ promoter	This study
pCS-TPTA	pCS27, pL_{lacO1}; *tyrA*^{fbr}-*ppsA-tktA-aroG*^{fbr}	(Huang et al., 2013)
pZE-RgTAL	pZE12luc, pL_{lacO1}; TAL from *Rhodotorula glutinis*	(Huang et al., 2013)
pCS-TPTA-RgTAL	pCS-TPTA harboring the pL_{lacO1}-controlled expression operon of RgTAL	This study
pCS-TPTA-RgTAL-asppc	pCS-TPTA-RgTAL harboring the asppc transcriptional cassette under the control of pL_{lacO1} promoter	This study

Table S2.4. DNA sequences of components/genes used in the study

Name	DNA sequence (5'-3')
RNAI (ColE1)	ACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGG ATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTC AGCAGAGCGCAGATACCAAATACT
RNAI (p15A)	ACAAGTTTTGGTGACTGCGCTCCTCCAAGCCAGTTACCTCG GTTCAAAGAGTTGGTAGCTCAGAGAACCTTCGAAAAACCG CCCTGCAAGGCGGTTTTTTCGTTTT
PT_RNAI (p15A)	**AGGAGGAATTAACCATGCAGTGGTGGTGGTGGTGGTG**G GTACCACAAGTTTTGGTGACTGCGCTCCTCCAAGCCAGTTA CCTCGGTTCAAAGAGTTGGTAGCTCAGAGAACCTTCGAAA AACCGCCCTGCAAGGCGGTTTTTTCGTTTTGGATCC**CACCA CCACCACCACCACTGCATGGTTAATTCCTCCT** (sequence of PT structure was shown in bold)
RNAIopt (p15A)	ACAAGTTTTGGTGACTGGCTCCTCCAAGCCAGTTACCTCGG TTCTCTGAGTTGGTAGCTCAGAGAACCTTCGAAAAACCGCC CTGCAAGGCGGTTTTTCGTTTT
sgRNA-TIS (ColE1)	**CAGATTACGCGCAGAAAAAA**GTTTTAGAGCTAGAAATAG CAAGTTAAAATAAGGCTAGTCCG (20 bp spacer sequence of the target was shown in bold)
sgRNA-CDS (ColE1)	**AGAGTTGGTAGCTCTTGATC**GTTTTAGAGCTAGAAATAG CAAGTTAAAATAAGGCTAGTCCG

	(20 bp spacer sequence of the target was shown in bold)
sgRNA-TIS (p15A)	**TGATCTTCTTGAGATCGTTT**GTTTTAGAGCTAGAAATAGC AAGTTAAAATAAGGCTAGTCCG (20 bp spacer sequence of the target was shown in bold)
sgRNA-CDS (p15A)	**CGAAAAACCGCCCTGCAAGG**GTTTTAGAGCTAGAAATAG CAAGTTAAAATAAGGCTAGTCCG (20 bp spacer sequence of the target was shown in bold)
asRNAI (p15A)	AAACGAAAAAACCGCCTTGCAGGGCGGTTTTTCGAAGGTT CTCTGAGCTACCAACTCTTTGAACCGAGGTAACTGGCTTGG AGGAGCGCAGTCACCAAAACTTGT
PT_asRNAI (p15A)	**AGGAGGAATTAACCATGCAGTGGTGGTGGTGGTGGTG**G GTACCAAACGAAAAAACCGCCTTGCAGGGCGGTTTTTCGA AGGTTCTCTGAGCTACCAACTCTTTGAACCGAGGTAACTGG CTTGGAGGAGCGCAGTCACCAAAACTTGTGGATCC**CACCA CCACCACCACCACTGCATGGTTAATTCCTCCT** (sequence of PT structure was shown in bold)
sgColE1	**CTTCTAGTGTAGCCGTAGTT**GTTTTAGAGCTAGAAATAGC AAGTTAAAATAAGGCTAGTCCG (20 bp spacer sequence of the target was shown in bold)
sgp15A	**CCTTTCAGTTTAGCCTTAAC**GTTTTAGAGCTAGAAATAGC AAGTTAAAATAAGGCTAGTCCG (20 bp spacer sequence of the target was shown in bold)

P9 promoter	TAAATTATCTCTGGCGGTGTTGACATAAATACCACTGGCGG TGATACTGAACATGTAAATAGTTACATGAT
$P_{padC\text{-}FG}$ promoter	GGACTGTCTTCAAACAGTCCTTGTTTTTTTATGTTCCTATTG TTTGACAGTTAACTGCAATGGTGTTAAAGTGAACATGTAAA TAGTTACATGATTTTTTCTGAAGGTGAGGTGGTTCTCGTGA AGAAGCCGGTTTTAAAACCATTCGCCTCTTTAGAAATCAAG GTTGATCCGCCTATCACGATTGGTGAGACAAGCCTGGGACT GAGATGATTCATTCCGATTCGTTCGGGAACAATAACCGGG GAAGTAAAGGGACGTATTTTGCCGGGCGGTGCCGATTCAC AAATGATTCGCGCTAACGGCAGAACAGATTTATCTGCCAG GTATGTGATTGAAACAGCAGATCATGAACTGATTTACATTG AAAACAATGGAATACGGCAAGTCAGCAAGCCGTTTCGAAA ACAAGCGGCAGCCGGGGAAATTATTGAACCGGAGCATGTT TATTTTCGTACGGTACCGACGTTTGAAACAGGCAGTGAAGT CTATCAATGGCTCCATGACCGCTTGTTTATCGGTTCCGCAG AAAGAACCCCTGATTACGTTCTACTAGACATTTATGAAGTA CAGTAAAAGACTAAGGAGAGTGTGTAAG
padRwt	ATGAGAGTATTAAAATACGCCATATTAGGGCTTTTGCGAA AAGGCGAATTGAGTGGATACGATATTACGAGTTATTTTAA AGAGGAGCTCGGCCAGTTTTGGAGCGCCAAGCACAGCCAG ATTTACCCTGAGCTCAAAAAGCTGACGGATGAAGGATTTA TCACGTTCCGGACAACGATTCAGGGCACAAAGCTGGAGAA AAAGATGTACACCCTGACAGACAGCGGAAAGCAGGAGCTG

	CATGACTGGCTGATCCGCCACCAGCCGATACCCGAGACGG TGAAGGATGAATTTATGCTGAAGGCTTATTTCATTTCTTGT CTATCGCGCCAGGAGGCTTCTGATTTGTTCAAAGACCAGCT GCAAAAACGCCAGGCCAAGCTTTCTGACTTACAGGGAAGC TATGAAAAGCTTATGGCTTCAGCAGAGCCGATGTCATTTTC TTCACCGGACTTCGGCCACTATCTTGTGCTAACGAAAGCGC TGGAGCGGGAGAAAAATTACGTTTCTTGGCTGGAATCAAT TTTAGCTATGATAGATAAGGATTAA
padR (K64A) (The K64A mutation was in bold and highlighted in yellow)	ATGAGAGTATTAAAATACGCCATATTAGGGCTTTTGCGAA AAGGCGAATTGAGTGGATACGATATTACGAGTTATTTTAA AGAGGAGCTCGGCCAGTTTTGGAGCGCCAAGCACAGCCAG ATTTACCCTGAGCTCAAAAAGCTGACGGATGAAGGATTTA TCACGTTCCGGACAACGATTCAGGGCACA**GCC**CTGGAGAA AAAGATGTACACCCTGACAGACAGCGGAAAGCAGGAGCTG CATGACTGGCTGATCCGCCACCAGCCGATACCCGAGACGG TGAAGGATGAATTTATGCTGAAGGCTTATTTCATTTCTTGT CTATCGCGCCAGGAGGCTTCTGATTTGTTCAAAGACCAGCT GCAAAAACGCCAGGCCAAGCTTTCTGACTTACAGGGAAGC TATGAAAAGCTTATGGCTTCAGCAGAGCCGATGTCATTTTC TTCACCGGACTTCGGCCACTATCTTGTGCTAACGAAAGCGC TGGAGCGGGAGAAAAATTACGTTTCTTGGCTGGAATCAAT TTTAGCTATGATAGATAAGGATTAA

tetR	ATGTCTCGTTTAGATAAAAGTAAAGTGATTAACAGCGCATT
	AGAGCTGCTTAATGAGGTCGGAATCGAAGGTTTAACAACC
	CGTAAACTCGCCCAGAAGCTAGGTGTAGAGCAGCCTACAT
	TGTATTGGCATGTAAAAAATAAGCGGGCTTTGCTCGACGCC
	TTAGCCATTGAGATGTTAGATAGGCACCATACTCACTTTTG
	CCCTTTAGAAGGGGAAAGCTGGCAAGATTTTTTACGTAATA
	ACGCTAAAAGTTTTAGATGTGCTTTACTAAGTCATCGCGAT
	GGAGCAAAAGTACATTTAGGTACACGGCCTACAGAAAAAC
	AGTATGAAACTCTCGAAAATCAATTAGCCTTTTTATGCCAA
	CAAGGTTTTTCACTAGAGAATGCATTATATGCACTCAGCGC
	TGTGGGGCATTTTACTTTAGGTTGCGTATTGGAAGATCAAG
	AGCATCAAGTCGCTAAAGAAGAAAGGGAAACACCTACTAC
	TGATAGTATGCCGCCATTATTACGACAAGCTATCGAATTAT
	TTGATCACCAAGGTGCAGAGCCAGCCTTCTTATTCGGCCTT
	GAATTGATCATATGCGGATTAGAAAAACAACTTAAATGTG
	AAAGTGGGTCTTAA
asppc	**AGGAGGAATTAACCATGCAGTGGTGGTGGTGGTGGTG**G
	GTACCTGATGGTTTCTCCCAGCACTTTGCCGAGCATACTGA
	CATTACTACGCAATGCGGAATATTGTTCGTTCATATTACCC
	CAGACACCCCATCTTATCGTTTGGGATCC**CACCACCACCA**
	CCACCACTGCATGGTTAATTCCTCCT

(sequence of PT structure was shown in bold)

Table S2.5. Primer sequences for quantitative PCR

Primer name	DNA sequence (5'-3')
Amp^R-F	GCGCAACGTTGTTGCCATTG
Amp^R-R	CCGGAGCTGAATGAAGCCAT
Kan^R-F	ATACTTTCTCGGCAGGAGCA
Kan^R-R	TGAATGAACTGCAGGACGAG
Chl^R-F	TTTGCCCATGGTGAAAACGG
Chl^R-R	TGTTACGGTGAAAACCTGGC

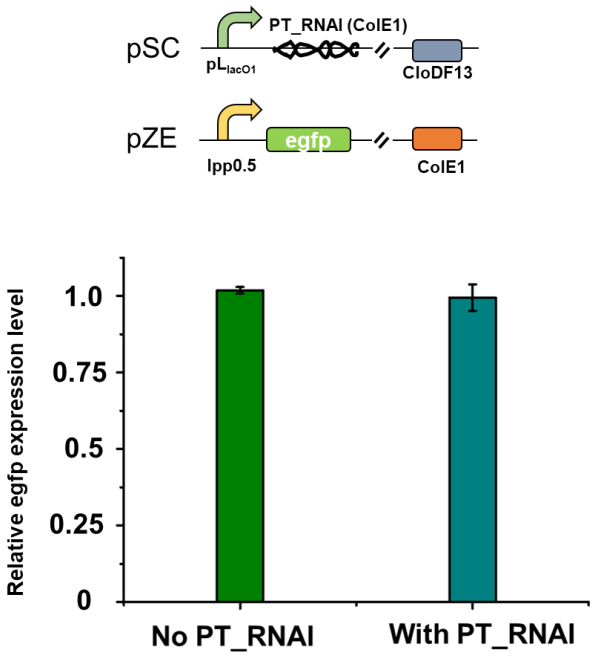
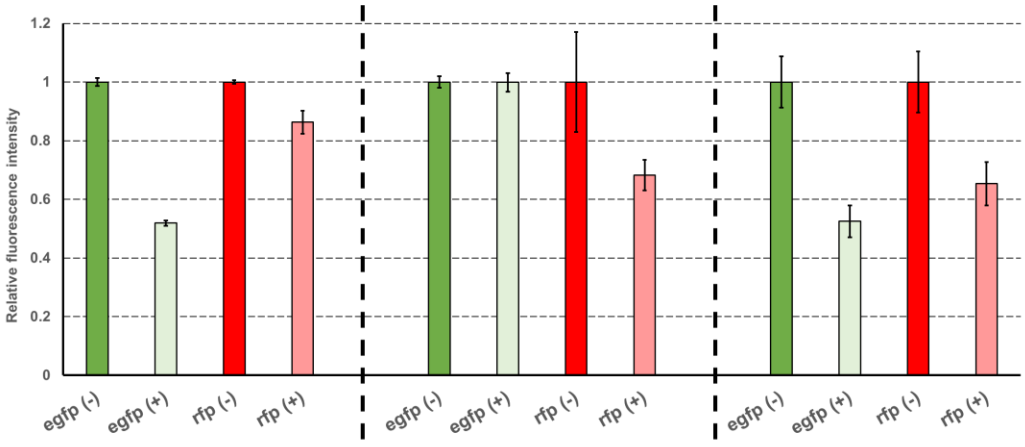


Fig. S2.1. Regulating the replication of the high-copy origin ColE1 using the pSC-PT_RNAI (ColE1). All data represent the mean of 3 biologically independent samples and error bars show standard deviation.



	Group 1	Group 2	Group 3
Inhibition on p15A (egfp)	Yes	No	Yes
Inhibition on ColE1 (rfp)	No	Yes	Yes

Fig. S2.2 Co-regulation of both p15A and ColE1 origin. The strains in all groups harbor both reporter plasmid pZE-lpp0.5-RFP and pCS-lpp0.5-egfp. The strain in group 1 only harbors the sgRNA target the p15A origin (sgRNA-TIS (p15A)). The strain in group 2 only harbors the sgRNA target the ColE1 origin (sgRNA-TIS (ColE1)). The strain in group 3 harbors both sgRNAs and can inhibit the replication of both origins at the same time. The "(-)" in the horizontal axis indicate that the strain was not induced by IPTG, and the "(+)" indicate the strain was induced by 0.5 mM IPTG. All data represent the mean of 3 biologically independent samples and error bars show standard deviation.

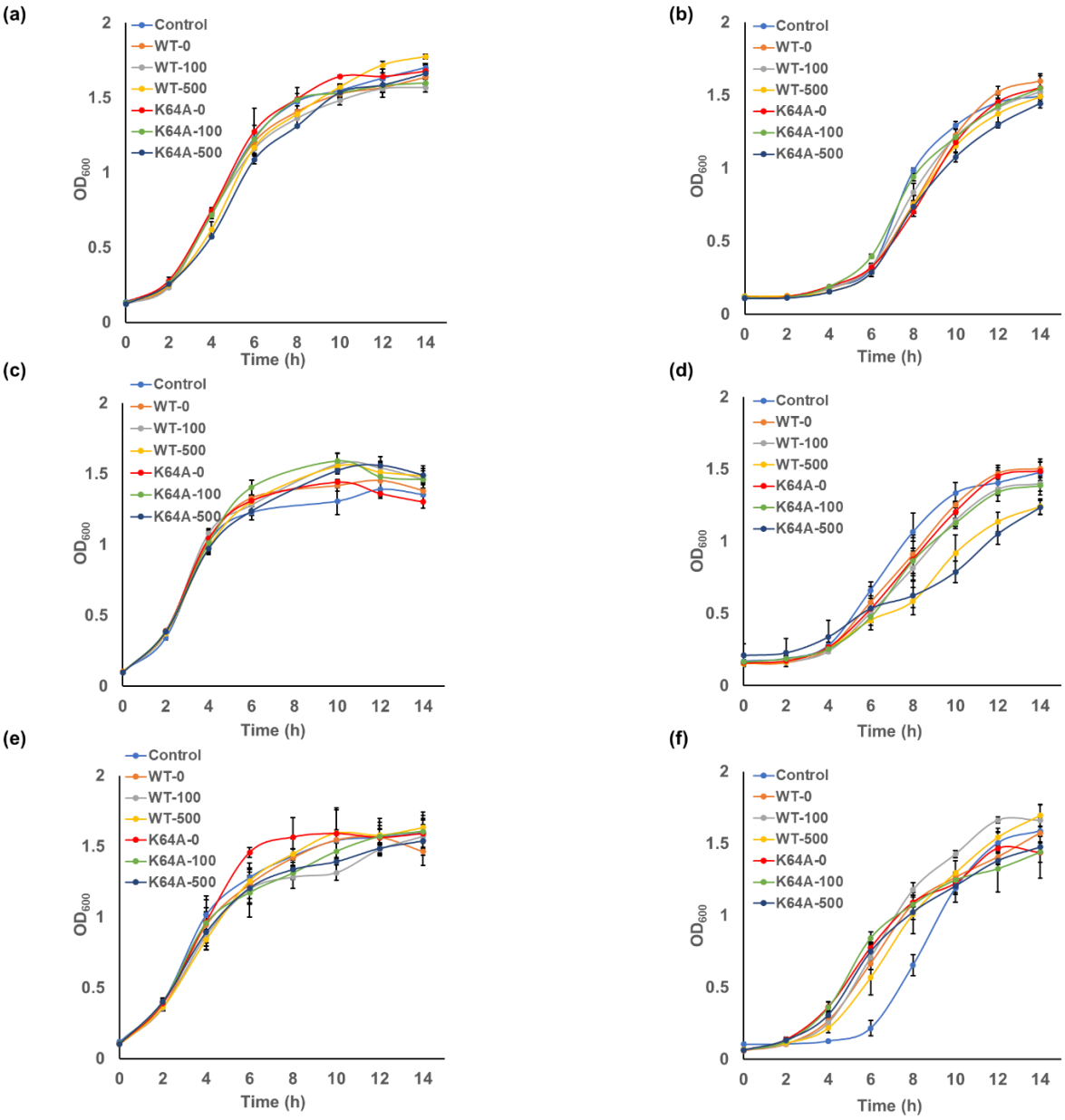


Fig. S2.3 Growth curves for strains harboring the circuit with the first control logic targeting p15A origin **(a)** and ColE1 origin **(b)**, circuit with the second control logic targeting p15A origin **(c)** and ColE1 origin **(d)**, and circuit with the third control logic targeting p15A origin **(e)** and ColE1 origin **(f)**. All data represent the mean of 3 biologically independent samples and error bars show standard deviation.

CHAPTER 3

BIOSENSOR-ASSISTED TITRATABLE CRISPRI HIGH-THROUGHPUT (BATCH) SCREENING FOR OVER-PRODUCTION PHENOTYPES [1]

1 Li, C., Wang, J., Jiang, T., Yan, Y., 2022. Biosensor-assisted titratable CRISPRi high-throughput (BATCH) screening for over-production phenotypes. *Submitted to Metabolic Engineering*

3.1 Abstract

With rapid advances in the development of metabolic pathways and synthetic biology toolkits, a persisting challenge in microbial bioproduction is on how to optimally rewire the metabolic flux and accelerate the concomitant high-throughput phenotype screening. Here we developed a biosensor-assisted titratable CRISPRi high-throughput (BATCH) screening approach that combines a titratable mismatch CRISPR interference and a biosensor mediated screening for high-production phenotypes in *Escherichia coli*. We first developed a programmable mismatch CRISPRi that could afford multiple levels of interference efficacy with an one-pot sgRNA pool (a total of 16 variants for each target gene) harboring two consecutive random mismatches in the seed region. The mismatch CRISPRi was demonstrated to enable almost a full range of gene knockdown when targeting different positions on genes. As a proof-of-principle demonstration of the BATCH screening system, we designed doubly mismatched sgRNA pools targeting 20 relevant genes in *E. coli* and optimized a PadR-based *p*-coumaric acid biosensor with broad dynamic range for the eGFP fluorescence guided high-production screening. Using sgRNA variants for the combinatorial knockdown of *pfkA* and *ptsI*, the *p*-coumaric acid titer was increased by 40.6% to o 1308.6 mg/l from glycerol in shake flasks. To further demonstrate the general applicability of the BATCH screening system, we recruited a HpdR-based butyrate biosensor that facilitated the screening of *E. coli* strains achieving 19.0% and 25.2% increase of butyrate titer in shake flasks with sgRNA variants targeting *sucA* and *ldhA*, respectively. This work reported the establishment of a plug-and-play approach that enables multilevel modulation of metabolic flux and high-throughput screening of high-production phenotypes.

3.2 Introduction

Metabolic engineering harnesses the ability of living microorganisms to degrade renewable and low-cost carbon feedstocks to produce pharmaceutically or industrially important chemicals (Keasling, 2010; Khosla and Keasling, 2003; Lee et al., 2019; Nielsen and Keasling, 2011). However, the sophisticatedly regulated carbon metabolisms naturally coined in the genomes of most microbes always devote to cell survival and produce target chemicals non-efficiently (Montaño López et al., 2022; Zha et al., 2009; Zu et al., 2020). To obtain high-production strains satisfying commercial relevance, traditional approach relies on either domestication of microbes or deletion of specific non-essential competing genes (Gallone et al., 2016; Park et al., 2007; Steensels et al., 2019). With increasingly available genomic or transcriptomic data and better understanding of cellular metabolisms, more efforts have been made to systematically interrogate and rewire metabolic network in both "localized" and "global" scales (Andreozzi et al., 2016; Feist and Palsson, 2008; Lee et al., 2021; Park et al., 2009).

To that end, programmable synthetic biology toolkits have been developed to expedite the genome-scale perturbation of metabolic fluxes, such as synthetic DNA oligonucleotides or CRISPR mediated multiplex genome engineering (Cong et al., 2013; Nyerges et al., 2016; Wang et al., 2009), retron-enabled genome editing (Lopez et al., 2022; Schubert et al., 2021; Simon et al., 2019), synthetic regulatory RNA mediated post-transcriptional interference (RNAi) (Na et al., 2013; Yang et al., 2015), and CRISPR-based transcriptional interference (CRISPRi) (Bikard et al., 2013; Gilbert et al., 2014; Hawkins et al., 2015). Among those, CRISPRi has been extensively used in systematic metabolic engineering due to its ease of programmability and multiplexity, robust activity, and host flexibility (Banerjee et al., 2020; Liu et al., 2022; Schultenkämper et al.,

2020; Wang, T. et al., 2018). However, the near knockout effect of CRISPRi always imposes static gene repression and causes cell fitness issue especially when targeting growth-essential genes (Cui et al., 2018; de Bakker et al., 2022; Wang et al., 2021a), rendering the approach unable to thoroughly investigate the impact of the expression levels of individual genes on the desired high-production phenotypes.

To address this limitation, one ideal strategy would be to design massively pooled sgRNA libraries that guide dCas9 only to their respective gene targets while simultaneously afford parallel interrogation of individual genes on microbial production when repressed to different extents. The CRISPRi efficacy could be modulated by engineered sgRNAs with random single mismatches (Jost et al., 2020; Qi et al., 2013), with interspersed double or multiple mismatches (Feng et al., 2021; Hawkins et al., 2020), with inducible hairpin or aptazyme-embedded structures (Ferry et al., 2017; Kocak et al., 2019; Siu and Chen, 2019; Tang et al., 2017), or with truncated spacers (Qi et al., 2013; Wang et al., 2021b). All these strategies require either lab-intensive synthesis of large sgRNA libraries or empirically tested sgRNA structures or specialized molecule inducers that have prevented their practical applications in metabolic engineering practices. Moreover, concomitant with large-scale sgRNA libraries, one recurring challenge would be the lack of efficient and cost-effective high-throughput CRISPRi screening methods for the desired production phenotypes.

In this work, we developed a biosensor-assisted titratable CRISPRi high-throughput (BATCH) screening approach that combines a titratable mismatch CRISPR interference and a biosensor mediated screening for high-production phenotypes in *Escherichia coli*. Based on the fact that either mismatch position, type or combination could impair CRISPRi activity, we first defined the

rule for a cost-effective one-pot mismatch sgRNA design that could enabled a full range of CRISPRi efficacy. Then, to use *p*-coumaric acid and butyrate as two proof-of-principle applications, we recruited their respective PadR and HpdR based biosensor-reporter system with broad dynamic properties. With mismatch sgRNA libraries targeting central carbon metabolisms, the BATCH screening system identified multiple target genes that enhanced the titer of both products. This work established a readily appliable approach of high-throughput and multilevel CRISPRi screen in metabolic engineering to obtaining desired high-production phenotypes with programmable and customizable mismatch sgRNA libraries.

3.3 Results

3.3.1 Design of BATCH screening system

The BATCH screening system is composed of two critical parts: (1) a titratable CRISPRi for customizable and multilevel gene repression in microbial chassis, and (2) chemical inducible biosensors with broad dynamic ranges for phenotype screening (**Fig. 3.1**). The titratable CRISPRi represses all relevant gene targets within the metabolic network or competing genes of the target biosynthetic pathways to variable extents. The biosensors are committed to sensing intracellular product chemicals and inducing the reporter gene eGFP expression, thus linking the product concentration with fluorescence intensity. Implementation of the BATCH screening system involves two rounds of selection: the first round selection of fluorescent single colonies on plates for test tube cultivation and the second round selection of highly fluorescent transformants from test tube cultures for shake flask cultivation validation. Combination of selected beneficial repressions could be iteratively launched to screen hyper-production phenotypes.

3.3.2 Establishing one-pot mismatch sgRNA libraries for titratable CRISPRi

First, to establish a titratable CRISPRi, we sought to alter the repression strength of target genes simply by engineering customizable sgRNA libraries. Since CRISPRi activity is highly sensitive to mismatch positions, types and combinations between the sgRNA and target DNA (Feng et al., 2021; Gilbert et al., 2014), we herein endeavored to establish a titratable CRISPRi method with a facile one-pot mismatch sgRNA library design.

To facilitate the CRISPRi system construction and also avoid a potential toxicity of SpdCas9 overexpression in *E. coli* cells (Calvo-Villamañán et al., 2020; Zhang and Voigt, 2018), we constructed a strain *E. coli*::dCas9 carrying SpdCas9 under the control of P_L*lacO1* promoter integrated into the genome at the low-expression *dkgB* locus (Bryant et al., 2014; Wang et al., 2021b). To generate mismatch sgRNAs, the inverse PCR was applied for introducing mutations in sgRNA spacers with degenerate primers (**Fig. 3.2A**). We first tested the single mismatch sensitivity of sgRNA on eGFP repression, by transforming *E. coli*::dCas9 with pCS-eGFP and pZE12-sgRNAs harboring mismatch sgRNAs targeting 5' coding sequence (+39 bp) of eGFP (**Fig. 3.2B and C**). The eGFP repression assay indicated that single mismatch outside a 7bp PAM-proximal seed position are tolerated without significantly affecting the CRISPRi efficiency (**Fig. 3.2C**). This was consistent with previous findings that mismatches within the generally recognized ~8bp seed region greatly compromised or abolished the CRISPR/(d)Cas9 functionality, at least in *E. coli* (Datsenko et al., 2012; Gilbert et al., 2014; Qi et al., 2013; Sternberg et al., 2015). To further investigate the sensitivity of double mismatches, we narrowed down to construct sgRNA variants harboring saturated double substitutions (NN) in PAM-proximal 7-8th bp seed region. The repression assay showed that the 16 possible sgRNA variants almost enabled a broad repression

profile of eGFP from 14.6% to 95.6%, suggesting that mismatching the 7-8th bp seed region could afford a multilevel gene repression (**Fig. 3.2D**). To examine the impact of target positions on the repression profile, we therefore designed double mismatch sgRNAs with random mutations of the 7-8th bp seed region that targeted both ribosome binding site (RBS) and alternate coding sequence (+378 bp) of eGFP with degenerate primers. In both cases, out of 32 randomly picked transformants, 12 different mismatch sgRNAs were obtained, accounting a coverage of 75% of total variants. Both sgRNA variants enabled a full range of knockdown levels of eGFP (**Fig. 3.2E and F**), further confirming the generable applicability of the one-pot mismatch sgRNA libraries for programmable and titratable CRISPRi.

3.3.3 Optimizing the *p*-coumaric acid biosensor PadR with broad dynamic behavior

For a proof-of-principle demonstration, we chose to target the *p*-coumaric acid biosynthetic pathway and optimized a PadR-based *p*-coumaric acid biosensor system for its high-production screening from glycerol. The *p*-coumaric acid has extensive applications in the food, pharmaceutical and cosmetic industry, and is also serving as an important precursor for more complex natural products such as flavonoids and polyphenols (Trantas et al., 2015; Wang et al., 2016).

To assist the high-production phenotype screening, natural or engineered chemical-inducible biosensors with broad dynamic range would be critical. An engineered *p*-coumaric acid-responsive biosensor system PadR-P_{padC} from *Bacillus subtilis* was established in our previous study (Jiang et al., 2021). The transcriptional repressor PadR could repress the expression of downstream gene controlled by P_{padC} promoter, while the presence of *p*-coumaric acid would activate the

transcription initiation (Jiang et al., 2021). To enable biosensor-assisted screening, four engineered hybrid P_{padC} promoters (P1, P2, P7, and P9) with optimized dynamic ranges and strong output strengths were selected to pair with the PadR. The PadR under control of the constitutive Plpp0.2 promoter and eGFP reporter gene under control of the hybrid promoters were assembled into pZE12-luc, resulting pZE-Plpp0.2-PadR-P1-eGFP, pZE-Plpp0.2-PadR-P2-eGFP, pZE-Plpp0.2-PadR-P7-eGFP, and pZE-Plpp0.2-PadR-P9-eGFP (**Fig. 3.3A**). To test the dynamic performance, all plasmids harboring the biosensor systems were transferred into *E. coli* BW25113(F′), and the transformants were cultivated with supplementation of gradient concentrations of *p*-coumaric acid (0, 100, 200, 400, 600, and 1000 mg/l). The PadR-P9 biosensor system exhibited the best performance in the selected concentration range, resulting in a 14.2-fold eGFP induction with 1000 mg/l *p*-coumaric acid (**Fig. 3.3B**). Thus, this combination was selected for the following BATCH screening of *p*-coumaric acid high-production phenotypes.

3.3.4 BATCH screening of *p*-coumaric acid over-production phenotypes

The *p*-coumaric acid pathway initiates from the aromatic amino acid (AAA) pathway and ends up with deamination of tyrosine to *p*-coumaric acid via tyrosine ammonia-lyase (TAL) (**Fig. 3.4A**). Extending from the central carbon metabolism, the *p*-coumaric acid pathway suffers from multiple potential competing pathways that divert the carbon fluxes. To redirect carbon flux to the *p*-coumaric acid pathway, we designed random double mismatch sgRNAs targeting 20 relevant genes in the glycolysis pathway (*ptsI*, *pgi*, *pfkA*, *fbaA*, *phoA*, *gapA*, *pgk*, *gpmA*, *gpmM*, *eno*, *pykA*, *pykF*, and *csrA*), the TCA cycle (*aceE*, *gltA*, and *ppc*), the fatty acid biosynthesis pathway (*accA*, and *fabD*), and the AAA pathway (*ubiC*, and *trpE*) (**Fig. 3.4A**).

To implement the BATCH screening for *p*-coumaric acid over-production phenotype, *E. coli*::dCas9 Δ*tyrR* Δ*pheA* with knockouts of the tyrosine repressor (*tyrR*) and the chorismate mutase/prephenate dehydratase (*pheA*) was applied as the host. *E. coli*::dCas9 Δ*tyrR* Δ*pheA* was co-transformed with plasmids pZE-CA (harboring RgTAL and PadR biosensor) and pCS-sgRNAs harboring mismatch sgRNA libraries targeting each gene (**Fig. 3.4B**). After transformation and growing on LB plates supplemented with 20 g/l glycerol and 0.5 mM IPTG, a total of 284 colonies with notably higher green fluorescence intensity on the plates were picked and cultivated in test tubes with M9 minimal medium containing 20 g/l glycerol (**Fig. S3.1**). Determination of the eGFP fluorescence intensities and *p*-coumaric acid concentrations of all selected transformants exhibited a moderate positive correlation (Pearson correlation coefficient r = 0.34) (**Fig. 3.4B**). For each target gene, the difference of eGFP fluorescence intensity and *p*-coumaric acid concentration among mismatch sgRNA variants underlined the impact of gene repression level on *p*-coumaric acid production (**Fig. S3.1**). Among all targets, sgRNA variants towards *csrA*, *fbaA*, *gapA*, *pfkA*, *phoA*, *ppc*, *ptsI*, and *pykF* produced more hits with increased *p*-coumaric acid titers (**Fig. S3.1**). Shake flask-based production tests with selected colonies from the test tube screening validated that mismatch sgRNAs towards *csrA*, *ptsI*, *pfkF*, *ppc*, and *pfkA* afforded more than 25% increase of *p*-coumaric acid titers (**Fig. 3.4C, and Fig. S3.1**). Especially, compared with the control strain (930.6 mg/l), sgpfkA-3 permitted the highest production of *p*-coumaric acid, reaching a peak titer of 1232.0 mg/l (**Fig. 3.4C**).

One notable advantage of mismatch sgRNAs in BATCH screening over authentic sgRNAs in conventional CRISPRi is the non-deleterious effect to cell fitness especially when targeting essential genes. To corroborate that, we chose to compare the functionality of the mismatch

sgRNAs and authentic sgRNAs towards growth-essential genes such as *ppc*, *ptsI* and *csrA*. As expected, authentic sgRNAs significantly impaired both cell growth (approximately 50% decrease) and *p*-coumaric acid titer, while selected sgRNA variants could improve *p*-coumaric acid production respectively by 28.4%, 25.0% and 24.6%, without sacrificing cell growth (**Fig. 3.4D**). This further underlined that conditional and optimal gene repression with BATCH screening system would expedite the rational high-production phenotype screening. One other advantage of BATCH screening is for easy testing of combinatorial effects of beneficial repressions from initial rounds. We therefore combined selected mismatch sgRNAs to further boost *p*-coumaric acid production. The combination of *sgpfkA*-3 and *sgptsI*-5 increased *p*-coumaric acid titer to 1308.6 mg/l, accounting for a 40.6% increase over the control strain (**Fig. 3.4E**).

3.3.5 Application of the BATCH system for screening of butyrate over-production phenotypes

To explore the applicability of the BATCH screening system to other production scenarios, we then implemented it to the screening of butyrate over-production phenotypes. Butyrate is a C4 fatty acid and a commodity chemical with wide applications in the field of chemical industry, food manufacturing, pharmaceuticals, perfume, and animal feed supplements (Guo et al., 2021; Jang et al., 2014). The butyrate pathway initiates from condensation of two acetyl-CoA to acetoacetyl-CoA followed by its reduction to 3-hydroxybutyryl-CoA with the enzyme thiolase (*thl*) and hydroxybutyryl-CoA dehydrogenase (*hbd*) (**Fig. 3.5A**). The 3-hydroxybutyryl-CoA was then reduced to butyryl-CoA by crotonase (*crt*) andtransenoyl-CoA reductase (*ter*), followed by the release of butyrate from butyryl-CoA with acetyl-CoA:acetoacetyl-CoA synthase (*atoDA*) (Choi et al., 2012; Saini et al., 2014). As the direct precursor for butyrate and also a critical central

metabolite, increasing the acetyl-CoA supply or reducing its consumption could enhance the butyrate titers.

To deploy the BATCH screening, we first sought to identify and characterize a butyrate biosensor with broad dynamic range. We previously engineered a promiscuous 3-hydroxybutyrate responsive HpdR-P_{hpdH} system on pCS-V2 plasmid, which harbors a deletion of a palindromic sequence (-209 to -118 bp) on the P_{hpdH} promoter and expresses the HpdR from *Pseudomonas putida* under control of Plpp1.0 promoter (Hanko et al., 2017; Wang et al., 2021b). To investigate the ligand scope of the engineered HpdR-P_{hpdH} system, *E. coli* BW25113(F′) transformed with pCS-V2 was subjected to C3-C5 fatty acids, hydroxy acids, and diacids with gradient concentrations (0-2.5 g/l). The eGFP fluorescence assay showed that the engineered HpdR-P_{hpdH} system exerted high sensitivity and broad dynamic range towards all C3-C5 fatty acids (propionate, butyrate, and valerate) and hydroxy acids (3-hydroxypropionate, 4-hydroxybutyrate, and 5-hydroxyvalerate) but not diacids (**Fig. 3.5B**). Particularly, among all ligands, butyrate afforded the highest eGFP induction fold (41-fold at 2.5 g/l) (**Fig. 3.5B**), indicating that the engineered HpdR-P_{hpdH} system could serve as an ideal butyrate biosensor.

To experimentally implement the BATCH screening, the engineered HpdR-P_{hpdH} system was assembled with the butyrate pathway, resulting in the plasmid pCS-HPE-THCTA (**Fig. 3.5C**). To reserve acetyl-CoA for butyrate biosynthesis, 14 mismatch sgRNA libraries involved in pyruvate consumption (*ldhA*), acetyl-CoA consumption (*ackA*, and *adhE*), fatty acid pathway (*accA* and *fabD*), TCA cycle (*gltA*, *acnA*, *acnB*, *icd*, *sucA*, *sucC*, *sdhC*, *sdhD*, *fumA* and *mdh*) were constructed into the pZE12-luc plasmid (resultant pZE-sgRNAs). *E. coli*::dCas9 was co-

transformed with pCS-HPE-THCTA and pZE-sgRNAs and subjected to BATCH screening. For over 150 variants screened in test tubes, the eGFP fluorescence intensities and butyrate titers were strongly correlated (Pearson correlation coefficient r = 0.69), confirming the applicability of biosensor-based screening of butyrate over-production phenotypes (**Fig. 3.5D, and Fig. S3.2**). The initial screening underlined that sgRNA variants targeting each gene enabled differential butyrate titers, and sgRNAs targeting *acnB, ldhA, sucA* and *sucC* generally afforded more sgRNA variants with enhanced butyrate titers (**Fig. S3.2**). Further validation of 27 selected over-production colonies in shake flask cultivations showed that the best sgRNA variants sgsucA-7 and sgldhA-10 increased butyrate titer by 19.0% and 25.2% over the control strain (2.90 g/l), reaching 3.45 and 3.63 g/l, respectively (**Fig. 3.5E**). These results together demonstrated the general applicability of BATCH screening in different production scenarios.

3.4 Discussion

Identification of competing gene targets and optimal control of their expression are critical in achieving high production of desired biochemicals in metabolic engineering (Ko et al., 2020; Na et al., 2013). Due to the intricately regulated nature and the essentiality in cellular maintenance of metabolic network, conventional gene knockout strategies would miss identification of potential targets and are often labor intensive. The RNAi or CRISPRi based screening has expedited large-scale profiling and repression of competing gene targets without modifying the genomes. However, the static gene repression imposed by RNAi or CRISPRi is sometimes suboptimal and even detrimental to metabolic pathways in production scenarios (Crook et al., 2016; Rousset et al., 2018). This work established a superior BATCH screening system that renders (1) titratable

CRISPRi using customizable mismatch sgRNA libraries for full-range gene repression, and (2) pinpointing the optimal gene repression with product biosensor based high-production screening.

The sensitivity of CRSIPRi towards mismatches enabled the fine-tuning of CRISPRi efficacy. In this work, we first located the mismatching window to the 7-8th bp within the seed region of sgRNA, whose random mutation could afford almost a full-scale of eGFP repression regardless of the targeting positions (**Fig. 3.2**). With designable degenerate primers, the mismatches could be easily created with one-pot PCR and incorporated into the sgRNA plasmid libraries, as indicated by the 75% coverage from 32 randomly picked colonies (**Fig. 3.2**). These two features add to the general applicability and manipulation simplicity of the BATCH screening system in metabolic engineering practices. The deployment of the BATCH system in *p*-coumaric acid production validated our hypothesis that repressing gene targets at different extents could dynamically affect the cell fitness and product formation. As demonstrated by repressing growth-essential genes including *ppc, ptsI* and *csrA*, repression of all three genes with authentic sgRNAs significantly impaired cell growth and *p*-coumaric acid production, while mismatch sgRNAs marginally affected cell growth but significantly increased *p*-coumaric acid production (**Fig. 3.4D**). Thus, the BATCH screening system allows the production-based screening of the competing gene targets and their optimal repressions in production scenarios.

The BATCH screening system relies on the biosensors of target chemicals with broad dynamic ranges. Chemical-inducible biosensors have been widely used to accelerate high-production screening and implemented into dynamic control paradigms for carbon metabolism rewiring (Seok et al., 2021; Zhang et al., 2015). However, the lack of biosensors or the poor dynamic performance

always limited their wide applications. In this study, we recruited two engineered biosensors, the PadR-P_{padC} system for *p*-coumaric acid and the HpdR-P_{hpdH} system for butyrate. When placed on the high-copy plasmid, while the native PadR-P_{padC} (P1) showed low-induction fold (3.9-fold) and eGFP expression even in the presence of 1 g/l *p*-coumaric acid, the engineered PadR-P_{padC} (P9) with mutations in P_{PadC} promoter showed markedly increased induction fold (14.2-fold) and elevated eGFP expression (**Fig. 3.3**). Particularly, biosensors for short-chain fatty acids (C3-C5) are rarely documented. We previously engineered a promiscuous 3-hydroxypropionate-responsive HpdR-P_{hpdH} system with significantly improved sensitivity and dynamic range towards 3-hydroxybutyrate (Wang et al., 2021b). Ligand profiling substantiated that the engineered HpdR-P_{hpdH} system could actually accept all C3-C5 fatty acids and counterpart hydroxy acids including butyrate (**Fig. 3.5B**). With broad dynamic ranges, both the engineered PadR-P_{padC} and HpdR-P_{hpdH} system rendered modest to strong positive correlations between eGFP fluorescence and product titers, which successfully guided the screening of their high-production *E. coli* cells. These chemical-inducible biosensors demonstrated here in phenotype screening could also be readily translated into other applications like sensor-mediated dynamic gene regulations.

In conclusion, we developed a new BATCH screening system that could be readily applied for large-scale target identification and multilevel gene repression with one-pot mismatch sgRNA libraries. We demonstrated its applicability to identify multiple gene targets for production improvement of *p*-coumaric acid and butyrate within single round of selection. Ideally, efficient assembly of sgRNA variants targeting multiple potential genes could also allow investigating the combinatorial effects with the BATCH screening system. In summary, the screening system

presented here would expedite the development of high-production microbial strains without genome intervention.

3.5 Materials and methods

3.5.1 Strains, medium and chemicals

All strains and plasmids used in this study are listed in **Table 3.1**. The *E. coli* XL1-Blue (Stratagene, La Jolla, CA) was used for plasmid construction and storage. The *E. coli*::dCas9 was created by integrating the P_L*lacO1*-dCas9 cassette into the genome of *E. coli* BW25113(F′) at *dkgB* locus and was used as the host strain for CRISPRi test and butyrate production. Further deletion of *pheA* and *tyrR* in *E. coli*::dCas9 resulted in the strain *E. coli*::dCas9 Δ*pheA* Δ*tyrR*, which was used as the host strain for *p*-coumaric acid production. The gene knockouts were conducted via P1 phage transduction method according to the standard protocol (Thomason et al., 2007). Luria-Bertani (LB) medium (10 g/l tryptone, 5 g/l yeast extract and 10 g/l NaCl) was used for seed culture preparation. The M9Y medium, prepared from M9 minimal medium (6 g/l Na_2HPO_4, 0.5 g/l NaCl, 3 g/l KH_2PO_4, 1 g/l NH_4Cl, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$) containing 5 g/l yeast extract and 20 g/l glycerol or 20 g/l glucose, was used for shake flask experiments for producing *p*-coumaric acid or butyrate, respectively. The antibiotics including ampicillin (100 μg/ml) and kanamycin (50 μg/ml) were added to the medium when necessary. Butyrate and *p*-coumaric acid standards were purchased from Sigma-Aldrich and MP Biomedical LLC, respectively.

3.5.2 Plasmid and strain construction

All plasmid constructions were carried out following the standard molecular cloning protocols (Sambrook et al., 1989). Phusion high-fidelity DNA polymerase, restriction DNA endonucleases and Quick Ligation kit were purchased from New England Biolabs (Beverly, MA, USA). The high-copy plasmid pHA-MCS were derived from pZE12-luc by replacing the *luc* gene with a multi-cloning site (MCS) and was applied as a negative control when needed. The fully matched sgRNA and rationally designed mismatch sgRNA variants targeting the eGFP coding sequence (starting from +39 bp) were constructed into the pCS-sgRNA scaffold plasmid using with *Apa*LI and *Bam*HI, resulting in plasmids pCS-sgegf*p*-cds1, and pCS-sg-m7 to m16. The inverse PCR was applied to introduce random mismatches in sgRNA variants targeting gene of interests. To create mismatch sgRNA libraries targeting eGFP, the forward primers sgegf*p*-rbs-F (*Apa*LI) or sgegf*p*-cds2-F (*Apa*LI) was individually paired with the reverse primer sgRNA-R (*Apa*LI) (**Table S3.1**) to amplify the whole template plasmid pCS-sgRNA, resulting in linear plasmids containing the sgRNA variants targeting the RBS region or the coding sequence (+378 bp) of the *eGFP*, respectively. The linear plasmids were then digested by *Apa*LI and *Dpn*I, and subjected to ligation. The ligation products were co-transferred into *E. coli*::dCas9 with the reporter plasmid pZE-eGFP. When plated on LB plates, transformants were randomly picked for the eGFP fluorescence assay by cultivating in LB tubes and plasmids were extracted for gene sequencing analysis. To construct mismatch sgRNA variants targeting genes in the *E. coli* genome for improved production of *p*-coumaric acid or butyrate, the inverse PCR templates were pCS-sgRNA and pZE-sgRNA, respectively. The selected targets and corresponding primers were listed in **Table S3.1**. The spacer of the selected sgRNA variants that enabled high-producing phenotypes were later sequenced and included in **Table S3.2**.

For constructing the *p*-coumaric acid responsive biosensor systems, the eGFP expression cassettes controlled by hybrid promoters (P1, P2, P7, and P9) were amplified from plasmids pZE-P1-eGFP, pZE-P2-eGFP, pZE-P7-eGFP, and pZE-P9-eGFP, respectively (Jiang et al., 2021). These cassettes were then inserted into the plasmid pZE-Plpp0.2-PadR (Li et al., 2022) using *Aat*II and *Xho*I, resulting in plasmids pZE-Plpp0.2-PadR-P1-eGFP, pZE-Plpp0.2-PadR-P2-eGFP, pZE-Plpp0.2-PadR-P7-eGFP, and pZE-Plpp0.2-PadR-P9-eGFP. The P_{L}*lacO1*-controlled RgTAL expression cassette was amplified from the plasmid pZE-RgTAL (Huang et al., 2013) and was constructed to the plasmid pZE-Plpp0.2-PadR-P9-eGFP using *Bam*HI and *Avr*II, yielding the plasmid pZE-CA. The biosynthetic pathway of butyrate consisting of *thl*, *hbd*, *crt*, *ter*, and *atoDA*, were sequentially constructed to the plasmid pCS27 to form plasmid pCS-THCTA using *Acc*65I, *Nde*I, *Sal*I, *Eco*RI, *Bam*HI, and *Mlu*I. The Plpp1-HpdR-P_{hpdH}-eGFP (HPE) biosensor cassette was amplified from the plasmid pCS-V2 (Wang et al., 2021b) and constructed to the plasmid pCS-THCTA using *Xho*I and *Xba*I, yielding the pCS-HPE-THCTA.

3.5.3 Biosensor dynamic range test and fluorescence assay

To test the dynamic range or substrate scope of engineered biosensor systems, the plasmids harboring the biosensor systems were transferred into *E. coli* BW25113(F′). Three independent colonies of each testing group were randomly picked and inoculated into 3.5 ml fresh LB medium (with appropriate antibiotics) as the seeds. After overnight (usually around 10-12 h) cultivation in a 37℃ rotatory shaker at a speed of 270 rpm, the seeds were transferred into fresh 3.5 ml LB medium with a ratio of 5% (175 μl), and the new cultures were grown in shaker under 37℃ with a speed of 270 rpm. After 1 h of cultivation, gradient concentrations of corresponding substrates were fed into the cultures. After 12 h of cultivation post induction, the cultures were sampled and

diluted by 10-fold before subjected to measurement of cell density (OD_{600}) and green fluorescence intensities using the 96-well plate reader (Synergy HT, BioTek). The green fluorescence intensities were determined using an excitation filter of 485/20 nm and an emission filter of 528/20 nm.

3.5.4 Screening high-production phenotypes in test tubes

For screening of *p*-coumaric acid high-production phenotypes, the constructed mismatch sgRNA libraries on pCS27 were co-transferred into *E. coli*::dCas9 Δ*pheA* Δ*tyrR* with the plasmid pZE-CA. For screening of butyrate high-production phenotypes, the constructed mismatched sgRNA libraries on pZE12-luc were co-transferred into *E. coli*::dCas9 with the plasmid pCS-HPE-THCTA. The transformants with high green fluorescence intensities on LB plates with appropriate antibiotics and 0.5 mM IPTG were selected and inoculated in 3.5 ml M9Y medium containing 20 g/l glycerol (for *p*-coumaric acid) or 20 g/l glucose (for butyrate). The IPTG was added in the medium to a final concentration of 0.5 mM at the beginning to induce the expression of CRISPRi system and biosynthesis pathways. Samples were taken after 24 h of induction and were subjected to measurement of cell densities, green fluorescence intensities, and product accumulations. The cell density and green fluorescence intensity were quantitively analyzed using the 96-well plate reader (Synergy HT, BioTek) with the same protocol described earlier. The production of *p*-coumaric acid or butyrate was analyzed by high-performance liquid chromatography (HPLC). The selected high-production strains were streaked in the LB plate (with appropriate antibiotics) and the plasmids mixture were extracted for DNA sequencing of the variant sgRNA spacers.

3.5.5 Validation of selected sgRNA variants via shake flasks fermentation

To validate the production performance of selected colonies from test tubes, three colonies of the strains streaked on the petri dish were inoculated in 3.5 ml LB medium and grown at 37 ℃ for 8-10 h. Then, 1 ml of each seed culture was transferred to 20 ml fresh M9Y medium (containing 20 g/l glycerol or glucose) in 125-ml shake flasks (Chemglass Life Sciences LLC) and grown at 30 ℃ for 48 h. The shake flask experiments were performed in a 30 ℃ rotary shaker with a speed of 270 rpm. The glycerol was used as carbon source for *p*-coumaric acid production while the glucose was used for butyrate production. IPTG was added in the medium with a final concentration of 0.5 mM. Samples were collected after 48 h post induction and were subjected for measurement of cell density (OD_{600}) and product accumulation. The cell density (OD_{600}) was measured using the spectrophotometer (VWR), and the products *p*-coumaric acid or butyrate were analyzed by HPLC.

3.5.6 Product analysis

The concentrations of *p*-coumaric acid were quantitatively analyzed by Agilent HPLC 1260 Infinity II (1260 Infinity II Diode Array Detector WR) with a reverse-phase ZORBAX SB-C18 column. A methanol-water (containing 0.1% trifluoroacetic acid) gradient system with a flow rate of 1 ml/min was used. The analyzing method, adopted from our previous study (Li et al., 2022), was set as follows: 5% methanol from 0 to 2 min, 5% to 80% methanol from 2 to 10 min, 80% methanol from 10 to 14 min, 80% to 5% methanol from 14 to 18 min, and 5% methanol from 18 to 20 min. Standard curve was drawn using gradient concentrations of *p*-coumaric acid solution prepared from *p*-coumaric acid standard (MP Biomedical LLC). The retention time of *p*-coumaric acid is at 8.89 min.

The butyrate was analyzed by Dionex Ultimate 3000 HPLC equipped with a Coregel-64H column (Transgenomic, Omaha, NE). Collected samples (1 ml) were centrifuged at 12,000 rpm for 15 minutes, and the supernatants were filtered through 0.22 μm film before applied to HPLC analysis. The HPLC method was modified from our previous research (Wang, Jian et al., 2018). Briefly, 4 mN sulfuric acid was used as the mobile phase with a flow rate of 0.4 ml/min. The column temperature was set at 45 ℃. The retention time of butyrate is at 38.897 min.

3.6 Tables and Figures

Table 3.1. Strains and plasmids

Strains	**Description**	**Reference**
E. coli XL1-Blue	*recA1 endA1gyrA*96*thi*-1*hsdR*17*supE*44*relA*1*lac*	Stratagene
E. coli BW25113 (F′)	*rrnBT14 ΔlacZWJ16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 F'[traD36 proAB lacIqZΔM15 Tn10(Tetr)]*	(Atsumi et al., 2008)
E. coli::dCas9	*E. coli* BW25113(F′) with the integration of the P_L*lacO1*-dCas9 cassette into the genome at *dkgB* locus	This study
E. coli::dCas9 Δ*pheA* Δ*tyrR*	*E. coli*::dCas9 with the deletion of *pheA* and *tyrR*	This study
Plasmids	**Description**	**Reference**
pZE12-luc	P_L*lacO1*; *luc*; *ColE1 ori*; Amp^R	(Lutz and Bujard, 1997)
pCS27	P_L*lacO1*; *p15A ori*; Kan^R	(Shen and Liao, 2008)
pZE-eGFP	pZE12luc harboring the P_L*lacO1*-controlled eGFP expression cassette	(Wang et al., 2021b)
pCS-sgRNA	pCS27 harboring sgRNA scaffold targeting eGFP	(Wang, J. et al., 2017b)
pZE-sgRNA	pZE12-luc harboring the sgRNA scaffold amplified from pCS27-sgRNA and inserted in between *Xho*I and *Avr*II	This study

pHA-MCS	pZE12luc with the luc gene replaced by a newly designed multi-cloning site	This study
pCS-sgegf*p*-cds1	pCS27 harboring the *$P_L$lacO1*-controlled transcriptional cassette of a fully matched sgRNA targeting the +39 bp of *eGFP*	This study
pCS-sgegf*p*-rbs	pCS-sgegfp targeting the RBS region of the *eGFP*	This study
pCS-sgegf*p*-cds2	pCS-sgegfp targeting the coding sequence (starting from + 378 bp)	This study
pZE-Plpp0.2-PadR-P1-eGFP	pZE-Plpp0.2-PadR harboring the hybrid promoter P1-controlled *eGFP* expression cassette	This study
pZE-Plpp0.2-PadR-P2-eGFP	pZE-Plpp0.2-PadR harboring the hybrid promoter P2-controlled *eGFP* expression cassette	This study
pZE-Plpp0.2-PadR-P7-eGFP	pZE-Plpp0.2-PadR harboring the hybrid promoter P7-controlled *eGFP* expression cassette	This study
pZE-Plpp0.2-PadR-P9-eGFP	pZE-Plpp0.2-PadR harboring the hybrid promoter P9-controlled *eGFP* expression cassette	This study
pZE-CA	pZE-Plpp0.2-PadR-P9-eGFP harboring the codon-optimized RgTAL from *Rhodotorula glutinis* under control of the *$P_L$lacO1* promoter	This study
pCS-V2	pCS27-Plpp1-HpdR-P_{hpdH}-eGFP with deletion of a palindromic sequence (-209 to -118 bp) on the P_{hpdH} promoter and HpdR under control of Plpp1.0 promoter	(Wang et al., 2021b)

pCS-THCTA	pCS27 harbors the butyrate synthetic pathway which includes the *thl*, *hbd*, *crt*, *ter*, and *atoDA*.	This study
pCS-HPE-THCTA	pCS-THCTA harboring Plpp1-HpdR-P_{hpdH}-eGFP (HPE) biosensor cassette from pCS-V2	This study

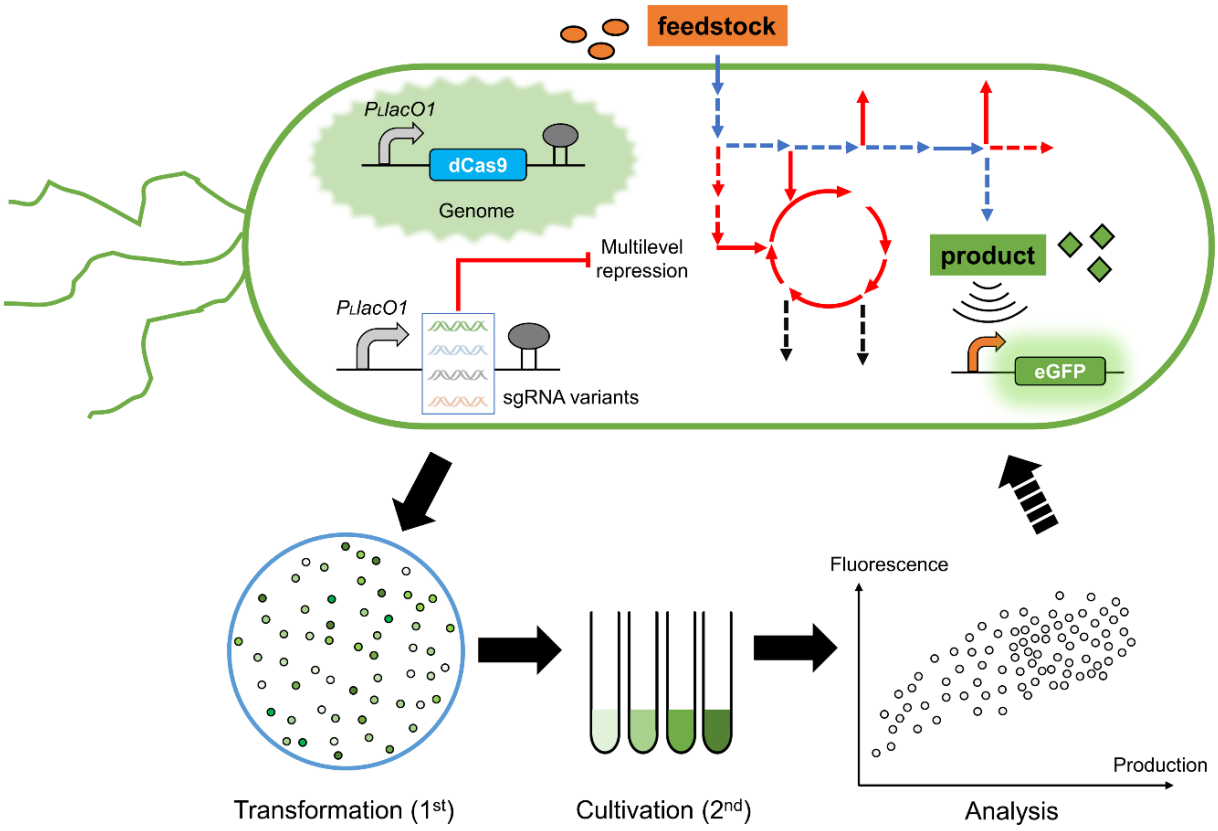


Figure 3.1. Schematic diagram of the BATCH screening workflow. The BATCH screening involves sgRNA variants mediated multilevel CRISPRi of gene of interests within metabolic network to redirect carbon flux to target biosynthetic pathways. The improved production phenotypes will be reflected by the target chemical inducible reporter GFP fluorescence. The BATCH screening consists of the first-round plate-based selection and the second-round test tube-based selection followed by final shake flask-based validation. The blue lines indicate biosynthetic pathways and the red lines indicate relevant or potential competing pathways.

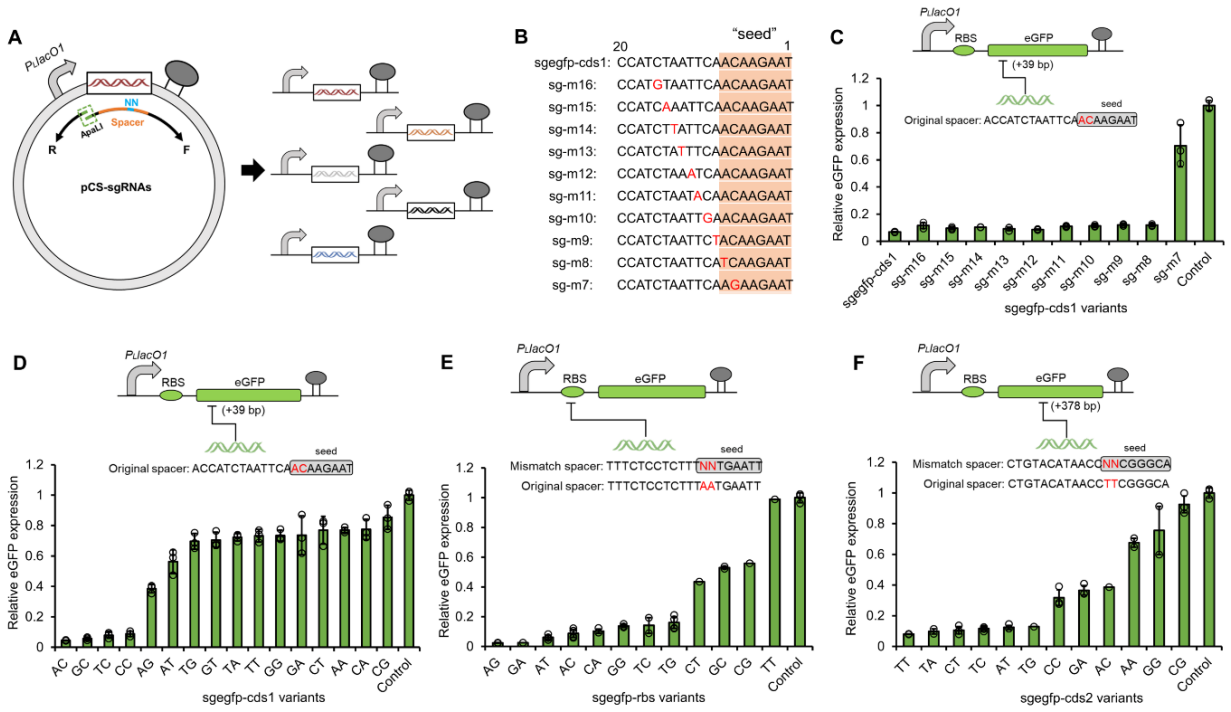


Figure 3.2. Establishing titratable CRISPRi with mismatch sgRNA libraries. (**A**) The design of inverse PCR for one-pot construction of mismatch sgRNAs. (**B**) The spacer sequences (5'-3') of single mismatch sgegf*p*-cds1 variants targeting the coding sequence (+39 bp) of eGFP. (**C**) The CRISPRi efficiencies of sgRNAs with single mismatch at different positions in the sgegf*p*-cds1 spacer (**c**), with rationally designed double mismatches in the 7-8th bp seed region of sgegf*p*-cds1 spacer (**D**), with randomly mutated double mismatches in the 7-8th bp seed region of sgegf*p*-rbs spacer (targeting the RBS region eGFP promoter) (**E**), and with randomly mutated double mismatches in the 7-8th bp seed region of sgegf*p*-cds2 spacer (targeting the coding sequence (+378 bp) of eGFP) (**F**). In **E** and **F**, 32 single colonies were randomly picked from each plate and were subjected to both eGFP fluorescence assay and spacer determination via plasmid sequencing. All data in **C** and **D** represent the mean of three biologically independent samples and error bars show standard deviation.

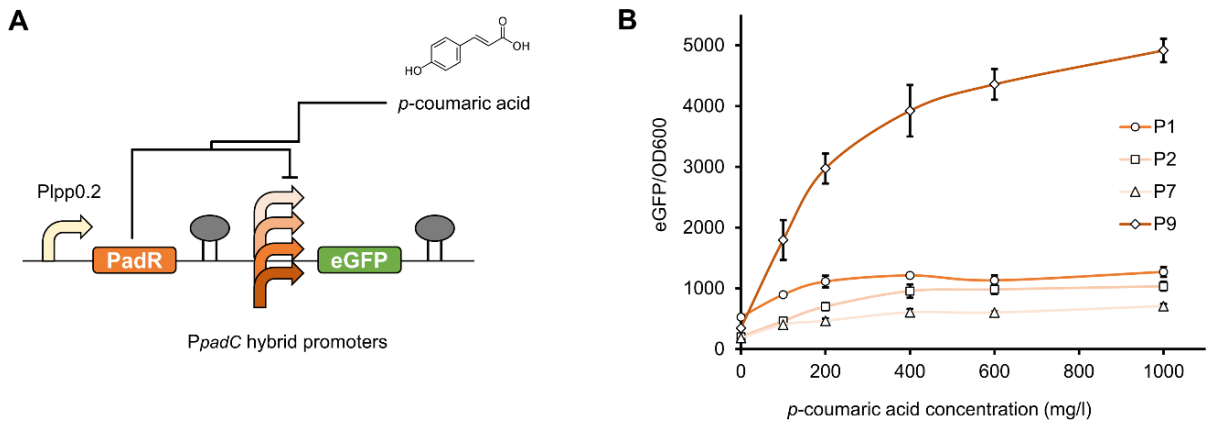


Figure 3.3. Optimizing the *p*-coumaric acid biosensor PadR with broad dynamic behavior. (**A**) The plasmid configuration for the PadR-based biosensor system. PadR was constitutively expressed under the control of the Plpp0.2 promoter and the reporter gene eGFP was controlled by the hybrid P_{padC} promoters (P1, P2, P7, and P9). (**B**) The dynamic ranges of the PadR-P_{padC} biosensor system harboring different hybrid promoters against different concentrations of *p*-coumaric acid (0, 100, 200, 400, 600, and 1000 mg/l). All data represent the mean of three biologically independent samples and error bars show standard deviation.

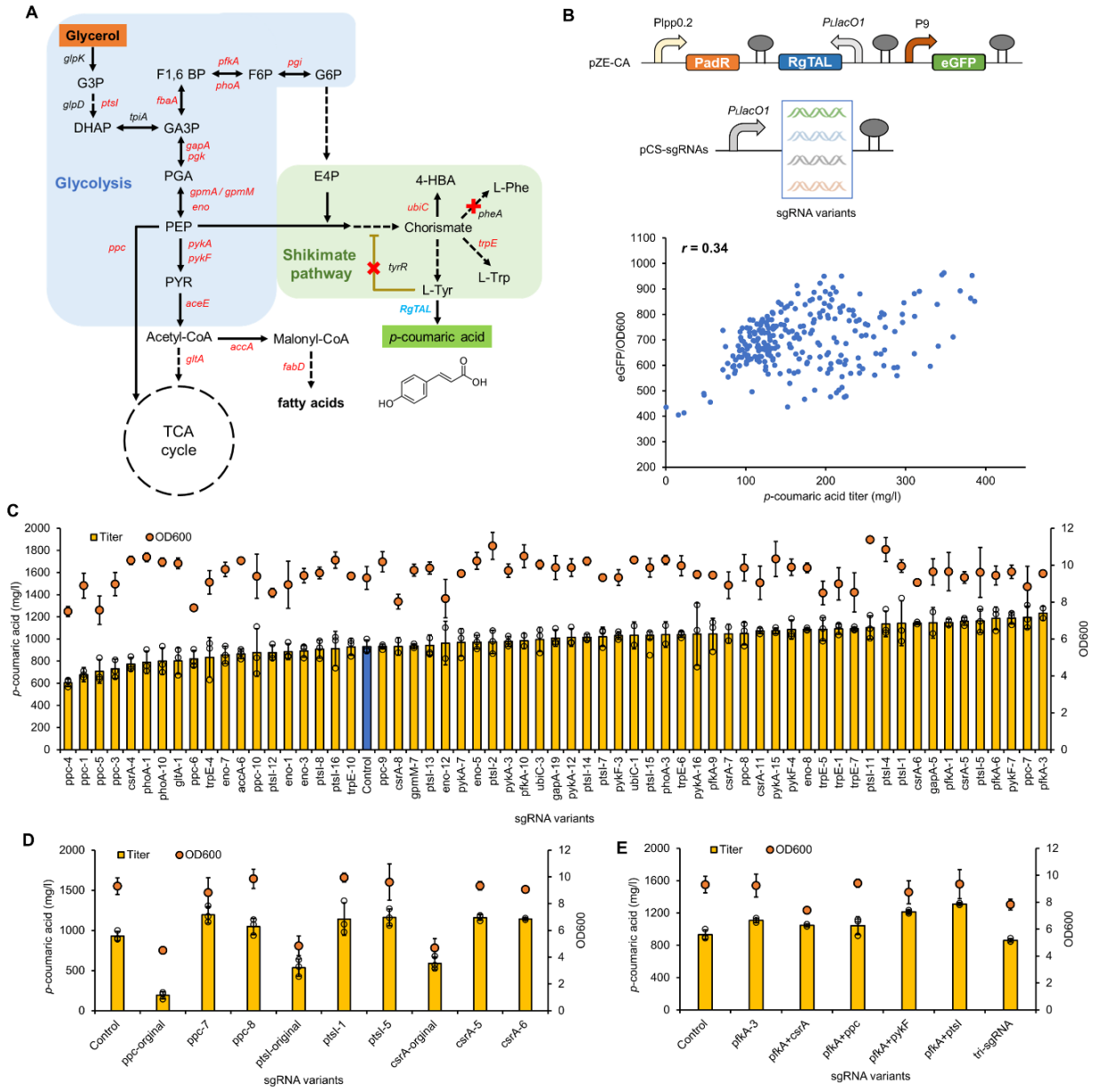


Figure 3.4. BATCH screening of *p*-coumaric acid over-production phenotypes. (**A**) The biosynthetic pathway of *p*-coumaric acid from glycerol. The genes marked in red are CRISPRi targets and the red crosses are genes to be knocked out. (**B**) The correlation between the *p*-coumaric acid titers (in test tubes) and the green fluorescence intensities in *E. coli*::dCas9 Δ*pheA* Δ*tyrR* with pZE-CA and pCS-sgRNA mismatch variants. (**C**) The titers of *p*-coumaric acid and cell densities

(OD_{600}) of *p*-coumaric acid producers with selected sgRNA variants in shake flasks. The blue bar represents the control strain without any sgRNA. (**D**) The comparison of *p*-coumaric acid production between authentic (fully matched) sgRNAs targeting *ppc*, *ptsI* and *csrA*, and their counterpart mismatch sgRNA variants in shake flasks. (**E**) The *p*-coumaric acid production test with combination of selected sgRNA variants in shake flasks. All data represent the mean of three biologically independent samples and error bars show standard deviation.

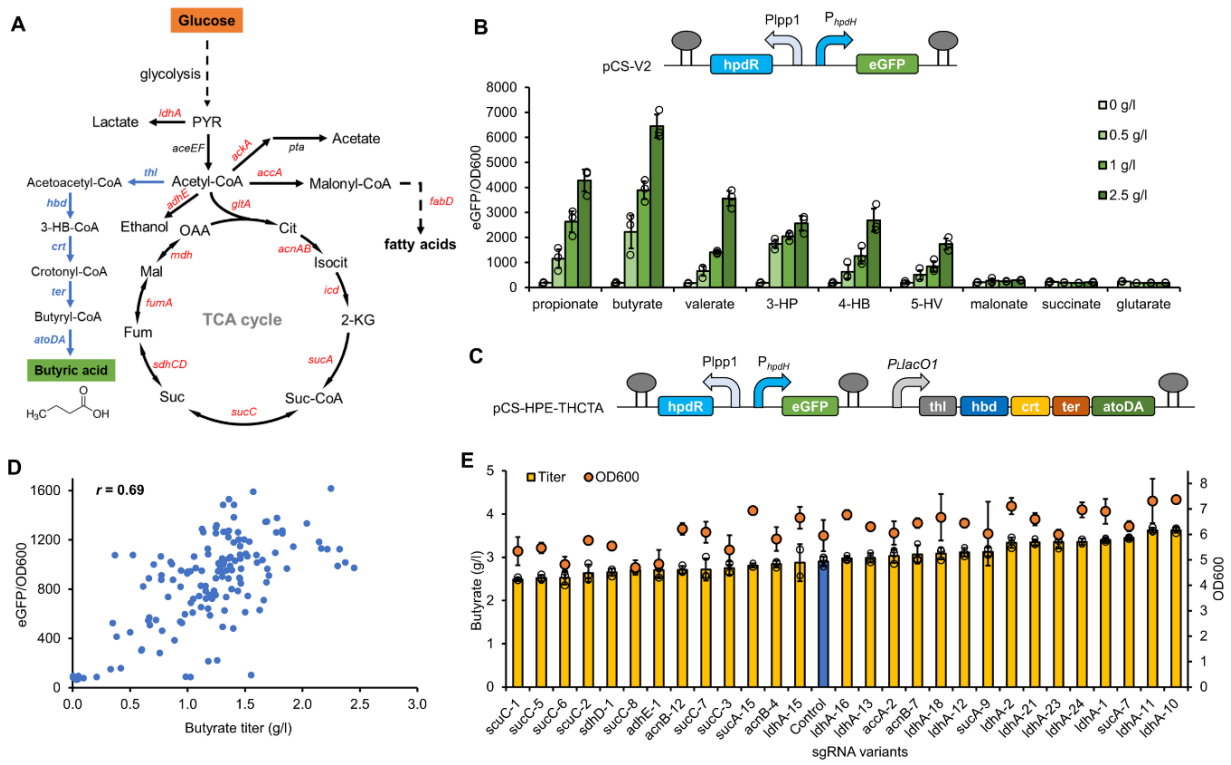


Figure 3.5. BATCH screening for butyrate over-production phenotypes. (**A**) The biosynthetic pathway for butyrate from glucose. The genes shown in blue are butyrate pathway genes and the genes shown in red are CRISPRi targets. (**B**) The ligand profiling of the engineered HpdR-P$_{hpdH}$ biosensor against C3-C5 fatty acids (propionate, butyrate, and valerate), hydroxy acids (3-hydroxypropionate (3-HP), 4-hydroxybutyrate (4-HB), and 5-hydroxyvalerate (5-HV)), and diacids (malonate, succinate, and glutarate) with gradient concentrations of 0-2.5 g/l. (**C**) The plasmid scheme of pCS-HPE-HCTA for BATCH screening of butyrate high-production phenotypes. (**D**) The correlation between the butyrate titers (in test tubes) and the green fluorescence intensities in *E. coli*::dCas9 with pCS-HPE-THCTA and pZE-sgRNA mismatch variants. (**E**) The titers of butyrate and cell densities (OD_{600}) of butyrate producers with selected sgRNA variants in shake flasks. The blue bar represents the control strain without any sgRNA. All

data represent the mean of three biologically independent samples and error bars show standard deviation.

3.7 Supplementary information

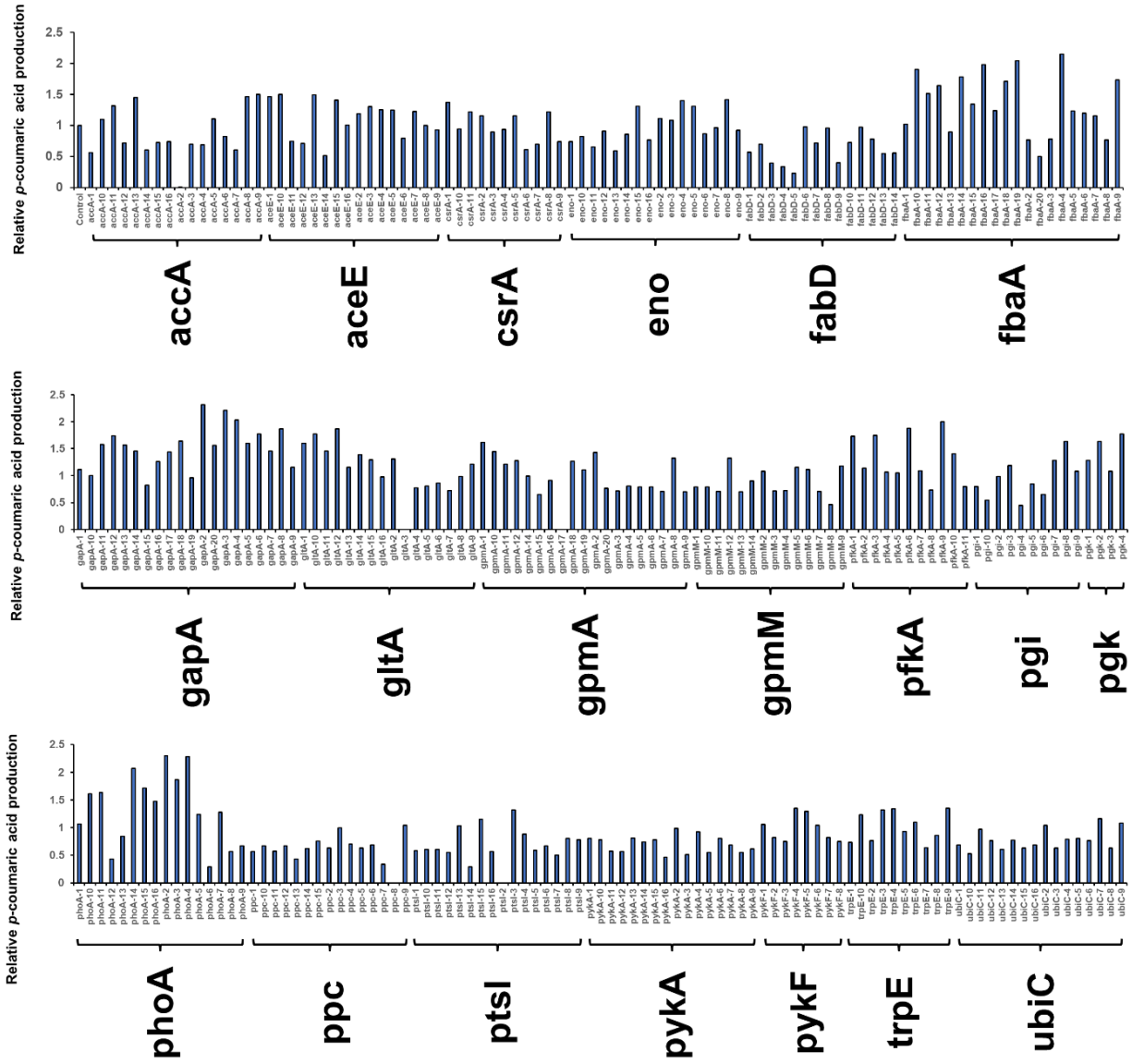


Figure S3.1. Relative *p*-coumaric acid titer of selected sgRNA variants in test tubes. The medium used for screening was M9Y with 20 g/L glycerol.

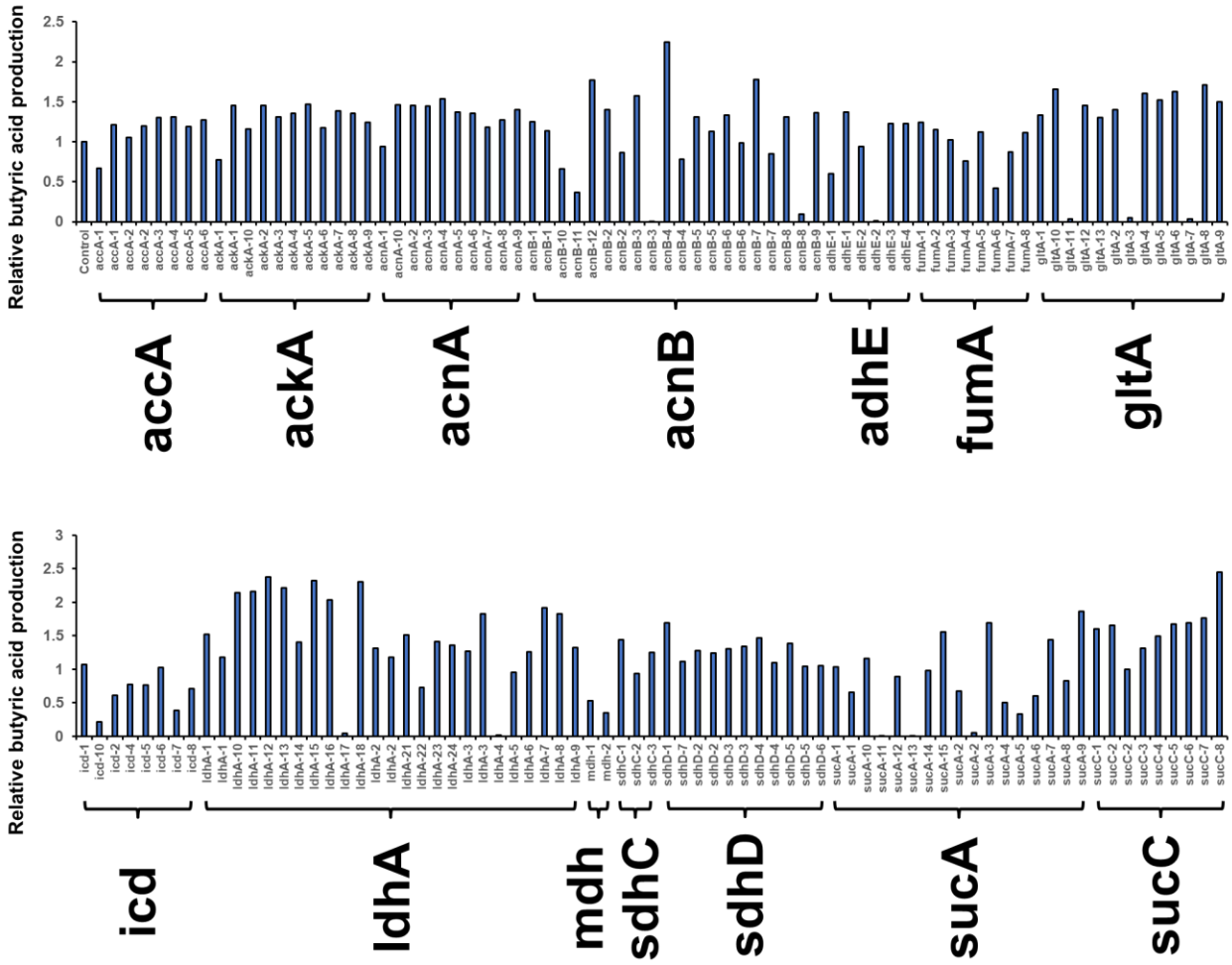


Figure S3.2. Relative butyric acid titer of selected sgRNA variants in test tubes. The medium used for screening was M9Y with 20 g/L glucose.

Table S3.1. Primers for constructing mismatched sgRNA arrays

Name	Sequence (5'-3')
sgegf*p*-rbs-F (ApaLI)	ggaaaGTGCACATTTCTCCTCTTTNNTGAATTGTTTTAGAG CTAGAAATAGCAAGTT
sgegf*p*-cds2-F (ApaLI)	ggaaaGTGCACACTGTACATAACCNNCGGGCAGTTTTAGA GCTAGAAATAGCAAGTT
sgpgi-F(ApaLI)	ggaaaGTGCACATGCCAGGCAGCGNNCTGCGTGTTTTAGA GCTAGAAATAGCAAGTT
sgpfkA-F(ApaLI)	ggaaaGTGCACACGAATTGCGGCGNNCATGCCGTTTTAGA GCTAGAAATAGCAAGTT
sgfbaA-F(ApaLI)	ggaaaGTGCACATCATCACCAGTGNNTACGCCGTTTTAGA GCTAGAAATAGCAAGTT
sggapA-F(ApaLI)	ggaaaGTGCACAGGAAAACAATGCNNCCGATAGTTTTAG AGCTAGAAATAGCAAGTT
sgpgk-F(ApaLI)	ggaaaGTGCACATTTCCCAGCAAGNNCCAGATGTTTTAGA GCTAGAAATAGCAAGTT
sggpmA-F(ApaLI)	ggaaaGTGCACACGTCGTACCAACNNGTGAAAGTTTTAGA GCTAGAAATAGCAAGTT
sggpmM-F(ApaLI)	ggaaaGTGCACATCCAGAATCACCNNTACCATGTTTTAGA GCTAGAAATAGCAAGTT
sgeno-F(ApaLI)	ggaaaGTGCACAACCGATGATTTTNNCGATTTGTTTTAGA GCTAGAAATAGCAAGTT

sgpykA-F(ApaLI)	ggaaaGTGCACATTTTGTTCTGCGNNGCCTTCGTTTTAGAG CTAGAAATAGCAAGTT
sgpykF-F(ApaLI)	ggaaaGTGCACATCCGATGGTGCANNCAATTTGTTTTAGA GCTAGAAATAGCAAGTT
sgaceE-F(ApaLI)	ggaaaGTGCACACGATCGGATCCANNTCATTTGTTTTAGA GCTAGAAATAGCAAGTT
sggltA-F(ApaLI)	ggaaaGTGCACAAACAGCTGTATCNNCGTTGAGTTTTAGA GCTAGAAATAGCAAGTT
sgppc-F(ApaLI)	ggaaaGTGCACAACTGACATTACTNNGCAATGGTTTTAGA GCTAGAAATAGCAAGTT
sgaccA-F(ApaLI)	ggaaaGTGCACACAATCGGCTGTTNNAAATCAGTTTTAGA GCTAGAAATAGCAAGTT
sgfabD-F(ApaLI)	ggaaaGTGCACACGGTTTGAGAACNNTGTCCAGTTTTAGA GCTAGAAATAGCAAGTT
sgcsrA-F(ApaLI)	ggaaaGTGCACAGACCTCATCCCCNNTCATGAGTTTTAGA GCTAGAAATAGCAAGTT
sgmurA-F(ApaLI)	ggaaaGTGCACAACTTCGCCCTGGNNCTTCGTGTTTTAGA GCTAGAAATAGCAAGTT
sgphoA-F(ApaLI)	ggaaaGTGCACAGTCACAGGGGTANNCAGTAAGTTTTAGA GCTAGAAATAGCAAGTT
sgptsI-F(ApaLI)	ggaaaGTGCACATTTACCGAAAGCNNTACCCGGTTTTAGA GCTAGAAATAGCAAGTT

sgubiC-F(ApaLI)	ggaaaGTGCACACACGCAGTTGCGNNAACGCGGTTTTAGA GCTAGAAATAGCAAGTT
sgtrpE-F(ApaLI)	ggaaaGTGCACACAGGTTAGCAGTNNGAGAGTGTTTTAGA GCTAGAAATAGCAAGTT
sgldhA-F(ApaLI)	ggaaaGTGCACAGTACTGTTTTGTNNTATAAAGTTTTAGA GCTAGAAATAGCAAGTT
sgadhE-F(ApaLI)	ggaaaGTGCACAGAAACTGGCATANNCACGCTGTTTTAGA GCTAGAAATAGCAAGTT
sgackA-F(ApaLI)	ggaaaGTGCACAACCATTTACTGCNNCGATGAGTTTTAGA GCTAGAAATAGCAAGTT
sgacnA-F(ApaLI)	ggaaaGTGCACAGTCCTTACTGGCNNCTCGTAGTTTTAGA GCTAGAAATAGCAAGTT
sgacnB-F(ApaLI)	ggaaaGTGCACACACGCTCAGCTANNTGCTTAGTTTTAGA GCTAGAAATAGCAAGTT
sgicd-F(ApaLI)	ggaaaGTGCACAGTGATCTTCTTGNNTTGTGCGTTTTAGA GCTAGAAATAGCAAGTT
sgsucA-F(ApaLI)	ggaaaGTGCACAGAGGTAAGAAGANNCCAACCGTTTTAG AGCTAGAAATAGCAAGTT
sgsucC-F(ApaLI)	ggaaaGTGCACACGGTGCTGGTAANNCATAGCGTTTTAGA GCTAGAAATAGCAAGTT
sgsdhC-F(ApaLI)	ggaaaGTGCACAGTCTGTAGGTCCNNATTAACGTTTTAGA GCTAGAAATAGCAAGTT

sgsdhD-F(ApaLI)	ggaaaGTGCACAGCCATTGCGTCCNNATGCGGGTTTTAGAGCTAGAAATAGCAAGTT
sgfumA-F(ApaLI)	ggaaaGTGCACAAAAGGAGCCTGANNATGAAAGTTTTAGAGCTAGAAATAGCAAGTT
sgmdh-F(ApaLI)	ggaaaGTGCACACAATACCGCCAGNNGCGCCGGTTTTAGAGCTAGAAATAGCAAGTT
sgRNA-R(ApaLI)	ggaaagtgcAcagtatcttgttatccgctcacaat

Table S3.2. DNA sequences of the screened sgRNA variants

sgRNA variants	Spacer DNA sequence (fully matched base pair)
	For *p*-coumaric acid
csrA-5	GACCTCATCCCCTCTCATGA (AA)
csrA-6	GACCTCATCCCCTCTCATGA (AA)
csrA-7	GACCTCATCCCCATTCATGA (AA)
csrA-8	GACCTCATCCCCTGTCATGA (AA)
csrA-11	GACCTCATCCCCCGTCATGA (AA)
eno-1	ACCGATGATTTTTCCGATTT (TA)
eno-3	ACCGATGATTTTGGCGATTT (TA)
eno-5	ACCGATGATTTTTCCGATTT (TA)
eno-7	ACCGATGATTTTTCCGATTT (TA)
eno-8	ACCGATGATTTTTTCGATTT (TA)
eno-12	ACCGATGATTTTGGCGATTT (TA)
gapA-5	GGAAAACAATGCAGCCGATA (GA)
gapA-19	GGAAAACAATGCCGCCGATA (GA)
gltA-1	AACAGCTGTATCAACGTTGA (CC)
gpmM-7	TCCAGAATCACCTATACCAT (AG)
pfkA-1	CGAATTGCGGCGTTCATGCC (TT)
pfkA-3	CGAATTGCGGCGTTCATGCC (TT)
pfkA-6	CGAATTGCGGCGTGCATGCC (TT)
pfkA-9	CGAATTGCGGCGTTCATGCC (TT)
pfkA-10	CGAATTGCGGCGTTCATGCC (TT)

phoA-3	GTCACAGGGGTAAACAGTAA (AA)
ppc-7	ACTGACATTACTTGGCAATG (AC)
ppc-8	ACTGACATTACTTGGCAATG (AC)
ppc-9	ACTGACATTACTATGCAATG (AC)
ptsI-1	TTTACCGAAAGCATTACCCG (GA)
ptsI-2	TTTACCGAAAGCTTTACCCG (GA)
ptsI-4	TTTACCGAAAGCCTTACCCG (GA)
ptsI-5	TTTACCGAAAGCCGTACCCG (GA)
ptsI-7	TTTACCGAAAGCTTTACCCG (GA)
ptsI-11	TTTACCGAAAGCTGTACCCG (GA)
ptsI-13	TTTACCGAAAGCTTTACCCG (GA)
ptsI-14	TTTACCGAAAGCAGTACCCG (GA)
ptsI-15	TTTACCGAAAGCTTTACCCG (GA)
pykA-3	TTTTGTTCTGCGTGGCCTTC (AA)
pykA-7	TTTTGTTCTGCGGAGCCTTC (AA)
pykA-12	TTTTGTTCTGCGGAGCCTTC (AA)
pykA-15	TTTTGTTCTGCGGGGCCTTC (AA)
pykA-16	TTTTGTTCTGCGGCGCCTTC (AA)
pykF-3	TCCGATGGTGCAGACAATTT (AA)
pykF-4	TCCGATGGTGCAGGCAATTT (AA)
pykF-7	TCCGATGGTGCAATCAATTT (AA)
trpE-1	CAGGTTAGCAGTTCGAGAGT (TC)
trpE-5	CAGGTTAGCAGTGCGAGAGT (TC)

trpE-6	CAGGTTAGCAGTCGGAGAGT (TC)
trpE-7	CAGGTTAGCAGTCGGAGAGT (TC)
ubiC-1	CACGCAGTTGCGGCAACGCG (TT)
ubiC-3	CACGCAGTTGCGGCAACGCG (TT)
	For butyric acid
accA-2	CAATCGGCTGTTATAAATCA (CA)
acnB-4	CACGCTCAGCTAAATGCTTA (CG)
acnB-7	CACGCTCAGCTAGATGCTTA (CG)
acnB-12	CACGCTCAGCTATTTGCTTA (CG)
ldhA-1	GTACTGTTTTGTCGTATAAA (GC)
ldhA-2	GTACTGTTTTGTTATATAAA (GC)
ldhA-10	GTACTGTTTTGTCGTATAAA (GC)
ldhA-11	GTACTGTTTTGTCCTATAAA (GC)
ldhA-12	GTACTGTTTTGTGCTATAAA (GC)
ldhA-13	GTACTGTTTTGTAATATAAA (GC)
ldhA-15	GTACTGTTTTGTAATATAAA (GC)
ldhA-16	GTACTGTTTTGTAATATAAA (GC)
ldhA-18	GTACTGTTTTGTGTTATAAA (GC)
ldhA-21	GTACTGTTTTGTATTATAAA (GC)
ldhA-23	GTACTGTTTTGTTTTATAAA (GC)
ldhA-24	GTACTGTTTTGTTGTATAAA (GC)
sdhD-1	GCCATTGCGTCCGCATGCGG (TA)
sucA-7	GAGGTAAGAAGAATCCAACC (GT)

sucA-9	GAGGTAAGAAGATCCCAACC (GT)
sucC-1	CGGTGCTGGTAATGCATAGC (GC)
sucC-2	CGGTGCTGGTAACGCATAGC (GC)
sucC-3	CGGTGCTGGTAAATCATAGC (GC)
sucC-5	CGGTGCTGGTAAAGCATAGC (GC)
sucC-6	CGGTGCTGGTAAGGCATAGC (GC)
sucC-7	CGGTGCTGGTAAAACATAGC (GC)
sucC-8	CGGTGCTGGTAAGGCATAGC (GC)

CHAPTER 4

EXPANDING THE SUBSTRATE SPECTRUM OF A PHENOLIC ACID-RESPONSIVE REGULATOR PADR FROM BACILLUS SUBTILIS

---

1 Li, C., Zou, Y., Jiang, T., Yan, Y., 2022. Expanding the substrate spectrum of a phenolic acid-responsive regulator PadR from *Bacillus subtilis*. To be s*ubmitted to ACS Synthetic Biology*.

4.1 Abstract

Transcriptional factors are commonly involved in genetically encoded biosensors and have been widely applied to metabolic engineering and synthetic biology. Engineering existing transcriptional factors to respond new chemicals can broaden the substrate repertoire of current transcriptional factors and is necessary for expanding their applicability in metabolic engineering and synthetic biology. PadR is a transcriptional repressor that is responsive towards phenolic acids which are important precursors for value-added flavonoids and coumarins. In this study, we modified the substrate specificity of a phenolic acid-responsive regulator PadR from *Bacillus subtilis 168* through both rational and semi-rational approaches. The PadR was engineered to respond to a series of aromatic compounds including caffeic acid, cinnamic acid, salicylic acid, anthranilic acid, vanillic acid, 4-aminobenzoic acid, and 4-hydroxycoumarin. The PadR variants developed in this study expanded its substrate repertoire, and would greatly benefit biosensor-based applications in future studies.

4.2 Introduction

Transcriptional factors (TFs) responding to small molecules or environmental signals have been widely applied in biosensor-based real-time environmental monitoring, gene expression control, genetic circuits construction, cellular regulatory network establishment, dynamic pathway regulation, and high-throughput screening (Ding et al., 2021; Farmer and Liao, 2000; Inda and Lu, 2020; Li et al., 2021; Mitchler et al., 2021; Nielsen and Keasling, 2016; Rogers and Church, 2016; Zhang et al., 2012; Zhang et al., 2015). While many natural and engineered transcriptional factors have been identified, characterized, and developed (Hossain et al., 2020; Jiang et al., 2021; Koch

et al., 2019; Li et al., 2020; Zhang et al., 2021; Zou et al., 2021), the substrate spectrum of existing TFs is still narrow, as many small molecules still cannot be recognized by TFs (Koch et al., 2019).

To address this challenge, one solution is the continuously mining and characterization of new TFs from nature. With the development of new mining algorithms, increased data generation and analysis speed, and establishment of user-friendly databases, this strategy becomes increasingly important and efficient in uncovering new TF-based biosensors with novel substrate scopes (Novichkov et al., 2013; Shlomi et al., 2007; Sun et al., 2020). However, the natural transcriptional factors and their effectors are often closely related to the hosts metabolism. Thus, it poses some difficulties to further expand the substrate spectra of TFs. As a promising alternative approach, engineering and evolving existing TFs to broaden their substrate scopes is also explored and has been demonstrated efficient in enabling new responsiveness of current biosensors (Chen, W. et al., 2015; Li et al., 2020; Lin et al., 2017; Tang et al., 2008; Tang and Cirino, 2011; Tang et al., 2013; Taylor et al., 2016). Two notable examples are the commonly used lactose-responsive repressor LacI and arabinose-responsive activator AraC. The engineered LacI variants were able to recognize gentiobiose, fucose, lactitol, and sucralose after the alteration of substrate binding pocket (Taylor et al., 2016). The engineering of AraC enabled an even broader substrate scope, making it responsive towards mevalonate, ectoine, and triacetic acid lactone. Notably, these compounds did not share a very similar structure to the native effector (L-arabinose) of AraC (Chen, W. et al., 2015; Tang and Cirino, 2011; Tang et al., 2013).

PadR is a phenolic acid responsive transcriptional repressor discovered from *Bacillus subtilis*(Jiang et al., 2021; Nguyen Thi Kim et al., 2011; Park et al., 2017). PadR would inhibit the

promoter P_{padC} and repress the transcription of downstream *padC* gene. However, when the cells were exposed to an environment with high concentration of phenolic acid, the phenolic acid would bind with the PadR and result in conformational change of PadR, releasing the inhibition on *padC* transcription and activating the decarboxylation of phenolic acids(Jiang et al., 2021; Nguyen Thi Kim et al., 2011; Park et al., 2017). Due to the capability of sensing phenolic acids such as *p*-coumaric acid and ferulic acid, the important precursors for a series of valuable flavonoids and coumarins(Huang et al., 2013; Lin et al., 2013b; Yang et al., 2015), this regulator was extensively studied (Jiang et al., 2021; Siedler et al., 2017). In 2017, this expression level of PadR was optimized in yeast through RBS engineering, and the engineered sensor system was used to screen high producing strains of *p*-coumaric acid(Siedler et al., 2017). In our previous studies, the PadR was further optimized with increased sensitivity, broader dynamic ranges, and expanded operational ranges(Jiang et al., 2021). The engineered variants were then applied in establishing the dynamic pathway control at gene copy level through regulating the plasmid replication(Li et al., 2022).

In this chapter, we investigated the substrate binding of the phenolic acid-responsive transcriptional repressor PadR from the *Bacillus subtilis 168* that was engineered and optimized in our previous study (Jiang et al., 2021). Based on the analysis of the simulated substrate binding pocket, we designed four rational mutations (H154Y, H154W, K127Y, and K127R), and one variant (K127Y) enabled increased responsiveness towards the *p*-coumaric acid, ferulic acid, and caffeic acid. Two variants, K127Y and H154W, also exhibited new responsiveness towards cinnamic acid. We also tested the dynamic performance of these variants against smaller aromatic compounds and found that some variants exhibited a good indication of responsiveness. Further

semi-rational site-saturated mutagenesis expanded the ligand scope of the PadR as the variants become responsive towards vanillic acid, 4-aminobenzoic acid, and 4-hydroxycoumarin. These novel PadR variants developed in this study would be beneficial for future metabolic engineering and synthetic biology.

4.3 Results

4.3.1 Analyzing substrate binding pocket of PadR

With the goal of expanding the substrate scope of PadR, direct investigation of its substrate binding is required. While the crystal structure of the PadR from *Bacillus subtilis 168* is not available, a close homologous protein (with a sequence similarity of 98.4%) from *Bacillus subtilis subsp. spizizenii str. W23* has been crystallized in a previous study (Park et al., 2017). Thus, the structure of the PadR from *Bacillus subtilis 168* was simulated with SWISS-Model (https://swissmodel.expasy.org/) using the crystal structure of PadR from *Bacillus subtilis subsp. spizizenii str. W23* (PDB ID: 5Y8T) as the template (Park et al., 2017). The simulated structure was used to analyze the substrate binding of PadR in Pymol (**Fig. 4.1a**). There are two residues, H154 and R164, that directly interacting with the native substrate *p*-coumaric acid of PadR. The residue H154 can form a hydrogen bond with the hydroxyl group of *p*-coumaric acid, while two hydrogen bonds can be formed between the R164 residue and the carboxylate group of *p*-coumaric acid (**Fig. 4.1a**). Aside from the H154 and R164, three additional residues, K127, L131, and S134 that located at the boundary of the substrate binding pocket, also caught our attention, because the alteration of these amino acids would reshape the pocket and possibly change the substrate preference of PadR.

4.3.2 Rational engineering to modify the PadR substrate specificity

Our first goal is to reduce the size of the substrate binding pocket and see if PadR can respond to cinnamic acid (without the hydroxy group at the *para* position compared with *p*-coumaric acid). Since the H154 directly interacted with the hydroxy group (**Fig. 4.1a**), we hypothesized that the substitution of histidine to larger amino acids might reduce the pocket size. As a result, the PadR would be able to interact with the benzene ring of cinnamic acid, eliminating the necessity of the hydroxy group for hydrogen bond formation. However, a caveat for this strategy is that the alteration of H154 might cause loss-of-function for the PadR, because both the *p*-coumaric acid and ferulic acid, the native substrates of the PadR, contain the hydroxy group at the *para*-position in the benzene ring. Therefore, it is reasonable to assume the H154 is important in substrate recognition.

To test this hypothesis, the H154 was mutated to larger residues like Trp (H154W) and Tyr (H154Y), and the dynamic responsiveness of these two variants against four structurally similar phenolic acids (*p*-coumaric acid, ferulic acid, caffeic acid, and cinnamic acid) were tested using the previously constructed hybrid promoter P7 as the responding promoter (Jiang et al., 2021) (**Fig. 4.2**). The dynamic performance of the wild type (WT) PadR was also tested and used as the control. The constitutively expressed PadR would bind with the P7 promoter and inhibit the transcription of downstream reporter gene, for which we chose the commonly used enhanced green fluorescent protein encoding gene *egfp*. The addition of *p*-coumaric acid would release such inhibition and restore the expression of the *egfp*. Therefore, the green fluorescence intensity normalized with cell density (OD_{600}) can be used to reflect the dynamics of the PadR variants. The PadR variants were placed under the control of the lpp0.2 promoter which is a constitutive promoter constructed in our

previous study (Wang, J. et al., 2017a). The plasmid containing the PadR variant pCS-lpp0.2-PadR and the plasmid containing the P7 promoter pZE-P7-egfp (Jiang et al., 2021) were co-transferred in to *E. coli* BW25113 F'. Gradient concentrations (0 mg/L, 300 mg/L, and 600 mg/L) of the aromatic compounds were added to induce the sensor system. As we would expect, the H154W and H154Y were unable to be activated by *p*-coumaric acid and ferulic acid (**Fig. 4.2a-c**), but the H154W exhibited responsiveness towards cinnamic acid albeit low induction level (**Fig. 4.2d**). While the *p*-coumaric acid and caffeic acid were unable to activate the H154Y, the increased concentration of *p*-coumaric acid or caffeic acid caused a reduced output activity when using the H154Y to regulate the P7 promoter, with 31.33% and 59.47% activity can be inhibited by 600 mg/L of *p*-coumaric acid and caffeic acid, respectively, compared to that of when no inducer was present (**Fig. 4.2a,c**). One thing to be noted is that the WT PadR also exhibited a slight responsiveness towards caffeic acid, which has not been reported in previous studies (**Fig. 4.2c**).

To avoid potential loss-of-function mutations, we selected the K127, instead of R164, as our next target because the K127 does not directly interact with the substrate. We hypothesized that, by replacing the K127 to larger amino acids, especially those that would significantly reshape the substrate binding pocket and push the substrate towards the H154, the PadR would be able to recognize cinnamic acid or even smaller aromatic compounds. Also, the substrate binding pocket would be narrowed and it would likely improve the binding affinity between PadR and the native substrates like *p*-coumaric acid or ferulic acid. To test our hypothesis, the residue K127 was mutated to Tyr (K127Y) and Arg (K127R) to narrow the substrate binding pocket, and the dynamic performance of these two variants was tested against the aforementioned four phenolic acids. The K127Y, as expected, exhibited improved responsiveness and sensitivity towards *p*-coumaric acid,

ferulic acid, and caffeic acid compared with the WT PadR (**Fig. 4.2**). It also recognized cinnamic acid as an inducer, with a much better induction fold (2.50-fold) than that of H154W (1.45-fold) (**Fig. 4.2d**). The K127R, on the other hand, showed reduced responsiveness towards *p*-coumaric acid and ferulic acid, and it was unable to be activated by caffeic acid and cinnamic acid (**Fig. 4.2**). The improved responsiveness of K127Y would expand its applications in biosensor-based dynamic pathway regulation and high-throughput screening, but its relaxed substrate specificity may cause cross-talk effects inside the cells. The variant K127R, though showing reduced responsiveness, exhibited an improved substrate specificity towards *p*-coumaric acid and ferulic acid, and thus is also suitable to be applied as a biosensor in metabolic engineering and synthetic biology, especially if cross-talk effects would be a concern for the related applications. The improved substrate specificity of variant K127R was likely because the slightly increased size of arginine (R) compared with lysine (K) and the change of polar charge. While both lysine and arginine are positively charged, basic, polar, and hydrophilic amino acids, the 3-carbon aliphatic straight chain ending in a guanidino group possibly provided more blockages for substrate entrance, and it will no longer be positively charged compared with lysine that harbors a positively charged ending ammonia group.

4.3.3 Profiling the substrate scope of PadR variants with smaller aromatic compounds

Observing the encouraging results against phenolic acids, we wondered whether these PadR variants can respond to aromatic acids with smaller structures, such as the 4-hydroxybenzoic acid, anthranilic acid, salicylic acid, and vanillic acid. These aromatic compounds were important precursors for value-added natural products such as coumarins, vanillin, and alkaloids. As the

induction folds were not very striking in the previous test (**Fig. 4.2**), we replaced the promoter P7 with a stronger and more sensitive hybrid promoter P9 developed in our previous study (Jiang et al., 2021). The dynamic performance of the four PadR variants pairing with the P9 hybrid promoter was tested. Surprisingly, only the K127Y showed noticeable activation when induced with 300 mg/L 4-hydroxybenzoic acid, and further increase of 4-hydroxybenzoic acid concentration did not elevate the responsiveness (**Fig. 4.3a**). As for the anthranilic acid and salicylic acid, three variants, H154W, H154Y, and K127Y, can be partially activated by these two substrates but only with low induction folds (**Fig. 4.3b,c**). While the slightly increased fluorescence intensity observed in WT PadR may also indicate a responsiveness towards these two substrates, such increase is too low and not significant enough to demonstrate its activity against anthranilic acid and salicylic acid. For the vanillic acid, interestingly, none of the variants showed observable responsiveness (**Fig. 4.3d**), but the H154W showed a slightly decreased fluorescence intensity (approximately a 18.12% reduction in *egfp* expression level) when induced with 600 mg/L vanillic acid (**Fig. 4.3d**). Notably, the K127R still did not show any responsiveness towards these four aromatic acids, which further confirmed its improved specificity towards *p*-coumaric acid and ferulic acid.

4.3.4 Site-saturated mutagenesis on key residues in the substrate binding pocket

The rational engineering of PadR expanded its substrate scope to cinnamic acid, 4-hydroxybenzoic acid, anthranilic acid, and salicylic acid. To further broaden the substrate spectrum of PadR, we sought to perform site-saturated mutagenesis on the PadR binding pockets and see whether the regulator can be engineered to respond more aromatic compounds. Three aromatic compounds, 4-aminobenzoic acid, 4-hydroxycoumarin, and vanillic acid, were chosen as the target effectors.

There are no reported biosensor for the former two chemicals but one biosensor has been reported for the vanillic acid (Kunjapur and Prather, 2019). To enable a quick selection process, a high-throughput screening workflow was established (**Fig. 4.4a**). The PadR variants generated from site-saturated mutagenesis would be under the control of lpp0.2 promoter and the expression cassette would be constructed to pCS27 plasmid (Shen and Liao, 2008). The library will then be co-transferred with the pZE-P7-egfp, and the strain would grow on LB plate with no inducer. In this process, transformants with no or low leaky expression (colonies with dim green fluorescence intensity) would be picked and streaked in LB plates with 400 mg/L inducers (namely the 4-aminobenzoic acid, 4-hydroxycoumarin, and vanillic acid). The strains with noticeable increase in green fluorescence intensity would be selected for further testing. Finally, the variants exhibiting responsiveness towards the target substrates would be sequenced and the specific mutations would be revealed (**Fig. 4.4a**).

As observed in the rational engineering, the substitution of the conserved residues (like H154) would likely abolish the function of PadR. Therefore, we sought to target the three boundary residues on the substrate binding pocket (K127, L131, and S134) simultaneously instead of the H154 and R164. Over 9,000 colonies were generated through site-saturated mutagenesis and screened by the aforementioned high-throughput workflow (**Fig. 4.4a**). Most of the variants became insensitive to the PadR repression and exhibited prominent leaking activities after the site-saturated mutagenesis. Thus, some colonies with visible green fluorescence even without the inducer were also randomly picked up. Further screening using streaked strains on substrate-containing LB plates revealed several interesting variants that showed notable responsiveness towards 4-hydroxycoumarin, 4-aminobenzoic acid, and vanillic acid. The responsiveness of these

selected variants as well as the WT PadR against the 4-hydroxycoumarin, 4-aminobenzoic acid, and vanillic acid were further quantified in test tube with gradient effector concentrations (0, 300, and 600 mg/L) (**Fig. 4.4c-f**). Interestingly, while the WT PadR did not respond to 4-aminobenzoic acid and vanillic acid, it exhibited responsiveness towards 4-hydroxycoumarin, but this was only a very slight in terms of fluorescence intensity increase (from 522 to 752 a.u.) (**Fig. 4.4b**). After the mutagenesis, the selected variants still all retained this responsiveness towards 4-hydroxycoumarin (**Fig. 4.4c-f**). The variant 3M4145, later sequenced to be PadR (K127R, L131T, S134R) followed this responsive pattern but enabled an increased egfp expression level change (from 1514 to 2040 a.u.) when induced by 4-hydroxycoumarin (**Fig. 4.4e**). The variant 3M1875 (K127N, L131K, S134Q) showed slight induction when induced with 600 mg/L 4-aminobenzoic acid but maintained its responsiveness towards 4-hydroxycoumarin (**Fig. 4.4c**). Similarly, the variant 3M3287 (K127L, L131V, S134Q) kept the responsiveness against 4-hydroxycoumarin but started to show new yet slight responsiveness to vanillic acid (**Fig. 4.4d**). The most interesting variant was the 3M6657. The DNA sequencing revealed that the K127 residue in this variant was mutated to the stop codon TGA. In other words, the 3M6657 was actually a truncated PadR (Δ127-182). This truncation resulted in prominent leaking of the sensor system, but exhibited apparent responsiveness towards all three substrates (**Fig. 4.4f**). The random mutagenesis further expanded the substrate scope of PadR to 4-hydroxycoumarin, 4-aminobenzoic acid, and vanillic acid. Nonetheless, the variants selected from the semi-rational engineering still suffered from the high leaky activities or low induction fold. Future optimization can be conducted to reduce the leaking expression and improve the dynamic ranges.

4.4 Discussion

Biosensor-enabled strategies such as dynamic pathway regulations and high-throughput screenings are widely used in metabolic engineering and synthetic biology. Broadening the responsive substrate scopes of the biosensors would largely improve the applicability of these biosensor-based strategies. In this study, through both rational and semi-rational approaches, the phenolic acid responsive repressor PadR from the *Bacillus subtilis 168* was engineered to respond eight new aromatic substrates (caffeic acid, cinnamic acid, 4-hydroxybenzoic acid, anthranilic acid, salicylic acid, vanillic acid, 4-aminobenzoic acid, and 4-hydroxycoumarin). Among these new effectors, there were no reported transcriptional factor that can respond to caffeic acid, 4-aminobenzoic acid, and 4-hydroxycoumarin. While the substrate scope of PadR was expanded, some variants did not show ideal induction folds or substantial leaky activities when induced by the new effectors. Plenty of work and strategies have been developed to improve the dynamic ranges of biosensor system, such as promoter engineering, DNA binding affinity engineering, and modulating the expression level of regulator (Ding et al., 2021; Jiang et al., 2021; Siedler et al., 2017; Zou et al., 2021). Future efforts can be devoted to further improving the dynamic properties of the variants obtained in this study. The PadR with expanded substrates would provide new opportunities for designing genetic circuits, performing dynamic pathway regulations, or developing high-throughput screening strategies in metabolic engineering and synthetic biology.

4.5 Materials and Methods

4.5.1 Strains, medium, and chemicals

The *E. coli* strain XL1-Blue (Stratagene, La Jolla, CA) was used for plasmid construction and *E. coli* BW25113 (F') was used for characterizing the dynamic performance of the PadR and its variants. Regarding the strain incubation and plasmid propagation, Luria-Bertani (LB) medium containing 10 g/L NaCl, 5 g/L yeast extract, and 10 g/L tryptone was selected. The antibiotics ampicillin and kanamycin were added in the culture, if necessary, with the concentration of 100 μg/mL and 50 μg/mL, respectively. The methanol (purchased from Fisher Scientific) was chosen as the dissolving reagent to prepare master stocks of *p*-coumaric acid (MP Biomedical LLC), ferulic acid (Sigma-Aldrich), cinnamic acid (Sigma-Aldrich), caffeic acid (Sigma-Aldrich), 4-hydroxybenzoic acid (Sigma-Aldrich), anthranilic acid (Sigma-Aldrich), salicylic acid (Alfa Aesar), vanillic acid (Alfa Aesar), 4-aminobenzoic acid (BeanTown Chemical), and 4-hydroxycoumarin (ACROS Organics) with a concentration of 100 g/L. For plasmid construction, the High-Fidelity Phusion DNA polymerase, restriction enzymes, and Quick Ligation Kit from New England Biolabs (Beverly, MA, USA) were used. The Zyppy Plasmid Miniprep Kit, Zymoclean Gel DNA Recovery Kit and DNA Clean and Concentrator-5 were purchased from Zymo Research (Irvine, CA, USA) were used to extract plasmids and purify DNA products, respectively.

4.5.2 DNA manipulation

All plasmids and strains used in this study were listed in **Table 4.1**. Plasmids pZE12-luc (high-copy number) and pCS27 (medium-copy number) were employed for constructing the biosensor system in this work (see Table 1 for a list of all strains and plasmids used in this study). The pCS27-

lpp0.2-PadR was used as the template for both rational and semi-rational mutagenesis of PadR (Jiang et al., 2021). The rational mutagenesis of PadR was carried out using overla*p*-extension PCR (OE-PCR). The gene fragment of mutated PadR variants generated by OE-PCR were digested by Acc65I and BamHI, and constructed to pCS-lpp0.2-egfp (Wang, J. et al., 2017a) to replace the eGFP gene, resulting in pCS-lpp0.2-H154W, pCS-lpp0.2-H154Y, pCS-lpp0.2-K127Y, and pCS-lpp0.2-K127R.

4.5.3 Site-saturated mutagenesis and high-throughput screening

Site-saturated mutagenesis was carried out using the strategy illustrated in **Fig. 4.4a**. by introducing random mutations in the reverse primer named 3M-HindIII-R with a DNA sequence (5'-3') of "GGAAAAGCTTTTCATANNNTCCCTGNNNGTCAGATAANNNGGCCTGGCGTTTTTGCAGCTGG". The codons encoding the K127, L131, and S134 residues were replaced with NNN to randomly generate the amino acid substitutions. The reverse primer 3M-HindIII-R was paired with the forward primer 3M-HindIII-F (with a DNA sequence of GGAAAAGCTTATGGCTTCAGCAGAGCCGATGTCATTTTC) to amplify the entire plasmid using pCS-lpp0.2-PadRwt (Jiang et al., 2021) as the template. This would lead to linear plasmids harboring the saturated mutagenesis library of PadR, which were then digested by HindIII and DpnI. The digestion products were purified and used for DNA ligation by Quick Ligation Kit purchased from New England Biolabs (Beverly, MA, USA). The ligation products will then be co-transferred with the plasmid pZE-P7-egfp into *E. coli* BW25113 (F'). The transformants were spread evenly on the LB plates with appropriate antibiotics but without any inducers. After overnight cell growth, the colonies on the plates exhibited different level of green fluorescence

intensities. These colonies were picked and streaked on a secondary LB plate containing 400 mg/L corresponding substrate (4-aminobenzoic acid, 4-hydroxycoumarin, or vanillic acid). The streaked colonies exhibited noticeable increase in fluorescence intensities after growing on the plate containing the substrate would be selected for further dynamic performance characterization.

4.5.4 Dynamic performance characterization and fluorescence assay

The *E. coli* BW25113(F') was used for characterizing the dynamic performance of the biosensors harboring PadR variants. Three independent transformants were randomly picked and cultivated in 3.5 mL of LB medium with appropriate antibiotics. The seeds were incubated at 37 °C with a shaking speed at 270 rpm in the New Brunswick Excella E24 shaker for around 12 h. 150 µl seeds were transferred into fresh 3.5 mL LB medium. When OD_{600} reached around 0.4 (roughly after 1h of cultivation), different concentrations of inducers/effectors were added into the medium. After 12 h of cultivation, all cultures were sampled and the fluorescence intensity and cell density of the samples were measured, for which the Synergy HT plate reader from Biotek was used. The samples were diluted by 5 times (40 µL sample with 160 µL DI water) and transferred into a 96-well plate (Corning® 96-well Flat Clear Bottom Black Polystyrene TC-treated Microplates, Corning 3603). The green fluorescence intensity was detected by using an excitation filter of 485/20 nm and an emission filter of 528/20 nm. The cell densities (OD_{600}) were also collected using this plate reader. The green fluorescence intensities were normalized with the corresponding cell densities to calculate the unit egfp expression levels (RFU/OD_{600}).

4.6 Tables and Figures

Table 4.1 Strains and plasmids used in this study

Name	Description	Source
E. coli XL1-Blue	*recA1 endA1gyrA*96*thi*-1*hsdR*17*supE*44*relA*1*lac*	Stratagene
E. coli BW25113 (F')	*rrnBT14 ⊿lacZWJ16 hsdR514 ⊿araBADAH33 ⊿ rhaBADLD78 F′ [traD36 proAB lacIqZ⊿M15 Tn10(Tetr)]*	(Atsumi et al., 2008)
Plasmids		
pZE-P7-egfp	pZE12luc harboring the hybrid promoter P7-controlled egfp expression cassette	(Jiang et al., 2021)
pZE-P9-egfp	pZE12luc harboring the hybrid promoter P9-controlled egfp expression cassette	(Jiang et al., 2021)
pCS-lpp0.2-egfp	pCS27 harboring a lpp0.2-controlled egfp expression cassette	(Wang, J. et al., 2017a)
pCS-lpp0.2-PadRwt	pCS27 harboring a lpp0.2-controlled wild type PadR expression cassette	(Jiang et al., 2021)
pCS-lpp0.2-H154W	pCS27 harboring a lpp0.2-controlled PadR (H154W) expression cassette	This study
pCS-lpp0.2-H154Y	pCS27 harboring a lpp0.2-controlled PadR (H154Y) expression cassette	This study

pCS-lpp0.2-K127Y	pCS27 harboring a lpp0.2-controlled PadR (K127Y) expression cassette	This study
pCS-lpp0.2-K127R	pCS27 harboring a lpp0.2-controlled PadR (K127R) expression cassette	This study
pCS-lpp0.2-3M1875	pCS27 harboring a lpp0.2-controlled PadR (K127N, L131K, S134Q) expression cassette	This study
pCS-lpp0.2-3M3287	pCS27 harboring a lpp0.2-controlled PadR (K127L, L131V, S134Q) expression cassette	This study
pCS-lpp0.2-3M4145	pCS27 harboring a lpp0.2-controlled PadR (K127R, L131T, S134R) expression cassette	This study
pCS-lpp0.2-3M6657	pCS27 harboring a lpp0.2-controlled PadR (Δ127-182) expression cassette	This study

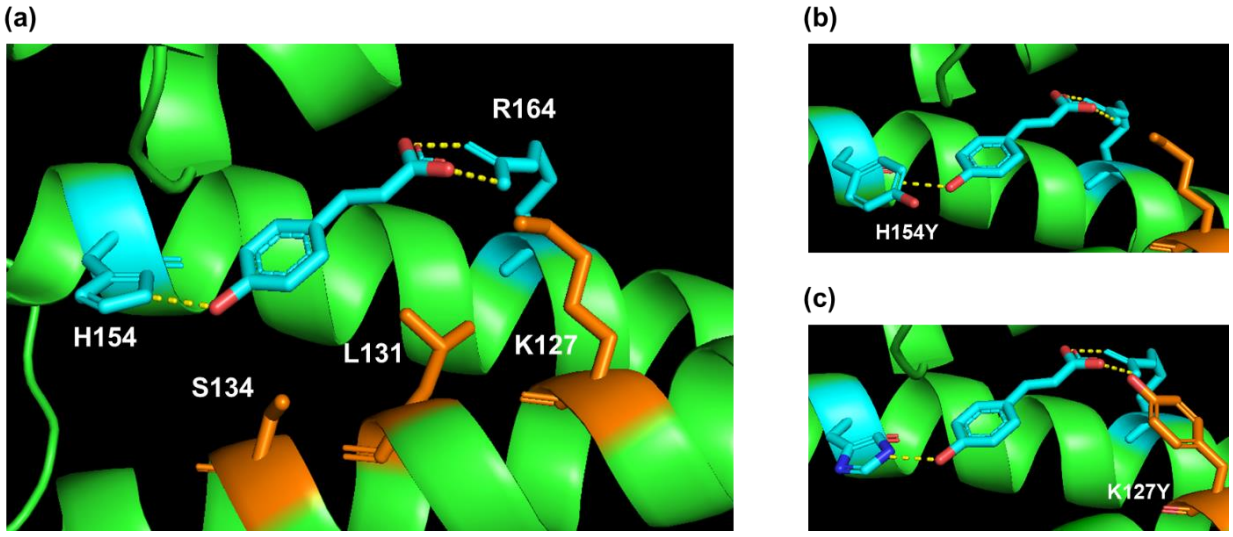


Figure 4.1. The simulated substrate binding pocket of PadR from *Bacillus subtilis 168*. (a) Wild type (WT) PadR substrate binding pocket with key residues marked up in white. The two residues H154 and R164 that directly interact with the substrate *p*-coumaric acid were colored in cyan. The three residues K127, L131, and S134 that located at the boundary of the binding pocket but did not directly interact with the substrate were colored in orange. **(b)** PadR substrate binding pocket with H154Y mutation. **(c)** PadR substrate binding pocket with K127Y mutation.

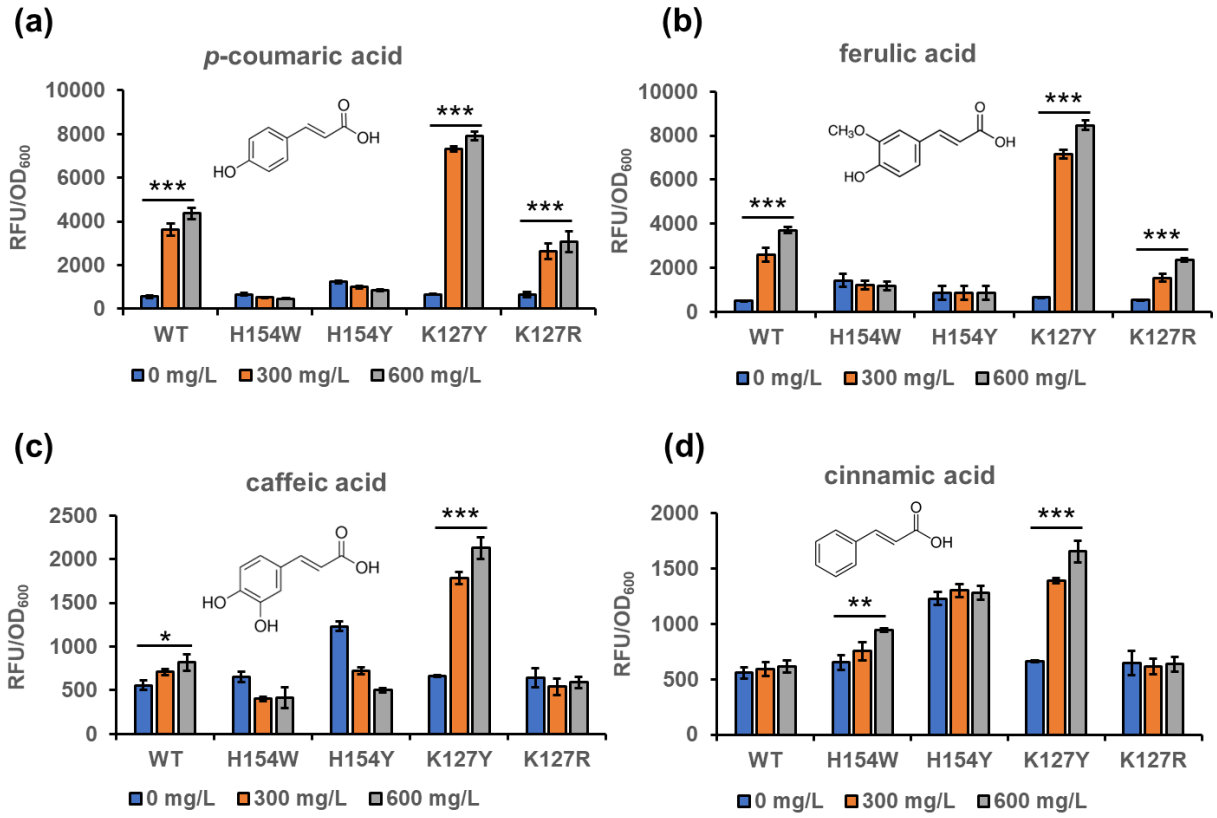


Figure 4.2. Dynamic performance of PadR variants towards different phenolic acids. (a) The dynamic responsiveness of PadR variants against *p*-coumaric acid. **(b)** The dynamic responsiveness of PadR variants against ferulic acid. **(c)** The dynamic responsiveness of PadR variants against caffeic acid. **(d)** The dynamic responsiveness of PadR variants against cinnamic acid. All data represent the mean of 3 biologically independent samples and error bars show standard deviation. Statistical analysis was performed by using Student's t test (two-tailed; unpaired; *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$).

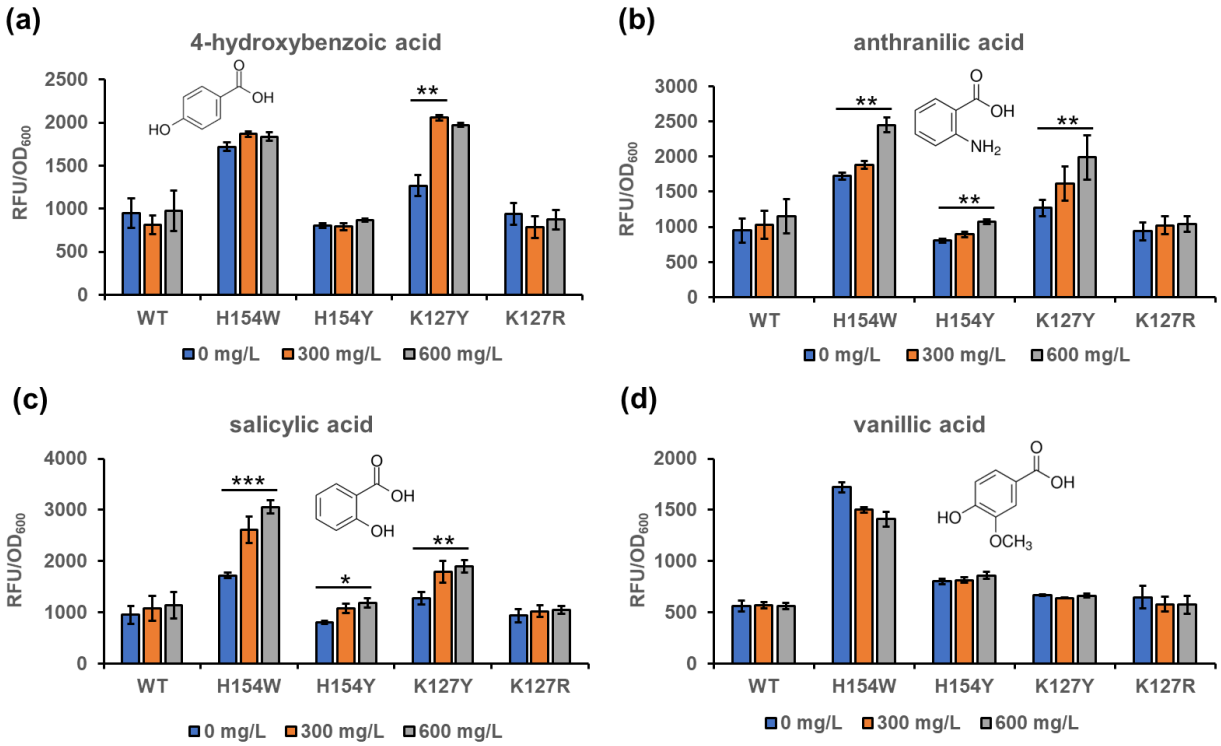


Figure 4.3. Dynamic performance of PadR variants towards smaller aromatic acids. (a) The dynamic responsiveness of PadR variants against 4-hydroxybenzoic acid. **(b)** The dynamic responsiveness of PadR variants against anthranilic acid. **(c)** The dynamic responsiveness of PadR variants against salicylic acid. **(d)** The dynamic responsiveness of PadR variants against vanillic acid. All data represent the mean of 3 biologically independent samples and error bars show standard deviation. Statistical analysis was performed by using Student's t test (two-tailed; unpaired; *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$).

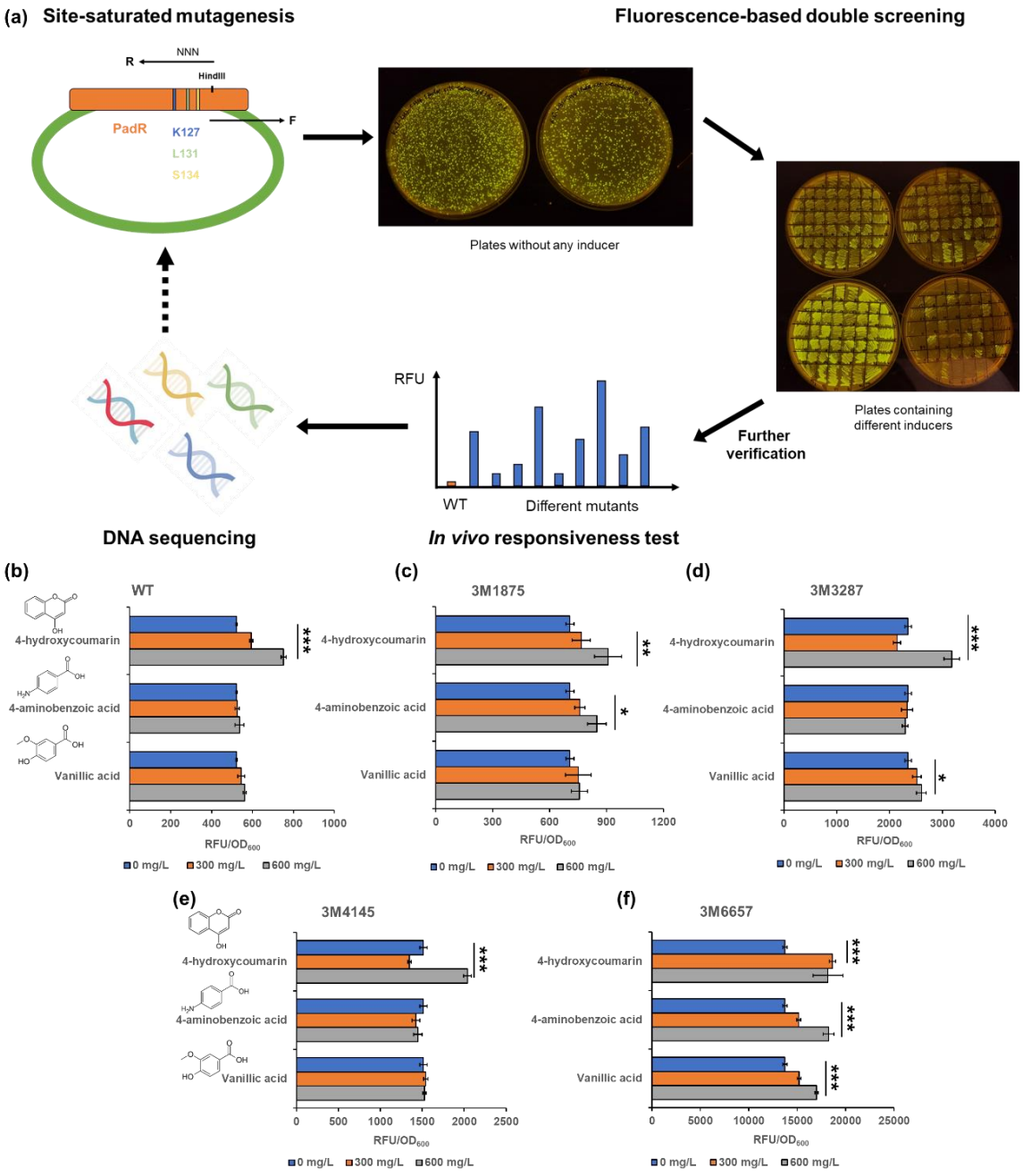


Figure 4.4. **Semi-rational engineering of PadR dynamic performance of selected mutants towards different substrates. (a)** The schematic illustration of high-throughput screening workflow. **(b)** Dynamic responsiveness of wild type PadR against 4-hydroxycoumarin, 4-

aminobenzoic acid, and vanillic acid. **(c)** Dynamic responsiveness of PadR variant 3M1875 against 4-hydroxycoumarin, 4-aminobenzoic acid, and vanillic acid. **(d)** Dynamic responsiveness of PadR variant 3M3287 against 4-hydroxycoumarin, 4-aminobenzoic acid, and vanillic acid. **(e)** Dynamic responsiveness of PadR variant 3M4145 against 4-hydroxycoumarin, 4-aminobenzoic acid, and vanillic acid. **(f)** Dynamic responsiveness of PadR variant 3M6657 against 4-hydroxycoumarin, 4-aminobenzoic acid, and vanillic acid. All data represent the mean of 3 biologically independent samples and error bars show standard deviation. Statistical analysis was performed by using Student's t test (two-tailed; unpaired; *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$).

CHAPTER 5

CONCLUSION

Establishing microbial cell factories for biosynthesis of value-added compounds, such as commodity chemicals, biofuels, and biopharmaceuticals, have become increasingly important due to the short doubling-time, ease of manipulation and cultivation, clear genetic background, and large amount of available genetic engineering toolsets. To improve the biosynthesis efficiencies of microbial cell factories, plenty of metabolic engineering strategies have been applied, such as adaptive laboratory evolution to obtain high-producing phenotypes, protein/enzyme engineering to improve the activity of rate-limiting biocatalysts, dynamic pathway regulation to minimize the competition between cell growth and target production, and host engineering to eliminate the competing pathways. In these processes, genetically-encoded biosensors have been widely applied to dynamically rewire carbon fluxes or establish high-throughput screening to reduce the workload when facing large size of libraries. Among the genetically encoded biosensors, transcriptional factors-based biosensors are most commonly used due to their great availabilities, relatively broader ligand spectra, and ease of optimization and engineering. As demonstrated in chapter 2 and chapter 3, the employment of biosensors enabled the dynamic pathway regulation via plasmid replication control, and screening of high-producing strains in a high-throughput manner, respectively. These results have demonstrated how the involvement of biosensors can effectively accelerate the engineering processes of microbial cell factories.

Nonetheless, existing transcriptional factors-based biosensors sometimes suffer from a limited scope of ligands, which substantially hindered the general applicability of biosensor-assisted metabolic engineering strategies. Thus, engineering existing biosensors to broaden the substrate scope has become one of the major approaches to further exploit the potential of transcriptional factors-based biosensors. In chapter 4, we engineered the phenolic acid-responsive regulator PadR from *Bacillus subtilis* to broaden its ligand spectrum, with the new and improved responsiveness towards cinnamic acid, caffeic acid, ferulic acid, 4-hydroxybenzoic acid, anthranilic acid, salicylic acid, vanillic acid, 4-aimnobenzoic acid, and 4-hydroxycoumarin. The novel biosensor variants obtained in this work with expanded ligand spectra and diversified dynamic properties provide versatile choices for fulfilling different application scenarios and requirements.

Harnessing plasmid replication mechanism to enable dynamic control of gene copy in bacteria

Biosensor-enabled dynamic pathway regulations have been widely applied in metabolic engineering to appropriately rewire carbon fluxes and meanwhile minimize the defects on cell growth. A series of gene expression regulation tools have been developed and optimized to achieve efficient dynamic pathway regulations from transcriptional level to post-translational level. While these regulation strategies have been demonstrated efficient in dynamic regulations of gene expression, the regulation at gene replication level was only exploited in a static manner by substituting the pathway-carrying plasmids with different copy numbers (Ajikumar et al., 2010; Jones et al., 2000).

In Chapter 2, we established dynamic regulation of gene copy numbers by tuning the replication of ColE1-derived plasmids in accordance to the inducer concentration, and expanded the dynamic regulation, for the first time, to gene replication level. Based on the understanding of the replication mechanism, the tunable parameter was determined to be the intracellular availability of RNAI and RNAII. By controlling the transcription of the RNAI and RNAII, we were able to establish the dynamically regulated gene replication in ColE1-derived plasmids. With the establishment of inducible gene replication control, this strategy was further diversified with different dynamic behaviors by constructing three prototypic genetic circuits with a *p*-coumaric acid biosensor. To demonstrate the real-world applicability of this strategy, the developed genetic circuits were then applied to improve the *p*-coumaric acid biosynthesis. The dynamic gene copy regulation enabled a better *p*-coumaric acid production, with the titer reaching 1.69 g/L in shake flasks. The unregulated control accumulated about 1.29 g/L *p*-coumaric acid, but the static control (the gene replication enhancement was constitutively executed) only enabled a titer of 0.97 g/L and a defected cell growth (**Fig. 2.6d**).

Besides validating the applicability of the dynamic gene copy control in microbial biosynthesis, our results also highlighted the superiority of dynamic pathway regulation. When the gene copy numbers were constantly maintained at normal level (unregulated control), the cells cannot benefit from the later boost on pathway expression level to direct more carbon fluxes to the product, but when the gene copy numbers were constitutively kept on a high level (static regulation control), it hampered the cell growth and negatively affected the cell fitness, and the production of target compound plummeted to undesired level.

The strategy developed in this work expanded the dynamic gene expression regulation to gene replication level, which has not been achieved previously. This provides an extra layer of regulation on gene expression and can offer more choices when selecting toolsets for regulating cellular metabolism. An additional benefit from the gene copy regulation is the simultaneous u*p*- or down-regulation of multiple targets, because the inhibition or enhancement of plasmid replication would decrease or increase the copy numbers of all the genes in the same plasmid, respectively. This would be especially helpful when regulating complex pathways with multiple regulation targets.

While the applicability of dynamic gene copy regulation has been demonstrated in real-world applications, one recurring challenge hindering the plasmid-based dynamic regulation is the instability of plasmids, which has been a long-talked question in biotechnology industry when scaling up the microbial productions (Rugbjerg et al., 2018). The stains suffering from loss-of-function plasmids or even complete plasmid loss would not benefit from the dynamic gene copy regulation and may even completely lose the ability for producing the target compounds. To solve this problem, there are many ways to construct long-lasting plasmids which can be stably maintained in cells during the fermentation. For example, our lab has streamlined the workflow to establish stable plasmids by sequestrating the growth-essential gene *folP*. The expression cassette for the *folP* of *E. coli* was constructed to the plasmid containing the biosynthetic pathways while the endogenous genome copy of *folP* was deleted. Thus, only the strains harboring the fully functional plasmid can survive during the fermentation, and as a result, such plasmids were able to be stably maintained in cells over 80 generations of cultivation (Zhang et al., 2019). Therefore, further integration of such strategy with the dynamic gene copy regulation in future studies would

increase the robustness and stability of the plasmid-based dynamic regulation, and thus improve the applicability of the dynamic gene copy regulation in biotechnology industry.

Besides the plasmid instability, the performance and applicable scenarios of dynamic gene copy regulation strongly rely on the dynamic properties of biosensors. Hence, another potential limiting factor that may constrain the applications of this strategy is the repertoire of usable genetically encoded biosensors to control the gene expression, which is also a common constraint for implementing all biosensor-related strategies. The ligand spectra of the biosensors will largely determine the applicable scenarios while the dynamic performance (including both the dynamic ranges and responsive ranges) of biosensors would affect the effectiveness of biosensor-enabled strategies. There are two major approaches to broaden the ligand spectra or improve the dynamic performance of the biosensors: one is the continuous mining and characterization of novel biosensors from the nature. The nature provides a substantial reservoir of biosensors awaiting to be discovered. With the rapid development of bioinformatic techniques and the increased accuracy of genetic annotations, we can anticipate that more biosensors with versatile dynamic performance and novel ligand preferences can be revealed. As a more straightforward approach, engineering existing biosensors to expand their ligand spectra or optimize their dynamic performance have been demonstrated efficient in recent years (Ding et al., 2021; Jiang et al., 2021; Li et al., 2020; Zou et al., 2021). Promoter engineering and RBS substitution were applied to fine-tune the regulator expression and alter the dynamic performance of the biosensors. Besides, protein engineering to modify the regulator structures has been demonstrated efficient in both improving the dynamic performance and broadening the substrate spectra of biosensors. Therefore, with the

expansion of the repertoire of usable biosensors, the applicability of dynamic gene copy regulation would be continuously improved.

Biosensor-assisted titratable CRISPRi high-throughput (BATCH) screening for over-production phenotypes

Rewiring carbon flux to target production by down-regulating competitive pathways was a common strategy to enhance the microbial biosynthesis, but excessive inhibition of growth-essential genes would reduce the flux to biomass generation and thus impair the cell growth, which in turns may negatively affect the productivities of the microbial cell factories. Thus, appropriate level of regulation, especially against growth-essential genes, to rewire carbon flux with minimized disturbance of cell growth is necessary for establishing effective microbial cell factories.

In this work, we first established the modified CRISPRi system with mismatched sgRNA arrays that can afford varied inhibition on target genes by locating the sensitive mismatch window (PAM-proximal 7-8th bp region) (**Fig. 3.2B,C**). Then, we demonstrated that the system was able to achieve a varied inhibition efficiency on egfp expression level ranging from 14.6% to 95.6% (**Fig. 3.2D**). Next, a one-pot sgRNA construction approach that can randomly introduce the double mismatches in the desired window was adopted to assembly the sgRNA arrays. We demonstrated the applicability of this strategy by constructing sgRNA arrays targeting the RBS region and coding sequence of the egfp expression cassette, and in both case the mismatched CRISPRi system have validated a wide range of inhibition (**Fig. 3.2E,F**). On locating the mismatch window and validating the varied inhibition efficiencies of mismatched CRISPRi, we then applied this system

to systematically target the central metabolism of *E. coli*, rewiring carbon flux to improve the biosynthesis of *p*-coumaric acid, the important precursor for a series of valuable natural products. To accelerate the workflow and improve the screening efficiency, the *p*-coumaric acid responsive sensor system PadR-P_{padC} that was engineered in our previous study (Jiang et al., 2021) was optimized and applied to assist the screening (**Fig. 3.3**). A total of 284 colonies with notably higher green fluorescence intensities on the plates were picked and screened. Among them, 62 variants exhibiting significant increase in *p*-coumaric acid production were selected for further validation in shake flasks fermentation. The sgRNAs targeting the csrA, ptsI, pykF, pfkA, or ppc was demonstrated to be beneficial for *p*-coumaric acid production. Further combining of multiple sgRNAs boosted the *p*-coumaric acid titer to 1308.6 g/L, which was 1.4-fold of the titer of control strain (**Fig. 3.4**). To demonstrate the general applicability of this strategy, we aimed to target the acetyl-CoA derived products as a secondary proof-of-concept demonstration. Acetyl-CoA is one of the most important central metabolites and many valuable products such as terpenoids, polyketides, and fatty acids, are derived from acetyl-CoA. To this end, the hpdR-P_{hpdH} biosensor system developed previously in the lab (Wang et al., 2021b) was adopted and the screening of the ligand spectrum for the biosensor system revealed that the butyric acid was the most effective ligand among all tested substrates (**Fig. 3.5B**). Thus, the BATCH screening was performed to screen potential targets that would improve the production of butyric acid (**Fig. 3.5**). The improved production of butyric acid further demonstrated the general applicability of the BATCH strategy.

One superiority using mismatched sgRNAs was claimed to be the fine-tuning of gene expression level to appropriately inhibit competitive pathway with less disturbance on cell growth. We demonstrated this by comparing the *p*-coumaric acid production when using fully matched sgRNA

and the mismatched sgRNAs targeting the *ppc*, *ptsI*, or *csrA*. The fully matched sgRNAs resulted in obvious growth defects and lowered *p*-coumaric acid production, while the mismatched sgRNA variants enabled improved p-coumaric acid titers without greatly affecting the cell growth (**Fig 3.4D**). These results again highlighted the importance of fine-tuning gene expression in constructing efficient microbial cell factories.

In the BATCH screening, the introduction of genetically encoded biosensors greatly improved the screening efficiencies. However, during the BATCH screening, we noticed that only a moderate correlation was formed between the fluorescence intensities and the titers of *p*-coumaric acid in test tube screening (**Fig. 3.4B**). Such correlation was obviously stronger when using the hpdR-based sensor in performing BATCH screening of butyric acid high-producers (**Fig. 3.5D**). While the batch-to-batch variations may also be a contributing factor to such differences, this was more likely because the hpdR-based biosensor exhibited less prominent leaky activity and a higher dynamic range compared to the PadR-based biosensor. Also, because the application of BATCH screening depends on the employment of biosensors, if a desired products cannot be recognized by current biosensors, it would be difficult to implement BATCH screening for improving its biosynthesis. Therefore, in the future, engineering the biosensors to reduce their leaky activities, improve the dynamic ranges and induction folds, and broaden the ligand spectra, would be beneficial for applying the biosensor-assisted high-throughput screenings. Specifically, fine-tuning the expression of the regulators via promoter engineering or RBS engineering can be conducted to expanding the dynamic ranges and improve the induction folds, while the expansion of ligand spectra can be explored via altering the substrate recognition domain of the transcriptional factors by protein engineering. The continuous alteration and optimization of

biosensors would greatly increase the potential of BATCH screening and accelerate the design and engineering of efficient microbial cell factories.

Expanding the substrate spectrum of a phenolic acid-responsive regulator PadR from *Bacillus subtilis*

The ligand preferences of natural transcriptional factors-based biosensors are usually identified and characterized to be closely related to the host metabolism. In the case of PadR from *Bacillus subtilis* (BsPadR), the BsPadR was utilized by the host for monitoring the concentration of phenolic acids and activate the degradation of such compounds when the concentrations are over the thresholds. As a result, the preferred substrates of PadR are phenolic acids, especially the *p*-coumaric acid and ferulic acid. Aiming at broadening the spectrum of recognizable ligands of BsPadR, we investigated the substrate binding pocket of PadR in chapter 4 and successfully enabled new responsiveness of BsPadR towards several aromatic compounds.

Guided by the structural information, we first rationally modified the binding pocket of BsPadR to reduce the pocket size. A total of five important residues were identified, including two conserved substrate binding residues H154 and R164, and three boundary residues K127, L131, and S134. The rationally designed variant K127Y enabled improved responsiveness of BsPadR towards its native substrates *p*-coumaric acid, ferulic acid, and caffeic acid, while the variant K127R showed improved ligand specificity towards *p*-coumaric acid and ferulic acid despite a slight decrease in responsiveness. The variants K127Y and H154W enabled new activities against cinnamic acid, which was the first report for such activity of PadR. Further substrate profiling of

the rationally designed variants also indicated the potentials of engineering the BsPadR for sensing the 4-hydroxybenzoic acid, anthranilic acid, and salicylic acid.

To further expand the ligand spectrum of the BsPadR, we targeted the three boundary residues K127, L131, and S134, for site-saturated mutagenesis. The semi-rational approach resulted in a big library containing thousands of BsPadR variants. To improve the screening efficiencies, a high-throughput screening workflow has been built based on the dynamic properties of BsPadR. All mutated variants will be processed through a two-step screening procedure. The first step is a negative selection which removes the variants that showed substantial leaky activities. The selected variants would go through the second screening in which their responsiveness against the substrate would be qualitatively evaluated. The dynamic performance of the final variants would be tested *in vivo* to quantitatively assess their activities towards the desired substrates, which in this case are the 4-aminobenzoic acid, vanillic acid, and 4-hydroxycoumarin. Unexpectedly, we observed a great number of variants that displayed significant leaking activities even without any induction, and thus we also included some of these variants for the next-round screening. The selection revealed several interesting variants that enabled new or improved activities against the desired substrates.

While the ligand spectrum of BsPadR was successfully expanded, the induction fold-changes of some variants obtained in this work were not ideal and thus were not ready for use in metabolic engineering and synthetic biology. Another issue lying in the selected variants from site-saturated mutagenesis was the high leaky activities when no inducer was added. Thus, following studies can explore more potential amino acid residues in the PadR substrate binding pocket and the site-

saturated mutagenesis can be carried out to investigate the effects on ligand recognition of these new variants.

Moreover, while the dynamic performance of the selected variants was not optimal, the observed induction suggested there are potentials, with future engineering and optimization, to further modify these variants and improve their usability in dynamic pathway regulation or high-throughput screening. As in the case of BsPadR, we have seen successful examples by modulating its expression level to enable different dynamic ranges, and alteration of its DNA binding regions via protein engineering to achieve relaxed or stringent DNA-regulator binding (Jiang et al., 2021; Siedler et al., 2017). Besides, promoter engineering and manipulation of the PadR binding boxes to construct hybrid promoters were also demonstrated to be efficient in reducing the leaky activities and improving the induction fold-changes (Jiang et al., 2021; Zou et al., 2021). Thus, with appropriate optimizations and engineering, the dynamic performance of the selected variants exhibiting novel responsiveness can be improved, and these variants would be readily applicable in metabolic engineering and synthetic biology.

Conclusion and outlooks

In conclusion, we exploited and demonstrated how the genetically-encoded biosensors (or more specifically the transcriptional factors-based biosensors) can assist the establishment of productive microbial cell factories by employing the biosensors in both dynamic pathway regulation and high-throughput screening. The dynamic gene copy control strategy developed in chapter 2 expanded the dynamic gene expression regulation to gene replication level, The application of the dynamic

gene copy control significantly improved the production of the value-added compound *p*-coumaric acid, which further substantiated the real-world potential of this regulatory strategy. The BATCH screening developed in chapter 3 demonstrated how the biosensor can be integrated to improve the selection processes and improve the screening efficiencies. The applications explored in these two chapters in turns demonstrated how the dynamic properties of biosensors would influence the effectiveness of the strategies. For example, the different dynamic ranges of wild type PadR and PadR (K64A) variant resulted in different performance of the genetic circuits (**Fig. 2.5**), and the different dynamic performance between the PadR-based biosensor and HpdR-based biosensor led to varied correlation between fluorescence and production (**Fig. 3.4 & Fig. 3.5**). Moreover, as biosensors can only respond to specific signals, either it was chemical or environmental, the implementation of biosensor-assisted metabolic engineering strategies would suffer from limited ligand spectra of existing biosensors. Therefore, to further expand the applicable scenarios of biosensors, we engineered the phenolic acid responsive regulator PadR in chapter 4 and make the repressor responsive for seven additional aromatic compounds. With appropriate optimization, these variants obtained in this work can be used to in biosensor-assisted metabolic engineering of the production of these chemicals and their derivatives.

As we introduced before, establishing highly efficient microbial cell factories capable of converting cheap renewable feedstocks into valuable products like biofuels, food ingredients, and pharmaceuticals, has become increasingly promising and has drawn considerable attentions in the biotechnology industry. Besides the biosensor-assisted strategies explored in this work, a series of strategies incorporating other metabolic engineering approaches and synthetic biology toolsets have also been developed to streamline the construction of effective microbial cell factories, such

as protein engineering to improve the catalytic activities of key enzymes, subcellular co-localization/compartmentalization for better substrate channeling or flux, and host engineering to eliminate non-essential competitive pathways. Biosensor-assisted dynamic pathway regulation and high-throughput screening serve as effective complementary strategies in metabolic engineering. With the continuous development and enlargement of current biosensor repertoire, systematic optimization of microbial cell factories integrating biosensors and other engineering approaches would be more commonly used in future metabolic engineering.

REFERENCES

Abernathy, M.H., He, L., Tang, Y.J., 2017. Channeling in native microbial pathways: Implications and challenges for metabolic engineering. Biotechnology Advances 35(6), 805-814.

Adegboye, M.F., Ojuederie, O.B., Talia, P.M., Babalola, O.O., 2021. Bioprospecting of microbial strains for biofuel production: metabolic engineering, applications, and challenges. Biotechnology for Biofuels 14(1), 5.

Aiba, H., 2007. Mechanism of RNA silencing by Hfq-binding small RNAs. Curr Opin Microbiol 10(2), 134-139.

Ajikumar, P.K., Xiao, W.-H., Tyo, K.E., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T.H., Pfeifer, B., Stephanopoulos, G., 2010. Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. Science 330(6000), 70-74.

Alper, H., Fischer, C., Nevoigt, E., Stephanopoulos, G., 2005. Tuning genetic control through promoter engineering. Proceedings of the National Academy of Sciences of the United States of America 102(36), 12678.

Andreozzi, S., Chakrabarti, A., Soh, K.C., Burgard, A., Yang, T.H., Van Dien, S., Miskovic, L., Hatzimanikatis, V., 2016. Identification of metabolic engineering targets for the enhancement of 1, 4-butanediol production in recombinant E. coli using large-scale kinetic models. Metabolic engineering 35, 148-159.

Atsumi, S., Cann, A.F., Connor, M.R., Shen, C.R., Smith, K.M., Brynildsen, M.P., Chou, K.J.Y., Hanai, T., Liao, J.C., 2008. Metabolic engineering of Escherichia coli for 1-butanol production. Metabolic Engineering 10(6), 305-311.

Balbás, P., Bolívar, F., 2004. pBR322 and Protein Expression Systems in E. coli. Recombinant Gene Expression(267), 77-90.

Banerjee, D., Eng, T., Lau, A.K., Sasaki, Y., Wang, B., Chen, Y., Prahl, J.-P., Singan, V.R., Herbert, R.A., Liu, Y., 2020. Genome-scale metabolic rewiring improves titers rates and yields of the non-native product indigoidine at scale. Nature communications 11(1), 1-11.

Barrick, J.E., Breaker, R.R., 2007. The distributions, mechanisms, and structures of metabolite-binding riboswitches. Genome Biol 8(11), R239.

Bekker, V., Dodd, A., Brady, D., Rumbold, K., 2014. Tools for metabolic engineering in Streptomyces. Bioengineered 5(5), 293-299.

Bester, A.C., Lee, J.D., Chavez, A., Lee, Y.-R., Nachmani, D., Vora, S., Victor, J., Sauvageau, M., Monteleone, E., Rinn, J.L., 2018. An integrated genome-wide CRISPRa approach to functionalize lncRNAs in drug resistance. Cell 173(3), 649-664. e620.

Biggs, B.W., Lim, C.G., Sagliani, K., Shankar, S., Stephanopoulos, G., De Mey, M., Ajikumar, P.K., 2016. Overcoming heterologous protein interdependency to optimize P450-mediated Taxol precursor synthesis in Escherichia coli. Proc Natl Acad Sci U S A 113(12), 3209-3214.

Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., Marraffini, L.A., 2013. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic acids research 41(15), 7429-7437.

Black, W.B., Zhang, L., Mak, W.S., Maxel, S., Cui, Y., King, E., Fong, B., Sanchez Martinez, A., Siegel, J.B., Li, H., 2020. Engineering a nicotinamide mononucleotide redox cofactor system for biocatalysis. Nature Chemical Biology 16(1), 87-94.

Brockman, I.M., Prather, K.L.J., 2015. Dynamic knockdown of E. coli central metabolism for redirecting fluxes of primary metabolites. Metab Eng 28, 104-113.

Bryant, J.A., Sellars, L.E., Busby, S.J., Lee, D.J., 2014. Chromosome position effects on gene expression in Escherichia coli K-12. Nucleic acids research 42(18), 11383-11392.

Calero, P., Nikel, P.I., 2019. Chasing bacterial chassis for metabolic engineering: a perspective review from classical to non-traditional microorganisms. Microbial Biotechnology 12(1), 98-124.

Calles, B., Goñi-Moreno, Á., de Lorenzo, V., 2019. Digitalizing heterologous gene expression in Gram-negative bacteria with a portable ON/OFF module. Molecular Systems Biology 15(12), e8777.

Calvo-Villamañán, A., Ng, J.W., Planel, R., Ménager, H., Chen, A., Cui, L., Bikard, D., 2020. On-target activity predictions enable improved CRISPR–dCas9 screens in bacteria. Nucleic acids research 48(11), e64-e64.

Cameron, D.E., Collins, J.J., 2014. Tunable protein degradation in bacteria. Nature Biotechnology 32(12), 1276-1281.

Camps, M., 2010. Modulation of ColE1-Like Plasmid Replication for Recombinant Gene Expression. Recent Patents on DNA & Gene Sequences 4(1), 58-73.

Carroll, A.L., Case, A.E., Zhang, A., Atsumi, S., 2018. Metabolic engineering tools in model cyanobacteria. Metabolic Engineering 50, 47-56.

Cesareni, G., Helmer-Citterich, M., Castagnoli, L., 1991. Control of ColE1 plasmid replication by antisense RNA. Trends in Genetics 7(7), 230-235.

Cesareni, G., Muesing, M.A., Polisky, B., 1982. Control of ColE1 DNA replication: the rop gene product negatively affects transcription from the replication primer promoter. Proceedings of the National Academy of Sciences 79(20), 6313-6317.

Chatterjee, P., Jakimo, N., Lee, J., Amrani, N., Rodríguez, T., Koseki, S.R.T., Tysinger, E., Qing, R., Hao, S., Sontheimer, E.J., Jacobson, J., 2020. An engineered ScCas9 with broad PAM range and high specificity and activity. Nature Biotechnology 38(10), 1154-1158.

Chen, J., Li, W., Yao, H., Xu, J., 2015. Insights into drug discovery from natural products through structural modification. Fitoterapia 103, 231-241.

Chen, W., Zhang, S., Jiang, P., Yao, J., He, Y., Chen, L., Gui, X., Dong, Z., Tang, S.Y., 2015. Design of an ectoine-responsive AraC mutant and its application in metabolic engineering of ectoine biosynthesis. Metab Eng 30, 149-155.

Chen, X., Liu, L., 2018. Gene Circuits for Dynamically Regulating Metabolism. Trends Biotechnol 36(8), 751-754.

Chen, Y., Ho, J.M.L., Shis, D.L., Gupta, C., Long, J., Wagner, D.S., Ott, W., Josic, K., Bennett, M.R., 2018. Tuning the dynamic range of bacterial promoters regulated by ligand-inducible transcription factors. Nat Commun 9(1), 64.

Choi, K.R., Jiao, S., Lee, S.Y., 2020. Metabolic engineering strategies toward production of biofuels. Current Opinion in Chemical Biology 59, 1-14.

Choi, Y.J., Park, J.H., Kim, T.Y., Lee, S.Y., 2012. Metabolic engineering of Escherichia coli for the production of 1-propanol. Metabolic engineering 14(5), 477-486.

Chouhan, S., Sharma, K., Zha, J., Guleria, S., Koffas, M.A.G., 2017. Recent Advances in the Recombinant Biosynthesis of Polyphenols. Frontiers in Microbiology 8, 2259.

Christiano, R., Nagaraj, N., Fröhlich, F., Walther, Tobias C., 2014. Global Proteome Turnover Analyses of the Yeasts S. cerevisiae and S. pombe. Cell Reports 9(5), 1959-1965.

Chung, H.K., Jacobs, C.L., Huo, Y., Yang, J., Krumm, S.A., Plemper, R.K., Tsien, R.Y., Lin, M.Z., 2015. Tunable and reversible drug control of protein production via a self-excising degron. Nature Chemical Biology 11(9), 713-720.

Clancy, S., 2008. DNA transcription. Nature education 1(1), 41.

Collias, D., Beisel, C.L., 2021. CRISPR technologies and the search for the PAM-free nuclease. Nature Communications 12(1), 555.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., 2013. Multiplex genome engineering using CRISPR/Cas systems. Science 339(6121), 819-823.

Cravens, A., Payne, J., Smolke, C.D., 2019. Synthetic biology strategies for microbial biosynthesis of plant natural products. Nat Commun 10(1), 2142.

Crocker, J., Stern, D.L., 2013. TALE-mediated modulation of transcriptional enhancers in vivo. Nature methods 10(8), 762.

Crook, N., Sun, J., Morse, N., Schmitz, A., 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. Applied microbiology and biotechnology 100(23), 10005-10018.

Cui, L., Vigouroux, A., Rousset, F., Varet, H., Khanna, V., Bikard, D., 2018. A CRISPRi screen in E. coli reveals sequence-specific toxicity of dCas9. Nature communications 9(1), 1-10.

Culver, G.M., 2001. Meanderings of the mRNA through the Ribosome. Structure 9(9), 751-758.

Dahl, R.H., Zhang, F., Alonso-Gutierrez, J., Baidoo, E., Batth, T.S., Redding-Johanson, A.M., Petzold, C.J., Mukhopadhyay, A., Lee, T.S., Adams, P.D., Keasling, J.D., 2013. Engineering dynamic pathway regulation using stress-response promoters. Nat Biotechnol 31(11), 1039-1046.

Datsenko, K.A., Pougach, K., Tikhonov, A., Wanner, B.L., Severinov, K., Semenova, E., 2012. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. Nature communications 3(1), 1-7.

de Bakker, V., Liu, X., Bravo, A.M., Veening, J.-W., 2022. CRISPRi-seq for genome-wide fitness quantification in bacteria. Nature protocols, 1-30.

Del Solar, G., Giraldo, R., Ruiz-Echevarría, M.J., Espinosa, M., Díaz-Orejas, R., 1998. Replication and control of circular bacterial plasmids. Microbiol. Mol. Biol. Rev. 62(2), 434-464.

Deng, J., Chen, C., Gu, Y., Lv, X., Liu, Y., Li, J., Ledesma-Amaro, R., Du, G., Liu, L., 2019. Creating an in vivo bifunctional gene expression circuit through an aptamer-based regulatory mechanism for dynamic metabolic engineering in Bacillus subtilis. Metabolic Engineering 55, 179-190.

Ding, N., Zhou, S., Deng, Y., 2021. Transcription-Factor-based Biosensor Engineering for Applications in Synthetic Biology. ACS Synthetic Biology 10(5), 911-922.

Dinh, C.V., Chen, X., Prather, K.L.J., 2020. Development of a Quorum-Sensing Based Circuit for Control of Coculture Population Composition in a Naringenin Production System. ACS Synthetic Biology 9(3), 590-597.

Dinh, C.V., Prather, K.L., 2019. Development of an autonomous and bifunctional quorum-sensing circuit for metabolic flux control in engineered Escherichia coli. Proceedings of the National Academy of Sciences 116(51), 25562-25568.

Doong, S.J., Gupta, A., Prather, K.L.J., 2018. Layered dynamic regulation for improving metabolic pathway productivity in Escherichia coli. Proc Natl Acad Sci U S A 115(12), 2964-2969.

Eggeling, L., Bott, M., Marienhagen, J., 2015. Novel screening methods--biosensors. Curr Opin Biotechnol 35, 30-36.

Ehrenworth, A.M., Peralta-Yahya, P., 2017. Accelerating the semisynthesis of alkaloid-based drugs through metabolic engineering. Nature chemical biology 13(3), 249-258.

Farmer, W.R., Liao, J.C., 2000. Improving lycopene production in Escherichia coli by engineering metabolic control. Nature biotechnology 18(5), 533-537.

Feist, A.M., Palsson, B.Ø., 2008. The growing scope of applications of genome-scale metabolic reconstructions using Escherichia coli. Nature biotechnology 26(6), 659-667.

Feng, H., Guo, J., Wang, T., Zhang, C., Xing, X.-h., 2021. Guide-target mismatch effects on dCas9–sgRNA binding activity in living bacterial cells. Nucleic Acids Research 49(3), 1263-1277.

Fernandez-Rodriguez, J., Voigt, C.A., 2016. Post-translational control of genetic circuits using Potyvirus proteases. Nucleic acids research 44(13), 6493-6502.

Ferry, Q.R., Lyutova, R., Fulga, T.A., 2017. Rational design of inducible CRISPR guide RNAs for de novo assembly of transcriptional programs. Nature communications 8(1), 1-10.

Fletcher, E., Krivoruchko, A., Nielsen, J., 2016. Industrial systems biology and its impact on synthetic biology of yeast cell factories. Biotechnology and Bioengineering 113(6), 1164-1170.

Fontana, J., Sparkman-Yager, D., Zalatan, J.G., Carothers, J.M., 2020. Challenges and opportunities with CRISPR activation in bacteria for data-driven metabolic engineering. Current Opinion in Biotechnology 64, 190-198.

Galanie, S., Thodey, K., Trenchard, I.J., Filsinger Interrante, M., Smolke, C.D., 2015. Complete biosynthesis of opioids in yeast. Science 349(6252), 1095-1100.

Gallone, B., Steensels, J., Prahl, T., Soriaga, L., Saels, V., Herrera-Malaver, B., Merlevede, A., Roncoroni, M., Voordeckers, K., Miraglia, L., 2016. Domestication and divergence of Saccharomyces cerevisiae beer yeasts. Cell 166(6), 1397-1410. e1316.

Ganesh, I., Ravikumar, S., Lee, S.H., Park, S.J., Hong, S.H., 2013. Engineered fumarate sensing Escherichia coli based on novel chimeric two-component system. J Biotechnol 168(4), 560-566.

Ganesh, I., Ravikumar, S., Yoo, I.K., Hong, S.H., 2015. Construction of malate-sensing Escherichia coli by introduction of a novel chimeric two-component system. Bioprocess Biosyst Eng 38(4), 797-804.

Gao, C., Hou, J., Xu, P., Guo, L., Chen, X., Hu, G., Ye, C., Edwards, H., Chen, J., Chen, W., Liu, L., 2019. Programmable biomolecular switches for rewiring flux in Escherichia coli. Nat Commun 10(1), 3751.

Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., 2014. Genome-scale CRISPR-mediated control of gene repression and activation. Cell 159(3), 647-661.

Greco, F.V., Pandi, A., Erb, T.J., Grierson, C.S., Gorochowski, T.E., 2021. Harnessing the central dogma for stringent multi-level control of gene expression. Nature Communications 12(1), 1738.

Guil, S., Esteller, M., 2015. RNA-RNA interactions in gene regulation: the coding and noncoding players. Trends Biochem Sci 40(5), 248-256.

Guo, L., Lu, J., Gao, C., Zhang, L., Liu, L., Chen, X., 2021. Dynamic control of the distribution of carbon flux between cell growth and butyrate biosynthesis in Escherichia coli. Applied Microbiology and Biotechnology 105(12), 5173-5187.

Gupta, A., Reizman, I.M., Reisch, C.R., Prather, K.L., 2017. Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. Nat Biotechnol 35(3), 273-279.

Hammer, S.K., Avalos, J.L., 2017. Harnessing yeast organelles for metabolic engineering. Nature Chemical Biology 13(8), 823-832.

Hanko, E.K., Minton, N.P., Malys, N., 2017. Characterisation of a 3-hydroxypropionic acid-inducible system from Pseudomonas putida for orthogonal gene expression control in Escherichia coli and Cupriavidus necator. Scientific reports 7(1), 1-13.

Hawkins, J.S., Silvis, M.R., Koo, B.-M., Peters, J.M., Osadnik, H., Jost, M., Hearne, C.C., Weissman, J.S., Todor, H., Gross, C.A., 2020. Mismatch-CRISPRi reveals the co-varying expression-fitness relationships of essential genes in Escherichia coli and Bacillus subtilis. Cell systems 11(5), 523-535. e529.

Hawkins, J.S., Wong, S., Peters, J.M., Almeida, R., Qi, L.S., 2015. Targeted transcriptional repression in bacteria using CRISPR interference (CRISPRi), CRISPR. Springer, pp. 349-362.

Holmqvist, E., Wright, P.R., Li, L., Bischler, T., Barquist, L., Reinhardt, R., Backofen, R., Vogel, J., 2016. Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo. EMBO J 35(9), 991-1011.

Hossain, G.S., Saini, M., Miyake, R., Ling, H., Chang, M.W., 2020. Genetic Biosensor Design for Natural Product Biosynthesis in Microorganisms. Trends in Biotechnology 38(7), 797-810.

Huang, Q., Lin, Y., Yan, Y., 2013. Caffeic acid production enhancement by engineering a phenylalanine over-producing Escherichia coli strain. Biotechnol Bioeng 110(12), 3188-3196.

Huo, Y.-X., Cho, K.M., Rivera, J.G.L., Monte, E., Shen, C.R., Yan, Y., Liao, J.C., 2011. Conversion of proteins into biofuels by engineering nitrogen flux. Nature Biotechnology 29(4), 346-351.

Inda, M.E., Lu, T.K., 2020. Microbes as Biosensors. Annual Review of Microbiology 74(1), 337-359.

Jang, S., Jang, S., Xiu, Y., Kang, T.J., Lee, S.-H., Koffas, M.A.G., Jung, G.Y., 2017. Development of Artificial Riboswitches for Monitoring of Naringenin In Vivo. ACS Synthetic Biology 6(11), 2077-2085.

Jang, Y.-S., Im, J.A., Choi, S.Y., Im Lee, J., Lee, S.Y., 2014. Metabolic engineering of Clostridium acetobutylicum for butyric acid production with high butyric acid selectivity. Metabolic engineering 23, 165-174.

Jeschek, M., Gerngross, D., Panke, S., 2016. Rationally reduced libraries for combinatorial pathway optimization minimizing experimental effort. Nature Communications 7(1), 11163.

Jiang, T., Li, C., Teng, Y., Zhang, R., Yan, Y., 2020. Recent advances in improving metabolic robustness of microbial cell factories. Current Opinion in Biotechnology 66, 69-77.

Jiang, T., Li, C., Yan, Y., 2021. Optimization of a p-Coumaric Acid Biosensor System for Versatile Dynamic Performance. ACS Synthetic Biology 10(1), 132-144.

Jones, J.A., Toparlak, O.D., Koffas, M.A., 2015. Metabolic pathway balancing and its role in the production of biofuels and chemicals. Curr Opin Biotechnol 33, 52-59.

Jones, J.A., Vernacchio, V.R., Lachance, D.M., Lebovich, M., Fu, L., Shirke, A.N., Schultz, V.L., Cress, B., Linhardt, R.J., Koffas, M.A.G., 2015. ePathOptimize: A Combinatorial Approach for Transcriptional Balancing of Metabolic Pathways. Scientific Reports 5(1), 11301.

Jones, K.L., Kim, S.-W., Keasling, J., 2000. Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria. Metabolic engineering 2(4), 328-338.

Jost, M., Santos, D.A., Saunders, R.A., Horlbeck, M.A., Hawkins, J.S., Scaria, S.M., Norman, T.M., Hussmann, J.A., Liem, C.R., Gross, C.A., 2020. Titrating gene expression using libraries of systematically attenuated CRISPR guide RNAs. Nature biotechnology 38(3), 355-364.

Kaern, M., Elston, T.C., Blake, W.J., Collins, J.J., 2005. Stochasticity in gene expression: from theories to phenotyoes. Nature Reviews Genetics 6(6), 451-464.

Kang, C.W., Lim, H.G., Yang, J., Noh, M.H., Seo, S.W., Jung, G.Y., 2018. Synthetic auxotrophs for stable and tunable maintenance of plasmid copy number. Metabolic Engineering 48, 121-128.

Kang, H.-S., Charlop-Powers, Z., Brady, S.F., 2016. Multiplexed CRISPR/Cas9- and TAR-Mediated Promoter Engineering of Natural Product Biosynthetic Gene Clusters in Yeast. ACS Synthetic Biology 5(9), 1002-1010.

Karin, M., 1990. Too many transcription factors: positive and negative interactions. The New Biologist 2(2), 126-131.

Keasling, J.D., 2010. Manufacturing molecules through metabolic engineering. Science 330(6009), 1355-1358.

Khosla, C., Keasling, J.D., 2003. Metabolic engineering for drug discovery and development. Nature Reviews Drug Discovery 2(12), 1019-1025.

Kiattisewee, C., Dong, C., Fontana, J., Sugianto, W., Peralta-Yahya, P., Carothers, J.M., Zalatan, J.G., 2021. Portable bacterial CRISPR transcriptional activation enables metabolic engineering in Pseudomonas putida. Metabolic Engineering 66, 283-295.

Kim, E.M., Woo, H.M., Tian, T., Yilmaz, S., Javidpour, P., Keasling, J.D., Lee, T.S., 2017. Autonomous control of metabolic state by a quorum sensing (QS)-mediated regulator for bisabolene production in engineered E. coli. Metab Eng 44, 325-336.

Ko, Y.-S., Kim, J.W., Lee, J.A., Han, T., Kim, G.B., Park, J.E., Lee, S.Y., 2020. Tools and strategies of systems metabolic engineering for the development of microbial cell factories for chemical production. Chemical Society Reviews 49(14), 4615-4636.

Kocak, D.D., Josephs, E.A., Bhandarkar, V., Adkar, S.S., Kwon, J.B., Gersbach, C.A., 2019. Increasing the specificity of CRISPR systems with engineered RNA secondary structures. Nature biotechnology 37(6), 657-666.

Koch, M., Pandi, A., Borkowski, O., Batista, A.C., Faulon, J.-L., 2019. Custom-made transcriptional biosensors for metabolic engineering. Current Opinion in Biotechnology 59, 78-84.

Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., 2015. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature 517(7536), 583.

Kunjapur, A.M., Prather, K.L.J., 2019. Development of a Vanillate Biosensor for the Vanillin Biosynthesis Pathway in E. coli. ACS Synth Biol.

Kusumawati, I., Indrayanto, G., 2013. Natural antioxidants in cosmetics, Studies in natural products chemistry. Elsevier, pp. 485-505.

Laity, J.H., Lee, B.M., Wright, P.E., 2001. Zinc finger proteins: new insights into structural and functional diversity. Current Opinion in Structural Biology 11(1), 39-46.

Larson, M.H., Gilbert, L.A., Wang, X., Lim, W.A., Weissman, J.S., Qi, L.S., 2013. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nature Protocols 8(11), 2180-2196.

Lasa, I., Toledo-Arana, A., Dobin, A., Villanueva, M., de los Mozos, I.R., Vergara-Irigaray, M., Segura, V., Fagegaltier, D., Penadés, J.R., Valle, J., 2011. Genome-wide antisense transcription drives mRNA processing in bacteria. Proceedings of the National Academy of Sciences 108(50), 20172-20177.

Latchman, D.S., 1993. Transcription factors: an overview. Int J Exp Pathol 74(5), 417-422.

Lee, C., Kim, J., Shin, S.G., Hwang, S., 2006. Absolute and relative QPCR quantification of plasmid copy number in Escherichia coli. Journal of Biotechnology 123(3), 273-280.

Lee, M.E., Aswani, A., Han, A.S., Tomlin, C.J., Dueber, J.E., 2013. Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. Nucleic Acids Research 41(22), 10668-10678.

Lee, S.-W., Oh, M.-K., 2015. A synthetic suicide riboswitch for the high-throughput screening of metabolite production in Saccharomyces cerevisiae. Metabolic Engineering 28, 143-150.

Lee, S.Y., Kim, H.U., Chae, T.U., Cho, J.S., Kim, J.W., Shin, J.H., Kim, D.I., Ko, Y.-S., Jang, W.D., Jang, Y.-S., 2019. A comprehensive metabolic map for production of bio-based chemicals. Nature Catalysis 2(1), 18-33.

Lee, S.Y., Nielsen, J., Stephanopoulos, G., 2021. Metabolic Engineering: Concepts and Applications. John Wiley & Sons.

Li, C., Jiang, T., Li, M., Zou, Y., Yan, Y., 2021. Fine-tuning gene expression for improved biosynthesis of natural products: From transcriptional to post-translational regulation. Biotechnology Advances, 107853.

Li, C., Zhang, R., Wang, J., Wilson, L.M., Yan, Y., 2020. Protein engineering for improving and diversifying natural product biosynthesis. Trends in biotechnology 38(7), 729-744.

Li, C., Zou, Y., Jiang, T., Zhang, J., Yan, Y., 2022. Harnessing plasmid replication mechanism to enable dynamic control of gene copy in bacteria. Metabolic Engineering 70, 67-78.

Li, S., Prasanna, X., Salo, V.T., Vattulainen, I., Ikonen, E., 2019. An efficient auxin-inducible degron system with low basal degradation in human cells. Nature methods 16(9), 866-869.

Lian, J., HamediRad, M., Hu, S., Zhao, H., 2017. Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. Nat Commun 8(1), 1688.

Lian, J., Schultz, C., Cao, M., HamediRad, M., Zhao, H., 2019. Multi-functional genome-wide CRISPR system for high throughput genotype–phenotype mapping. Nature Communications 10(1), 5794.

Liang, C., Zhang, X., Wu, J., Mu, S., Wu, Z., Jin, J.-M., Tang, S.-Y., 2020. Dynamic control of toxic natural product biosynthesis by an artificial regulatory circuit. Metabolic engineering 57, 239-246.

Libis, V., Delepine, B., Faulon, J.L., 2016. Sensing new chemicals with bacterial transcription factors. Curr Opin Microbiol 33, 105-112.

Lim Chin, G., Fowler Zachary, L., Hueller, T., Schaffer, S., Koffas Mattheos, A.G., 2011. High-Yield Resveratrol Production in Engineered Escherichia coli. Applied and Environmental Microbiology 77(10), 3451-3460.

Lin, J.-L., Wagner, J.M., Alper, H.S., 2017. Enabling tools for high-throughput detection of metabolites: Metabolic engineering and directed evolution applications. Biotechnology Advances 35(8), 950-970.

Lin, Y., Shen, X., Yuan, Q., Yan, Y., 2013a. Microbial biosynthesis of the anticoagulant precursor 4-hydroxycoumarin. Nat Commun 4, 2603.

Lin, Y., Sun, X., Yuan, Q., Yan, Y., 2013b. Combinatorial biosynthesis of plant-specific coumarins in bacteria. Metab Eng 18, 69-77.

Liu, C., Zhang, B., Liu, Y.M., Yang, K.Q., Liu, S.J., 2018. New Intracellular Shikimic Acid Biosensor for Monitoring Shikimate Synthesis in Corynebacterium glutamicum. ACS Synth Biol 7(2), 591-601.

Liu, C.L., Cai, J.Y., Bi, H.R., Tan, T.W., 2018. A novel DMAPP-responding genetic circuit sensor for high-throughput screening and evolving isoprene synthase. Appl Microbiol Biotechnol 102(3), 1381-1391.

Liu, D., Mannan, A.A., Han, Y., Oyarzun, D.A., Zhang, F., 2018. Dynamic metabolic control: towards precision engineering of metabolism. J Ind Microbiol Biotechnol 45(7), 535-543.

Liu, D., Xiao, Y., Evans, B.S., Zhang, F., 2015. Negative feedback regulation of fatty acid production based on a malonyl-CoA sensor-actuator. ACS Synth Biol 4(2), 132-140.

Liu, J., Liu, M., Shi, T., Sun, G., Gao, N., Zhao, X., Guo, X., Ni, X., Yuan, Q., Feng, J., 2022. CRISPR-assisted rational flux-tuning and arrayed CRISPRi screening of an l-proline exporter for l-proline hyperproduction. Nature Communications 13(1), 1-16.

Liu, Q., Yu, T., Li, X., Chen, Y., Campbell, K., Nielsen, J., Chen, Y., 2019. Rewiring carbon metabolism in yeast for high level production of aromatic chemicals. Nature Communications 10(1), 1-13.

Liu, X., Cheng, J., Zhang, G., Ding, W., Duan, L., Yang, J., Kui, L., Cheng, X., Ruan, J., Fan, W., Chen, J., Long, G., Zhao, Y., Cai, J., Wang, W., Ma, Y., Dong, Y., Yang, S., Jiang, H., 2018. Engineering yeast for the production of breviscapine by genomic analysis and synthetic biology approaches. Nature Communications 9(1).

Liu, Y., Nielsen, J., 2019. Recent trends in metabolic engineering of microbial chemical factories. Current Opinion in Biotechnology 60, 188-197.

Liu, Y., Wan, X., Wang, B., 2019. Engineered CRISPRa enables programmable eukaryote-like gene activation in bacteria. Nature Communications 10(1), 3693.

Lopez, S.C., Crawford, K.D., Lear, S.K., Bhattarai-Kline, S., Shipman, S.L., 2022. Precise genome editing across kingdoms of life using retron-derived DNA. Nature chemical biology 18(2), 199-206.

Lu, Z., Peng, B., Ebert, B.E., Dumsday, G., Vickers, C.E., 2021. Auxin-mediated protein depletion for metabolic engineering in terpene-producing yeast. Nature Communications 12(1), 1051.

Lutz, R., 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 Research 25(6), 1203-1210.

Mahr, R., Frunzke, J., 2016. Transcription factor-based biosensors in biotechnology: current state and future prospects. Appl Microbiol Biotechnol 100(1), 79-90.

Mandal, M., Boese, B., Barrick, J.E., Winkler, W.C., Breaker, R.R., 2003. Riboswitches Control Fundamental Biochemical Pathways in Bacillus subtilis and Other Bacteria. Cell 113(5), 577-586.

Martínez, V., Lauritsen, I., Hobel, T., Li, S., Nielsen, A.T., Nørholm, Morten H.H., 2017. CRISPR/Cas9-based genome editing for simultaneous interference with gene expression and protein stability. Nucleic Acids Research 45(20), e171-e171.

Mitchler, M.M., Garcia, J.M., Montero, N.E., Williams, G.J., 2021. Transcription factor-based biosensors: a molecular-guided approach for natural product engineering. Current Opinion in Biotechnology 69, 172-181.

Miyahisa, I., Funa, N., Ohnishi, Y., Martens, S., Moriguchi, T., Horinouchi, S., 2006. Combinatorial biosynthesis of flavones and flavonols in Escherichia coli. Appl Microbiol Biotechnol 71(1), 53-58.

Mizuno, T., Chou, M.Y., Inouye, M., 1984. A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). Proceedings of the National Academy of Sciences 81(7), 1966-1970.

Montaño López, J., Duran, L., Avalos, J.L., 2022. Physiological limitations and opportunities in microbial metabolic engineering. Nature Reviews Microbiology 20(1), 35-48.

Moradali, M.F., Rehm, B.H.A., 2020. Bacterial biopolymers: from pathogenesis to advanced materials. Nature Reviews Microbiology 18(4), 195-210.

Muranaka, N., Abe, K., Yokobayashi, Y., 2009. Mechanism-guided library design and dual genetic selection of synthetic OFF riboswitches. Chembiochem 10(14), 2375-2381.

Na, D., Yoo, S.M., Chung, H., Park, H., Park, J.H., Lee, S.Y., 2013. Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs. Nat Biotechnol 31(2), 170-174.

Naseri, G., Behrend, J., Rieper, L., Mueller-Roeber, B., 2019. COMPASS for rapid combinatorial optimization of biochemical pathways based on artificial transcription factors. Nature Communications 10(1), 2615.

Nguyen Thi Kim, C., Tran Ngoc, P., Cavin, J.-F., 2011. Genetic and Biochemical Analysis of PadR-padC Promoter Interactions during the Phenolic Acid Stress Response in Bacillus subtilis 168. Journal of Bacteriology 193(16), 4180-4191.

Ni, J., Zhang, G., Qin, L., Li, J., Li, C., 2019. Simultaneously down-regulation of multiplex branch pathways using CRISPRi and fermentation optimization for enhancing β-amyrin production in Saccharomyces cerevisiae. Synthetic and Systems Biotechnology 4(2), 79-85.

Nielsen, J., 2013. Production of biopharmaceutical proteins by yeast. Bioengineered 4(4), 207-211.

Nielsen, J., Keasling, J.D., 2011. Synergies between synthetic biology and metabolic engineering. Nature biotechnology 29(8), 693-695.

Nielsen, J., Keasling, J.D., 2016. Engineering Cellular Metabolism. Cell 164(6), 1185-1197.

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., Kanemaki, M., 2009. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nature methods 6(12), 917-922.

Novichkov, P.S., Kazakov, A.E., Ravcheev, D.A., Leyn, S.A., Kovaleva, G.Y., Sutormin, R.A., Kazanov, M.D., Riehl, W., Arkin, A.P., Dubchak, I., Rodionov, D.A., 2013. RegPrecise 3.0 – A resource for genome-scale exploration of transcriptional regulation in bacteria. BMC Genomics 14(1), 745.

Nyerges, Á., Csörgő, B., Nagy, I., Bálint, B., Bihari, P., Lázár, V., Apjok, G., Umenhoffer, K., Bogos, B., Pósfai, G., 2016. A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species. Proceedings of the National Academy of Sciences 113(9), 2502-2507.

Palmer, C.M., Miller, K.K., Nguyen, A., Alper, H.S., 2019. Engineering 4-coumaroyl-CoA derived polyketide production in Yarrowia lipolytica through a β-oxidation mediated strategy. Metabolic Engineering.

Park, J.H., Lee, K.H., Kim, T.Y., Lee, S.Y., 2007. Metabolic engineering of Escherichia coli for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. Proceedings of the national academy of sciences 104(19), 7797-7802.

Park, J.M., Kim, T.Y., Lee, S.Y., 2009. Constraints-based genome-scale metabolic simulation for systems metabolic engineering. Biotechnology advances 27(6), 979-988.

Park, S.C., Kwak, Y.M., Song, W.S., Hong, M., Yoon, S.-i., 2017. Structural basis of effector and operator recognition by the phenolic acid-responsive transcriptional regulator PadR. Nucleic Acids Research 45(22), 13080-13093.

Park, S.Y., Yang, D., Ha, S.H., Lee, S.Y., 2018. Metabolic Engineering of Microorganisms for the Production of Natural Compounds. Advanced Biosystems 2(1).

Pelechano, V., Steinmetz, L.M., 2013. Gene regulation by antisense transcription. Nature Reviews Genetics 14(12), 880-893.

Peng, B., Nielsen, L.K., Kampranis, S.C., Vickers, C.E., 2018. Engineered protein degradation of farnesyl pyrophosphate synthase is an effective regulatory mechanism to increase monoterpene production in Saccharomyces cerevisiae. Metabolic Engineering 47(June 2017), 83-93.

Peng, B., Plan, M.R., Chrysanthopoulos, P., Hodson, M.P., Nielsen, L.K., Vickers, C.E., 2017. A squalene synthase protein degradation method for improved sesquiterpene production in Saccharomyces cerevisiae. Metab Eng 39, 209-219.

Peng, R., Wang, Y., Feng, W.-w., Yue, X.-j., Chen, J.-h., Hu, X.-z., Li, Z.-f., Sheng, D.-h., Zhang, Y.-m., Li, Y.-z., 2018. CRISPR/dCas9-mediated transcriptional improvement of the biosynthetic gene cluster for the epothilone production in Myxococcus xanthus. Microbial Cell Factories 17(1), 15.

Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., 2013. RNA-guided gene activation by CRISPR-Cas9–based transcription factors. Nature methods 10(10), 973.

Pfeiffer, F., Mayer, G., 2016. Selection and biosensor application of aptamers for small molecules. Frontiers in chemistry 4, 25.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., Lim, W.A., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152(5), 1173-1183.

Rhie, M.N., Kim, H.T., Jo, S.Y., Chu, L.L., Baritugo, K.-A., Baylon, M.G., Lee, J., Na, J.-G., Kim, L.H., Kim, T.W., Park, C., Hong, S.H., Joo, J.C., Park, S.J., 2019. Recent Advances in the Metabolic Engineering of Klebsiella pneumoniae: A Potential Platform Microorganism for Biorefineries. Biotechnology and Bioprocess Engineering 24(1), 48-64.

Rodrigues, T., Reker, D., Schneider, P., Schneider, G., 2016. Counting on natural products for drug design. Nat Chem 8(6), 531-541.

Rogers, J.K., Church, G.M., 2016. Genetically encoded sensors enable real-time observation of metabolite production. Proc Natl Acad Sci U S A 113(9), 2388-2393.

Rogers, J.K., Taylor, N.D., Church, G.M., 2016. Biosensor-based engineering of biosynthetic pathways. Curr Opin Biotechnol 42, 84-91.

Rousset, F., Cui, L., Siouve, E., Becavin, C., Depardieu, F., Bikard, D., 2018. Genome-wide CRISPR-dCas9 screens in E. coli identify essential genes and phage host factors. PLoS genetics 14(11), e1007749.

Rugbjerg, P., Myling-Petersen, N., Porse, A., Sarup-Lytzen, K., Sommer, M.O.A., 2018. Diverse genetic error modes constrain large-scale bio-based production. Nat Commun 9(1), 787.

Ruscito, A., DeRosa, M.C., 2016. Small-molecule binding aptamers: Selection strategies, characterization, and applications. Frontiers in chemistry 4, 14.

Saini, M., Wang, Z.W., Chiang, C.-J., Chao, Y.-P., 2014. Metabolic engineering of Escherichia coli for production of butyric acid. Journal of agricultural and food chemistry 62(19), 4342-4348.

Salehi, B., Fokou, P.V., Sharifi-Rad, M., Zucca, P., Pezzani, R., Martins, N., Sharifi-Rad, J., 2019. The Therapeutic Potential of Naringenin: A Review of Clinical Trials. Pharmaceuticals 12(1).

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular cloning. Cold spring harbor laboratory press New York.

Sandberg, T.E., Salazar, M.J., Weng, L.L., Palsson, B.O., Feist, A.M., 2019. The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. Metabolic Engineering 56, 1-16.

Sander, T., Wang, C.Y., Glatter, T., Link, H., 2019. CRISPRi-based downregulation of transcriptional feedback improves growth and metabolism of arginine overproducing E. coli. ACS synthetic biology 8(9), 1983-1990.

Schallmey, M., Frunzke, J., Eggeling, L., Marienhagen, J., 2014. Looking for the pick of the bunch: high-throughput screening of producing microorganisms with biosensors. Curr Opin Biotechnol 26, 148-154.

Schubert, M.G., Goodman, D.B., Wannier, T.M., Kaur, D., Farzadfard, F., Lu, T.K., Shipman, S.L., Church, G.M., 2021. High-throughput functional variant screens via in vivo production of single-stranded DNA. Proceedings of the National Academy of Sciences 118(18).

Schultenkämper, K., Brito, L.F., Wendisch, V.F., 2020. Impact of CRISPR interference on strain development in biotechnology. Biotechnology and Applied Biochemistry 67(1), 7-21.

Sekar, K., Gentile, A.M., Bostick, J.W., Tyo, K.E., 2016. N-terminal-based targeted, inducible protein degradation in Escherichia coli. PLoS One 11(2), e0149746.

Selzer, G., Som, T., Itoh, T., Tomizawa, J.i., 1983. The origin of replication of plasmid p15A and comparative studies on the nucleotide sequences around the origin of related plasmids. Cell 32(1), 119-129.

Seok, J.Y., Han, Y.H., Yang, J.-S., Yang, J., Lim, H.G., Kim, S.G., Seo, S.W., Jung, G.Y., 2021. Synthetic biosensor accelerates evolution by rewiring carbon metabolism toward a specific metabolite. Cell Reports 36(8), 109589.

Shen, C.R., Liao, J.C., 2008. Metabolic engineering of Escherichia coli for 1-butanol and 1-propanol production via the keto-acid pathways. Metabolic engineering 10(6), 312-320.

Shen, X., Wang, J., Li, C., Yuan, Q., Yan, Y., 2019. Dynamic gene expression engineering as a tool in pathway engineering. Curr Opin Biotechnol 59, 122-129.

Shlomi, T., Eisenberg, Y., Sharan, R., Ruppin, E., 2007. A genome-scale computational study of the interplay between transcriptional regulation and metabolism. Molecular Systems Biology 3(1), 101.

Siedler, S., Khatri, N.K., Zsohár, A., Kjærbølling, I., Vogt, M., Hammar, P., Nielsen, C.F., Marienhagen, J., Sommer, M.O.A., Joensson, H.N., 2017. Development of a Bacterial Biosensor for Rapid Screening of Yeast p-Coumaric Acid Production. ACS Synthetic Biology 6(10), 1860-1869.

Siedler, S., Stahlhut, S.G., Malla, S., Maury, J., Neves, A.R., 2014. Novel biosensors based on flavonoid-responsive transcriptional regulators introduced into Escherichia coli. Metab Eng 21, 2-8.

Silva, F., Queiroz, J.A., Domingues, F.C., 2012. Evaluating metabolic stress and plasmid stability in plasmid DNA production by Escherichia coli. Biotechnol Adv 30(3), 691-708.

Simon, A.J., Ellington, A.D., Finkelstein, I.J., 2019. Retrons and their applications in genome engineering. Nucleic acids research 47(21), 11007-11019.

Siu, K.-H., Chen, W., 2019. Riboregulated toehold-gated gRNA for programmable CRISPR–Cas9 function. Nature chemical biology 15(3), 217-220.

Skjoedt, M.L., Snoek, T., Kildegaard, K.R., Arsovska, D., Eichenberger, M., Goedecke, T.J., Rajkumar, A.S., Zhang, J., Kristensen, M., Lehka, B.J., 2016. Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast. Nature chemical biology 12(11), 951.

Skulj, M., Okrslar, V., Jalen, S., Jevsevar, S., Slanc, P., Strukelj, B., Menart, V., 2008. Improved determination of plasmid copy number using quantitative real-time PCR for monitoring fermentation processes. Microb Cell Fact 7, 6.

Sohn, Y.J., Kim, H.T., Jo, S.Y., Song, H.M., Baritugo, K.-A., Pyo, J., Choi, J.-i., Joo, J.C., Park, S.J., 2020. Recent Advances in Systems Metabolic Engineering Strategies for the Production of Biopolymers. Biotechnology and Bioprocess Engineering 25(6), 848-861.

Soma, Y., Fujiwara, Y., Nakagawa, T., Tsuruno, K., Hanai, T., 2017. Reconstruction of a metabolic regulatory network in Escherichia coli for purposeful switching from cell growth mode to production mode in direct GABA fermentation from glucose. Metabolic Engineering 43(July), 54-63.

Steensels, J., Gallone, B., Voordeckers, K., Verstrepen, K.J., 2019. Domestication of industrial microbes. Current biology 29(10), R381-R393.

Sternberg, S.H., LaFrance, B., Kaplan, M., Doudna, J.A., 2015. Conformational control of DNA target cleavage by CRISPR–Cas9. Nature 527(7576), 110-113.

Stuitje, A.R., Veltkamp, E., Nijkamp, H.J.J., Weijers, P.J., 1979. Origin and direction of replication of the bacteriocinogenic plasmid Clo DF13. Nucleic Acids Research 6(1), 71-80.

Sun, H., Zhao, H., Ang, E.L., 2020. A New Biosensor for Stilbenes and a Cannabinoid Enabled by Genome Mining of a Transcriptional Regulator. ACS Synthetic Biology 9(4), 698-705.

Svensson, S.L., Sharma, C.M., 2016. Small RNAs in bacterial virulence and communication. Virulence Mechanisms of Bacterial Pathogens, 169-212.

Tan, S.Z., Prather, K.L., 2017. Dynamic pathway regulation: recent advances and methods of construction. Curr Opin Chem Biol 41, 28-35.

Tang, S.-Y., Fazelinia, H., Cirino, P.C., 2008. AraC regulatory protein mutants with altered effector specificity. Journal of the American Chemical Society 130(15), 5267-5271.

Tang, S.Y., Cirino, P.C., 2011. Design and application of a mevalonate-responsive regulatory protein. Angew Chem Int Ed Engl 50(5), 1084-1086.

Tang, S.Y., Qian, S., Akinterinwa, O., Frei, C.S., Gredell, J.A., Cirino, P.C., 2013. Screening for enhanced triacetic acid lactone production by recombinant Escherichia coli expressing a designed triacetic acid lactone reporter. J Am Chem Soc 135(27), 10099-10103.

Tang, W., Hu, J.H., Liu, D.R., 2017. Aptazyme-embedded guide RNAs enable ligand-responsive genome editing and transcriptional activation. Nature communications 8(1), 1-8.

Taylor, N.D., Garruss, A.S., Moretti, R., Chan, S., Arbing, M.A., Cascio, D., Rogers, J.K., Isaacs, F.J., Kosuri, S., Baker, D., Fields, S., Church, G.M., Raman, S., 2016. Engineering an allosteric transcription factor to respond to new ligands. Nat Methods 13(2), 177-183.

Thomason, L.C., Costantino, N., Court, D.L., 2007. E. coli genome manipulation by P1 transduction. Current protocols in molecular biology, 1.17. 11-11.17. 18.

Trantas, E.A., Koffas, M.A., Xu, P., Ververidis, F., 2015. When plants produce not enough or at all: metabolic engineering of flavonoids in microbial hosts. Frontiers in plant science 6, 7.

Tsui, H.-C.T., Leung, H.-C.E., Winkler, M.E., 1994. Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in Escherichia coli K-12. Molecular Microbiology 13(1), 35-49.

Ververidis, F., Trantas, E., Douglas, C., Vollmer, G., Kretzschmar, G., Panopoulos, N., 2007. Biotechnology of flavonoids and other phenylpropanoid-derived natural products. Part I: Chemical diversity, impacts on plant biology and human health. Biotechnol J 2(10), 1214-1234.

Vogel, J., Wagner, E.G.H., 2007. Target identification of small noncoding RNAs in bacteria. Current Opinion in Microbiology 10(3), 262-270.

Wang, H.H., Isaacs, F.J., Carr, P.A., Sun, Z.Z., Xu, G., Forest, C.R., Church, G.M., 2009. Programming cells by multiplex genome engineering and accelerated evolution. Nature 460(7257), 894-898.

Wang, J., Cui, X., Yang, L., Zhang, Z., Lv, L., Wang, H., Zhao, Z., Guan, N., Dong, L., Chen, R., 2017. A real-time control system of gene expression using ligand-bound nucleic acid aptamer for metabolic engineering. Metabolic Engineering 42, 85-97.

Wang, J., Gao, D., Yu, X., Li, W., Qi, Q., 2015. Evolution of a chimeric aspartate kinase for L-lysine production using a synthetic RNA device. Applied microbiology and biotechnology 99(20), 8527-8536.

Wang, J., Guleria, S., Koffas, M.A., Yan, Y., 2016. Microbial production of value-added nutraceuticals. Current opinion in biotechnology 37, 97-104.

Wang, J., Li, C., Zou, Y., Yan, Y., 2020. Bacterial synthesis of C3-C5 diols via extending amino acid catabolism. Proceedings of the National Academy of Sciences 117(32), 19159.

Wang, J., Mahajani, M., Jackson, S.L., Yang, Y., Chen, M., Ferreira, E.M., Lin, Y., Yan, Y., 2017a. Engineering a bacterial platform for total biosynthesis of caffeic acid derived phenethyl esters and amides. Metab Eng 44, 89-99.

Wang, J., Shen, X., Rey, J., Yuan, Q., Yan, Y., 2018. Recent advances in microbial production of aromatic natural products and their derivatives. Appl Microbiol Biotechnol 102(1), 47-61.

Wang, J., Teng, Y., Zhang, R., Wu, Y., Lou, L., Zou, Y., Li, M., Xie, Z.-R., Yan, Y., 2021a. Engineering a PAM-flexible SpdCas9 variant as a universal gene repressor. Nature communications 12(1), 1-10.

Wang, J., Wu, Y., Sun, X., Yuan, Q., Yan, Y., 2017b. De Novo Biosynthesis of Glutarate via alpha-Keto Acid Carbon Chain Extension and Decarboxylation Pathway in Escherichia coli. ACS Synth Biol 6(10), 1922-1930.

Wang, J., Yang, Y., Zhang, R., Shen, X., Chen, Z., Wang, J., Yuan, Q., Yan, Y., 2018. Microbial production of branched-chain dicarboxylate 2-methylsuccinic acid via enoate reductase-mediated bioreduction. Metabolic engineering 45, 1-10.

Wang, J., Zhang, R., Zhang, J., Gong, X., Jiang, T., Sun, X., Shen, X., Wang, J., Yuan, Q., Yan, Y., 2021b. Tunable hybrid carbon metabolism coordination for the carbon-efficient biosynthesis of 1,3-butanediol in Escherichia coli. Green Chemistry 23(21), 8694-8706.

Wang, R., Cress, B.F., Yang, Z., Hordines, J.C., Zhao, S., Jung, G.Y., Wang, Z., Koffas, M.A.G., 2019. Design and Characterization of Biosensors for the Screening of Modular Assembled Naringenin Biosynthetic Library in Saccharomyces cerevisiae. ACS Synthetic Biology 8(9), 2121-2130.

Wang, R., Zhao, S., Wang, Z., Koffas, M.A.G., 2020. Recent advances in modular co-culture engineering for synthesis of natural products. Current Opinion in Biotechnology 62, 65-71.

Wang, T., Guan, C., Guo, J., Liu, B., Wu, Y., Xie, Z., Zhang, C., Xing, X.-H., 2018. Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance. Nature communications 9(1), 1-15.

Wang, T., Zheng, X., Ji, H., Wang, T.-L., Xing, X.-H., Zhang, C., 2019. Dynamics of transcription–translation coordination tune bacterial indole signaling. Nature Chemical Biology 2019, 1-10.

Wang, W., Li, S., Li, Z., Zhang, J., Fan, K., Tan, G., Ai, G., Lam, S.M., Shui, G., Yang, Z., Lu, H., Jin, P., Li, Y., Chen, X., Xia, X., Liu, X., Dannelly, H.K., Yang, C., Yang, Y., Zhang, S.,

Alterovitz, G., Xiang, W., Zhang, L., 2020. Harnessing the intracellular triacylglycerols for titer improvement of polyketides in Streptomyces. Nature Biotechnology 38(1), 76-83.

Wang, Z., Cirino, P.C., 2016. New and improved tools and methods for enhanced biosynthesis of natural products in microorganisms. Current Opinion in Biotechnology 42, 159-168.

Weissman, K.J., 2016. Genetic engineering of modular PKSs: from combinatorial biosynthesis to synthetic biology. Nat. Prod. Rep. 33(2), 203-230.

Westbrook, A.M., Lucks, J.B., 2017. Achieving large dynamic range control of gene expression with a compact RNA transcription–translation regulator. Nucleic Acids Research 45(9), 5614-5624.

Wu, G., Yan, Q., Jones, J.A., Tang, Y.J., Fong, S.S., Koffas, M.A., 2016. Metabolic burden: cornerstones in synthetic biology and metabolic engineering applications. Trends in biotechnology 34(8), 652-664.

Wu, J., Du, G., Chen, J., Zhou, J., 2015. Enhancing flavonoid production by systematically tuning the central metabolic pathways based on a CRISPR interference system in Escherichia coli. Scientific reports 5(1), 1-14.

Wu, J., Yu, O., Du, G., Zhou, J., Chen, J., 2014. Fine-Tuning of the Fatty Acid Pathway by Synthetic Antisense RNA for Enhanced (2S)-Naringenin Production from l-Tyrosine in Escherichia coli. Appl Environ Microbiol 80(23), 7283-7292.

Wu, Y., Chen, T., Liu, Y., Tian, R., Lv, X., Li, J., Du, G., Chen, J., Ledesma-Amaro, R., Liu, L., 2020. Design of a programmable biosensor-CRISPRi genetic circuits for dynamic and autonomous dual-control of metabolic flux in Bacillus subtilis. Nucleic Acids Research 48(2), 996-1009.

Xiong, D., Lu, S., Wu, J., Liang, C., Wang, W., Wang, W., Jin, J.M., Tang, S.Y., 2017. Improving key enzyme activity in phenylpropanoid pathway with a designed biosensor. Metab Eng 40, 115-123.

Xiu, Y., Jang, S., Jones, J.A., Zill, N.A., Linhardt, R.J., Yuan, Q., Jung, G.Y., Koffas, M.A.G., 2017. Naringenin-responsive riboswitch-based fluorescent biosensor module for Escherichia coli co-cultures. Biotechnology and Bioengineering 114(10), 2235-2244.

Xu, P., Li, L., Zhang, F., Stephanopoulos, G., Koffas, M., 2014. Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. Proc Natl Acad Sci U S A 111(31), 11299-11304.

Yang, D., Park, S.Y., Park, Y.S., Eun, H., Lee, S.Y., 2020. Metabolic Engineering of Escherichia coli for Natural Product Biosynthesis. Trends in Biotechnology 38(7), 745-765.

Yang, D., Yoo, S.M., Gu, C., Ryu, J.Y., Lee, J.E., Lee, S.Y., 2019. Expanded synthetic small regulatory RNA expression platforms for rapid and multiplex gene expression knockdown. Metabolic Engineering 54, 180-190.

Yang, J., Seo, S.W., Jang, S., Shin, S.-I., Lim, C.H., Roh, T.-Y., Jung, G.Y., 2013. Synthetic RNA devices to expedite the evolution of metabolite-producing microbes. Nature Communications 4(1), 1413.

Yang, X., Liu, J., Zhang, J., Shen, Y., Qi, Q., Bao, X., Hou, J., 2021. Quorum sensing-mediated protein degradation for dynamic metabolic pathway control in Saccharomyces cerevisiae. Metabolic Engineering 64, 85-94.

Yang, Y., Lin, Y., Li, L., Linhardt, R.J., Yan, Y., 2015. Regulating malonyl-CoA metabolism via synthetic antisense RNAs for enhanced biosynthesis of natural products. Metab Eng 29, 217-226.

Yang, Y., Lin, Y., Wang, J., Wu, Y., Zhang, R., Cheng, M., Shen, X., Wang, J., Chen, Z., Li, C., 2018. Sensor-regulator and RNAi based bifunctional dynamic control network for engineered microbial synthesis. Nature communications 9(1), 1-10.

Yanisch-Perron, C., Vieira, J., Messing, J., Chambers, S., Prior, S., Barstow, D., Minton, N., Gilbert, W., Messing, J., Messing, J., 1985. Improved M13 phage cloning vectors and host strains: nucleotide. Gene 33(1), 103-119.

Yoo, S.M., Na, D., Lee, S.Y., 2013. Design and use of synthetic regulatory small RNAs to control gene expression in Escherichia coli. Nature Protocols 8(9), 1694-1707.

Yuan, J., Ching, C.-B., 2015. Dynamic control of ERG9 expression for improved amorpha-4,11-diene production in Saccharomyces cerevisiae. Microbial Cell Factories 14(1), 38.

Zha, J., Wu, X., Koffas, M.A.G., 2020. Making brilliant colors by microorganisms. Current Opinion in Biotechnology 61, 135-141.

Zha, W., Rubin-Pitel, S.B., Shao, Z., Zhao, H., 2009. Improving cellular malonyl-CoA level in Escherichia coli via metabolic engineering. Metabolic engineering 11(3), 192-198.

Zhang, F., Carothers, J.M., Keasling, J.D., 2012. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. Nat Biotechnol 30(4), 354-359.

Zhang, F., Keasling, J., 2011. Biosensors and their applications in microbial metabolic engineering. Trends Microbiol 19(7), 323-329.

Zhang, J., Jensen, M.K., Keasling, J.D., 2015. Development of biosensors and their application in metabolic engineering. Curr Opin Chem Biol 28, 1-8.

Zhang, J., Pang, Q., Wang, Q., Qi, Q., Wang, Q., 2021. Modular tuning engineering and versatile applications of genetically encoded biosensors. Critical Reviews in Biotechnology, 1-18.

Zhang, R., Li, C., Wang, J., Yang, Y., Yan, Y., 2018. Microbial production of small medicinal molecules and biologics: From nature to synthetic pathways. Biotechnol Adv 36(8), 2219-2231.

Zhang, R., Yang, Y., Wang, J., Lin, Y., Yan, Y., 2019. Synthetic symbiosis combining plasmid displacement enables rapid construction of phenotype-stable strains. Metabolic Engineering 55, 85-91.

Zhang, S., Voigt, C.A., 2018. Engineered dCas9 with reduced toxicity in bacteria: implications for genetic circuit design. Nucleic acids research 46(20), 11115-11125.

Zhao, P., Li, Q., Tian, P., Tan, T., 2021. Switching metabolic flux by engineering tryptophan operon-assisted CRISPR interference system in Klebsiella pneumoniae. Metabolic Engineering.

Zhao, S., Jones, J.A., Lachance, D.M., Bhan, N., Khalidi, O., Venkataraman, S., Wang, Z., Koffas, M.A.G., 2015. Improvement of catechin production in Escherichia coli through combinatorial metabolic engineering. Metabolic Engineering 28, 43-53.

Zhao, X., Shi, F., Zhan, W., 2015. Overexpression of ZWF1 and POS5 improves carotenoid biosynthesis in recombinant Saccharomyces cerevisiae. Letters in Applied Microbiology 61(4), 354-360.

Zhou, L.-B., Zeng, A.-P., 2015. Engineering a Lysine-ON Riboswitch for Metabolic Control of Lysine Production in Corynebacterium glutamicum. ACS Synthetic Biology 4(12), 1335-1340.

Zhou, S., Lyu, Y., Li, H., Koffas, M.A.G., Zhou, J., 2019. Fine-tuning the (2S)-naringenin synthetic pathway using an iterative high-throughput balancing strategy. Biotechnology and Bioengineering 116(6), 1392-1404.

Zhu, L.-y., Qiu, X.-y., Zhu, L.-y., Wu, X.-m., Zhang, Y., Zhu, Q.-h., Fan, D.-y., Zhu, C.-s., Zhang, D.-y., 2016. Spatial organization of heterologous metabolic system in vivo based on TALE. Scientific Reports 6(1), 26065.

Zhu, Q., Jackson, E.N., 2015. Metabolic engineering of Yarrowia lipolytica for industrial applications. Current Opinion in Biotechnology 36, 65-72.

Zou, Y., Li, C., Zhang, R., Jiang, T., Liu, N., Wang, J., Wang, X., Yan, Y., 2021. Exploring the Tunability and Dynamic Properties of MarR-PmarO Sensor System in Escherichia coli. ACS Synthetic Biology 10(8), 2076-2086.

Zu, Y., Prather, K.L., Stephanopoulos, G., 2020. Metabolic engineering strategies to overcome precursor limitations in isoprenoid biosynthesis. Current Opinion in Biotechnology 66, 171-178.

APPENDIX

LIST OF PUBLICATIONS

Peer-reviewed journal articles:

1. **Li, C.,** Zou, Y., Jiang, T., Zhang, J., Yan, Y., 2022. Harnessing plasmid replication mechanism to enable dynamic control of gene copy in bacteria. ***Metabolic Engineering*** 70, 67-78.
2. **Li, C.**, Jiang, T., Li, M., Zou, Y., Yan, Y., 2021. Fine-tuning gene expression for improved biosynthesis of natural products: From transcriptional to post-translational regulation. ***Biotechnology Advances***, 107853.
3. **Li, C.**, Zhang, R., Wang, J., Wilson, L.M., Yan, Y., 2020. Protein engineering for improving and diversifying natural product biosynthesis. ***Trends in biotechnology*** 38(7), 729-744.
4. **Li, C.**, Gao, X., Peng, X., Li, J., Bai, W., Zhong, J., He, M., Xu, K., Wang, Y., Li, C., 2020. Intelligent microbial cell factory with genetic pH shooting (GPS) for cell self-responsive base/acid regulation. ***Microbial cell factories*** 19(1), 1-13.
5. Wang, J., **Li, C.** (co-first author), Jiang, T., Yan, Y., 2022. Biosensor-assisted titratable CRISPRi high-throughput (BATCH) screening for over-production phenotypes. Submitted to ***Metabolic Engineering***.
6. Wang, J., **Li, C.** (co-first author), Zou, Y., Yan, Y., 2020. Bacterial synthesis of C3-C5 diols via extending amino acid catabolism. ***Proceedings of the National Academy of Sciences*** 117(32), 19159-19167.

7. Zou, Y., **Li, C.** (co-first author), Zhang, R., Jiang, T., Liu, N., Wang, J., Wang, X., Yan, Y., 2021. Exploring the Tunability and Dynamic Properties of MarR-PmarO Sensor System in Escherichia coli. ***ACS Synthetic Biology*** 10(8), 2076-2086.
8. Jiang, T., **Li, C.**, Yan, Y., 2020. Optimization of a *p*-Coumaric Acid Biosensor System for Versatile Dynamic Performance. ***ACS Synthetic Biology*** 10(1), 132-144.
9. Jiang, T., **Li, C.**, Teng, Y., Zhang, R., Yan, Y., 2020. Recent advances in improving metabolic robustness of microbial cell factories. ***Current Opinion in Biotechnology*** 66, 69-77.
10. Zhang, R., **Li, C.**, Wang, J., Yan, Y., 2018. Microbial ligninolysis: toward a bottom-up approach for lignin upgrading. ***Biochemistry*** 58(11), 1501-1510.
11. Zhang, R., **Li, C**., Wang, J., Yang, Y., Yan, Y., 2018. Microbial production of small medicinal molecules and biologics: from nature to synthetic pathways. ***Biotechnology advances*** 36(8), 2219-2231.
12. Shen, X., Wang, J., **Li, C**., Yuan, Q., Yan, Y., 2019. Dynamic gene expression engineering as a tool in pathway engineering. ***Current opinion in biotechnology*** 59, 122-129.
13. Yang, Y., Lin, Y., Wang, J., Wu, Y., Zhang, R., Cheng, M., Shen, X., Wang, J., Chen, Z., **Li, C**., Yuan, Q., Yan, Y., 2018. Sensor-regulator and RNAi based bifunctional dynamic control network for engineered microbial synthesis. ***Nature Communications*** 9(1), 3043.
14. Jia, H., Sun, X., Sun, H., **Li, C**., Wang, Y., Feng, X., Li, C., 2016. Intelligent microbial heat-regulating engine (IMHeRE) for improved thermo-robustness and efficiency of bioconversion. ***ACS synthetic biology*** 5(4), 312-320