Model-driven DBTL cycle acceleration with broad-host-range bacterial CRISPRa/i circuits

Cholpisit Kiattisewee

A dissertation
Submitted in partial fulfillment of the
Requirements for the degree of

Doctor of Philosophy

University of Washington
2023

Reading Committee:
Jesse G. Zalatan, Chair
James M. Carothers, Chair
Jorge A. Marchand
Caroline S. Harwood

Program Authorized to Offer Degree
Molecular Engineering and Sciences

University of Washington

Abstract

Model-driven DBTL cycle acceleration with broad-host-range bacterial CRISPRa/i circuits

Cholpisit Kiattisewee

Chairs of the Supervisory Committee:

Jesse G. Zalatan
Department of Chemistry

James M. Carothers
Department of Chemical Engineering

CRISPR-Cas gene regulatory tools have revolutionized biological network programming. Recently, we developed CRISPR gene activation (CRISPRa) tools that demonstrate broad applicability across various bacteria. With thorough characterization of bacterial CRISPRa, we found that the design rules for effective CRISPRa are stringent and highly context-dependent. Thus, we developed numerous strategies to overcome existing limitations — including DNA context engineering, utilization of engineered proteins to bypass target sites requirements, and characterization of multiple bacterial CRISPRa systems for alternative design rules. Implementation of CRISPRa tools in chemical bioproduction enabled synthesis of *p*-aminocinnamic acid, a precursor to various functional polymer materials, which was previously difficult to synthesize through conventional routes. In combination with CRISPR gene interference (CRISPRi), we further explore the capability of a combined CRISPRa/i platform to regulate expression of both foreign genes and native genes intricately involved in bacterial metabolism. Since programmability of CRISPRa/i relies on guide RNA sequence, we found that multiple guide RNAs could be implemented simultaneously to regulate multiple genes in the same system. Engineering at the RNA level could also provide tunable regulation for each gene target. Furthermore, when combined with genome-scale metabolic models, this system accelerates the Design-Build-Test-Learn (DBTL) process for microbial strain optimization, bypassing stepwise genetic reconstruction through the use of trans-acting CRISPRa/i circuits. The resulting constructs can be comprehensively investigated using a multi-omics platform, providing detailed information to improve subsequent DBTL cycles. By coupling programmable and tunable gene regulatory tools with large metabolic models informed by omics data, our platform establishes a foundation for non-canonical microbial strain engineering that benefits diverse disciplines from industrial biotechnology to therapeutic discovery.

To my brother Napaphon (Act) Lertprotsombat,
whose aspirations continue to resonate lively within me

Table of Contents

Table of Figures

Preface

As written above, my name is Cholpisit Kiattisewee which appears to be challenging to pronounce for most English speakers. I was born and raised in Thailand and usually go by "Ice" as a nickname. This dissertation would not be completed without countless continuous support, both technically and emotionally. I expressed my gratitude to everyone in the Acknowledgement section. A brief summary of each technical chapter was listed below.

Chapter 1 is the introduction about CRISPR-based genetic circuits in bacteria entitled "Bacterial CRISPR circuits: Portable Tools for Genetic Engineering and Manipulation" which is part of the preparation from our team for the *Annual Review in Chemical and Biomolecular Engineering*.

Chapter 2 contains our manuscript published in *Metabolic Engineering*, "Portable bacterial CRISPR transcriptional activation enables metabolic engineering in *Pseudomonas putida*". In this paper, we transferred the CRISPRa/i tools previously developed in *E. coli* into *P. putida*. We found good correlation in design rules of CRISPRa in these two organisms which showcased the broad-host-range applicability of CRISPR gene regulatory tools.

Chapter 3 contains our manuscript in preparation under the title "Bioproduction of aromatic amines enabled by stress-resistant bacteria". This work described the long challenge from our team to biochemically synthesize aromatic amine in microbial hosts. We found that *P. putida*, with developed CRISPR tools, enables production of *p*-aminocinnamic acid directly from a simple feedstock — glucose.

Chapter 4 is the on-going application of the bacterial CRISPR circuits and is the core of my dissertation "Model-driven DBTL cycle acceleration with bacterial CRISPRa/i". We repurposed CRISPR tools for Design-Build-Test-Learn strain optimization in industrial biotechnology. This state-of-the-art technology will accelerate the development of the field by allowing rapid prototyping in a broad range of microbial hosts.

Chapter 5 contains our manuscript published in *ACS Synthetic Biology*, "Expanding the scope of bacterial CRISPR activation with PAM-flexible dCas9 variants". Here, we tackle one challenge in bacterial CRISPRa, stringent target sites, via bypassing PAM requirement for CRISPRa programming.

Appendices contain abstracts of research where Ice contributed as a co-author that are not the key narrative of this dissertation.

Chapter 1:

Bacterial CRISPR Circuits: Portable Tools for Genetic Engineering and Manipulation

Part of preparation for a review article for publication in Annual Review in Chemical and Biomolecular Engineering

In an era where genome-scale DNA assembly is increasingly routine, we lack robust approaches to design large, functional genetic programs to effectively engineer microbial biosystems. Engineered microbes have the potential to produce a vast array of valuable chemicals from biomass feedstocks (Intasian et al., 2021; Lee et al., 2019; Nielsen and Keasling, 2016). However, because the underlying metabolic and regulatory networks are large and complex, the development of microorganisms for bioproduction typically requires lengthy study and extensive manipulation. In principle, large gene regulatory networks (GRNs) targeting dozens of genomic and non-native genes could be constructed and implemented to engineer microbes for chemical bioproduction. Central and peripheral metabolisms can be rewired for increased yield, efficient carbon-conserving processes, and CO_2 sequestration (Westenberg and Peralta-Yahya, 2023). In practice, this task is difficult and resource-intensive due to limited understanding in the underlying networks and inability to test large genetic programs with predictable functions.

In the past decade, CRISPR (clustered regularly interspaced short palindromic repeats) technologies have revolutionized the basis of experimental biological research. The discovery of new Cas (CRISPR-associated proteins) and re-discovery of existing Cas functions have expanded the programmability of CRISPR tools beyond gene editing (Collias and Beisel, 2021; Wang et al., 2022; Wang and Doudna, 2023). Most CRISPR systems are composed of two major functions. First, a recognition and binding domain that allowed specific targeting of DNA, RNA, or proteins. Second, an effector domain that brings about the molecular function spanning from inherent nuclease activity, gene regulation (Vigouroux and Bikard, 2020), nucleobase editing (Chen and Liu, 2023; Pickar-Oliver and Gersbach, 2019), gene insertion (Tou et al., 2023), protease activity (Hu et al., 2022), or other programmed molecular recognition (Alba Burbano et al., 2023; Cunningham-Bryant et al., 2019; Kirkpatrick et al., 2020; Villegas Kcam et al., 2021). A combination of target choice (DNA, RNA, Protein) and engineered molecular functions (cut, regulate, insert, etc.) allowed limitless space for molecular programming that can be repurposed for various applications in different sectors, including molecular biology tools, chemical production, health-related issues.

In this chapter, we will focus on the coupling of CRISPR tools with other molecular biology parts to uncover new technologies for new functions spanning from CRISPR regulation circuitry and the implementation of CRISPR circuitry in diverse bacteria. These perspectives will serve as a foundation for CRISPR technologies described in the following chapters.

1.1 CRISPR tools for central dogma control

Central dogma of molecular biology describes the flow of genetic information from DNA to RNA and then protein through three key processes — Replication (DNA-to-DNA), Transcription (DNA-to-RNA), and Translation (RNA-to-protein). The CRISPR system is well known for its application in DNA editing. CRISPR tools are guided by short RNA to generate various effects to the central dogma (Jinek et al., 2012). After its first discovery and repurpose for DNA-editing tools, a plethora of CRISPR tools based on engineered Cas proteins have been developed. Class II (e.g. Cas9), V (e.g. Cas12), and VI (e.g. Cas13) Cas systems exhibit distinct properties and are widely used in diverse applications. They possess either DNA-recognition (class II and V) or RNA-recognition (class VI) features. Some additional features could be repurposed for novel applications, such as collateral nucleases activity for CRISPR diagnostics (Cas12 and Cas13) and transposase activity for guided DNA insertion (Cas12k) (Abudayyeh and Gootenberg, 2021; Kaminski et al., 2021; Tou et al., 2023). Engineering of Cas nucleases also generates new types of functions for these proteins. The most common engineering for Cas enzymes is at the nuclease active sites. Removing nuclease activity results in deactivated Cas proteins (dCas) possessing DNA/RNA recognition property without DNA/RNA cleavage property. These dCas proteins bring about novel applications especially for gene regulation — repression and activation. These DNA/RNA recognitions could be applied for the control of all processes in central dogma.

Using the CRISPR-Cas9 gene editing platform and accompanying sgRNA design as a basis, deactivated Cas9 (dCas9) was repurposed for programmable gene repression in bacteria, termed CRISPR interference (CRISPRi), by physically blocking the transcription initiation or elongation processes (Qi et al., 2013). Other dCas proteins were also implemented in the same fashion, including a repertoire of dCas9s from different species (Rock et al., 2017) and dCas12a (dCpf1) whose CRISPR RNA (crRNA) array processing was maintained which enabled rapid construction of multi-gene repression (Banerjee et al., 2020; Fonfara et al., 2016; Zetsche et al., 2015). Various design rules of each CRISPRi system were investigated covering the strand preference, target site preference, mismatch tolerance, and guide RNA multiplexing. The gene activation counterpart, termed CRISPR activation (CRISPRa), presents a greater challenge due to the diverse mode of gene activation found across different organisms (Figure 1.1). In eukaryotes, CRISPRa was developed based on transcription factors, e.g. VP64, HSF1, and p65 domains, which are widely applicable in many eukaryotic systems, ranging from yeasts, mammalian cells, and plants (Casas-Mollano et al., 2023; Chavez et al., 2015; Konermann et al., 2014; Zalatan et al., 2015). However, prokaryotic CRISPRa is less developed since prokaryotes lack a universal transcriptional activator. The first report of bacterial CRISPRa was established by fusing an RNA polymerase (RNAP) omega subunit (RpoZ) to dCas9, but knockout of rpoZ was found to be necessary for high gene activation (Bikard et al., 2013). Recently, several bacterial CRISPRa systems with high fold-activation have been developed based on engineered bacterial transcription factors, e.g. SoxS (Dong et al., 2018), PspF (Y. Liu et al., 2019), and AsiA (Ho et al., 2020). Current bacterial CRISPRa systems are listed in Table 1.1. However, in these systems, bacterial CRISPRa still suffers from stringent design rules (Fontana et al., 2020), especially on narrow functional target sites restricted by availability of the Protospacer Adjacent Motifs (PAM) sequences. Based on the characterized reporter, shifting the target site by 1–3 bp caused significant decreases in activation and CRISPRa effects were negligible at 4–9 bp shifts.

However, by further shifting to 10–11 bp, the CRISPR-mediated gene activation was recovered. These findings suggested that CRISPRa is highly dependent on 3D positioning of the CRISPRa complex relative to target promoter since 10–11 bp periodicity corresponds to one full turn of a DNA helix. Several alternative solutions have been developed to overcome these limitations, from circular permutations of dCas9 to alter the 3D position of activator relative to dCas9 (Villegas Kcam et al., 2021) or using engineered dCas9 with PAM-flexibility to broaden the target site accessibility (Kiattisewee et al., 2022; Klanschnig et al., 2022). CRISPRa based on dCas12a (Schilling et al., 2020) and type I-E Cascade (Villegas Kcam et al., 2022) were also developed, adding orthogonal tools to the CRISPRa collections. The development of CRISPRa in addition to thorough characterizations of CRISPRi has enabled complex, programmable gene regulatory networks (GRNs) and led to numerous applications which will be described further in this dissertation.

Given the ability of dCas9 to target any DNA sequences, it could also be repurposed for DNA replication control (Figure 1.1). In bacterial systems, most heterologous genes could be delivered on a replicable plasmid or inserted into a bacterial chromosome. Since the bacterial genome is tightly maintained as a single-copy, interfering with the replication of genomic DNA could result in a change in growth rate. CRISPRi targeting at the origin of replication (*oriC*) or DNA replication machinery (*dnaA*) was shown to result in a slower growth rate in *E. coli* and *P. putida* (Liu et al., 2020). Plasmid replication usually relies on host replication machinery and specific non-coding elements installed in the plasmid sequence (Kües and Stahl, 1989). Therefore, CRISPRi could be used to interfere with the specific plasmid parts and could modulate the replication process without perturbing the host chromosomal replication. For ColE1-family plasmid where replication initiation relies on RNA primer and antisense RNA, CRISPRi repression could either yield Copy-UP or Copy-DOWN trait depending on the target sites (Li et al., 2022). On the other hand, for the pSC101-family where replication protein (RepA) is necessary for plasmid maintenance, CRISPRi could be used to regulate expression of RepA, yielding a controllable copy-number (Joshi et al., 2022).

Unlike DNA replication and transcription, the translation process uses RNA as a template instead (Figure 1.1). Therefore, an RNA-targeting Cas protein (class VI) is necessary for this modulation. Several Cas13 proteins have been characterized for various applications from RNA cleavage (Abudayyeh et al., 2016; East-Seletsky et al., 2016), RNA editing (Cox et al., 2017; Qu et al., 2019), RNA detection (Kellner et al., 2019), and live-cell imaging (Yang et al., 2019). Some Cas13 enzymes are less restricted on the protospacer flanking site (PFS), an RNA counterpart of PAM in DNA-targeting Cas, making them highly programmable. Two notable Cas13 variants are Cas13a (C2c2) and Cas13d (CasRx) representing the first and most recent variants with foci in diagnostic and therapeutic, respectively (Kiga et al., 2020; Konermann et al., 2018). The deactivated versions of Cas13 have also been investigated for gene repression by physically blocking ribosomes during the translational elongation process (Charles et al., 2021). For gene activation, dCas13 could be repurposed for mRNA stabilization or used for recruiting a translational initiation factor to increase the magnitude of protein expression (Otoupal et al., 2022).

New types of Cas proteins are actively being discovered and may provide novel functions and applications beyond three classes of Cas proteins mentioned above (class II, V, and VI). However, the general design rules for central dogma programming should remain applicable

independent of DNA or RNA modulating machinery. As most CRISPR systems use short spacer sequences (20-28nt) to identify the target, a massive-scale gRNA library could be constructed at accessible cost for parallel testing of different applications, ranging from optimization of bioproduction to identification of therapeutic targets (Fenster et al., 2022; Jiang et al., 2020; Liu et al., 2021; Peng et al., 2018). These emerging and developing technologies are changing the way scientists and engineers tackle fundamental research questions that are previously technically-difficult or cost-inhibitive. In this dissertation, we will describe the new applications in metabolic engineering enabled by central dogma regulation technologies.

1.2 Implementation of CRISPR circuits in bacteria

The simplest method of implementing genetic circuits in bacteria is to deliver the corresponding DNA that can be further expressed as RNA and protein machineries. Foreign DNA can be maintained in bacteria either as a self-replicating plasmid or integrated into the host genome (Figure 2). The plasmid-based strategy serves as a hallmark for bacterial synthetic biology as it is much easier to assemble DNA in a plasmid format compared to genome integration. However, there are several challenges coming together with plasmid-based gene expression. First, the plasmid DNA replication mechanism is usually intertwined with host DNA replication using the host's machinery. Therefore, the host-range of bacterial plasmids is usually limited to a narrow range, e.g. ColE1-related plasmids are usually replicable in *E. coli* and its close relatives (Ares-Arroyo et al., 2021). Second, as the necessary number of heterologous genes increases, the payload should be distributed to an increasing number of plasmids and bigger plasmid sizes which are limited depending on several factors. Multiple plasmids can be simultaneously maintained in *E. coli* (Tomoiaga et al., 2022; Villegas Kcam et al., 2021). However, selection pressure has to be maintained for each unique plasmid, usually through antibiotic selections, and could cause accumulated metabolic burden and then decreased gene expression from each node (Silva et al., 2012). On the other hand, if heterologous genes were maintained on a minimal number of plasmids, the plasmid will grow bigger and thus lead to decreasing transformation efficiency in bacteria or inability to pack in a viral vector for eukaryotic DNA delivery systems (Wu et al., 2010). Although naturally-occuring plasmids could be as big as 200k, the plasmid which is smaller than 10kb is much more stable and easy to be transformed (Rodríguez-Beltrán et al., 2021; Sheng et al., 1995). Therefore, foreign genes with big DNA payloads are usually integrated into the host genome instead of being maintained as replicative plasmids. Recently, various genetic tools were developed in diverse bacterial hosts (Choi and Lee, 2020; Wirth et al., 2019). The rise of CRISPR gene editing platform also accelerates novel method developments (Biggs et al., 2020; Collias et al., 2023; Cook et al., 2018). Moreover, the new design based on integrase and pre-installed landing pads allowed highly-efficient genome integration methods (Elmore et al., 2017; Wang et al., 2019). The further coupling of these integrase-based technologies with combinatorial DNA library allowed massively-parallel screening of diverse genotypes (Elmore et al., n.d.; Ke et al., 2022). In this section, we will discuss DNA delivery methods used in CRISPR technologies development and explore the plausible design for modular program construction in prokaryotic systems.

To implement CRISPR circuits, at least two components — a Cas protein and a guide RNA — are necessary. The initial reports of CRISPRi repression or CRISPRa activation in *E. coli* are usually implemented by expressing dCas9 or engineered dCas9 on a medium-copy plasmid

(p15A) while expressing guide RNA or modified guide RNA on a high-copy plasmid (ColE1) (Bikard et al., 2013; Qi et al., 2013). The target genes are either residing in the genome or expressed in a low-copy plasmid (pSC101) (Dong et al., 2018; Ho et al., 2020). In the advent of new CRISPRa systems where the activator part is separated from that of the Cas protein, modified activator domain is expressed on another compatible plasmid (Villegas Kcam et al., 2021) or on the same plasmid as the Cas protein (Dong et al., 2018; Y. Liu et al., 2019). Therefore, the number of plasmids used in bacterial expression could total up to four plasmids simultaneously. However, these plasmid borne strategies are not directly transferable to non-model organisms for several reasons. First, the plasmids used for gene expression might be narrow range and are incompatible with the chassis of choice. Second, the gene expression derived from the plasmid of different copy-numbers could lead to improper expression levels and lead to impaired CRISPRa or CRISPRi efficiency. In *P. putida* CRISPRa/i, expression of dCas9 on medium-copy plasmids exhibited severe growth defects and CRISPR programs were optimal with dCas9 and activator expressed from the single-copy genome while keeping other parts on the plasmids (Kiattisewee et al., 2021). It is also possible to express all parts from the CRISPR machinery to targeted metabolic genes from a bacterial genome where expression levels are relatively more consistent (Ke et al., 2022).

Based on the theoretical evaluations of CRISPRa and CRISPRi (Clamons and Murray, 2019), the number of target gene copy-numbers significantly affects the fold-change derived from CRISPRa/i circuits. Other parameters ranging from concentrations of each component and binding affinity between each part also play an important role. Thus, programmability of CRISPR circuits requires proper tuning of those parameters starting from the expression levels of each component, whether from a multi-copy plasmid or a single-copy genome. In this dissertation, multiple CRISPR circuits were designed and optimized based on the specific use case which will be described in detail in each chapter.

1.3 Tables and Figures

Table 1.1: Reported CRISPR activation systems in bacteria

CRISPRa systems	Mechanism	Optimal sites[a]	Strand	Organisms	References
Recruitment via dCas proteins					
dCas9-RpoZ	RNAP assembly	-60 to -100	NT	*E. coli, B. subtilis, L. enzymogenes, M. xanthus*	(Bikard et al., 2013; Lu et al., 2019; Peng et al., 2018; Yu et al., 2018)
dCas9-RpoA	RNAP assembly	-267 to -415	T	*B. subtilis*	(Lu et al., 2019)
dCas9-SYNZIP-αNTD	RNAP assembly	-60 to -100	NT	*E. coli*	(Ameruoso et al., 2022; Villegas Kcam et al., 2021)
dCas9-RpoD	RNAP assembly	-199 to -216	NT	*S. oneidensis*	(Chen et al., 2022)
dCas9-AsiA[b]	RpoD and RpoB binding	-182 to -252	T	*E. coli, S. enterica, K. oxytoca*	(Ho et al., 2020)
dCas12a-SoxS	RpoA and RpoD binding	-50 to -150	T	*P. polymyxa*	(Schilling et al., 2020)
dCas12a-RpoZ	RNAP assembly	-183 to -328	T	*C. glutamicum*	(W. Liu et al., 2019)
dCas12a-RemA	Unknown	-90	NT	*B. subtilis*	(Wu et al., 2020)
Cascade-αNTD	RNAP assembly	-100 to -120	T	*E. coli*	(Villegas Kcam et al., 2022)
dCas13d-IF	Ribosome recruitment	5'-UTR	-	*E. coli*	(Otoupal et al., 2022)

CRISPRa systems	Mechanism	Optimal sites[a]	Strand	Organisms	References
Recruitment via guide RNA					
SoxS[b] (MCP or PCP or MCP + SYNZIP)	RpoA and RpoD binding	-60 to -100	NT	*E. coli,* *P. putida,* Cell-Free System, Hydrogel	(Alba Burbano et al., 2023; Dong et al., 2018; Fontana et al., 2020; Kiattisewee et al., 2021; Tickman et al., 2021)
TetD (MCP)	RpoA and RpoD binding	-60 to -100	NT	*E. coli*	(Dong et al., 2018; Fontana et al., 2020)
RpoZ (MCP)	RNAP assembly	-91	NT	*E. coli*	(Dong et al., 2018; Fontana et al., 2020)
αNTD (MCP)	RNAP assembly	-60 to -100	T	*E. coli*	(Dong et al., 2018; Fontana et al., 2020)
PspF[b] (λN22 or MCP)	RpoN recruitment[c]	-91 to -131[d]	NT	*E. coli,* *K. oxytoca,* Cell-Free System	(Liu et al., 2022; Y. Liu et al., 2019)

Notes: [a]Relative to transcription start site (TSS), [b]Engineered proteins, [c]RpoN promoter is distinct to other bacterial promoters of RpoD-promoter superfamily, [d]DNA looping is present in the promoter

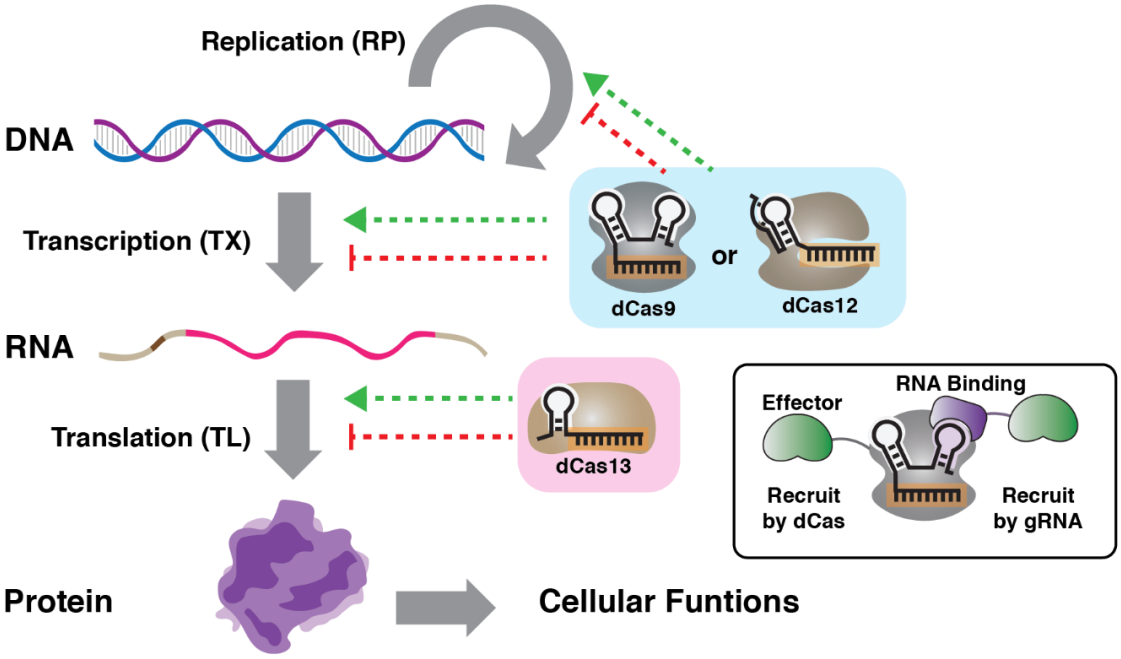


Figure 1.1: Central Dogma Control by CRISPR tools

The central dogma of molecular biology could be modulated by CRISPR-Cas systems. Transcription process by reading DNA into RNA could be regulated by interfering with RNA polymerase (RNAP) binding or elongation, termed CRISPR interference (CRISPRi), or recruitment of RNAP to the promoter, termed CRISPR activation (CRISPRa), using the transcription factor fused or append to the CRISPR complex (bracket). dCas9 and dCas12 are major Cas proteins used in CRISPRa and CRISPRi processes for DNA-targeting activity. These DNA-targeting modules could also be repurposed for controlling DNA-to-DNA replication. Targeting at the origin of replication could slow down the replication process and could also be further applied for targeting DNA replicating machines using the design rules of well-established CRISPRa/i. Finally, RNA-targeting dCas13 is a new hallmark of translational modification. A similar strategy previously explored in the transcription control could be adapted to develop translational repression or activation. RNA stabilization is also unique to translational control as protein production rate could be manipulated independently of the DNA processing.

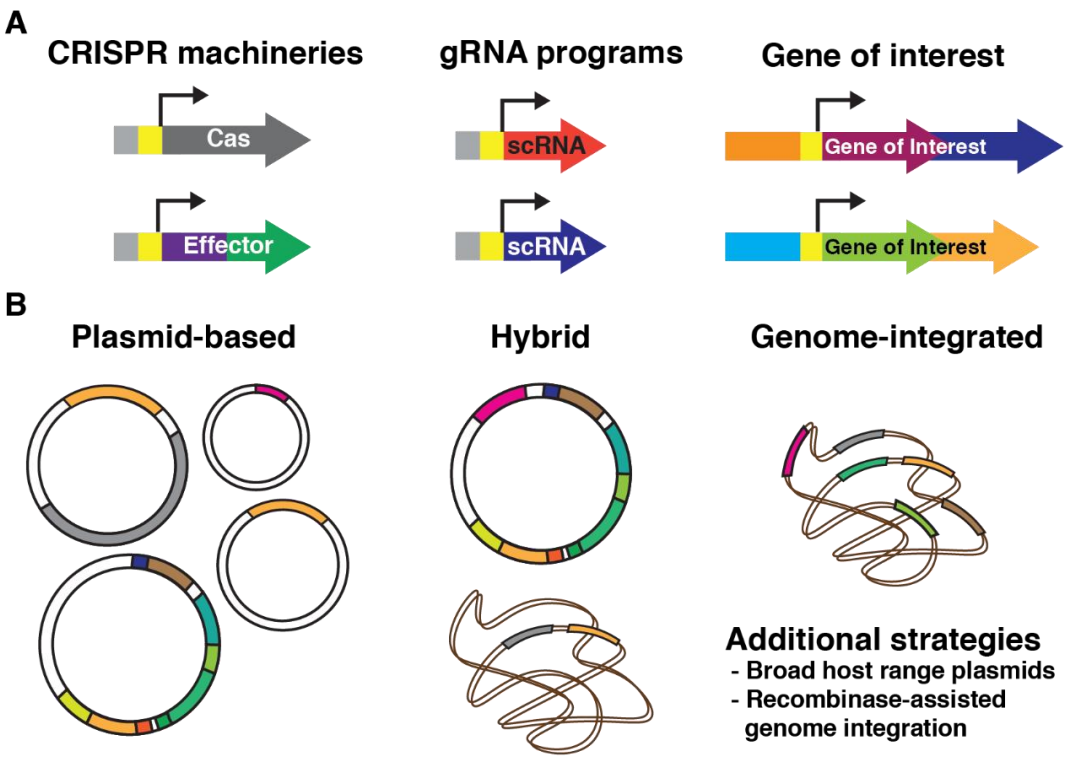


Figure 1.2: Bacterial Expression Strategies of CRISPR Circuits Machinery

(A) Bacterial CRISPR circuits could be employed by expressing CRISPR machinery, CRISPR gRNA programs, and heterologous genes of interest. (B) Delivery of CRISPR tools and heterologous genes could be categorized into three common strategies — i) plasmid-based expression, ii) genome-integration, and iii) hybrid of plasmid expression and genome-integration. Plasmid-based gene expression is the simplest form of heterologous DNA delivery which could be used to alter metabolism of microbes. However, incremental burden derived from plasmid replication and expression of selection markers poses disadvantages for optimal gene expression. It is also possible to manipulate the copy-number of plasmid in the plasmid replication processes using CRISPR tools. On the other hand, genome-integrated machinery provides stable gene expression from a single copy of DNA which has minimal burden but suffers from a slower genetic engineering process. Finally, hybrid gene expression strategy is possible by delivering big genetic payload onto the genome while keeping programmable parts on the easy-to-build plasmids. This strategy allowed minimal burden coming from replication of big plasmid while keeping the agile construction processes at the plasmid level. Other strategies including broad host range plasmids and recombinase-assisted genome integration could also be employed to accelerate genetic engineering processes.

1.4 References

Abudayyeh, O.O., Gootenberg, J.S., **2021**. CRISPR diagnostics. *Science* **372**, 914–915. https://doi.org/10.1126/science.abi9335

Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B.T., Shmakov, S., Makarova, K.S., Semenova, E., Minakhin, L., Severinov, K., Regev, A., Lander, E.S., Koonin, E.V., Zhang, F., **2016**. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353**, aaf5573. https://doi.org/10.1126/science.aaf5573

Alba Burbano, D., Cardiff, R.A.L., Tickman, B.I., Kiattisewee, C., Maranas, C.J., Zalatan, J.G., Carothers, J.M., **2023**. Engineering activatable promoters for scalable and multi-input CRISPRa/i circuits. *Proc. Natl. Acad. Sci.* **120**, e2220358120. https://doi.org/10.1073/pnas.2220358120

Ameruoso, A., Villegas Kcam, M.C., Cohen, K.P., Chappell, J., **2022**. Activating natural product synthesis using CRISPR interference and activation systems in *Streptomyces*. *Nucleic Acids Res.*

Ares-Arroyo, M., Rocha, E.P.C., Gonzalez-Zorn, B., **2021**. Evolution of ColE1-like plasmids across γ-Proteobacteria: From bacteriocin production to antimicrobial resistance. *PLOS Genet.* **17**, e1009919. https://doi.org/10.1371/journal.pgen.1009919

Banerjee, D., Eng, T., Lau, A.K., Sasaki, Y., Wang, B., Chen, Y., Prahl, J.-P., Singan, V.R., Herbert, R.A., Liu, Y., Tanjore, D., Petzold, C.J., Keasling, J.D., Mukhopadhyay, A., **2020**. Genome-scale metabolic rewiring improves titers rates and yields of the non-native product indigoidine at scale. *Nat. Commun.* **11**, 5385. https://doi.org/10.1038/s41467-020-19171-4

Biggs, B.W., Bedore, S.R., Arvay, E., Huang, S., Subramanian, H., McIntyre, E.A., Duscent-Maitland, C.V., Neidle, E.L., Tyo, K.E.J., **2020**. Development of a genetic toolset for the highly engineerable and metabolically versatile *Acinetobacter baylyi* ADP1. *Nucleic Acids Res.* **48**, 5169–5182. https://doi.org/10.1093/nar/gkaa167

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 Res* **41**, 7429–37. https://doi.org/10.1093/nar/gkt520

Casas-Mollano, J.A., Zinselmeier, M.H., Sychla, A., Smanski, M.J., **2023**. Efficient gene activation in plants by the MoonTag programmable transcriptional activator. *Nucleic Acids Res.* gkad458. https://doi.org/10.1093/nar/gkad458

Charles, E.J., Kim, S.E., Knott, G.J., Smock, D., Doudna, J., Savage, D.F., **2021**. Engineering improved Cas13 effectors for targeted post-transcriptional regulation of gene expression. *bioRxiv*.

Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., E, P.R.I., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J., Ter-Ovanesyan, D., Braff, J.L., Davidsohn, N., Housden, B.E., Perrimon, N., Weiss, R., Aach, J., Collins, J.J., Church, G.M., **2015**. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* **12**, 326–8. https://doi.org/10.1038/nmeth.3312

Chen, P.J., Liu, D.R., **2023**. Prime editing for precise and highly versatile genome manipulation. *Nat. Rev. Genet.* **24**, 161–177. https://doi.org/10.1038/s41576-022-00541-1

Chen, Y., Niu, X., Cheng, M., Wang, L., Sun, P., Song, H., Cao, Y., **2022**. CRISPR/dCas9-RpoD-Mediated Simultaneous Transcriptional Activation and Repression in *Shewanella oneidensis* MR-1. *ACS Synth. Biol.* **11**, 2184–2192. https://doi.org/10.1021/acssynbio.2c00149

Choi, K.R., Lee, S.Y., **2020**. Protocols for RecET-based markerless gene knockout and integration to express heterologous biosynthetic gene clusters in *Pseudomonas putida*.

Microb. Biotechnol. **13**, 199–209. https://doi.org/10.1111/1751-7915.13374
Clamons, S., Murray, R., **2019**. Modeling predicts that CRISPR-based activators, unlike CRISPR-based repressors, scale well with increasing gRNA competition and dCas9 bottlenecking. *bioRxiv* 719278. https://doi.org/10.1101/719278
Collias, D., Beisel, C.L., **2021**. CRISPR technologies and the search for the PAM-free nuclease. *Nat. Commun.* **12**, 555. https://doi.org/10.1038/s41467-020-20633-y
Collias, D., Vialetto, E., Yu, J., Co, K., Almási, É. d. H., Rüttiger, A.-S., Achmedov, T., Strowig, T., Beisel, C.L., **2023**. Systematically attenuating DNA targeting enables CRISPR-driven editing in bacteria. *Nat. Commun.* **14**, 680. https://doi.org/10.1038/s41467-023-36283-9
Cook, T.B., Rand, J.M., Nurani, W., Courtney, D.K., Liu, S.A., Pfleger, B.F., **2018**. Genetic tools for reliable gene expression and recombineering in *Pseudomonas putida*. *J. Ind. Microbiol. Biotechnol.* **45**, 517–527. https://doi.org/10.1007/s10295-017-2001-5
Cox, D.B.T., Gootenberg, J.S., Abudayyeh, O.O., Franklin, B., Kellner, M.J., Joung, J., Zhang, F., **2017**. RNA editing with CRISPR-Cas13. *Science* **358**, 1019–1027. https://doi.org/10.1126/science.aaq0180
Cunningham-Bryant, D., Sun, J., Fernandez, B., Zalatan, J.G., **2019**. CRISPR–Cas-Mediated Chemical Control of Transcriptional Dynamics in Yeast. *ChemBioChem* **20**, 1519–1523. https://doi.org/10.1002/cbic.201800823
Dong, C., Fontana, J., Patel, A., Carothers, J.M., Zalatan, J.G., **2018**. Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nat Commun* **9**, 2489. https://doi.org/10.1038/s41467-018-04901-6
East-Seletsky, A., O'Connell, M.R., Knight, S.C., Burstein, D., Cate, J.H.D., Tjian, R., Doudna, J.A., **2016**. Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature* **538**, 270–273. https://doi.org/10.1038/nature19802
Elmore, J.R., Dexter, G.N., Baldino, H., Huenemann, J.D., Francis, R., Peabody, G.L., Martinez-Baird, J., Riley, L.A., Simmons, T., Coleman-Derr, D., Guss, A.M., Egbert, R.G., **2023**. High-throughput genetic engineering of nonmodel and undomesticated bacteria via iterative site-specific genome integration. *Sci. Adv.* **9**, eade1285. https://doi.org/10.1126/sciadv.ade1285
Elmore, J.R., Furches, A., Wolff, G.N., Gorday, K., Guss, A.M., **2017**. Development of a high efficiency integration system and promoter library for rapid modification of *Pseudomonas putida* KT2440. *Metab. Eng. Commun.* **5**, 1–8. https://doi.org/10.1016/j.meteno.2017.04.001
Fenster, J.A., Werner, A.Z., Tay, J.W., Gillen, M., Schirokauer, L., Hill, N.C., Watson, A., Ramirez, K.J., Johnson, C.W., Beckham, G.T., Cameron, J.C., Eckert, C.A., **2022**. Dynamic and single cell characterization of a CRISPR-interference toolset in *Pseudomonas putida* KT2440 for β-ketoadipate production from p-coumarate. *Metab. Eng. Commun.* **15**, e00204. https://doi.org/10.1016/j.mec.2022.e00204
Fonfara, I., Richter, H., Bratovič, M., Le Rhun, A., Charpentier, E., **2016**. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* **532**, 517–521. https://doi.org/10.1038/nature17945
Fontana, J., Dong, C., Kiattisewee, C., Chavali, V.P., Tickman, B.I., Carothers, J.M., Zalatan, J.G., **2020**. Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. *Nat. Commun.* **11**, 1618. https://doi.org/10.1038/s41467-020-15454-y
Ho, H.-I., Fang, J.R., Cheung, J., Wang, H.H., **2020**. Programmable CRISPR-Cas transcriptional activation in bacteria. *Mol. Syst. Biol.* **16**, e9427. https://doi.org/10.15252/msb.20199427
Hu, C., van Beljouw, S.P.B., Nam, K.H., Schuler, G., Ding, F., Cui, Y., Rodríguez-Molina, A., Haagsma, A.C., Valk, M., Pabst, M., Brouns, S.J.J., Ke, A., **2022**. Craspase is a CRISPR RNA-guided, RNA-activated protease. *Science* **377**, 1278–1285.

https://doi.org/10.1126/science.add5064
Intasian, P., Prakinee, K., Phintha, A., Trisrivirat, D., Weeranoppanant, N., Wongnate, T., Chaiyen, P., **2021**. Enzymes, In Vivo Biocatalysis, and Metabolic Engineering for Enabling a Circular Economy and Sustainability. *Chem. Rev.* **121**, 10367–10451. https://doi.org/10.1021/acs.chemrev.1c00121
Jiang, W., Oikonomou, P., Tavazoie, S., **2020**. Comprehensive Genome-wide Perturbations via CRISPR Adaptation Reveal Complex Genetics of Antibiotic Sensitivity. *Cell* **180**, 1002-1017.e31. https://doi.org/10.1016/j.cell.2020.02.007
Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., **2012**. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816–821. https://doi.org/10.1126/science.1225829
Joshi, S.H.-N., Yong, C., Gyorgy, A., **2022**. Inducible plasmid copy number control for synthetic biology in commonly used *E. coli* strains. *Nat. Commun.* **13**, 6691. https://doi.org/10.1038/s41467-022-34390-7
Kaminski, M.M., Abudayyeh, O.O., Gootenberg, J.S., Zhang, F., Collins, J.J., **2021**. CRISPR-based diagnostics. *Nat. Biomed. Eng.* **5**, 643–656. https://doi.org/10.1038/s41551-021-00760-7
Ke, J., Robinson, D., Wu, Z.-Y., Kuftin, A., Louie, K., Kosina, S., Northen, T., Cheng, J.-F., Yoshikuni, Y., **2022**. CRAGE-CRISPR facilitates rapid activation of secondary metabolite biosynthetic gene clusters in bacteria. *Cell Chem. Biol.* **29**, 696-710.e4. https://doi.org/10.1016/j.chembiol.2021.08.009
Kellner, M.J., Koob, J.G., Gootenberg, J.S., Abudayyeh, O.O., Zhang, F., **2019**. SHERLOCK: nucleic acid detection with CRISPR nucleases. *Nat. Protoc.* **14**, 2986–3012. https://doi.org/10.1038/s41596-019-0210-2
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*. *Metab. Eng.* **66**, 283–295. https://doi.org/10.1016/j.ymben.2021.04.002
Kiattisewee, C., Karanjia, A.V., Legut, M., Daniloski, Z., Koplik, S.E., Nelson, J., Kleinstiver, B.P., Sanjana, N.E., Carothers, J.M., Zalatan, J.G., **2022**. Expanding the Scope of Bacterial CRISPR Activation with PAM-Flexible dCas9 Variants. *ACS Synth. Biol.* **11**, 4103–4112. https://doi.org/10.1021/acssynbio.2c00405
Kiga, K., Tan, X.-E., Ibarra-Chávez, R., Watanabe, S., Aiba, Y., Sato'o, Y., Li, F.-Y., Sasahara, T., Cui, B., Kawauchi, M., Boonsiri, T., Thitiananpakorn, K., Taki, Y., Azam, A.H., Suzuki, M., Penadés, J.R., Cui, L., **2020**. Development of CRISPR-Cas13a-based antimicrobials capable of sequence-specific killing of target bacteria. *Nat. Commun.* **11**, 2934. https://doi.org/10.1038/s41467-020-16731-6
Kirkpatrick, R.L., Lewis, K., Langan, R.A., Lajoie, M.J., Boyken, S.E., Eakman, M., Baker, D., Zalatan, J.G., **2020**. Conditional Recruitment to a DNA-Bound CRISPR–Cas Complex Using a Colocalization-Dependent Protein Switch. *ACS Synth. Biol.* **9**, 2316–2323. https://doi.org/10.1021/acssynbio.0c00012
Klanschnig, M., Cserjan-Puschmann, M., Striedner, G., Grabherr, R., **2022**. CRISPRactivation-SMS, a message for PAM sequence independent gene up-regulation in *Escherichia coli*. *Nucleic Acids Res.* gkac804. https://doi.org/10.1093/nar/gkac804
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., Nureki, O., Zhang, F., **2014**. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583. https://doi.org/10.1038/nature14136
Konermann, S., Lotfy, P., Brideau, N.J., Oki, J., Shokhirev, M.N., Hsu, P.D., **2018**. Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. *Cell* **173**, 665-676.e14. https://doi.org/10.1016/j.cell.2018.02.033

Kües, U., Stahl, U., **1989**. Replication of plasmids in gram-negative bacteria. *Microbiol. Rev.* **53**, 491–516.
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. *Nat. Catal.* **2**, 18–33. https://doi.org/10.1038/s41929-018-0212-4
Li, C., Zou, Y., Jiang, T., Zhang, J., Yan, Y., **2022**. Harnessing plasmid replication mechanism to enable dynamic control of gene copy in bacteria. *Metab. Eng.* **70**, 67–78. https://doi.org/10.1016/j.ymben.2022.01.003
Liu, W., Tang, D., Wang, H., Lian, J., Huang, L., Xu, Z., **2019**. Combined genome editing and transcriptional repression for metabolic pathway engineering in *Corynebacterium glutamicum* using a catalytically active Cas12a. *Appl. Microbiol. Biotechnol.* **103**, 8911–8922. https://doi.org/10.1007/s00253-019-10118-4
Liu, X., Kimmey, J.M., Matarazzo, L., de Bakker, V., Van Maele, L., Sirard, J.-C., Nizet, V., Veening, J.-W., **2021**. Exploration of Bacterial Bottlenecks and *Streptococcus pneumoniae* Pathogenesis by CRISPRi-Seq. *Cell Host Microbe* **29**, 107-120.e6. https://doi.org/10.1016/j.chom.2020.10.001
Liu, Y., Chen, J., Crisante, D., Jaramillo Lopez, J.M., Mahadevan, R., **2020**. Dynamic Cell Programming with Quorum Sensing-Controlled CRISPRi Circuit. *ACS Synth. Biol.* **9**, 1284–1291. https://doi.org/10.1021/acssynbio.0c00148
Liu, Y., Pinto, F., Wan, X., Yang, Z., Peng, S., Li, M., Cooper, J.M., Xie, Z., French, C.E., Wang, B., **2022**. Reprogrammed tracrRNAs enable repurposing of RNAs as crRNAs and sequence-specific RNA biosensors. *Nat. Commun.* **13**, 1937. https://doi.org/10.1038/s41467-022-29604-x
Liu, Y., Wan, X., Wang, B., **2019**. Engineered CRISPRa enables programmable eukaryote-like gene activation in bacteria. *Nat. Commun.* **10**, 3693. https://doi.org/10.1038/s41467-019-11479-0
Lu, Z., Yang, S., Yuan, X., Shi, Y., Ouyang, L., Jiang, S., Yi, L., Zhang, G., **2019**. CRISPR-assisted multi-dimensional regulation for fine-tuning gene expression in *Bacillus subtilis*. *Nucleic Acids Res.* **47**, e40–e40. https://doi.org/10.1093/nar/gkz072
Nielsen, J., Keasling, J.D., **2016**. Engineering Cellular Metabolism. *Cell* **164**, 1185–1197. https://doi.org/10.1016/j.cell.2016.02.004
Otoupal, P.B., Cress, B.F., Doudna, J.A., Schoeniger, J.S., **2022**. CRISPR-RNAa: targeted activation of translation using dCas13 fusions to translation initiation factors. *Nucleic Acids Res.* **50**, 8986–8998. https://doi.org/10.1093/nar/gkac680
Peng, R., Wang, Y., Feng, W., Yue, X., Chen, J., Hu, X., Li, Z., Sheng, D., Zhang, Y., Li, Y., **2018**. CRISPR/dCas9-mediated transcriptional improvement of the biosynthetic gene cluster for the epothilone production in *Myxococcus xanthus*. *Microb. Cell Factories* **17**, 15. https://doi.org/10.1186/s12934-018-0867-1
Pickar-Oliver, A., Gersbach, C.A., **2019**. The next generation of CRISPR–Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.* **20**, 490–507. https://doi.org/10.1038/s41580-019-0131-5
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**, 1173–83. https://doi.org/10.1016/j.cell.2013.02.022
Qu, L., Yi, Z., Zhu, S., Wang, C., Cao, Z., Zhou, Z., Yuan, P., Yu, Y., Tian, F., Liu, Z., Bao, Y., Zhao, Y., Wei, W., **2019**. Programmable RNA editing by recruiting endogenous ADAR using engineered RNAs. *Nat. Biotechnol.* **37**, 1059–1069. https://doi.org/10.1038/s41587-019-0178-z
Rock, J.M., Hopkins, F.F., Chavez, A., Diallo, M., Chase, M.R., Gerrick, E.R., Pritchard, J.R., Church, G.M., Rubin, E.J., Sassetti, C.M., Schnappinger, D., Fortune, S.M., **2017**. Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR

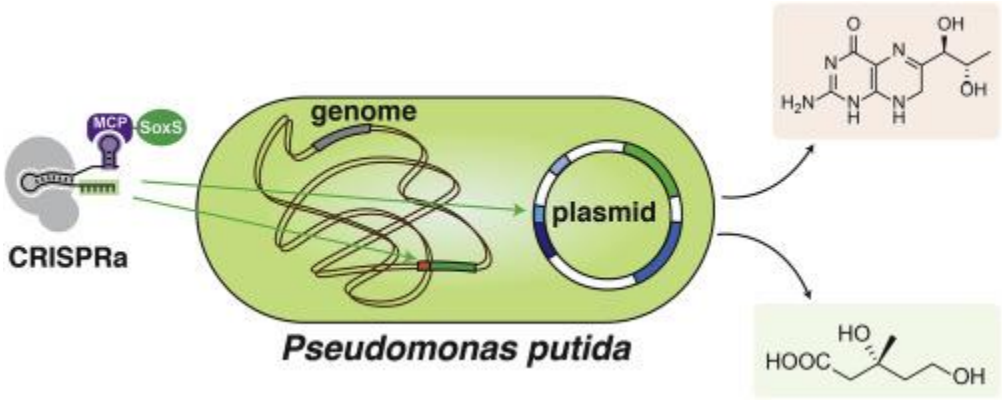
interference platform. *Nat. Microbiol.* **2**, 16274. https://doi.org/10.1038/nmicrobiol.2016.274
Rodríguez-Beltrán, J., DelaFuente, J., León-Sampedro, R., MacLean, R.C., San Millán, Á., **2021**. Beyond horizontal gene transfer: the role of plasmids in bacterial evolution. *Nat. Rev. Microbiol.* **19**, 347–359. https://doi.org/10.1038/s41579-020-00497-1
Schilling, C., Koffas, M.A.G., Sieber, V., Schmid, J., **2020**. Novel Prokaryotic CRISPR-Cas12a-Based Tool for Programmable Transcriptional Activation and Repression. *ACS Synth. Biol.* **9**, 3353–3363. https://doi.org/10.1021/acssynbio.0c00424
Sheng, Y., Mancino, V., Birren, B., **1995**. Transformation of *Escherichia coli* with large DNA molecules by electroporation. *Nucleic Acids Res.* **23**, 1990–1996. https://doi.org/10.1093/nar/23.11.1990
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**, 691–708. https://doi.org/10.1016/j.biotechadv.2011.12.005
Tickman, B.I., Burbano, D.A., Chavali, V.P., Kiattisewee, C., Fontana, J., Khakimzhan, A., Noireaux, V., Zalatan, J.G., Carothers, J.M., **2021**. Multi-layer CRISPRa/i circuits for dynamic genetic programs in cell-free and bacterial systems. *Cell Syst.* S2405471221004191. https://doi.org/10.1016/j.cels.2021.10.008
Tomoiaga, D., Bubnell, J., Herndon, L., Feinstein, P., **2022**. High rates of plasmid cotransformation in *E. coli* overturn the clonality myth and reveal colony development. *Sci. Rep.* **12**, 11515. https://doi.org/10.1038/s41598-022-14598-9
Tou, C.J., Orr, B., Kleinstiver, B.P., **2023**. Precise cut-and-paste DNA insertion using engineered type V-K CRISPR-associated transposases. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-022-01574-x
Vigouroux, A., Bikard, D., **2020**. CRISPR Tools To Control Gene Expression in Bacteria. *Microbiol. Mol. Biol. Rev.* **84**, e00077-19. https://doi.org/10.1128/mmbr.00077-19
Villegas Kcam, M.C., Tsong, A.J., Chappell, J., **2022**. Uncovering the Distinct Properties of a Bacterial Type I-E CRISPR Activation System. *ACS Synth. Biol.* **11**, 1000–1003. https://doi.org/10.1021/acssynbio.1c00496
Villegas Kcam, M.C., Tsong, A.J., Chappell, J., **2021**. Rational engineering of a modular bacterial CRISPR–Cas activation platform with expanded target range. *Nucleic Acids Res.* **49**, 4793–4802. https://doi.org/10.1093/nar/gkab211
Wang, G., Zhao, Z., Ke, J., Engel, Y., Shi, Y.-M., Robinson, D., Bingol, K., Zhang, Z., Bowen, B., Louie, K., Wang, B., Evans, R., Miyamoto, Y., Cheng, K., Kosina, S., De Raad, M., Silva, L., Luhrs, A., Lubbe, A., Hoyt, D.W., Francavilla, C., Otani, H., Deutsch, S., Washton, N.M., Rubin, E.M., Mouncey, N.J., Visel, A., Northen, T., Cheng, J.-F., Bode, H.B., Yoshikuni, Y., **2019**. CRAGE enables rapid activation of biosynthetic gene clusters in undomesticated bacteria. *Nat. Microbiol.* **4**, 2498–2510. https://doi.org/10.1038/s41564-019-0573-8
Wang, J.Y., Doudna, J.A., **2023**. CRISPR technology: A decade of genome editing is only the beginning. *Science* **379**, eadd8643. https://doi.org/10.1126/science.add8643
Wang, J.Y., Pausch, P., Doudna, J.A., **2022**. Structural biology of CRISPR–Cas immunity and genome editing enzymes. *Nat. Rev. Microbiol.* **20**, 641–656. https://doi.org/10.1038/s41579-022-00739-4
Westenberg, R., Peralta-Yahya, P., **2023**. Toward implementation of carbon-conservation networks in nonmodel organisms. *Curr. Opin. Biotechnol.* **81**, 102949. https://doi.org/10.1016/j.copbio.2023.102949
Wirth, N.T., Kozaeva, E., Nikel, P.I., **2019**. Accelerated genome engineering of *Pseudomonas putida* by I-SceI-mediated recombination and CRISPR-Cas9 counterselection. *Microb Biotechnol* **n/a**. https://doi.org/10.1111/1751-7915.13396
Wu, Y., Liu, Y., Lv, X., Li, J., Du, G., Liu, L., **2020**. CAMERS-B: CRISPR/Cpf1 assisted multiple-

genes editing and regulation system for *Bacillus subtilis*. *Biotechnol Bioeng* **117**, 1817–1825. https://doi.org/10.1002/bit.27322
Wu, Z., Yang, H., Colosi, P., **2010**. Effect of Genome Size on AAV Vector Packaging. *Mol. Ther.* **18**, 80–86. https://doi.org/10.1038/mt.2009.255
Yang, L.-Z., Wang, Y., Li, S.-Q., Yao, R.-W., Luan, P.-F., Wu, H., Carmichael, G.G., Chen, L.-L., **2019**. Dynamic Imaging of RNA in Living Cells by CRISPR-Cas13 Systems. *Mol. Cell* **76**, 981-997.e7. https://doi.org/10.1016/j.molcel.2019.10.024
Yu, L., Su, W., Fey, P.D., Liu, F., Du, L., **2018**. Yield Improvement of the Anti-MRSA Antibiotics WAP-8294A by CRISPR/dCas9 Combined with Refactoring Self-Protection Genes in *Lysobacter enzymogenes* OH11. *ACS Synth. Biol.* **7**, 258–266. https://doi.org/10.1021/acssynbio.7b00293
Zalatan, J.G., Lee, M.E., Almeida, R., Gilbert, L.A., Whitehead, E.H., La Russa, M., Tsai, J.C., Weissman, J.S., Dueber, J.E., Qi, L.S., Lim, W.A., **2015**. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* **160**, 339–50. https://doi.org/10.1016/j.cell.2014.11.052
Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., Koonin, E.V., Zhang, F., **2015**. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* **163**, 759–771. https://doi.org/10.1016/j.cell.2015.09.038

Chapter 2:

Portable bacterial CRISPR transcriptional activation enables metabolic engineering in *Pseudomonas putida*

Cholpisit Kiattisewee, Chen Dong, Jason Fontana, Widianti Sugianto, Pamela Peralta-Yahya, James M. Carothers*, Jesse G. Zalatan*



Abstract

CRISPR-Cas transcriptional programming in bacteria is an emerging tool to regulate gene expression for metabolic pathway engineering. Here we implement CRISPR-Cas transcriptional activation (CRISPRa) in *P. putida* using a system previously developed in *E. coli*. We provide a methodology to transfer CRISPRa to a new host by first optimizing expression levels for the CRISPRa system components, and then applying rules for effective CRISPRa based on a systematic characterization of promoter features. Using this optimized system, we regulate biosynthesis in the biopterin and mevalonate pathways. We demonstrate that multiple genes can be activated simultaneously by targeting multiple promoters or by targeting a single promoter in a multi-gene operon. This work will enable new metabolic engineering strategies in *P. putida* and pave the way for CRISPR-Cas transcriptional programming in other bacterial species.

This work is published in Metabolic Engineering on April 4th, 2021
DOI: https://doi.org/10.1016/j.ymben.2021.04.002

2.1 Introduction

The development of microbial platforms for industrial chemical production frequently requires optimizing the expression levels of multiple genes (Lee et al., 2019; Nielsen and Keasling, 2016). The advent of CRISPR-Cas tools that can be used to rapidly program gene expression promises to accelerate pathway engineering for the efficient production of high-value compounds (Fontana et al., 2020b). The application of CRISPR-Cas tools for transcriptional repression (CRISPRi) in bacterial metabolic engineering is well-established (Banerjee et al., 2020; Kim et al., 2019; Qi et al., 2013; Tian et al., 2019; Zheng et al., 2019). By comparison, the development of CRISPR-Cas tools for programmable transcriptional activation (CRISPRa) has lagged due to the paucity of effective transcriptional activators (Dong et al., 2018), and the complexity of the rules governing CRISPRa-directed transcription from bacterial promoters (Fontana et al., 2020a). Despite these challenges, the potential for using CRISPRa to program gene expression has been demonstrated through the successful implementation in *E. coli*, *M. xanthus*, *K. oxytoca*, and *S. enterica* (Bikard et al., 2013; Dong et al., 2018; Ho et al., 2020; Liu et al., 2019; Peng et al., 2018). Determining how to strategically port CRISPRa systems into other microbes could significantly improve our metabolic engineering capabilities.

Pseudomonas putida is a gram-negative soil bacterium that has recently received attention as a potential chassis for bioproduction due to desirable metabolic capabilities and the capacity to survive harsh bioprocessing conditions (Nikel et al., 2016; Nikel and de Lorenzo, 2018) *P. putida* has high reducing power (Chavarría et al., 2013) and the ability to metabolize a broad range of feedstocks, from glucose to the toxic products of aromatic lignin degradation (Elmore et al., 2020; Johnson and Beckham, 2015; Kim et al., 2000; Loeschcke and Thies, 2015). The successful implementation of CRISPR genome editing and CRISPRi in *P. putida* (Aparicio et al., 2018; Banerjee et al., 2020; Kim et al., 2019; Tan et al., 2018; Wirth et al., 2019), shows that CRISPR gene targeting can be effective in *P. putida* and provides a starting point to assess whether gene activation with a CRISPRa system can be achieved.

CRISPR-Cas transcriptional control typically uses the catalytically inactive Cas9 protein (dCas9) with programmable guide RNAs that recognize DNA targets through Watson-Crick base pairing (Xu and Qi, 2019). Recently, we identified and optimized a variant of the transcriptional activator SoxS (R93A/S101A) that can be linked to a programmable CRISPR-Cas DNA binding domain to activate gene expression in *E. coli* (Dong et al., 2018; Fontana et al., 2020a). SoxS interacts with an interface on the α-subunit of RNA polymerase (RpoA) that is widely conserved throughout bacterial species, including in *P. putida*, suggesting that the CRISPRa system we developed in *E. coli* should also be effective in *P. putida* and other bacteria. However, in contrast to the relative permissiveness of CRISPRi (and CRISPRa in eukaryotes) (Gilbert et al., 2014; Konermann et al., 2014; Qi et al., 2013), CRISPRa in bacteria is known to be sensitive to several features of target promoters, including the precise distance from the transcription start site and the intervening sequence composition. (Fontana et al., 2020a; Ho et al., 2020; Liu et al., 2019). It is not known to what extent the rules characterized in one bacterial species are generalizable in others.

In this paper, we developed CRISPRa for programming heterologous gene expression in *P. putida* KT2440. We first constructed genetic components and established experimental

approaches to permit CRISPRa machinery developed in *E. coli* to be expressed and utilized in *P. putida*. By investigating promoter features that impact CRISPRa, such as guide RNA target sites and promoter strengths, we identified designs permitting 30- to 100-fold activation of heterologous reporter gene expression. We also demonstrated that CRISPRa can be coupled with CRISPRi for multi-gene programming and endogenous gene activation. Using an inducible system derived from *P. putida*, we have developed an inducible CRISPRa/CRISPRi platform with low leakage in the uninduced state. We showed that CRISPRa can drive the expression of heterologous genes to produce desirable metabolic products including biopterin derivatives and mevalonate. We further showed that the inducible CRISPRa system can generate 40-fold increases in mevalonate production, achieving titers comparable to those from a previously reported IPTG-inducible system. Taken together, this work provides a toolbox of components and validated workflows for implementing CRISPRa to program heterologous gene expression in *P. putida*. More broadly, these efforts establish a framework for the further development of CRISPRa tools for programming gene expression in industrially-promising bacteria.

2.2 Materials and methods

General Procedures

Plasmids pBBR1-MCS2(pBBR1-KmR), pBBR1-MCS5(pBBR1-GmR) (Kovach et al., 1995), pTNS1, pUC18T-miniTn7T-GmR (Choi and Schweizer, 2006), pRK2013, pFLP2, and *P. putida* KT2440 were a gift from the Harwood lab at the University of Washington. pRK2-AraE (Cook et al., 2018) was a gift from the Pfleger lab at the University of Wisconsin-Madison (Addgene #110141). pMVA2RBS035 (Jervis et al., 2019) was a gift from the Scrutton lab at the University of Manchester (Addgene #121051). *S. pyogenes* dCas9 (*Sp*-dCas9) was expressed from the endogenous *Sp*.pCas9 promoter and the MCP-SoxS (R93A, S101A) (abbreviated MCP-SoxS) transcriptional activator fusion protein was expressed from the BBa_J23107 promoter (Fontana et al., 2020a) (http://parts.igem.org). The modified single guide RNAs (sgRNA) (Dong et al., 2018), scaffold RNAs b2.1xMS2 (scRNAs), were expressed from the BBa_J23119 promoter in the pBBR1-GmR plasmid, unless specified. 20 bp scRNA/sgRNA target sequences are provided in Table S4. mRFP1 and sfGFP reporters were expressed from the weak BBa_J23117 minimal promoter (http://parts.igem.org), unless specified, either by integrating into the genome or in the pBBR1-GmR plasmid together with the scRNA(s). All plasmids were constructed and propagated in *E. coli* NEB turbo cells (New England Biolabs). All *P. putida* strains were constructed from the wild type strain KT2440. DNA sequences are provided in the Supplementary Material. See Table 1 for a complete list of bacterial strains and plasmid constructs.

Plasmid Construction

All PCR fragments were amplified with Phusion DNA Polymerase (Thermo-Fisher Scientific) for Infusion Cloning (Takara Bio). Transformants were cultured or selected either on Lysogeny Broth (LB) or agar plates, with appropriate antibiotics, used in the following concentrations: 100 µg/mL Carbenicillin, 25 µg/mL Chloramphenicol, 30 µg/mL Kanamycin, 30 µg/mL Gentamicin. Successful constructs were confirmed by Sanger sequencing (GENEWIZ). Details for cloning strategies are described in the Supplementary Methods and Tables S1-S3. sgRNA/scRNA target sequences are provided in Table S4.

Pseudomonas putida Strain Construction

Pseudomonas putida genome integrations were performed using the tri-parental conjugation for the mini-Tn7 method (Choi and Schweizer, 2006) or electroporation for the pGNW2 method (Wirth et al., 2019). Plasmid transformations into *P. putida* were performed either by electroporation (Choi and Schweizer, 2006) or heat-shock of $CaCl_2$ chemically competent cells (Zhao et al., 2013). Detailed methods for the preparation and transformation of chemically competent cells are described in the Supplementary Methods.

Fluorescence measurements of reporter gene expression

Fluorescence measurements of reporter gene expression were carried out either by flow cytometry or plate reader. Single colonies from LB plates were inoculated in 500 µL of EZ-RDM (Teknova) supplemented with the appropriate antibiotics and grown in 96-deep-well plates at 30 °C with shaking overnight 225 rpm. For small-molecule induction, overnight cultures were diluted 100-fold into a new culture with appropriate antibiotics and inducers, then shaken overnight at 30 °C, 225 rpm. For flow cytometry, overnight cultures were diluted 1:50 in Dulbecco's phosphate-buffered saline (PBS) and analyzed on a MACSQuant VYB flow cytometer with the MACSQuantify 2.8 software (Miltenyi Biotec) using the methods and instruments settings as described (Dong et al., 2018). For plate reader measurements, 150 µL of overnight culture were transferred into a flat, clear-bottomed black 96-well plate. OD_{600} and fluorescence values were measured in a Biotek Synergy HTX plate reader and analyzed using the BioTek Gen5 2.07.17 software. For mRFP1 detection, the excitation wavelength was 540 nm and emission wavelength was 600 nm. For sfGFP detection, the excitation wavelength was 485 nm and the emission wavelength was 528 nm. Data were plotted using Prism (GraphPad).

Mevalonate production and quantitation by GC-MS

For mevalonate production experiments, the GC-MS method was adapted from prior methods (Pfleger et al., 2007). Single colonies from LB plates were inoculated in 500 µL of EZ-RDM (Teknova) supplemented with the appropriate antibiotics and grown in 96-deep-well plates at 30 °C with shaking overnight at 225 rpm. Overnight cultures were subcultured by 1:100 dilution into 3 mL of EZ-RDM media with 1% glucose as the carbon source, supplemented with the appropriate antibiotics, and shaken at 225 rpm for 72 hours at 30 °C. After 72 hours, 560 µL of cell suspension was acidified with 140 µL of 0.5 M HCl and vortexed. 700 µL ethyl acetate was added and samples were then vortexed again vigorously for 3 minutes and centrifuged at maximum speed in a benchtop centrifuge (15,000 rcf) for 10 min. The organic phase was then transferred into GC-MS vials for analysis. GC-MS analysis was performed using an Agilent 5973 instrument with a temperature program as follows. The inlet temperature was 250 °C (splitless mode). The column flow was kept at 1 mL/min in HP-5MS (Agilent). The temperature cycle started at 80 °C and was followed by a gradient of 20 °C/min to 260 °C, a second gradient of 40 °C/min to 300 °C, and a hold at 300 °C for 2 min. m/z = 71, the second most abundant peak corresponding to mevalonolactone, was used for quantitation (Pfleger et al., 2007). A calibration curve was generated using freshly-prepared D,L-mevalonolactone (Sigma) dissolved in ethyl acetate. The calculated concentration was adjusted by the addition of HCl. Data were plotted using Prism (GraphPad).

Biopterin production and measurement

For the biopterin production experiments, single colonies from LB plates were inoculated in 500 µL of EZ-RDM supplemented with the appropriate antibiotics and grown in 96-deep-well plates at 30 °C with shaking overnight. Each sample was then sub-cultured at 100-fold dilution in 5 mL of EZ-RDM supplemented with the appropriate antibiotics and grown in 14 mL culture tubes at 30 °C and shaking for 24 hours. The overnight cultures were spun down and pteridine concentrations were determined by measuring the OD_{340} and comparing the results to a standard calibration curve prepared with purchased reagents (Cayman Chemical). The HPLC-MS measurements were performed as described (Ehrenworth et al., 2015). A detailed HPLC-MS protocol is provided in the Supplementary Methods. Data were plotted using Prism (GraphPad).

2.3 Results

Enabling CRISPRa in *P. putida*

Plasmid-based CRISPRa in P. putida

The first challenge to enable a CRISPRa system in *P. putida* is to express the components from *E. coli* in *P. putida*. The bacterial CRISPRa system developed in *E. coli* consists of three components, dCas9, MCP-SoxS, and scRNA (Dong et al., 2018), delivered in a p15A plasmid that is present at ~10 copies/cell (Shetty et al., 2008) (Figure 2.1A). The scRNA is a modified sgRNA with a 3' MS2 hairpin to recruit the MCP-SoxS activator. The reporter gene(s) were delivered in a pSC101** plasmid which is present at ~5 copies/cell (Lee et al., 2011). We used *E. coli* SoxS as the activator domain because it recognizes a motif on RpoA that is conserved between *E. coli* and *P. putida* (Dong et al., 2018), and there is no direct homolog of SoxS in *P. putida* (Park et al., 2006). To test this system in *P. putida*, the three CRISPRa components need to be expressed at levels sufficient to activate the target gene without dCas9 expression being so high that cellular functions are inhibited (Depardieu and Bikard, 2020; Zhang and Voigt, 2018). We first moved components from two *E. coli* plasmid constructs, a CRISPRa system plasmid and a reporter plasmid, directly into two *P. putida* expression plasmids, pBBR1 and pRK2 (each present at 25-30 copies/cell according to (Cook et al., 2018) (Figure 2.1B). We observed reporter gene expression that depended on the presence of an on-target scRNA (Figure 2.1C). Reporter gene expression in the presence of an off-target scRNA was indistinguishable from a strain without scRNA/sgRNA present (Figure 2.S1).

Growth-defect mitigation elevates CRISPRa efficiency

P. putida strains with the initial implementation of the CRISPRa system grew poorly on both agar and liquid media (Figure 2.S2B). To mitigate the growth defect, we tested multiple different plasmid and genome-integrated delivery methods for the CRISPRa components. We first reduced the expression levels of dCas9 and MCP-SoxS by moving these genes from the pBBR1 plasmid to the pRK2 plasmid, which expresses transgenes at a lower level in *P. putida* (Damalas et al., 2020) (Figure 2.S1). This change partially mitigated the growth defect and improved the CRISPRa reporter gene expression (Figure 2.S2). We reduced the expression levels of dCas9 and MCP-SoxS further by integrating the dCas9/MCP-SoxS cassette into the *P. putida* KT2440 genome (generating strain PPC01). We then delivered the scRNA and reporter gene cassettes on plasmids with different combinations of two origins of replication (pBBR1 and pRK2) and two

antibiotic markers (GmR and KmR) to test whether variations in the plasmid backbones impart different metabolic burdens (Mi et al., 2016). We observed the highest level of activation (~5-fold) with the scRNA and reporter both expressed from a single pBBR1-GmR backbone, while the plasmid with either pRK2 origin or KmR marker yielded weaker activation (~2-fold) (Figure 2.1C and Figure 2.S2A). The presence of the second plasmid reduced both fold-activation by CRISPRa and basal expression of mRFP significantly (Figure 2.S3). In general, both CRISPRa fold-activation and the corresponding basal mRFP expression (off-target control) increased in strains that grew faster (Figure 2.S2 & 2.S3), suggesting that there are different metabolic burdens associated with different delivery methods and plasmid expression systems. Taken together, these results suggest that optimizing expression levels will be important for implementing CRISPRa in new bacterial species. To improve *P. putida* CRISPRa beyond the five-fold activation obtained in Figure 2.1C, we proceeded with the genomically integrated dCas9/MCP-SoxS strain (PPC01) for further optimization. While there is no specific target value for fold-activation, we aim for the largest dynamic range possible to provide the highest possible tunable range in future applications.

Characterization of promoter elements for optimal CRISPRa efficiency in *P. putida*

To improve the fold-activation of CRISPRa in *P. putida*, we investigated the criteria for effective CRISPRa that we previously observed in *E. coli* (Fontana et al., 2020a). Specifically, factors known to affect CRISPRa efficiency in *E. coli* include i) the distance of target sequence to transcription start-site (TSS), ii) the sequence composition of the 20 bp scRNA targeting sequence, iii) the basal minimal promoter strength, and iv) the 5'-proximal sequence composition between target sequence and minimal promoter (Figure 2.2A).

Distance to transcription start-site (TSS)

In *E. coli*, the most effective CRISPRa target sites are in the region of -60 to -100 bp before the TSS, with sharp peaks of activity every 10 bases, separated by regions of inactivity (Fontana et al., 2020a). We constructed an integrated reporter that can be targeted at multiple sites (J1-sfGFP, previously characterized in *E. coli* (Fontana et al., 2020a) and delivered plasmids with scRNAs targeting different sites as shown in Figure 2.2B. With target sites spaced 10 bp apart, the optimal sites in *P. putida* were located in the -60 to -100 bp range before the TSS, similar to that in *E. coli*. When we tested sites at single base resolution between -81 to -93 bp, we observed peaks of activity ~10-11 bases apart, similar to what we observed in *E. coli* (Figure 2.2C). The efficiency of CRISPRa diminished after a 4-5 bp shift and was recovered at 10-12 bp. This finding suggests that CRISPRa has a periodic dependence on distance from the TSS, and similar effects have been observed in multiple bacterial species (Fontana et al., 2020a; Ho et al., 2020).

scRNA target sequences

Next, we examined the 20 bp target sequence that is recognized by the scRNA. The experiments described above were performed with the J1 promoter, which contains an array of 20 base target sites. We proceeded to test an alternative promoter, termed J3, that has a different set of 20 base target sites (see Methods and Supplemental Information for complete sequence details). We tested multiple target sites in the J3 promoter and found that the J306 site, located 81 bases upstream of the TSS, yielded the highest fold-activation (Figure 2.S4). Compared to the J1 promoter, where we observed 4-fold activation (J106 target site), the fold-activation with the

J3 promoter increased to 34-fold (J306 target site) (Figure 2.2D). For both J1 and J3, the CRISPRa-induced expression levels were similar. The large difference in fold-activation results from an unexpected difference in basal expression levels. The basal expression of J3-mRFP is 11-fold lower than that of J1-mRFP, which leads to much higher fold-activation. This difference in basal expression was not observed in *E. coli*, where J1 and J3 reporters produced 27-fold and 36-fold activation, respectively (Fontana et al., 2020a).

To test whether the different basal expression levels were due to differences in the 20 base target sites or to other features of the promoters, we constructed hybrid promoters where the 20 base J106 target site in J1 was replaced by J306 (J1(306)) and vice versa (J3(106)). We observed low basal expression only with the hybrid J3(106) promoter (Figure 2.2D), suggesting that other sequence features of the J3 promoter besides the 20 base target site are responsible for the low basal expression of the J3 promoter. These sequence features could be upstream of the target sequence or between the target sequence and the minimal BBa_J23117 promoter. We therefore turned our focus to the J3 upstream sequence as a basis for further optimization of CRISPRa, as it yielded the best dynamic range from the promoter sequences tested.

Minimal promoter strength

The promoters that we tested in this work consist of a 35 base minimal promoter that binds the sigma subunit of RNA polymerase and an upstream 170 base sequence region with scRNA target sites. The 35 bp minimal promoter sequence is also a key factor that governs the dynamic range of CRISPRa. In *E. coli*, we found that minimal promoter strength and the sigma factor regulating the promoter have large effects on CRISPRa (Fontana et al., 2020a). However, the alternative sigma factor regulons in *P. putida* are less characterized compared to those in *E. coli*. We therefore decided to focus on the sigma-70 regulon, the house-keeping sigma factor, that covers the majority of *E. coli* and *P. putida* endogenous promoters (Fujita et al., 1995). To test the effects of promoter strength, we introduced 11 minimal 35 base promoters from the Anderson promoter collection (BBa_J231XX, http://parts.igem.org) into the J3-mRFP reporter (Figure 2.3A). The variations in promoter strength arise from point mutations in the -10 and -35 sites that tune transcriptional activity; no significant changes in the transcription start sites (TSS) were detected when these promoters were experimentally characterized (Kosuri et al., 2013) (see Supplemental Information for annotated sequences).

CRISPRa-mediated expression from the Anderson promoter series followed trends similar to that previously observed in *E. coli* (Fontana et al., 2020a). When the promoter strengths are extremely weak (BBa_J23109 and BBa_J23113), the CRISPRa fold-activation dropped significantly to 3.1-fold and 1.4-fold compared to 27-fold with the moderately weak BBa_J23117 minimal promoter. As promoter strength increases from BBa_J23117 to the strong BBa_J23110 promoter, CRISPRa fold-activation decreases because basal expression increases ~10-fold, while the maximal CRISPRa output varies by <4-fold (Figure 2.3A). CRISPRa with the strongest promoter tested (BBa_J23111), could not be measured because no colonies were obtained when the CRISPR machinery was delivered to *P. putida* with this reporter, possibly due to the metabolic burden of expressing high levels of mRFP and the CRISPR system at the same time. The minimal BBa_J23117 promoter yields the highest fold-activation in both *E. coli* and *P. putida* (36-fold and 27-fold, respectively) presumably because basal expression is weak enough that significant activation is possible, but not so weak that the promoter is difficult to activate. Thus, we used

reporters with the BBa_J23117 minimal promoter for further characterization and application. We note that if a higher absolute expression level is preferred, the stronger BBa_J23106 promoter yielded the highest absolute expression level (2.4-fold higher than CRISPRa-mediated activation of BBa_J23117), although the fold-activation was smaller (Figure 2.3A).

5'-Proximal sequences

The last factor we tested is the intervening sequence between the 20 base target site and the 35 base minimal promoter, termed the 5'-proximal sequence. This sequence is 26 bp long when using an optimal target site located at -81 bp from the TSS. We constructed a pooled library of reporter gene plasmids with variable 26 base 5' proximal sequences using a randomized oligo pool (see Supplemental Methods). Each reporter retains the same 20 base J306 scRNA target site and the BBa_J23117 minimal promoter. We transformed this library into a *P. putida* reporter strain and functionally characterized a large number of single colonies without sequencing each colony. The random 5'-proximal sequences led to a broad range of mRFP levels from CRISPRa (Figure 2.3B), similar to what we observed in *E. coli* (Fontana et al., 2020a). Random 5' proximal sequences also affect basal expression levels in the absence of CRISPRa, although these effects are relatively small (Figure 2.S5A). To determine if 5'-proximal sequence preferences are correlated between *E. coli* and *P. putida*, we also tested several known sequences previously characterized in *E. coli*. We observed that high-efficiency 26 bp sequences from *E. coli* yield high CRISPRa efficiency in *P. putida* while a weak sequence from *E. coli* remains weak in *P. putida* (Figure 2.3C). Across the set of sequences we analyzed, one of these (5'-PS5) exhibited a higher fold activation (32-fold) compared to the J3-BBa_J23117 promoter (27-fold). The basal expression in 5'-PS5 is 15% lower than J3-BBa_J23117, and both sequences gave similar activated levels. We also tested whether the 26 bp 5' proximal sequence from the J1 promoter was responsible for the high basal activity of the J1 promoter (Figure 2.2D). When the 26 bp 5' proximal sequence from J1 was inserted into the J3 promoter, we observed relatively low basal expression (5'-PS2 in Figure 2.3C), similar to the J3 promoter. This result suggests that the 5' proximal sequence of J1 is not the cause of the high basal activity of the complete J1 promoter, and that sequence features upstream of the 5' proximal sequence and the 20 bp target site could be responsible.

The variation in CRISPRa outputs with different promoter features suggests that a set of distinct and orthogonal heterologous promoters could be developed for tunable control of gene expression. Promoters with orthogonal 20 base target sequences, together with different 5' proximal sequences, minimal promoters, and target site positions could be used to access a broad range of CRISPRa-mediated gene expression levels. Further, systematically varying the 5'-proximal sequence could allow us to identify promoters with lower basal expression and higher dynamic range of activation, similar to the case of the 5'-PS5 sequence mentioned above. We expect to be able to construct combinatorial libraries of multi-gene programs to explore how independently tuning gene expression levels in metabolic pathways affects product titers.

Correlation of CRISPRa efficiency between organisms

To compare CRISPRa in *P. putida* to that in *E. coli*, we constructed a correlation plot of mRFP expression from CRISPRa strains with different promoter sequence variations (Figure 2.3D). This plot indicates that the expression level induced by CRISPRa in *E. coli* correlates well

with CRISPRa in *P. putida* (R^2 = 0.80). The fold-activation is also correlated (R^2 = 0.69 Figure 2.S5B), although the fold-activation of *P. putida* CRISPRa tends to be lower than that of *E. coli*. The discrepancies across organisms might arise from variations in genetic context, transcription machinery, or cellular compositions between bacterial species. Despite these modest discrepancies, CRISPRa behaves largely similarly in *E. coli* and *P. putida*, suggesting that optimized CRISPRa circuits will be portable between species and that further modifications and improvements to CRISPRa systems should be readily transferable. While we do not expect these trends to be generalizable across all bacterial species, the metrics that we describe here can be systematically evaluated in alternative bacterial hosts to assess whether design principles and optimized CRISPRa circuits can be easily ported to new hosts.

Using *P. putida* CRISPRa for sophisticated transcriptional control strategies

With an optimized CRISPRa system in *P. putida*, we explored several strategies to enable more sophisticated control over gene expression programs. We constructed multi-gene CRISPRa/CRISPRi programs, demonstrated endogenous gene activation, and developed an inducible CRISPRa system for tunable, dynamically-regulated expression. These strategies will enable the construction of multi-gene programs to rewire metabolic networks for optimal biosynthesis in *P. putida*.

Multi-gene regulation by CRISPRa and CRISPRi

With optimized expression levels and a delivery strategy for the CRISPRa system in *P. putida* in place, we tested whether CRISPRa and CRISPRi can be used together to activate and repress multiple genes. This strategy has been previously successful in *E. coli* (Dong et al., 2018). We constructed a dual-reporter plasmid with weakly expressed mRFP (J3-BBa_J23117-mRFP) and highly expressed sfGFP (J3(106)-BBa_J23111-sfGFP). We inserted a dual scRNA/sgRNA cassette in this plasmid with a J306 scRNA for mRFP activation and an sgRNA that targets within the sfGFP open reading frame (ORF) for repression. We delivered this plasmid to a *P. putida* strain with integrated dCas9/MCP-SoxS and observed simultaneous activation of mRFP (6.6-fold) and repression of sfGFP (13-fold) (Figure 2.4). The magnitude of CRISPRa fold-activation in simultaneous CRISPRa/i was weaker than the 15-fold activation that was observed if just a single scRNA was delivered to activate the mRFP reporter, possibly due to competition between multiple scRNA/sgRNA cassettes for a limited pool of dCas9.

To determine if CRISPRa can be used to activate multiple genes simultaneously, we constructed a dual-reporter plasmid with weakly expressed mRFP (J3-BBa_J23117-mRFP) and weakly expressed sfGFP (J3(106)-BBa_J23117-sfGFP). We inserted a dual scRNA cassette into this plasmid with scRNAs that target mRFP and sfGFP for activation and delivered it to a *P. putida* strain with integrated dCas9/MCP-SoxS. We observed simultaneous activation of mRFP (19-fold) and sfGFP (69-fold) (Figure 2.S6). As seen with simultaneous CRISPRa/CRISPRi, the CRISPRa effects with dual activation were weaker than those observed if each reporter was targeted individually (41-fold activation for mRFP and 105-fold activation for sfGFP), consistent with the idea that competition for dCas9 among multiple species of sgRNA/scRNA may be an issue for multi-gene programs (Huang et al., 2020). We also observed simultaneous CRISPRa at multiple genes using the weak mRFP/strong sfGFP reporter described above; the strong sfGFP could be activated a further 2-3-fold when activated by an upstream scRNA (Figure 2.4).

We also demonstrated simultaneous CRISPRa/CRISPRi and dual CRISPRa on multi-gene reporters with integrated genomic reporters. The general trends were similar to what we observed with plasmid-based reporters (Figure 2.S7-2.S8), but the magnitudes of the effects were smaller, likely due to the lower copy number of the reporter gene. The ability to activate genomically-integrated heterologous reporters suggests that CRISPRa may be effective at endogenous genomic targets in *P. putida*.

CRISPRa on P. putida endogenous promoters

To determine if CRISPRa can activate endogenous promoters, we identified a set of endogenous genes with appropriate upstream scRNA target sites. We analyzed thousands of reported TSSs for *P. putida* (D'Arrigo et al., 2016) and selected ten promoters with potentially activatable target sites located at the proper distance from the TSS. Specifically, we identified NGG protospacer adjacent motifs (PAMs), which are required for recognition of *Sp*-dCas9/guide-RNA complex (Qi et al., 2013), at distances corresponding to the J105-J112 target sites (Figure 2.2B) with ± 2 bp flexibility (Figure 2.2C). For each endogenous promoter, we built a reporter cassette with the promoter, flanking sequences, and an mRFP reporter gene following a strategy previously described for an *E. coli* endogenous promoter library (Zaslaver et al., 2006). We introduced on-target or off-target scRNAs into the reporter plasmid and delivered it to a *P. putida* strain with integrated dCas9/MCP-SoxS. We observed >1.5-fold activation at 4 of the 10 promoters tested, with the highest fold-activation (2.8-fold) from scRNA G2 targeting *katG* (PP_3668) promoter (Figure 2.5A & 2.S9). The magnitudes of fold-activation from endogenous promoters are significantly lower than those under control of synthetic heterologous promoters (up to 40-fold and 100-fold for mRFP and sfGFP, respectively) (Table S5). Although higher fold-activation values may be desirable for future applications, we note that relatively modest effects can still be physiologically significant. For example, external stresses can produce a wide range of expression changes in stress-responsive genes in *P. putida*. While some changes are quite large, others are in the 2-fold to 5-fold range (Bojanovič et al., 2017; Molina-Santiago et al., 2017). We suggest that tools to perturb endogenous gene expression in this range may still be effective for modulating bacterial physiology and redirecting metabolic flux. Further, we note that the ability to combine endogenous gene activation with heterologous gene activation and CRISPRi repression enables access to a vastly expanded space of gene expression programs compared to other synthetic gene regulatory methods.

This success rate and the magnitude of gene activation at endogenous targets in *P. putida* was similar to that observed previously in *E. coli* (Fontana et al., 2020a). To predictably activate any endogenous gene, we expect that it will be necessary to further elucidate the rules for effective CRISPRa. Accurate annotations of TSSs and PAM-flexible dCas9 variants to precisely target the optimal distance upstream of the endogenous gene may improve activation (Fontana et al., 2020a). Alternative bacterial activation domains are also available with different properties (Ho et al., 2020; Liu et al., 2019), and it may be possible to combine multiple activators as has been previously reported in eukaryotic systems (Chavez et al., 2015; Konermann et al., 2014).

Tunability of CRISPRa and CRISPRi with inducible promoter

To tune expression levels with CRISPRa and CRISPRi, we placed the CRISPR system components under the control of a small-molecule inducible promoter. We expressed dCas9 and/or MCP-SoxS using XylS-Pm, an inducible promoter system from the *P. putida* mt-2 toluene degradation pathway (Wirth et al., 2019). XylS-Pm provides a higher dynamic range compared to the widely-used LacI-Ptrc system (Figure 2.S10A & 2.S10B). We constructed strains with inducible dCas9, inducible MCP-SoxS, or double-inducible dCas9/MCP-SoxS (PPC08-PPC10). Using a weak J3-BBa_J23117-mRFP reporter, we induced with *m*-toluic acid (0-5 mM) and observed tunable gene activation as a function of inducer concentration in all three inducible strains (Figure 2.S11). This approach will enable tunable and dynamically-regulatable expression control for further applications in metabolic engineering.

Using a strong reporter (J3-BBa_J23110-mRFP) that can be either activated or repressed, we showed that the extent of CRISPRa or CRISPRi could be tuned with different inducer levels. We delivered this reporter with either an activating scRNA or a repressing sgRNA to the inducible dCas9 strain (PPC08) and observed 3-fold activation with CRISPRa or 7-fold repression with CRISPRi at 1 mM *m*-toluic acid (Figure 2.5B). This result suggests another potential strategy for improving the dynamic range of activation from heterologous genes. By targeting CRISPRi and CRISPRa to the same locus, we may be able to obtain lower basal expression and higher induced expression. Such a strategy would require expression of only the sgRNA for repression in the off state and only the scRNA for activation in the on state, which could potentially be achieved with orthogonal induction systems or with multi-layer CRISPR circuits.

Metabolic Engineering with CRISPRa

Biopterin pathway activation

By characterizing the promoter features necessary for effective CRISPRa in *P. putida*, we were able to apply CRISPRa for metabolic pathway engineering. We used the J3-BBa_J23117 promoter described in the previous section to place genes of interest under the control of a CRISPRa system. In a strain with integrated dCas9/MCP-SoxS (PPC01), transcriptional units controlled by J3-BBa_J23117 can be activated by the cognate J306 scRNA (Figure 2.6A). Using this approach, we demonstrated that CRISPRa can be used to switch on two different heterologous biosynthesis pathways, one for tetrahydrobiopterin (BH4) production with multiple transcriptional units activated by the same scRNA and one for mevalonate production as a multi-gene transcriptional unit under a single promoter.

BH4 is an important cofactor in aromatic amino acid biosynthesis that can be produced from a three-enzyme pathway (Figure 2.6B). BH4 has been previously produced in yeast using the *E. coli* GTPCH enzyme and the *M. alpina* PTPS and SR enzymes (Ehrenworth et al., 2015; Trenchard et al., 2015). We used the *gtpch* gene from *E. coli* MG1655 and *ptps/sr* genes from *M. alpina* that were codon-optimized for expression in *E. coli*. Each gene was placed under control of the J3-BBa_J23117 promoter in a *P. putida* compatible plasmid (Figure 2.6C). Because BH4 can be readily oxidized by atmospheric oxygen into dihydrobiopterin (BH2) and then biopterin in yeast (Ehrenworth et al., 2015), we initially screened for pathway output by absorbance at 340 nm, which reports on BH2 and biopterin. We observed a significant increase in OD_{340} when the pathway was switched on with the cognate scRNA (Figure 2.S12A). Subsequent analysis by

HPLC-MS to identify specific parental ions confirmed that BH2 is the major product (Figure 2.6D & FIgure 2.S12B and Figure 2.S13). BH2 was also detected in the off-target scRNA sample (Figure 2.6D), likely due to basal expression of the biopterin pathway enzymes. When the last gene in the pathway (*sr*) was omitted, no biopterin derivatives were detected by HPLC-MS, confirming that the full pathway is necessary for heterologous biopterin production (Figure 2.S12A-B). Thus, biopterin pathway activation by CRISPRa was able to significantly increase heterologous production. We note that in some metabolic engineering applications, basal production may be problematic and pathway promoters may need to be modified to minimize leaky expression of the heterologous pathway genes. In future experiments, the CRISPRa system can be used to test whether product titers can be further optimized by independently activating biopterin pathway genes with orthogonal scRNAs and tuning their expression to different levels.

The major product of the biopterin pathway in *P. putida* is BH2, in contrast to *S. cerevisiae* where fully oxidized biopterin is the major product (Ehrenworth et al., 2015). The finding that BH2 is the major product suggests that the reducing potential of *P. putida* prevented BH2 from further oxidation. In *E. coli*, BH2 is the major product but the ratio of BH2:biopterin is significantly lower than in *P. putida* (Figure 2.S12C). Even though the fully reduced BH4, which is the desired product, was not observed in our system, the low biopterin level in *P. putida* suggests that its reducing power is advantageous for biosynthesis of oxidation-sensitive compounds.

Mevalonate pathway activation

We next tested if CRISPRa could be used to produce mevalonic acid, a precursor to terpenoid natural products including fine chemicals, biofuels, and therapeutics (Anthony et al., 2009; Jervis et al., 2019; Peralta-Yahya et al., 2011) Mevalonate has previously been produced in *P. putida* using two genes, *mvaE* and *mvaS*, expressed in a single operon under the control of LacI-Ptrc (Figure 2.7A) (Kim et al., 2019). We placed the *mvaES* operon under the control of J3-BBa_J23117 synthetic promoter (Figure 2.7B). The constitutively-active CRISPRa-regulated mevalonate production strain was cultured side-by-side with the LacI-Ptrc regulated *mvaES* strain as a control. We observed that the CRISPRa strain yielded 402 ± 21 mg/L mevalonate, which is similar to the highest mevalonate titer of 459 mg/L obtained with LacI-Ptrc after IPTG induction (Figure 2.7C). The CRISPRa-regulated *mvaES* operon enables tight control of mevalonate production, with basal mevalonate production from the off-target CRISPRa control strain indistinguishable from the empty plasmid control (Figure 2.7C). In contrast, the uninduced LacI-Ptrc strain produced mevalonate levels up to 214 ± 57 mg/L and yielded highly variable mevalonate levels in every IPTG concentration (ranging from 66 to 459 mg/L). We also observed highly variable IPTG-induced mRFP expression, suggesting that expression from the LacI-Ptrc promoter may be unstable in *P. putida* (Figure 2.S10C). Taken together, our results demonstrate that we can effectively activate multi-gene biosynthesis pathways using a single operon (>40-fold increase in mevalonate biosynthesis, Figure 2.7) or with each enzyme produced from a separate transcriptional unit with its own CRISPRa-responsive promoter (5-fold increase in BH2 production, Figure 2.6).

To determine if an inducible CRISPRa system could effectively regulate mevalonate production, we tested a strain with toluic acid-inducible CRISPRa machinery (dCas9, MCP-SoxS, or both). In the absence of inducer we observed 84 ± 11 mg/L mevalonate from the inducible dCas9 strain. With inducer added to this strain (0.01 to 1.0 mM), we observed a similar

mevalonate level to that with constitutively expressed dCas9 (345 to 397 mg/L and 402 ± 21 mg/L, respectively) (Figure 2.7). The inducible MCP-SoxS strain appeared to be leaky in the absence of inducer (112 ± 2 mg/L) and gave a lower mevalonate titer when induced (254 ± 9 mg/L). The double-inducible strain, with both dCas9 and MCP-SoxS controlled by XylS-Pm, had no significant leaky production in the absence of inducer but yielded the lowest mevalonate titer (199 ± 20 mg/L). The off-target scRNA yielded a level of mevalonate indistinguishable from the empty plasmid controls (less than 10 mg/L in Figure 2.S14). The inducible CRISPRa system provides an additional layer of control that can be switched on at different growth phases and could be coupled with an inducible CRISPRi system for multi-gene programs with both activation and repression. Compared to the LacI-Ptrc regulated *mvaES* strain, which showed significant leaky production, the inducible dCas9 CRISPRa-regulated *mvaES* strain had minimal leakage and could provide advantages in situations where leaky metabolic gene expression could be toxic or burdensome to the cell.

2. 4 Conclusions

In this work, we have ported a CRISPRa system from *E. coli* to *P. putida*. We optimized the expression methods of dCas9, MCP-SoxS, and scRNA in *P. putida* and demonstrated that the criteria for effective CRISPRa target sites in *P. putida* are similar to that in *E. coli*. We anticipate that a similar process of optimizing expression systems will enable effective CRISPRa-regulated gene expression in a wide range of bacterial species to enable complex CRISPR-based transcriptional programming in other industrially-relevant microbes.

As reported previously in *E. coli* and in many eukaryotic systems, CRISPRa and CRISPRi can be used to target multiple genes simultaneously for activation or repression. Further, the CRISPRa system can be induced with small molecules, which will enable dynamic control of heterologous pathway activation. In *P. putida*, we applied CRISPRa to metabolic pathway engineering for tetrahydrobiopterin and mevalonate biosynthesis, providing a proof-of-concept that CRISPRa-mediated gene regulation can be used to activate heterologous biosynthetic pathways.

In future work, we expect that an inducible CRISPR-Cas transcriptional control system will enable the rapid exploration of large combinatorial spaces of gene expression levels. A key advantage of CRISPR-Cas-mediated control is that, in principle, each gene of interest can be targeted by an orthogonal guide RNA and its expression level can be independently tuned. We can target endogenous genes for both activation and repression to redirect metabolic flux towards the desired pathway precursors, and we can tunably activate heterologous pathways to optimal expression levels to maximize the production of desired biosynthetic products. By learning the design principles for how to rewire metabolic networks, we expect to enable more efficient biosynthetic production pathways for valuable chemical products.

Author contributions

C.K., C.D., J.F., P.P.-Y., J.M.C., and J.G.Z. designed experiments and analyzed data. C.K., C.D., J.F., and W.S. performed experiments. C.K., C.D., J.F., J.M.C., and J.G.Z. wrote the manuscript with input from all of the authors.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgments

We thank Mary Lidstrom, Caroline Harwood, Amy Schaefer, Martin Sadilek, Luke Zhu, Nick Kruyer and members of the Zalatan, Carothers and Peralta-Yahya groups for technical assistance, advice, and helpful discussions. This work was supported by NSF Award 1817623 (J.M.C, J.G.Z.) and NSF Award 1844152 (J.M.C., P.P.Y) and DOE Award DE-EE0008927 (J.M.C, J.G.Z.)

2.5 Tables and Figures

Table 2.1: Bacterial strains and plasmids used in this study

Strains/Plasmids	Features	Sources
Strains		
P. putida KT2440	Wildtype strain	Harwood lab
PPC01	KT2440 with integrated *Sp*.pCas9-dCas9 and BBa_J23107-MCP-SoxS made from pPPC001	This study
PPC02	KT2440 with integrated J1-BBa_J23117-sfGFP, *Sp*.pCas9-dCas9, and BBa_J23107-MCP-SoxS, made from pPPC002	This study
PPC03.N	KT2440 with integrated J1(+N)-BBa_J23117-sfGFP, *Sp*.pCas9-dCas9, and BBa_J23107-MCP-SoxS, made from pPPC003.N	This study
PPC04	KT2440 with integrated BBa_J23111-mRFP, BBa_J1-J23117-sfGFP, *Sp*.pCas9-dCas9, and BBa_J23107-MCP-SoxS, made from pPPC004	This study
PPC05	PPC01 with integrated J3-BBa_J23117-mRFP and BBa_J23111-sfGFP, made from pPPC031 and pPPC032	This study
PPC06	PPC01 with integrated J3-BBa_J23117-mRFP and J3(106)-BBa_J23117-sfGFP, made from pPPC031 and pPPC033	This study
PPC07	PPC01 with integrated J3-BBa_J23117-mRFP and J3(106)-BBa_J23111-sfGFP, made from pPPC031 and pPPC034	This study
PPC08	KT2440 with integrated XylS-Pm-dCas9, BBa_J23107-MCP-SoxS made from pPPC005	This study
PPC09	KT2440 with integrated *Sp*.pCas9-dCas9, XylS-Pm-MCP-SoxS made from pPPC006	This study
PPC10	KT2440 with integrated XylS-Pm-dCas9, XylS-Pm-MCP-SoxS made from pPPC007	This study
Plasmids		
pUC18T-miniTn7T-Gm	Plasmid backbone for integration into *P. putida* genome, GmR/AmpR	(Choi and Schweizer, 2006)
pTNS1	Tn7 transposase (tnsABCD) expressing plasmid, R6K origin of replication, AmpR	(Choi and Schweizer, 2006)
pRK2013	Helper plasmid for triparental mating, KmR	(Choi and Schweizer, 2006)
pFLP2	*S. cerevisiae* Flippase expression plasmid for marker deletion, AmpR	(Choi and Schweizer, 2006)
pBBR1-MCS5 (pBBR1-GmR)	Broad-host-range plasmid backbone with multiple cloning site, GmR	(Kovach et al., 1995)
pBBR1-MCS2 (pBBR1-KmR)	Broad-host-range plasmid backbone with multiple cloning site, KmR	(Kovach et al., 1995)

pRK2-AraE	Broad-host-range plasmid backbone with AraE expressing cassette, GmR	(Cook et al., 2018)
pRK2-GmR	Broad-host-range plasmid backbone with multiple cloning site, GmR	This study
pRK2-KmR	Broad-host-range plasmid backbone with multiple cloning site, KmR	This study
pGNW2	Integrative vector carrying P14g-msfGFP, KmR	(Wirth et al., 2019)
pS448-CsR	CRISPR/Cas9 counterselection in Gram-negative bacteria with XylS/Pm promoter, SmR	(Wirth et al., 2019)
pSEVA1213S	pRK2, PEM7-I-SceI; AmpR	(Wirth et al., 2019)
pGNW2-pp1	pGNW2 derivative with integration site at prophage1, KmR	This study
pGNW2-pp2	pGNW2 derivative with integration site at prophage2, KmR	This study
pCK241	pBBR1 bearing LacI-Ptrc-mRFP, GmR	This study
pCK243	pBBR1 bearing XylS-Pm-mRFP, GmR	This study
pCK255	pBBR1 bearing I-SceI and *sacB* genes, GmR	This study
pMVA2RBS035	p15A, LacI-Ptrc *mvaE*, *mvaS*, *mvaK1*, *mvaK2*, and *mvaD* from *E. faecalis*, and *idi* gene from *E. coli*, KmR	(Jervis et al., 2019)
pCD442	p15A, *Sp*.pCas9-dCas9, BBa_J23107-MCP-SoxS, CmR	(Fontana et al., 2020a)
pPPC001	pUC18T-miniTn7T, *Sp*.pCas9-dCas9, BBa_J23107-MCP-SoxS, AmpR/GmR	This study
pPPC002	pUC18T-miniTn7T, J1-BBa_J23117-sfGFP, *Sp*.pCas9-dCas9, and BBa_J23107-MCP-SoxS, AmpR/GmR	This study
pPPC003.N	pUC18T-miniTn7T, J1(+N)-BBa_J23117-sfGFP, *Sp*.pCas9-dCas9, and BBa_J23107-MCP-SoxS, AmpR/GmR	This study
pPPC004	pUC18T-miniTn7T, BBa_J23111-mRFP, J1-BBa_J23117-sfGFP, *Sp*.pCas9-dCas9, and BBa_J23107-MCP-SoxS, AmpR/GmR	This study
pPPC005	pUC18T-miniTn7T, XylS-Pm-dCas9, BBa_J23107-MCP-SoxS, AmpR/GmR	This study
pPPC006	pUC18T-miniTn7T, *Sp*.pCas9-dCas9, XylS-Pm-MCP-SoxS, AmpR/GmR	This study
pPPC007	pUC18T-miniTn7T, XylS-Pm-dCas9, XylS-Pm-MCP-SoxS, AmpR/GmR	This study
pPPC008	pBBR1, sgRNA or scRNA, GmR	This study
pPPC009	pBBR1, sgRNA or scRNA, KmR	This study
pPPC010	pBBR1, *Sp*.pCas9-dCas9, BBa_J23107-MCP-SoxS, scRNA, KmR	This study
pPPC011	pRK2, *Sp*.pCas9-dCas9, BBa_J23107-MCP-SoxS, KmR	This study
pPPC012	pBBR1, J1-BBa_J23117-mRFP, GmR	This study
pPPC013	pBBR1, J1-BBa_J23117-mRFP, KmR	This study
pPPC014	pRK2, J1-BBa_J23117-mRFP, GmR	This study
pPPC015	pRK2, J1-BBa_J23117-mRFP, KmR	This study

pPPC016	pBBR1, J1-BBa_J23117-mRFP, scRNA, GmR	This study
pPPC016(306)	pBBR1, J1-BBa_J23117-mRFP where J106 was replaced with J306, scRNA, GmR	This study
pPPC017	pBBR1, J1-BBa_J23117-mRFP, scRNA, KmR	This study
pPPC018	pRK2, J1-BBa_J23117-mRFP, scRNA, GmR	This study
pPPC019	pRK2, J1-BBa_J23117-mRFP, scRNA, KmR	This study
pPPC020	pBBR1, J3-BBa_J23117-mRFP, scRNA, GmR	This study
pPPC020(106)	pBBR1, J3-BBa_J23117-mRFP where J306 was replaced with J106, scRNA, GmR	This study
pPPC021.J231XX	pBBR1, J3-BBa_J231XX-mRFP, scRNA, GmR	This study
pPPC022.5PS	pBBR1, J3-Random-5PS-BBa_J23117-mRFP, scRNA-J306, GmR	This study
pPPC023.5PSN	pBBR1, J3-Ec-5PS-BBa_J23117-mRFP, scRNA, GmR	This study
pPPC024	pBBR1, J3(106)-BBa_J23111-sfGFP, J3-BBa_J23117-mRFP, scRNA, GmR	This study
pPPC025	pBBR1, J3(106)-BBa_J23117-sfGFP, J3-BBa_J23117-mRFP, scRNA, GmR	This study
pPPC026.XN	pBBR1, PP_NNNN-mRFP, scRNA, GmR where PP_NNNN is an endogenous promoter	This study
pPPC027	pBBR1, J3-BBa_J23117-GTPCH, J3-J23117-PTPS, J3-J23117-SR, scRNA, GmR	This study
pPPC028	pBBR1, J3-BBa_J23117-GTPCH, J3-J23117-PTPS, scRNA, GmR	This study
pPPC029	pBBR1, LacI-Ptrc-*mvaES*, GmR	This study
pPPC030	pBBR1, J3-BBa_J23117-*mvaES*, scRNA, GmR	This study
pPPC031	pGNW2 derivative with integration site at prophage1 for integration of J3-BBa_J23117-mRFP cassette, KmR	This study
pPPC032	pGNW2 derivative with integration site at prophage2 for integration of BBa_J23111-sfGFP, KmR	This study
pPPC033	pGNW2 derivative with integration site at prophage2 for integration of J3(106)-BBa_J23117-sfGFP, KmR	This study
pPPC034	pGNW2 derivative with integration site at prophage2 for integration of J3(106)-BBa_J23111-sfGFP, KmR	This study

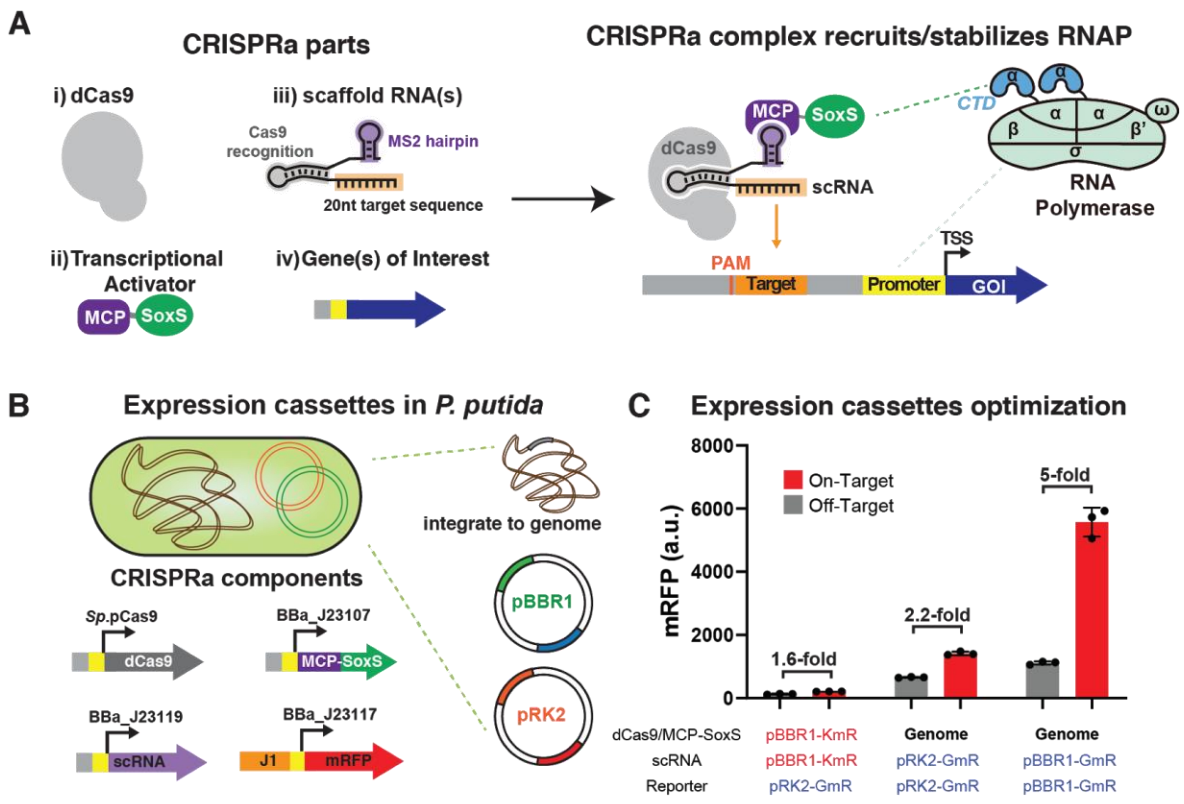


Figure 2.1. Configuring CRISPRa in *P. putida*

(A) CRISPRa components (i-iii) are necessary to activate the gene of interest (iv). The CRISPRa ternary complex recruits and stabilizes RNA polymerase at the promoter region. (B) Available gene expression tools in *P. putida* include pBBR1 plasmid, pRK2 plasmid, and genome integration. We used two antibiotic selection markers, Gentamicin (GmR) and/or Kanamycin (KmR). (C) Testing CRISPRa in different expression systems. The CRISPRa fold-activation is highest when dCas9/MCP-SoxS were integrated into the genome and the scRNA/reporter genes were expressed on pBBR1-GmR plasmid. The J109 scRNA was used for activation and hAAVS1 was used as an off-target scRNA. Values in panel C represent the mean ± standard deviation calculated from $n = 3$ independent biological replicates.

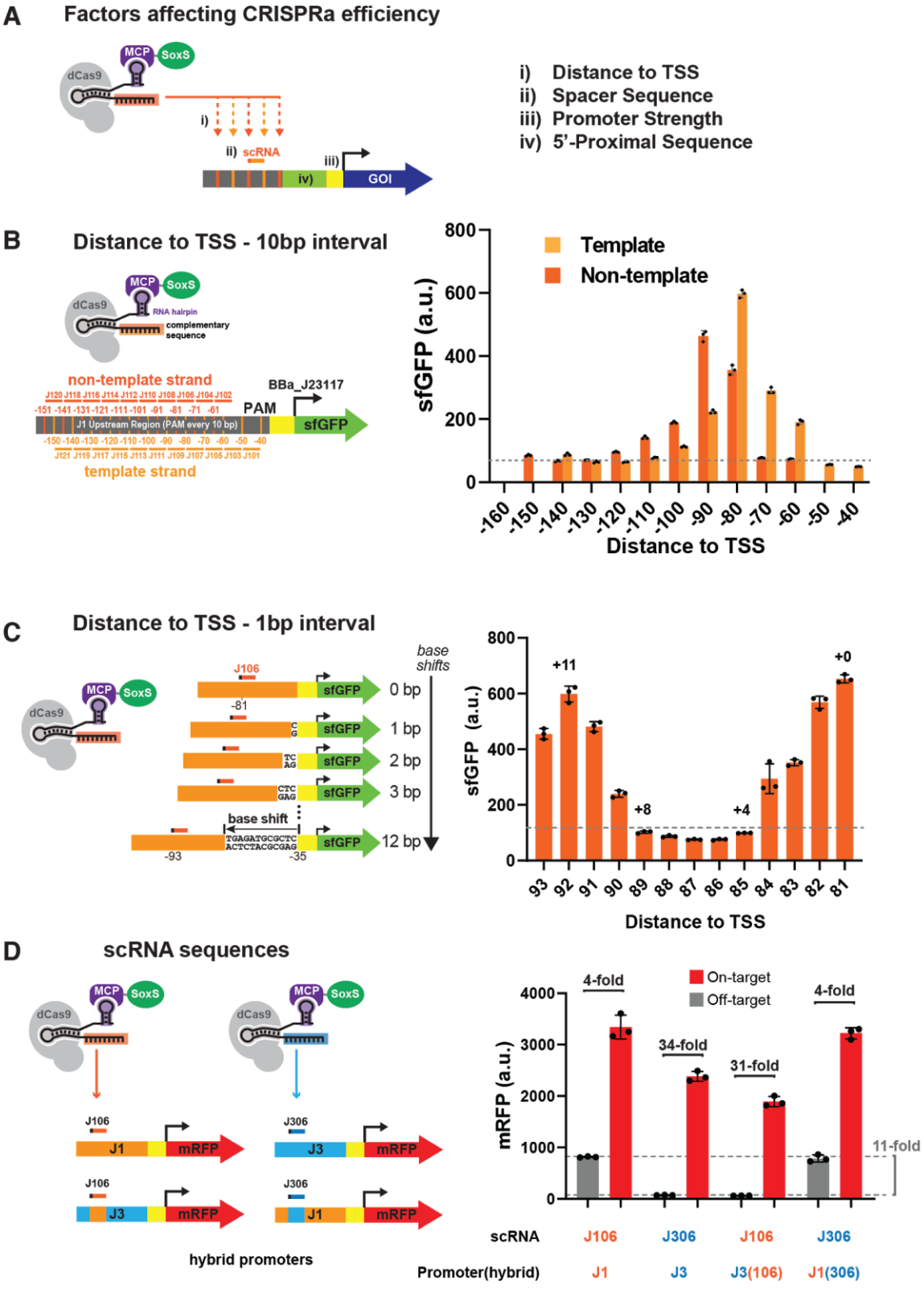

A
Factors affecting CRISPRa efficiency
MCP
SoxS
dCas9
i)
ii) scRNA
iii)
iv)
GOI
i) Distance to TSS
ii) Spacer Sequence
iii) Promoter Strength
iv) 5'-Proximal Sequence
B
Distance to TSS - 10bp interval
RNA hairpin
complementary sequence
non-template strand
BBa_J23117
PAM
sfGFP
J1 Upstream Region (PAM every 10 bp)
template strand
Template
Non-template
sfGFP (a.u.)
Distance to TSS
C
Distance to TSS - 1bp interval
base shifts
J106
-81
0 bp
1 bp
2 bp
3 bp
base shift
12 bp
-93
-35
+11
+0
+8
+4
D
scRNA sequences
J306
J1
J3
mRFP
hybrid promoters
mRFP (a.u.)
4-fold
On-target
Off-target
34-fold
31-fold
11-fold
scRNA
Promoter(hybrid)
J1
J3
J3(106)
J1(306)

Figure 2.2. Sensitivity of CRISPRa to distance from the TSS and promoter sequence composition in *P. putida*

(A) Factors known to affect CRISPRa efficiency in *E. coli* include i) distance to TSS, ii) scRNA target sequence, iii) minimal promoter strength, and 5'-proximal sequence between target sequence and minimal promoter. (B) Effect of distance to TSS on CRISPRa efficiency at 10 bp resolution. The J1 synthetic sequence upstream of the minimal promoter includes target sites every 10 bp in both the template strand (light orange), and the non-template strand (orange). scRNAs J101-J121 were expressed in the pBBR1-GmR backbone. The observed peaks of activation are slightly offset on the template and non-template strands because the distance is defined from the TSS to the PAM sites, which is proximal to the TSS on template strand targets and distal to the TSS on non-template strands. The most effective sites at -91 on the template strand (J108) and -80 on the non-template strand (J109) target overlapping 20-base sites. (C) Effect of distance to TSS on CRISPRa efficiency at single bp resolution. N bases were added upstream of the minimal promoter (N = 1 - 12), and the J106 scRNA was used to target sites at -81 to -93 upstream of the TSS. (D) The J3 upstream sequence has lower basal expression (11-fold) and higher fold-activation by CRISPRa than the J1 sequence. When the 20 bp target sequence J106 was inserted into the J3 promoter, the basal expression remains low. When the 20 bp target sequence J306 was inserted into the J1 promoter, basal expression remains high. See Methods and Supplemental Information for detailed descriptions of the J1 and J3 sequences. Values in panel B, C, and D represent the mean ± standard deviation calculated from $n = 3$ independent biological replicates.

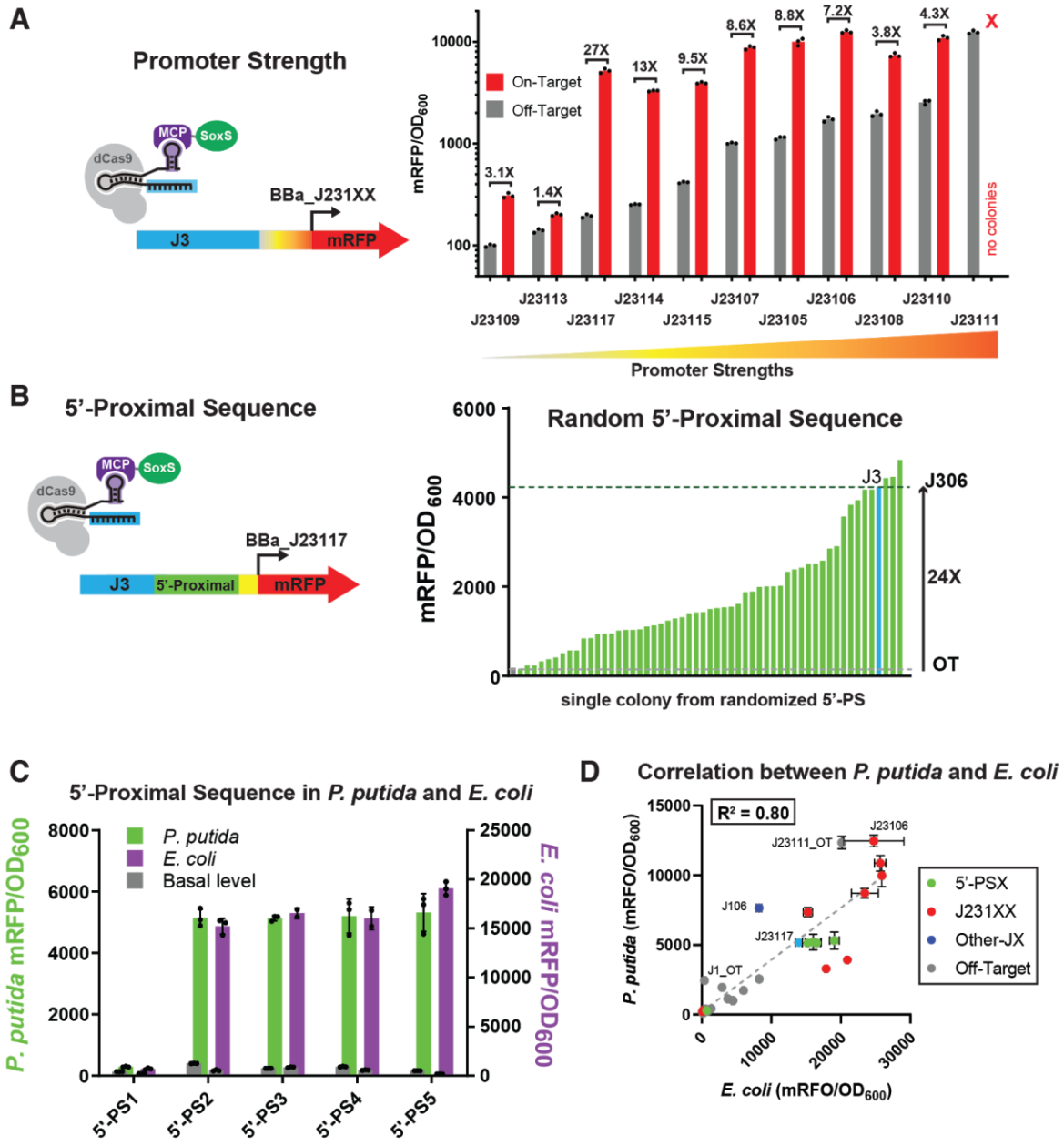


Figure 2.3. Sensitivity of CRISPRa to promoter strength and 5' upstream sequence in *P. putida*

(A) CRISPRa is sensitive to basal promoter strength. Variants of pPPC021.J231XX were constructed by changing the BBa_J23117 promoter into ten other minimal promoters. The promoters weaker than BBa_J23117 exhibited low CRISPRa efficiency while the fold-activation was maximized at BBa_J23117 and decreased as the promoter strength increased beyond that point. (B) CRISPRa is sensitive to the sequence composition of the 26 bp 5'-proximal sequence between the scRNA target site and the minimal BBa_J23117 promoter. (C) Comparison of

CRISPRa-induced expression with different 5'-proximal sequences characterized in *E. coli* and *P. putida*. Sequences are available in the Supplemental Information. (D) Correlation between CRISPRa-induced mRFP expression levels from different promoter contexts in *E. coli* and *P. putida* (R^2 = 0.80). Values in panel A, C, D represent the mean ± standard deviation calculated from *n = 3* independent biological replicates. Bars in panel B represent the value of one (*n = 1*) sample.

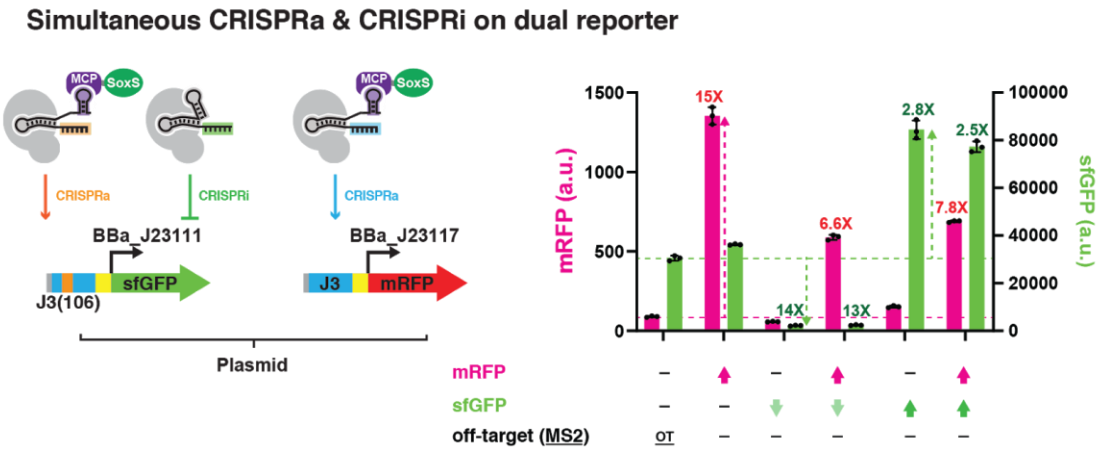


Figure 2.4. Multi-gene CRISPRa/CRISPRi in the plasmid-borne dual reporters

A multi-gene CRISPRa/CRISPRi reporter with weakly expressed mRFP (J3-BBa_J23117) and highly expressed sfGFP (J3(106)-BBa_J23111) shows simultaneous activation and repression when an activator scRNA for mRFP and a repressor sgRNA for sfGFP are delivered. The strong sfGFP reporter can also be further activated ~2-3-fold if targeted by an upstream activating scRNA. This strain exhibits noticeably slower growth in both agar and liquid media (data not shown). Values represent the mean ± standard deviation calculated from *n = 3* independent biological replicates.

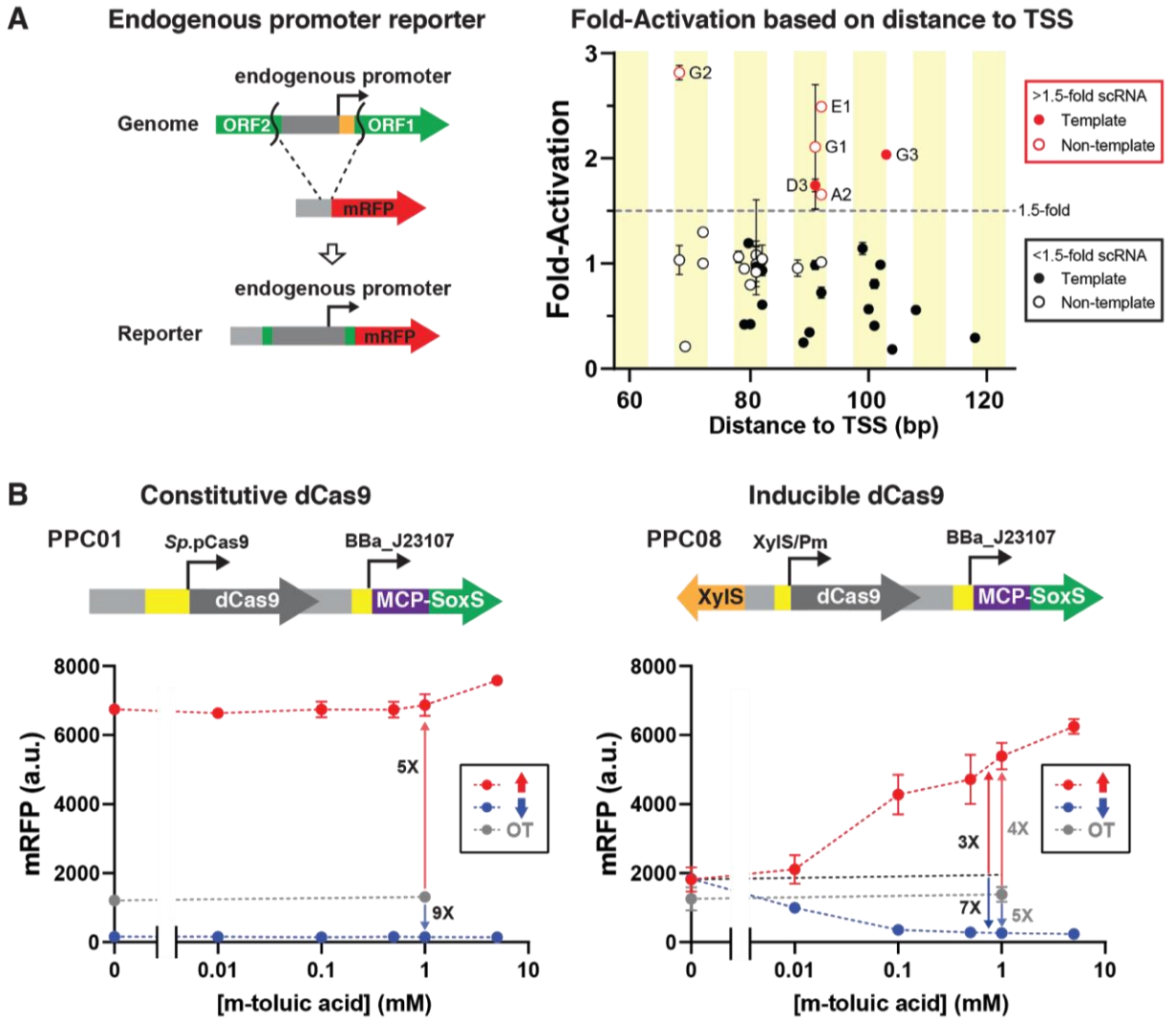


Figure 2.5. CRISPRa with endogenous promoters and inducible CRISPRa/CRISPRi in *P. putida*

(A) The putative promoter sequences between two open reading frames (ORFs) with 60 bp flanking sequences were incorporated into the mRFP reporter. scRNAs were introduced for all potentially activatable target sites corresponding to the effective distances in Figure 2.2B & 2.C. Precise distances from the target site to the TSS are listed in Table S4. hAAVS1 was used as an off-target scRNA for all ten promoters. (B) Tunable activation of mRFP expression with CRISPRa and tunable repression of mRFP expression with CRISPRi were achieved with different inducer concentrations (0-5 mM *m*-toluic acid) in the inducible-dCas9 strain (right). The inducible-dCas9 strain yielded 3-fold activation with CRISPRa or 7-fold repression with CRISPRi at 1 mM m-toluic acid compared to the no-inducer control. Fold-changes compared to the off-target control were 4-fold and 5-fold, respectively. The constitutively expressed dCas9 strain (left) showed little to no response to inducer concentration. Values in panel A and B represent the mean ± standard deviation calculated from *n = 3* independent biological replicates.

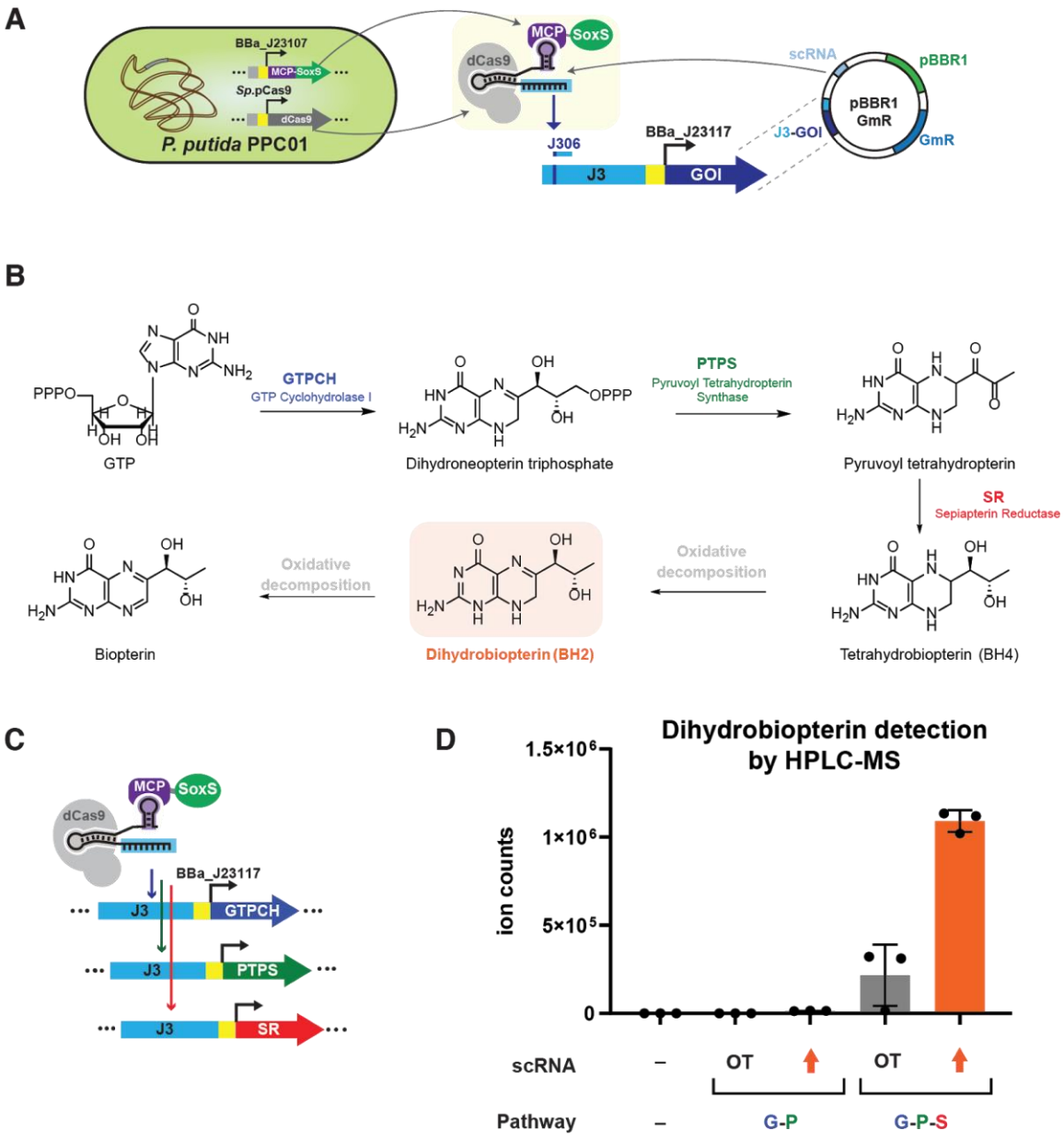


Figure 2.6. Multi-target CRISPR activation on a biopterin pathway

(A) Graphical depiction of CRISPRa implementation to any gene of interest by integrated dCas9/MCP-SoxS strains (PPC001) where scRNA(s) and heterologous genes were delivered on pBBR1-GmR plasmid. (B) Biosynthetic and spontaneous oxidation scheme from GTP into tetrahydrobiopterin (BH4) and its oxidized variants. The three-enzyme pathway consisted of *E. coli gtpch, M. alpina ptps,* and *M. alpina sr*. Tetrahydrobiopterin is reactive towards ambient oxygen and is readily oxidized into dihydrobiopterin (BH2) and biopterin, respectively. (C) Graphical depiction of CRISPRa activating three genes with a single scRNA. (D) Dihydrobiopterin (BH2) levels observed by HPLC-MS of PPC01 strains bearing pPPC024 (3-gene pathway) or pPPC025 (absence of sr gene). HPLC-MS data of three biopterin species are shown in Figure

2.S13. Values in panel D represent the mean ± standard deviation calculated from *n = 3* technical replicates.

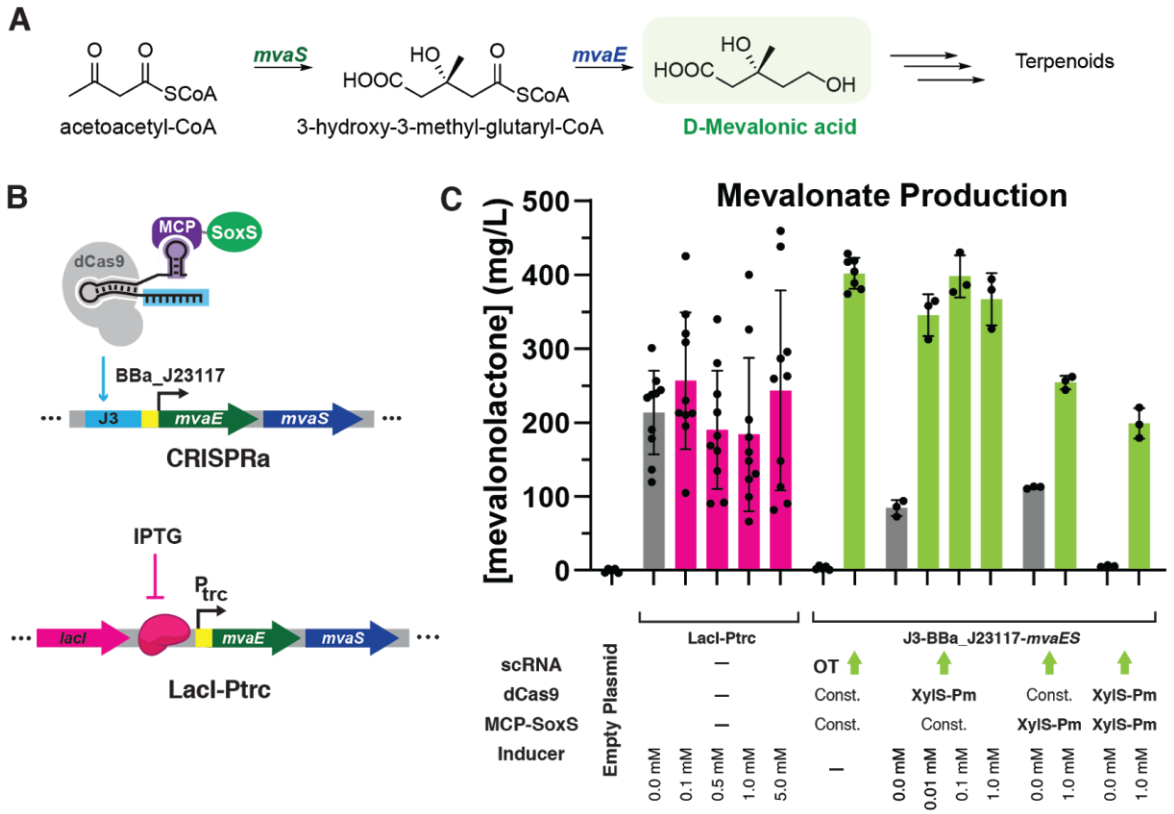


Figure 2.7. CRISPR activation on Mevalonate Production Operon

(A) Biosynthetic pathway for D-mevalonic acid production from acetoacetyl-CoA with heterologous *mvaS* and *mvaE* genes from *Enterococcus faecalis*. (B) Graphical depiction comparing CRISPR activation complex (pPPC029) with the LacI-Ptrc inducible system (pPPC030) for gene activation. (C) Titer of mevalonate calculated based on GC-MS detection of cyclized mevalonolactone (m/z = 71). The J306 scRNA was used as an on-target CRISPRa scRNA while hAAVS1 was used as an off-target scRNA. Up to 5.0 mM IPTG was used for induction of LacI-Ptrc and up to 1 mM *m*-toluic acid was used for induction of XylS-Pm. The off-target scRNA produced a mevalonate titer indistinguishable from the no plasmid control (less than 10 mg/L, see Figure 2.S14). Values in panel C represent the mean ± standard deviation calculated from *n = 3* independent biological replicates, *n = 5* for the no plasmid control and off-target scRNA, *n = 7* for the constitutively expressed dCas9/MCP-SoxS strain, and *n = 10* for the LacI-Ptrc strain.

2.6 References

Anthony, J.R., Anthony, L.C., Nowroozi, F., Kwon, G., Newman, J.D., Keasling, J.D., **2009**. Optimization of the mevalonate-based isoprenoid biosynthetic pathway in *Escherichia coli* for production of the anti-malarial drug precursor amorpha-4,11-diene. *Metab. Eng.* **11**, 13–19. https://doi.org/10.1016/j.ymben.2008.07.007

Aparicio, T., de Lorenzo, V., Martínez-García, E., **2018**. CRISPR/Cas9-Based Counterselection Boosts Recombineering Efficiency in *Pseudomonas putida*. *Biotechnol. J.* **13**, 1700161. https://doi.org/10.1002/biot.201700161

Banerjee, D., Eng, T., Lau, A.K., Sasaki, Y., Wang, B., Chen, Y., Prahl, J.-P., Singan, V.R., Herbert, R.A., Liu, Y., Tanjore, D., Petzold, C.J., Keasling, J.D., Mukhopadhyay, A., **2020**. Genome-scale metabolic rewiring improves titers rates and yields of the non-native product indigoidine at scale. *Nat. Commun.* **11**, 5385. https://doi.org/10.1038/s41467-020-19171-4

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 Res* **41**, 7429–37. https://doi.org/10.1093/nar/gkt520

Bojanovič, K., D'Arrigo, I., Long, K.S., **2017**. Global transcriptional responses to osmotic, oxidative, and imipenem stress conditions in *Pseudomonas putida*. *Appl. Environ. Microbiol.* AEM.03236-16. https://doi.org/10.1128/AEM.03236-16

Chavarría, M., Nikel, P.I., Pérez-Pantoja, D., de Lorenzo, V., **2013**. The Entner–Doudoroff pathway empowers *Pseudomonas putida* KT2440 with a high tolerance to oxidative stress. *Environ. Microbiol.* **15**, 1772–1785. https://doi.org/10.1111/1462-2920.12069

Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., E, P.R.I., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J., Ter-Ovanesyan, D., Braff, J.L., Davidsohn, N., Housden, B.E., Perrimon, N., Weiss, R., Aach, J., Collins, J.J., Church, G.M., **2015**. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* **12**, 326–8. https://doi.org/10.1038/nmeth.3312

Choi, K.-H., Schweizer, H.P., **2006**. mini-Tn7 insertion in bacteria with single attTn7 sites: example *Pseudomonas aeruginosa*. *Nat. Protoc.* **1**, 153–161. https://doi.org/10.1038/nprot.2006.24

Cook, T.B., Rand, J.M., Nurani, W., Courtney, D.K., Liu, S.A., Pfleger, B.F., **2018**. Genetic tools for reliable gene expression and recombineering in *Pseudomonas putida*. *J. Ind. Microbiol. Biotechnol.* **45**, 517–527. https://doi.org/10.1007/s10295-017-2001-5

Damalas, S.G., Batianis, C., Martin-Pascual, M., de Lorenzo, V., Martins Dos Santos, V.A.P., **2020**. SEVA 3.1: enabling interoperability of DNA assembly among the SEVA, BioBricks and Type IIS restriction enzyme standards. *Microb Biotechnol* **n/a**. https://doi.org/10.1111/1751-7915.13609

D'Arrigo, I., Bojanovič, K., Yang, X., Holm Rau, M., Long, K.S., **2016**. Genome-wide mapping of transcription start sites yields novel insights into the primary transcriptome of *Pseudomonas putida*. *Environ. Microbiol.* **18**, 3466–3481. https://doi.org/10.1111/1462-2920.13326

Depardieu, F., Bikard, D., **2020**. Gene silencing with CRISPRi in bacteria and optimization of dCas9 expression levels. *Methods* **172**, 61–75. https://doi.org/10.1016/j.ymeth.2019.07.024

Dong, C., Fontana, J., Patel, A., Carothers, J.M., Zalatan, J.G., **2018**. Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nat Commun* **9**, 2489. https://doi.org/10.1038/s41467-018-04901-6

Ehrenworth, A.M., Sarria, S., Peralta-Yahya, P., **2015**. Pterin-Dependent Mono-oxidation for the Microbial Synthesis of a Modified Monoterpene Indole Alkaloid. *ACS Synth. Biol.* **4**,

1295–1307. https://doi.org/10.1021/acssynbio.5b00025
Elmore, J.R., Dexter, G.N., Salvachúa, D., O'Brien, M., Klingeman, D.M., Gorday, K., Michener, J.K., Peterson, D.J., Beckham, G.T., Guss, A.M., **2020**. Engineered *Pseudomonas putida* simultaneously catabolizes five major components of corn stover lignocellulose: Glucose, xylose, arabinose, p-coumaric acid, and acetic acid. *Metab. Eng.* **62**, 62–71. https://doi.org/10.1016/j.ymben.2020.08.001
Fontana, J., Dong, C., Kiattisewee, C., Chavali, V.P., Tickman, B.I., Carothers, J.M., Zalatan, J.G., **2020a**. Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. *Nat. Commun.* **11**, 1618. https://doi.org/10.1038/s41467-020-15454-y
Fontana, J., Sparkman-Yager, D., Zalatan, J.G., Carothers, J.M., **2020b**. Challenges and opportunities with CRISPR activation in bacteria for data-driven metabolic engineering. *Curr. Opin. Biotechnol.* **64**, 190–198. https://doi.org/10.1016/j.copbio.2020.04.005
Fujita, M., Hanaura, Y., Amemura, A., **1995**. Analysis of the rpoD gene encoding the principal sigma factor of *Pseudomonas putida*. *Gene* **167**, 93–98. https://doi.org/10.1016/0378-1119(95)00675-3
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., Qi, L.S., Kampmann, M., Weissman, J.S., **2014**. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647–661. https://doi.org/10.1016/j.cell.2014.09.029
Ho, H.-I., Fang, J.R., Cheung, J., Wang, H.H., **2020**. Programmable CRISPR-Cas transcriptional activation in bacteria. *Mol. Syst. Biol.* **16**, e9427. https://doi.org/10.15252/msb.20199427
Huang, H.-H., Bellato, M., Qian, Y., Cárdenas, P., Pasotti, L., Magni, P., Del Vecchio, D., **2020**. dCas9 regulator to neutralize competition in CRISPRi circuits. *bioRxiv* 2020.08.11.246561. https://doi.org/10.1101/2020.08.11.246561
Jervis, A.J., Carbonell, P., Vinaixa, M., Dunstan, M.S., Hollywood, K.A., Robinson, C.J., Rattray, N.J.W., Yan, C., Swainston, N., Currin, A., Sung, R., Toogood, H., Taylor, S., Faulon, J.-L., Breitling, R., Takano, E., Scrutton, N.S., **2019**. Machine Learning of Designed Translational Control Allows Predictive Pathway Optimization in *Escherichia coli*. *ACS Synth. Biol.* **8**, 127–136. https://doi.org/10.1021/acssynbio.8b00398
Johnson, C.W., Beckham, G.T., **2015**. Aromatic catabolic pathway selection for optimal production of pyruvate and lactate from lignin. *Metab. Eng.* **28**, 240–247. https://doi.org/10.1016/j.ymben.2015.01.005
Kim, D.Y., Kim, Y.B., Rhee, Y.H., **2000**. Evaluation of various carbon substrates for the biosynthesis of polyhydroxyalkanoates bearing functional groups by *Pseudomonas putida*. *Int. J. Biol. Macromol.* **28**, 23–29. https://doi.org/10.1016/S0141-8130(00)00150-1
Kim, S.K., Yoon, P.K., Kim, S.J., Woo, S.G., Rha, E., Lee, H., Yeom, S.J., Kim, H., Lee, D.H., Lee, S.G., **2019**. CRISPR interference-mediated gene regulation in *Pseudomonas putida* KT2440. *Microb Biotechnol* **0**. https://doi.org/10.1111/1751-7915.13382
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., Nureki, O., Zhang, F., **2014**. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583. https://doi.org/10.1038/nature14136
Kosuri, S., Goodman, D.B., Cambray, G., Mutalik, V.K., Gao, Y., Arkin, A.P., Endy, D., Church, G.M., **2013**. Composability of regulatory sequences controlling transcription and translation in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **110**, 14024. https://doi.org/10.1073/pnas.1301301110
Kovach, M.E., Elzer, P.H., Steven Hill, D., Robertson, G.T., Farris, M.A., Roop, R.M., Peterson, K.M., **1995**. Four new derivatives of the broad-host-range cloning vector pBBR1MCS,

carrying different antibiotic-resistance cassettes. *Gene* **166**, 175–176. https://doi.org/10.1016/0378-1119(95)00584-1
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. *Nat. Catal.* **2**, 18–33. https://doi.org/10.1038/s41929-018-0212-4
Lee, T.S., Krupa, R.A., Zhang, F., Hajimorad, M., Holtz, W.J., Prasad, N., Lee, S.K., Keasling, J.D., **2011**. BglBrick vectors and datasheets: A synthetic biology platform for gene expression. *J. Biol. Eng.* **5**, 12. https://doi.org/10.1186/1754-1611-5-12
Liu, Y., Wan, X., Wang, B., **2019**. Engineered CRISPRa enables programmable eukaryote-like gene activation in bacteria. *Nat. Commun.* **10**, 3693. https://doi.org/10.1038/s41467-019-11479-0
Loeschcke, A., Thies, S., **2015**. *Pseudomonas putida*—a versatile host for the production of natural products. *Appl. Microbiol. Biotechnol.* **99**, 6197–6214. https://doi.org/10.1007/s00253-015-6745-4
Mi, J., Sydow, A., Schempp, F., Becher, D., Schewe, H., Schrader, J., Buchhaupt, M., **2016**. Investigation of plasmid-induced growth defect in *Pseudomonas putida*. *J. Biotechnol.* **231**, 167–173. https://doi.org/10.1016/j.jbiotec.2016.06.001
Molina-Santiago, C., Udaondo, Z., Gómez-Lozano, M., Molin, S., Ramos, J.-L., **2017**. Global transcriptional response of solvent-sensitive and solvent-tolerant *Pseudomonas putida* strains exposed to toluene. *Environ. Microbiol.* **19**, 645–658. https://doi.org/10.1111/1462-2920.13585
Nielsen, J., Keasling, J.D., **2016**. Engineering Cellular Metabolism. *Cell* **164**, 1185–1197. https://doi.org/10.1016/j.cell.2016.02.004
Nikel, P.I., Chavarría, M., Danchin, A., de Lorenzo, V., **2016**. From dirt to industrial applications: *Pseudomonas putida* as a Synthetic Biology chassis for hosting harsh biochemical reactions. *Curr. Opin. Chem. Biol.* **34**, 20–29. https://doi.org/10.1016/j.cbpa.2016.05.011
Nikel, P.I., de Lorenzo, V., **2018**. *Pseudomonas putida* as a functional chassis for industrial biocatalysis: From native biochemistry to trans-metabolism. *Metab. Eng.* **50**, 142–155. https://doi.org/10.1016/j.ymben.2018.05.005
Park, W., Peña-Llopis, S., Lee, Y., Demple, B., **2006**. Regulation of superoxide stress in *Pseudomonas putida* KT2440 is different from the SoxR paradigm in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **341**, 51–56. https://doi.org/10.1016/j.bbrc.2005.12.142
Peng, R., Wang, Y., Feng, W., Yue, X., Chen, J., Hu, X., Li, Z., Sheng, D., Zhang, Y., Li, Y., **2018**. CRISPR/dCas9-mediated transcriptional improvement of the biosynthetic gene cluster for the epothilone production in *Myxococcus xanthus*. *Microb. Cell Factories* **17**, 15. https://doi.org/10.1186/s12934-018-0867-1
Peralta-Yahya, P.P., Ouellet, M., Chan, R., Mukhopadhyay, A., Keasling, J.D., Lee, T.S., **2011**. Identification and microbial production of a terpene-based advanced biofuel. *Nat. Commun.* **2**, 483. https://doi.org/10.1038/ncomms1494
Pfleger, B.F., Pitera, D.J., Newman, J.D., Martin, V.J.J., Keasling, J.D., **2007**. Microbial sensors for small molecules: Development of a mevalonate biosensor. *Metab. Eng.* **9**, 30–38. https://doi.org/10.1016/j.ymben.2006.08.002
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**, 1173–83. https://doi.org/10.1016/j.cell.2013.02.022
Shetty, R.P., Endy, D., Knight, T.F., **2008**. Engineering BioBrick vectors from BioBrick parts. *J. Biol. Eng.* **2**, 5. https://doi.org/10.1186/1754-1611-2-5
Tan, S.Z., Reisch, C.R., Prather, K.L.J., **2018**. A Robust CRISPR Interference Gene Repression System in *Pseudomonas*. *J. Bacteriol.* **200**, e00575-17. https://doi.org/10.1128/jb.00575-17

Tian, T., Kang, J.W., Kang, A., Lee, T.S., **2019**. Redirecting Metabolic Flux via Combinatorial Multiplex CRISPRi-Mediated Repression for Isopentenol Production in *Escherichia coli*. *ACS Synth. Biol.* **8**, 391–402. https://doi.org/10.1021/acssynbio.8b00429

Trenchard, I.J., Siddiqui, M.S., Thodey, K., Smolke, C.D., **2015**. De novo production of the key branch point benzylisoquinoline alkaloid reticuline in yeast. *Metab. Eng.* **31**, 74–83. https://doi.org/10.1016/j.ymben.2015.06.010

Wirth, N.T., Kozaeva, E., Nikel, P.I., **2019**. Accelerated genome engineering of *Pseudomonas putida* by I-SceI-mediated recombination and CRISPR-Cas9 counterselection. *Microb. Biotechnol.* **13**. https://doi.org/10.1111/1751-7915.13396

Xu, X., Qi, L.S., **2019**. A CRISPR–dCas Toolbox for Genetic Engineering and Synthetic Biology. *CRISPR Basic Biol. Its Technol. Appl.* **431**, 34–47. https://doi.org/10.1016/j.jmb.2018.06.037

Zaslaver, A., Bren, A., Ronen, M., Itzkovitz, S., Kikoin, I., Shavit, S., Liebermeister, W., Surette, M.G., Alon, U., **2006**. A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nat. Methods* **3**, 623–628. https://doi.org/10.1038/nmeth895

Zhang, S., Voigt, C.A., **2018**. Engineered dCas9 with reduced toxicity in bacteria: implications for genetic circuit design. *Nucleic Acids Res.* **46**, 11115–11125. https://doi.org/10.1093/nar/gky884

Zhao, F., Zhang, Y., Li, H., Shi, R.J., Han, S.Q., **2013**. [CaCl2-heat shock preparation of competent cells of three *Pseudomonas strains* and related transformation conditions]. *Ying Yong Sheng Tai Xue Bao* **24**, 788–94.

Zheng, Y., Su, T., Qi, Q., **2019**. Microbial CRISPRi and CRISPRa Systems for Metabolic Engineering. *Biotechnol. Bioprocess Eng.* **24**, 579–591. https://doi.org/10.1007/s12257-019-0107-5

Supplementary Information

Supplementary Figures

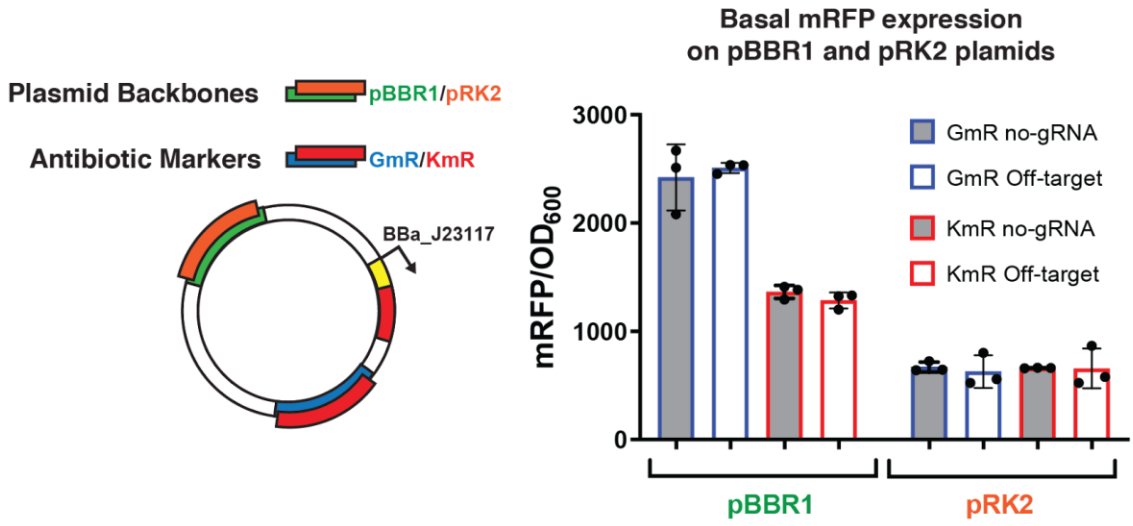
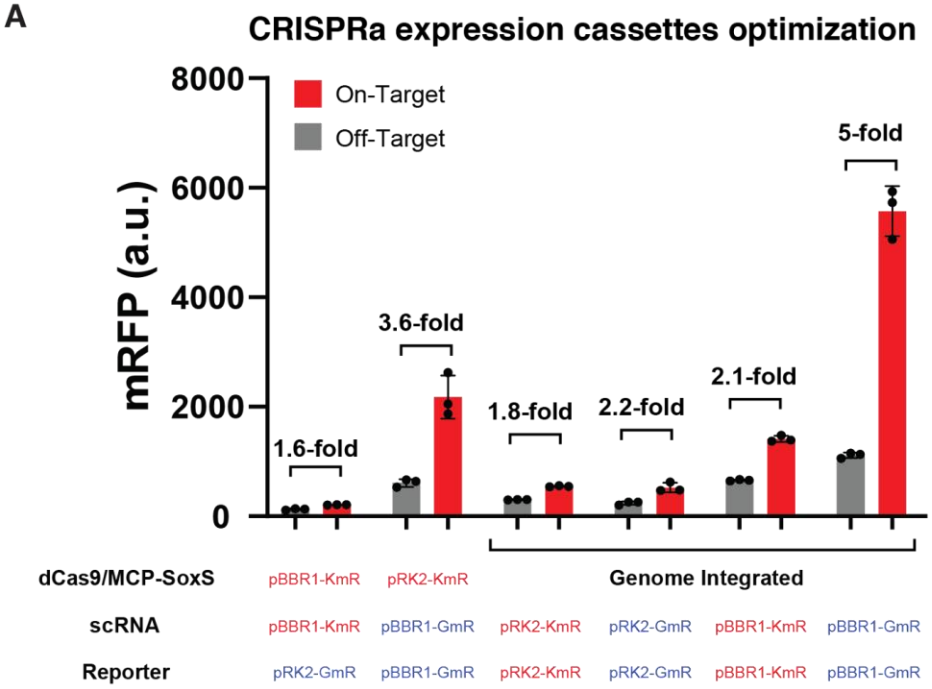


Figure 2.S1: Basal mRFP expression on pBBR1 and pRK2 plasmids

Basal expression of mRFP reporter gene from J1-BBa_J23117-mRFP in different plasmid backbones (pBBR1 or pRK2 origins), with either GmR (blue) or KmR (red) antibiotics. The plasmids expressing an off-target scRNA were tested side-by-side to the no-gRNA control and exhibited indistinguishable expression levels. See Table S2 for plasmid constructs used here. Values represent the mean ± standard deviation calculated from *n = 3* independent biological replicates. independent samples.

A


CRISPRa expression cassettes optimization
mRFP (a.u.)
8000
6000
4000
2000
0
On-Target
Off-Target
1.6-fold
3.6-fold
1.8-fold
2.2-fold
2.1-fold
5-fold
dCas9/MCP-SoxS
pBBR1-KmR
pRK2-KmR
Genome Integrated
scRNA
pBBR1-KmR
pBBR1-GmR
pRK2-KmR
pRK2-GmR
pBBR1-KmR
pBBR1-GmR
Reporter
pRK2-GmR
pBBR1-GmR
pRK2-KmR
pRK2-GmR
pBBR1-KmR
pBBR1-GmR

B

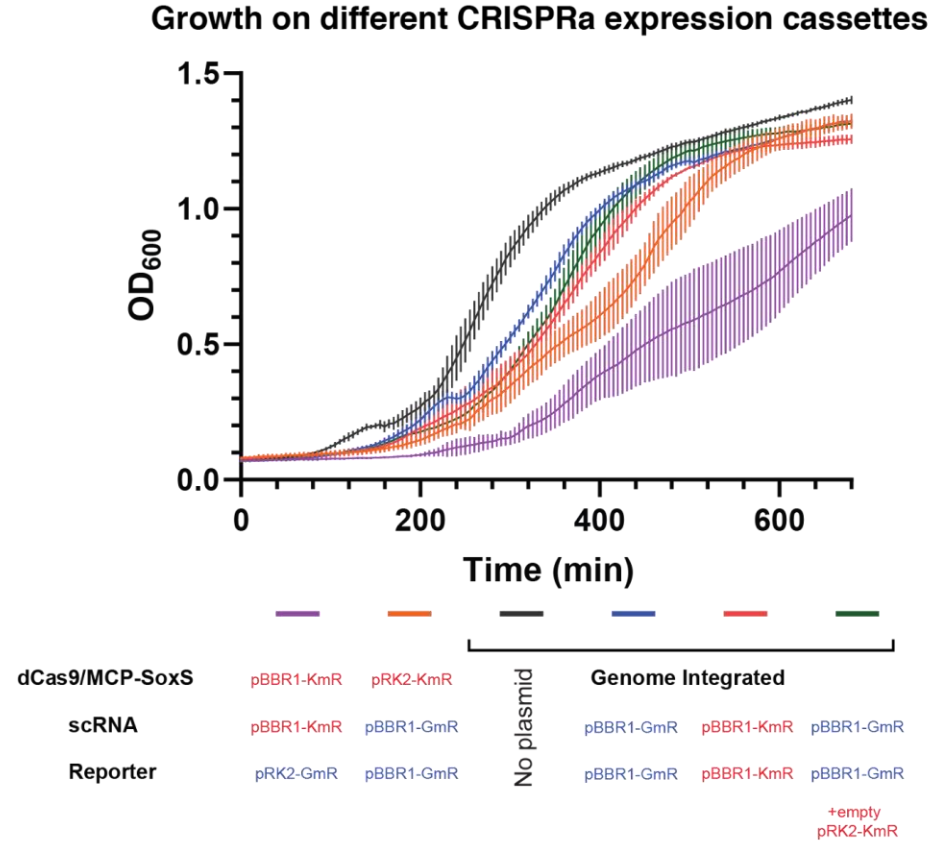

Growth on different CRISPRa expression cassettes
OD600
1.5
1.0
0.5
0.0
0
200
400
600
Time (min)
dCas9/MCP-SoxS
pBBR1-KmR
pRK2-KmR
No plasmid
Genome Integrated
scRNA
pBBR1-KmR
pBBR1-GmR
pBBR1-GmR
pBBR1-KmR
pBBR1-GmR
Reporter
pRK2-GmR
pBBR1-GmR
pBBR1-GmR
pBBR1-KmR
pBBR1-GmR
+empty
pRK2-KmR

Figure 2.S2: Expression cassettes affect CRISPRa efficiency and growth of *P. putida*

(A) CRISPRa from different expression methods for dCas9/MCP-SoxS, scRNA, and reporter. *P. putida* compatible plasmids are pBBR1 and pRK2 in which GmR (blue text) and KmR (red text) can be used as antibiotic selection markers. Both basal expressions and fold-activation varied with different expression systems. The genomically-integrated dCas9/MCP-SoxS cassettes give the highest fold-activation in pBBR1-GmR (5-fold compared to that of an off-target scRNA control). (B) Growth curve (OD_{600} vs. time) of *P. putida* strains in liquid culture with dCas9/MCP-SoxS cassette either on plasmid or integrated into the genome. Every strain expresses an off-target scRNA. Expressing dCas9/MCP-SoxS on plasmid systems (purple or orange lines) significantly reduces the growth rate compared to that of 2-plasmid strains with integrated dCas9/MCP-SoxS (green line). Qualitatively similar growth defects were observed when colonies were grown on agar plates. Values in panel A and B represent the mean ± standard deviation calculated from $n = 3$ independent biological replicates.

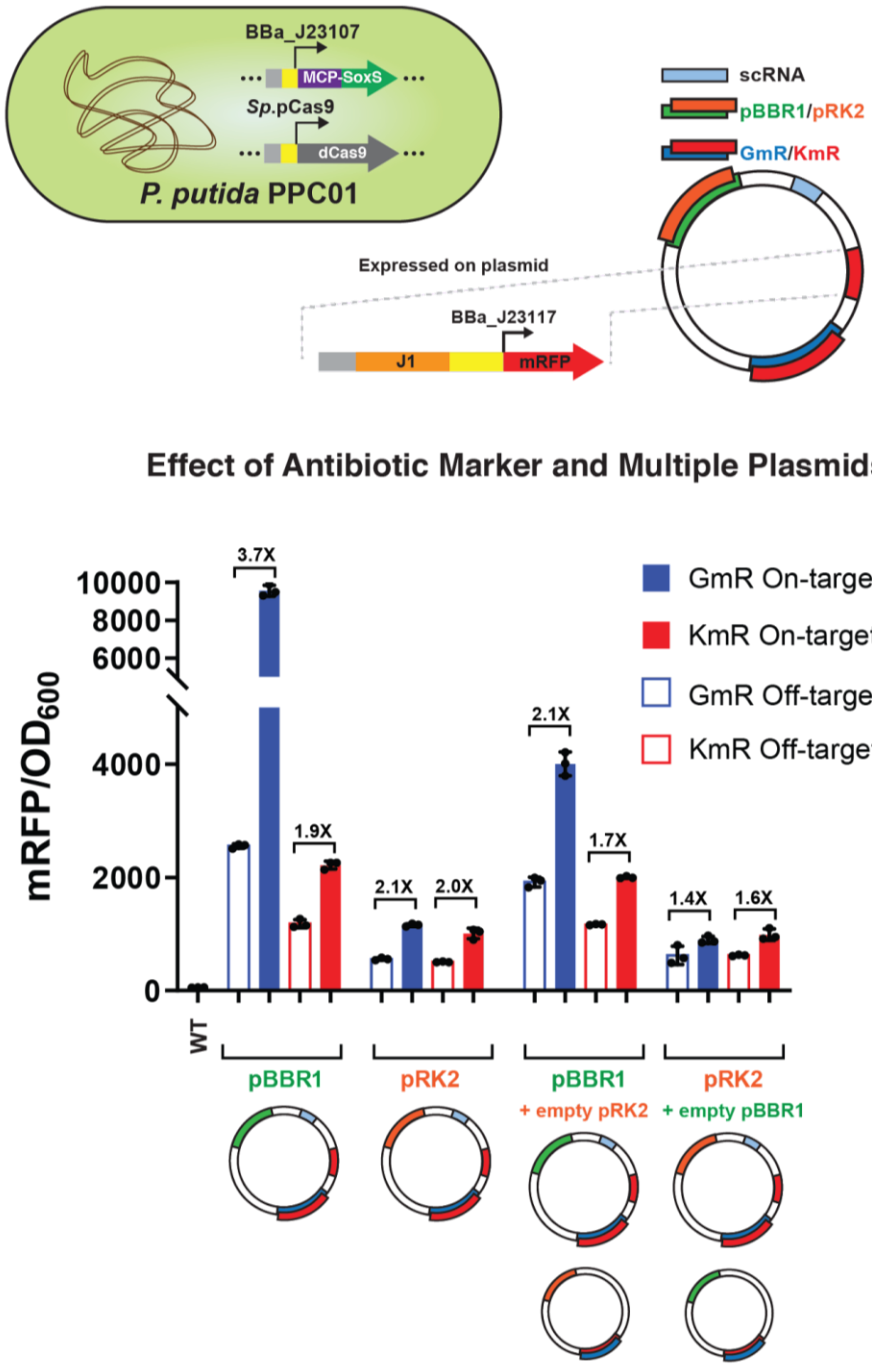


Figure 2.S3: Additional plasmid decreases CRISPRa efficiency

P. putida (PPC01) strains bearing J1-mRFP and scRNA plasmids with different origin of replications and antibiotic markers (Table S1 and S2) were further transformed with a second empty plasmid of different origin of replication and antibiotic marker. Expressing a second empty plasmid led to significant drops in both basal expression levels and fold-activation. Values represent the mean ± standard deviation calculated from $n = 3$ independent biological replicates.

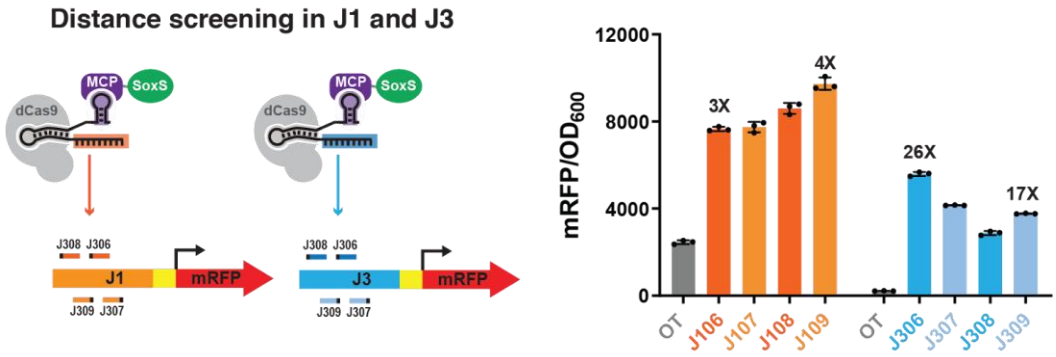


Figure 2.S4: Distance effect of CRISPRa in the J3 promoter

Comparison of optimal target sites of the J1 promoter (J106-J109) and the J3 promoter (J306-J309). The highest fold-activation was obtained with the J306 scRNA. Values represent the mean ± standard deviation calculated from *n* = *3* independent biological replicates.

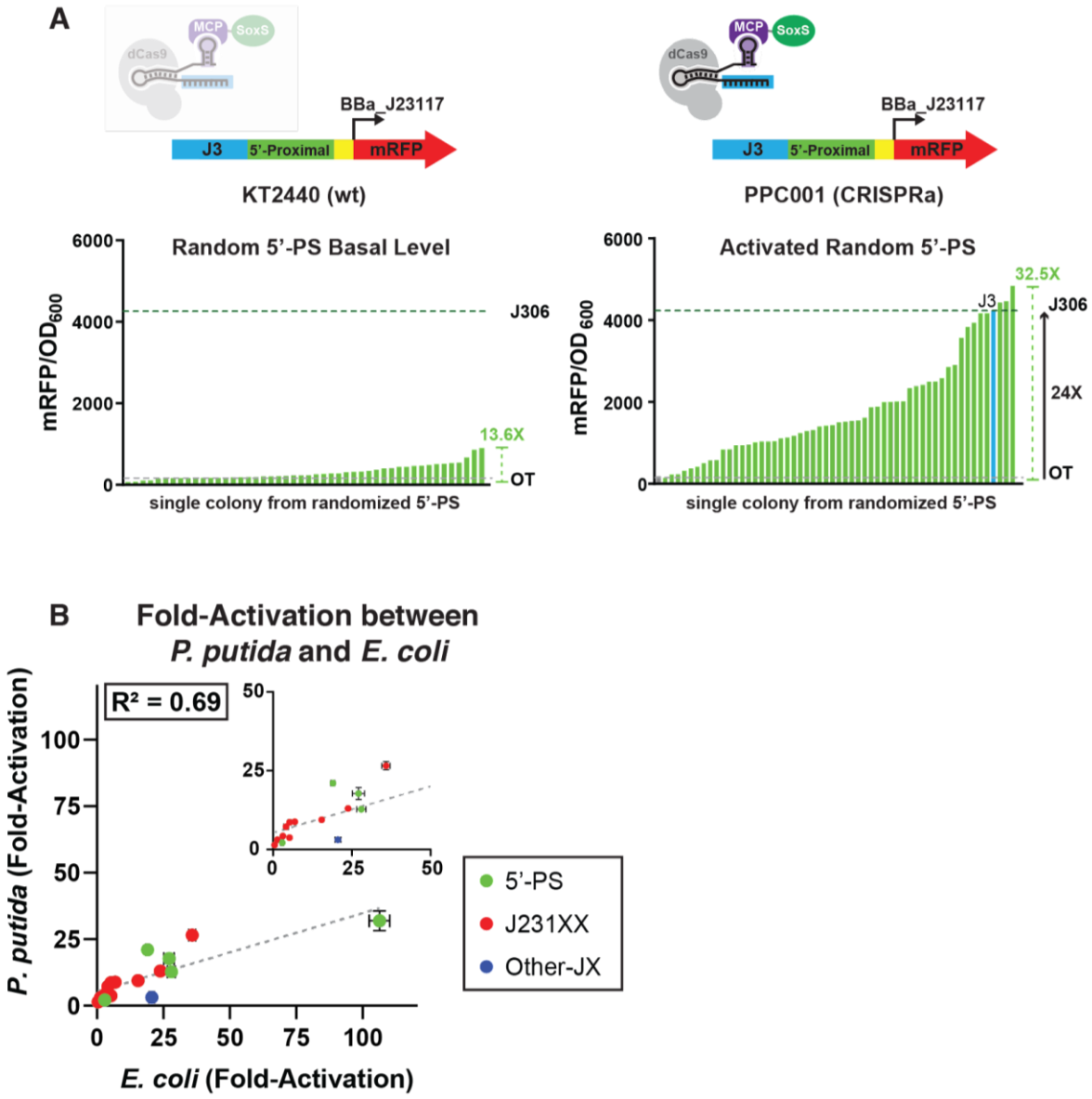


Figure 2.S5: Correlation plot of fold-activation between *P. putida* and *E. coli*

(A) Randomized 5'-proximal sequences (5'-PS) were transformed into either a no-CRISPR (KT2440) strain or a CRISPRa (PPC01) strain. Differences in basal expression with variable 5'-PS are detectable but are relatively small compared to the effects on CRISPRa fold-activation. (B) Correlation plot of fold-activation between CRISPRa strains with cognate scRNAs and off-target scRNAs (R^2 = 0.69). *E. coli* data are from (Fontana et al., 2020a). Bars in panel A represent the value of *n = 1* independent biological replicates. Values in panel B represent the mean ± standard deviation calculated from *n = 3* independent biological replicates.

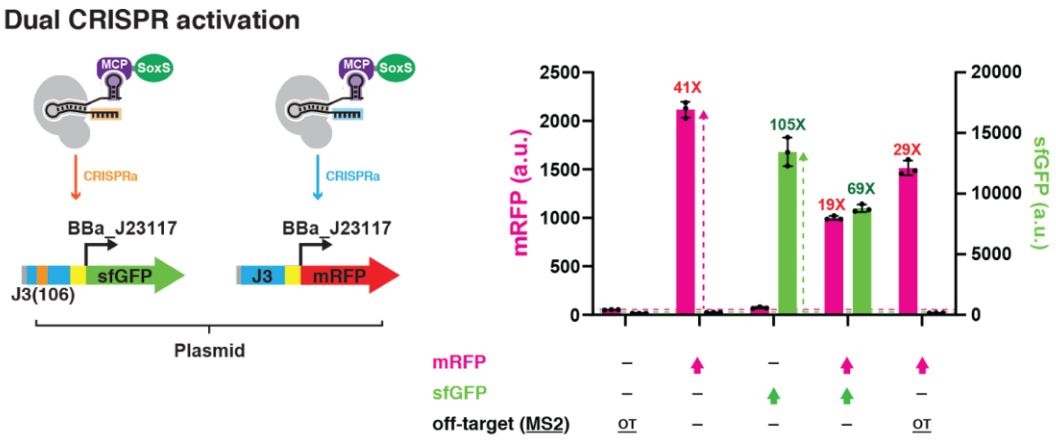


Figure 2.S6: Dual CRISPR activation on the plasmid-borne dual reporter

A multi-gene CRISPRa reporter with weakly expressed mRFP (J3-BBa_J23117) and weakly expressed sfGFP (J3(106)-BBa_J23117) can be simultaneously activated by targeting the reporters with two cognate scRNAs. The observed fold-activations with two scRNAs expressed are weaker than with only one scRNA expressed, possibly due to competition for a limited pool of dCas9. A similar effect is observed with one on-target and one off-target scRNA. Values represent the mean ± standard deviation calculated from $n = 3$ independent biological replicates.

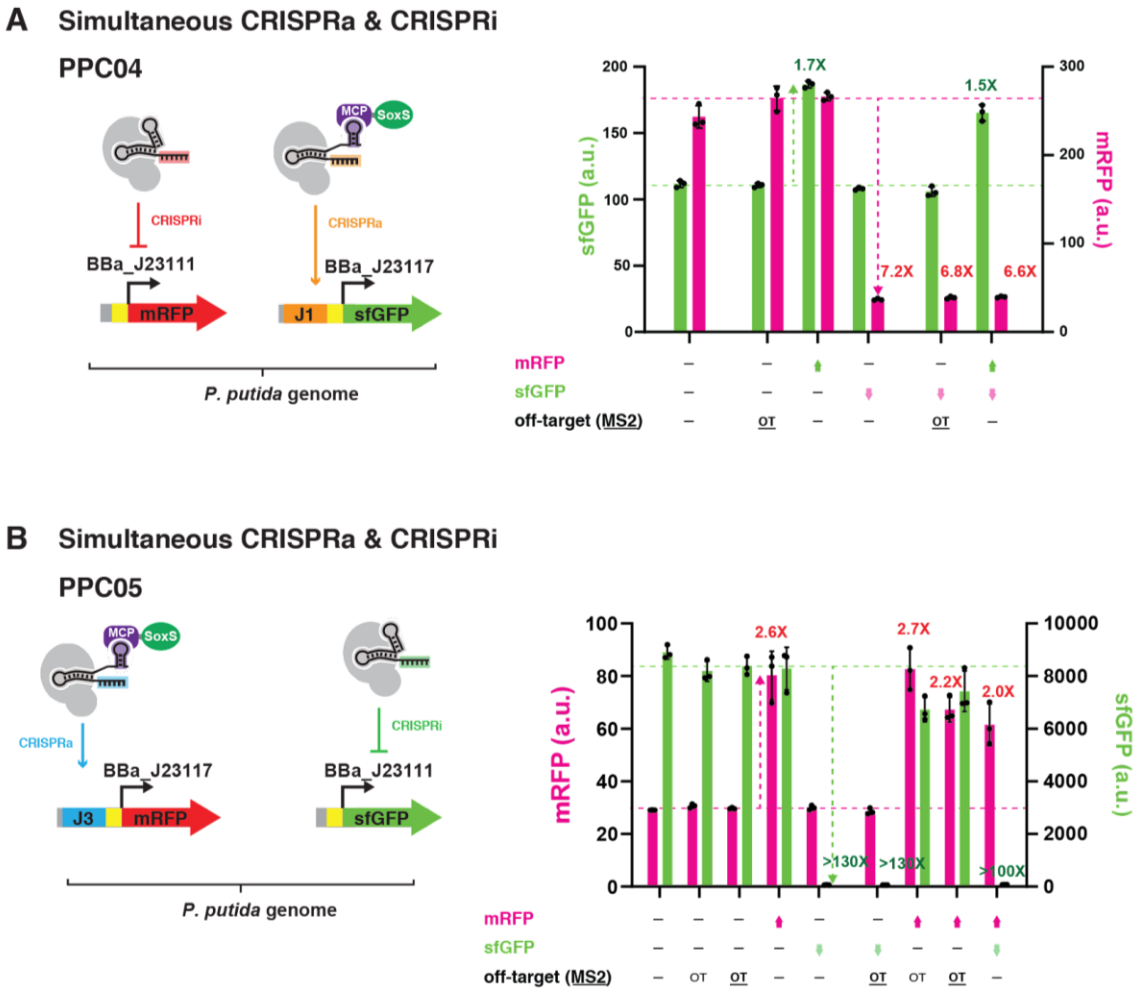


Figure 2.S7: Simultaneous CRISPRa/CRISPRi in strains with integrated dual reporters

(A) Highly expressed mRFP (BBa_J23111) and weakly expressed sfGFP (J1-BBa_J23117) were integrated together with a dCas9/MCP-SoxS construct to produce strain PPC04. Simultaneous CRISRPa on sfGFP and CRISPRi on mRFP are detectable, but the CRISPRa fold-activation is modest compared to that observed with the plasmid reporter (main text Figure 2.4). (B) Weakly expressed mRFP (J3-BBa_J23117) and highly expressed sfGFP (BBa_J23111) were integrated into the PPC01 strain at the pp1 and pp2 sites, respectively, to produce strain PPC05. The fold-change from CRISPRa and CRISPRi marginally improved compared to the PPC04 strain (panel A). The magnitude of CRISPRa fold-activation in simultaneous CRISPRa/CRISPRi was weaker than that observed if just a single scRNA was delivered to activate the mRFP reporter, possibly due to competition between multiple scRNA/gRNA cassettes for a limited pool of dCas9. A similar effect was observed with one activating scRNA and one off-target scRNA. Values in panel A and B represent the mean ± standard deviation calculated from $n = 3$ independent biological replicates.

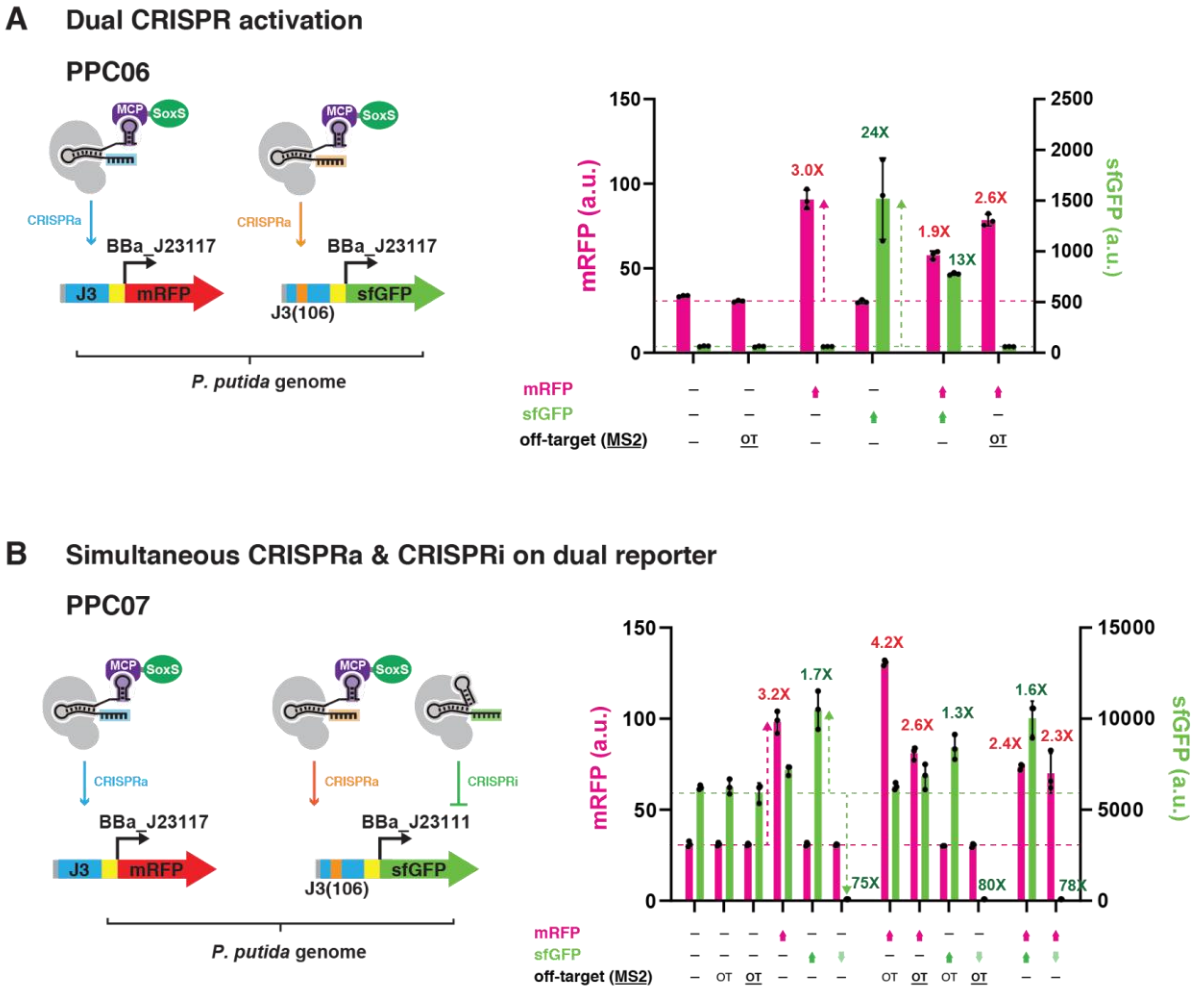


Figure 2.S8: Multi-gene CRISPRa/CRISPRi regulation in the integrated dual reporters

(A) An integrated dual reporter with weakly expressed mRFP (J3-BBa_J23117) and weakly expressed sfGFP (J3(106)-BBa_J23117) can be activated with two scRNAs, one targeting each reporter. The presence of the second sgRNA/scRNA reduced CRISPRa efficiency compared to single gene activation. (B) An integrated dual reporter with weakly expressed mRFP (J3-BBa_J23117) and a highly expressed sfGFP (J3(106)-BBa_J23111) can be activated or repressed at the sfGFP reporter. Simultaneous activation of mRFP and repression of sfGFP occurs with a J306 scRNA for mRFP activation and an sgRNA that targets within the sfGFP ORF for repression. Activation of both mRFP and sfGFP occurs with a J306 scRNA for mRFP and a J106 scRNA for sfGFP. An unexpected improvement in CRISPRa mRFP expression from the second off-target sgRNA was observed, while sfGFP CRISPRa suffered from the presence of the second guide-RNA, similar to previous conditions. Values in panel A and B represent the mean ± standard deviation calculated from *n = 3* independent biological replicates.

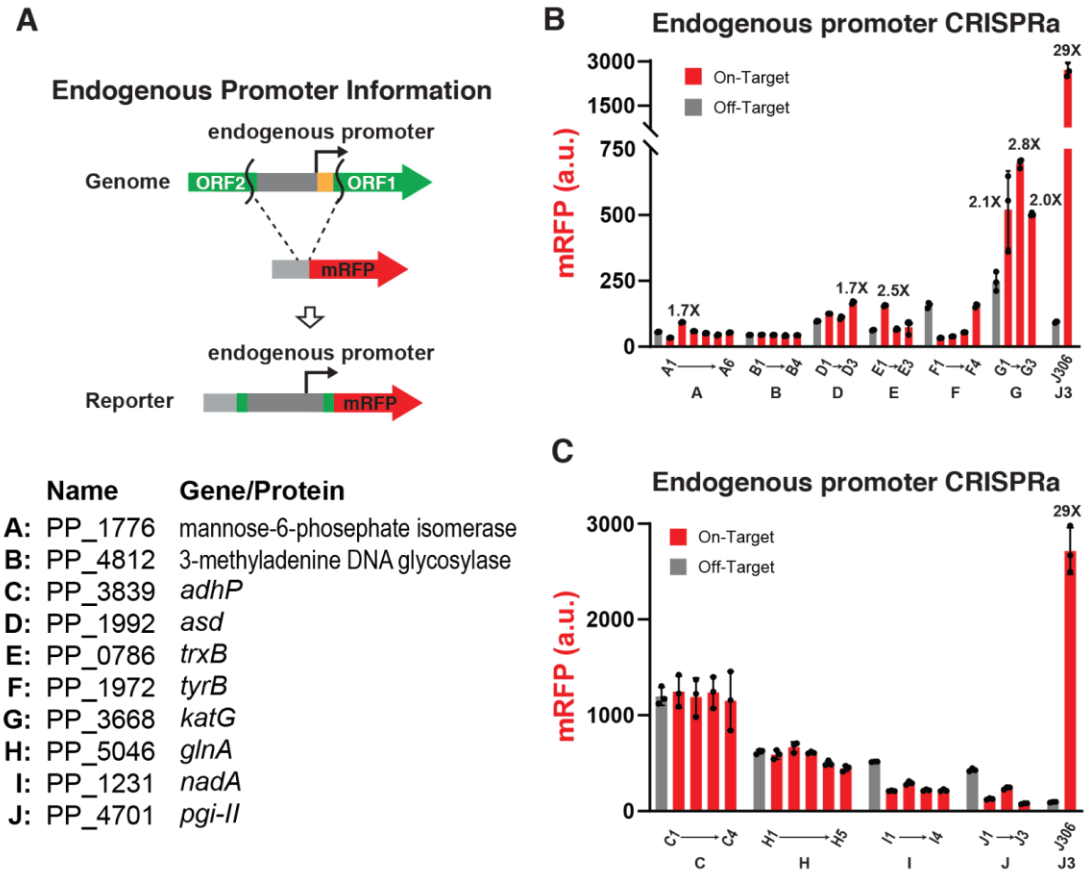
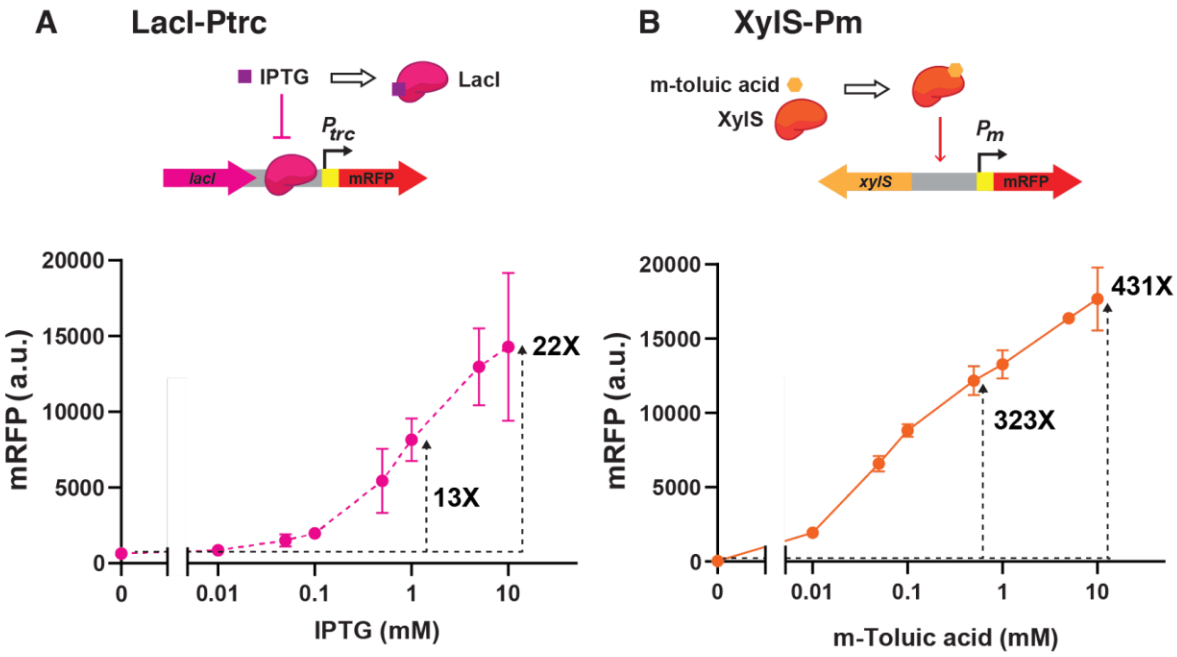


Figure 2.S9: CRISPR activation of *P. putida* endogenous promoters

(A) Ten *P. putida* endogenous promoters were selected based on available scRNA target sites at the appropriate phase and distance from reported transcription start sites (TSSs; see main text) and were coupled with mRFP reporter. Each gene was given an abbreviated code (A-J). (B) Activation profiles of CRISPRa on the endogenous promoters with relatively low basal expression (promoters A-G except promoter C) were plotted with the corresponding scRNAs (A1-A6, for example). Fold-changes were provided for instances where >1.5-fold activation was observed compared to an off-target scRNA (hAAVS1). (C) Activation profiles of CRISPRa on the endogenous promoters with relatively high basal expression promoters (promoters C and H-J) revealed no significant CRISPRa activity with the scRNAs that were tested. The J3-BBa_J23117 promoter with J306 scRNA was included as a positive CRISPRa control. Values in panel B and C represent the mean ± standard deviation calculated from $n = 3$ independent biological replicates.



C Fluorescence distributions measured by flow cytometry

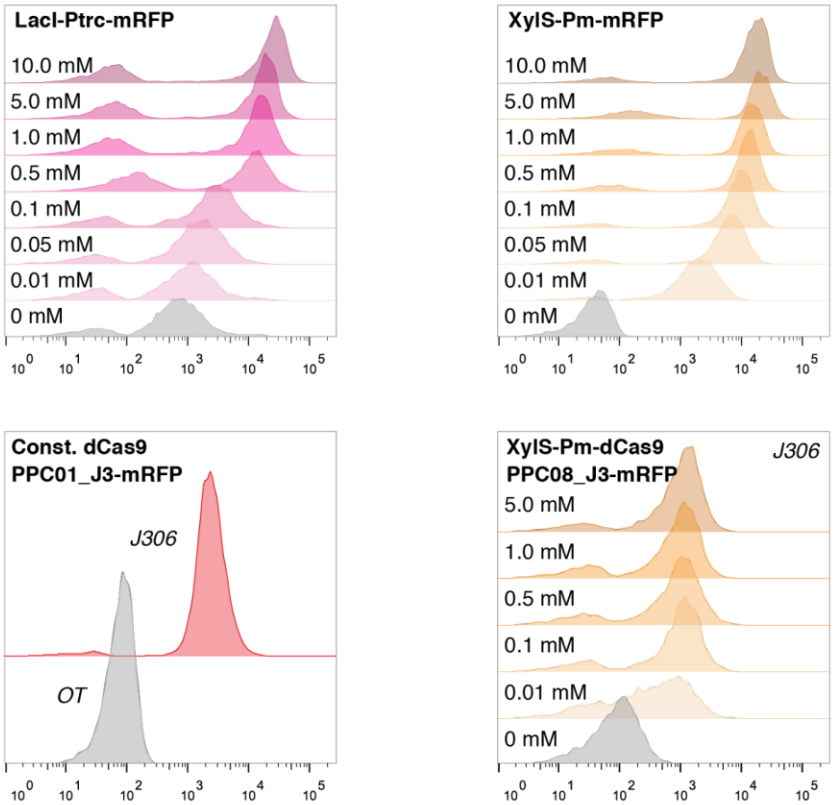


Figure 2.S10: Inducible promoters in *P. putida*

(A) Inducible LacI-Ptrc-mediated activation of an mRFP reporter gene in *P. putida* KT2440 with 0-1 mM IPTG. We observed 13-fold activation at 1 mM IPTG compared to the no inducer condition. (B) Inducible XylS-Pm-mediated activation of an mRFP reporter gene in *P. putida* KT2440 with 0-5 mM *m*-toluic acid. We observed >300-fold activation at 1 mM *m*-toluic acid compared to the no inducer condition. The XylS-Pm promoter provides a better dynamic range compared to the LacI-Ptrc promoter mainly due to its relatively low basal expression. (C) mRFP reporter gene fluorescence distributions measured by flow cytometry. At every inducer concentration, the LacI-Ptrc promoter demonstrated higher variability within the population than XylS-Pm. We observed bimodal population distributions with LacI-Ptrc, suggesting that this promoter is unstable in *P. putida*. No bimodal distributions were detected with constitutive CRISPRa, and only a small bimodal population was observed with XylS-Pm inducible mRFP or dCas9. The strains shown for constitutive CRISPRa are on-target (J306) and off-target (OT) in the PPC01 background. For inducible CRISPRa, the strain background is PPC08. Values in panel A represent the mean ± standard deviation calculated from $n = 6$ independent biological replicates.. Values in panel B represent the mean ± standard deviation calculated from $n = 3$ independent biological replicates. Values in panel C represent the data (n =1) from different inducer concentrations or scRNAs.

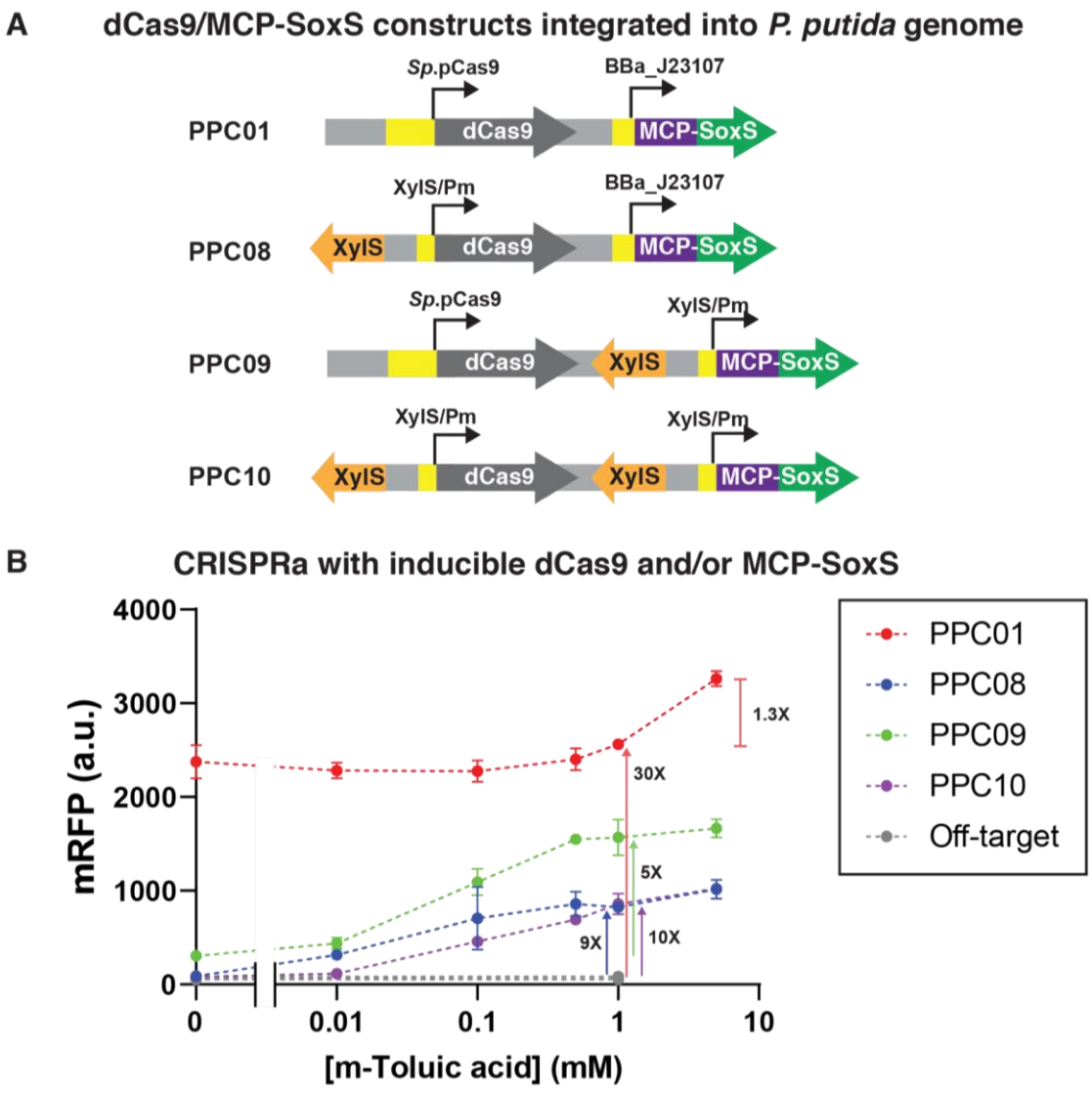


Figure 2.S11: Inducible CRISPRa by XylS-Pm on CRISPRa machinery

(A) Inducible dCas9/MCP-SoxS constructs were integrated into the *P. putida* genome with the mini-Tn7 method. XylS-Pm promoters were introduced in place of the *Sp*.pCas9 promoter regulating dCas9 and/or the J23107 promoter for MCP-SoxS. (B) Strains were transformed with a vector (pPPC020) carrying either an off-target scRNA or a J306 scRNA. The fold-activation in the presence of *m*-toluic acid as an inducer are 9-fold, 5-fold, and 10-fold from PPC08, PPC09, and PPC10, respectively. Strains with inducible dCas9 only, or both dCas9 and MCP-SoxS inducible, showed minimal leaky activation in the absence of inducer. If only MCP-SoxS is under control of the inducible promoter, leaky activation is detectable. In a strain with constitutively expressed CRISPRa machinery that should not be responsive to inducer, high inducer concentrations (5 mM) also led to a modest increase in mRFP expression (1.3-fold). Values in panel B represent the mean ± standard deviation calculated from $n = 3$ independent biological replicates.

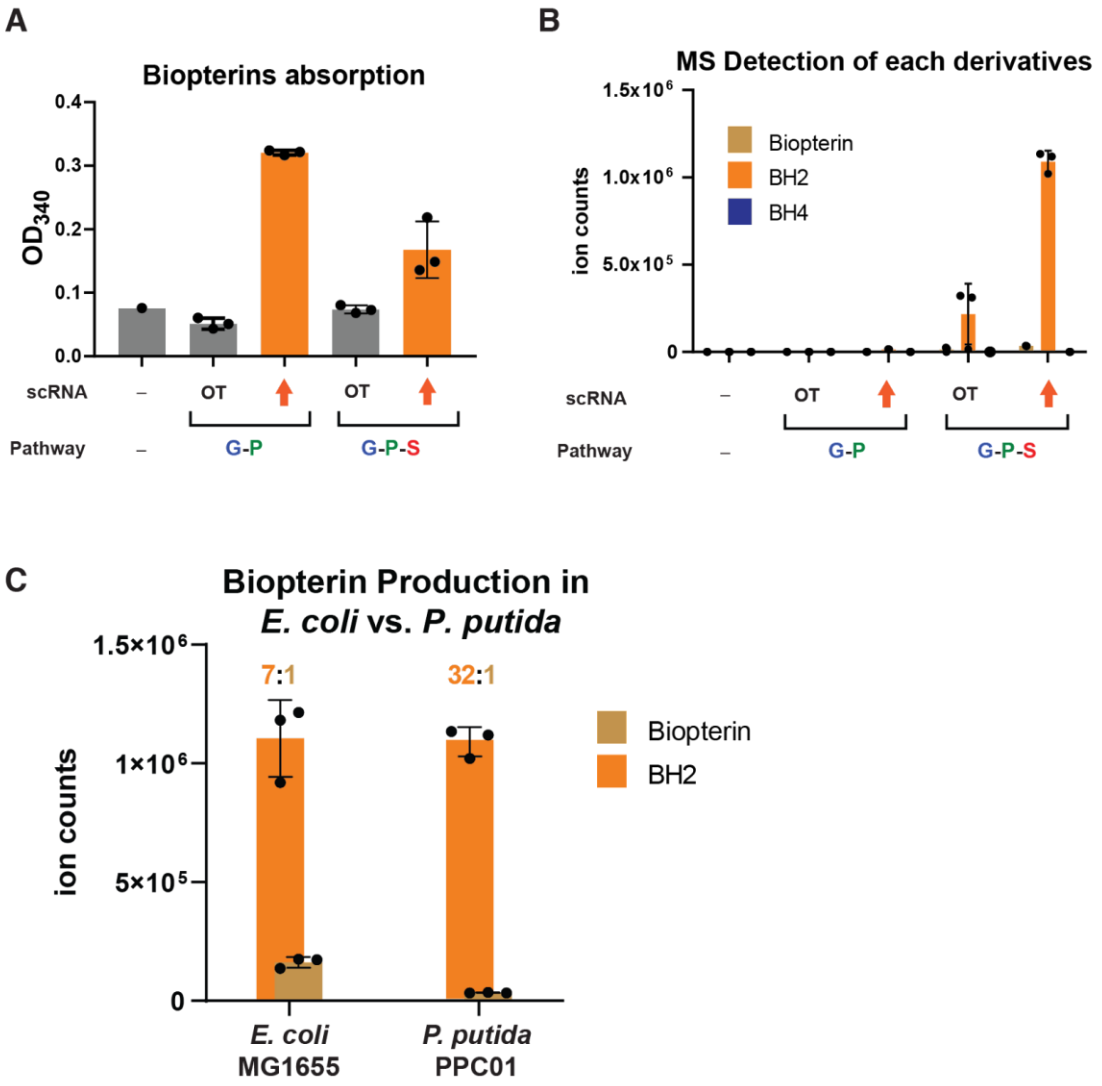


Figure 2.S12: Biopterins Production in *P. putida* with CRISPRa tool

(A) OD_{340} absorption of *P. putida* supernatant, which corresponds to absorption of dihydrobiopterin and biopterin. (B) The HPLC-MS signals of three biopterin derivatives with pathway genes regulated by different CRISPRa programs. (C) Comparison of biopterin and dihydrobiopterin (BH2) production from a CRISPR-activated tetrahydrobiopterin pathway in *E. coli* MG1655 (transformed with pCK015 and pCK005.AAV/pCD581 bearing hsAAV1/J306 scRNA) and *P. putida* PPC01 (transformed with pPPC027 bearing either hsAAV1 or J306 scRNA). The ratio of signal between BH2 and biopterin produced from *P. putida* (32:1) is higher than that of *E. coli* (7:1). Values in panel A represent the mean ± standard deviation calculated from *n = 3* independent biological replicates. The no pathway control in panel A represents one (*n = 1*) sample. Values in panel B, and C represent the mean ± standard deviation calculated from *n = 3* technical replicates.

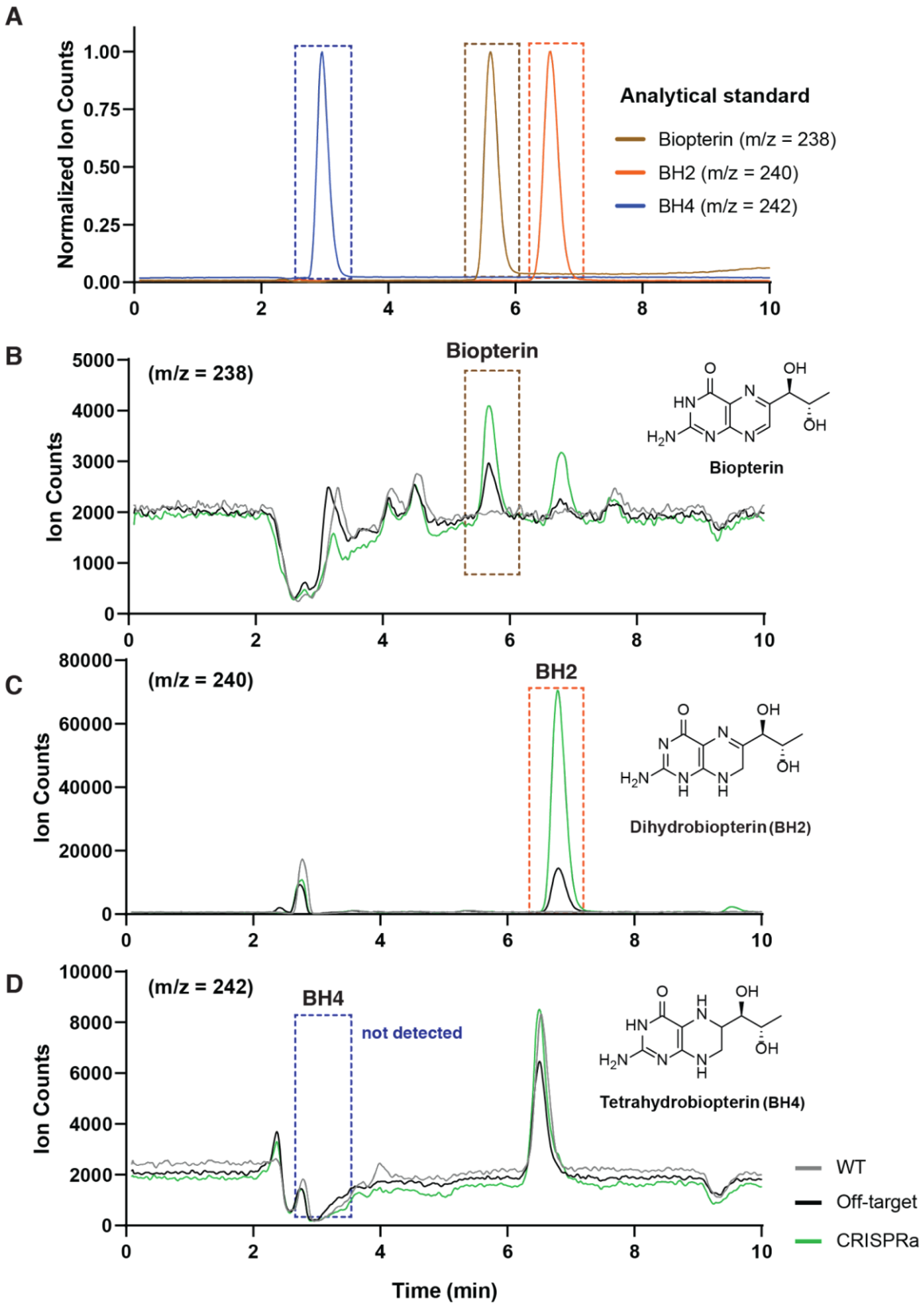

A
Normalized Ion Counts
1.00
0.75
0.50
0.25
0.00
0
2
4
6
8
10
Analytical standard
Biopterin (m/z = 238)
BH2 (m/z = 240)
BH4 (m/z = 242)
B
(m/z = 238)
Biopterin
Ion Counts
5000
4000
3000
2000
1000
0
Biopterin
C
(m/z = 240)
BH2
80000
60000
40000
20000
0
Dihydrobiopterin (BH2)
D
(m/z = 242)
BH4
not detected
10000
8000
6000
4000
2000
0
Tetrahydrobiopterin (BH4)
WT
Off-target
CRISPRa
Time (min)

Figure 2.S13: HPLC-MS Spectra of Biopterin Products in *P. putida*

(A) Overlaid chromatograms of commercial standards for biopterin, BH2, and BH4 normalized to maximum signal of each corresponding ion count. (B-D) Biopterin pathway outputs from a *P. putida* strain with CRISPRa-mediated activation of the metabolic pathway. Panels correspond to m/z ion count signals for Biopterin (B), BH2 (C), and BH4 (D). The parental strain KT2440 (gray) was used as a negative control (no heterologous pathway). PPC01 carrying pPPC027 with a J306 scRNA (green) showed significant improvement in BH2 product compared to that of an off-target scRNA (black). BH4 (r.t. ~ 3 min) was not observed in any tested condition.

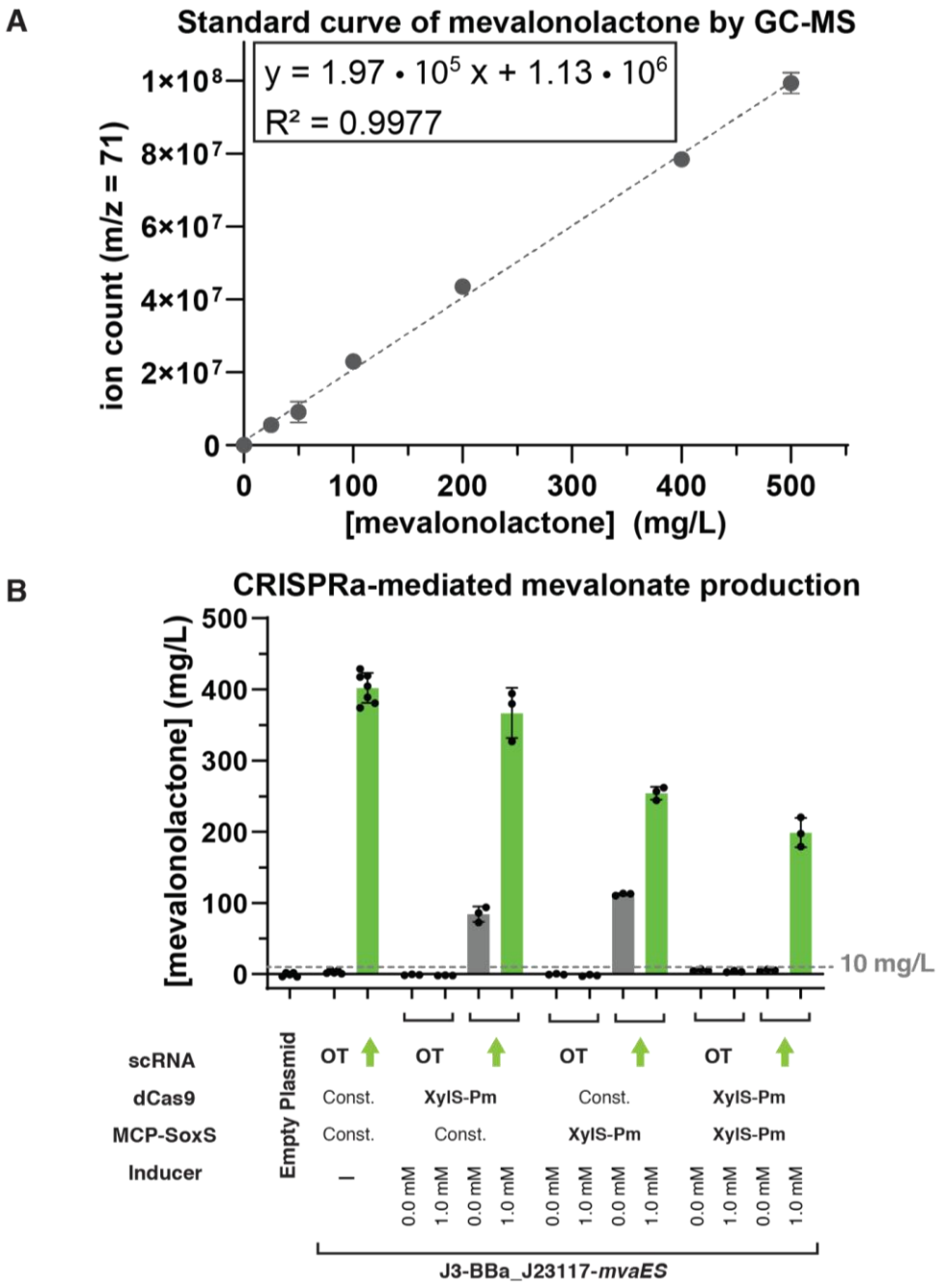


Figure 2.S14: GC-MS Detection of Mevalonic Acid

(A) A representative standard curve of mevalonolactone in ethyl acetate at different concentrations (0, 25, 50, 100, 200, 400, and 500 mg/mL) measured by GC-MS at m/z = 71. (B) CRISPRa-mediated mevalonate production with additional off-target controls (see also main text Figure 2.7). Strains with an off-target scRNA yielded mevalonate levels that were indistinguishable to that of an empty plasmid control (less than 10 mg/L). Values in panel A represent the mean ± standard deviation calculated from *n = 3* technical replicates. Values in panel B represent the mean ± standard deviation calculated from *n = 3* independent biological

replicates, *n* = *5* for the no plasmid control and off-target scRNA of constitutively expressed dCas9/MCP-SoxS strain, and *n* = *7* for the J306 scRNA.

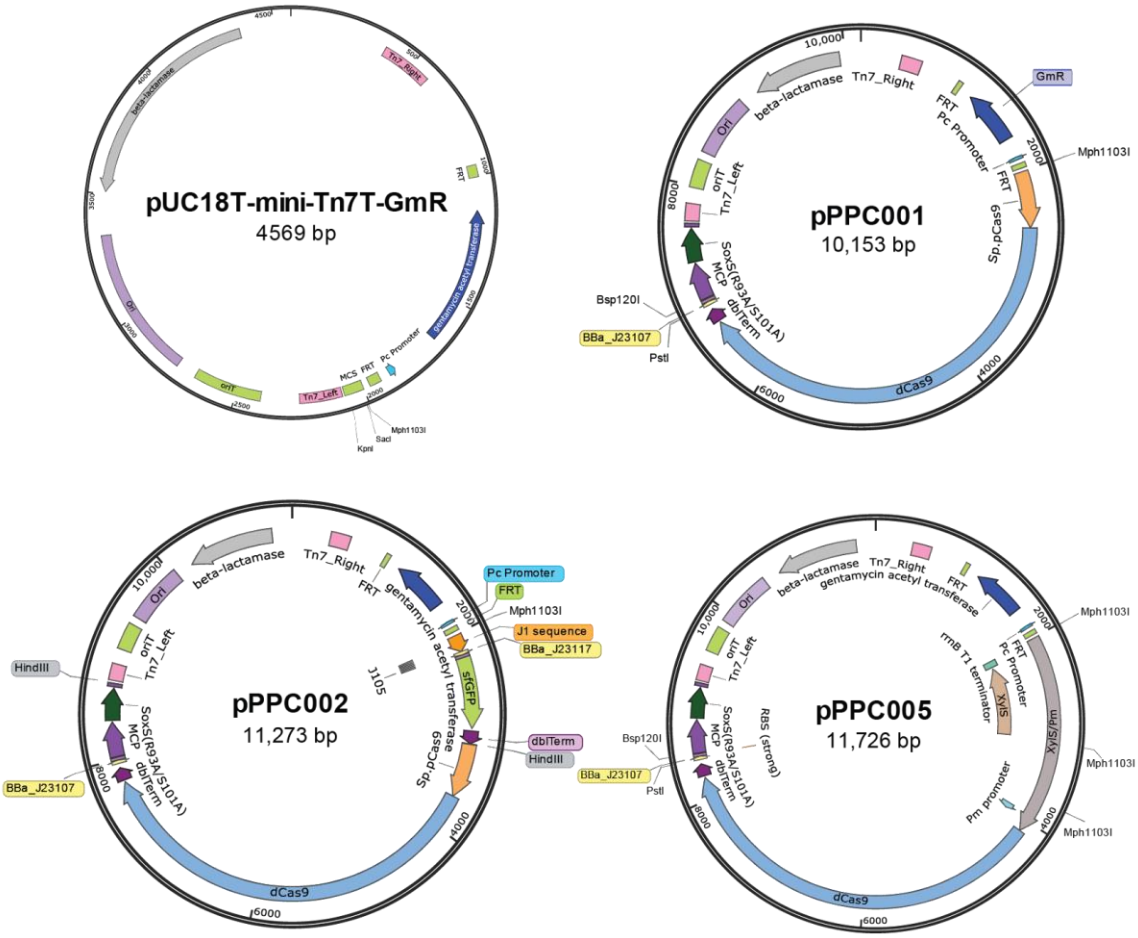


Figure 2.S15: Mini-Tn7T Plasmid Maps

Selected examples of integration plasmid maps with labeled important parts and restriction sites of pPPC001, pPPC002, and pPPC005.

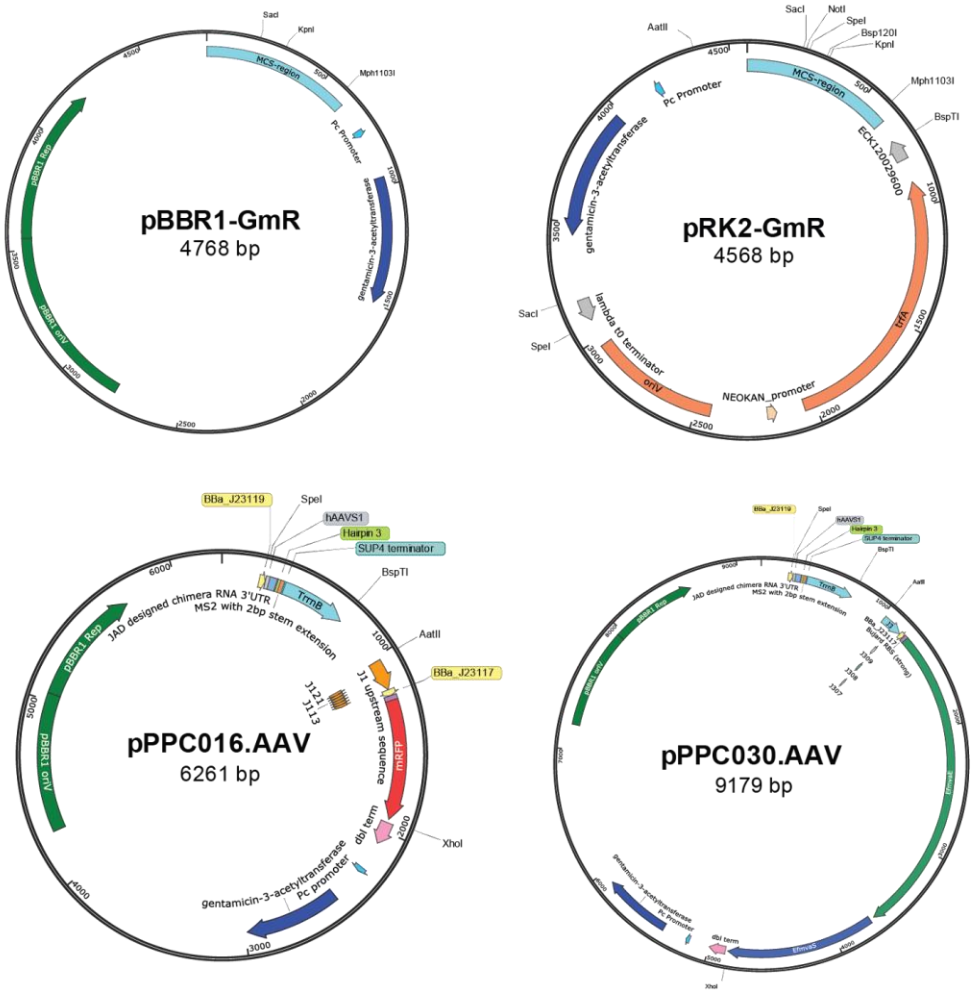


Figure 2.S16: Replicable Plasmid Maps

Selected examples of replicable plasmid maps with labeled important parts and restriction sites of pBBR1-GmR, pRK2-GmR, pPPC016, and pPPC030.

Supplementary Tables

Table 2.S1: Description of bacterial strains and plasmids used in each figure

Figures	Strains	Plasmid	scRNA target (or sgRNA)
1C	KT2440, PPC01	pPPC010 + pPPC014, pPPC014, pPPC016	hAAVS1, J109
2B	PPC02	pPPC008	hAAVS1, J101-121
2C	PPC02, PPC03.1-12	pPPC008	hAAVS1, J106
2D	PPC01	pPPC016, pPPC020, pPPC020.106, pPPC016.306,	hAAVS1, J106, J306
3A	PPC01	pPPC020, pPPC021.J231XX (10 promoters)	hAAVS1, J306
3B	PPC01	pPPC022.5PS	J306
3C	PPC01, MG1655	pPPC023.5PSN (PS1 to PS5), co-transform with pCD442 in *E. coli*	hAAVS1, J306
3D	PPC01, MG1655	pPPC016, pPPC020, pPPC021.J231XX, pPPC023.5PSN *E. coli* plasmids according to (Fontana et al., 2020a)	hAAVS1, J106, J306
4	PPC01	pPPC024	hAAVS1, J306, jfGFP1, J306-jfGFP1, J106, J306-J106
5A	PPC01	pPPC026.XN	hAAVS1, AN-JN as listed in Table S4
5B	PPC01, PPC08	pPPC021.J23110	hAAVS1, J306, RR2
6D	PPC01	pPPC027, pPPC028	hAAVS1, J306
7C	KT2440, PPC01, PPC08, PPC09, PPC10	pBBR1-GmR, pPPC029, pPPC030	hAAVS1, J306
S1	KT2440	pPPC012, pPPC013, pPPC014, pPPC015, pPPC016, pPPC017, pPPC018, pPPC019	hAAVS1
S2A	KT2440, PPC01	pPPC010 + pPPC014, pPPC011 + pPPC016, pPPC019, pPPC018, pPPC017, pPPC016	hAAVS1, J109
S2B	KT2440, PPC01	pPPC010 + pPPC014, pPPC011 + pPPC016, pPPC016, pPPC017, pPPC016 + pRK2-KmR	hAAVS1
S3	PPC01	pPPC016, pPPC017, pPPC018, pPPC019, pBBR1-GmR, pBBR1-KmR, pRK2-GmR, pRK2-KmR	hAAVS1, J109
S4	PPC01	pPPC016, pPPC020	hAAVS1, J106, J107, J108, J109, J306, J307, J308, J309
S5A	KT2440, PPC01	pPPC022.5PS	J306

S5B	PPC01, MG1655	pPPC016, pPPC020, pPPC021.J231XX, pPPC023.5PSN *E. coli* plasmids according to (Fontana et al., 2020a)	hAAVS1, J106, J306
S6	PPC01	pPPC025	hAAVS1, J306, J106, J306-J106, J306-hAAVS1
S7A	PPC04	pBBR1-KmR, pPPC009	hAAVS1, J106, RR2, hAAVS1-RR2, J106-RR2
S7B	PPC05	pBBR1-GmR, pPPC008	hAAVS1(sgRNA), hAAVS1, J306, jfGFP1, hAAVS1-jfGFP1, J306-hAAVS1(sgRNA), J306-hAAVS1, J306-jfGFP1
S8A	PPC06	pBBR1-GmR, pPPC008	hAAVS1, J306, J106, J306-J106, J306-hAAVS1
S8B	PPC07	pBBR1-GmR, pPPC008	hAAVS1(sgRNA), hAAVS1, J306, J106, jfGFP1, J306-hAAVS1(sgRNA), J306-hAAVS1, hAAVS1(sgRNA)-J106, hAAVS1-jfGFP1, J306-J106, J306-jfGFP1
S9	PPC01	pPPC026.XN	hAAVS1, AN-JN as listed in Table S4
S10A	KT2440	pCK241	
S10B	KT2440	pCK243	
S10C	KT2440, PPC01, PPC08	pCK241, pCK243, pPPC020	hAAVS1, J306
S11	PPC01, PPC08, PPC09, PPC10	pPPC020	hAAVS1, J306
S12A, S12B	PPC01	pPPC027, pPPC028	hAAVS1, J306
S12C	PPC01, MG1655	pPPC027, pCD442, pCD581, pCK015	hAAVS1, J306
S13	KT2440, PPC01	pPPC027	hAAVS1, J306
S14	KT2440, PPC01, PPC08, PPC09, PPC10	pBBR1-GmR, pPPC030	hAAVS1, J306

Table 2.S2: Description of biological parts in each plasmid

Integration plasmid	**Backbone**	**Mph1103I**	**SacI/KpnI**	**HindIII**
pPPC001	pUC18T-miniTn7T-Gm		*Sp*.pCas9-dCas9/BBa_J23107-MCP-SoxS	
pPPC002	pUC18T-miniTn7T-Gm		J1-BBa_J23117-sfGFP	*Sp*.pCas9-dCas9/BBa_J23107-MCP-SoxS
pPPC003.N	pUC18T-miniTn7T-Gm		J1(+N)-BBa_J23117-sfGFP	*Sp*.pCas9-dCas9/BBa_J23107-MCP-SoxS
pPPC004	pUC18T-miniTn7T-Gm	BBa_J23111-mRFP	J1-BBa_J23117-sfGFP	*Sp*.pCas9-dCas9/BBa_J23107-MCP-SoxS
pPPC005	pUC18T-miniTn7T-Gm		XylS-Pm-dCas9/BBa_J23107-MCP-SoxS	
pPPC006	pUC18T-miniTn7T-Gm		*Sp*.pCas9-dCas9/XylS-Pm-MCP-SoxS	
pPPC007	pUC18T-miniTn7T-Gm		XylS-Pm-dCas9/XylS-Pm-MCP-SoxS	
Integration plasmid	**Backbone**	**Integration site**	**Gene Island**	
pGNW2-pp1	pGNW2	pp1		
pPPC031	pGNW2	pp1	J3-BBa_J23117-mRFP	
pGNW2-pp2	pGNW2	pp2		
pPPC032	pGNW2	pp2	BBa_J23111-sfGFP	
pPPC033	pGNW2	pp2	J3(106)-BBa_J23117-sfGFP	
pPPC034	pGNW2	pp2	J3(106)-BBa_J23111-sfGFP	
Replicable plasmid	**Backbone**	**SacI/KpnI (NotI/Bsp120I)**	**Mph1103I (AatII/XhoI or KpnI/XhoI)**	**scRNA or sgRNA**
pPPC008	pBBR1-GmR	scRNA, sgRNA, scRNA/sgRNAs		hAAVS1, J101-121, J306, jfGFP1, hAAVS1(sgRNA), hAAVS1-jfGFP1, J306-hAAVS1(sgRNA), J306-hAAVS1, J306-jfGFP1,

				J306-J106, hAAVS1(sgRNA)-J106
pPPC009	pBBR1-KmR	scRNA, sgRNA, scRNA/sgRNAs		hAAVS1, J106, RR2, hAAVS1-RR2, J106-RR2
pPPC010	pBBR1-KmR	scRNA	*Sp*.pCas9-dCas9/BBa_J23107-MCP-SoxS	hAAVS1, J109,
pPPC011	pRK2-KmR		*Sp*.pCas9-dCas9/BBa_J23107-MCP-SoxS	
pPPC012	pBBR1-GmR		J1-BBa_J23117-mRFP	
pPPC013	pBBR1-KmR		J1-BBa_J23117-mRFP	
pPPC014	pRK2-GmR		J1-BBa_J23117-mRFP	
pPPC015	pRK2-KmR		J1-BBa_J23117-mRFP	
pPPC016	pBBR1-GmR	scRNA	J1-BBa_J23117-mRFP	hAAVS1, J106,J107, J108, J109
pPPC016.306	pBBR1-GmR	scRNA	J1-BBa_J23117-mRFP (swap J106 to J306)	hAAVS1, J306
pPPC017	pBBR1-KmR	scRNA	J1-BBa_J23117-mRFP	hAAVS1, J109
pPPC018	pRK2-GmR	scRNA	J1-BBa_J23117-mRFP	hAAVS1, J109
pPPC019	pRK2-KmR	scRNA	J1-BBa_J23117-mRFP	hAAVS1, J109
pPPC020	pBBR1-GmR	scRNA	J3-BBa_J23117-mRFP	hAAVS1, J306, J307, J308, J309
pPPC020.106	pBBR1-GmR	scRNA	J1-BBa_J23117-mRFP (swap J306 to J106)	hAAVS1, J106
pPPC021.J231XX	pBBR1-GmR	scRNA	J3-BBa_J231XX-mRFP where BBa_J231XX refers to either of J23109, J23113, J23114, J23115, J23107, J23105, J23106, J23108, J23111	hAAVS1, J306
pPPC021.J23110	pBBR1-GmR	scRNA	J3-BBa_J23110-mRFP	hAAVS1, J306, RR2
pPPC022.5PS	pBBR1-GmR	scRNA	J3(random-5PS)-BBa_J23117-mRFP	J306
pPPC023.5PSN	pBBR1-GmR	scRNA	J3(5PSN)-BBa_J23117-mRFP	hAAVS1, J306
pPPC024	pBBR1-GmR	scRNA	J3(106)-BBa_J23111-sfGFP_J3-BBa_J23117-mRFP	hAAVS1, J306, jfGFP1, J306-jfGFP1, J106, J306-J106

pPPC025	pBBR1-GmR	scRNA	J3(106)-BBa_J23117-sfGFP_J3-BBa_J23117-mRFP	hAAVS1, J306, J106, J306-J106, J306-hAAVS1
pPPC026.XN	pBBR1-GmR	scRNA	PP_NNNN-mRFP where PP_NNNN refers to 10 endogenous promoters as listed in DNA Sequences	hAAVVS1, AN-JN as listed in Table S4
pPPC027	pBBR1-GmR	scRNA	J3-BBa_J23117-GTPCH, J3-BBa_J23117-PTPS, J3-BBa_J23117-SR	hAAVS1, J306
pPPC028	pBBR1-GmR	scRNA	J3-BBa_J23117-GTPCH, J3-BBa_J23117-PTPS	hAAVS1, J306
pPPC029	pBBR1-GmR		LacI-Ptrc-mvaE-mvaS	
pPPC030	pBBR1-GmR	scRNA	J3-J23117-mvaE-mvaS	hAAVS1, J306
pCK241	pBBR1-GmR		LacI-Ptrc-mRFP	
pCK243	pBBR1-GmR		XylS-Pm-mRFP	
pCK255	pBBR1-GmR		I-SceI_sacB	

Table 2.S3: Cloning Primers

Name	Sequence	Descriptive name
oCDP057	ATTCGATCATGCATGTTACGAAATCATCCTGTGGAGCTT	pPPC001_*Sp*.pCas9_F
oCDP058	CAAGGCCTTCGCGAGGCGAAAAAACCCCGCCG	pPPC001_BBa_B1002_R
oCDP003	attcgatcatgcatgCTGCAGGCCTACGGTATCCACCGG	pPPC002_J1
oCDP002	caaggccttcgcgagAAGCTTtataaacgcagaaaggcccac	pPPC002_dblTerm_R
oCDP021	tgcgtttataAAGCTTTTACGAAATCATCCTGTGGAGC	pPPC002_*Sp*.pCas9_F
oCDP022	ccttcgcgagAAGCTTgcgaaaaaaccccgccg	pPPC002_BBa_B1102_R
oCDP061	AAGCTAATTCGATCatgcatttgacggctagctcagtcc	pPPC003_BBa_J23111_F
oCDP062	CGTAGGCCTGCAGcatataaacgcagaaaggcccacc	pPPC003_dblTerm_F
oCK257	gaagctaattcgatcatgcatgatttgtcctactcaggagagcg	pPPC005_XylS_F
oCK258	attgagtatttcttatccatatgtttttcctcc	pPPC005_Pm_R
oCK259	cggtttgcgtattgggcgcaatttgtcctactcaggagagcg	pPPC006_XylS_F
oCK260	ACGTCTTCGCTACTCGCCATatgtttttcctcctaaccgcg	pPPC006_Pm_R
oCK101	ccagctggcaattccgacgtcgtcgaatttgctttcgaatttctgc	pRK2-GmR_MCS_F
oCK102	aagaccggcggtcttaagtttttttggctgaagaattcgcaaatattatacgcaaggcgac	pRK2-GmR_MCS_R
oCK085	aacaggagtccaagcgcatggaagccatcacaaacg	pRK2-KmR_KmR_F
oCK086_short	gacgtcggaattgccagctggg	pRK2-KmR_KmR_R
oCDP023	tatagggcgaattggagctcTTGACAGCTAGCTCAGTCC	pPPC008_scRNA_F
oCDP008	gggaacaaaagctggTCTAGCTTAAGAGTTCACCGACAAACAACAGATA	pPPC008_scRNA_R
oCDP056	TCGGTGAACTCTTAAcGGATCCTTGACAGCTAGCTCAGTCCTAGG	2nd-gRNA_F
oCDP051	gctggTCTAGCTTAAGAGTTCACCGACAAACAACAGAT	2nd-gRNA_R
oCK079	GCTCAGTCCTAGGTATAATACTAGT	change-gRNA_F
oCK287	TAGGTATAATACTAGTNNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGT	new-gRNA_F
oCK077	ttgcgtattgggcgcaTTACGAAATCATCCTGTGGAGCTT	pPPC010_*Sp*.pCas9_F
oCK078	attacaacagtttttagcgaaaaaaccccgccg	pPPC010_*Sp*.pCas9_R
oCK097	ttgcgtattgggcgcatggcaattccgacgtc	pPPC012_J1_F
oCK098	attacaacagtttttaTATAAACGCAGAAAGGCCC	pPPC012_dblTerm_R
oCK237	GCGTTCTGGACACAATTGGGTTCCACCGGATACCTCCGGACttgacagctagctcagtcc	pPPC016_J106-to-J306_F

oCK279	TTGTGTCCAGAACGCTCCGTAGGACACCGCAGGATACCTGAGGTCGCCCG	pPPC016_J106-to-J306_R
oCK130	ccaccgcggtggcggccgcTTGACAGCTAGCTCAGTCCTAG	pPPC018_gRNA_F
oCK146	agctgggtaccgggcccAGTTCACCGACAAACAACAGATA	pPPC018_gRNA_R
oJF365	GCGGTTACCAAAGGCGTCCTCGTCGTCTTGAAGTTGCG	pPPC020_J306-to-J106_F
oJF366	GCCTTTGGTAACCGCAGGAGAAGTGAGGAGACGAGC	pPPC020_J306-to-J106_R
oCK177	ttgggcgcatggcaattccg	pPPC021_J3_F
oCK219	GCCTGGagatccttactcga	pPPC021_dblTerm_R
oJF447	GCGTTCTGGACACAANNNNNNNNNNNNNNNNNNNNNNNNNNNNttgacagctagctcagtcc	pPPC022_BBa_J23117_F
oJF448	TTGTGTCCAGAACGCTCCGTAGGAGAAG	pPPC022_J3_R
oCK084	cggtgcttaaaaactcgagtaaggatctCCAGG	pPPC022_dblTerm_F
oBT110	gctagcactatacctaggactgagctagccgtcaa	pPPC024_BBa_J23111_R
oBT072	CAGTCCTAGGTATAGTGCTAGCGAATTCATTAAAGAGGAG	pPPC024_BBa_J23111-RBS_F
oCK383	ATGATCGCAAATGCTgAGTACTtataaacgcagaaaggcccacccg	pPPC024_dblT_R
oCDP059	TTAAAGAGGAGAAAGGTACCATGGCGAGTAGCGAAGACG	pPPC026_mRFP_F
oCK251	tgcgcccaatacgcaaaccg	pPPC026_MCS_R
oCK253	cggtttgcgtattgggcgcagttctcagggctcgccgaga	pPPC026_PP_1776-A_F
oCK354	GGTACCTTTCTCCTCTTTAATGAATTCtgatactggccacagacggct	pPPC026_PP_1776-A_R
oCK169	gcgcatggcaattccgatatcAGCATTT	pPPC027_J3_F
oCK170	GGagatccttactcgagtttTTATTCGTCGTAGAAATCAATGTGG	pPPC027_SR_R
oCK073	ccagctggcaattccgacgtc	pPPC029_LacI_F
oCK072_Short	agatccttactcgagtttttaattacgatagctacgcacgg	pPPC029_mvaS_R
oCDP046	TTAAAGAGGAGAAAGGTACCatgaaaaccgtggtgattattgatg	pPPC030_mvaE_F
oCK325	tgcctgcaggtcgactctagatgaccgacctgatcgaagtgaagac	pGNW2-pp1_HR1_F
oCK326	atggcggATGCATgggctcggttctctactggcg	pGNW2-pp1_HR1_R
oCK327	cgagcccATGCATccgccattacgatctgacttgcc	pGNW2-pp1_HR2_F
oCK328	ataacagggtaatctgaatttcacttcactcgggaaaaatcagggg	pGNW2-pp1_HR2_R
oCK344	ccagtagagaaccgagcccAgcatggcaattccgacgtc	pPPC031_J3_F
oCK345	agtcagatcgtaatggcggATATAAACGCAGAAAGGCCC	pPPC031_dblT_R
oCK369	tgcctgcaggtcgactctagaccaggatgaataccttaaggacgcc	pGNW2-pp2_HR1_F
oCK370	acaccttATGCATgcgcgtgatgcgctgcacac	pGNW2-pp2_HR1_R
oCK371	cacgcgcATGCATaaggtgtattcccccggcattca	pGNW2-pp2_HR2_F

oCK372	ataacagggtaatctgaatttccgctcggcgcgtgatccgc	pGNW2-pp2_HR2_R
oCK373	agcgcatcacgcgcATGCATttgacggctagctcagtcctaggt	pPPC032_BBa_J23111_F
oCK374	cgggggaatacaccttAGTACTtataaacgcagaaaggcccacccg	pPPC032_dblT_R
oCK384	cagcgcatcacgcgcATGCATAGCATTTGCGATCATTCACGCAGC	pPPC033_J3_F
oCDP004	GGTACCTTTCTCCTCTTTAATGAATTC	RBS_R
oCD292	GAATTCATTAAAGAGGAGAAAGGTACC	RBS_F
oCD281	ATGGCGAGTAGCGAAGACGT	mRFP_F
oCK320	TTCGCTACTCGCCATGGTACCTTTCTCCTCTTTAATGAATTCtgaaattgttatccgctc	Ptrc_RBS_R
oCK259	cggtttgcgtattgggcgcaatttgtcctactcaggagagcg	XylS-Pm_F
oCK277	GGTACCTTTCTCCTCTTTAATGAATTCttgcataaagcctaaggggtaggccttactaga	Pm_RBS_R

Table 2.S4: Target sequences of scRNA and sgRNA

Name	Sequence	Target Promoter/Gene	Target Strand	Distance to TSS
sgRNA	NNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGC AAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGT GGCACCGAGTCGGTGCTTTTTTT			
scRNA_1x MS2.b2	NNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGC AAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGT GGCACATGAGGATCACCCATGTGCTTTTTTT			
hAAVS1	GGGGCCACTAGGGACAGGAT	Off-target	N.A.	N.A.
RR2	TGGAACCGTACTGGAACTGC	mRFP (CRISPRi)	Template	185 from ATG
jfGFP1	CATCTAATTCAACAAGAATT	sfGFP (CRISPRi)	Template	38 from ATG
J101	TGGGTTCCACCGGATACCTC	J1	Non-template	40
J102	AGGTATCCGGTGGAACCCAA	J1	Template	61
J103	AGGCGTCCTTTGGGTTCCAC	J1	Non-template	50
J104	TGGAACCCAAAGGACGCCTT	J1	Template	71
J105	CGGTTACCAAAGGCGTCCTT	J1	Non-template	60
J106	AGGACGCCTTTGGTAACCGC	J1, J3(106)	Template	81
J107	CGGTGTCCTGCGGTTACCAA	J1	Non-template	70
J108	TGGTAACCGCAGGACACCGC	J1	Template	91
J109	AGGTATCCTGCGGTGTCCTG	J1	Non-template	80
J110	AGGACACCGCAGGATACCTG	J1	Template	101
J111	GGGCGACCTCAGGTATCCTG	J1	Non-template	90
J112	AGGATACCTGAGGTCGCCCG	J1	Template	111
J113	GGGCCACCACGGGCGACCTC	J1	Non-template	100
J114	AGGTCGCCCGTGGTGGCCCA	J1	Template	121
J115	TGGTGACCATGGGCCACCAC	J1	Non-template	110
J116	TGGTGGCCCATGGTCACCAT	J1	Template	131
J117	GGGTGACCTATGGTGACCAT	J1	Non-template	120
J118	TGGTCACCATAGGTCACCCT	J1	Template	141
J119	TGGTTGCCAAGGGTGACCTA	J1	Non-template	130
J120	AGGTCACCCTTGGCAACCAA	J1	Template	151
J121	AGGACACCTTTGGTTGCCAA	J1	Non-template	140

J306	TTGTGTCCAGAACGCTCCGT	J3, J1(306)	Template	81
J307	ACTTCTCCTACGGAGCGTTC	J3	Non-template	70
J308	AACGCTCCGTAGGAGAAGTG	J3	Template	91
J309	TCGTCTCCTCACTTCTCCTA	J3	Non-template	80
A1	TTCATGTAGCTTGTCCCCCG	PP_1776-A	Template	82
A2	CCGACTGAAGATGCGCTCTC	PP_1776-A	Non-template	92
A3	ATGCGCTCTCTGGCGCTCCT	PP_1776-A	Non-template	82
A4	TGCGCTCTCTGGCGCTCCTC	PP_1776-A	Non-template	81
A5	GCGCTCTCTGGCGCTCCTCG	PP_1776-A	Non-template	80
A6	CGCTCTCTGGCGCTCCTCGG	PP_1776-A	Non-template	79
B1	ACTGGGATTTGTGTAGGAGC	PP_4812-B	Non-template	92
B2	GGGTTTACCCGCGAAAGGGC	PP_4812-B	Non-template	72
B3	GGCAGTGCCGGCCCTTTCGC	PP_4812-B	Template	82
B4	TGGCAGTGCCGGCCCTTTCG	PP_4812-B	Template	81
C1	AATGCGTGGTCGCTTAATCC	PP_3839-C	Non-template	82
C2	ATGCGTGGTCGCTTAATCCT	PP_3839-C	Non-template	81
C3	TAATCCTGGGTTAACCGGAC	PP_3839-C	Non-template	68
C4	TGCGCCGGTCCGGTTAACCC	PP_3839-C	Template	81
D1	GGCCCCTGCGCTGCGCTCCG	PP_1992-D	Non-template	72
D2	CAGCGCAGGGGCCGGATGAT	PP_1992-D	Template	99
D3	CCGGAGCGCAGCGCAGGGGC	PP_1992-D	Template	91
E1	TATCGATGAAATCGCAGCAT	PP_0786-E	Non-template	92
E2	CAGCATAGGCGATGCCTATG	PP_0786-E	Non-template	78
E3	CCTTAGACAATCCACCTCAT	PP_0786-E	Template	81
F1	AAAGCTGCGCCAGAGTGTCG	PP_1972-F	Non-template	69
F2	ACACTCTGGCGCAGCTTTTG	PP_1972-F	Template	89
F3	GACACTCTGGCGCAGCTTTT	PP_1972-F	Template	90
F4	CGACACTCTGGCGCAGCTTT	PP_1972-F	Template	91
G1	GGCGTCCTGGGCAAAGGGTA	PP_3668-G	Non-template	91
G2	CTGTGTATTGAAGCATGGCG	PP_3668-G	Non-template	68
G3	CACAGCCATACCCTTTGCCC	PP_3668-G	Template	103
H1	TATCCCACCCTCGCCATTTT	PP_5046-H	Non-template	88

H2	CCCTCGCCATTTTCGGGCAC	PP_5046-H	Non-template	81
H3	TGCCCGAAAATGGCGAGGGT	PP_5046-H	Template	102
H4	GTGCCCGAAAATGGCGAGGG	PP_5046-H	Template	101
H5	TGCATGCCAGTGCCCGAAAA	PP_5046-H	Template	92
I1	GGTTTTTGTAGTGCTTGTGC	PP_1231-I	Template	101
I2	GGGTTTTTGTAGTGCTTGTG	PP_1231-I	Template	100
I3	CAATCCAGCGATTACTAAAG	PP_1231-I	Template	80
I4	ACAATCCAGCGATTACTAAA	PP_1231-I	Template	79
J1	TGGGTATGGCAGGGGGATTT	PP_4701-J	Template	118
J2	GTGCTGGGAATGGGTATGGC	PP_4701-J	Template	108
J3	CCACGTGCTGGGAATGGGTA	PP_4701-J	Template	104

Table 2.S5: Summary of CRISPRa-mediated fold-changes in gene expression and metabolite production

Description	Reporter gene/ Metabolic genes	scRNA	Fold-change	Figure
Activation of plasmid-bourne mRFP with J1 promoter using a 2-plasmid system	J1-mRFP (2-plasmid)	J109	1.6-fold	Figure 1C
Activation of plasmid-bourne mRFP with J1 promoter using a strain with integrated dCas9/MCP-SoxS (PPC01)	J1-mRFP (pBBR1-GmR)	J109	5-fold	Figure 1C
Activation of plasmid-bourne mRFP with J3 promoter	J3-mRFP (pBBR1-GmR)	J306	34-fold	Figure 2D
Dual activation of mRFP and sfGFP reporters (plasmid-bourne)	J3-mRFP (pBBR1-GmR)	J306	41-fold	Figure S6
	J3(106)-sfGFP (pBBR1-GmR)	J106	105-fold	Figure S6
Dual activation of mRFP and sfGFP reporters (genomically-integrated)	J3-mRFP (integrated)	J306	3-fold	Figure S8A
	J3(106)-sfGFP (integrated)	J106	24-fold	Figure S8A
Activation of plasmid-bourne endogenous promoter reporter constructs (mRFP)	PP_1776-A-mRFP (pBBR1-GmR)	A2	1.7-fold	Figure 5A
	PP_1992-D-mRFP (pBBR1-GmR)	D3	1.7-fold	Figure 5A
	PP_0786-E-mRFP (pBBR1-GmR)	E1	2.5-fold	Figure 5A
	PP_3668-G-mRFP (pBBR1-GmR)	G2	2.8-fold	Figure 5A
Metabolite production upon activation of biopterin pathway genes	J3-GTPCH, J3-PTPS, J3-SR (pBBR1-GmR)	J306	5-fold	Figure 6D
Metabolite production upon activation of mevalonate pathway genes	J3-mvaES (pBBR1-GmR)	J306	>40-fold	Figure 7B

Fold-change values were calculated relative to a strain with an off-target scRNA. The J1 and J3 promoters shown in this table contain the BBa_J23117 minimal promoter.

Supplementary Methods

Media and Chemicals

E. coli and *P. putida* culture and engineering were generally performed in LB media. *Pseudomonas* isolation agar (Difco) was used in the tri-parental mating for mini-Tn7 cloning (Choi and Schweizer, 2006). Fluorescent protein reporter gene activation and metabolic engineering experiments were performed in EZ rich-defined media (Teknova) with 0.2% glucose as the carbon source, unless specified. Appropriate concentration of antibiotics were included for plasmid maintenance: 100 μg/mL for carbenicillin, 25 μg/mL for chloramphenicol, 30 μg/mL for gentamicin, 30 μg/mL for kanamycin. For two-plasmid transformations in this work, the antibiotic concentration was reduced by half to 15 μg/mL each of gentamicin and kanamycin. IPTG was prepared in water as a 1 M stock solution prior to use. *m*-Toluic acid was prepared as a 0.5 M stock solution in 50% DMSO/water. Biopterin, BH2, and BH4 (Cayman Chemical) were stored in DMSO and diluted into water prior to use. D,L-mevalonolactone (Sigma) was freshly prepared in ethanol as a 20 mg/mL solution before dilution in ethyl acetate.

Plasmids construction strategy

Plasmids used in this study can be separated into genome integration plasmids and replicable plasmids. Integration plasmids were made based on pUC18T-mini-Tn7T-Gm. Replicable plasmids were constructed from pBBR1_MCS2 (named pBBR1-KmR in this study), pBBR1_MCS5 (named pBBR1-GmR in this study), and pRK2-AraE (Table 1). The AraE cassette from pRK2-AraE (bearing GmR marker) was replaced with multiple-cloning-site regions of pBBR1 to generate pRK2-GmR. pRK2-KmR was made by replacing the GmR marker and AraE cassette with the KmR marker and its multiple-cloning-site from pBBR1-KmR. The CRISPRa components and genes of interest were incorporated into each backbone at the multiple-cloning-site region. The detailed methodology for construction of each backbone is provided below. Table S1 shows the list of strains and plasmids used in each figure. Plasmids descriptions are listed in Table S2.

MCP-SoxS(R93A/S101A) was used in this study and will be abbreviated as MCP-SoxS. Both dCas9 and MCP-SoxS were obtained from the pCD442 plasmid (Fontana et al., 2020a). The 1xMS2 scRNA.b2 was used in this study with variable 20 bp target sequences (Dong et al., 2018). The full sequence of sgRNA and scRNA are provided in the DNA sequences section. Any plasmid with sg/scRNA has different 20bp target sequences according to Table S2. sg/scRNA sequences were provided in Table S4.

Integration plasmids

<u>pPPC001 and pPPC005-007</u>

For pPPC001, the dCas9/MCP-SoxS cassette was amplified from pCD442 and inserted into pUC18T-miniTn7T-Gm with KpnI/SacI. For pPPC005, the dCas9/MCP-SoxS coding sequence was amplified from pCD442 and inserted together with XylS-Pm as a promoter of dCas9, amplified from pS448-CsR (Wirth et al., 2019). Further modification of the MCP-SoxS promoter was achieved by digestion of pPPC001/pPPC05 with PstI/Bsp120I and XylS-Pm was inserted into the corresponding site to give pPPC006/007, respectively. See Figure 2.S9 for representative plasmid maps.

pPPC002-004

For integration plasmids with the reporter gene included, J1-BBa_J23117-sfGFP was amplified from pJF076Sa (Fontana et al., 2020a) and inserted into the KpnI/SacI site of pUT18T-miniTn7T-Gm. Then, dCas9/MCP-SoxS was added to the HindIII site to give pPPC002. For pPPC003.N, the J1(+N)-BBa_J23117-sfGFP fragments were amplified from pJF155.1-12 (Fontana et al., 2020a) instead of pJF076Sa. In the case of pPPC004 with additional BBa_J23111-mRFP reporter, the corresponding reporter fragment was amplified from the pJF143.J3.J23111 (Fontana et al., 2020a), with BBa_J23111 promoter instead of a BBa_J23117, and inserted into pPPC002 at the Mph1103I cut site.

Replicable plasmids

pRK2-GmR was made by digesting pRK2-AraE (containing GmR marker) with AatII/BspTI and the multiple-cloning-site (MCS) from pBBR1 was inserted into the pRK2 backbone. pRK2-KmR was made by digestion of pRK2-AraE with SacI/BspTI and insert KmR and MCS fragments from pBBR1-KmR. Then, the further modification of these plasmids followed the general manipulation at the MCS. See Figure 2.S10 for representative plasmid maps.

scRNA (or sgRNA) was inserted into the replicable plasmid at the SacI/KpnI site of the MCS. Then, the reporter fragment was inserted at the Mph1103I region. The dCas9/MCP-SoxS cassette was inserted into Mph1103I. For pRK2-GmR and pRK2-KmR, the scRNA fragment was amplified from pPPC013 and inserted into pRK2 backbones at NotI/Bsp120I site due to conflicting SacI/KpnI cut sites in the pRK2 backbone.

To change the scRNA target sequence, the existing scRNA cassette was excised with SpeI/BspTI and the new scRNA fragment was inserted. To express multiple scRNAs from the same plasmid, additional scRNA (or sgRNA) cassettes can be inserted at the BspTI site. To generate a new scRNA fragment, any existing scRNA construct can be amplified with a forward primer binding at the promoter region, oCK079_GCTCAGTCCTAGGTATAATACTAGT. To introduce a new 20 base target sequence, a forward primer with the same overhang can be used, oCK287_TAGGTATAATACTAGTNNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGT, where the variable 20nt in oCK287 can be replaced with the desired target sequence.

To insert J1-mRFP reporter cassettes, the PCR fragment was amplified from pJF076Sa (Fontana et al., 2020a) as a template and cloned into the Mph1103I site. The J3-mRFP variants were constructed in the same manner with pJF143.J3 (Fontana et al., 2020a) as a PCR template. To insert other genes of interest under control of J1 or J3 promoters, several approaches are available. AatII was introduced upstream of J1 and J3 sequences. KpnI was added at the end of strong RBS and XhoI was added between the stop codon and terminator. The desired cassette can be cloned into the Mph1103I site directly or inserted at the aforementioned sites. In our hands, biopterin pathway genes were inserted with AatII/XhoI using pCK015 (J3-GTPCH-J3-PTPS-J3-SR) and pCK014 (J3-GTPCH-J3-PTPS) as templates for pPPC027-028. LacI-Ptrc was added into AatII/XhoI and mvaES was added into KpnI/XhoI using pMVA2RBS035 as a template for PCR to give pPPC029-030 respectively.

pCK014 and pCK015 were analogs of pPPC028 and pPPC027, respectively, in pSC101** origin for *E. coli* experiment which can be double transformed with pCK005.AAV and pCD581

(Fontana et al., 2020a). The *gtpch* gene was amplified from the *E. coli* MG1655 genome. *ptps* and *sr* from *M. alpina* were synthesized from GeneArt (Thermo-Fisher) with codon-optimization for expression in *E. coli* using Gene Designer (Atum). Each J3-CDS was individually added into the reporter cassette. Then, an additional J3-CDS construct was inserted into the existing one at the EcoRV site (altered from AatII due to the presence of cut-site in *sr* gene) to get pCK014 and pCK015.

Changing scRNA target sequence of J1 or J3

To alter the scRNA 20 base target sequence, we used a single-fragment PCR to change the existing 20 bp target of J106 in pPPC016 to the desired J306 using oCK237/oCK279 (Table S3). Then, the fragment was treated with DpnI, gel purified, and circularized with Infusion. The same method as used for converting J306 to J106 with oJF365/oJF366.

Construction of 5'-proximal sequence library (pPPC022)

To generate a library of different 26bp sequences upstream of a minimal promoter, a fragment with randomized 26bp region (5'-PS-BBa_J23117-mRFP) was constructed with the oJF447 and oCK219 primers (Table S3). pPPC020 bearing J306 scRNA was linearized by PCR with oJF448/oCK084 and treated with DpnI to remove the parent vector. Then the linearized pPPC016 backbone fragment and a randomized 26bp library fragment were assembled with Infusion.

Construction of 5'-proximal sequence variants characterized in *E. coli* (pPPC023)

pPPC023 was constructed similarly to pPPC022 as described above. Five oJF447 variants with known 26bp sequences (provided in the DNA Sequences section) were used to generate 5'-PSN-BBa_J23117-mRFP fragments (PS1 to PS5) for insertion into the linearized backbone.

Construction of dual reporter plasmids (pPPC024-025, pPPC031-034)

For the plasmid-based dual reporter for multi-gene CRISPRa with two strongly expressed fluorescent reporters, a J3(106)-BBa_J23117-sfGFP cassette was inserted at the AatII site of J3-BBa_J23117-mRFP (pPPC020) to generate pPPC024. The plasmid-based dual reporter for CRISPRi/a with weakly expressed mRFP and strongly expressed sfGFP was constructed by delivering J3(106)-BBa_J23111-sfGFP to pPPC020 to generate pPPC025. Multiple sgRNA/scRNA cassettes were delivered as described above in the Replicable plasmids section.

The genomically-integrated dual reporter strains were constructed by sequentially integrating separate mRFP and sfGFP reporters at different genomic sites. Plasmids, pGNW2-pp1 and pGNW2-pp2 were constructed from pGNW2 (Wirth et al., 2019) by addition of prophage1 (pp1) or prophage2 (pp2) regions into the XbaI/EcoRI site. Flanking homology sites (HR1 and HR2) were separated by an Mph1103I site for insertion of the desired heterologous gene. J3-BBa_J23117-mRFP was inserted into pGNW2-pp1 at the Mph1103I site to construct pPPC031. sfGFP constructs with different promoters were cloned into pGNW2-pp2 at the Mph1103I site to generate pPPC032-034.

Construction of endogenous promoter reporter (pPPC026)

The J3-BBa-J23117 reporter (pPPC020) was modified into an endogenous promoter reporter by replacing the J3-BBa_J23117 promoter with an intergenic region from each gene of interest. The intergenic region contained 60 bases from the ORF of interest on the 3' end. On the 5' end, the intergenic region extended 60 bp into the next upstream ORF, following a previously reported strategy (Zaslaver et al., 2006). The mRFP cassette, along with its original strong RBS, was included downstream of the 60 bp fragment of the ORF of interest. Complete sequences are provided below.

$CaCl_2$ chemically competent cell preparation

The chemically competent cell preparation was adapted from a prior method (Zhao et al., 2013). From an overnight culture seeded from a single colony of a *P. putida* strain in LB, the cell suspension was 100-fold diluted into 50 mL LB without antibiotic in a 250 mL Erlenmeyer flask. The culture was incubated at 30 °C to OD_{600} = 0.8 - 1.0, transferred to 2 x 50 mL conical tubes, and placed on ice for 5 minutes. The cell suspension was centrifuged at 4 °C for 10 min at 5000 rpm. After discarding the supernatant, the cells were washed with an ice-cold solution of 50 mM $MgCl_2$ + 10 mM $CaCl_2$ twice. The final pellets were resuspended in 15% glycerol + 100 mM $CaCl_2$ solution to give chemically competent cells. The competent cells can be stored at -80 °C for a month with negligible loss of activity.

For transformation, 50 ng of a *P. putida* compatible plasmid was added to 100 µL of $CaCl_2$ chemically competent cells in a 1.5 mL microcentrifuge tube. Cells were mixed gently and incubated on ice for 30 minutes. The incubated competent cells were subjected to heat-shock at 42 °C for 3 minutes and cooled on ice for another 5 minutes. Then, 900 µL of LB was added to the competent cells and cultures were shaken at 30 °C for 1.5 hours. The outgrowth competent cells were spun down at 10000 rcf, room temperature for 1 minute. After discarding ~900 µL of supernatant, cells were resuspended in residual media for plating on a pre-warmed agar plate with appropriate antibiotic selection.

Biopterin Quantification by HPLC-MS

The LC-MS quantification was adapted from the prior method (Ehrenworth et al., 2015). LC−MS analysis was completed using an Agilent 1100/1260 series system equipped with a 1260 ALS autosampler and a 6120 Single Quadrupole LC−MS with a Poroshell 120 SB-Aq 3.0 mm × 100 mm × 2.7 µm column and an electrospray ion source. LC conditions: solvent A—150 mM acetic acid with 0.1% formic acid; solvent B—methanol with 0.1% formic acid. Gradient: 4 min ramp from 95%:5%:0.2 (A:B:flow rate in mL/min) to 70%:30%:0.2, 6 min ramp to 40%:60%:0.2, 2 min ramp to 2%:98%:0.2, 2 min ramp to 2%:98%:0.5, 4 min at 2%:98%:0.5, 1 min ramp to 95%:5%:0.5, 7 min at 95%:5%:0.5, and 1.5 min post time. MS acquisition (positive ion mode) included 25% scan from m/z 100−600, 25% scan from m/z 230−260, 25% scan from m/z 145−165, and 25% selective ion monitoring (SIM) for BH4 (m/z 242.1), dihydrobiopterin (m/z 240.1), and biopterin (m/z 238.1). Retention times were determined using commercially available standards (BH4, BH2, and biopterin from Cayman Chemical).

Determination of *P. putida* growth rate

Single colonies from LB plates were inoculated in 500 µL of EZ-RDM (Teknova) supplemented with the appropriate antibiotics and grown in 96-deep-well plates at 30 °C with shaking overnight. From the overnight cultures, OD_{600} of each replicate was measured in a 1-cm cuvette, then diluted to OD_{600} = 0.1 (30-50 fold dilution) and 200 µL of each diluted culture were grown in flat bottom microplate at 30 °C in a Biotek Synergy HTX plate reader for 16 hours with continuous slow orbital shaking.

DNA sequences

The DNA sequence of biological parts used in this study were provided below in the 5'-to-3' format of the non-template strand. Each part was color-coded based on its function: red-promoter, blue-CDS, brown-upstream sequence, dark green-5'-UTR, purple-terminator, and others as specified. The **transcription start site** (TSS), **start codon** and **stop codon** are **bolded**.

>*Sp*.pCas9-dCas9-dblTerm

TTACGAAATCATCCTGTGGAGCTTAGTAGGTTTAGCAAGATGGCAGCGCCTAAATGTAGAATGATAAAAGGATTAAG
AGATTAATTTCCCTAAAAATGATAAAACAAGCGTTTTGAAAGCGCTTGTTTTTTTGGTTTGCAGTCAGAGTAGAATA
GAAGTATCAAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATG
CTGTTTTGAATGGTTCCAACAAGATTATTTTATAACTTTTATAACAAATAATCAAGGAGAAATTCAAAGAAATTTAT
CAGCCATAAAACAATACTTAATACTATAGAATGATAACAAAATAAACTACTTTTTAAAAGAATTTTGTGTTATAATC
TATTTATTATTAAGTATTGGGTAATATTTTTTGAAGAGATATTTTGAAAAAGAAAAATTAAAGCATATTAAACTAAT
TTCGGAGGTCATTAAAACTATTATTGAAATCATCAAACTCATTATGGATTTAATTTAAACTTTTTATTTTAGGAGGC
AAAA**ATG**GATAAGAAATACTCAATAGGCTTAGcTATCGGCACAAATAGCGTCGGATGGGCGGTGATCACTGATGAAT
ATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATACAGACCGCCACAGTATCAAAAAAAATCTTATAGGGGCT
CTTTTATTTGACAGTGGAGAGACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAA
GAATCGTATTTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTCTTTCATCGACTTG
AAGAGTCTTTTTTGGTGGAAGAAGACAAGAAGCATGAACGTCATCCTATTTTTGGAAATATAGTAGATGAAGTTGCT
TATCATGAGAAATATCCAACTATCTATCATCTGCGAAAAAAATTGGTAGATTCTACTGATAAAGCGGATTTGCGCTT
AATCTATTTGGCCTTAGCGCATATGATTAAGTTTCGTGGTCATTTTTTGATTGAGGGAGATTTAAATCCTGATAATA
GTGATGTGGACAAACTATTTATCCAGTTGGTACAAACCTACAATCAATTATTTGAAGAAAACCCTATTAACGCAAGT
GGAGTAGATGCTAAAGCGATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATCTCATTGCTCAGCTCCC
CGGTGAGAAGAAAAATGGCTTATTTGGGAATCTCATTGCTTTGTCATTGGGTTTGACCCCTAATTTTAAATCAAATT
TTGATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAAGATACTTACGATGATGATTTAGATAATTTATTGGCGCAA
ATTGGAGATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATTTATCAGATGCTATTTTACTTTCAGATATCCTAAG
AGTAAATACTGAAATAACTAAGGCTCCCCTATCAGCTTCAATGATTAAACGCTACGATGAACATCATCAAGACTTGA
CTCTTTTAAAAGCTTTAGTTCGACAACAACTTCCAGAAAAGTATAAAGAAATCTTTTTTGATCAATCAAAAAACGGA
TATGCAGGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTATAAATTTATCAAACCAATTTTAGAAAAAATGGA
TGGTACTGAGGAATTATTGGTGAAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGGACCTTTGACAACGGCTCTA
TTCCCCATCAAATTCACTTGGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAAGAC
AATCGTGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCCATTGGCGCGTGGCAATAGTCG
TTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTT
CAGCTCAATCATTTATTGAACGCATGACAAACTTTGATAAAAATCTTCCAAATGAAAAAGTACTACCAAAACATAGT
TTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTGAAGGAATGCGAAAACCAGC
ATTTCTTTCAGGTGAACAGAAGAAAGCCATTGTTGATTTACTCTTCAAAACAAATCGAAAAGTAACCGTTAAGCAAT
TAAAAGAAGATTATTTCAAAAAAATAGAATGTTTTGATAGTGTTGAAATTTCAGGAGTTGAAGATAGATTTAATGCT
TCATTAGGTACCTACCATGATTTGCTAAAAATTATTAAAGATAAAGATTTTTTGGATAATGAAGAAAATGAAGATAT
CTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGAGATGATTGAGGAAAGACTTAAAACATATGCTC
ACCTCTTTGATGATAAGGTGATGAAACAGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTG
ATTAATGGTATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTGAAATCAGATGGTTTTGCCAATCGCAA
TTTTATGCAGCTGATCCATGATGATAGTTTGACATTTAAAGAAGACATTCAAAAAGCACAAGTGTCTGGACAAGGCG
ATAGTTTACATGAACATATTGCAAATTTAGCTGGTAGCCCTGCTATTAAAAAAGGTATTTTACAGACTGTAAAAGTT
GTTGATGAATTGGTCAAAGTAATGGGGCGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATCAGAC
AACTCAAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGAAGAAGGTATCAAAGAATTAGGAAGTCAGA
TTCTTAAAGAGCATCCTGTTGAAAATACTCAATTGCAAAATGAAAAGCTCTATCTCTATTATCTCCAAAATGGAAGA
GACATGTATGTGGACCAAGAATTAGATATTAATCGTTTAAGTGATTATGATGTCGATgcCATTGTTCCACAAAGTTT
CCTTAAAGACGATTCAATAGACAATAAGGTCTTAACGCGTTCTGATAAAAATCGTGGTAAATCGGATAACGTTCCAA
GTGAAGAAGTAGTCAAAAAGATGAAAAACTATTGGAGACAACTTCTAAACGCCAAGTTAATCACTCAACGTAAGTTT
GATAATTTAACGAAAGCTGAACGTGGAGGTTTGAGTGAACTTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGA
AACTCGCCAAATCACTAAGCATGTGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGATGAAAATGATAAAC
TTATTCGAGAGGTTAAAGTGATTACCTTAAAATCTAAATTAGTTTCTGACTTCCGAAAAGATTTCCAATTCTATAAA
GTACGTGAGATTAACAATTACCATCATGCCCATGATGCGTATCTAAATGCCGTCGTTGGAACTGCTTTGATTAAGAA
ATATCCAAAACTTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTCGTAAAATGATTGCTAAGTCTG

AGCAAGAAATAGGCAAAGCAACCGCAAAATATTTCTTTTACTCTAATATCATGAACTTCTTCAAAACAGAAATTACA
CTTGCAAATGGAGAGATTCGCAAACGCCCTCTAATCGAAACTAATGGGGAAACTGGAGAAATTGTCTGGGATAAAGG
GCGAGATTTTGCCACAGTGCGCAAAGTATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAGAAGTACAGACAG
GCGGATTCTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTATTGCTCGTAAAAAAGACTGGGATCCA
AAAAAATATGGTGGTTTTGATAGTCCAACGGTAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAATC
GAAGAAGTTAAAATCCGTTAAAGAGTTACTAGGGATCACAATTATGGAAAGAAGTTCCTTTGAAAAAAATCCGATTG
ACTTTTTAGAAGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAATCATTAAACTACCTAAATATAGTCTTTTTGAG
TTAGAAAACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAATTACAAAAAGGAAATGAGCTGGCTCTGCCAAGCAA
ATATGTGAATTTTTTATATTTAGCTAGTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAAT
TGTTTGTGGAGCAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCGTGTTATTTTA
GCAGATGCCAATTTAGATAAAGTTCTTAGTGCATATAACAAACATAGAGACAAACCAATACGTGAACAAGCAGAAAA
TATTATTCATTTATTTACGTTGACGAATCTTGGAGCTCCCGCTGCTTTTAAATATTTTGATACAACAATTGATCGTA
AACGATATACGTCTACAAAAGAAGTTTTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTTATGAAACACGC
ATTGATTTGAGTCAGCTAGGAGGTGAC**TAA**CTCGAGTAAGGATCTCCAGGCATCAAATAAAACGAAAGGCTCAGTCG
AAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGG
TGGGCCTTTCTGCGTTTATA

>BBa_J23107-MCP-GGGGS_linker-SoxS (R93A, S101A) (5'-UTR)

TTTACGGCTAGCTCAGCCCTAGGTATTATGCTAGC**G**AATTCATTAAAGAGGAGAAAGGTACC**ATG**GGGCCCGCTTCT
AACTTTACTCAGTTCGTTCTCGTCGACAATGGCGGAACTGGCGACGTGACTGTCGCCCCAAGCAACTTCGCTAACGG
GATCGCTGAATGGATCAGCTCTAACTCGCGTTCACAGGCTTACAAAGTAACCTGTAGCGTTCGTCAGAGCTCTGCGC
AGAATCGCAAATACACCATCAAAGTCGAGGTGCCTAAAGGCGCCTGGCGTTCGTACTTAAATATGGAACTAACCATT
CCAATTTTCGCCACGAATTCCGACTGCGAGCTTATTGTTAAGGCAATGCAAGGTCTCCTAAAAGATGGAAACCCGAT
TCCCTCAGCAATCGCAGCAAACTCCGGCATCTACGGTGGCGGAGGTAGCATGTCCCATCAGAAAATTATTCAGGATC
TTATCGCATGGATTGACGAGCATATTGACCAGCCGCTTAACATTGATGTAGTCGCAAAAAAATCAGGCTATTCAAAG
TGGTACTTGCAACGAATGTTCCGCACGGTGACGCATCAGACGCTTGGCGATTACATTCGCCAACGCCGCCTGTTACT
GGCCGCCGTTGAGTTGCGCACCACCGAGCGTCCGATTTTTGATATCGCAATGGACCTGGGTTATGTCTCGCAGCAGA
CCTTCTCCCGCGTTTTCGCGCGGCAGTTTGATCGCACTCCCGCGGATTATCGCCACCGCCTG**TAA**GCGGCCGCCACG
CAAAAAACCCCGCTTCGGCGGGGTTTTTTCGC

>XylS-Pm-CDS (5'-UTR)

ATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAACAGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCTTTTC
GTTTTATTTGATGCCTTTAATTAAACGTTCGTAATCAAGCCACTTCCTTTTTGCATTGACGCAGGGTGTCGGAAGGC
AACTCGCCGAACGCGCTCCTATAGTTTTCAGCGAAGCGTCCCAAATGTAAGAAGCCGTAGTCTAGGGCTATCTCAGT
TATACTACGCACATTGGCACTGGGATCGTTCAAGCAGGCGCGGATGCTTTCGAGCTTGCGGTTGCGGATGTAGTTCT
TCGGCGTGGTGCCGGCGTGCTTCTCGAACAAATTGTAGAGCGAGCGTGGACTCATCATCGCCAGCTCCGCTAACCGC
TCAAGGCTGATATTCCGTTTGAGATTCTCCTCAATGAATTGAACGACTCGCTCGAAAGACGGGTTACCTTTGCTGAA
AATTTCACGGCTGACATTGCTGCCCAGCATTTCGAGCAGCTTGGAAGCGATGATCCCCGCATAGTGCTCTTGGACCC
GAGGCATCGACTTTGTATGTTCCGCTTCGTCACAAACTAACCCGAGTAGATTGATAAAGCCATCGAGTTGCTGGAGA
TTGTGTCGCGCGGCGAAACGGATACCCTCCCTCGGCTTGTGCCAATTGTTGTCACTGCACGCCCGATCAAGGACCAC
TGAGGGCAATTTAACGATAAATTTCTCGCAATCTTCTGAATAGGTCAGGTCGGCTTGGTCATCCGGATTGAGCAGCA
ATAGTTCGCCCGGCGCAAAATAGTGCTCCTGGCCATGGCCACGCCACAGGCAATGGCCTTTGAGTATTATTTGCAGA
TGATAACAGGTTTCTAATCCAGGCGAGATTACCCTCACGCTACCGCCGTAGCTGATTCGACACAGATCGAGGCATCC
GAAGATTCTGTGGTGCAGCCTGCCTGCCGGGCGCCCGCCCTTGGGCAGGCGAATAGAGTGCGTACCGACATACTGGT
TAACATAATCGGAGACTGCATAGGGCTCGGCGTGGACGAAGATCTGACTTTTCTCGTTCAATAAGCAAAAATCCATA
GTTCACGGTTCTCTTATTTTAATGTGGGCTGCTTGGTGTGATGTAGAAAGGCGCCAAGTCGATGAAAATGCATCTCG
ACGTGATGCGTATACGGGTTACCCCCATTGCCACGTTGCGCCATCCTTTTTGCAATCAGTGACCACTTTTCCAAGCA
AAAATAACGCCAAGCAGAACGAAGACGTTCTTTTTAAGAAGCGAGAACACCAGAAGTTCGTGCTGTCGGGGCATGGG
GCGACGAATTGGCGGATAAAGGGGATCTGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAG
TTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGTAATTACGTATGGCAATGAAAGACGG
TGAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCT
GGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTG

GCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGA
TTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGCATGGGAATTAGCTTGATCTGACCAACG
ACCGGTAGCGGAGCTATCCAACGGCGGTATACCAGGAAAACACACAGCAGGTACATCAGAACAGTACCATGACTGAA
GAACAAATAGTTTTTTCCTGATCCATAAAGCAGAACGGCCTGCTCCATGACAAATCTGGCTCCCCAACTAATGCCCC
ATGCAGCCAGCATAACCAGCATAAAGTGCAGTGTCCGGTTTGATAGGGATAAGTCCAGCCTTGCAAGAAGCGGATAC
AGGAGTGCAAAAAATGGCTATCTCTAGTAAGGCCTACCCCTTAGGCTTTATGCAACAGAAACAATAATAATGGAGTC
ATGACCATGCCTAGGCCGCGGTTAGGAGGAAAAACAT**ATG___**

>J1-BBa_J23117-sfGFP (5'-UTR)

GCCTACGGTATCCACCGGAGACCTATGGCAGCCTCCGGCCGCCATAGGACACCTTTGGTTGCCAAGGGTGACCTATG
GTGACCATGGGCCACCACGGGCGACCTCAGGTATCCTGCGGTGTCCTGCGGTTACCAAAGGCGTCCTTTGGGTTCCA
CCGGATACCTCCGGACTTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGC**G**AATTCATTAAAGAGGAGAAAGGTAC
C**ATG**AGCAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACA
AATTTTCTGTCCGTGGAGAGGGTGAAGGTGATGCTACAAACGGAAAACTCACCCTTAAATTTATTTGCACTACTGGA
AAACTACCTGTTCCGTGGCCAACACTTGTCACTACTCTGACCTATGGTGTTCAATGCTTTTCCCGTTATCCGGATCA
CATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAACGCACTATATCTTTCAAAGATG
ACGGGACCTACAAGACGCGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATCGTATCGAGTTAAAGGGTATT
GATTTTAAAGAAGATGGAAACATTCTTGGACACAAACTCGAGTACAACTTTAACTCACACAATGTATACATCACGGC
AGACAAACAAAAGAATGGAATCAAAGCTAACTTCAAAATTCGCCACAACGTTGAAGATGGTTCCGTTCAACTAGCAG
ACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCGACACAATCT
GTCCTTTCGAAAGATCCCAACGAAAAGCGTGACCACATGGTCCTTCTTGAGTTTGTAACTGCTGCTGGGATTACACA
TGGCATGGATGAGCTCTACAAA**TAA**

>J3-BBa_J23111-mRFP (5'-UTR)

AGCATTTGCGATCATTCACGCAGCGCTTATTCAGTTGCTCACTGCGATGTCATAATCATCGCTACGAGCTGTGAAAG
ATGCATAAAGCTCGTACGACGCGTTCGCTCGTCTCCTCACTTCTCCTACGGAGCGTTCTGGACACAACGTCGTCTTG
AAGTTGCGATTATAGATTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGC**G**AATTCATTAAAGAGGAGAAAGGTAC
C**ATG**GCGAGTAGCGAAGACGTTATCAAAGAGTTCATGCGTTTCAAAGTTCGTATGGAAGGTTCCGTTAACGGTCACG
AGTTCGAAATCGAAGGTGAAGGTGAAGGTCGTCCGTACGAAGGTACCCAGACCGCTAAACTGAAAGTTACCAAAGGT
GGTCCGCTGCCGTTCGCTTGGGACATCCTGTCCCCGCAGTTCCAGTACGGTTCCAAAGCTTACGTTAAACACCCGGC
TGACATCCCGGACTACCTGAAACTGTCCTTCCCGGAAGGTTTCAAATGGGAACGTGTTATGAACTTCGAAGACGGTG
GTGTTGTTACCGTTACCCAGGACTCCTCCCTGCAAGACGGTGAGTTCATCTACAAAGTTAAACTGCGTGGTACCAAC
TTCCCGTCCGACGGTCCGGTTATGCAGAAAAAAACCATGGGTTGGGAAGCTTCCACCGAACGTATGTACCCGGAAGA
CGGTGCTCTGAAAGGTGAAATCAAAATGCGTCTGAAACTGAAAGACGGTGGTCACTACGACGCTGAAGTTAAAACCA
CCTACATGGCTAAAAAACCGGTTCAGCTGCCGGGTGCTTACAAAACCGACATCAAACTGGACATCACCTCCCACAAC
GAAGACTACACCATCGTTGAACAGTACGAACGTGCTGAAGGTCGTCACTCCACCGGTGCT**TAA**

>J3-BBa_J23117-EcGTPCH (5'-UTR)

AGCATTTGCGATCATTCACGCAGCGCTTATTCAGTTGCTCACTGCGATGTCATAATCATCGCTACGAGCTGTGAAAG
ATGCATAAAGCTCGTACGACGCGTTCGCTCGTCTCCTCACTTCTCCTACGGAGCGTTCTGGACACAACGTCGTCTTG
AAGTTGCGATTATAGATTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGC**G**AATTCATTAAAGAGGAGAAAGGTAC
C**ATG**CATCACCATCACCATCACCCATCACTCAGTAAAGAAGCGGCCCTGGTTCATGAAGCGTTAGTTGCGCGAGGAC
TGGAAACACCGCTGCGCCCGCCCGTGCATGAAATGGATAACGAAACGCGCAAAAGCCTTATTGCTGGTCATATGACC
GAAATCATGCAGCTGCTGAATCTCGACCTGGCTGATGACAGTTTGATGGAAACGCCGCATCGCATCGCTAAAATGTA
TGTCGATGAAATTTTCTCCGGTCTGGATTACGCCAATTTCCCGAAAATCACCCTCATTGAAAACAAAATGAAGGTCG
ATGAAATGGTCACCGTGCGCGATATCACTCTGACCAGCACCTGTGAACACCATTTTGTTACCATCGATGGCAAAGCG
ACGGTGGCCTATATCCCGAAAGATTCGGTGATCGGTCTGTCAAAAATTAACCGCATTGTGCAGTTCTTTGCCCAGCG
TCCGCAGGTGCAGGAACGTCTGACGCAGCAAATTCTTATTGCGCTACAAACGCTGCTGGGCACCAATAACGTGGCTG
TCTCGATCGACGCGGTGCATTACTGCGTGAAGGCGCGTGGCATCCGCGATGCAACCAGTGCCACGACAACGACCTCT
CTTGGTGGATTGTTCAAATCCAGTCAGAATACGCGCCACGAGTTTCTGCGCGCTGTGCGTCATCACAAC**TAA**

>J3-BBa_J23117-MaPTPS (5'-UTR)

AGCATTTGCGATCATTCACGCAGCGCTTATTCAGTTGCTCACTGCGATGTCATAATCATCGCTACGAGCTGTGAAAG
ATGCATAAAGCTCGTACGACGCGTTCGCTCGTCTCCTCACTTCTCCTACGGAGCGTTCTGGACACAACGTCGTCTTG

AAGTTGCGATTATAGATTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGC**G**AATTCATTAAAGAGGAGAAAGGTAC
C**ATG**CATCACCACCATCACCATACGTCCTCAACTCCAGTTAGAACTGCTTACGTTACCAGGATCGAACACTTCTCCG
CTGCGCACAGATTGAACTCCGTCCACCTCTCGCCTGCTGAGAACGTCAAGCTCTTCGGTAAGTGCAACCACACTTCC
GGTCACGGTCACAACTACAAGGTCGAGGTGACCATCAAGGGTCAGATCAACCCACAATCCGGCATGGTCATCAACAT
CACCGATCTTAAGAAGACTTTGCAAGTCGCTGTCATGGACCCTTGTGACCATAGAAACTTGGATATAGACGTCCCAT
ACTTCGAGTCCAGACCCTCCACTACTGAGAACCTCGCTGTCTTCTTGTGGGAGAATATCAAGAGCCACTTGCCACCT
TCCGACGCGTACGATTTGTACGAGATCAAGTTGCACGAAACCGACAAGAACGTTGTCGTTTACAGAGGTGAA**TAA**

>J3-BBa_J23117-MaSR (5'-UTR)

AGCATTTGCGATCATTCACGCAGCGCTTATTCAGTTGCTCACTGCGATGTCATAATCATCGCTACGAGCTGTGAAAG
ATGCATAAAGCTCGTACGACGCGTTCGCTCGTCTCCTCACTTCTCCTACGGAGCGTTCTGGACACAACGTCGTCTTG
AAGTTGCGATTATAGATTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGC**G**AATTCATTAAAGAGGAGAAAGGTAC
C**ATG**CATCACCATCACCACCATAGCAGTAAAGAACATCATTTGGTTATTATTAACGGTGTTAATAGAGGTTTCGGGC
ACTCCGTCGCGTTGGATTACATAAGACACTCAGGTGCTCACGCGGTGTCCTTTGTCTTGGTTGGTAGAACCCAGCAT
TCCTTGGAGCAAGTCTTAACGGAGCTGCACGAGGCTGCATCCCACGCTGGTGTCGTCTTCAAGGGTGTCGTTGTGTC
CGAGGTCGACCTGGCTCACTTGAACTCCCTCGACTCCAACCTCGCGAGGATACAGTCCGCCGCCGCTGACCTAAGAG
ACGAGGCGGCGCAAAGCACCAGAACTATCACTAAGTCGGTCCTCTTCAACAACGCGGGTAGCTTGGGTGACTTGTCC
AAGACTGTTAAGGAGTTCACCTGGCAAGAGGCTCGTTCCTACCTCGATTTCAACGTCGTGTCCCTCGTTGGTTTGTG
CTCCATGTTCTTGAAGGATACCCTCGAAGCATTCCCAAAGGAACAATACCCAGATCATAGAACTGTGGTCGTGTCCA
TCTCTTCCCTATTAGCTGTTCAGGCTTTCCCAAACTGGGGTTTGTACGCTGCTGGTAAGGCAGCTAGAGATAGACTA
TTAGGTGTTATTGCTCTCGAAGAAGCAGCTAATAACGTAAAGACCTTGAACTACGCTCCAGGTCCATTGGATAACGA
AATGCAGGCTGACGTCCGCAGAACTTTGGGTGATAAGGAACAACTGAAGATCTACGACGACATGCATAAGTCTGGTT
CCTTGGTGAAGATGGAGGACTCCTCTAGAAAGTTGATTCATTTGTTAAAGGCTGACACCTTCACCTCCGGTGGCCAC
ATTGATTTCTACGACGAA**TAA**

>LacI-Ptrc-*mvaE-mvaS* (5'-UTR)

CCAGCTGGCAATTCCGACGTCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGA
GAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGA
CCGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAG
CTGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAG
TCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGG
TGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGG
CTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATT
TCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGC
ATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTG
GCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTC
CGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGG
CGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGATATCTCGGTAGTGGGATACGACGAT
ACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGT
GGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAA
AAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAG
GTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAAGTTAGCGCGAATTGATCTGGTTTGACAGCT
TATCATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCG
TAAATCACTGCATAATTCGTGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTGCGCCGACATCATAACGG
TTCTGGCAAATATTCTGAAATGAGCTGTTGACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATA
ACAATTTCAGAATTCAAAAGATCTTTTAAGGACGAAACGTACAT**ATG**AAAACCGTGGTGATTATTGATGCACTGCGT
ACCCCGATTGGTAAATACAAAGGTAGCCTGAGCCAGGTTAGCGCAGTTGATCTGGGCACCCATGTTACCACCCAGCT
GCTGAAACGTCATAGCACCATTAGCGAAGAAATTGATCAGGTGATTTTTGGCAATGTTCTGCAGGCAGGTAATGGTC
AGAATCCGGCACGTCAGATTGCAATTAATAGCGGTCTGAGCCATGAAATTCCGGCAATGACCGTTAATGAAGTTTGT
GGTAGCGGTATGAAAGCAGTTATTCTGGCAAAACAGCTGATCCAGCTGGGCGAAGCCGAAGTTCTGATTGCCGGTGG
TATTGAAAATATGAGCCAGGCACCGAAACTGCAGCGTTTCAATTATGAAACCGAAAGCTATGATGCACCGTTTAGCA
GCATGATGTATGATGGTCTGACCGATGCATTTAGCGGTCAGGCAATGGGTCTGACAGCAGAAAATGTTGCAGAAAAA
TATCATGTGACCCGTGAAGAACAGGATCAGTTTAGCGTTCATAGCCAGCTGAAAGCAGCACAGGCACAGGCCGAAGG
TATTTTCGCAGATGAAATTGCACCGCTGGAAGTTAGCGGCACCCTGGTTGAAAAAGATGAAGGTATTCGTCCGAATA
GCAGCGTTGAAAAACTGGGTACACTGAAAACGGTGTTTAAAGAAGATGGCACCGTTACCGCAGGCAATGCAAGTACC
ATTAATGATGGTGCAAGCGCACTGATTATTGCCAGCCAAGAATATGCCGAAGCACATGGTCTGCCGTATCTGGCAAT

TATTCGTGATAGCGTTGAAGTTGGTATTGATCCGGCATATATGGGTATTAGCCCGATTAAAGCAATTCAGAAACTGC
TGGCACGTAATCAGCTGACCACCGAAGAAATCGACCTGTACGAAATTAATGAAGCATTTGCCGCAACCAGCATTGTT
GTTCAGCGTGAACTGGCACTGCCGGAAGAAAAAGTTAACATTTATGGCGGTGGCATCAGCCTGGGTCATGCAATTGG
TGCAACCGGTGCACGTCTGCTGACCAGCCTGAGCTATCAGCTGAATCAGAAAGAGAAAAAATACGGCGTTGCAAGCC
TGTGTATTGGTGGTGGCCTGGGTCTGGCAATGCTGCTGGAACGCCCTCAACAGAAAAAAAACAGCCGTTTTTATCAG
ATGAGTCCGGAAGAACGTCTGGCCAGCCTGCTGAATGAAGGTCAGATTAGCGCAGATACCAAAAAAGAATTTGAAAA
CACCGCACTGAGCAGCCAGATTGCCAACCACATGATTGAAAATCAGATCAGCGAAACCGAAGTGCCGATGGGTGTTG
GTCTGCATCTGACCGTGGATGAAACGGATTATCTGGTTCCGATGGCAACCGAAGAACCGAGCGTTATTGCAGCCCTG
AGCAATGGTGCAAAAATTGCACAGGGCTTTAAAACCGTGAATCAGCAGCGTCTGATGCGTGGTCAGATTGTTTTTTA
TGATGTTGCCGATGCAGAAAGCCTGATTGATGAACTGCAGGTTCGTGAAACAGAAATTTTCCAGCAGGCAGAACTGA
GTTATCCGAGCATTGTTAAACGCGGTGGTGGTCTGCGTGATCTGCAGTATCGTGCATTTGATGAAAGTTTTGTTAGC
GTGGATTTTCTGGTGGATGTTAAAGACGCAATGGGTGCCAATATTGTTAATGCAATGCTGGAAGGTGTTGCCGAACT
GTTTCGTGAATGGTTTGCAGAACAAAAAATCCTGTTTAGCATCCTGAGTAACTATGCCACCGAAAGCGTTGTTACCA
TGAAAACAGCAATTCCGGTTAGCCGTCTGAGCAAAGGTAGTAATGGTCGTGAAATTGCCGAAAAAATTGTTCTGGCA
AGCCGTTATGCCAGCCTGGATCCGTATCGTGCCGTTACCCATAATAAAGGTATTATGAATGGCATTGAAGCAGTTGT
GCTGGCCACCGGTAATGATACCCGTGCAGTTAGCGCAAGCTGTCATGCATTTGCAGTTAAAGAAGGTCGTTATCAGG
GTCTGACCAGCTGGACCCTGGATGGTGAGCAGCTGATTGGTGAAATTAGCGTTCCGCTGGCACTGGCAACCGTTGGT
GGTGCCACCAAAGTTCTGCCGAAAAGCCAGGCAGCAGCCGATCTGCTGGCAGTTACCGATGCAAAAGAACTGAGCCG
TGTTGTTGCAGCAGTTGGTCTGGCACAGAATCTGGCAGCACTGCGTGCACTGGTTAGCGAAGGCATTCAGAAAGGTC
ACATGGCACTGCAGGCACGTTCACTGGCCATGACCGTGGGTGCGACCGGTAAAGAAGTTGAAGCCGTTGCACAGCAA
CTGAAACGCCAGAAAACAATGAATCAGGATCGTGCCCTGGCAATTCTGAATGATCTGCGTAAACAG**TAA**TGATTAGC
GACAAAATATGAGGAGTGCAAAAA**ATG**ACCATTGGCATCGACAAAATCAGCTTTTTTGTTCCGCCTTACTATATCGA
CATGACCGCACTGGCCGAAGCACGTAATGTTGATCCGGGTAAATTTCATATTGGTATTGGTCAGGATCAGATGGCCG
TTAATCCGATTAGCCAGGATATTGTTACCTTTGCAGCAAATGCAGCAGAAGCAATTCTGACCAAAGAAGATAAAGAA
GCCATCGATATGGTTATTGTTGGCACCGAAAGCAGCATTGATGAAAGCAAAGCAGCCGCAGTTGTTCTGCATCGTCT
GATGGGTATTCAGCCGTTTGCACGTAGCTTTGAAATTAAAGAAGGTTGTTACGGCGCAACCGCAGGTCTGCAGCTGG
CAAAAAATCATGTTGCACTGCATCCGGATAAAAAAGTTCTGGTTGTTGCAGCAGATATCGCCAAATATGGTCTGAAT
AGCGGTGGTGAACCGACCCAGGGTGCCGGTGCAGTTGCAATGCTGGTTGCAAGCGAACCGCGTATTCTGGCACTGAA
AGAGGATAATGTTATGCTGACGCAGGATATCTATGATTTTTGGCGTCCGACCGGTCATCCGTATCCGATGGTTGATG
GTCCGCTGAGCAATGAAACCTATATTCAGAGCTTTGCACAGGTGTGGGATGAACATAAAAAACGTACCGGTCTGGAT
TTCGCAGATTATGATGCACTGGCCTTTCATATTCCGTATACCAAAATGGGTAAAAAAGCACTGCTGGCGAAAATTAG
CGATCAGACCGAAGCCGAACAAGAACGTATCCTGGCACGTTATGAAGAAAGCATTATCTATAGCCGTCGTGTGGGTA
ATCTGTATACCGGTAGCCTGTATCTGGGTCTGATTAGCCTGCTGGAAAATGCAACCACCCTGACCGCTGGTAATCAG
ATTGGTCTGTTTAGCTATGGTAGCGGTGCCGTTGCAGAATTCTTTACCGGTGAACTGGTTGCAGGTTATCAGAATCA
TCTGCAGAAAGAAACCCATCTGGCCCTGCTGGATAATCGTACCGAACTGAGCATTGCAGAATATGAAGCAATGTTTG
CAGAAACCCTGGATACCGATATTGATCAGACCCTGGAAGACGAATTAAAATATAGCATTAGCGCCATTAATAACACC
GTGCGTAGCTATCGTAAT**TAA**

>J3-BBa_J23117-*mvaE-mvaS* (5'-UTR)

AGCATTTGCGATCATTCACGCAGCGCTTATTCAGTTGCTCACTGCGATGTCATAATCATCGCTACGAGCTGTGAAAG
ATGCATAAAGCTCGTACGACGCGTTCGCTCGTCTCCTCACTTCTCCTACGGAGCGTTCTGGACACAACGTCGTCTTG
AAGTTGCGATTATAGATTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCGAATTCATTAAAGAGGAGAAAGGTAC
C**ATG**AAAACCGTGGTGATTATTGATGCACTGCGTACCCCGATTGGTAAATACAAAGGTAGCCTGAGCCAGGTTAGCG
CAGTTGATCTGGGCACCCATGTTACCACCCAGCTGCTGAAACGTCATAGCACCATTAGCGAAGAAATTGATCAGGTG
ATTTTTGGCAATGTTCTGCAGGCAGGTAATGGTCAGAATCCGGCACGTCAGATTGCAATTAATAGCGGTCTGAGCCA
TGAAATTCCGGCAATGACCGTTAATGAAGTTTGTGGTAGCGGTATGAAAGCAGTTATTCTGGCAAAACAGCTGATCC
AGCTGGGCGAAGCCGAAGTTCTGATTGCCGGTGGTATTGAAAATATGAGCCAGGCACCGAAACTGCAGCGTTTCAAT
TATGAAACCGAAAGCTATGATGCACCGTTTAGCAGCATGATGTATGATGGTCTGACCGATGCATTTAGCGGTCAGGC
AATGGGTCTGACAGCAGAAAATGTTGCAGAAAAATATCATGTGACCCGTGAAGAACAGGATCAGTTTAGCGTTCATA
GCCAGCTGAAAGCAGCACAGGCACAGGCCGAAGGTATTTTCGCAGATGAAATTGCACCGCTGGAAGTTAGCGGCACC
CTGGTTGAAAAAGATGAAGGTATTCGTCCGAATAGCAGCGTTGAAAAACTGGGTACACTGAAAACGGTGTTTAAAGA
AGATGGCACCGTTACCGCAGGCAATGCAAGTACCATTAATGATGGTGCAAGCGCACTGATTATTGCCAGCCAAGAAT
ATGCCGAAGCACATGGTCTGCCGTATCTGGCAATTATTCGTGATAGCGTTGAAGTTGGTATTGATCCGGCATATATG
GGTATTAGCCCGATTAAAGCAATTCAGAAACTGCTGGCACGTAATCAGCTGACCACCGAAGAAATCGACCTGTACGA
AATTAATGAAGCATTTGCCGCAACCAGCATTGTTGTTCAGCGTGAACTGGCACTGCCGGAAGAAAAAGTTAACATTT
ATGGCGGTGGCATCAGCCTGGGTCATGCAATTGGTGCAACCGGTGCACGTCTGCTGACCAGCCTGAGCTATCAGCTG
AATCAGAAAGAGAAAAAATACGGCGTTGCAAGCCTGTGTATTGGTGGTGGCCTGGGTCTGGCAATGCTGCTGGAACG
CCCTCAACAGAAAAAAAACAGCCGTTTTTATCAGATGAGTCCGGAAGAACGTCTGGCCAGCCTGCTGAATGAAGGTC
AGATTAGCGCAGATACCAAAAAAGAATTTGAAAACACCGCACTGAGCAGCCAGATTGCCAACCACATGATTGAAAAT
CAGATCAGCGAAACCGAAGTGCCGATGGGTGTTGGTCTGCATCTGACCGTGGATGAAACGGATTATCTGGTTCCGAT
GGCAACCGAAGAACCGAGCGTTATTGCAGCCCTGAGCAATGGTGCAAAAATTGCACAGGGCTTTAAAACCGTGAATC
AGCAGCGTCTGATGCGTGGTCAGATTGTTTTTTATGATGTTGCCGATGCAGAAAGCCTGATTGATGAACTGCAGGTT
CGTGAAACAGAAATTTTCCAGCAGGCAGAACTGAGTTATCCGAGCATTGTTAAACGCGGTGGTGGTCTGCGTGATCT
GCAGTATCGTGCATTTGATGAAAGTTTTGTTAGCGTGGATTTTCTGGTGGATGTTAAAGACGCAATGGGTGCCAATA
TTGTTAATGCAATGCTGGAAGGTGTTGCCGAACTGTTTCGTGAATGGTTTGCAGAACAAAAAATCCTGTTTAGCATC
CTGAGTAACTATGCCACCGAAAGCGTTGTTACCATGAAAACAGCAATTCCGGTTAGCCGTCTGAGCAAAGGTAGTAA
TGGTCGTGAAATTGCCGAAAAAATTGTTCTGGCAAGCCGTTATGCCAGCCTGGATCCGTATCGTGCCGTTACCCATA
ATAAAGGTATTATGAATGGCATTGAAGCAGTTGTGCTGGCCACCGGTAATGATACCCGTGCAGTTAGCGCAAGCTGT
CATGCATTTGCAGTTAAAGAAGGTCGTTATCAGGGTCTGACCAGCTGGACCCTGGATGGTGAGCAGCTGATTGGTGA
AATTAGCGTTCCGCTGGCACTGGCAACCGTTGGTGGTGCCACCAAAGTTCTGCCGAAAAGCCAGGCAGCAGCCGATC
TGCTGGCAGTTACCGATGCAAAAGAACTGAGCCGTGTTGTTGCAGCAGTTGGTCTGGCACAGAATCTGGCAGCACTG
CGTGCACTGGTTAGCGAAGGCATTCAGAAAGGTCACATGGCACTGCAGGCACGTTCACTGGCCATGACCGTGGGTGC
GACCGGTAAAGAAGTTGAAGCCGTTGCACAGCAACTGAAACGCCAGAAAACAATGAATCAGGATCGTGCCCTGGCAA
TTCTGAATGATCTGCGTAAACAG**TAA**TGATTAGCGACAAAATATGAGGAGTGCAAAAA**ATG**ACCATTGGCATCGACA
AAATCAGCTTTTTTGTTCCGCCTTACTATATCGACATGACCGCACTGGCCGAAGCACGTAATGTTGATCCGGGTAAA
TTTCATATTGGTATTGGTCAGGATCAGATGGCCGTTAATCCGATTAGCCAGGATATTGTTACCTTTGCAGCAAATGC
AGCAGAAGCAATTCTGACCAAAGAAGATAAAGAAGCCATCGATATGGTTATTGTTGGCACCGAAAGCAGCATTGATG
AAAGCAAAGCAGCCGCAGTTGTTCTGCATCGTCTGATGGGTATTCAGCCGTTTGCACGTAGCTTTGAAATTAAAGAA
GGTTGTTACGGCGCAACCGCAGGTCTGCAGCTGGCAAAAAATCATGTTGCACTGCATCCGGATAAAAAAGTTCTGGT
TGTTGCAGCAGATATCGCCAAATATGGTCTGAATAGCGGTGGTGAACCGACCCAGGGTGCCGGTGCAGTTGCAATGC
TGGTTGCAAGCGAACCGCGTATTCTGGCACTGAAAGAGGATAATGTTATGCTGACGCAGGATATCTATGATTTTTGG
CGTCCGACCGGTCATCCGTATCCGATGGTTGATGGTCCGCTGAGCAATGAAACCTATATTCAGAGCTTTGCACAGGT
GTGGGATGAACATAAAAAACGTACCGGTCTGGATTTCGCAGATTATGATGCACTGGCCTTTCATATTCCGTATACCA
AAATGGGTAAAAAAGCACTGCTGGCGAAAATTAGCGATCAGACCGAAGCCGAACAAGAACGTATCCTGGCACGTTAT
GAAGAAAGCATTATCTATAGCCGTCGTGTGGGTAATCTGTATACCGGTAGCCTGTATCTGGGTCTGATTAGCCTGCT
GGAAAATGCAACCACCCTGACCGCTGGTAATCAGATTGGTCTGTTTAGCTATGGTAGCGGTGCCGTTGCAGAATTCT
TTACCGGTGAACTGGTTGCAGGTTATCAGAATCATCTGCAGAAAGAAACCCATCTGGCCCTGCTGGATAATCGTACC
GAACTGAGCATTGCAGAATATGAAGCAATGTTTGCAGAAACCCTGGATACCGATATTGATCAGACCCTGGAAGACGA
ATTAAAATATAGCATTAGCGCCATTAATAACACCGTGCGTAGCTATCGTAAT**TAA**

>dblTerm
TAAAAACTCGAGTAAGGATCTCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATC
TGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATA

Endogenous promoters

Underlined sequences are 60bp from two ORF adjacent to the promoter. **Bolded** nucleotides are **TSS** according to (D'Arrigo et al., 2016) or **start codon** of the gene or mRFP.

>PP_1776-A-mRFP
GTTCTCAGGGCTCGCCGAGAAACGCATAACCCATGCTTTGAGGTAATTATTCCTGAATAAAGCGGGTTGGCCATTGA
ACGTTCACGCGCGCAGTTGTCTCAAACCTGCCATTTGAGTTTCGCCGCCCGACGGTGCAGTTGCTAAAACGGCGGTT
GAACAGCCGACTGAAGATGCGCTCTCTGGCGCTCCTCGGGGGACAAGCTACATGAAAAAAACTCTGGTATTCTACGA
GCCTCAGCTCGGGTCTGGAGCCTGGATGGCCAGGAATTGCC**T**CGCTGGGCGCGTATATTTATTGCTGCTGCACCCGA
CCGCAGCGGCTGTCAGATATTAAGATACATGCAGGTTTCCTGAGGTTTGAAACTTCAAGTGGCCGTTAAGGACTCAA
AT**ATG**GAATTGATCCCGGTAATTTTATCCGGTGGCGTTGGTAGCCGTCTGTGGCCAGTATCAGAATTCATTAAAGAG
GAGAAAGGTACC**ATG**GCGAGTAGCGAAGACGTTATCAAAGAGTTCATGCGTTTCAAAGTTCGTATGGAAGGTTCCGT
TAACGGTCACGAGTTCGAAATCGAAGGTGAAGGTGAAGGTCGTCCGTACGAAGGTACCCAGACCGCTAAACTGAAAG
TTACCAAAGGTGGTCCGCTGCCGTTCGCTTGGGACATCCTGTCCCCGCAGTTCCAGTACGGTTCCAAAGCTTACGTT
AAACACCCGGCTGACATCCCGGACTACCTGAAACTGTCCTTCCCGGAAGGTTTCAAATGGGAACGTGTTATGAACTT
CGAAGACGGTGGTGTTGTTACCGTTACCCAGGACTCCTCCCTGCAAGACGGTGAGTTCATCTACAAAGTTAAACTGC
GTGGTACCAACTTCCCGTCCGACGGTCCGGTTATGCAGAAAAAAACCATGGGTTGGGAAGCTTCCACCGAACGTATG
TACCCGGAAGACGGTGCTCTGAAAGGTGAAATCAAAATGCGTCTGAAACTGAAAGACGGTGGTCACTACGACGCTGA
AGTTAAAACCACCTACATGGCTAAAAAACCGGTTCAGCTGCCGGGTGCTTACAAAACCGACATCAAACTGGACATCA
CCTCCCACAACGAAGACTACACCATCGTTGAACAGTACGAACGTGCTGAAGGTCGTCACTCCACCGGTGCT**TAA**

>PP_4812-B
GGAAGCCTCACGGGCAGCGCGACCCAAACGGGTCATATAGTCAAGAACGGACTCAGTCATGGGTTCGGTGTCTTGGC
GAAGGGGAAATCGGCTGATTATAACTGCCGCGCAGGTGTACGCCCAGCGGCGGGTGGCGGATGGTAGAAAATGGATG
GGGCAATGTGTAGGAAAGATGTAACCGGGGTTATCGAGATTTCCATCTCCTGTCGCGGCCCTTTAGCGGGCGCGCCC
GCTCCCACTGGGATTTGTGTAGGAGCGGGTTTACCCGCGAAAGGGCCGGCACTGCCAACATCACCGTCCAATTCAGC
CTCGATTAAGCCATTCATTGCTATCATCCCCGCCTCTCCAG**C**CACGAACTGCCCCGC**ATG**CCAGCCCTGCCCGACAG
CTTTTTCGACCGCGACGCCCAGACCCTGGCCAAGGCCCTGGAATTCATTAAAGAGGAGAAAGGTACC**ATG**___

>PP_3839-C
AAATTCGTGGTCAAGCAGGTGATCGTGCTCGGGTTCGTGCAGGTCGTCTTCGTCGCGCATGCTGGCTCCTGGGATAA
TGGCAGGCCTGTATAGGCTGATCAGGGTGCCGGGTCAATGCGTGGTCGCTTAATCCTGGGTTAACCGGACCGGCGCA
ACCTGCAGGCTCTCCCTTCAAACGTCTTTTGCCCGCCTTCCATGGCGGGCTTTTTTATGAC**C**CTGCGCACTTCGTTG
CAGATGATGACAGCCTCATGACCTGACCTTCACGAAATGTCGACATCCGGATGGGCACACTGGCCCTGCTCCGTATG
TTCTTGCAGCCCGCCGCATCCTTGCCGCGGGCTTTTCTTTTTTCCGCAAAGGCCAGCCAGGCATACGCAGGAATTTT
GTGGAAGCGCCCACCTTGACCATGACCTGAAGCAGTTTTGCCTGGGCCGGCAAGGTCTATGCTATCCCGACGATCAC
CCAAGCTCACATCGGATAGACACGGAGGCTCTC**ATG**AAAGCTGCTGTCGTTGCACCAGGCCGTCGCGTGGACGTGAT
AGAGAAAAGCCTGCGCGAATTCATTAAAGAGGAGAAAGGTACC**ATG**___

>PP_1992-D
GTACATGCTCATGTTTCGACGGATCGGGTCATCAAAGCCGCCAAAACCAAAAGACCTTGAAAGGGATGGCCCTTACG
GCGAGTCACTTTTTGTCAAACGCGACAAAAAGTAACCAAAAAACGCTGCGCTCCCATCATCCGGCCCCTGCGCTGCG
CTCCGGGGTCCCCTCACTCCGGCCTTGCTCCCGGCAGGACCGCGCCGAAGGCCCCATCCTGGGGCCTCAGCGCTTGA
CGGGCATCCATGCCCGTCACCTGCCTCCGCAAGGCCTGCGTTCGGCCTCCTGAAGTCGCGAAGATCAAGATCAAGAT

CAAGATCAAGATCAAGATCAAGATCAACAGCAACAGCAACAGCAATAGCTACAGCAATAGGCAACGCTGGGGGAGGC
ACAGACAAAACCCTACAATCCGAGTAAAGTGCCGCCCCCGCCGTTTACAGCAAAAGGATCCCTTCC**ATG**ACACACCC
TTTGGATATCGCCGTCGTCGGCGCCACCGGCAGCGTCGGTGAAGCCTTGGAATTCATTAAAGAGGAGAAAGGTACC**A**
TG___

>PP_0786-E
CAGCACCAGTTTCAGGCCTTCGCGCAGGTATTGGGGGCCAGAGAGGACAGGGGCTTGCATCTAGAGACTCCGCAGAG
AAGGAAAACGCGCTGACCTTACCGAGTTTGCCGCGCCGACGAAAGGCACGCCGTGGGCGGGAAACGGGATGTAACAA
AAGCGCCCGGGGCTTGTTTATCGATGAAATCGCAGCATAGGCGATGCCTATGAGGTGGATTGTCTAAGGATATTTCC
TTAATCTTTGCGACCTCGCTACAGTGCGCTCAACTTTTCGCTTTGCGGGCCTG**C**GCGTTTTAGCCTTCCCCAAGTGC
TGCGTGGGTCCTTTTTAATTTCTTGGCTGGCGCAGCCGGTACACCGATGCCGGCCCTGCGGCCCGCTCGACAGGAGT
TCGAC**ATG**TCTGAAGTACGTCATTCGCGCGTCATCATTCTCGGTTCCGGCCCTGCCGGTTACAGCGAATTCATTAAA
GAGGAGAAAGGTACC**ATG**___

>PP_1972-F
GATAGCCTCGGGCTGGTCGCCAGCCGGCTGGAAACGTGTGACGAGCTGGAACTCGGACATGAAGGACCTCGCGGTGC
AGCAGCTGTAAATTTAGCCAGTAGTCTATACCCAAATGCGCCCGTTCGGGGGCGCTTTCAAGGCACAGGTGTGGGCG
CCTGCAACGGTGGACACGGCATTTAGACCAATGGTCGAAAAATATTTGCGCAAATAGCCCCAAAAGCTGCGCCAGAG
TGTCGCGGTGACCGGTCGGTATCACTATACTGACTCCCCGTTTGTGCACCGCTTCAGTGCATTCGGCTGGAGCG**T**GT
ACGCCCTATCACACTCCAATCAGAGCCAAGGTAACA**ATG**AGCCTGTTTTCCGCTGTCGAGCTGGCACCCCGCGACCC
TATTCTGGGCCTCAACGAAGAATTCATTAAAGAGGAGAAAGGTACC**ATG**___

>PP_3668-G
CAGGGCTTCGAACAGCACCAGCAGGTTCATGTCGACGCGGCGCAGGTCGTTGCGGTTCATGAGGGCTCGGCGTCCTG
GGCAAAGGGTATGGCTGTGTATTGAAGCATGGCGCGGGGCCTTGTGCCTTGTATGGATGTGCCGGCCTCATCGCGGG
CTTGACCGCGATGAGGCCGATACAG**G**CCTGCGCCTGACAGAGCCAGCCTATCAGGCTCCAACAGCCACTCTATTAGA
CCTCTGCCCAAGCTCGGCTAGTCTTTCTCGTGGCCCCGCGAATTCCGCGACAGGGCAGCGCGTCAGCACCTGCGTGC
AAGACCGTGCCCCCTCGCCGTGACAGCTTCGCAAGCCCAGTGTACACCTGATGAGGGGTAGTACGAGCCCACCCGCT
CGGCTGAAAGAACACTGGCATAGACCGGAAATCTGGATAACCGACCCAAAGGTACCCGCAG**ATG**TCGAACGAATCGA
AATGCCCGTTCCATCAAACCGCAGGTGGCGGCACCACCAACCGTGAATTCATTAAAGAGGAGAAAGGTACC**ATG**___

>PP_5046-H
CTTGCGCACCGGCCGGCTTTTGATGGTGATTTCTGGGAAGACTTTGACGATAAGTTTCATTGGTTAACAGCGCGCGC
AGGGCCTGCCGAAAATGAGGGGCGCGAATTATATCGGAAATTGCTCAGGATTTGACCAACTTTTGATCAGAAGCTTT
GAGATAAATGCAGAGGCCAGGTTTTGCAGCGCCTGTTCCGGCCCTTTCGCGGGTAAACCCGCGCCTACAGGCGGTGC
AAAGCCCGTAGGAGCGGGTTTACCCGCGAAGAGGCCCTAGAACCTGGCACACCACTCACGCCACGCACCCTATTGGT
GCGACACGATCAAAAAACGCATCGTAAGGGTGCACTTTCACCCGCTAACCCAACGCCAAATGCACGCAAACGCCCCC
TTTTATCCCACCCTCGCCATTTTCGGGCACTGGCATGCAATTTGCTCCCTTGTGAGGCAGGTAAGCTTGGCCGACTA
TCCGCGCCCGGCAACACCCTTTTTCCAGGGCAGC**G**GCCCACCGCGCTCTAGACCATCCGGAGGACAAC**ATG**TCGAAG
TCGGTTCAACTCATCAAAGATCATGACGTCAAGTGGATTGATCTGCGTTTCGAATTCATTAAAGAGGAGAAAGGTAC
C**ATG**___

>PP_1231-I
GTTGAAGAATTTGCCATGTTCATGGCGGCGCATCTGGGGGCCCTGTCGCCGCAGGGGTGATGGTTTCTCCTGTTGCG
GCCTCTTCGCGGGTGAACCCGCTCCTACGAGGATTTCACAGTGTAGGGGCGGGTTCACCCGCGAAGGGGCCCGCACA
AGCACTACAAAAACCCCTTTAGTAATCGCTGGATTGTCTGTAGCCTTCGGCCTCCGATAATAATCCGCGCCCGCAGA
GGGCCGTGGCGGTCAACAT**G**CCGTCCGGCACCCTTAGCCGATCTGCCAGGGCCGGGTGATACACTCGATTTGACGCG

CAAGCGCGCCTGCAGGTCCAGCGATC**ATG**ACCCAGATTTCCGAACGCCTTTTGGTTCAGGCCCACCTCGACGCCAAG
CAGCCCAACGAATTCATTAAAGAGGAGAAAGGTACC**ATG**___

>PP_4701-J
CTGGGTAACACGCGGTTGATCGACAACCTCTACCTGCATTTGGAAGAGAAGACCGCATAACGGTCTTGGGCTGCCTT
GCAGCCCATTCGCGGGCAAGCCCGCTCCTGCACGTTTACCTGTAGGGGCGGGCTTGCTTGCAATGCCCCCAAAATCC
CCCTGCCATACCCATTCCCAGCACGTGGCCTTTGCCTATAATGGTGCCAGCCTGAACCCGGCAACGACTGCCGTGTC
CAAGCCCTCA**C**CACGCACCAAGGGAACCCCGCGCA**ATG**GCGTATTACCGTACACCCCACGATGTGACGGCCCTGCCC
GCCTGGCAGGCGCTTCAGGAATTCATTAAAGAGGAGAAAGGTACC**ATG**___

Other sequences

>pBBR1-MCS
GTCGAATTTGCTTTCGAATTTCTGCCATTCATCCGCTTATTATCACTTATTCAGGCGTAGCAACCAGGCGTTTAAGG
GCACCAATAACTGCCTTAAAAAAATTACGCCCCGCCCTGCCACTCATCGCAGTACGGCCTATTGGTTAAAAAATGAG
CTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTCCATTCGCCATTCAGGCTGCGC
AACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCG
ATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACT
CACTATAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTC
GATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAA
TTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATA
CGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACT
GCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGC
GTATTGGGCGCATGCATAAAAACTGTTGTAATTCATTAAGCATTCTGCCGACATGGAAGCCATCACAAACGGCATGA
TGAACCTGAATCGCCAGCGGCATCAGCACCTTGTCGCCTTGCGTATAATATTTGC

>pp1_HR1-J3-BBa_J23117-mRFP1-dblT-pp1_HR2
ATGACCGACCTGATCGAAGTGAAGACGGCAGACCTGGTGGGCGAGGCGCTTGGGTGGGCCGTGGGCACGGCGGAAGG
CCTGGACCTGTTCATGGCGCCGCCGGAGTACGGCAACCCACACCGAGTGTTCGCCCGCTACCAGGGCCAGGCCATCG
AGCACACCAAGCGCTTCAACCCGTGGGAAGACTGGGCGGTTGGCGGGCCGATCATGCAGAAGCACAACGTCAGCCTG
CACTGCCCGCAGCCAGAGTGGGACTACTGGGCAGCCTGGATAACCGATAACGGCAAGGACGTCGCCCAGGGCGCTGA
TCTGCCGTTGCCGGCGGCGTGCCGGGCCATAGTCGCCCACCAGCTCGGCGATACCGTCCAGGTGCCGAAGGAGCTGA
TGCCATGACCGTGATCCTTCCCCTCGCCTACATGGCCTACCTGATCTACAGGGGGCTTCTCGGTGAGGGAGGCGCCT
GCAAGCAAAGGGCACGACATGACCTGACGACAGCACGGCAAAAAACAAACTCGAAAGGATCATCCACAAGATCAAGC
GCTGCCTGGCGCTATTCAAAAGCTCGAATGAATATGAGAGAGTCTAGGCCCACCCGCCGATTACGAAGGTCTTCGCT
CGGAGCACACCCCAGACCAAGGCTCGACTCATAGTTTCGCTTGGTCTGGTGCTGTAAGCCTCTTCTACAATTCGGTC
CCCGCTTTTGGAGTACACCCCGATGAAGAGCTGCGTTTCGCCTGTCCGCGAAAGACGGGTTTGCACGTCGATACTCC
TGCCGTCCTCAAGGATTTCGTCGTGATGACGAAGGTGAAGCGCTGGGTCTGCCCAGGTCCAGAATTTTTCGCCGCGA
CATCTCATATCAATCTCCTCTTACTTATCCCAGTAGGCGCGGTAAAGAGAGGGATAGATATCCATTTCGCTTAAATG
CGACCGGTGGAAAATGATCGGCCCTAATCCTTGCTGATAGATATCAGCGGGACAGCGCCAGTAGAGAACCGAGCCCA
GCATGGCAATTCCGACGTCAGCATTTGCGATCATTCACGCAGCGCTTATTCAGTTGCTCACTGCGATGTCATAATCA
TCGCTACGAGCTGTGAAAGATGCATAAAGCTCGTACGACGCGTTCGCTCGTCTCCTCACTTCTCCTACGGAGCGTTC
TGGACACAACGTCGTCTTGAAGTTGCGATTATAGATTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGC**G**AATTCA
TTAAAGAGGAGAAAGGTACC**ATG**GCGAGTAGCGAAGACGTTATCAAAGAGTTCATGCGTTTCAAAGTTCGTATGGAA
GGTTCCGTTAACGGTCACGAGTTCGAAATCGAAGGTGAAGGTGAAGGTCGTCCGTACGAAGGTACCCAGACCGCTAA
ACTGAAAGTTACCAAAGGTGGTCCGCTGCCGTTCGCTTGGGACATCCTGTCCCCGCAGTTCCAGTACGGTTCCAAAG
CTTACGTTAAACACCCGGCTGACATCCCGGACTACCTGAAACTGTCCTTCCCGGAAGGTTTCAAATGGGAACGTGTT
ATGAACTTCGAAGACGGTGGTGTTGTTACCGTTACCCAGGACTCCTCCCTGCAAGACGGTGAGTTCATCTACAAAGT

TAAACTGCGTGGTACCAACTTCCCGTCCGACGGTCCGGTTATGCAGAAAAAAACCATGGGTTGGGAAGCTTCCACCG
AACGTATGTACCCGGAAGACGGTGCTCTGAAAGGTGAAATCAAAATGCGTCTGAAACTGAAAGACGGTGGTCACTAC
GACGCTGAAGTTAAAACCACCTACATGGCTAAAAAACCGGTTCAGCTGCCGGGTGCTTACAAAACCGACATCAAACT
GGACATCACCTCCCACAACGAAGACTACACCATCGTTGAACAGTACGAACGTGCTGAAGGTCGTCACTCCACCGGTG
CT**TAA**AAACTCGAGTAAGGATCTCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTA
TCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATATC
CGCCATTACGATCTGACTTGCCACTTTGGCCGTTATCCCTATCTCTGGTCATACGCACCTCTCAATGCTCAGAAGTC
CTGAGCTAAATTGAGTGTGCGCCTACAGAGACTAGAGAAGATTTATATCAGACCGACGGGTACAAACGAAAAAATCA
ATAACCGGCTCCTGTCAGCACCTCGCAGAGAAAACAGATCGCAGATATTCAGCATGCGTAGCTGATGCTCGACCGAG
ATCGCACCTGACATGCACCGCTGCGCGGCACATCACACTGAGTAGAGAGCGTTCCGCGACGCCGATGGCGTGAAGCG
CGAGCTGGAAAACGTCGTGCCGTTCAAGGGCGCCTAACCCCTCCCCATACAACTCAAGCCCGCCGACATGCGCGGGC
GAGGATTACCTATGTCCAATTTCCTGACTCGCTGGCTCAAGCGCAAGAAGAAGCCCGGGCCGCGTCCAACACTGGCA
CCAACTGGATTTGCTCGCGGCCACAGCCCGGCAGTCGGCAGGCTAGATCCGATGCTTGATCCGCTCAACCCGTTGAG
CCCCGTCAGCCCGTTGCATCCCGCCTACCAGGCCGACAGCTACGAACCACCGCGCAGCACCAGCAGTTCCTGCTCCA
GCCGTGATTACAGCAGCTACGACTGCGGCAGCAGCTACTCGTCGAGCGACAGCAGCAGTTCCAGCGATAGCGGATCC
AGCTCCAGCAGCTGCGACTGACCACCAACCTGCCGCCACCGGCGGCGTGGAGACCATCCCATGGAAACCGAAATCCT
TTCGGACGAAGAGCTGGTGGCGATCACCGGCTACAAACCCCGGGCGTGGCAGCGCCGTTGGCTAACAGAAAAAGGCT
GGCACTTCGTCGAGAGCCGCGGCGGCCGGCCACTGGTTGGCCGCCAGTACGCCCGCCAGAAGCTCAGCGGCGTGGTG
ATCGACACCTTGCCGGTCGCACCAGCCCCACCACCAACGCCCGCCTGGACCCCTGATTTTTCCCGAGTGAAGTGA

>pp2_HR1-J3(106)-BBa_J23111-sfGFP1-dblT-pp2_HR2

ACCAGGATGAATACCTTAAGGACGCCACCGGTAACCGGCGTTATTGGCCGGTCGCTTGCGTCAAGGTGGACCTTGAA
GCATTGCGTCGCGCTCGTGACCAGCTGTGGGCTGAGGCCATGTTCTGCTACCAGGCCGGTGATATCTGGTGGGTGAC
CCGTGAGGAGGAAGAACTGTTCACTGCAGAGCAGGAAGAGCGCTTCGTGGTAGATGAATGGGAGGGGCCGATCCTGA
AATGGTTGGAGGAATCCCAGGCCGGCGAGACGGTCACCGGAAGCGAAGTGTTGGGGCAGGCATTGAACCTTGACCCT
GGCCACTGGGGCAAGCCTGAGCAGATGCGGGTGGGATCGATCATGCACCGCCTAGGTTGGCGGCGTCGCAGGCTGGC
TGCGCTGCCGAAGAGCGGTAAGCGCCCTTGGGCATATCAGAAGCCTGATGGTTGGGGGCGCAGCGCCTTGGAGCAGT
CCACGCAGCCGAAGGAGGAGTGCTTTTGATCAAACACATAGATGAGATGCTGAAACTGTGGGCTCAAGAGCTCCATG
CGCCGGAGCCCTGTCATTCGGCGGGCGGTGTTGGTAGCATGCTCGGCCTGTTGATCGAGTGCAAGGGTGACCTTGTG
CGTGGCACCCGAGGCAGCAAGGTGCTACTGGACGAGTCTGCGGACATCGAGATCATTGTGAATAAGCATCTGGCACC
GGAGCTCTACCTGGTGGTTCGTGAGCACTACTGTAACGCCGACAGCGAGCTGTACCAAAAGTACCGGCACTGTGGGT
GTAGCCGGGATACCTATTACAAGCGCTTACATGAGGCGCACGTCTGTATCGCAGGCTTGCTCTTGGGGCGAGCGGCA
TGATTCGCCATCCACGGTCCTACCGTCCCGCTGCCGTCCTGCCGCGTTTGATGCAGGTCAGCCTAGCTCAAGCCCGC
GCAGGTCGCGGGCTGTCCCACCGTCCCACCGTACACACGAAGGCGCGCACATAGGCGTGTGCAGCGCATCACGCGCA
TGCATAGCATTTGCGATCATTCACGCAGCGCTTATTCAGTTGCTCACTGCGATGTCATAATCATCGCTACGAGCTGT
GAAAGATGCATAAAGCTCGTACGACGCGTTCGCTCGTCTCCTCACTTCTCCTGCGGTTACCAAAGGCGTCCTCGTCG
TCTTGAAGTTGCGATTATAGATTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGC**G**AATTCATTAAAGAGGAGAAA
GGTACC**ATG**AGCAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGG
GCACAAATTTTCTGTCCGTGGAGAGGGTGAAGGTGATGCTACAAACGGAAAACTCACCCTTAAATTTATTTGCACTA
CTGGAAAACTACCTGTTCCGTGGCCAACACTTGTCACTACTCTGACCTATGGTGTTCAATGCTTTTCCCGTTATCCG
GATCACATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAACGCACTATATCTTTCAA
AGATGACGGGACCTACAAGACGCGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATCGTATCGAGTTAAAGG
GTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAACTCGAGTACAACTTTAACTCACACAATGTATACATC
ACGGCAGACAAACAAAAGAATGGAATCAAAGCTAACTTCAAAATTCGCCACAACGTTGAAGATGGTTCCGTTCAACT
AGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCGACAC
AATCTGTCCTTTCGAAAGATCCCAACGAAAAGCGTGACCACATGGTCCTTCTTGAGTTTGTAACTGCTGCTGGGATT
ACACATGGCATGGATGAGCTCTACAAA**TAA**GGATCCAAACTCGAGTAAGGATCTCCAGGCATCAAATAAAACGAAAG
GCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTC
ACCTTCGGGTGGGCCTTTCTGCGTTTATAAGTACTAAGGTGTATTCCCCCGGCATTCAACAGACATTCCCCGGACGT

TTACCACTTATTGCCGGGTGGCATTAAAACTCGCTTGCTGCCACCGGAATCGACCTGTAAAAAGTACCCATCTTCGA
GACGTGCGGGCGCACAAAGCGGCCCGCCAAACACTGAAAACCCCGGCCCTGGCGCCGGGGTTTTTGCCTTTGGGGGA
TTCGATGAACAGCGAGCAGCAAGCGTTAGTTGAGATGCCAATCTGGTTGGTGATCTTCCTGTCCCTGGTCGGCGGGG
TGTCAGGCGAGATGTGGCGGGCCGACATGGCCGGCGTTAGCGGCTGGTTCATTTTCCGCCAGGTGCTGCTGCGCTCC
GGTGCCTGCGTCGTATGCGGACTGTCGACCATCATGCTGCTGTACTCGGCGGGCATGTCGATGTGGTCGGCCAGTGC
CATTGGTTGTCTCACTGCCACTGCCGGTGCGGATGTGGCCATAGGGTTGTACAAGCGTTGGGTCGCCAAGCGGCTGG
GCGTCTGCGATGTCACGTCCCGTAGCGGCGAACCTGGACAGTGACCCGATCGCCAGCCTCGGTGGGGTCGGGGACCC
TGGCGATATGGCCGGGTTACGGGGCAGGAAACCCGCGGCTCTTCGCTAGCGGACAGTTCGCCAGCTTACTGAAATTC
AACCTGTTGAAATTGAAAGGTTGTTGGTTGAAATACCATTGAAATGGAGGGCTCATGACGGATTCGAACTTCTTGTC
AAAGAGCGCCTTCGCTGCTCGCATAGGGAGATCACCCAGCTACATCACCTGGTTGAAAGACAACGGCCGCCTGGTGC
TTTCACCCGATGGAAAATTGGTGGATGTGCTGGCCACCGAGGCCAAGATTCAGGAGACAGCTGATCCGGCCAAAGCA
GCCGTCGCGGCTCGGCATGAAGAAAACCGCATCGAGCGGGACGTCCGGGCCCACATCCAGCCTAGCGCCGACACACC
TGCGGTGCAGCCAGCGGATCACGCGCCGAGCGGA

Anderson Promoters

RNAP recognition sites (-35 and -10 elements) of the Anderson promoter series are underlined. TSSs are indicated in **bold** at the predicted site. TSSs of the Anderson promoter series were characterized in (Kosuri et al., 2013); most TSSs are located immediately after the 35 bp sequence as expected. Some weak promoters have undefined TSSs due to low RNAseq signal and some of the TSSs reported in (Kosuri et al., 2013) are shifted downstream by 1 base when **C** is the first nucleotide of the 5'-UTR. For all experiments reported in this manuscript, the first base of the 5'-UTR is **G** (see 5'-UTR sequence below) The promoters shown below appear in an order of ascending basal expression level in *P. putida* (weakest first).

>BBa_J23109-5'-UTR
TTTACAGCTAGCTCAGTCCTAGGGACTGTGCTAGC**G**AATTCATTAAAGAGGAGAAAGGTACC

>BBa_J23113
CTGATGGCTAGCTCAGTCCTAGGGATTATGCTAGC**G**

>BBa_J23117
TTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGC**G**

>BBa_J23114
TTTATGGCTAGCTCAGTCCTAGGTACAATGCTAGC**G**

>BBa_J23115
TTTATAGCTAGCTCAGCCCTTGGTACAATGCTAGC**G**

>BBa_J23107
TTTACGGCTAGCTCAGCCCTAGGTATTATGCTAGC**G**

>BBa_J23105
TTTACGGCTAGCTCAGTCCTAGGTACTATGCTAGC**G**

>BBa_J23106
TTTACGGCTAGCTCAGTCCTAGGTATAGTGCTAGC**G**

>BBa_J23108
CTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGC**G**

>BBa_J23110
TTTACGGCTAGCTCAGTCCTAGGTACAATGCTAGC**G**

>BBa_J23111
TTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGC**G**

>BBa_J23119
TTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGC**G**

>5'-UTR (Ribosome-Binding-Site)
GAATTCATTAAAGAGGAGAAAGGTACC

5'-Proximal sequences

>PS1
CAATATGACGTGTTGTTAATTTGGTT

>PS2 (same as J1)
TTGGGTTCCACCGGATACCTCCGGAC

>PS3
GTCGTAAATAAGTAAGTCACTCCCAC

>PS4
GTTGTCCTTCTAGTCGCCCATGACTC

>PS5
ACACCGACTACCCCTGCTGGGCCCAG

sgRNA/scRNA sequences

>BBa_J23119(SpeI)-sgRNA-rrnBTerm
TTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTNNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAA
GTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGAAGCTTGGGCCCGAA
CAAAAACTCATCTCAGAAGAGGATCTGAATAGCGCCGTCGACCATCATCATCATCATCATTGAGTTTAAACGGTCTC
CAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGCGGTCTGAT
AAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTA
GCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTC
GAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACT

>BBa_J23119(SpeI)-scRNA_1xMS2.b2-rrnBTerm
TTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTNNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAA
GTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACATGAGGATCACCCATGTGCTTTTTTTGAAGCTTG
GGCCCGAACAAAAACTCATCTCAGAAGAGGATCTGAATAGCGCCGTCGACCATCATCATCATCATCATTGAGTTTAA
ACGGTCTCCAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGC

GGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAA
ACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAG
GCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACT

Supplemental References

Choi, K.-H., Schweizer, H.P., **2006**. mini-Tn7 insertion in bacteria with single attTn7 sites: example *Pseudomonas aeruginosa*. *Nat. Protoc.* **1**, 153–161. https://doi.org/10.1038/nprot.2006.24

D'Arrigo, I., Bojanovič, K., Yang, X., Holm Rau, M., Long, K.S., **2016**. Genome-wide mapping of transcription start sites yields novel insights into the primary transcriptome of *Pseudomonas putida*. *Environ. Microbiol.* **18**, 3466–3481. https://doi.org/10.1111/1462-2920.13326

Dong, C., Fontana, J., Patel, A., Carothers, J.M., Zalatan, J.G., **2018**. Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nat. Commun.* **9**, 2489. https://doi.org/10.1038/s41467-018-04901-6

Ehrenworth, A.M., Sarria, S., Peralta-Yahya, P., **2015**. Pterin-Dependent Mono-oxidation for the Microbial Synthesis of a Modified Monoterpene Indole Alkaloid. *ACS Synth. Biol.* **4**, 1295–1307. https://doi.org/10.1021/acssynbio.5b00025

Fontana, J., Dong, C., Kiattisewee, C., Chavali, V.P., Tickman, B.I., Carothers, J.M., Zalatan, J.G., **2020**. Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. *Nat. Commun.* **11**, 1618. https://doi.org/10.1038/s41467-020-15454-y

Kosuri, S., Goodman, D.B., Cambray, G., Mutalik, V.K., Gao, Y., Arkin, A.P., Endy, D., Church, G.M., **2013**. Composability of regulatory sequences controlling transcription and translation in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **110**, 14024. https://doi.org/10.1073/pnas.1301301110

Wirth, N.T., Kozaeva, E., Nikel, P.I., **2019**. Accelerated genome engineering of *Pseudomonas putida* by I-SceI-mediated recombination and CRISPR-Cas9 counterselection. *Microb Biotechnol*. **13** https://doi.org/10.1111/1751-7915.13396

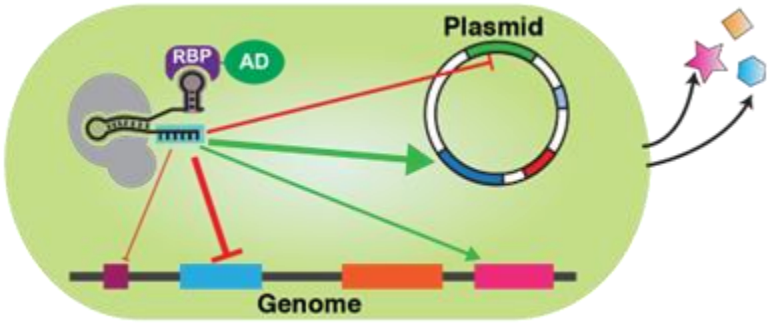
Zaslaver, A., Bren, A., Ronen, M., Itzkovitz, S., Kikoin, I., Shavit, S., Liebermeister, W., Surette, M.G., Alon, U., **2006**. A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli. Nat. Methods* **3**, 623–628. https://doi.org/10.1038/nmeth895

Zhao, F., Zhang, Y., Li, H., Shi, R.J., Han, S.Q., 2013. [CaCl2-heat shock preparation of competent cells of three *Pseudomonas* strains and related transformation conditions]. *Ying Yong Sheng Tai Xue Bao* **24**, 788–94.

Chapter 3:

Bioproduction of aromatic amines enabled by stress-resistant bacteria

Cholpisit Kiattisewee[+], Ian D. Faulkner[+], James M. Carothers*, Jesse G. Zalatan*



Abstract

Pseudomonas putida is a potential chassis for bioproduction due to its high tolerance against aromatic compounds, such as p-aminocinnamic acid (p-ACA) — a precursor for high-performance polymers synthesis. In this work, we demonstrated that p-ACA can be produced in the non-model bacteria *P. putida*. Production of p-ACA was previously challenging in *E. coli* due to its inherent toxicity. With optimization of enzymes of choice, production of p-ACA directly from glucose can be improved from a trace-level to more than 100 μM. We further fine-tuned enzyme expression number using CRISPRa/i transcriptional regulation and plasmids with varying copy-number, and therefore improved the p-ACA titer to 500 μM. Further investigation at different cell density suggested that conversion of p-ACA peaked later at the stationary stage when the other competing aromatic analogs were consumed. This finding demonstrates the potential of the transcriptional CRISPR tools for tuning of gene expression level, which could provide a new methodology to optimize metabolic engineering in non-model bacteria.

This work is in preparation for publication. Part of this work was presented at the AIChE Annual Meeting 2022 (ISBN: 978-0-8169-1118-9) on November 17th, 2022

3.1 Introduction

Bioproduction of fuels and chemicals from renewable biomass offers a promising alternative to a petroleum-dependent society, paving the way toward a more sustainable circular economy (Intasian et al., 2021). Recent advancements in metabolic engineering and synthetic biology have introduced a novel platform for biosynthesis, ranging from small molecules to biopolymers (Nielsen and Keasling, 2016; Volk et al., 2023). Given the global demand for polymeric materials, which now has exceeded 140 kg per capita (Filiciotto and Rothenberg, 2021; Han et al., 2022), the development of bio-based polymers from renewable biomass is crucial for establishing a carbon-conserving society (Bernhardsgrütter et al., 2021; Jatain et al., 2021; Westenberg and Peralta-Yahya, 2023).

The simplest route to siphon renewable biomass into the polymer industry is to biochemically synthesize the monomeric precursors of polymeric materials. Currently, bio-derived polymeric products, such as polyhydroxybutyrate and polylactic acid, served as biodegradable polyesters, mainly used for disposable supplies with minimal physical or thermal requirements (Singhvi et al., 2019; Urtuvia et al., 2014). However, the search for more efficient and robust bio-derived polymers with improved properties remains open.

Among the promising candidates, amine-containing polymers, particularly aromatic amines, stand out due to their potential applications in biomedical and conductive materials. (Cywar et al., 2022; Gonzalez de Gortari et al., 2022). Their properties originate from amine moieties, which provide pH responsiveness and hydrogen bonding capabilities at the molecular levels. Conventionally, aromatic amines are synthesized from chemical catalytic hydrogenation of petroleum-derived nitroarene precursors at elevated temperatures (Tadrent et al., 2018). On the other hand, synthesis of aromatic amines from biomass requires high temperature (Li et al., 2022). These approaches are neither sustainable nor eco-friendly, making enzymatic or fermentative routes a greener alternative method for aromatic amine production.

One noteworthy amine-functionalized polymer precursor is p-aminophenylalanine (p-AF), derived from aromatic amino acid biosynthetic pathways broadly available in microorganisms (Sariaslani, 2007). Expression of *papABC* genes from *Streptomyces venezuelae* or *Pseudomonas fluorescens* in *E. coli* offers a synthetic route for converting chorismate into p-AF, initially fermented from glucose (Masuo et al., 2016; Mehl et al., 2003). The subsequent deamination of p-AF using phenylalanine ammonia lyase (pal) from *Streptomyces lividans* or *Arabidopsis thaliana* could generate p-aminocinnamic acid (p-ACA) (Noda et al., 2011; Suvannasara et al., 2014). p-ACA contains a vinylic group adjacent to the phenyl ring resembling the core structure of styrene, a precursor of widely-used polymeric material polystyrene. This compound can be further transformed into p-aminostyrene (p-AS) via chemical catalysis (Sheng et al., 2020; van Schijndel et al., 2020) or enzymatic reactions (Grubbe et al., 2020; Mukai et al., 2010; Williamson et al., 2020). Moreover, p-ACA can undergo photodimerization and condensation with carboxylic acids to generate polyimides and polyamides, with thermal properties comparable to high-performance petroleum-based counterparts (Suvannasara et al., 2014; Tateyama et al., 2016).

However, existing microbial aromatic amine biosynthesis methods have encountered low efficiency and necessity for multi-step fermentation (Minakawa et al., 2019; Tateyama et al.,

2016). Model organisms, like *E. coli* and *S. cerevisiae*, pose several challenges ranging from foreign genes expressions, ability to use cost-effective feedstocks, and tolerance to rising metabolites concentrations from the fermentation process (Fatma et al., 2020; Sun et al., 2015; Yan and Fong, 2017). Particularly, bioproduction of aromatic compounds is challenging due to toxicity and solubility limits of desired products and intermediates (Qi et al., 2007; Suvannasara et al., 2014). Hence, non-conventional microbes that can withstand toxic environments and use diverse feedstocks are appealing for bioproduction of difficult-to-synthesize aromatic molecules.

One promising candidate, *Pseudomonas putida*, known for its tolerance to high concentrations of aromatic compounds and unique metabolisms, is attractive for aromatic amines bioproductions (Molina-Santiago et al., 2016; Nikel and de Lorenzo, 2018; Volke et al., 2020). Recent research in *P. putida* demonstrated its capability to utilize a wide range of feeds from lignocellulosic biomass to degraded polymeric materials (Kawaguchi et al., 2016; Tsuge et al., 2016; Werner et al., 2021). Advancements of CRISPR-based tools also significantly facilitate unconventional microbial research as they allow programmable control of gene expression in various bacteria, including *P. putida* (Elmore et al., 2023; Kiattisewee et al., 2021; Lu et al., 2022; Wang et al., 2019). Thus, bioproduction of aromatic amines in *P. putida* holds promise beyond the conventional method developed in model organisms.

In this work, we applied CRISPR activation (CRISPRa) gene regulatory tools previously developed in *P. putida* to enable direct production of p-ACA from simple feedstocks. We found that transferring p-AF biosynthetic clusters directly to *P. putida* yielded up to a millimolar level of p-AF (~180 mg/L). Subsequent provision of phenylalanine ammonia lyase (Pal) enzyme resulted in a trace amount of p-ACA, the first report of direct production to our knowledge. By altering the PAL enzyme, the production of p-ACA is significantly improved from trace level to 100 µM (~16 mg/L). Further fine-tuning of enzyme expression using CRISPRa/i tools further improved p-ACA titer to 500 µM (~82 mg/L). Further investigation at different cultivation times suggested that conversion of p-ACA peaked later at the stationary stage when the other competing aromatic analogs, phenylalanine and tyrosine, were depleted. This study showcases the potential of the CRISPR tools for tuning gene expression levels in bacteria which could enable a synthesis route to challenging production of biomolecules.

3.2 Materials and Methods

Plasmid and strain constructions

All plasmids were constructed using either Infusion cloning (Takara Bio) or Golden Gate Assembly (NEB & Promega). All cloning-related transformations were performed with *E. coli* NEB-turbo (NEB) or DH10b cells (Invitrogen). *E. coli* MG1655 was used for a bioproduction test. All engineered *P. putida* are derivatives of KT2440 (Kiattisewee et al., 2021). DNA oligonucleotides were purchased from IDT. Transform of *E. coli* was performed with heat-shock chemically competent cells method. *P. putida* transformation was performed as previously described with either $CaCl_2$-based chemically competent cells or electroporation (Kiattisewee et al., 2021). Sanger sequencing to verify constructs was performed by Genewiz/Azenta. DNA encoding the *papABC* operon responsible for p-AF production was taken from *Pseudomonas fluorescens* SBW25 (Masuo et al., 2016). The *E. coli* derived *aroGL* operon responsible for elevated chorismate levels, encoding AroL and a feedback-resistant mutant of AroG, were obtained from

a previous report (Juminaga Darmawi et al., 2012). The *A. thaliana* PAL2 was obtained from a previous report (Cochrane et al., 2004), while the *R. glutinis* Pal was synthesized as gene fragments by Twist Bioscience. TyrB used in this work was amplified from *E. coli* MG1655 without modification. Promoter sequences placing these outputs under CRISPRa control were obtained from Anderson Promoter (http://parts.igem.org). Integrations were achieved using an R6K delivery plasmid and an I-SceI-mediated protocol adapted from a previous report (Kiattisewee et al., 2021; Wirth et al., 2019).

Aromatic amines bioproduction in bacteria

The culture conditions used in this work for aromatic amine production were adapted from the previous literature (Burke et al., 2017). *E. coli* MG1655 chemically competent cells were transformed with metabolic pathway and CRISPR plasmids, while a control strain was transformed with similar plasmids without pathway genes and used for standard curve diluent. Single colonies were picked in triplicate and used to inoculate 2 mL of MOPS EZ-Rich defined media (Teknova) and appropriate antibiotics, in 14 mL polypropylene culture tubes. Cultures were grown at 30˚C and shaken at 200 rpm for 24 hours.

In the case of *P. putida*, constitutively expressed dCas9 and MCP-SoxS were previously integrated into *P. putida* KT2440 to generate CKPP002 (PPC001 in Kiattisewee et al.). This strain, or its derivatives with integrated metabolic pathway, was transformed with a pBBR1 plasmid containing either scRNAs only or pathway genes and scRNAs. Two-plasmid production strains including the additional pRK2 plasmid were doubly transformed in series, using competent cells containing the *papABC*/*aroGL* plasmid. A control strain for standard curve diluent was transformed with similar plasmids not containing pathway genes. Single colonies were picked in triplicate and used to inoculate 2 mL of MOPS EZ-Rich defined media (Teknova), supplemented with appropriate antibiotics, in 14 mL polypropylene culture tubes. Cultures were grown at 30˚C and shaken at 200 rpm for 24 hours, unless specified.

p-ACA tolerance assay using time-course growth experiments

Triplicates of overnight LB (Teknova) cultures of *E. coli* MG1655 and *P. putida* KT2440 were inoculated 1:100 into fresh MOPS EZ-Rich defined media (Teknova) and allowed to reach an OD_{600} of 0.05 – 0.1, as measured in a cuvette with 1 cm pathlength on a Nanodrop spectrophotometer (Nanodrop 2000c) spectrophotometer. 196 uL of culture was then placed into wells of a black 96-well plate with a clear bottom (Corning) and 4 uL of DMSO containing varying amounts of p-ACA were added to each culture. Noted concentrations are final p-ACA concentrations in the cultures. Final DMSO concentration in media (vol/vol) was kept at 2% to equalize DMSO toxicity. The plate was then placed into a Biotek Synergy HT plate reader with orbital shaking set to medium and temperature set to 30 ˚C for 16 hours. OD_{600} measurements were taken every 5 minutes. Error bars are standard deviation across three platewell replicates from the same overnight culture. Percent reduction was calculated for each timepoint as (OD_0mM - OD_20mM) / OD_0mM within each replicate and then averaged.

Investigation of p-ACA catabolic activity

Starting with KT2440 or derivative overnight cultures in LB, 1:100 dilution was performed to 2 mL of M9 minimal media supplemented with 2 g/L carbon sources — glucose, p-CA (p-coumaric acid), and p-ACA. The cultures were incubated at 30 ˚C with continuous shaking at 200 rpm. The cultures were checked after 2 days post inoculation. If the OD_{600} remains unchanged after 48 hours, it is assumed that the corresponding feed cannot be used as a sole carbon source.

To test the persistence of p-ACA in *P. putida* cultures. 50 μM of p-ACA was added to *P. putida* cultures together with the 1:100 subculture in 2 mL of EZ Rich Medium. The bacterial cultures were incubated at 30 ˚C, 200 rpm for 24 hours, and the supernatants were subjected to HPLC analysis with the method described below.

HPLC for production analysis

The HPLC method was adapted from the previous literature (Burke et al., 2017). Culture supernatants were filtered by centrifuging at 14000 RCF for 20 minutes using an Amicon Ultracel 10K centrifuge filter (Millipore). Filtered supernatants were supplemented with 0.2% trifluoroacetic acid (TFA) and assessed using an Agilent HPLC with a diode array detector set at 210 nm and 280 nm. p-AF, p-ACA, and other components were separated using a ZORBAX Eclipse Plus phenyl-hexyl column (Agilent) with water plus 0.2% TFA as solvent A and methanol plus 0.2% TFA as solvent B. The mobile phase gradient was as follows: 100% solvent A at 1 mL/min from 0 to 4 min, ratio increased to 95% solvent B and 5% solvent A at 1 mL/min from 4 to 18 min, 100% solvent A at 1 mL/min from 18 to 22 min. Sample concentrations were determined by interpolation from standard curves ranging from about 10 μM to 1 mM. p-AF, p-ACA, p-CA (p-coumaric acid), t-CA (trans-cinnamic acid), and p-ABA (p-aminobenzoic acid) for standard curves were obtained from Santa Cruz Biotechnology, TCI-America, or Sigma-Aldrich.

Design of CRISPRa program to regulate gene expression

CRISPRa tools enable programmable expression of heterologous genes in bacteria using synthetic promoters containing cognitive DNA target of CRISPR guide RNA (gRNA) (Dong et al., 2018; Fontana et al., 2020; Kiattisewee et al., 2021). The CRISPRa programs in this study follow the design described in Kiattisewee et al. while using up to three orthogonal scRNAs simultaneously targeting three separate transcriptional units.

3.3 Results

P. putida is tolerant to high p-ACA concentration

Attempts to produce p-ACA in growing *E. coli* from various groups have only managed to biosynthesize the p-aminophenylalanine (p-AF) intermediate (Figure 3.1A) (Burke et al., 2017; Masuo et al., 2016; Stevens and Carothers, 2015). Recent work was able to convert p-AF to p-ACA using *E. coli* resting cells (Minakawa et al., 2019), demonstrating the feasibility of bioconversion in bacterial cultures. One potential challenge in direct fermentation of p-ACA from glucose is the plausible p-ACA induced growth defects. To validate if p-ACA was indeed toxic to *E. coli*, we performed a p-ACA tolerance test of *E. coli* MG1655 in comparison with *P. putida* KT2440 (Figure 3.1B). We found that *E. coli* suffered from a severe growth defect with increasing p-ACA concentration (Figure 3.1C). Conversely, even at higher levels, p-ACA has only a marginal

effect on *P. putida* growth, indicating that accumulation of p-ACA in *P. putida* cultures is possible up to at least 20 mM (~320 mg/L). Moreover, intracellular concentration of metabolites produced inside the microbial cytoplasm could be much higher than that of external concentration (Mutanda et al., 2022), suggesting that *P. putida* is the appropriate host for p-ACA production in subsequent studies.

Furthermore, to rule out the possibility that p-ACA tolerance is due to catabolic activity of *P. putida*, we tested the ability of *P. putida* to grow on p-ACA as a sole carbon source. M9 media supplemented with 2 g/L of various carbon sources were employed for this test (Figure 3.S1A). In addition to wildtype strain, we also include the fcs-operon knock-out variant (KT2440::Δfcs-Δech-Δvdh, CKPP028) to the catabolic activity investigation based on an assumption that p-ACA catabolism might be similar to p-coumaric acid (p-CA), whose structure contains a hydroxyl group (-OH) instead of an amino group (-NH_2) (Williamson et al., 2020). Interestingly, we found that p-ACA cannot be used as a sole carbon source in both wildtype *P. putida* KT2440 and fcs-operon knockout strains. Apart from catabolism of p-ACA directly, it is also possible that p-ACA was degraded into inactive byproducts. To investigate potential p-ACA degradation, we measured p-ACA concentration over time in the growing cultures of *P. putida* in rich media (Figure 3.S1B). We found that p-ACA concentration remains unchanged in both wildtype and fcs-operon knockout strains, suggesting that p-ACA is metabolically stable and that accumulation of p-ACA in *P. putida* culture is achievable. With this knowledge, we move forward to express heterologous metabolic pathways in *P. putida* for p-ACA bioproduction.

p-AF bioproduction in *P. putida*

p-AF biosynthesis has previously been demonstrated in *E. coli* by heterologous expression of the *papABC* pathway, a three-gene operon from *P. fluorescens* (Masuo et al., 2016). PapABC pathway is capable of converting chorismate into p-aminophenylpyruvic acid (p-APP) which then could be converted into p-AF with endogenous or heterologous transaminase (TyrB) (Mohammadi Nargesi et al., 2019). As an initial test, we cloned the *papABC* operon pathway into *P. putida* compatible plasmids (pBBR1-GmR). To increase production of chorismate, a precursor to p-AF, we also overexpress a feed-back resistant AroG from *E. coli* together with AroL in the same operon to increase the production of chorismate (Ger et al., 1994; Juminaga Darmawi et al., 2012). These *aroGL* and *papABC* operons were placed under control of CRISPRa activatable promoters, J3-BBa_J23117 and J5-BBa_J23117, where gene expressions could be programmed using modified scaffold RNAs (scRNAs), a strategy previously developed for CRISPRa-mediated gene overexpression in *P. putida* (Kiattisewee et al., 2021).

CRISPRa programming with scRNA allows overexpression of the target gene in *P. putida* when scRNA targeting its cognate DNA pair is present. When the CRISPR-Cas complex is bound to the target DNA, the MS2-hairpin-containing scRNA recruits MS2-coated protein (MCP) that was covalently fused to an effector domain (SoxS) (Dong et al., 2018). The effector domain then recruits or stabilizes the RNA polymerase at the target promoter, leading to an increased transcription level. We designed a platform to generate orthogonal scRNAs/promoters pair for tunable overexpression of different transcriptional units which could be implemented in both *E. coli* and *P. putida* (Appendix H). Through scRNAs programming to upregulate *aroGL* and *papABC* expression (Figure 3.2A), we observed more than 1.3 mM (~238 mg/L) of p-AF in *P. putida* while *E. coli* yielded only up to 200 µM (~36 mg/L) using a similar expression strategy (Figures 3.2B &

3.2C). These findings imply that *P. putida* may possess a distinct and favorable metabolic background advantageous for aromatic amino acid pathway overproduction.

Direct p-ACA production from glucose using a growing *P. putida*

To complete the p-ACA biosynthetic pathway, we introduced a phenylalanine ammonia lyase enzyme (Pal) from *Arabidopsis thaliana* (At-PAL2, also referred to as At-Pal in this study) (Cochrane et al., 2004) alongside *E. coli* aromatic amino acid aminotransferase (TyrB). These enzymes were heterologously expressed using a second *P. putida* compatible plasmid and placed under control of CRISPRa scRNA (pRK2-KmR containing J6-BBa_J23117-pal2-tyrB as shown in Figure 3.3A). While previous efforts in *E. coli* had failed to produce detectable p-ACA, we surprisingly observed trace amounts of p-ACA (<5 μM) being produced in *P. putida* using this pathway (Figure 3.3B). To our knowledge, this marks the first instance of detectable p-ACA fermentation directly from glucose in a live bacterial culture. In these conditions, we still observed high accumulation of p-AF, the precursor of p-ACA, in the supernatant. The presence of p-AF in the supernatant suggested that the Pal reaction for converting p-AF into p-ACA might be inefficient. We also observed a high level of trans-cinnamic acid (t-CA) which suggests that phenylalanine, the native substrate of Pal, is being converted with high efficacy (Figure 3.S2). Since we saw high accumulation of p-AF, we assumed that endogenous TyrB is sufficient for transamination reaction and heterologous TyrB could be omitted to reduce metabolic burden from protein expression (Wu et al., 2016). We found that this strategy slightly improved p-ACA production and significantly reduced the accumulation of p-AF (Figure 3.3B). From these results, we hypothesized that this low p-ACA production could be due to inefficient catalytic properties inherent to the Pal enzyme from *A. thaliana*.

Despite several reports on enzyme discovery and engineering on the substrate scope of Pal, there is no report on the variant that exhibits high catalytic properties toward p-amino substituent. One study found that amine moiety at the para position of the phenyl ring introduced deactivating effects in the Pal reaction (Bartsch and Bornscheuer, 2010). A recent study investigated the substrate competition and inhibition effect of phenylalanine for p-ACA production in the whole-cell biocatalysis and suggested that Pal from *Rhodotorula glutinis* (Rg-Pal), with a weaker phenylalanine catalytic property, could be a potential candidate (Konishi et al., 2018). Here, we replaced At-Pal with the reported Rg-Pal enzyme that exhibited lower k_{cat} for phenylalanine (3.4 s^{-1} for Rg-pal compared to 5.1 s^{-1} for At-Pal) while maintaining similar k_{cat} for p-AF in an in vitro assay (0.44 s^{-1} and 0.40 s-1 for Rg-pal and At-Pal, respectively) (Figure 3.3C) (Konishi et al., 2018; Sariaslani, 2007). To our surprise, replacing At-Pal with Rg-Pal led to significant improvement in p-ACA level from less than 5 μM to 73.8 ± 9.8 μM (Figure 3.3D). The absence of *E. coli* TyrB also leveraged conversion of p-AF into p-ACA, yielding a 67% increase in p-ACA production (123 ± 0.9 μM). This result suggested that endogenous TyrB from *P. putida* is sufficient for conversion of p-APP into p-AF, rendering additional TyrB an expression burden without positive effects. Therefore, we chose the six-heterologous-gene pathway (*aroGL* from *E. coli*, *papABC* from *P. fluorescens*, and *pal* from *R. glutinis*) without heterologous *tyrB* for further optimization of p-ACA bioproduction in *P. putida*.

Optimizing CRISPRa-based gene expression for increasing p-ACA titer

As suggested by the decreased total metabolite levels in the two-plasmid p-ACA production system (~300 μM) compared to the single-plasmid p-AF production strain (1.3 mM), we hypothesized that the flux to aromatic amino acid pathway was decreasing due to an increased metabolic burden from the second plasmid incorporation (Kiattisewee et al., 2021). Therefore, optimizing the heterologous gene expression strategies alone might provide improvement to p-ACA titer. Here, we compared p-AF productions from three different strategies as previously tested by Kiattisewee and coworkers — i) one-plasmid system, ii) two-plasmid system, and iii) genome-integrated pathway. It was found that introduction of the second plasmid reduced the p-AF production from 1.3 mM to less than 100 μM at the same culture condition (Figure 3.S3). Similarly, when the same metabolic pathways were integrated into the genome, p-AF production was reduced to less than 20 μM. Thus, decreasing p-AF production in *P. putida* was likely due to decreasing expression levels of pathway enzymes.

Another factor that contributes to metabolic burden is the copy-number of heterologous genes containing plasmids. Previous literature suggested that high-copy pBBR1 plasmid could be burdensome and affect the growth rate of *P. putida* (Mi et al., 2016). By mutating the Rep protein (G159S) to reduce copy-number to approximately 25% of the original copy-number (25-30 copies, Cook et al., 2018), the burden from expressing the plasmid and relevant protein is completely alleviated (Mi et al., 2016). We also previously observed that moving the expression of a potentially toxic protein, e.g. dCas9, from the medium-copy plasmid to that of the genome-integrated version significantly reduced growth defects in *P. putida* (Kiattisewee et al., 2021). Here, we investigate gene expressions in three different copy-numbers — i) medium-copy pBBR1 plasmid, ii) low-copy pBBR1-G159S plasmid (Mi et al., 2016), and iii) single-copy genome integration (Figure 3.4A). Using the previously characterized CRISPRa-mediated mRFP expression under different minimal promoters, we test the dynamic range of expression level from basal (Off-target scRNA) and CRISPR-activated (J306) with three described copy-number (Figure 3.S4A). We found that fold-change derived from pBBR1 plasmid yielded the largest fold-change and dynamic range (27-fold) from the weakly-expressed minimal promoter in all cases. When the copy-number decreased to low-copy and single-copy, the fold-change significantly decreased altogether with the expression level of the CRISPR-activated level (24-fold for pBBR1-G159S and 4.6-fold for genome integration). Even though fold-change in the medium-copy and low-copy plasmids are comparable (27-fold vs. 24-fold), we observed a big difference in the fluorescence level of mRFP (~5900 a.u. vs. ~2000 a.u.) that might affect total expressions of metabolic enzymes for metabolic engineering applications. These results underscore the potential of combining genetic payloads from the two-plasmid system into a single medium-copy plasmid, exhibiting high expression levels of pathway enzymes and high dynamic range from the CRISPRa program.

Consequently, we constructed a relatively large (13kb) medium-copy plasmid carrying both metabolic genes and a scRNA program, resulting in an increased p-ACA titer from 123 ± 0.9 μM to 473 ± 13 μM (Figure 3.4B). However, this also led to a significant growth defect noticeable from smaller colony sizes compared to no pathway control, requiring two-day incubation at 30 ˚C to be visible. To address this point, we introduced G159S mutation to the pBBR1 for copy-number reduction and found that growth defects could be mitigated (Figure 3.S5A). Unexpectedly, the

titer of p-AF significantly decreased from 713 ± 217 μM to 163 ± 203 μM likely due to reduced expression levels of pathway enzymes and also yielded high variation across replicates (Figure 3.S5B). We believe that moving the metabolic pathway to the genome will substantially decrease the burden associated with plasmid replication even better than that of copy-number lowering but gene expression level from decreasing copy-number should be considered. Therefore, we move forward to construct the genome-integrated variant of p-ACA production *P. putida* strains with increasing basal promoter strengths to match the expression level achieved from the plasmid-based system.

Informed by the different expression levels derived from different minimal promoter strengths (Figure 3.S4B), we chose medium strength BBa_J23105 and BBa_J23110 as minimal promoter candidates for CRISPR-mediated gene expression from the genomic single-copy. Due to the toxicity of p-ACA pathway in *E. coli*, the construction of donor plasmids require a reduced expression of pathway genes during cloning in an *E. coli* host. We shuffled the upstream sequence to low basal expression variants as suggested by our reporter engineering works together with auxiliary CRISPRi plasmid that suppresses pathway genes expression in the cloning phase (Alba Burbano et al., 2023; Wen et al., 2021). We first tested the expression of AroGL and PapABC from the genome for the p-AF production (Figure 3.S6A). Expressing metabolic genes from a weaker BBa_J23105 appeared to be more stable and yielded higher p-AF production. We also tested integration of Rg-Pal into the genome while supplying a p-AF production pathway on the plasmid (Figure 3.S6B). We observed stable production of p-ACA as well as accumulation of p-AF, suggesting that more Pal is needed to fully convert p-AF to p-ACA. By combinatorially expressing AroGL/PapABC and Pal from either genomic or plasmid levels, we observed full conversion of p-AF to p-ACA in all conditions tested (Figure 3.S6B). Despite negligible growth defects, the genome-integrated variant with the highest p-ACA titer provided only half of the level from that of the single-plasmid system. This decrease in p-ACA production could also be attributed to limiting resource utilization where fast-growing microbes used up significant nutrients for biomass synthesis rather than desired chemical reactions (Lipson, 2015). In summary, these results overall suggested that a single-plasmid expression system is the most suitable expression approach for p-ACA production in *P. putida*.

Further validation of p-ACA bioproduction toward sustainable p-AS production

With the successful establishment of p-ACA production in *P. putida*, we shifted our focus to the utilization of cost-effective feedstocks (Molina-Santiago et al., 2016; Nikel et al., 2016; Nikel and de Lorenzo, 2018; Volke et al., 2020; Weimer et al., 2020). Recent development in *P. putida* engineering allowed it to efficiently catabolize various feedstocks including lignocellulosic biomass derivatives (Dvořák and de Lorenzo, 2018; Elmore et al., 2020; Eng et al., 2021) and waste-derived feedstocks (Franden et al., 2018; Werner et al., 2021). To demonstrate that the p-ACA production is sustained in different carbon sources, we investigated the p-ACA bioproduction of the *P. putida* strain in minimal media (M9) instead of rich media (EZ) using glucose as a main carbon source. Here, we found that p-ACA production strain grows significantly slower in M9 minimal medium and requires further investigation on the culture time. We stepped back to investigate the production rate of p-ACA in rich media (EZ) and found that p-AF accumulation occurred at an earlier stage without conversion to p-ACA. This agrees with a previous report on substrate competition with the native substrate of Pal, phenylalanine, which is always present

during the growth of *P. putida* (Minakawa et al., 2019). Only after phenylalanine was depleted through conversion into trans-cinnamic acid (t-CA), the reaction of p-AF to p-ACA would initiate (Figures 3.S7A & 3.S7C). Another aromatic amino acid, tyrosine, was also transformed into p-coumaric acid (p-CA) via a Pal-mediated reaction but ended up being consumed by *P. putida* at the later stages. This finding agreed with the previous analysis that *P. putida* is capable of catabolizing p-CA (Elmore et al., 2020), and could also contribute to the bioproduction of p-ACA. Based on these insights, p-ACA production in M9 likely required an extended culture time to compromise with the substrate competition with phenylalanine and tyrosine. By extending the culture time from 26 h to 72 h, we found that *P. putida* cultures reached a stationary state and yielded ~70 μM (~11 mg/L) of p-ACA with negligible p-AF leftover (Figure 3.5B and Figure 3.S8). This p-ACA production in the minimal medium suggested that p-ACA production could be performed with other complex media of choice as long as the feed can be reintroduced into the central metabolism for aromatic amino acid biosynthesis.

3.4 Discussion and Conclusions

Industrial bioproduction of chemicals from renewable feedstocks has served as an alternative route toward a sustainable society. In the past decades, developments in genetic engineering and strains discovery have transformed the field toward unconventional organisms. The early work on genetic engineering encompassed mainly model organisms which are not always transferable to other microbes. The advent of novel tools in the CRISPR era allowed rapid development in non-model organisms which granted a shortcut for engineers on pathway engineering in the metabolically versatile host such as *Pseudomonas putida*. In this work, we found that the recently developed CRISPRa tools could be applied for bioproduction of p-aminocinnamic acid (p-ACA), previously challenging in *E. coli*, the model organism for bacteria.

p-ACA can be enzymatically synthesized from the non-canonical amino acid, p-aminophenylalanine (p-AF), in a single-step reaction. It can be coupled with carboxylic acids to generate polyimides and polyamides, with properties comparable to petroleum-based polymers (Suvannasara et al., 2014; Tateyama et al., 2016). Transformation of p-ACA into p-aminostyrene (p-AS) could also be achieved through chemical catalysis which is a precursor of functionalized polystyrene (Sheng et al., 2020; van Schijndel et al., 2020). In this work, we used a multi-gene CRISPRa program to regulate three transcription units of six to seven genes relevant to p-ACA biosynthesis. Given the programmability and expandability of CRISPRa/i circuits, a larger circuit could be further constructed where additional guide RNAs could be used to control additional genes in this aromatic amine pathway or used for other metabolic pathways of interest. Since CRISPRa/i tools are applicable for regulating endogenous genes (Kiattisewee et al., 2021), it is possible to couple the endogenous and heterologous genes together and rewire the native metabolism to accommodate engineered programs. Such targets could be central carbon metabolism to uplift the shikimate and chorismate pathway upstream of aromatic amino acid biosynthesis, optimize the relevant pathway in cofactor regeneration, or interfere with the competing pathway that draws the flux away for desired products.

These CRISPRa-mediated expression levels generally served as a prototype for programmable gene expression that can be reprogrammed using arbitrary tools both with cis- and trans-acting mechanisms. However, the delayed production of p-ACA also poses a requirement

of dynamic programming of gene expression where Pal enzyme can be expressed only at the later stage when productions of phenylalanine subsided. Based on our observation on metabolic burden in *P. putida*, CRISPRa-based delayed expression could provide an alternative strategy for expression dynamics. With programmability of CRISPRa/i, it is possible to control the gene expression dynamics further by equipping genetic circuitry with other tools such as metabolite-responsive RNA aptamers or transcription factors (Hwang and Carothers, 2016; Rottinghaus et al., 2022). Previous works in *E. coli* have demonstrated that dynamic genetic programming can optimize resource utilization to increase titer of the desired compounds (Dinh and Prather, 2019; Gao et al., 2021; Kim et al., 2017; Tian et al., 2020; Wu et al., 2020; Yang et al., 2018).

Apart from genetic programming, bioproduction could also be enhanced via enzyme selection and engineering. Screening of enzyme orthologs based on relevant kinetic parameters showed successful improvement in this work using substrate specificity to guide promiscuity. It is also possible to engineer or evolve the enzyme further based on side-product selection or biosensor (Mays et al., 2020). Previous work has shown that adaptive laboratory evolution strategies could provide optimal genetic context to heterologous metabolic pathways (Lim et al., 2021). The existing strategies to remove carboxyl groups from p-ACA are prevalent in the chemical processes (Sheng et al., 2020; van Schijndel et al., 2020). However, p-ACA enzymatic decarboxylation has not been demonstrated despite successful conversion of various cinnamic acid derivatives into corresponding styrenes by different decarboxylase enzymes (e.g. PadC, FDC1, PAD1) (Grubbe et al., 2020; Lee et al., 2019; Liu et al., 2018; McKenna and Nielsen, 2011; Mukai et al., 2010; Tomek et al., 2015; Verhoef Suzanne et al., 2009; Williamson et al., 2020). Therefore, biochemical conversion of p-ACA to p-aminostyrene (p-AS) could also be further explored with an aid of enzyme engineering for substrate compatibility (Duță et al., 2022; Li et al., 2021; Liu et al., 2015; Shen et al., 2020).

Finally, the recent advancement in feedstock utilization has also paved the way for a sustainable bioeconomy. By engineering the metabolism of microorganisms with robust capability, a wider range of compounds could be fed to the engineered organisms and further valorize the accumulated biological wastes (Dvořák and de Lorenzo, 2018; Elmore et al., 2020; Erickson et al., 2022; Franden et al., 2018; Werner et al., 2021). Non-biological treatment also plays a role in simplifying the complex feedstocks to accessible molecules for bacteria (Abu-Omar et al., 2021; Erickson et al., 2022; Schutyser et al., 2018; Sun et al., 2018). From this work, it is evident that aromatic compounds such as p-coumaric acid (p-CA), produced from a side reaction of Pal, could be fed back to *P. putida* metabolism to then synthesize compounds of interest. By combining genetic engineering strategies (knock-out, knock-in) with that of trans-acting genetic regulatory circuits developed in this work, we foresee that the bioproduction research in industrial microbiology could be accelerated in a plug-and-play manner, regardless of microbial hosts.

Author contributions

C.K. and I.D.F. designed experiments and analyzed data. C.K. and I.D.F. performed experiments. C.K., I.D.F., J.M.C., and J.G.Z. wrote the manuscript.

Declaration of competing interest

J.G.Z. and J.M.C are members of the Wayfinder Biosciences scientific advisory board.

Acknowledgments

We thank members of the Zalatan and Carothers groups for technical assistance, advice, and helpful discussions. This work was supported by NSF Award 1817623 (J.M.C, J.G.Z.), NSF Award 2225632 (J.G.Z. and J.M.C.), DOE Award DE-EE0008927 (J.M.C, J.G.Z.) and DOE Award DE-SC0023091 (J.M.C, J.G.Z.).

Supporting Information

Supplementary data associated with this work is provided at the end of this chapter.

3.5 Tables and Figures

Table 3.1: Selected p-AF and p-ACA production titers in this work

Figure	***aroGL* expression**	***papABC* expression**	***pal* expression**	**p-AF in μM (mg/L[a])**	**p-ACA in μM (mg/L[a])**
2B (*E. coli*)	pBBR1-GmR J3-BBa_J23117-*aroGL*	pBBR1-GmR J5-BBa_J23117-*papABC*	–	209 ± 55 (37.7 ± 9.8)	N.D.
2C (*P. putida*)	pBBR1-GmR J3-BBa_J23117-*aroGL*	pBBR1-GmR J5-BBa_J23117-*papABC*	–	1322 ± 153 (238 ± 27.5)	N.D.
3B	pBBR1-GmR J3-BBa_J23117-*aroGL*	pBBR1-GmR J5-BBa_J23117-*papABC*	pRK2-KmR J6-BBa_J23117-*At-pal-Ec-tyrB*	117 ± 7.5 (21.0 ± 1.4)	<1 (<1)
3B	pBBR1-GmR J3-BBa_J23117-*aroGL*	pBBR1-GmR J5-BBa_J23117-*papABC*	pRK2-KmR J6-BBa_J23117-*At-pal*	20.8 ± 4.0 (3.7 ± 0.7)	2.3 ± 0.26 (<1)
3D	pBBR1-GmR J3-BBa_J23117-*aroGL*	pBBR1-GmR J5-BBa_J23117-*papABC*	pRK2-KmR J6-BBa_J23117-*Rg-pal-Ec-tyrB*	269 ± 54 (48.6 ± 9.8)	73.9 ± 9.8 (12.1 ± 1.6)
3D & 4B (2-plasmid)	pBBR1-GmR J3-BBa_J23117-*aroGL*	pBBR1-GmR J5-BBa_J23117-*papABC*	pRK2-KmR J6-BBa_J23117-***Rg-pal***	151 ± 21 (27.3 ± 3.9)	123 ± 0.9 (20.1 ± 0.1)
4B (1-plasmid)	pBBR1-GmR J3-BBa_J23117-*aroGL*	pBBR1-GmR J5-BBa_J23117-*papABC*	pBBR1-GmR J6-BBa_J23117-*Rg-pal*	13.4 ± 5.3 (2.4 ± 0.95)	473 ± 13 (77 ± 2.1)
4B (genomic)	Genomic J3(LL)-BBa_J23105-*aroGL*	Genomic J5(LL)-BBa_J23105-*papABC*	Genomic J6-BBa_J23105-*Rg-pal*	N.D.	127 ± 18 (20.7 ± 3.0)
5 (M9, 72h)	pBBR1-GmR J3-BBa_J23117-*aroGL*	pBBR1-GmR J5-BBa_J23117-*papABC*	pBBR1-GmR J6-BBa_J23117-*Rg-pal*	N.D.	67.8 ± 17.2 (11.1 ± 2.8)

Notes: [a]Production titers in mg/L were shown in parentheses, N.D. stands for Not Detectable

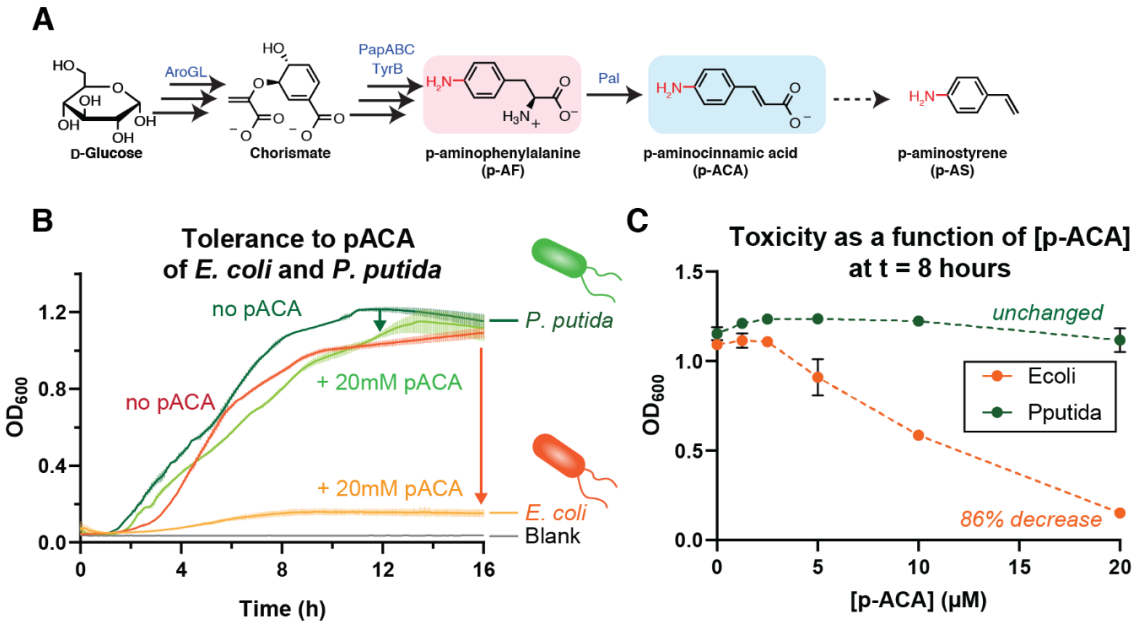


Figure 3.1: *P. putida* is resistant to high concentrations of aromatic amines

(A) Biosynthetic pathway of p-aminostyrene starting from glucose. *aroGL* operon was retrieved from *E. coli* with a feedback-resistant mutant introduced in *aroG*. *papABC* operon was reconstructed from *P. fluorescens*. pal enzymes were selected from *A. thaliana* or *R. glutinis*. Conversion of p-ACA to p-AS was left to be further explored biochemically or fed through abiotic process. (B) Growth curve of *E. coli* and *P. putida* with and without p-ACA in EZ-RDM culture. (C) OD_{600} at 8 hours after subculture showed resistance of *P. putida* across a wide range of p-ACA concentrations while *E. coli* suffered from growth defects as p-ACA concentration increased.

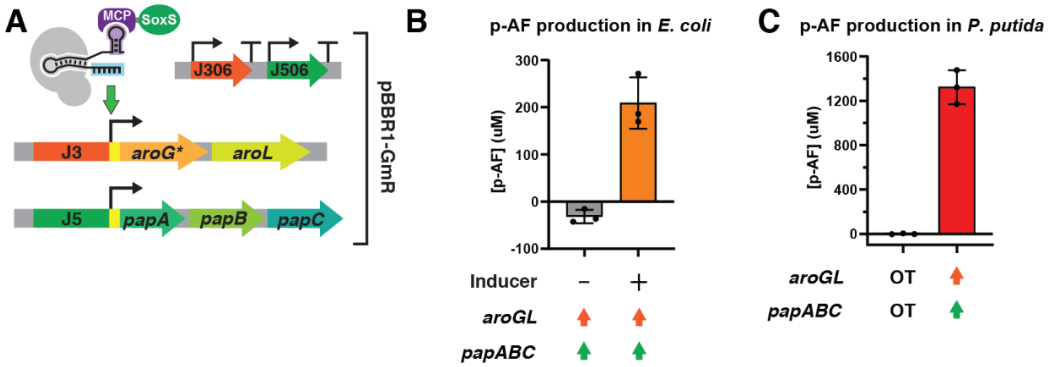


Figure 3.2: p-AF production in *E. coli* and *P. putida*

(A) Operon organization of *aroGL* and *papABC* under control of CRISPR-mediated promoters. (B) p-AF production in *E. coli* under aTc-inducible conditions. (C) p-AF production in *P. putida* showed significantly higher efficiency compared to that of *E. coli*.

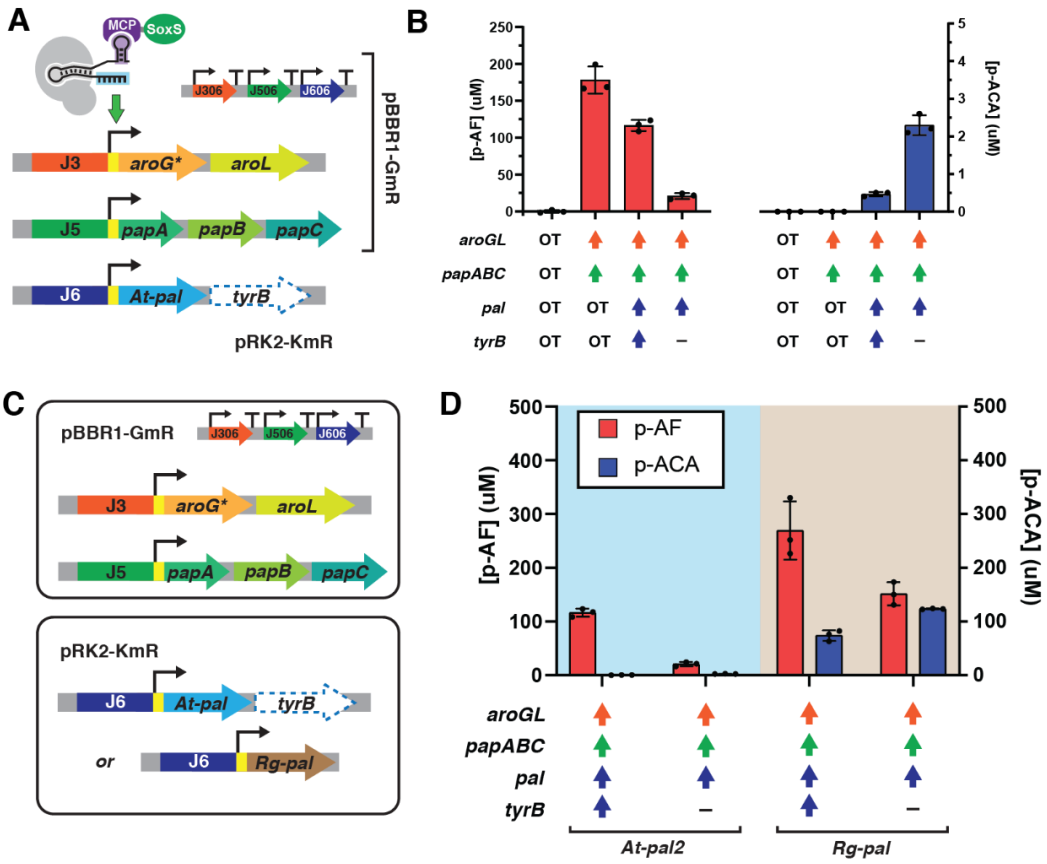


Figure 3.3: Optimizing p-ACA bioproduction in *P. putida* with pathway engineering

(A) Introduction of *At-pal* with and without *Ec-tyrB* in the second plasmid, pRK2-KmR, to implement p-ACA production. (B) p-AF and p-ACA titers in different metabolic pathway constructions. Trace level of p-ACA could be observed. (C) Variation of Pal cassette from At-pal to Rg-pal. (D) p-AF and p-ACA titer in different pathway construction. Rg-pal without Ec-tyrB yielded the best p-ACA level across all conditions.

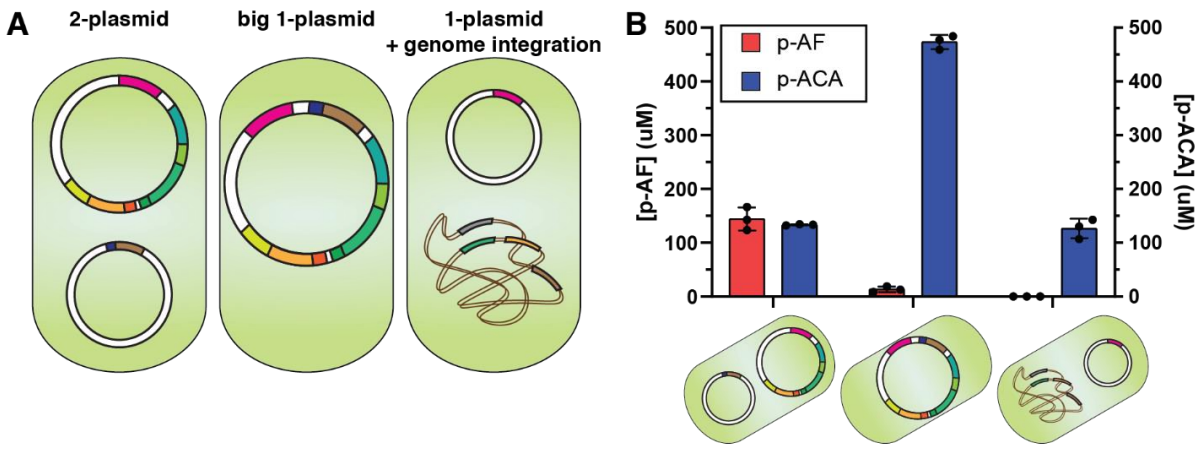


Figure 3.4: Expression strategies optimization

(A) Heterologous genes delivery were tested in three conditions: two plasmids, one big plasmid, and integrating pathway genes to the genome (CRISPR gRNA program was expressed from a small plasmid). (B) p-AF and p-ACA productions in three different conditions. Expression levels in the genome-integrated strategy were screened for optimal titer (See Supplementary Figure 3.S6). One plasmid system yielded the highest titer. See Table 3.1 for listed titers from different conditions.

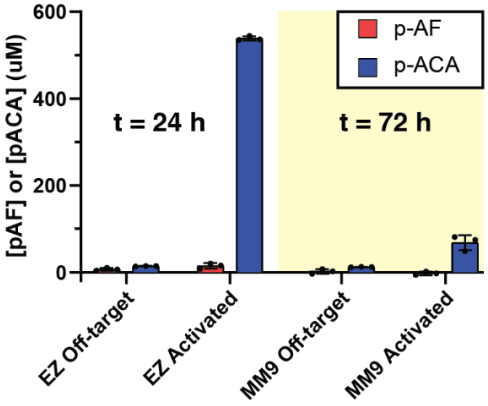


Figure 3.5: p-ACA production in M9 minimal media

The one-plasmid expression strain was investigated in the culture of M9 minimal media with 0.2% glucose as a carbon source. Growth rate of p-ACA producing strain in M9 appeared to be significantly slower than EZ-RDM (see Supplementary Figure 3.S7-3.S8 for further details).

3.6 References

Abu-Omar, M.M., Barta, K., Beckham, G.T., Luterbacher, J.S., Ralph, J., Rinaldi, R., Román-Leshkov, Y., Samec, J.S.M., Sels, B.F., Wang, F., **2021**. Guidelines for performing lignin-first biorefining. *Energy Environ. Sci.* **14**, 262–292. https://doi.org/10.1039/D0EE02870C

Alba Burbano, D., Cardiff, R.A.L., Tickman, B.I., Kiattisewee, C., Maranas, C.J., Zalatan, J.G., Carothers, J.M., **2023**. Engineering activatable promoters for scalable and multi-input CRISPRa/i circuits. *Proc. Natl. Acad. Sci.* **120**, e2220358120. https://doi.org/10.1073/pnas.2220358120

Bartsch, S., Bornscheuer, U.T., **2010**. Mutational analysis of phenylalanine ammonia lyase to improve reactions rates for various substrates. *Protein Eng. Des. Sel.* **23**, 929–933. https://doi.org/10.1093/protein/gzq089

Bernhardsgrütter, I., Stoffel, G.M., Miller, T.E., Erb, T.J., **2021**. CO2-converting enzymes for sustainable biotechnology: from mechanisms to application. *Curr. Opin. Biotechnol.* **67**, 80–87. https://doi.org/10.1016/j.copbio.2021.01.003

Burke, C.R., Sparkman-Yager, D., Carothers, J.M., **2017**. Multi--state design of kinetically--controlled RNA aptamer ribosensors. *bioRxiv*. https://doi.org/10.1101/213538

Cochrane, F.C., Davin, L.B., Lewis, N.G., **2004**. The *Arabidopsis* phenylalanine ammonia lyase gene family: kinetic characterization of the four PAL isoforms. *Proteomics 1* **65**, 1557–1564. https://doi.org/10.1016/j.phytochem.2004.05.006

Cook, T.B., Rand, J.M., Nurani, W., Courtney, D.K., Liu, S.A., Pfleger, B.F., **2018**. Genetic tools for reliable gene expression and recombineering in *Pseudomonas putida*. *J. Ind. Microbiol. Biotechnol.* **45**, 517–527. https://doi.org/10.1007/s10295-017-2001-5

Cywar, R.M., Rorrer, N.A., Hoyt, C.B., Beckham, G.T., Chen, E.Y.-X., **2022**. Bio-based polymers with performance-advantaged properties. *Nat. Rev. Mater.* **7**, 83–103. https://doi.org/10.1038/s41578-021-00363-3

Dinh, C.V., Prather, K.L.J., **2019**. Development of an autonomous and bifunctional quorum-sensing circuit for metabolic flux control in engineered *Escherichia coli*. *Proc. Natl. Acad. Sci.* **116**, 25562–25568. https://doi.org/10.1073/pnas.1911144116

Dong, C., Fontana, J., Patel, A., Carothers, J.M., Zalatan, J.G., **2018**. Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nat Commun* **9**, 2489. https://doi.org/10.1038/s41467-018-04901-6

Duță, H., Filip, A., Nagy, L.C., Nagy, E.Z.A., Tőtős, R., Bencze, L.C., **2022**. Toolbox for the structure-guided evolution of ferulic acid decarboxylase (FDC). *Sci. Rep.* **12**, 3347. https://doi.org/10.1038/s41598-022-07110-w

Dvořák, P., de Lorenzo, V., **2018**. Refactoring the upper sugar metabolism of *Pseudomonas putida* for co-utilization of cellobiose, xylose, and glucose. *Metab. Eng.* **48**, 94–108. https://doi.org/10.1016/j.ymben.2018.05.019

Elmore, J.R., Dexter, G.N., Baldino, H., Huenemann, J.D., Francis, R., Peabody, G.L., Martinez-Baird, J., Riley, L.A., Simmons, T., Coleman-Derr, D., Guss, A.M., Egbert, R.G., **2023**. High-throughput genetic engineering of nonmodel and undomesticated bacteria via iterative site-specific genome integration. *Sci. Adv.* **9**, eade1285. https://doi.org/10.1126/sciadv.ade1285

Elmore, J.R., Dexter, G.N., Salvachúa, D., O'Brien, M., Klingeman, D.M., Gorday, K., Michener, J.K., Peterson, D.J., Beckham, G.T., Guss, A.M., **2020**. Engineered *Pseudomonas putida* simultaneously catabolizes five major components of corn stover lignocellulose: Glucose, xylose, arabinose, p-coumaric acid, and acetic acid. *Metab. Eng.* **62**, 62–71. https://doi.org/10.1016/j.ymben.2020.08.001

Eng, T., Banerjee, D., Lau, A.K., Bowden, E., Herbert, R.A., Trinh, J., Prahl, J.-P., Deutschbauer, A., Tanjore, D., Mukhopadhyay, A., **2021**. Engineering *Pseudomonas*

putida for efficient aromatic conversion to bioproduct using high throughput screening in a bioreactor. *Metab. Eng.* **66**, 229–238. https://doi.org/10.1016/j.ymben.2021.04.015
Erickson, E., Bleem, A., Kuatsjah, E., Werner, A.Z., DuBois, J.L., McGeehan, J.E., Eltis, L.D., Beckham, G.T., **2022**. Critical enzyme reactions in aromatic catabolism for microbial lignin conversion. *Nat. Catal.* **5**, 86–98. https://doi.org/10.1038/s41929-022-00747-w
Fatma, Z., Schultz, J.C., Zhao, H., **2020**. Recent advances in domesticating non-model microorganisms. *Biotechnol. Prog.* **36**, e3008. https://doi.org/10.1002/btpr.3008
Filiciotto, L., Rothenberg, G., **2021**. Biodegradable Plastics: Standards, Policies, and Impacts. *ChemSusChem* **14**, 56–72. https://doi.org/10.1002/cssc.202002044
Fontana, J., Dong, C., Kiattisewee, C., Chavali, V.P., Tickman, B.I., Carothers, J.M., Zalatan, J.G., **2020**. Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. *Nat. Commun.* **11**, 1618. https://doi.org/10.1038/s41467-020-15454-y
Franden, M.A., Jayakody, L.N., Li, W.-J., Wagner, N.J., Cleveland, N.S., Michener, W.E., Hauer, B., Blank, L.M., Wierckx, N., Klebensberger, J., Beckham, G.T., **2018**. Engineering *Pseudomonas putida* KT2440 for efficient ethylene glycol utilization. *Metab. Eng.* **48**, 197–207. https://doi.org/10.1016/j.ymben.2018.06.003
Gao, C., Guo, L., Hu, G., Liu, J., Chen, X., Xia, X., Liu, L., **2021**. Engineering a CRISPRi Circuit for Autonomous Control of Metabolic Flux in *Escherichia coli*. *ACS Synth. Biol.* **10**, 2661–2671. https://doi.org/10.1021/acssynbio.1c00294
Ger, Y.-M., Chen, S.-L., Chiang, H.-J., Shiuan, D., **1994**. A Single Ser-180 Mutation Desensitizes Feedback Inhibition of the Phenylalanine-Sensitive 3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate (DAHP) Synthetase in *Escherichia coli*. *J. Biochem. (Tokyo)* **116**, 986–990. https://doi.org/10.1093/oxfordjournals.jbchem.a124657
Gonzalez de Gortari, M., Misra, M., Mohanty, A.K., **2022**. Polyphthalamide polymers: A review on synthesis, properties, and advance manufacturing and emerging applications. *J. Appl. Polym. Sci.* **139**, e52965. https://doi.org/10.1002/app.52965
Grubbe, W.S., Rasor, B.J., Krüger, A., Jewett, M.C., Karim, A.S., **2020**. Cell-free styrene biosynthesis at high titers. *Metab. Eng.* **61**, 89–95. https://doi.org/10.1016/j.ymben.2020.05.009
Han, X., Liu, J., Tian, S., Tao, F., Xu, P., **2022**. Microbial cell factories for bio-based biodegradable plastics production. *iScience* **25**, 105462. https://doi.org/10.1016/j.isci.2022.105462
Hwang, C., Carothers, J.M., **2016**. Label-free selection of RNA aptamers for metabolic engineering. *Vitro Sel. Evol.* **106**, 37–41. https://doi.org/10.1016/j.ymeth.2016.06.016
Intasian, P., Prakinee, K., Phintha, A., Trisrivirat, D., Weeranoppanant, N., Wongnate, T., Chaiyen, P., **2021**. Enzymes, In Vivo Biocatalysis, and Metabolic Engineering for Enabling a Circular Economy and Sustainability. *Chem. Rev.* **121**, 10367–10451. https://doi.org/10.1021/acs.chemrev.1c00121
Jatain, I., Dubey, K.K., Sharma, Manisha, Usmani, Z., Sharma, Minaxi, Gupta, V.K., **2021**. Synthetic biology potential for carbon sequestration into biocommodities. *J. Clean. Prod.* **323**, 129176. https://doi.org/10.1016/j.jclepro.2021.129176
Juminaga Darmawi, Baidoo Edward E. K., Redding-Johanson Alyssa M., Batth Tanveer S., Burd Helcio, Mukhopadhyay Aindrila, Petzold Christopher J., Keasling Jay D., **2012**. Modular Engineering of l-Tyrosine Production in *Escherichia coli*. *Appl. Environ. Microbiol.* **78**, 89–98. https://doi.org/10.1128/AEM.06017-11
Kawaguchi, H., Hasunuma, T., Ogino, C., Kondo, A., **2016**. Bioprocessing of bio-based chemicals produced from lignocellulosic feedstocks. *Chem. Biotechnol. • Pharm. Biotechnol.* **42**, 30–39. https://doi.org/10.1016/j.copbio.2016.02.031
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*. *Metab. Eng.* **66**, 283–295. https://doi.org/10.1016/j.ymben.2021.04.002
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. https://doi.org/10.1016/j.ymben.2017.11.004
Konishi, K., Takaya, N., Masuo, S., Zhou, S., **2018**. 4-amino cinnamic acid production method using enzyme.
Lee, K., Bang, H.B., Lee, Y.H., Jeong, K.J., **2019**. Enhanced production of styrene by engineered *Escherichia coli* and in situ product recovery (ISPR) with an organic solvent. *Microb. Cell Factories* **18**, 79. https://doi.org/10.1186/s12934-019-1129-6
Li, J., Weng, Z., Cao, Q., Qi, Y., Lu, B., Zhang, S., Wang, J., Jian, X., **2022**. Synthesis of an aromatic amine derived from biomass and its use as a feedstock for versatile epoxy thermoset. *Chem. Eng. J.* **433**, 134512. https://doi.org/10.1016/j.cej.2022.134512
Li, Q., Xia, Y., Zhao, T., Gong, Y., Fang, S., Chen, M., **2021**. Improving the catalytic characteristics of phenolic acid decarboxylase from *Bacillus amyloliquefaciens* by the engineering of N-terminus and C-terminus. *BMC Biotechnol.* **21**, 44. https://doi.org/10.1186/s12896-021-00705-7
Lim, H.G., Eng, T., Banerjee, D., Alarcon, G., Lau, A.K., Park, M.-R., Simmons, B.A., Palsson, B.O., Singer, S.W., Mukhopadhyay, A., Feist, A.M., **2021**. Generation of *Pseudomonas putida* KT2440 Strains with Efficient Utilization of Xylose and Galactose via Adaptive Laboratory Evolution. *ACS Sustain. Chem. Eng.* **9**, 11512–11523. https://doi.org/10.1021/acssuschemeng.1c03765
Lipson, D.A., **2015**. The complex relationship between microbial growth rate and yield and its implications for ecosystem processes. *Front. Microbiol.* **6**.
Liu, C., Men, X., Chen, H., Li, M., Ding, Z., Chen, G., Wang, F., Liu, H., Wang, Q., Zhu, Y., Zhang, H., Xian, M., **2018**. A systematic optimization of styrene biosynthesis in *Escherichia coli* BL21(DE3). *Biotechnol. Biofuels* **11**, 14. https://doi.org/10.1186/s13068-018-1017-z
Liu, S.P., Zhang, L., Mao, J., Ding, Z.Y., Shi, G.Y., **2015**. Metabolic engineering of *Escherichia coli* for the production of phenylpyruvate derivatives. *Metab. Eng.* **32**, 55–65. https://doi.org/10.1016/j.ymben.2015.09.007
Lu, L., Shen, X., Sun, X., Yan, Y., Wang, J., Yuan, Q., **2022**. CRISPR-based metabolic engineering in non-model microorganisms. *Curr. Opin. Biotechnol.* **75**, 102698. https://doi.org/10.1016/j.copbio.2022.102698
Masuo, S., Zhou, S., Kaneko, T., Takaya, N., **2016**. Bacterial fermentation platform for producing artificial aromatic amines. *Sci. Rep.* **6**, 25764. https://doi.org/10.1038/srep25764
Mays, Z.J., Mohan, K., Trivedi, V.D., Chappell, T.C., Nair, N.U., **2020**. Directed evolution of Anabaena variabilis phenylalanine ammonia-lyase (PAL) identifies mutants with enhanced activities. *Chem. Commun.* **56**, 5255–5258. https://doi.org/10.1039/D0CC00783H
McKenna, R., Nielsen, D.R., **2011**. Styrene biosynthesis from glucose by engineered *E. coli*. *Metab. Eng.* **13**, 544–554. https://doi.org/10.1016/j.ymben.2011.06.005
Mehl, R.A., Anderson, J.C., Santoro, S.W., Wang, L., Martin, A.B., King, D.S., Horn, D.M., Schultz, P.G., **2003**. Generation of a Bacterium with a 21 Amino Acid Genetic Code. *J. Am. Chem. Soc.* **125**, 935–939. https://doi.org/10.1021/ja0284153
Mi, J., Sydow, A., Schempp, F., Becher, D., Schewe, H., Schrader, J., Buchhaupt, M., **2016**. Investigation of plasmid-induced growth defect in *Pseudomonas putida*. *J. Biotechnol.* **231**, 167–173. https://doi.org/10.1016/j.jbiotec.2016.06.001
Minakawa, H., Masuo, S., Kaneko, T., Takaya, N., **2019**. Fermentation and purification of

microbial monomer 4-amminocinnamic acid to produce ultra-high performance bioplastics. *Process Biochem.* **77**, 100–105. https://doi.org/10.1016/j.procbio.2018.11.021
Mohammadi Nargesi, B., Sprenger, G.A., Youn, J.-W., **2019**. Metabolic Engineering of *Escherichia coli* for para-Amino-Phenylethanol and para-Amino-Phenylacetic Acid Biosynthesis. *Front. Bioeng. Biotechnol.* **6**.
Molina-Santiago, C., Cordero, B.F., Daddaoua, A., Udaondo, Z., Manzano, J., Valdivia, M., Segura, A., Ramos, J.-L., Duque, E., **2016**. *Pseudomonas putida* as a platform for the synthesis of aromatic compounds. *Microbiology*. https://doi.org/10.1099/mic.0.000333
Mukai, N., Masaki, K., Fujii, T., Kawamukai, M., Iefuji, H., **2010**. PAD1 and FDC1 are essential for the decarboxylation of phenylacrylic acids in *Saccharomyces cerevisiae*. *J. Biosci. Bioeng.* **109**, 564–569. https://doi.org/10.1016/j.jbiosc.2009.11.011
Mutanda, I., Sun, J., Jiang, J., Zhu, D., **2022**. Bacterial membrane transporter systems for aromatic compounds: Regulation, engineering, and biotechnological applications. *Biotechnol. Adv.* **59**, 107952. https://doi.org/10.1016/j.biotechadv.2022.107952
Nielsen, J., Keasling, J.D., **2016**. Engineering Cellular Metabolism. *Cell* **164**, 1185–1197. https://doi.org/10.1016/j.cell.2016.02.004
Nikel, P.I., Chavarría, M., Danchin, A., de Lorenzo, V., **2016**. From dirt to industrial applications: *Pseudomonas putida* as a Synthetic Biology chassis for hosting harsh biochemical reactions. *Curr. Opin. Chem. Biol.* **34**, 20–29. https://doi.org/10.1016/j.cbpa.2016.05.011
Nikel, P.I., de Lorenzo, V., **2018**. *Pseudomonas putida* as a functional chassis for industrial biocatalysis: From native biochemistry to trans-metabolism. *Metab. Eng.* **50**, 142–155. https://doi.org/10.1016/j.ymben.2018.05.005
Noda, S., Miyazaki, T., Miyoshi, T., Miyake, M., Okai, N., Tanaka, T., Ogino, C., Kondo, A., **2011**. Cinnamic acid production using *Streptomyces lividans* expressing phenylalanine ammonia lyase. *J. Ind. Microbiol. Biotechnol.* **38**, 643–648. https://doi.org/10.1007/s10295-011-0955-2
Qi, W.W., Vannelli, T., Breinig, S., Ben-Bassat, A., Gatenby, A.A., Haynie, S.L., Sariaslani, F.S., **2007**. Functional expression of prokaryotic and eukaryotic genes in *Escherichia coli* for conversion of glucose to p-hydroxystyrene. *Metab. Eng.* **9**, 268–276. https://doi.org/10.1016/j.ymben.2007.01.002
Rottinghaus, A.G., Xi, C., Amrofell, M.B., Yi, H., Moon, T.S., **2022**. Engineering ligand-specific biosensors for aromatic amino acids and neurochemicals. *Cell Syst.* **13**, 204-214.e4. https://doi.org/10.1016/j.cels.2021.10.006
Sariaslani, F.S., **2007**. Development of a Combined Biological and Chemical Process for Production of Industrial Aromatics from Renewable Resources. *Annu. Rev. Microbiol.* **61**, 51–69. https://doi.org/10.1146/annurev.micro.61.080706.093248
Schutyser, W., Renders, T., Van den Bosch, S., Koelewijn, S.-F., Beckham, G.T., Sels, B.F., **2018**. Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. *Chem. Soc. Rev.* **47**, 852–908. https://doi.org/10.1039/C7CS00566K
Shen, Y.-P., Niu, F.-X., Yan, Z.-B., Fong, L.S., Huang, Y.-B., Liu, J.-Z., **2020**. Recent Advances in Metabolically Engineered Microorganisms for the Production of Aromatic Chemicals Derived From Aromatic Amino Acids. *Front. Bioeng. Biotechnol.* **8**.
Sheng, W.-B., Sumera, Y., Chen, C., Zhang, M., Lei, Y., Peng, C., Li, B., Jian, Y., Wang, W., **2020**. Metal-free and base-free decarboxylation of substituted cinnamic acids in a deep eutectic solvent. *Synth. Commun.* **50**, 558–563. https://doi.org/10.1080/00397911.2019.1708945
Singhvi, M.S., Zinjarde, S.S., Gokhale, D.V., **2019**. Polylactic acid: synthesis and biomedical applications. *J. Appl. Microbiol.* **127**, 1612–1626. https://doi.org/10.1111/jam.14290
Stevens, J.T., Carothers, J.M., **2015**. Designing RNA-based genetic control systems for efficient

production from engineered metabolic pathways. *ACS Synth Biol* **4**, 107–15. https://doi.org/10.1021/sb400201u
Sun, X., Shen, X., Jain, R., Lin, Y., Wang, Jian, Sun, J., Wang, Jia, Yan, Y., Yuan, Q., **2015**. Synthesis of chemicals by metabolic engineering of microbes. *Chem. Soc. Rev.* **44**, 3760–3785. https://doi.org/10.1039/C5CS00159E
Sun, Z., Fridrich, B., de Santi, A., Elangovan, S., Barta, K., **2018**. Bright Side of Lignin Depolymerization: Toward New Platform Chemicals. *Chem. Rev.* **118**, 614–678. https://doi.org/10.1021/acs.chemrev.7b00588
Suvannasara, P., Tateyama, S., Miyasato, A., Matsumura, K., Shimoda, T., Ito, T., Yamagata, Y., Fujita, T., Takaya, N., Kaneko, T., **2014**. Biobased Polyimides from 4-Aminocinnamic Acid Photodimer. *Macromolecules* **47**, 1586–1593. https://doi.org/10.1021/ma402499m
Tadrent, S., Luart, D., Bals, O., Khelfa, A., Luque, R., Len, C., **2018**. Metal-Free Reduction of Nitrobenzene to Aniline in Subcritical Water. *J. Org. Chem.* **83**, 7431–7437. https://doi.org/10.1021/acs.joc.8b00406
Tateyama, S., Masuo, S., Suvannasara, P., Oka, Y., Miyazato, A., Yasaki, K., Teerawatananond, T., Muangsin, N., Zhou, S., Kawasaki, Y., Zhu, L., Zhou, Z., Takaya, N., Kaneko, T., **2016**. Ultrastrong, Transparent Polytruxillamides Derived from Microbial Photodimers. *Macromolecules* **49**, 3336–3342. https://doi.org/10.1021/acs.macromol.6b00220
Tian, J., Yang, G., Gu, Y., Sun, X., Lu, Y., Jiang, W., **2020**. Developing an endogenous quorum-sensing based CRISPRi circuit for autonomous and tunable dynamic regulation of multiple targets in *Streptomyces*. *Nucleic Acids Res.* **48**, 8188–8202. https://doi.org/10.1093/nar/gkaa602
Tomek, K.J., Saldarriaga, C.R.C., Velasquez, F.P.C., Liu, T., Hodge, D.B., Whitehead, T.A., **2015**. Removal and upgrading of lignocellulosic fermentation inhibitors by in situ biocatalysis and liquid-liquid extraction. *Biotechnol. Bioeng.* **112**, 627–632. https://doi.org/10.1002/bit.25473
Tsuge, Y., Kawaguchi, H., Sasaki, K., Kondo, A., **2016**. Engineering cell factories for producing building block chemicals for bio-polymer synthesis. *Microb. Cell Factories* **15**, 19. https://doi.org/10.1186/s12934-016-0411-0
Urtuvia, V., Villegas, P., González, M., Seeger, M., **2014**. Bacterial production of the biodegradable plastics polyhydroxyalkanoates. *Int. J. Biol. Macromol.* **70**, 208–213. https://doi.org/10.1016/j.ijbiomac.2014.06.001
van Schijndel, J., Molendijk, D., van Beurden, K., Canalle, L.A., Noël, T., Meuldijk, J., **2020**. Preparation of bio-based styrene alternatives and their free radical polymerization. *Eur. Polym. J.* **125**, 109534. https://doi.org/10.1016/j.eurpolymj.2020.109534
Verhoef Suzanne, Wierckx Nick, Westerhof R. G. Maaike, de Winde Johannes H., Ruijssenaars Harald J., **2009**. Bioproduction of p-Hydroxystyrene from Glucose by the Solvent-Tolerant Bacterium *Pseudomonas putida* S12 in a Two-Phase Water-Decanol Fermentation. *Appl. Environ. Microbiol.* **75**, 931–936. https://doi.org/10.1128/AEM.02186-08
Volk, M.J., Tran, V.G., Tan, S.-I., Mishra, S., Fatma, Z., Boob, A., Li, H., Xue, P., Martin, T.A., Zhao, H., **2023**. Metabolic Engineering: Methodologies and Applications. *Chem. Rev.* **123**, 5521–5570. https://doi.org/10.1021/acs.chemrev.2c00403
Volke, D.C., Calero, P., Nikel, P.I., **2020**. *Pseudomonas putida*. *Trends Microbiol.* **28**, 512–513. https://doi.org/10.1016/j.tim.2020.02.015
Wang, G., Zhao, Z., Ke, J., Engel, Y., Shi, Y.-M., Robinson, D., Bingol, K., Zhang, Z., Bowen, B., Louie, K., Wang, B., Evans, R., Miyamoto, Y., Cheng, K., Kosina, S., De Raad, M., Silva, L., Luhrs, A., Lubbe, A., Hoyt, D.W., Francavilla, C., Otani, H., Deutsch, S., Washton, N.M., Rubin, E.M., Mouncey, N.J., Visel, A., Northen, T., Cheng, J.-F., Bode, H.B., Yoshikuni, Y., **2019**. CRAGE enables rapid activation of biosynthetic gene clusters

in undomesticated bacteria. *Nat. Microbiol.* **4**, 2498–2510. https://doi.org/10.1038/s41564-019-0573-8
Weimer, A., Kohlstedt, M., Volke, D.C., Nikel, P.I., Wittmann, C., **2020**. Industrial biotechnology of *Pseudomonas putida*: advances and prospects. *Appl. Microbiol. Biotechnol.* **104**, 7745–7766. https://doi.org/10.1007/s00253-020-10811-9
Wen, X., Zhang, Y., Cheng, H., An, J., Guo, Y., Wang, L., Wang, M., **2021**. A CRISPR/dCas9-assisted system to clone toxic genes in *Escherichia coli*. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1865**, 129994. https://doi.org/10.1016/j.bbagen.2021.129994
Werner, A.Z., Clare, R., Mand, T.D., Pardo, I., Ramirez, K.J., Haugen, S.J., Bratti, F., Dexter, G.N., Elmore, J.R., Huenemann, J.D., Peabody, G.L., Johnson, C.W., Rorrer, N.A., Salvachúa, D., Guss, A.M., Beckham, G.T., **2021**. Tandem chemical deconstruction and biological upcycling of poly(ethylene terephthalate) to β-ketoadipic acid by *Pseudomonas putida* KT2440. *Metab. Eng.* **67**, 250–261. https://doi.org/10.1016/j.ymben.2021.07.005
Westenberg, R., Peralta-Yahya, P., **2023**. Toward implementation of carbon-conservation networks in nonmodel organisms. *Curr. Opin. Biotechnol.* **81**, 102949. https://doi.org/10.1016/j.copbio.2023.102949
Williamson, J.J., Bahrin, N., Hardiman, E.M., Bugg, T.D.H., **2020**. Production of Substituted Styrene Bioproducts from Lignin and Lignocellulose Using Engineered *Pseudomonas putida* KT2440. *Biotechnol. J.* **15**, 1900571. https://doi.org/10.1002/biot.201900571
Wirth, N.T., Kozaeva, E., Nikel, P.I., **2019**. Accelerated genome engineering of *Pseudomonas putida* by I-SceI-mediated recombination and CRISPR-Cas9 counterselection. *Microb Biotechnol* **13**. https://doi.org/10.1111/1751-7915.13396
Wu, G., Yan, Q., Jones, J.A., Tang, Y.J., Fong, S.S., Koffas, M.A.G., **2016**. Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. *Trends Biotechnol.* **34**, 652–664. https://doi.org/10.1016/j.tibtech.2016.02.010
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 Res.* **48**, 996–1009. https://doi.org/10.1093/nar/gkz1123
Yan, Q., Fong, S.S., **2017**. Challenges and Advances for Genetic Engineering of Non-model Bacteria and Uses in Consolidated Bioprocessing. *Front. Microbiol.* **8**. https://doi.org/10.3389/fmicb.2017.02060
Yang, Y., Lin, Y., Wang, Jian, Wu, Y., Zhang, R., Cheng, M., Shen, X., Wang, Jia, Chen, Z., Li, C., Yuan, Q., Yan, Y., **2018**. Sensor-regulator and RNAi based bifunctional dynamic control network for engineered microbial synthesis. *Nat. Commun.* **9**, 3043. https://doi.org/10.1038/s41467-018-05466-0

Supplementary Information

Supplementary Figures

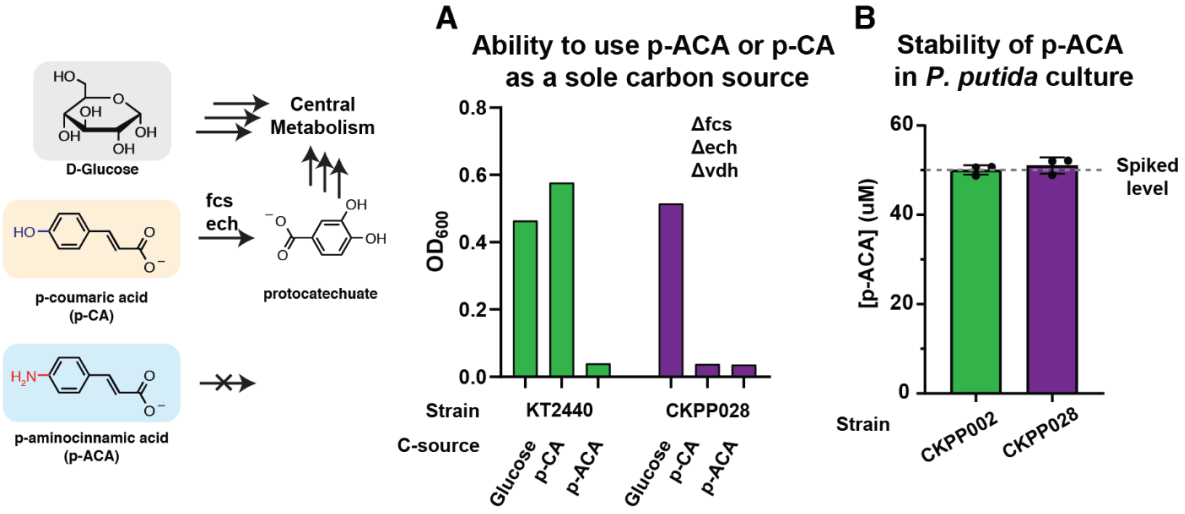


Figure 3.S1: Catabolic activity of *P. putida* toward p-ACA and related chemicals

(A) Test of *P. putida* ability to use p-ACA as a sole carbon-source using glucose and p-coumaric acid as controls. Experiments were performed in M9 with 0.2% of the selected carbon source. CKPP028 with p-CA catabolic genes knock-outwere tested side-by-side. (B) Investigation of p-ACA stability in the growing *P. putida* culture. 50 μM of p-ACA was spiked in the *P. putida* inoculation in EZ-RDM. No significant changes were observed.

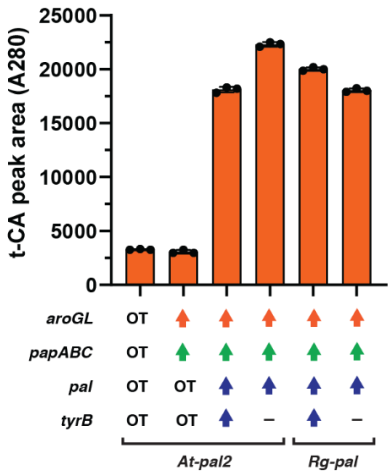


Figure 3.S2: Accumulation of trans-cinnamic acid during p-ACA bioproduction

Observation of trans-cinnamic acid (t-CA) when Pal enzyme was introduced into *P. putida*. t-CA is considered one of the key side-products in p-ACA biosynthesis.

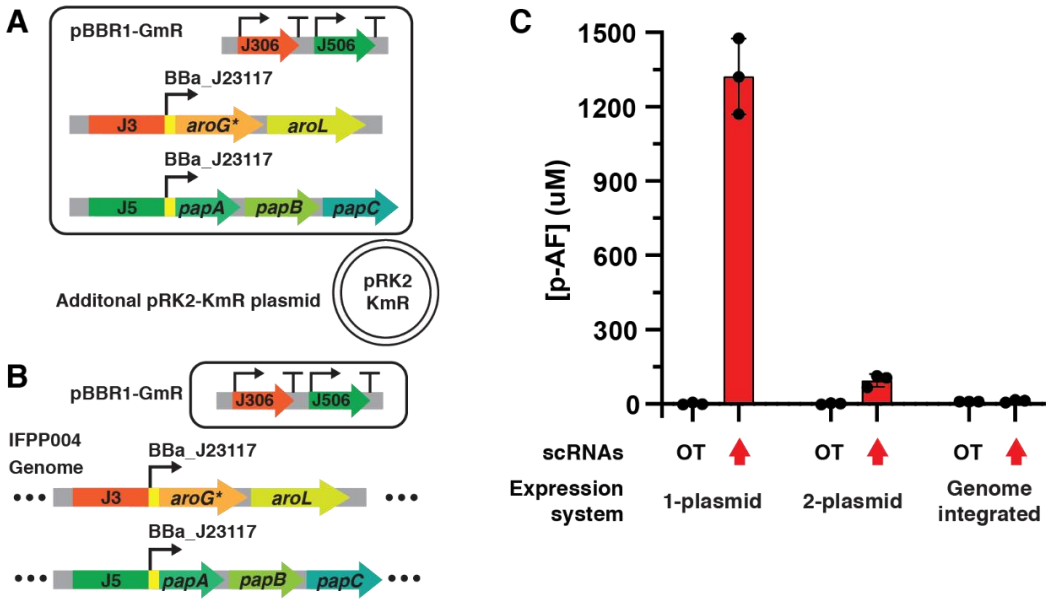


Figure 3.S3: p-AF bioproduction in the presence of the 2nd plasmid and from the genome integration

(A) Addition of the second pRK2-KmR plasmid (pCK334) as an expression cassette for further pathway incorporation. (B) Integration of p-AF biosynthetic genes cluster into the *P. putida* genome. scRNA programs were kept on the plasmid level to retain simple programmability. (C) p-AF titer from two-plasmid condition and genome-integrated systems in comparison to the single-plasmid system.

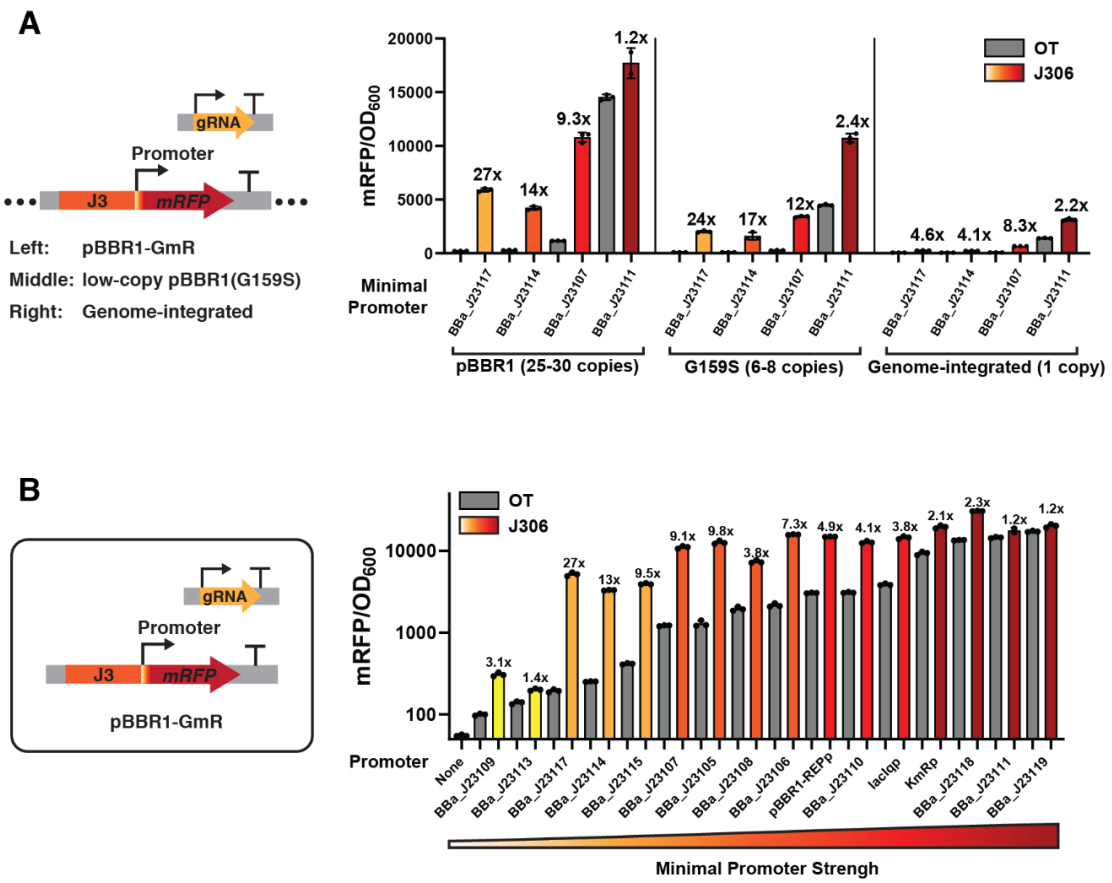


Figure 3.S4: Gene expression ranges and dynamic ranges from CRISPRa with different copy-number and minimal promoter strengths

(A) CRISPRa-mediated gene expression in different copy-numbers with distinct minimal promoter strengths. Decreasing copy-number led to lower gene expression ranges and dynamic ranges. (B) Full catalog of minimal promoter with CRISPRa-mediated gene expression. BBa_J23110 (moderately strong) and BBa_J23105 (moderately weak) were selected for screening in genomic expression.

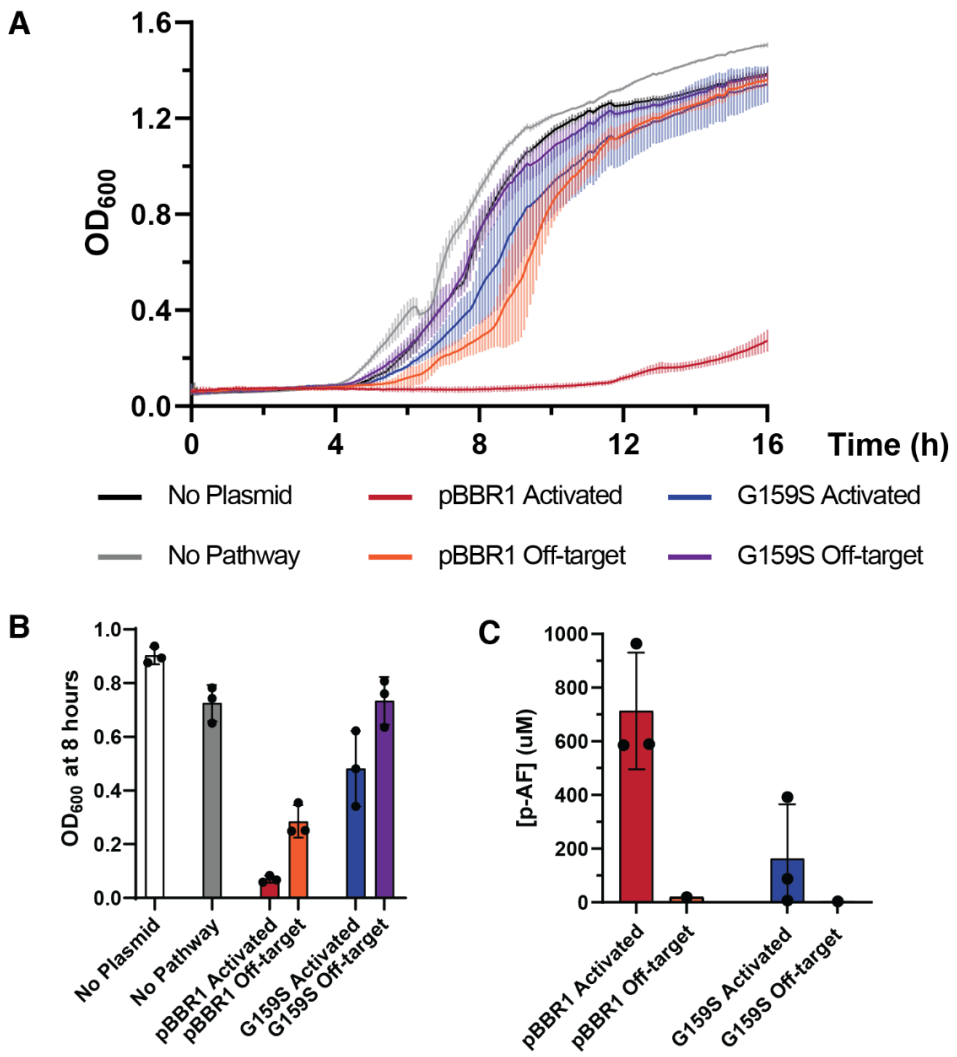


Figure 3.S5: Effect of pathway gene expression and copy-number on growth-rate and bioproduction of p-AF in *P. putida*

(A) Growth rate of *P. putida* bearing different metabolic programs and copy-number. (B) Growth of *P. putida* captured at 8 hours after inoculation. (C) p-AF production of the pathway in medium-copy plasmid (pBBR1) and low-copy plasmid (pBBR1-G159S).

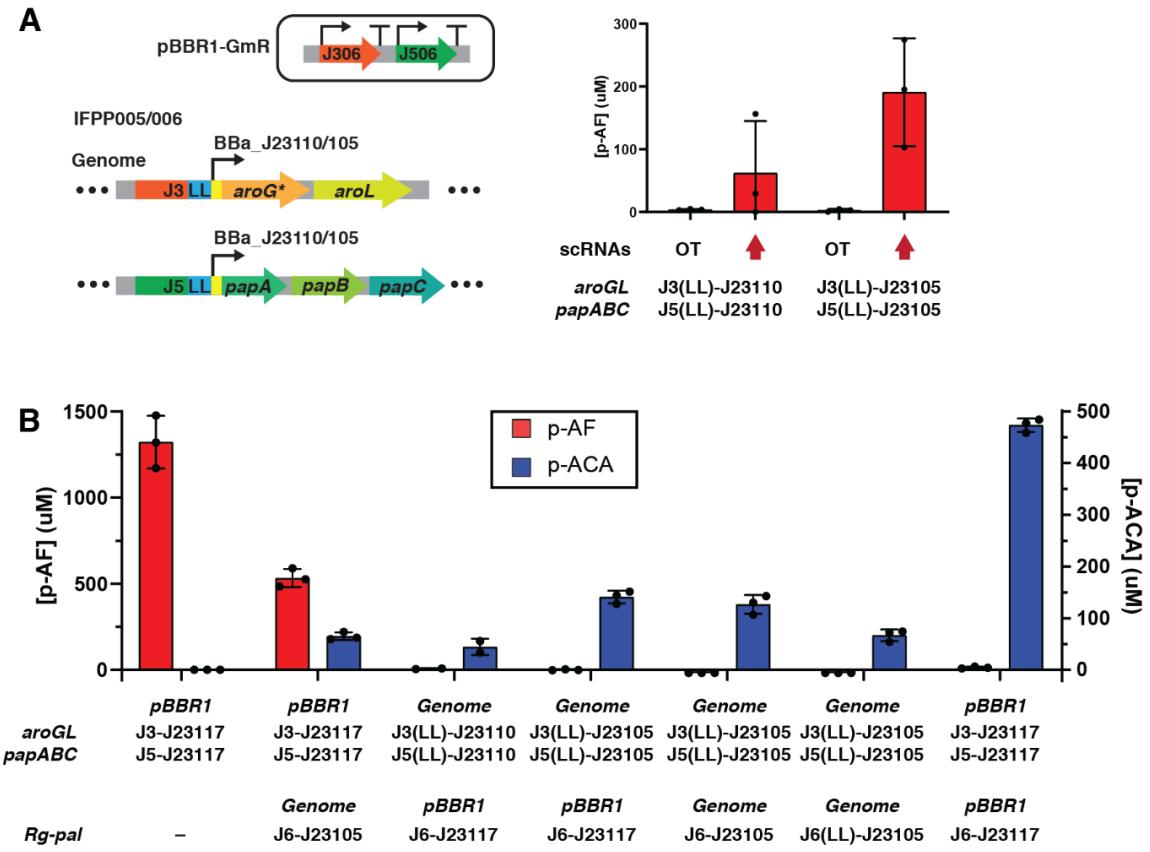


Figure 3.S6: p-AF and p-ACA production from genome-integrated pathways

(A) p-AF production with integrated *aroGL* and *papABC* cassettes under BBa_J23110 and BBa_J23105 minimal promoters. (B) Further investigation of p-ACA productions under different expression strategies of each cassette. Expressing all pathway genes on the plasmid yielded the highest p-ACA level.

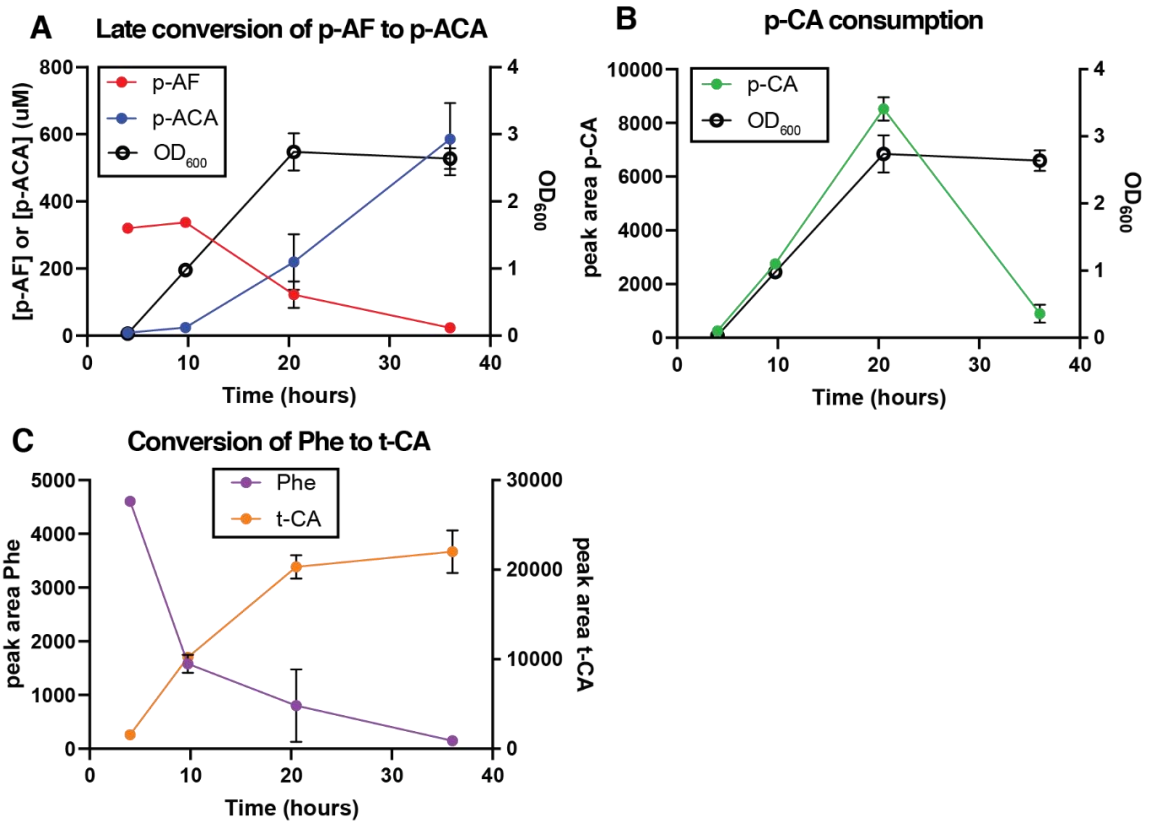


Figure 3.S7: Time-course measurement of aromatic compounds in EZ-RDM

(A) p-ACA production cultures were collected at different inoculation times — 4 h, 9.5 h, 20.5 h, and 36 h. The results suggested that p-AF accumulates at the early stages while slowly being converted into p-ACA at the later stages. (B) p-CA accumulation was shown until late stationary state where it was catabolized by *P. putida*. (C) Conversion of Phenylalanine (Phe) into t-CA was shown to align with the p-ACA production rate. This finding suggested that p-AF-to-p-ACA conversion occur later when Phe is depleted from the system.

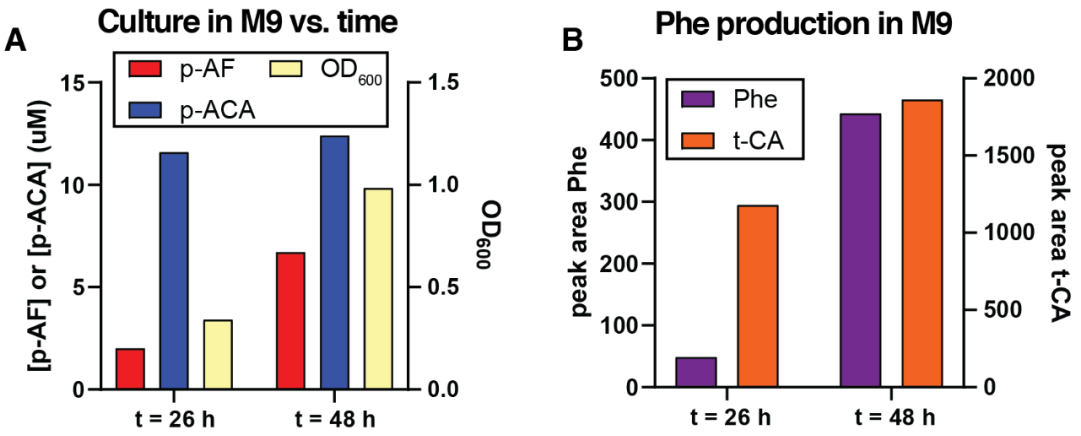


Figure 3.S8: Growth and aromatic compounds production in M9

(A) p-AF and p-ACA levels remain low in the early cultures in minimal medium (M9 with 0.2% glucose). (B) Phe and t-CA accumulates in the 48 hours collection time suggesting that competitive inhibition of Phe toward p-AF-to-p-ACA conversion is still in-progress.

Supplementary Tables

Table 3.S1: Engineered *P. putida* strains for p-ACA production

Strains	Features	Sources
P. putida KT2440	Wildtype strain	Harwood lab
CKPP002 (PPC01)	KT2440 with integrated *Sp*.pCas9-dCas9 and BBa_J23107-MCP-SoxS	Kiattisewee et al., 2021
CKPP028	CKPP002, *fcs*-operon knock-out	This study
IFPP002	CKPP002 with integration of J3-J23117-aroGL, J5-J23117-*papABC*, and J6-J23117-*At-pal-Ec-tyrB*	This study
IFPP004	CKPP002 with integration of J3-J23117-*aroGL* and J5-J23117-*papABC*	This study
CKPP031	CKPP002 with integration of J3-J23117-mRFP	This study
CPP0032	CKPP002 with integration of J3-J23114-mRFP	This study
CKPP033	CKPP002 with integration of J3-J23107-mRFP	This study
CKPP034	CKPP002 with integration of J3-J23111-mRFP	This study
IFPP005	CKPP002 with integration of J3(LL)-J23110-*aroGL* and J5(LL)-J23110-*papABC*	This study
IFPP006	CKPP002 with integration of J3(LL)-J23105-*aroGL* and J5(LL)-J23105-*papABC*	This study
IFPP007	CKPP002 with integration of J6-J23105-*Rg-pal*	This study
IFPP008	IFPP006 with integration of J6-J23105-*Rg-pal*	This study
IFPP009	IFPP006 with integration of J6(LL)-J23105-*Rg-pal*	This study

Table 3.S2: Plasmids used in this study

Plasmids	Features	Sources
pCK120.gRNA (pPPC020)	pBBR1, J3-BBa_J23117-mRFP, scRNA, GmR	Kiattisewee et al., 2021
pCK265.gRNA (pPPC008)	pBBR1, sgRNA or scRNA, GmR	Kiattisewee et al., 2021
pCK082.gRNA (pPPC019)	pRK2, J1-BBa_J23117-mRFP, scRNA, KmR	Kiattisewee et al., 2021
pGNW2-pp1	pGNW2 derivative with integration site at prophage1, KmR	Kiattisewee et al., 2021
pGNW2-pp2	pGNW2 derivative with integration site at prophage2, KmR	Kiattisewee et al., 2021
pPPC031	pGNW2 derivative with integration site at prophage1 for integration of J3-BBa_J23117-mRFP cassette, KmR	Kiattisewee et al., 2021
pCK255	pBBR1 bearing I-SceI and *sacB* genes, GmR	Kiattisewee et al., 2021
pCK354	p15A, *Sp*.pCas9-dCas9 and TetR-Ptet-MCP-SoxS(R93A/S101A), CmR	This study
pIDFP01.J3-J5-J6	J306, J506, J606 scRNAs expressed under BBa_J23105 promoter	This study
pIDFP01.OT-OT-OT	Three Off-target scRNAs expressed under BBa_J23105 promoter	This study
pJFP229B.J3-J5-J6	J3-J23117-*aroGL*, J5-J23117-*papABC*, J306, J506, J606 scRNAs expressed under BBa_J23105 promoter	This study
pJFP229B.OT-OT-OT	J3-J23117-*aroGL*, J5-J23117-*papABC*, Three Off-target scRNAs expressed under BBa_J23105 promoter	This study
pIDFP02	J3(LL)-J23110-*aroGL*, J5(LL)-J23110-*papABC*, or J23105 version	This study
pCK334	J6-*At-pal2-Ec-tyrB*_scRNA(hAAVS1)	This study
pCK388	J6-*At-pal2*_scRNA(hAAVS1)	This study
pCK425	J6-*Rg-pal*	This study
pCK426	J6-*Rg-pal-Ec-tyrB*	This study
pCK439	Expressing *Rg-pal* only	This study
pCK120.promoter	pBBR1, J3-promoter-mRFP, scRNA, GmR	This study
pCK248.promoter	pBBR1-G159S, J3-promoter-mRFP, scRNA, GmR	This study
pCK261.promoter	pGNW2 derivative with integration site at prophage1 for integration of J3-BBa_J23117-mRFP cassette, KmR	Kiattisewee et al., 2021 and this study
pCK440	Big plasmid with three pathways, J306/J506/J606	This study
pCK456	Big plasmid with three pathways, off-target scRNAs	This study
pCK443	Big plasmid with three pathways, with TyrB	This study

pCK232	*fcs*-operon knock-out	This study
pJFP229B-int	For integration of J3-J23117-*aroGL* and J5-J23117-*papABC*	This study
pIDFP02-105-int	For integration of J3(LL)-J23105-*aroGL* and J5(LL)-J23105-*papABC*	This study
pIDFP02-110-int	For integration of J3(LL)-J23110-*aroGL* and J5(LL)-J23110-*papABC*	This study
pCK520	For integration of J6-J23105-*Rg-pal*	This study
pCK520.E	For integration of J6(LL)-J23105-*Rg-pal*	This study

DNA sequences

The DNA sequence of biological parts used in this study were provided below in the 5'-to-3' format of the non-template strand. Each part was color-coded based on its function: red-promoter, blue-CDS, brown-upstream sequence, dark green-5'-UTR, purple-terminator, and others as specified. The **transcription start site** (TSS), **start codon** and **stop codon** are **bolded**. DNA sequences related to CRISPRa machinery can be found in the previously published manuscript (Kiattisewee et al., 2021).

Representative Metabolic Pathways

>J3-BBa_J23117-aroGL-dblT (RBS)

AGCATTTGCGATCATTCACGCAGCGCTTATTCAGTTGCTCACTGCGATGTCATAATCATCGCTACGAGCTGTGAAAG
ATGCATAAAGCTCGTACGACGCGTTCGCTCGTCTCCTCACTTCTCCTACGGAGCGTTCTGGACACAACGTCGTCTTG
AAGTTGCGATTATAGATTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGCGAATTCAAAAGATCTAAATAACCTAA
ACGAGAGGAAAGAATA**ATG**AATTATCAGAACGACGATTTACGCATCAAAGAAATCAAAGAGTTACTTCCTCCTGTCG
CATTGCTGGAAAAATTCCCCGCTACTGAAAATGCCGCGAATACGGTTGCCCATGCCCGAAAAGCGATCCATAAGATC
CTGAAAGGTAATGATGATCGCCTGTTGGTTGTGATTGGCCCATGCTCAATTCATGATCCTGTCGCGGCAAAAGAGTA
TGCCACTCGCTTGCTGGCGCTGCGTGAAGAGCTGAAAGATGAGCTGGAAATCGTAATGCGCGTCTATTTTGAAAAGC
CGCGTACCACGGTGGGCTGGAAAGGGCTGATTAACGATCCGCATATGGATAATAGCTTCCAGATCAACGACGGTCTG
CGTATAGCCCGTAAATTGCTGCTTGATATTAACGACAGCGGTCTGCCAGCGGCAGGTGAGTTTCTCAACATGATCAC
CCCACAATATCTCGCTGACCTGATGAGCTGGGGCGCAATTGGCGCACGTACCACCGAATCGCAGGTGCACCGCGAAC
TGGCATCAGGGCTTTCTTGTCCGGTCGGCTTCAAAAATGGCACCGACGGTACGATTAAAGTGGCTATCGATGCCATT
AATGCCGCCGGTGCGCCGCACTGCTTCCTGTCCGTAACGAAATGGGGGCATTCGGCGATTGTGAATACCAGCGGTAA
CGGCGATTGCCATATCATTCTGCGCGGCGGTAAAGAGCCTAACTACAGCGCGAAGCACGTTGCTGAAGTGAAAGAAG
GGCTGAACAAAGCAGGCCTGCCAGCACAGGTGATGATCGATTTCAGCCATGCTAACTCGTCCAAACAATTCAAAAAG
CAGATGGATGTTTGTGCTGACGTTTGCCAGCAGATTGCCGGTGGCGAAAAGGCCATTATTGGCGTGATGGTGGAAAG
CCATCTGGTGGAAGGCAATCAGAGCCTGGAGAGCGGGGAGCCGCTGGCCTACGGTAAGAGCATCACCGATGCCTGCA
TCGGCTGGGAAGATACCGATGCTCTGTTACGTCAACTGGCGAATGCAGTAAAAGCGCGTCGCGGG**TAA**GGATCTAAA
GGAGGCCATCC**ATG**ACACAACCTCTTTTTCTGATCGGGCCTCGGGGCTGTGGTAAAACAACGGTCGGAATGGCCCTT
GCCGATTCGCTTAACCGTCGGTTTGTCGATACCGATCAGTGGTTGCAATCACAGCTCAATATGACGGTCGCGGAGAT
CGTCGAAAGGGAAGAGTGGGCGGGATTTCGCGCCAGAGAAACGGCGGCGCTGGAAGCGGTAACTGCGCCATCCACCG
TTATCGCTACAGGCGGCGGCATTATTCTGACGGAATTTAATCGTCACTTCATGCAAAATAACGGGATCGTGGTTTAT
TTGTGTGCGCCAGTATCAGTCCTGGTTAACCGACTGCAAGCTGCACCGGAAGAAGATTTACGGCCAACCTTAACGGG
AAAACCGCTGAGCGAAGAAGTTCAGGAAGTGCTGGAAGAACGCGATGCGCTATATCGCGAAGTTGCGCATATTATCA
TCGACGCAACAAACGAACCCAGCCAGGTGATTTCTGAAATTCGCAGCGCCCTGGCACAGACGATCAATTGT**TAA**CCA
GGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTC
TACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATA

>J5-BBa_J23117-papABC-dblT

TATACATCGCATCACTACACTATTGATTATCATTGTGTACGTAACGAGCTTGCACAACGTGAAGTTCTTCGAGCACT
TCAGCTCGCAACGTAAATGACAGTTGCTGTTAAGTGACGTGAATCCTTCAATGCTGCTCATGCTGCTGTCGTAAATA
AGTAAGTCACTCCCACTTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGCGAATTCAAAAGATCTAACTGGTAATT
TGAGGAGGTAATTT**ATG**AAGATCCTGCTGATCGATAACTTTGATAGCTTCACCCAGAATATTGCCCAGTATCTGTAT
GAAGTTACCGGTATTTGTGCCGATATTGTTACCAATACCGTGACCTATGAACATCTGCAAATCGAACAGTATGATGC
CGTTGTTCTGAGCCCTGGTCCGGGTCATCCGGGTGAATATCTGGATTTTGGTGTTTGTGGTCAGGTGATTCTGCATA
GTCCGGTTCCGCTGCTGGGTATTTGTCTGGGTCATCAGGGTATTGCACAGTTTTTAGGTGGCACCGTTGGTCATGCA
CCGACACCGGTTCATGGTTATCGTAGCAAAATTACCCATAGCGGTAGCGGTCTGTTTCGTGATCTGCCGGAACAGTT
TGAAGTTGTTCGTTATCATAGCCTGATGTGTACCCATCTGCCGCAAGAACTGCGTTGTACCGCATGGACCGAAGAGG

GTGTTGTTATGGCAATTGAACATGAAAGCCGTCCGATTTGGGGTGTTCAGTTTCATCCGGAAAGCATTGATAGCGAA
TATGGTCATGCCCTGCTGAGCAACTTTATTGGTATGGCCATCGAACATAATGGCAATCATCGTACCAGCGCAACCCA
GAATCCGGATGCAAGCGCAAGCGCCAATGAACATTATCGTGCAGTTGGTGGTCTGCTGAATATGCAGCTGGCCTATC
GTACCTATCCGGGTCCGTTTGATCCGCTGGCACTGTTTACCCAGCGTTATGCACAGGATCATCATGCATTTTGGCTG
GATAGCGAAAAAAGCGAACGTCCGAATGCACGTTATAGCATTATGGGTAGTGGTCAGGCACAGGGTAGCATTCGTCT
GACATATGATGTTAATAGCGAAAGTCTGACCCTGGCAGGTCCGAAAGGTAGCCGTATTGTGACCGGTGATTTTTTCA
CCCTGTTTAGCCAGATTGTTGAAAGCGTTAATGTTGCAGTTCCGCAGTATCTGCCGTTTGAGTTTAAAGGTGGTTTT
GTGGGTTATATGGGCTATGAACTGAAAGCACTGACCGGTGGTAATAAAGTGTATCGTAGCGGTCAGCCGGATGCAGG
TTTTATGTTTGCACCGCATTTTTTTGTGTTCGATCATCACGATCAGACCGTGTATGAGTGCATGATTAGCGCAACCG
GTCAGAGTCCGCAGTGGCCTCAGCTGCTGACCAGCATGACCACACTGAATAATGCAACCGATCGTCGTCCGTTTGTT
CCTGGTGCAGTTGATGAACTGGAACTGAGCCTGGAAGATGGTCCGGATGATTATATTCGTAAAGTTAAACAGAGCCT
GCAGTATATTACCGATGGTGAAAGCTATGAAATCTGTCTGACCAATCGTGCACGTATGAGCTATAGCGGTGAACCGC
TGGCAGCATATCGTCGTATGCGTGAAGCATCACCGGTTCCGTATGGTGCATATCTGTGTTTTGATAGTTTTAGCGTT
CTGAGCGCAAGTCCGGAAACCTTTCTGCGTATTGATGAAGGTGGTCTGATTGAATCACGTCCGATTAAAGGCACCCG
TGCGCGTAGCAAAGATCCGAGCGAAGATCAGCGTCTGCGTAGCGATCTGCAGGCAAGCACCAAAGATCGTGCAGAAA
ATCTGATGATTGTTGATCTGGTTCGCCATGATCTGAATCAGGTTTGTCGTAGTGGTAGCGTTCATGTTCCGCATATT
TTTGCAGTTGAAAGCTTTAGCAGCGTTCATCAGCTGGTTAGCACCGTTCGTGGTCATCTGCGTAATGATATTAGCAC
CATGGAAGCAATTCGTGCCTGTTTTCCTGGTGGTAGTATGACAGGTGCACCGAAAAAACGTACCATGGAAATTATTG
ATGGTCTGGAAACCTGTGCACGTGGTGTTTATAGTGGTGCATTAGGTTGGATTAGCTTTAGCGGTAGTGCAGAACTG
AGCATTGTTATTCGTACCGCAGTGCTGCATAAACAGCAGGCAGAATTTGGTATTGGTGGTGCAATTGTTGCACATAG
CGATCCGAATGAAGAGCTTGAAGAGACGCTGGTCAAAGCCAGCGTGCCTTATTACAGTTTTTACGCAGGGAGCGAAA
AA**TGA**ATAGTAATCAGTAAGGAGATAAAGA**ATG**AACATGACCGAACATCGTCATATGAGCCCGACCACACCGAGCGC
AATTCTGCAGCCGCAGCGTGATCAGCTGGATCGTATTAACAATCATCTGGTTGATCTGCTGGGTGAACGTATGAGCG
TTTGTATGGATATTGCAGAACTGAAAGCAGCACATGATATTCCGATGATGCAGCCTCAGCGCATTGTTCAGGTTCTG
GATCAGCTGAAAGATAAAAGCAGTACCGTTGGTCTGCGTCCGGATTATGTTCAGAGCGTTTTTAAACTGATCATCGA
GGAAACCTGCATCCAAGAAGAACAGCTGATTCAGCGTCGTCGTAATCAGGGTCAGCGTAGC**TAA**CGTAAATATAAGG
AGGTCAAAC**ATG**AATACCAATACCGTTGTGGTTTTAGGTGGTGCAGGTCTGATTGGTAGCATGATTAGCCGTATTCT
GAAACAGTATGGTTATTTTGTTCGTGTGGTTGATCGTCGTCCGGCAGAATTTGAATGTGAATATCATGAAATGGACG
TGACCAAACCGTTTAATGATACCGGTGCAGTTTTTCGTAATGCAACCGCAGTTGTTTTTGCACTGCCGGAAAGCGTT
GCAGTTAGCGCAATTCCGTGGGTTACCACCTTTCTGAGCAGCGAAGTTGTTCTGATTCCGACCTGTAGCGTTCAGGG
TCCGTTCTATAAAGCACTGAAAGCAGCAGCACCGCGTCAGCCGTTTGTTGGTGTTAATCCGATGTTTAGCCCGAAAC
TGAGCGTGCAGGGTCGTAGCGTTGCCGTTTGTGTTGAAGATACCCAGGCAGCACAGACCTTTATTGAACGTCATCTG
ATGGAAGCCGGTATGAAAATTCGTCGTATGACCCCGAGCGCACATGATGAACTGATGGCACTGTGTCAGGCACTGCC
GCATGCAGCAATTTTAGGTTTTGGTATGGCACTGGCAAAAAGCAGCGTTGATATGGATATTGTTGCAGAAGTTATGC
CTCCGCCTATGCGTACCATGATGGCCCTGCTGAGTCGTATTCTGGTTAATCCGCCTGAAGTTTATTGGGATATTCAG
CTGGAAAATGATCAGGCAACCGCACAGCGTGATGCACTGGTTCATGGTCTGGAACGTCTGCAAGAAAATATTGTGGA
ACAGGATTATGAGCGCTTCAAAAGCGATCTGCAGAGCGTTAGCACCGCACTGGGTAAACGTCTGAATGCGGGTGCAG
TTGATTGTCAGCACCTGTTTAGCCTGCTGAAT**TAA**CCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGG
CCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTC
TGCGTTTATA

>J6-BBa_J23117-At-pal2-Ec-tyrB-ECK120033736

CTGCACGAGTTCGCTGTCGAGACAAGTCTCTTAGCGACGTATTACGAAGATCACATAGTCAGATGAAGCTATAGAGC
ACGACGCTAACGATTACGTCACGCTTGACACAACAGTTTCGCTACCTAGTGCTCGCGCGACTGCGACGTTGTCCTTC
TAGTCGCCCATGACTCTTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGC**A**ACTGGTAATTTGAGGAGGTAATTT**A**
TGGATCAAATCGAAGCAATGTTGTGCGGCGGAGGAGAGAAGACAAAAGTGGCGGTTACTACGAAGACTTTGGCAGAT
CCATTGAATTGGGGTTTAGCAGCGGATCAAATGAAAGGAAGTCATTTAGATGAAGTGAAGAAGATGGTCGAAGAGTA
TCGTAGACCAGTCGTGAATCTTGGCGGAGAAACACTGACGATCGGACAAGTTGCTGCCATCTCCACCGTAGGAGGCA
GCGTTAAGGTTGAGTTAGCGGAGACTTCAAGAGCCGGTGTGAAAGCTAGCAGTGATTGGGTTATGGAGAGCATGAAC

AAAGGTACTGACAGTTACGGAGTCACCACCGGCTTTGGTGCTACTTCTCACCGGAGAACCAAAAACGGCACCGCATT
ACAAACAGAACTCATTAGATTTTTGAACGCCGGAATATTCGGAAACACGAAGGAGACATGTCACACACTGCCGCAAT
CCGCCACAAGAGCCGCCATGCTCGTCAGAGTCAACACTCTTCTCCAAGGATACTCCGGGATCCGATTCGAGATCCTC
GAAGCGATTACAAGTCTCCTCAACCACAACATCTCTCCGTCACTACCTCTCCGTGGAACCATTACCGCCTCCGGCGA
TCTCGTTCCTCTCTCTTACATCGCCGGACTTCTCACCGGCCGTCCTAATTCCAAAGCCACCGGTCCCGACGGTGAAT
CGCTAACCGCGAAAGAAGCTTTTGAGAAAGCCGGAATCAGTACTGGATTCTTCGATTTACAACCTAAGGAAGGTTTA
GCTCTCGTTAATGGCACGGCGGTTGGATCTGGAATGGCGTCGATGGTTCTATTCGAAGCGAATGTCCAAGCGGTGTT
AGCGGAGGTTTTATCAGCGATCTTCGCGGAGGTTATGAGCGGGAAACCTGAGTTTACCGATCATCTGACTCATCGTT
TAAAACATCATCCCGGACAAATCGAAGCGGCGGCGATAATGGAGCACATACTCGACGGAAGCTCATACATGAAATTA
GCTCAAAAGGTTCACGAGATGGATCCATTGCAGAAACCAAAACAAGATCGTTACGCTCTTCGTACATCTCCTCAATG
GCTAGGTCCTCAAATTGAAGTAATCCGTCAAGCTACGAAATCGATAGAGCGTGAAATCAACTCCGTTAACGATAATC
CGTTGATCGATGTTTCGAGGAACAAGGCGATTCACGGTGGTAACTTCCAAGGAACACCAATCGGAGTTTCTATGGAT
AACACGAGATTGGCGATTGCTGCGATTGGGAAGCTAATGTTTGCTCAATTCTCTGAGCTTGTTAATGATTTCTACAA
CAATGGACTTCCTTCGAATCTAACTGCTTCGAGTAATCCAAGTTTGGATTATGGATTCAAAGGAGCAGAGATTGCTA
TGGCTTCTTATTGTTCTGAGCTTCAATACTTGGCTAATCCAGTCACAAGCCATGTTCAATCAGCTGAGCAACATAAT
CAAGATGTGAACTCTCTTGGTTTGATCTCGTCTCGTAAAACATCTGAAGCTGTGGATATTCTTAAGCTAATGTCAAC
AACGTTCCTTGTGGGGATATGTCAAGCTGTTGATTTGAGACATTTGGAGGAGAATCTGAGACAAACTGTGAAGAACA
CAGTTTCTCAAGTTGCTAAGAAAGTGTTAACCACTGGAATCAACGGTGAGTTACATCCGTCAAGGTTTTGCGAGAAG
GACTTGCTTAAGGTTGTTGATCGTGAGCAAGTGTTCACGTATGTGGATGATCCTTGTAGCGCTACGTACCCGTTGAT
GCAGAGACTAAGACAAGTTATTGTTGATCACGCTTTGTCCAACGGTGAGACTGAGAAGAATGCAGTGACTTCGATCT
TTCAAAAGATTGGAGCTTTTGAAGAGGAGCTTAAGGCTGTGCTTCCAAAGGAAGTTGAAGCGGCTAGAGCGGCTTAT
GGGAATGGAACTGCGCCGATTCCTAACCGGATTAAGGAATGTAGGTCGTATCCGTTGTATAGGTTCGTGAGGGAAGA
GCTTGGAACGAAGTTGTTGACTGGAGAAAAGGTTGTGTCTCCGGGAGAGGAGTTTGATAAGGTCTTCACTGCTATGT
GTGAAGGTAAACTTATTGATCCGTTGATGGATTGTCTCAAGGAATGGAACGGAGCTCCGATTCCGATTTGC**TAA**AAA
GAGGAGAAATTACAT**GTG**TTTCAAAAAGTTGACGCCTACGCTGGCGACCCGATTCTTACGCTTATGGAGCGTTTTAA
AGAAGACCCTCGCAGCGACAAAGTGAATTTAAGTATCGGTCTGTACTACAACGAAGACGGAATTATTCCACAACTGC
AAGCCGTGGCGGAGGCGGAAGCGCGCCTGAATGCGCAGCCTCATGGCGCTTCGCTTTATTTACCGATGGAAGGGCTT
AACTGCTATCGCCATGCCATTGCGCCGCTGCTGTTTGGTGCGGACCATCCGGTACTGAAACAACAGCGCGTAGCAAC
CATTCAAACCCTTGGCGGCTCCGGGGCATTGAAAGTGGGCGCGGATTTCCTGAAACGCTACTTCCCGGAATCAGGCG
TCTGGGTCAGCGATCCTACCTGGGAAAACCACGTAGCAATATTCGCCGGGGCTGGATTCGAAGTGAGTACTTACCCC
TGGTATGACGAAGCGACTAACGGCGTGCGCTTTAATGACCTGTTGGCGACGCTGAAAACATTACCTGCCCGCAGTAT
TGTGTTGCTGCATCCATGTTGCCACAACCCAACGGGTGCCGATCTCACTAATGATCAGTGGGATGCGGTGATTGAAA
TTCTCAAAGCCCGCGAGCTTATTCCATTCCTCGATATTGCCTATCAAGGATTTGGTGCCGGTATGGAAGAGGATGCC
TACGCTATTCGCGCCATTGCCAGCGCTGGATTACCCGCTCTGGTGAGCAATTCGTTCTCGAAAATTTTCTCCCTTTA
CGGCGAGCGCGTCGGCGGACTTTCTGTTATGTGTGAAGATGCCGAAGCCGCTGGCCGCGTACTGGGGCAATTGAAAG
CAACAGTTCGCCGCAACTACTCCAGCCCGCCGAATTTTGGTGCGCAGGTGGTGGCTGCAGTGCTGAATGACGAGGCA
TTGAAAGCCAGCTGGCTGGCGGAAGTAGAAGAGATGCGTACTCGCATTCTGGCAATGCGTCAGGAATTGGTGAAGGT
ATTAAGCACAGAGATGCCAGAACGCAATTTCGATTATCTGCTTAATCAGCGCGGCATGTTCAGTTATACCGGTTTAA
GTGCCGCTCAGGTTGACCGACTACGTGAAGAATTTGGTGTCTATCTCATCGCCAGCGGTCGCATGTGTGTCGCCGGG
TTAAATACGGCAAATGTACAACGTGTGGCAAAGGCGTTTGCTGCGGTGATG**TAA**GGATCCAAACTCGAGTAAGGATC
TGTGCTTTTTTTAACGCATGAGAAAGCCCCCGGAAGATCACCTTCCGGGGGCTTTTTTATTGCGC

>J6-BBa_J23117-Rg-pal-ECK120033736
CTGCACGAGTTCGCTGTCGAGACAAGTCTCTTAGCGACGTATTACGAAGATCACATAGTCAGATGAAGCTATAGAGC
ACGACGCTAACGATTACGTCACGCTTGACACAACAGTTTCGCTACCTAGTGCTCGCGCGACTGCGACGTTGTCCTTC
TAGTCGCCCATGACTCTTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGC**A**TCGGGCTGCCCCAGCCAATCACACC
ATACCATAAGGAGCATTTTTTT**ATG**GCGCCAAGCGTCGACTCCATCGCCACGTCCGTGGCCAACAGCTTGAGCAACGG
ACTGGCTGGCGATCTGCGGAAGAAAACCAGCGGCGCGGGAAGCCTGCTCCCGACAACCGAGACTACCCAGATCGACA
TCGTCGAGCGGATTCTCGCCGATGCCGGTGCTACTGACCAGATCAAGTTGGATGGCTATACCTTAACGCTCGGCGAT

GTGGTGGGCGCCGCGCGGCGAGGGAGGACCGTTAAAGTCGCCGACTCGCCGCAGATTCGCGAAAAGATCGATGCGTC
AGTGGAGTTTCTGCGCACTCAACTGGACAACAGTGTGTACGGCGTGACCACCGGCTTCGGCGGCTCGGCCGACACCC
GCACCGAAGACGCCATTTCGCTGCAAAAGGCGCTGCTGGAGCACCAGCTGTGCGGCGTTTTGCCAACCTCGATGGAC
GGCTTCGCCTTGGGGCGTGGTCTGGAGAACTCGTTGCCACTGGAAGTCGTGCGGGGCGCCATGACCATCAGAGTCAA
TAGCCTGACCCGCGGCCATTCGGCTGTCCGTATCGTCGTCCTGGAAGCCCTGACCAACTTCTTGAACCACGGCATCA
CCCCGATCGTGCCGCTGCGCGGGACCATCAGCGCGTCGGGCGACTTGAGCCCGCTCAGTTACATCGCTGCGTCGATC
ACCGGGCATCCGGACAGCAAGGTGCATGTAGACGGGCAAATCATGAGCGCCCAAGAAGCTATTGCACTGAAAGGCCT
GCAACCCGTCGTTCTTGGCCCGAAGGAAGGCTTGGGCCTGGTCAACGGCACCGCCGTGTCGGCCTCCATGGCCACGC
TCGCCCTGACCGACGCCCACGTGCTCAGCCTGCTCGCGCAGGCCAATACCGCACTGACGGTGGAAGCAATGGTGGGC
CATGCCGGCTCATTCCACCCGTTCCTGCATGATGTGACTCGCCCGCACCCCACCCAGATCGAGGTGGCCCGTAACAT
TCGCACGCTGCTGGAGGGCAGCAAGTATGCCGTACACCACGAAACCGAAGTAAAAGTGAAAGACGACGAGGGTATCC
TTCGCCAGGACCGCTACCCGCTTCGCTGCTCGCCTCAGTGGCTGGGTCCGCTGGTGAGCGACATGATCCACGCGCAC
AGCGTGCTGTCACTGGAAGCAGGCCAGAGTACGACCGACAACCCGCTGATTGACCTAGAAAACAAGATGACCCATCA
TGGTGGTGCCTTCATGGCCAGCAGCGTGGGTAACACCATGGAGAAGACCCGGCTGGCCGTAGCACTGATGGGCAAGG
TGTCGTTCACCCAATTGACAGAAATGCTGAACGCTGGCATGAACCGCGCCCTGCCCAGCTGCCTGGCAGCCGAGGAC
CCGTCACTGAGTTACCACTGCAAGGGCCTTGATATCGCCGCCGCAGCCTACACCAGCGAGCTGGGCCACCTGGCCAA
TCCTGTCAGCACGCATGTGCAGCCGGCCGAGATGGGTAACCAGGCAATCAATTCCCTGGCGCTCATATCTGCCCGCC
GTACAGCCGAGGCCAACGACGTGCTGTCTTTGCTGCTTGCAACTCACTTGTACTGTGTACTGCAAGCCGTTGATCTG
CGTGCGATGGAATTTGAACACACCAAGGAGTTCGAGCCGATGGTCACCGATTTGCTAAAACAGCACTTCGGCGCCCT
GGCCACAGCGGACGTGGAGGACAAGGTGCGCAAGAGCATCTACAAGCGCTTACAGCAGAACAATTCCTACGACCTGG
AGCAACGCTGGCACGACACCTTCAGCGTAGCGACCGGTGCGGTCGTCGAGGCGCTGGCCGGCAACGAAGTGTCTCTG
GCCTCCCTGAACGCCTGGAAGGTGGCGTGCGCCGAAAAAGCTATCGCCCTGACCCGTACCGTGCGCGACAGCTTCTG
GGCTGCGCCTAGCTCGGCCTCGCCGGCGCTCAAGTACCTGTCACCACGGACCCGCATCCTGTACAGCTTTGTCCGCG
AAGATGTGGGCGTCAAGGCCAGGCGCGGTGATGTTTATCTCGGCAAGCAGGAAGTGACGATCGGTACGAACGTGTCC
CGTATCTATGAGGCCATCAAAGACGGGCGCATTGCACCCGTGCTGGTCAAAATGATGGCA**TAA**AAACTCGAGTAAGG
ATCTAACGCATGAGAAAGCCCCCGGAAGATCACCTTCCGGGGGCTTTTTTATTGCGC

>pBBR1_oriV-pBBR1_REP

CGCCCTACGGGCTTGCTCTCCGGGCTTCGCCCTGCGCGGTCGCTGCGCTCCCTTGCCAGCCCGTGGATATGTGGACG
ATGGCCGCGAGCGGCCACCGGCTGGCTCGCTTCGCTCGGCCCGTGGACAACCCTGCTGGACAAGCTGATGGACAGGC
TGCGCCTGCCCACGAGCTTGACCACAGGGATTGCCCACCGGCTACCCAGCCTTCGACCACATACCCACCGGCTCCAA
CTGCGCGGCCTGCGGCCTTGCCCCATCAATTTTTTTAATTTTCTCTGGGGAAAAGCCTCCGGCCTGCGGCCTGCGCG
CTTCGCTTGCCGGTTGGACACCAAGTGGAAGGCGGGTCAAGGCTCGCGCAGCGACCGCGCAGCGGCTTGGCCTTGAC
GCGCCTGGAACGACCCAAGCCTATGCGAGTGGGGGCAGTCGAAGGGCGAAGCCCGCCCGCCTGCCCCCCGAGCCTCA
CGGCGGCGAGTGCGGGGGTTCCAAGGGGGCAGCGCCACCTTGGGCAAGGCCGAAGGCCGCGCAGTCGATCAACAAGC
CCCGGAGGGGCCACTTTTTGCCGGAGGGGGAGCCGCGCCGAAGGCGTGGGGGAACCCCGCAGGGGTGCCCTTCTTTG
GGCACCAAAGAACTAGATATAGGGCGAAATGCGAAAGACTTAAAAATCAACAACTTAAAAAAGGGGGGTACGCAACA
GCTCATTGCGGCACCCCCCGCAATAGCTCATTGCGTAGGTTAAAGAAAATCTGTAATTGACTGCCACTTTTACGCAA
CGCATAATTGTTGTCGCGCTGCCGAAAAGTTGCAGCTGATTGCGCATGGTGCCGCAACCGTGCGGCACCCCTACCGC
ATGGAGATAAGC**ATG**GCCACGCAGTCCAGAGAAATCGGCATTCAAGCCAAGAACAAGCCCGGTCACTGGGTGCAAAC
GGAACGCAAAGCGCATGAGGCGTGGGCCGGGCTTATTGCGAGGAAACCCACGGCGGCAATGCTGCTGCATCACCTCG
TGGCGCAGATGGGCCACCAGAACGCCGTGGTGGTCAGCCAGAAGACACTTTCCAAGCTCATCGGACGTTCTTTGCGG
ACGGTCCAATACGCAGTCAAGGACTTGGTGGCCGAGCGCTGGATCTCCGTCGTGAAGCTCAACGGCCCCGGCACCGT
GTCGGCCTACGTGGTCAATGACCGCGTGGCGTGGGGCCAGCCCCGCGACCAGTTGCGCCTGTCGGTGTTCAGTGCCG
CCGTGGTGGTTGATCACGACGACCAGGACGAATCGCTGTTGGGGCATGGCGACCTGCGCCGCATCCCGACCCTGTAT
CCGGGCGAGCAGCAACTACCGACCGGCCCCGGCGAGGAGCCGCCCAGCCAGCCCGGCATTCCGGGCATGGAACCAGA
CCTGCCAGCCTTGACCGAAACGGAGGAATGGGAACGGCGCGGGCAGCAGCGCCTGCCGATGCCCGATGAGCCGTGTT
TTCTGGACGATGGCGAGCCGTTGGAGCCGCCGACACGGGTCACGCTGCCGCGCCGG**TAG**

>pBBR1_REP(G159S)

ATGGCCACGCAGTCCAGAGAAATCGGCATTCAAGCCAAGAACAAGCCCGGTCACTGGGTGCAAACGGAACGCAAAGC
GCATGAGGCGTGGGCCGGGCTTATTGCGAGGAAACCCACGGCGGCAATGCTGCTGCATCACCTCGTGGCGCAGATGG
GCCACCAGAACGCCGTGGTGGTCAGCCAGAAGACACTTTCCAAGCTCATCGGACGTTCTTTGCGGACGGTCCAATAC
GCAGTCAAGGACTTGGTGGCCGAGCGCTGGATCTCCGTCGTGAAGCTCAACGGCCCCGGCACCGTGTCGGCCTACGT
GGTCAATGACCGCGTGGCGTGGGGCCAGCCCCGCGACCAGTTGCGCCTGTCGGTGTTCAGTGCCGCCGTGGTGGTTG
ATCACGACGACCAGGACGAATCGCTGTTGGGGCATGGCGACCTGCGCCGCATCCCGACCCTGTATCCGGGCGAGCAG
CAACTACCGACCGGCCCCGGCGAGGAGCCGCCCAGCCAGCCCGGCATTCCGGGCATGGAACCAGACCTGCCAGCCTT
GACCGAAACGGAGGAATGGGAACGGCGCGGGCAGCAGCGCCTGCCGATGCCCGATGAGCCGTGTTTTCTGGACGATG
GCGAGCCGTTGGAGCCGCCGACACGGGT**A**ACGCTGCCGCGCCGG**TAG**

>pp1_HR1-J3-BBa_J23111-mRFP1-dblT-pp1_HR2

ATGACCGACCTGATCGAAGTGAAGACGGCAGACCTGGTGGGCGAGGCGCTTGGGTGGGCCGTGGGCACGGCGGAAGG
CCTGGACCTGTTCATGGCGCCGCCGGAGTACGGCAACCCACACCGAGTGTTCGCCCGCTACCAGGGCCAGGCCATCG
AGCACACCAAGCGCTTCAACCCGTGGGAAGACTGGGCGGTTGGCGGGCCGATCATGCAGAAGCACAACGTCAGCCTG
CACTGCCCGCAGCCAGAGTGGGACTACTGGGCAGCCTGGATAACCGATAACGGCAAGGACGTCGCCCAGGGCGCTGA
TCTGCCGTTGCCGGCGGCGTGCCGGGCCATAGTCGCCCACCAGCTCGGCGATACCGTCCAGGTGCCGAAGGAGCTGA
TGCCATGACCGTGATCCTTCCCCTCGCCTACATGGCCTACCTGATCTACAGGGGGCTTCTCGGTGAGGGAGGCGCCT
GCAAGCAAAGGGCACGACATGACCTGACGACAGCACGGCAAAAAACAAACTCGAAAGGATCATCCACAAGATCAAGC
GCTGCCTGGCGCTATTCAAAAGCTCGAATGAATATGAGAGAGTCTAGGCCCACCCGCCGATTACGAAGGTCTTCGCT
CGGAGCACACCCCAGACCAAGGCTCGACTCATAGTTTCGCTTGGTCTGGTGCTGTAAGCCTCTTCTACAATTCGGTC
CCCGCTTTTGGAGTACACCCCGATGAAGAGCTGCGTTTCGCCTGTCCGCGAAAGACGGGTTTGCACGTCGATACTCC
TGCCGTCCTCAAGGATTTCGTCGTGATGACGAAGGTGAAGCGCTGGGTCTGCCCAGGTCCAGAATTTTTCGCCGCGA
CATCTCATATCAATCTCCTCTTACTTATCCCAGTAGGCGCGGTAAAGAGAGGGATAGATATCCATTTCGCTTAAATG
CGACCGGTGGAAAATGATCGGCCCTAATCCTTGCTGATAGATATCAGCGGGACAGCGCCAGTAGAGAACCGAGCCCA
GCATGGCAATTCCGACGTCAGCATTTGCGATCATTCACGCAGCGCTTATTCAGTTGCTCACTGCGATGTCATAATCA
TCGCTACGAGCTGTGAAAGATGCATAAAGCTCGTACGACGCGTTCGCTCGTCTCCTCACTTCTCCTACGGAGCGTTC
TGGACACAACGTCGTCTTGAAGTTGCGATTATAGATTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGC**G**AATTCA
TTAAAGAGGAGAAAGGTACC**ATG**GCGAGTAGCGAAGACGTTATCAAAGAGTTCATGCGTTTCAAAGTTCGTATGGAA
GGTTCCGTTAACGGTCACGAGTTCGAAATCGAAGGTGAAGGTGAAGGTCGTCCGTACGAAGGTACCCAGACCGCTAA
ACTGAAAGTTACCAAAGGTGGTCCGCTGCCGTTCGCTTGGGACATCCTGTCCCCGCAGTTCCAGTACGGTTCCAAAG
CTTACGTTAAACACCCGGCTGACATCCCGGACTACCTGAAACTGTCCTTCCCGGAAGGTTTCAAATGGGAACGTGTT
ATGAACTTCGAAGACGGTGGTGTTGTTACCGTTACCCAGGACTCCTCCCTGCAAGACGGTGAGTTCATCTACAAAGT
TAAACTGCGTGGTACCAACTTCCCGTCCGACGGTCCGGTTATGCAGAAAAAAACCATGGGTTGGGAAGCTTCCACCG
AACGTATGTACCCGGAAGACGGTGCTCTGAAAGGTGAAATCAAAATGCGTCTGAAACTGAAAGACGGTGGTCACTAC
GACGCTGAAGTTAAAACCACCTACATGGCTAAAAAACCGGTTCAGCTGCCGGGTGCTTACAAAACCGACATCAAACT
GGACATCACCTCCCACAACGAAGACTACACCATCGTTGAACAGTACGAACGTGCTGAAGGTCGTCACTCCACCGGTG
CT**TAA**AAACTCGAGTAAGGATCTCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTA
TCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATATC
CGCCATTACGATCTGACTTGCCACTTTGGCCGTTATCCCTATCTCTGGTCATACGCACCTCTCAATGCTCAGAAGTC
CTGAGCTAAATTGAGTGTGCGCCTACAGAGACTAGAGAAGATTTATATCAGACCGACGGGTACAAACGAAAAAATCA
ATAACCGGCTCCTGTCAGCACCTCGCAGAGAAAACAGATCGCAGATATTCAGCATGCGTAGCTGATGCTCGACCGAG
ATCGCACCTGACATGCACCGCTGCGCGGCACATCACACTGAGTAGAGAGCGTTCCGCGACGCCGATGGCGTGAAGCG
CGAGCTGGAAAACGTCGTGCCGTTCAAGGGCGCCTAACCCCTCCCCATACAACTCAAGCCCGCCGACATGCGCGGGC
GAGGATTACCTATGTCCAATTTCCTGACTCGCTGGCTCAAGCGCAAGAAGAAGCCCGGGCCGCGTCCAACACTGGCA
CCAACTGGATTTGCTCGCGGCCACAGCCCGGCAGTCGGCAGGCTAGATCCGATGCTTGATCCGCTCAACCCGTTGAG
CCCCGTCAGCCCGTTGCATCCCGCCTACCAGGCCGACAGCTACGAACCACCGCGCAGCACCAGCAGTTCCTGCTCCA
GCCGTGATTACAGCAGCTACGACTGCGGCAGCAGCTACTCGTCGAGCGACAGCAGCAGTTCCAGCGATAGCGGATCC
AGCTCCAGCAGCTGCGACTGACCACCAACCTGCCGCCACCGGCGGCGTGGAGACCATCCCATGGAAACCGAAATCCT

TTCGGACGAAGAGCTGGTGGCGATCACCGGCTACAAACCCCGGGCGTGGCAGCGCCGTTGGCTAACAGAAAAAGGCT
GGCACTTCGTCGAGAGCCGCGGCGGCCGGCCACTGGTTGGCCGCCAGTACGCCCGCCAGAAGCTCAGCGGCGTGGTG
ATCGACACCTTGCCGGTCGCACCAGCCCCACCACCAACGCCCGCCTGGACCCCTGATTTTTCCCGAGTGAAGTGA

>pp2_HR1-J6-BBa_J23105-Rg-pal-ECK120033736-pp2_HR2

ACCAGGATGAATACCTTAAGGACGCCACCGGTAACCGGCGTTATTGGCCGGTCGCTTGCGTCAAGGTGGACCTTGAA
GCATTGCGTCGCGCTCGTGACCAGCTGTGGGCTGAGGCCATGTTCTGCTACCAGGCCGGTGATATCTGGTGGGTGAC
CCGTGAGGAGGAAGAACTGTTCACTGCAGAGCAGGAAGAGCGCTTCGTGGTAGATGAATGGGAGGGGCCGATCCTGA
AATGGTTGGAGGAATCCCAGGCCGGCGAGACGGTCACCGGAAGCGAAGTGTTGGGGCAGGCATTGAACCTTGACCCT
GGCCACTGGGGCAAGCCTGAGCAGATGCGGGTGGGATCGATCATGCACCGCCTAGGTTGGCGGCGTCGCAGGCTGGC
TGCGCTGCCGAAGAGCGGTAAGCGCCCTTGGGCATATCAGAAGCCTGATGGTTGGGGGCGCAGCGCCTTGGAGCAGT
CCACGCAGCCGAAGGAGGAGTGCTTTTGATCAAACACATAGATGAGATGCTGAAACTGTGGGCTCAAGAGCTCCATG
CGCCGGAGCCCTGTCATTCGGCGGGCGGTGTTGGTAGCATGCTCGGCCTGTTGATCGAGTGCAAGGGTGACCTTGTG
CGTGGCACCCGAGGCAGCAAGGTGCTACTGGACGAGTCTGCGGACATCGAGATCATTGTGAATAAGCATCTGGCACC
GGAGCTCTACCTGGTGGTTCGTGAGCACTACTGTAACGCCGACAGCGAGCTGTACCAAAAGTACCGGCACTGTGGGT
GTAGCCGGGATACCTATTACAAGCGCTTACATGAGGCGCACGTCTGTATCGCAGGCTTGCTCTTGGGGCGAGCGGCA
TGATTCGCCATCCACGGTCCTACCGTCCCGCTGCCGTCCTGCCGCGTTTGATGCAGGTCAGCCTAGCTCAAGCCCGC
GCAGGTCGCGGGCTGTCCCACCGTCCCACCGTACACACGAAGGCGCGCACATAGGCGTGTGCAGCGCATCACGCGCA
TGCATCCTGCACGAGTTCGCTGTCGAGACAAGTCTCTTAGCGACGTATTACGAAGATCACATAGTCAGATGAAGCTA
TAGAGCACGACGCTAACGATTACGTCACGCTTGACACAACAGTTTCGCTACCTAGTGCTCGCGCGACTGCGACGTTG
TCCTTCTAGTCGCCCATGACTCTTTACGGCTAGCTCAGTCCTAGGTACTATGCTAGCATCGGGCTGCCCCAGCCAAT
CACACCATACCATAAGGAGCATTTTTTTATGGCGCCAAGCGTCGACTCCATCGCCACGTCCGTGGCCAACAGCTTGAG
CAACGGACTGGCTGGCGATCTGCGGAAGAAAACCAGCGGCGCGGGAAGCCTGCTCCCGACAACCGAGACTACCCAGA
TCGACATCGTCGAGCGGATTCTCGCCGATGCCGGTGCTACTGACCAGATCAAGTTGGATGGCTATACCTTAACGCTC
GGCGATGTGGTGGGCGCCGCGCGGCGAGGGAGGACCGTTAAAGTCGCCGACTCGCCGCAGATTCGCGAAAAGATCGA
TGCGTCAGTGGAGTTTCTGCGCACTCAACTGGACAACAGTGTGTACGGCGTGACCACCGGCTTCGGCGGCTCGGCCG
ACACCCGCACCGAAGACGCCATTTCGCTGCAAAAGGCGCTGCTGGAGCACCAGCTGTGCGGCGTTTTGCCAACCTCG
ATGGACGGCTTCGCCTTGGGGCGTGGTCTGGAGAACTCGTTGCCACTGGAAGTCGTGCGGGGCGCCATGACCATCAG
AGTCAATAGCCTGACCCGCGGCCATTCGGCTGTCCGTATCGTCGTCCTGGAAGCCCTGACCAACTTCTTGAACCACG
GCATCACCCCGATCGTGCCGCTGCGCGGGACCATCAGCGCGTCGGGCGACTTGAGCCCGCTCAGTTACATCGCTGCG
TCGATCACCGGGCATCCGGACAGCAAGGTGCATGTAGACGGGCAAATCATGAGCGCCCAAGAAGCTATTGCACTGAA
AGGCCTGCAACCCGTCGTTCTTGGCCCGAAGGAAGGCTTGGGCCTGGTCAACGGCACCGCCGTGTCGGCCTCCATGG
CCACGCTCGCCCTGACCGACGCCCACGTGCTCAGCCTGCTCGCGCAGGCCAATACCGCACTGACGGTGGAAGCAATG
GTGGGCCATGCCGGCTCATTCCACCCGTTCCTGCATGATGTGACTCGCCCGCACCCCACCCAGATCGAGGTGGCCCG
TAACATTCGCACGCTGCTGGAGGGCAGCAAGTATGCCGTACACCACGAAACCGAAGTAAAAGTGAAAGACGACGAGG
GTATCCTTCGCCAGGACCGCTACCCGCTTCGCTGCTCGCCTCAGTGGCTGGGTCCGCTGGTGAGCGACATGATCCAC
GCGCACAGCGTGCTGTCACTGGAAGCAGGCCAGAGTACGACCGACAACCCGCTGATTGACCTAGAAAACAAGATGAC
CCATCATGGTGGTGCCTTCATGGCCAGCAGCGTGGGTAACACCATGGAGAAGACCCGGCTGGCCGTAGCACTGATGG
GCAAGGTGTCGTTCACCCAATTGACAGAAATGCTGAACGCTGGCATGAACCGCGCCCTGCCCAGCTGCCTGGCAGCC
GAGGACCCGTCACTGAGTTACCACTGCAAGGGCCTTGATATCGCCGCCGCAGCCTACACCAGCGAGCTGGGCCACCT
GGCCAATCCTGTCAGCACGCATGTGCAGCCGGCCGAGATGGGTAACCAGGCAATCAATTCCCTGGCGCTCATATCTG
CCCGCCGTACAGCCGAGGCCAACGACGTGCTGTCTTTGCTGCTTGCAACTCACTTGTACTGTGTACTGCAAGCCGTT
GATCTGCGTGCGATGGAATTTGAACACACCAAGGAGTTCGAGCCGATGGTCACCGATTTGCTAAAACAGCACTTCGG
CGCCCTGGCCACAGCGGACGTGGAGGACAAGGTGCGCAAGAGCATCTACAAGCGCTTACAGCAGAACAATTCCTACG
ACCTGGAGCAACGCTGGCACGACACCTTCAGCGTAGCGACCGGTGCGGTCGTCGAGGCGCTGGCCGGCAACGAAGTG
TCTCTGGCCTCCCTGAACGCCTGGAAGGTGGCGTGCGCCGAAAAAGCTATCGCCCTGACCCGTACCGTGCGCGACAG
CTTCTGGGCTGCGCCTAGCTCGGCCTCGCCGGCGCTCAAGTACCTGTCACCACGGACCCGCATCCTGTACAGCTTTG
TCCGCGAAGATGTGGGCGTCAAGGCCAGGCGCGGTGATGTTTATCTCGGCAAGCAGGAAGTGACGATCGGTACGAAC
GTGTCCCGTATCTATGAGGCCATCAAAGACGGGCGCATTGCACCCGTGCTGGTCAAAATGATGGCATAAAACTCGA

GTAAGGATCTAACGCATGAGAAAGCCCCCGGAAGATCACCTTCCGGGGGCTTTTTTATTGCGCGGATCCTTACTCCG
TCGACAGTACTAAGGTGTATTCCCCCGGCATTCAACAGACATTCCCCGGACGTTTACCACTTATTGCCGGGTGGCAT
TAAAACTCGCTTGCTGCCACCGGAATCGACCTGTAAAAAGTACCCATCTTCGAGACGTGCGGGCGCACAAAGCGGCC
CGCCAAACACTGAAAACCCCGGCCCTGGCGCCGGGGTTTTTGCCTTTGGGGGATTCGATGAACAGCGAGCAGCAAGC
GTTAGTTGAGATGCCAATCTGGTTGGTGATCTTCCTGTCCCTGGTCGGCGGGGTGTCAGGCGAGATGTGGCGGGCCG
ACATGGCCGGCGTTAGCGGCTGGTTCATTTTCCGCCAGGTGCTGCTGCGCTCCGGTGCCTGCGTCGTATGCGGACTG
TCGACCATCATGCTGCTGTACTCGGCGGGCATGTCGATGTGGTCGGCCAGTGCCATTGGTTGTCTCACTGCCACTGC
CGGTGCGGATGTGGCCATAGGGTTGTACAAGCGTTGGGTCGCCAAGCGGCTGGGCGTCTGCGATGTCACGTCCCGTA
GCGGCGAACCTGGACAGTGACCCGATCGCCAGCCTCGGTGGGGTCGGGGACCCTGGCGATATGGCCGGGTTACGGGG
CAGGAAACCCGCGGCTCTTCGCTAGCGGACAGTTCGCCAGCTTACTGAAATTCAACCTGTTGAAATTGAAAGGTTGT
TGGTTGAAATACCATTGAAATGGAGGGCTCATGACGGATTCGAACTTCTTGTCAAAGAGCGCCTTCGCTGCTCGCAT
AGGGAGATCACCCAGCTACATCACCTGGTTGAAAGACAACGGCCGCCTGGTGCTTTCACCCGATGGAAAATTGGTGG
ATGTGCTGGCCACCGAGGCCAAGATTCAGGAGACAGCTGATCCGGCCAAAGCAGCCGTCGCGGCTCGGCATGAAGAA
AACCGCATCGAGCGGGACGTCCGGGCCCACATCCAGCCTAGCGCCGACACACCTGCGGTGCAGCCAGCGGATCACGC
GCCGAGCGGA

Minimal Promoters

RNAP recognition sites (-35 and -10 elements) of the Anderson promoter series are underlined. TSSs are indicated in **bold** at the predicted site. For all experiments reported in this manuscript, the first base of the 5'-UTR is **G** (see 5'-UTR sequence below). The promoters shown below appear in an order of ascending basal expression level in *P. putida* (weakest first).

>BBa_J23109-5'-UTR
TTTACAGCTAGCTCAGTCCTAGGGACTGTGCTAGC**G**AATTCATTAAAGAGGAGAAAGGTACC

>BBa_J23113
CTGATGGCTAGCTCAGTCCTAGGGATTATGCTAGC**G**

>BBa_J23117
TTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGC**G**

>BBa_J23114
TTTATGGCTAGCTCAGTCCTAGGTACAATGCTAGC**G**

>BBa_J23115
TTTATAGCTAGCTCAGCCCTTGGTACAATGCTAGC**G**

>BBa_J23107
TTTACGGCTAGCTCAGCCCTAGGTATTATGCTAGC**G**

>BBa_J23105
TTTACGGCTAGCTCAGTCCTAGGTACTATGCTAGC**G**

>BBa_J23108
CTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGC**G**

>BBa_J23106

TTTACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCG

>pBBR1-REPp
TTGACTGCCACTTTTACGCAACGCATAATTGTTGTG

>BBa_J23110
TTTACGGCTAGCTCAGTCCTAGGTACAATGCTAGCG

>lacIqp
GTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCG

>KmRp
TTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGGAAG

>BBa_J23118
TTGACGGCTAGCTCAGTCCTAGGTATTGTGCTAGCG

>BBa_J23111
TTGACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCG

>BBa_J23119
TTGACAGCTAGCTCAGTCCTAGGTATAATGCTAGCG

sgRNA/scRNA sequences

>BBa_J23119(SpeI)-scRNA_1xMS2.b2-rrnBTerm
TTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTNNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAA
GTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACATGAGGATCACCCATGTGCTTTTTTTGAAGCTTG
GGCCCGAACAAAAACTCATCTCAGAAGAGGATCTGAATAGCGCCGTCGACCATCATCATCATCATCATTGAGTTTAA
ACGGTCTCCAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGC
GGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAA
ACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAG
GCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACT

>BBa_J23105(SpeI)-scRNA_1xMS2.b2-ECK120033736
TTTACGGCTAGCTCAGTCCTAGGTACTATACTAGTNNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAA
GTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACATGAGGATCACCCATGTGCTTTTTTTAACGCATG
AGAAAGCCCCCGGAAGATCACCTTCCGGGGGCTTTTTTATTGCGC

>hAAVS1
GGGGCCACTAGGGACAGGAT

>J306
TTGTGTCCAGAACGCTCCGT

>J506
AGCAGCATGAGCAGCATTGA

>J606
GTCGCAGTCGCGCGAGCACT

> J119 (OT1)
TGGTTGCCAAGGGTGACCTA

> JW06 (OT2)
GATGATAGCGTGACTAGCTG

> JD_H (OT3)
ACAAGTATCGAAGCGAACTA

Low-leak Upstream Sequences

The 26nt proximal to the minimal promoter of upstream sequences were mutated to decrease the basal expression of synthetic promoters. The mutated region was underlined in the DNA sequences below.

>J3(LL-E)
AGCATTTGCGATCATTCACGCAGCGCTTATTCAGTTGCTCACTGCGATGTCATAATCATCGCTACGAGCTGTGAAAG
ATGCATAAAGCTCGTACGACGCGTTCGCTCGTCTCCTCACTTCTCCTACGGAGCGTTCTGGACACAACCGGCCCCCC
CCGCTGCCGCGGGCCG

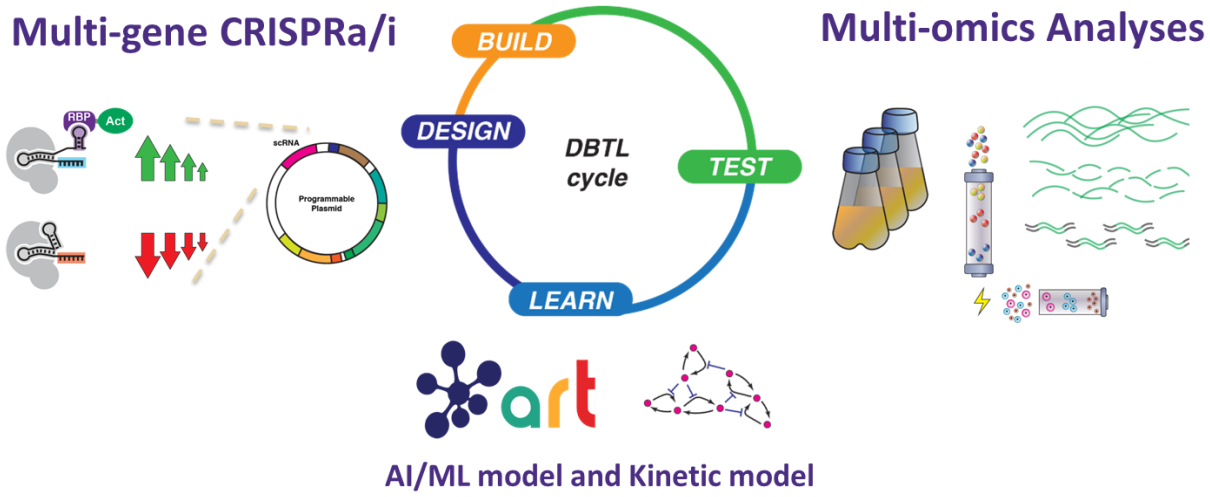
>J5(LL-F)
TATACATCGCATCACTACACTATTGATTATCATTGTGTACGTAACGAGCTTGCACAACGTGAAGTTCTTCGAGCACT
TCAGCTCGCAACGTAAATGACAGTTGCTGTTAAGTGACGTGAATCCTTCAATGCTGCTCATGCTGCTCGCCGCCCGG
CCCGTGCCCGCGCCCG

>J6(LL-E)
CTGCACGAGTTCGCTGTCGAGACAAGTCTCTTAGCGACGTATTACGAAGATCACATAGTCAGATGAAGCTATAGAGC
ACGACGCTAACGATTACGTCACGCTTGACACAACAGTTTCGCTACCTAGTGCTCGCGCGACTGCGACCCGGCCCCCC
CCGCTGCCGCGGGCCG

Chapter 4:

Model-driven DBTL cycle acceleration with bacterial CRISPRa/i

Cholpisit Kiattisewee[+],Ian Faulkner[+],Allan V. Scott, Tijana Radivojevic, Janis Shin, Ava V. Karanjia, Jeremy Zucker, Nathalie Munoz, Herbert Sauro, Hector Garcia Martin, Alex S. Beliaev, Jesse G. Zalatan*, James Carothers*



Abstract

CRISPR-Cas transcriptional tools have been widely applied to program complex biological networks. Recently, we developed CRISPR gene activation (CRISPRa) tools that demonstrate broad applicability across various bacteria. In combination with CRISPR interference (CRISPRi) for gene repression, we developed a multi-gene regulatory platform that can simultaneously manipulate gene expressions at a genome-wide level, ranging from endogenous genes to heterologous pathways. In this work, we apply this CRISPR technology platform to accelerate the Design-Build-Test-Learn (DBTL) processes for microbial strain optimization. We started by computationally designing guide RNA to fine-tune specific targets in expansive metabolic landscapes. Since CRISPRa/i regulates endogenous genes in trans and bypasses the tedious genome engineering steps, the time for build phase is significantly shortened. We acquired multi-omics data in the test phase and forwarded into kinetic-based and machine-learning models to learn the plausible improvements. By coupling programmable gene regulatory tools with large metabolic models informed by omics data, we can narrow down the large engineering landscape to a prioritized set of candidates sufficient for the learning process. With this platform, rapid prototyping in non-model bacteria could be streamlined and accelerated for diverse applications, including metabolic engineering and the discovery of therapeutic targets.

This work is an extended project from Chapter 3 where we applied CRISPRa/i tool in DBTL cycles for strain engineering applications

4.1 Introduction

Microbial metabolic engineering is a major area that contributes to a sustainable circular economy by utilizing renewable materials for bioproduction (Intasian et al., 2021; Nielsen and Keasling, 2016). Facilitated by advancements in synthetic biology, novel pathways have broadened chemical routes beyond what nature has achieved thus far (Lee et al., 2019a). However, the optimization processes to reach the target production levels in microbial strains are often costly and time-consuming. Due to complex genetic networks in biological metabolism, traditional maneuvering of a few selected genes fails to explore the full potential of production capacity. Several recent works have been developed to support genome-wide manipulation and detection of changes resulting from programmed manipulations, such as a genome-scale-model (Lawson et al., 2021), multi-omic analysis (Roy et al., 2021) coupled with a machine learning model (Radivojević et al., 2020), and a mechanistic model based on flux analysis (Choi et al., 2018; Shin et al., 2023). Furthermore, to efficiently navigate genome-scale metabolism, a tool capable of perturbing multiple genes with high programmability is necessary. This project attempts to integrate CRISPR-based gene regulatory tools with other big data strategies to accelerate the DBTL cycle processes, a systematic iterative approach consisting of Design-Build-Test-Learn steps to optimize engineering solutions, gearing toward bioproduction of interest.

CRISPR-Cas systems have been repurposed for several applications in the field of synthetic biology. For applications in transcriptional modifications, catalytically-dead Cas9 (dCas9), guide RNA, and other prosthetic machineries can be used (Dominguez et al., 2016; Vigouroux and Bikard, 2020). One such transcriptional control strategy, CRISPR interference (CRISPRi) can be achieved by having dCas9 block RNA polymerase function (Qi et al., 2013). In contrast, CRISPR activation (CRISPRa) requires an auxiliary component to recruit/stabilize RNA polymerase to the proper position to elevate expressions of designated genes (Bikard et al., 2013; Dong et al., 2018). Combined CRISPR activation/repression (CRISPRa/i) circuits can be programmed at the engineered guide-RNA(s) containing complementary sequences to the DNA target, single-guide-RNA (sgRNA) for CRISPRi and scaffold-RNA (scRNA) for CRISPRa (Fontana et al., 2020b). Therefore, CRISPRa/i can provide a programmable environment for genome-scale engineering to accelerate the DBTL optimization platform.

Recently, CRISPRa/i tools were recently demonstrated to also function in *Pseudomonas putida*, an emerging bacterial host that has different industrially relevant traits suitable for metabolic engineering applications (Kiattisewee et al., 2021). Hence, the accelerated DBTL platform will be applied to *P. putida* and other microbes of interest to explore biosynthesis space beyond benchmark *E. coli* or *S. cerevisiae* hosts. The biosynthetic pathway of p-aminocinnamic acid (Chapter 3) was chosen in this project due to difficulties observed in the *E. coli* system and longstanding interest in the challenging area of aromatic amine biosynthesis (Bartsch and Bornscheuer, 2010; Konishi et al., 2018; Minakawa et al., 2019; Suvannasara et al., 2014).

4.2 Methods

Plasmid and strain constructions

All plasmids were constructed using either Infusion cloning (Takara Bio) or Golden Gate Assembly (NEB & Promega). All cloning-related transformations were performed with *E. coli* NEB-turbo (NEB) or DH10b cells (Invitrogen). *E. coli* MG1655 was used for a bioproduction test. All

engineered *P. putida* are derivatives of KT2440 (Kiattisewee et al., 2021). DNA oligonucleotides were purchased from IDT. Transform of *E. coli* was performed with heat-shock chemically competent cells method. *P. putida* transformation was performed as previously described with either CaCl2-based chemically competent cells or electroporation (Kiattisewee et al., 2021). Sanger sequencing to verify constructs was performed by Genewiz/Azenta. DNA encoding the papABC operon responsible for p-AF production was taken from *Pseudomonas fluorescens* SBW25 (Masuo et al., 2016). The *aroGL* operon responsible for elevated chorismate levels, encoding AroL and a feedback-resistant mutant of AroG, were obtained from a previous report (Juminaga Darmawi et al., 2012). *R. glutinis* PAL was synthesized as gene fragments by Twist Bioscience. Promoter sequences placing these outputs under CRISPRa control were obtained from Anderson Promoter (http://parts.igem.org). Integrative strains were retrieved from previous works (Chapter 3) (Kiattisewee et al., 2021).

Construction of multiplex gRNA CRISPRa Gene Regulatory Networks

CRISPRa tools enable programmable expression of heterologous genes in bacteria using synthetic promoters containing cognitive DNA target of CRISPR guide RNA (gRNA) (Dong et al., 2018; Fontana et al., 2020a; Kiattisewee et al., 2021). We constructed the new design methodology and assembly protocol for this larger GRNs construction (Appendix C). The CRISPRa programs in this study follow the design described in Chapter 3 and Appendix C where up to six orthogonal scRNAs can be simultaneously implemented for regulating different transcriptional units.

Multi-fluorescent reporters assay via flow cytometry

Single colonies of engineered *P. putida* from LB-agar plates were inoculated in 500 µL EZ-RDM (Teknova) supplemented with appropriate antibiotics and grown in 96-deep-well plates at 37 °C and shaking. Cultures were grown overnight at 30 °C and shaking and then diluted in 1:100 in PBS and analyzed on a MACSQuant VYB flow cytometer (Miltenyi Biotec) using a previously described strategy to gate for single cells (Dong et al., 2018). A side scatter threshold trigger (SSC-H) was applied to enrich single cell events. A narrow gate along the diagonal line on the SSC-H vs SSC-A plot was selected to exclude the events where multiple cells were grouped together. Within the selected population, events that appeared on the edges of the FSC-A vs. SSC-A plot and the fluorescence histogram were excluded. For sfGFP detection, the excitation wavelength was 488 nm and emission wavelength was 525/50 nm. For mTagBFP2 detection, the excitation wavelength was 405 nm and emission wavelength was 452/45 nm. For mRFP1 detection, the excitation wavelength was 561 nm and emission wavelength was 615/20 nm.

Aromatic amines bioproduction in bacteria

The culture condition for p-AF and p-ACA biosynthesis in this work was adapted from the previous literature (Burke et al., 2017). Engineered *P. putida* was transformed with CRISPR-mediated pathway expression plasmids, while a control strain was transformed with similar plasmids without pathway genes and used for standard curve diluent. Single colonies in triplicates were inoculated to 2 mL of MOPS EZ-Rich defined media (Teknova) and appropriate antibiotics, in 14 mL polypropylene culture tubes and were grown at 30˚C, shaken at 200 rpm for 18 hours as starting cultures. Then, each tube was subcultured to a 30 mL culture in a customized platform equipped with automated optical density collection, made at PNNL. These cultures were

incubated to OD_{600} at 0.8 (1 cm path length) for RNA extraction and downstream transcriptomic analysis. The residual cultures were further incubated to reach a total of 36 hours to allow late p-ACA bioconversion. The final supernatants were further subjected to HPLC analysis for metabolite quantification. Aromatic amino acid derivatives were purchased from Sigma, TCI, and Santa Cruz to use as HPLC standards.

Identification of genomic targets to leverage p-ACA production in *P. putida*

Genomic targets for CRISPRa gene activation and CRISPRi gene repression were performed using Flux-RETAP (A REaction TArget Prioritization) based on the modified genome-scale metabolic model to include chemical reactions derived from heterologous p-ACA biosynthetic genes (Stevens and Carothers, 2015; Tokic et al., 2020). By varying endogenous gene expression levels, 10 native genes that yielded a high tendency for increasing p-ACA production were selected for CRISPRa screening. In the case of CRISPRi, 30 genes with a high tendency for increasing p-ACA production were selected for validation due to simpler CRISPRi design rules by ORF-targeting sgRNA. Additional native genes relevant to aromatic amino acid biosynthesis that potentially compete with p-ACA accumulations were added to downregulation candidates (Liu et al., 2019; Luo and Lee, 2020).

Investigation of CRISPRa/i for *P. putida* endogenous gene targets

The sfGFP reporter was modified to incorporate an endogenous promoter, as an intergenic region, to a generate transcriptional fusion of endogenous promoter and sfGFP on a pBBR1-GmR plasmid. For endogenous CRISPRa investigation, the intergenic region contained 60 bases from the ORF of interest on the 3' end. On the 5' end, the intergenic region extended 60 bp into the immediate upstream ORF, following a previously reported strategy (Zaslaver et al., 2006). Due to limited knowledge to predict the activatability of CRISPRa in endogenous promoters, we screened all scRNA with proper distance-to-TSS with reported and predicted TSS (Coppens and Lavigne, 2020; D'Arrigo et al., 2016). In the case of CRISPRi investigations, 300 bp of the ORF was incorporated to extend the target sites available for CRISPRi-based gene repression. If the gene of interest is part of an operon, we include upstream DNA up to 150 bases of the upstream ORF that is beyond the operon of interest (Caspi et al., 2018).

Transcriptomic analysis of CRISPRa/i perturbation in the p-ACA production strains

From cultures collected at OD_{600} 0.8 (1 cm path length), each sample was lysed and purified for total RNA using Aurum™ Total RNA Mini Kit (Bio-Rad). The triplicate samples conferring to the same genotypes were pooled before submitted to standard RNA sequencing by Genewiz/Azenta. The transcriptomic profiles were analyzed using *P. putida* KT2440 reference genome plus heterologous genes (CRISPR machinery and pathway genes) to calculate read counts corresponding to each gene and then normalized with the DESeq2 algorithm, performed by Ryan Mcclure (PNNL). Further analyses were performed using standard Python packages.

4.3 Preliminary Results

Multiplex CRISPRa to control metabolic genes

Microorganism-based chemical bioproduction plays an important role in building a sustainable circular bioeconomy as it can convert relatively cheap renewable feedstocks into valuable

chemicals. Recent developments in synthetic biology and metabolic engineering also expand chemical repertoires available through biosynthetic or semi-synthetic processes (Lee et al., 2019b) and also technology to engineer non-conventional organisms for the production of specialty chemicals (Fatma et al., 2020; Yan and Fong, 2017). In this study, we choose *P. putida* as a host due to its tolerance to harsh conditions which enables production of p-aminocinnamic acid (p-ACA) which was shown to be toxic to a model organism like *E. coli* (Chapter 3). Further investigation of p-ACA tolerance in additional industrially-relevant bacteria showed that *P. putida* showed the highest tolerance (Figure 4.S1). The previous characterization of CRISPRa in *P. putida* also made it attractive as a host for strain optimization via CRISPRa/i circuit (Figures 4.1A & 4.1B) (Kiattisewee et al., 2021) as previous investigations of multi-gene programming in bacteria usually made up to 2-3 nodes (Kiattisewee et al., 2021; Santos-Moreno et al., 2020; Tickman et al., 2021).

Therefore, we started by evaluating the potential of *P. putida* CRISPRa/i tools in a larger genetic circuit (Figures 4.1C-4.1F). Previous work has demonstrated the ability to multiplex CRISPRi in *P. putida*, but failed in significantly perturbing all the desired genes as programmed, likely due to complex regulation in the native metabolism (Banerjee et al., 2020). In terms of CRISPRa in *E. coli*, a study by Clamons and Murray postulated that CRISPRa fold activation could be sustained by up to dozens of scRNA species (Clamons and Murray, 2019). These studies suggested that proper characterization of multiple scRNA programming should be evaluated before further implementation of the DBTL study.

Since the sgRNA competition was previously observed in *E. coli* multi-gRNA programs (Huang et al., 2021; Zhang and Voigt, 2018), we started by varying scRNA concentration to obtain an optimal expression level of scRNA in *P. putida* system. Guided by a similar approach implemented in CRISPRi-mediated gene repression in *E. coli* (Fontana et al., 2018), we titrated the expression level of scRNA levels in *P. putida* CRISPR where dCas9 and activator cassettes were constitutively expressed from the genome (Figure 4.S2). We found that a moderately strong expression level (BBa_J23110) yielded the highest fold-activation for CRISPR-mediated fold-change, agreeing with other investigations by our team in *E. coli* (Alba Burbano et al., 2023). Therefore, we choose BBa_J23110 as an expression level for further constructions of multi-gRNA programs.

Using the high-throughput strategy developed by our lab (Appendix C), we can rapidly build a multi-gRNA program with 6 gRNA in a one-pot Golden-Gate assembly reaction. Through sequential construction, we can reach up to 9 gRNA in total, providing the possibility to build larger and more complex gene regulatory networks (GRNs). Using three fluorescent reporters previously examined in the *E. coli* system (Appendix H), we first demonstrated that multi-gene CRISPRa remains effective with the 3 scRNAs/reporters set-up (Figures 4.1C & 4.1E). By increasing the number of scRNA from 3 scRNAs to 6 scRNAs where the latter 3 scRNA are off-target, we found that fold-change from CRISPRa-mediated gene expression dropped as the number of scRNA increased (Appendix C). We then tested the effect of increasing scRNA number with a single plasmid system and genome-integrated system for in vivo bioproduction of p-ACA. We found that the titer of p-ACA slightly decreased as a result of increasing scRNA for the plasmid-based system (539 ± 80 μM to 433 ± 13 μM for 3-scRNA and 6-scRNA, respectively) while the genome-integrated system suffered greatly from additional scRNA (Figures 4.1D & 4.1F-4.1G). These

findings suggested that a larger CRISPRa/i genetic circuit could be applied for strain optimization where additional gRNA could be introduced together with the programmability of individual gRNA.

Tuning CRISPRa/i in *P. putida*

Next, we proceed to evaluate the tunability of CRISPRa/i circuit by gRNA engineering. Based on previous CRISPRa/i characterization and RNA engineering works performed by our team, two key factors are outstanding criteria for gene expression tuning — i) target sites' position relative to the promoter and ii) gRNA folding energetics. In the case of CRISPRa, we demonstrated that distance significantly affects the efficiency in *E. coli* and *P. putida* (Fontana et al., 2020a; Kiattisewee et al., 2021). For CRISPRi, distance screening has been performed in *E. coli* and *B. subtilis* (Peters et al., 2016; Qi et al., 2013). Here, we applied the same design rules for CRISPRi-mediated gene repression in *P. putida* with a highly-expressed sfGFP integrated into the genome (CKPP037) since CRISPRi tends to yield higher efficiency at a lower copy-number in contrast to CRISPRa where medium copy-number exhibited higher fold-change (Figure 4.S3) (Kiattisewee et al., 2021). By testing CRISPRa at 17 different target sites, we found a similar trend as was seen previously, with the exception of certain outliers from the distance factor (Figures 4.S4A & 4.S4B). This finding suggested that there could be other factors to explore beyond the distance that govern the efficiency of CRISPRi-mediated gene repression. We then applied the gRNA folding energetics algorithm developed by our lab (Appendix H) to the 17 sgRNAs tested. We found that the gRNA with previous selection criteria (New Net Energy below -27.5 kcal/mol and <10 kcal/mol Bind Barrier) agrees with the trend of highly-functional sgRNA tested in *P. putida* (Figures 4.S4C & 4.S4D). Therefore, these gRNA folding energetics algorithms could be applied for computational gRNA analysis beyond *E. coli* system.

We also further evaluated the potential of gRNA modifications to control CRISPRa/i efficiency via gRNA truncation (Appendix H, Figure 4.S5A). Similar characterization is also performed for tunable CRISPRi gene repression (Figure 4.S5B). We also evaluated the gRNA folding energetics of these truncated spacers and found that the binding energy of these truncated gRNAs are positively correlated with the spacer lengths. We found that variation of sgRNA spacer and truncation of sgRNA spacer can provide access to different gene expression levels. To evaluate the potential of further engineering to include an additional regulatory knob to gRNA, such as inclusion of a small-molecule responsive RNA aptamer (Hwang and Carothers, 2016), we tested 5'-modification by addition of a stable RNA hairpin upstream of the gRNA spacer. We found that 5'-modification significantly impaired the fold-change from both CRISPRa and CRISPRi (Figure 4.S6). We hypothesized that this 5'-modification approach is similar to introducing mismatches to the extended spacer target (Todor et al., 2021). To this point, we demonstrated that programming at the gRNA level could be acquired either by varying concentrations of gRNA (Figure 4.S2) or modification of gRNA spacer sequences (Figures 4.S3-4.S6).

Identification and evaluation of CRISPRa/i perturbation to leverage p-ACA production

Moving forward to the strain engineering via CRISPRa/i perturbation, genomic targets for CRISPRa/i were identified using Flux RETAP methodology (Figure 4.2 and Figure 4.S7) or suggested from competing aromatic biosynthetic pathway (Figure 4.S8). Out of >40 gene targets, most candidates were part of carbohydrate, amino acids, nucleic acids, and fatty acids metabolism. Then, we constructed a synthetic GFP-fusion reporter based on the previously used

strategy for all candidate genes (Figure 4.3A) (Zaslaver et al., 2006). For any CRISPRa target, all scRNAs with compatible PAM at the appropriate distance were selected for further investigation (ranging from 2 to 5 scRNA variants). scRNA that exhibited the highest fold-change in the transcriptional fusion reporter for each promoter was chosen for further consideration. For CRISPRi, due to ease of design for effective sgRNAs, only 2 sgRNAs per gene target with appropriate gRNA folding energetics were investigated and the one with higher fold-change was selected for further applications.

We also expected that CRISPRa/i will introduce polar effects to the adjacent genes, especially those in the same operon as the perturbed genes. Therefore, we investigated the polar effect from CRISPRa and CRISPRi of the first or second genes in the transcriptional fusion GFP-RFP (Figure 4.S9A) or RFP-GFP (Figure 4.S9B). It was found that CRISPRa is likely exhibiting similar effects to two genes in the operon while CRISPRi affects both genes in the operon but has more effect on the targeting gene. The polar effect was also tested in the synthetic promoter of the native oprB-II-gcd operon (Figure 4.S9C). Targeting gcd or sfGFP ORFs led to decreasing sfGFP expression and also relieved growth defects likely due to increasing decoy binding sites of putative transcription factors regulating oprB-II-gcd operon (Wang et al., 2021).

Throughout the screening of gRNAs targeting endogenous genes, CRISPRa/i perturbed strains with any observed growth defect were omitted from the set. We used a 1.5-fold dynamic range cut-off for both CRISPRa and CRISPRi. We curated 7 scRNAs with observed CRISPRa-based gene activation (64%) and 34 sgRNAs of accessible CRISPRi targets (94%) for further implementation in the metabolic engineering context (Figures 4.3B & 4.3C and Figures 4.S10 & 4.S11).

Construction and planning of the DBTL cycle

With endogenous perturbations identified and characterized, we proceed with applying the CRISPRa/i-based tools in the DBTL strain optimization (Figure 4.4A). The CRISPRa/i-based DBTL approach is described by each module — Design, Build, Test, and Learn. **Design:** endogenous genes were previously identified with genome-scale metabolic models and characterized with transcriptional fusion synthetic reporters. **Build:** CRISPRa/i circuits were incorporated into p-ACA production plasmid by appending 4th gRNA to the existing 3-scRNA p-ACA production plasmid using a sequential Golden-Gate Assembly strategy (Appendix C). **Test:** each strain with designated perturbations was cultivated in a MOPS-based medium (Chapter 3) and tested for transcriptomics and targeted metabolomics analyses. **Learn:** The resulting data were used to learn the outcome of endogenous perturbation which can then be fed into the models and design new combinatorial perturbations further.

During the TEST of p-ACA producing *P. putida*, three endogenous CRISPRi strains failed to grow in the MOPS-based medium. In terms of metabolites, we found that the introduction of the 4th gRNA with an off-target spacer sequence showed a high p-ACA production level (582 ± 49 μM) comparable to that of 3 scRNA programs (Chapter 3). However, we found that 20 out of 30 strains with CRISPRi programs exhibited decreased p-ACA levels to less than 10 μM (Figure 4.4B). These impairments could be due to strong gene repression toward metabolic genes, leading to disruption of p-ACA biosynthetic pathway. On the other hand, strains with CRISPRa

perturbations exhibited marginal changes in p-ACA concentration but possessed varying side-products — e.g. p-coumaric acid (p-CA) and p-aminobenzoic acid (p-ABA) (Figure 4.S12).

Then, we collected and analyzed transcriptomic data of each strain by finding the median expression level of each gene out of the 42 strains successfully acquired in this stage including two control strains (3 scRNA and 4th off-target scRNA strains). To our surprise, a large number of perturbations were not activated nor repressed as previously found in the synthetic reporter context (Figure 4.S13). Out of the perturbations with successful CRISPRa/i programming, the polar effect was also observed when the gene target is part of a multi-gene operon (Figure 4.S14). We further evaluate top gene candidates that were upregulated in the high p-ACA producing strains to draw correlations between CRISPRa/i programming and the resulting transcriptomic output. We found that symporters and transporters related to aromatic amino acids or nitrogen-containing compounds were upregulated when p-ACA was observed (Figure 4.4C). We postulated that CRISPRa/i perturbations could trigger native cellular regulation and resulted in feedback responses that led to an opposite result to the programmed input at the collection time for transcriptomic analysis (Figure 4.S15).

Out of 40 perturbations tested, we narrowed down target genes to 10 candidates (4 CRISPRa and 6 CRISPRi) for further construction of the new DBTL cycle using metabolomics and transcriptomics data as selection criteria (Figure 4.5). We focused on CRISPRa/i perturbing strains with matching activation or repression effects as programmed together with their production levels of p-ACA. The data acquired at this round will serve as a basis for designing new combinations with diverse programming. Using this knowledge in the current development of DBTL acceleration platform, we aim to incorporate machine learning based and mechanistic model based learning to provide appropriate suggestions for improvement in chemical bioproduction.

4.4 Discussions and Future Aspects

Based on >40 variations tested in the initial CRISPRa/i circuit, we found that the central metabolism of *P. putida* is potentially difficult to perturb and resistant to programming. Informed by the inconsistent CRISPR perturbation efficiency in the synthetic reporter context and the real changes derived from the native gene expression, we need alternative models to understand metabolism in hosts like *P. putida*. Such a mechanistic model could be further incorporated into the genome-scale analysis in which the parameters could be comprehensible (Shin et al., 2023). Toward this end, we will investigate the metabolic outcome of endogenous perturbations in the context of p-ACA metabolic pathway using the Bayesian metabolic model currently explored in our team. With these tools customized for our study, we can further elaborate the gene regulatory networks (GRNs) modeling for further application beyond metabolic engineering (Badia-i-Mompel et al., 2023; Tickman et al., 2021), such as microbial consortia modeling (Duncker et al., 2021), molecular recording and responses (Sheth and Wang, 2018), and therapeutic discovery (Anglada-Girotto et al., 2022). We will also further investigate the transcriptomic response at a single cell level to investigate the changes at a population level using a single-cell approach, e.g. microSPLiT (Kuchina et al., 2021).

Regarding the p-ACA production levels, only a few CRISPRa perturbations (*tal* and *gcl* activations) showed a marginal improvement in titer despite significantly lower side-product

accumulations. These findings suggested that there might be a bottleneck in the chemical reaction cascades that limit the maximum production of p-ACA to ~500-600 μM (~80-100 mg/L). Recent work on *pabABC* enzymes homologs from *C. glutamicum* (p-aminobenzoic acid biosynthesis, not to be confused with *papABC*) suggested that these native pathways are highly connected to the p-AF production by heterologous *papABC* pathway (Mohammadi Nargesi et al., 2019). Therefore, further modifications to the endogenous genes either by deletion or repression could redirect the flux from the p-ABA pathway to that of desired p-ACA synthesis. Further optimization at the *papABC* operon itself is also possible to fine-tune the individual reaction of *papA*, *papB*, and *papC* to prevent the accumulation of intermediates that could be fed into other pathways, especially conversion to p-ABA by *pabC* (O'Rourke et al., 2011). As the CRISPRa platform to facilitate variation of heterologous gene expression was developed in *E. coli* (Appendix H), *papABC* could also be refactored into its individual transcriptional uni for further fine-tuning by the scRNA program. Altogether, the CRISPRa/i DBTL acceleration approach also provides insightful information and could be combined with other strategies to improve the strain engineering platform.

In addition to the existing CRISPRa/i gene regulation technologies in *P. putida*, we can also expand the BUILD approach with emerging genetic engineering strategies based on high-performance integration through landing pad technologies (Elmore et al., 2023; Wang et al., 2019). Upon improved efficiency of gene delivery, a gRNA library could also be implemented for high-throughput screening (Fenster et al., 2022). Since CRISPRa-mediated gene expression could be a prototype for desired gene expression from a different copy-number expression strategy (Chapter 3), we can also diverge the p-ACA production pathway from the CRISPRa-mediated gene expression control which will prevent retroactivity originating from an expansion of gRNA programs and provide higher feasibility to the design space for CRISPRa/i-based genetic manipulation (Appendix C). Finally, these findings based on a study in *P. putida* could potentially be transferred to other organisms, especially to *A. baylyi* with p-ACA resistance, simple and rapid genetic manipulation method, and metabolic potentials for lignocellulosic biomass utilization (Arvay et al., 2021; Biggs et al., 2020). Demonstration and development of CRISPRa/i-based DBTL acceleration in the following cycles will pave the way for strain optimization for p-ACA biosynthesis and other bioproduction campaigns. Another modulation beyond transcriptional control, such as translational regulation, could also be incorporated to provide an alternative DBTL cycle strategy. The mRNA targeting CRISPR systems is attractive due to potentially lower polar effects compared to the transcriptional control that will affect adjacent genes in the same operon. Finally, these programmable and easy-to-build CRISPR-based platforms could also be adapted to study other systems ranging from studying complex biological networks to the discovery of novel therapeutic targets.

Author contributions

C.K. and I.D.F. designed CRISPRa/i-based experiments and analyzed data. C.K. and I.D.F. built CRISPR-based plasmids and strains. C.K. I.D.F. and A.V.S. performed microbial cultures. C.K. and I.D.F. performed HPLC-based targeted metabolomics. N. M. performed an in-depth metabolomic analysis. A.V.S. performed RNA collection, extraction, and data acquisition. C.K. analyzed transcriptomic data. T.R., J.S., J.Z., H.S., and H.G.M. conducted computational analyses and developed the models. C.K., I.D.F., and A.V.K. performed p-ACA tolerance

experiments. C.K. and S. S. B. developed Golden-Gate Assembly for genetic circuit construction. H.S., H.G.M., A.S.B., J.G.Z., and J.M.C. supervised the projects. C.K., I.D.F., J.G.Z., and J.M.C. wrote the manuscript with input from all authors.

Declaration of competing interest

J.G.Z. and J.M.C are members of the Wayfinder Biosciences scientific advisory board.

Acknowledgments

We thank Ryan McClure for assistance in transcriptomic analysis. We thank Eric Hill for assistance in customized microbial cultures. We thank members of the Zalatan and Carothers groups and collaborators from PNNL and LBNL for technical assistance, advice, and helpful discussions. This work was supported by NSF Award 1817623 (J.M.C, J.G.Z.), NSF Award 2225632 (J.G.Z. and J.M.C.), DOE Award DE-EE0008927 (J.M.C, J.G.Z.) and DOE Award DE-SC0023091 (J.M.C, J.G.Z.).

Supporting Information

Supplementary data associated with this work is provided at the end of this chapter.

4.5 Tables and Figures

Table 4.1: *P. putida* CRISPRa endogenous gene candidates recommended by Flux-RETAP

Index	Locus Tag	Gene	Metabolic Pathway	Basal Level (a.u.)	Fold-Change
1	PP_1972	*tyrB*	AAA[a]	5819.6	No activation
2	PP_2329	*pabB*	AAA[a]	342.8	4.76
3	PP_1830	*aroC*	AAA[a]	3139.5	1.80
4	PP_5079	*aroK*[b]	AAA[a]	332.6	2.48
5	PP_5078	*aroB*[b]	AAA[a]		
6	PP_4297	*gcl*	Glyoxylate metabolism	4829.0	3.90
7	PP_1075	*glpK*	Glycerol metabolism	2825.4	2.99
8	PP_4715	*tpiA*	Glycerol metabolism	31603.0	Growth defect
9	PP_4169	*gpsA*	Glycerol metabolism	24180.3	1.32
10	PP_4965	*tktA*	TCA cycle	4488.7	1.74
11	PP_2082	*ppsA*	TCA cycle	920.4	1.09
12	PP_2168	*tal*	TCA cycle	2244.6	2.11

Note: [a]AAA stands for Aromatic amino acid biosynthesis, [b]aroK and aroB are part of the same transcriptional unit and can be activated from the same promoter

Table 4.2: *P. putida* CRISPRi endogenous gene candidates recommended by Flux-RETAP

Index	Locus Tag	Gene	Metabolic Pathway	Basal Level (a.u.)	Fold-Change
1	PP_2000	*purF*	Nucleic acid metabolism	261.9	1.80
2	PP_4823	*purD*	Nucleic acid metabolism	435.7	5.27
3	PP_1037	*purL*	Nucleic acid metabolism	2756.5	10.16
4	PP_1665	*purM*	Nucleic acid metabolism	2113.0	22.14
5	PP_4016	*purB*	Nucleic acid metabolism	257.4	4.49
6	PP_5335	*purK*[a]	Nucleic acid metabolism	3319.4	11.19
7	PP_5336	*purE*[a]	Nucleic acid metabolism		
8	PP_4822	*purH*	Nucleic acid metabolism	680.8	9.59
9	PP_1032	*guaA*	Nucleic acid metabolism	3713.6	5.67
10	PP_4889	*purA*	Nucleic acid metabolism	397.7	1.75
11	PP_4678	*ilvC*	Amino acid metabolism	5986.2	20.67
12	PP_1025	*leuA*	Amino acid metabolism	8054.3	5.83
13	PP_1988	*leuB*	Amino acid metabolism	1071.4	16.25
14	PP_1768	*serC*	Amino acid metabolism	373.9	No Repression
15	PP_4909	*serB*	Amino acid metabolism	1309.2	No Repression
16	PP_5155	*serA*	Amino acid metabolism	1660.0	34.9
17	PP_5128	*ilvD*	Amino acid metabolism	3157.8	6.48
18	PP_1530	*dapD*	Amino acid metabolism	2507.1	10.72
19	PP_1525	*dapE*	Amino acid metabolism	1960.6	11.55
20	PP_4725	*dapB*	Amino acid metabolism	539.8	4.26
21	PP_1237	*dapA-I*	Amino acid metabolism	7526.7	45.64
22	PP_2639	*dapA-II*	Amino acid metabolism	400.1	7.45
23	PP_1989	*asd1*	Amino acid metabolism	3417.5	16.95
24	PP_1992	*asd2*	Amino acid metabolism	947.5	30.76

Index	Locus Tag	Gene	Metabolic Pathway	Basal Level (a.u.)	Fold-Change
25	PP_4473	*lysC*	Amino acid metabolism	374.9	3.36
26	PP_0966	*hisD*[b]	Amino acid metabolism	779.1	15.24
27	PP_0967	*hisC*[b]	Amino acid metabolism		
28	PP_0289	*hisB*	Amino acid metabolism	2270.0	13.74
29	PP_5015	*hisE*	Amino acid metabolism	4061.6	3.18
30	PP_1914	*fabG*	Fatty acid metabolism	4966.0	4.82
31	PP_0059	*gmbH*	Lipopolysaccharide metabolism	2533.2	48.22

Note: [a]purK and purE are part of the same transcriptional unit, [b]hisD and hisC are part of the same transcriptional unit

Table 4.3: *P. putida* CRISPRi endogenous gene candidates from competing pathways

Index	Locus Tag	Gene	Metabolic Pathway	Basal Level (a.u.)	Fold-Change
1	PP_0417	*trpE*	Tryptophan biosynthesis	1994.9	14.43
2	PP_0420	*pabA*[a] *(trpG)*	Tryptophan biosynthesis	2429.2	12.44
3	PP_0421	*trpD*	Tryptophan biosynthesis		
4	PP_1769	*pheA*	Phenylalanine biosynthesis	528.9	3.62
5	PP_1770	*tyrA*	Tyrosine biosynthesis	314.3	4.18
6	PP_1917	*pabC*	p-Aminobenzoate biosynthesis	238.9	2.73
7	PP_2554	*quiC*	Shikimate metabolism	501.9	5.29
8	PP_3569	*quiA*	Shikimate metabolism	716.0	20.51

Note: [a]pabA and trpD are part of the same transcriptional unit

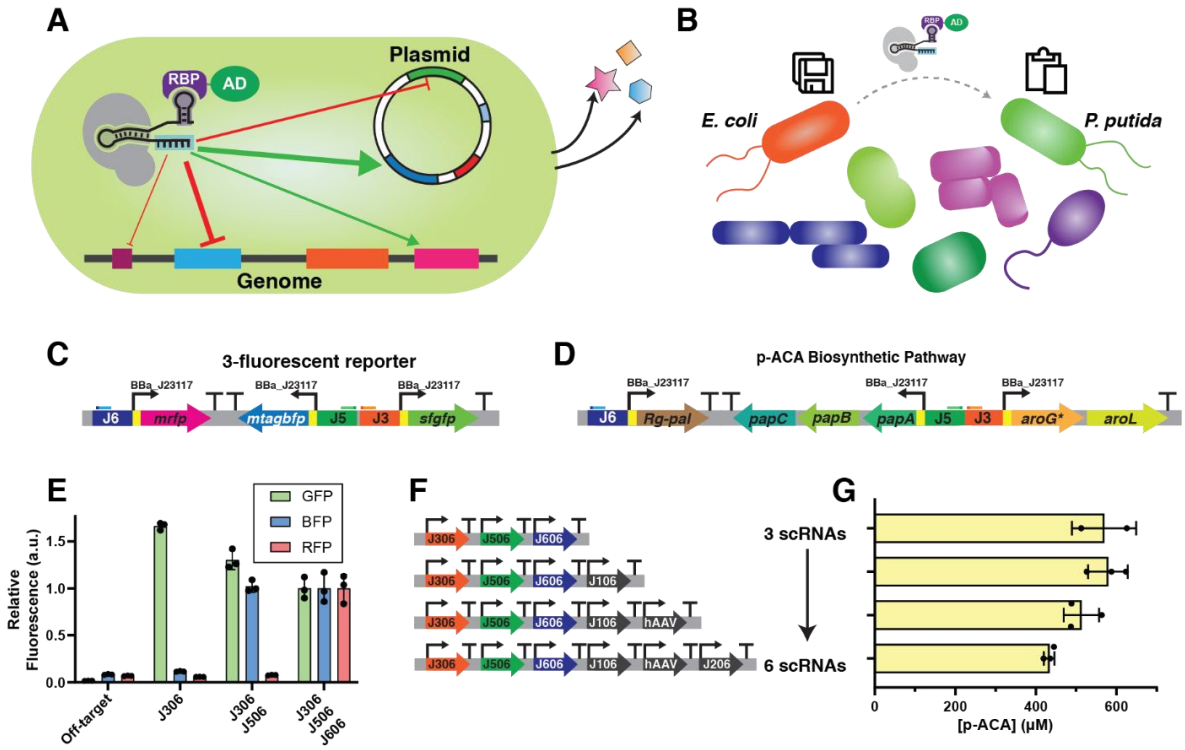


Figure 4.1: CRISPR circuits implementation in bacteria

(A) CRISPRa/i can be used for gene regulation of both heterologous genes and endogenous genes delivered on a plasmid or integrated into the genome. (B) CRISPRa/i tools originally developed in *E. coli* are portable to other bacteria and were extensively characterized in *P. putida*. (C) 3-fluorescent reporter consisted of sfGFP, mtagBFP, and mRFP under synthetic CRISPRa reporters. (D) p-ACA biosynthetic pathway under control of synthetic CRISPRa reporters. (E) CRISPRa-mediated gene expression of three fluorescent proteins as a function of scRNA input. (F) Organization of gRNA assembly for increasing the number of scRNA (G) p-ACA production with an increasing number of scRNA (3-scRNA to 6-scRNA).

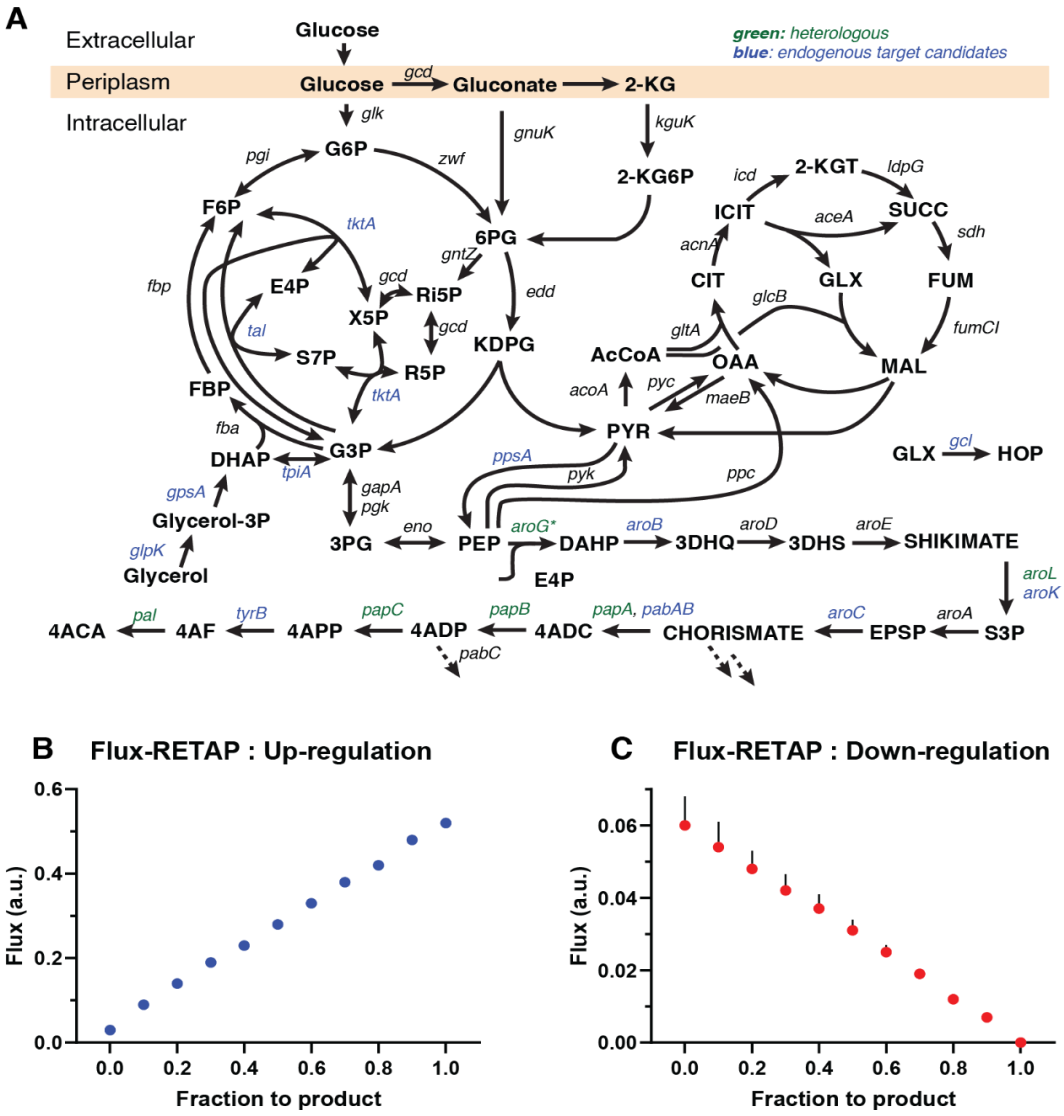


Figure 4.2: CRISPRa/i targets identification from metabolic model

(A) Identification of *P. putida* target genes for CRISPRa relevant to increased production of p-ACA via genome-scale metabolic model (Flux-RETAP). See Supplementary Figure 4.S7 & 4.S8 for CRISPRi target genes. (B & C) Examples of Flux-RETAP analysis of gene candidates for CRISPRa gene activation and CRISPRi gene repression.

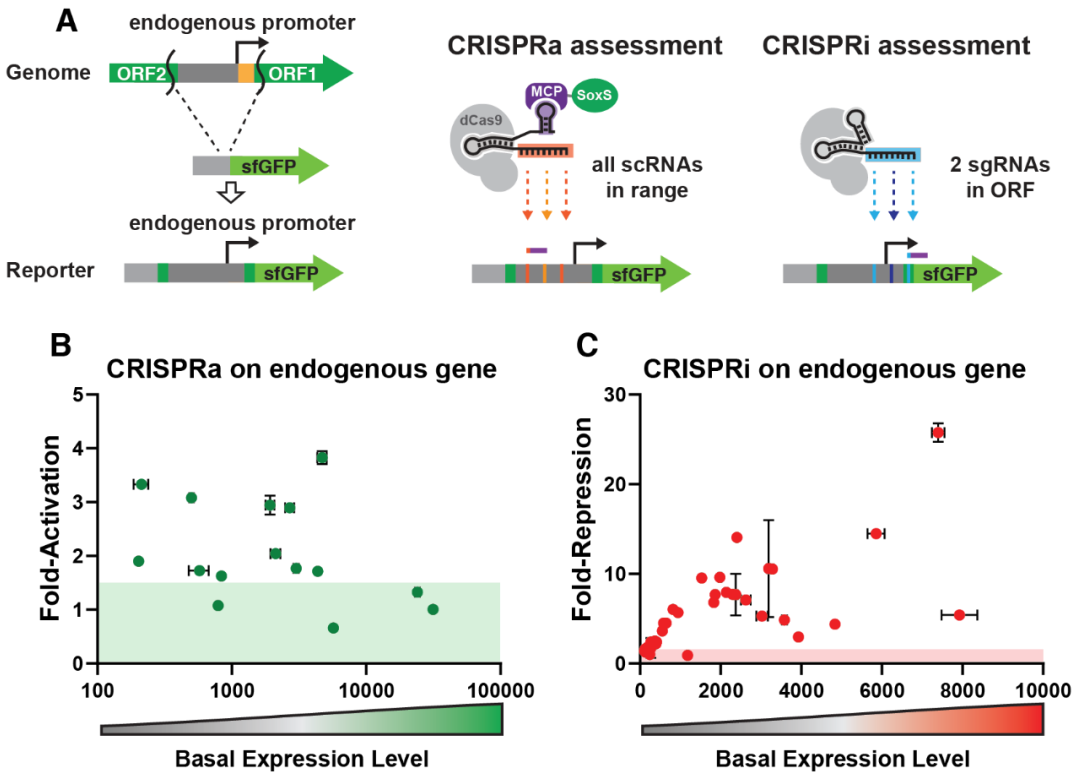


Figure 4.3: CRISPRa/i investigation of endogenous genes

(A) Schematic of CRISPRa/i capability using plasmid-based GFP-fusion reporter. For CRISPRa, all scRNAs in effective range were investigated while 2 sgRNAs with qualified RNA folding energetics were chosen for CRISPRi evaluation. (B) Highest CRISPRa fold-activation from each endogenous promoter plotted in a function of basal expression level. (C) Highest CRISPRi fold-repression from each endogenous promoter plotted in a function of basal expression level.

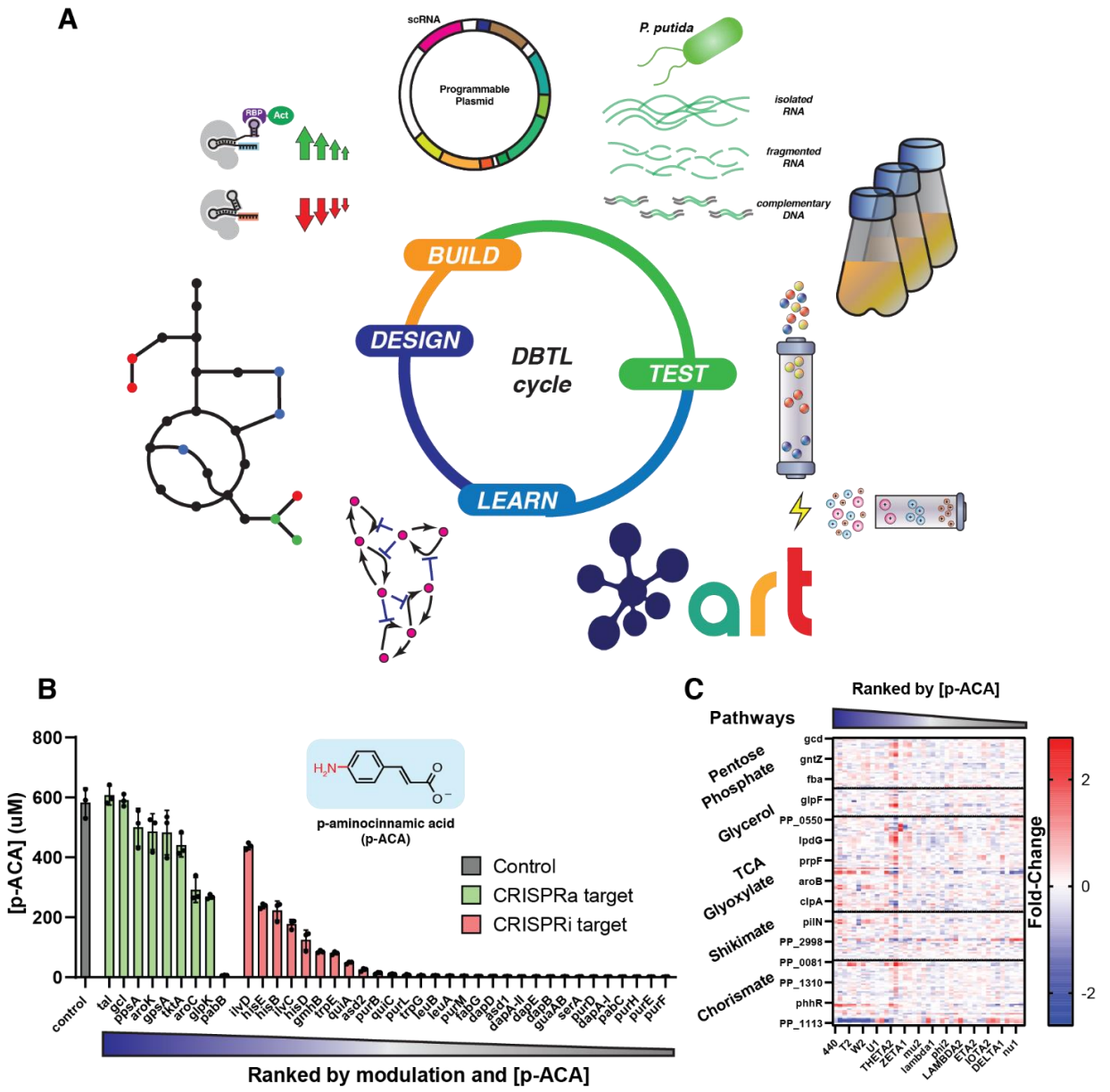


Figure 4.4: DBTL cycle approach for strain engineering accelerated by CRISPRa/i

(A) Schematic of CRISPRa/i-based DBTL cycle strain optimization. For designing, endogenous genes were identified with genome-scale metabolic models. For building, CRISPRa/i circuits were evaluated and incorporated into p-ACA production plasmid. Then, each strain with designated perturbations were cultivated and tested for transcriptomics and targeted metabolomics analyses. The resulting data were used to learn the outcome of endogenous perturbation which can then be fed into the models and design new combinatorial perturbations further on. (B) p-ACA productions from each perturbation were plotted. (C) Relative RNA levels of metabolic genes were plotted in a heatmap.

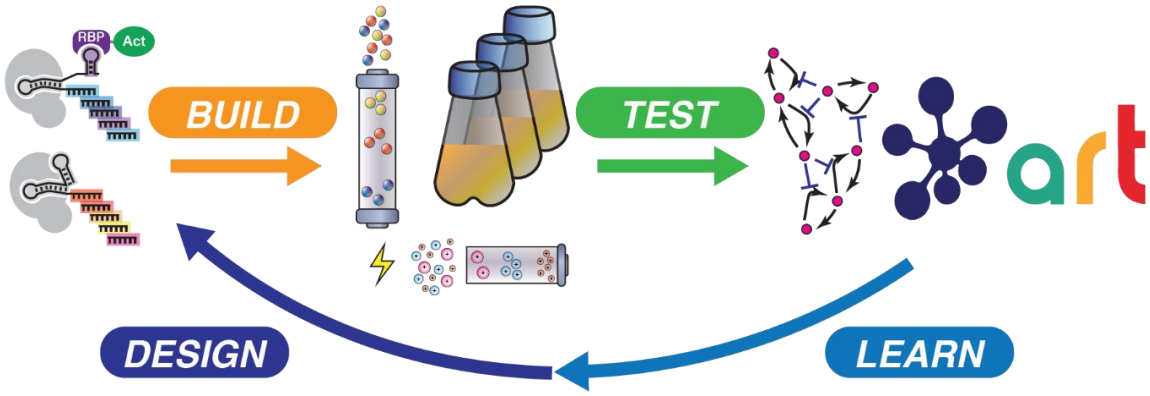


Figure 4.5: Next-step of DBTL cycle exploration

Further rounds of DBTL cycles were proposed as a combination of CRISPRa/i-based perturbations with 2 to 3 perturbations simultaneously. 4 CRISPRa targets and 6 CRISPRi targets from the single perturbation were chosen based on p-ACA levels (high production) and transcriptomic responses (activation or repression at the target gene).

4.6 References

Alba Burbano, D., Cardiff, R.A.L., Tickman, B.I., Kiattisewee, C., Maranas, C.J., Zalatan, J.G., Carothers, J.M., **2023**. Engineering activatable promoters for scalable and multi-input CRISPRa/i circuits. *Proc. Natl. Acad. Sci.* **120**, e2220358120. https://doi.org/10.1073/pnas.2220358120

Anglada-Girotto, M., Handschin, G., Ortmayr, K., Campos, A.I., Gillet, L., Manfredi, P., Mulholland, C.V., Berney, M., Jenal, U., Picotti, P., Zampieri, M., **2022**. Combining CRISPRi and metabolomics for functional annotation of compound libraries. *Nat. Chem. Biol.* **18**, 482–491. https://doi.org/10.1038/s41589-022-00970-3

Arvay, E., Biggs, B.W., Guerrero, L., Jiang, V., Tyo, K., **2021**. Engineering *Acinetobacter baylyi* ADP1 for mevalonate production from lignin-derived aromatic compounds. *Metab. Eng. Commun.* **13**, e00173. https://doi.org/10.1016/j.mec.2021.e00173

Badia-i-Mompel, P., Wessels, L., Müller-Dott, S., Trimbour, R., Ramirez Flores, R.O., Argelaguet, R., Saez-Rodriguez, J., **2023**. Gene regulatory network inference in the era of single-cell multi-omics. *Nat. Rev. Genet.* https://doi.org/10.1038/s41576-023-00618-5

Banerjee, D., Eng, T., Lau, A.K., Sasaki, Y., Wang, B., Chen, Y., Prahl, J.-P., Singan, V.R., Herbert, R.A., Liu, Y., Tanjore, D., Petzold, C.J., Keasling, J.D., Mukhopadhyay, A., **2020**. Genome-scale metabolic rewiring improves titers rates and yields of the non-native product indigoidine at scale. *Nat. Commun.* **11**, 5385. https://doi.org/10.1038/s41467-020-19171-4

Bartsch, S., Bornscheuer, U.T., **2010**. Mutational analysis of phenylalanine ammonia lyase to improve reactions rates for various substrates. *Protein Eng. Des. Sel.* **23**, 929–933. https://doi.org/10.1093/protein/gzq089

Biggs, B.W., Bedore, S.R., Arvay, E., Huang, S., Subramanian, H., McIntyre, E.A., Duscent-Maitland, C.V., Neidle, E.L., Tyo, K.E.J., **2020**. Development of a genetic toolset for the highly engineerable and metabolically versatile *Acinetobacter baylyi* ADP1. *Nucleic Acids Res.* **48**, 5169–5182. https://doi.org/10.1093/nar/gkaa167

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 Res* **41**, 7429–37. https://doi.org/10.1093/nar/gkt520

Burke, C.R., Sparkman-Yager, D., Carothers, J.M., **2017**. Multi--state design of kinetically--controlled RNA aptamer ribosensors. *bioRxiv*. https://doi.org/10.1101/213538

Caspi, R., Billington, R., Fulcher, C.A., Keseler, I.M., Kothari, A., Krummenacker, M., Latendresse, M., Midford, P.E., Ong, Q., Ong, W.K., Paley, S., Subhraveti, P., Karp, P.D., **2018**. The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res.* **46**, D633–D639. https://doi.org/10.1093/nar/gkx935

Choi, K., Medley, J.K., König, M., Stocking, K., Smith, L., Gu, S., Sauro, H.M., **2018**. Tellurium: An extensible python-based modeling environment for systems and synthetic biology. *Biosystems* **171**, 74–79. https://doi.org/10.1016/j.biosystems.2018.07.006

Clamons, S., Murray, R., **2019**. Modeling predicts that CRISPR-based activators, unlike CRISPR-based repressors, scale well with increasing gRNA competition and dCas9 bottlenecking. *bioRxiv* 719278. https://doi.org/10.1101/719278

Cook, T.B., Rand, J.M., Nurani, W., Courtney, D.K., Liu, S.A., Pfleger, B.F., **2018**. Genetic tools for reliable gene expression and recombineering in *Pseudomonas putida*. *J. Ind. Microbiol. Biotechnol.* **45**, 517–527. https://doi.org/10.1007/s10295-017-2001-5

Coppens, L., Lavigne, R., **2020**. SAPPHIRE: a neural network based classifier for σ70 promoter prediction in *Pseudomonas*. *BMC Bioinformatics* **21**, 415. https://doi.org/10.1186/s12859-020-03730-z

D'Arrigo, I., Bojanovič, K., Yang, X., Holm Rau, M., Long, K.S., **2016**. Genome-wide mapping of

transcription start sites yields novel insights into the primary transcriptome of *Pseudomonas putida*. *Environ. Microbiol.* **18**, 3466–3481. https://doi.org/10.1111/1462-2920.13326
Dominguez, A.A., Lim, W.A., Qi, L.S., **2016**. Beyond editing: repurposing CRISPR–Cas9 for precision genome regulation and interrogation. *Nat. Rev. Mol. Cell Biol.* **17**, 5–15. https://doi.org/10.1038/nrm.2015.2
Dong, C., Fontana, J., Patel, A., Carothers, J.M., Zalatan, J.G., **2018**. Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nat Commun* **9**, 2489. https://doi.org/10.1038/s41467-018-04901-6
Duncker, K.E., Holmes, Z.A., You, L., **2021**. Engineered microbial consortia: strategies and applications. *Microb. Cell Factories* **20**, 211. https://doi.org/10.1186/s12934-021-01699-9
Elmore, J.R., Dexter, G.N., Baldino, H., Huenemann, J.D., Francis, R., Peabody, G.L., Martinez-Baird, J., Riley, L.A., Simmons, T., Coleman-Derr, D., Guss, A.M., Egbert, R.G., **2023**. High-throughput genetic engineering of nonmodel and undomesticated bacteria via iterative site-specific genome integration. *Sci. Adv.* **9**, eade1285. https://doi.org/10.1126/sciadv.ade1285
Fatma, Z., Schultz, J.C., Zhao, H., **2020**. Recent advances in domesticating non-model microorganisms. *Biotechnol. Prog.* **36**, e3008. https://doi.org/10.1002/btpr.3008
Fenster, J.A., Werner, A.Z., Tay, J.W., Gillen, M., Schirokauer, L., Hill, N.C., Watson, A., Ramirez, K.J., Johnson, C.W., Beckham, G.T., Cameron, J.C., Eckert, C.A., **2022**. Dynamic and single cell characterization of a CRISPR-interference toolset in *Pseudomonas putida* KT2440 for β-ketoadipate production from p-coumarate. *Metab. Eng. Commun.* **15**, e00204. https://doi.org/10.1016/j.mec.2022.e00204
Fontana, J., Dong, C., Ham, J.Y., Zalatan, J.G., Carothers, J.M., **2018**. Regulated Expression of sgRNAs Tunes CRISPRi in *E. coli*. *Biotechnol J* **13**, e1800069. https://doi.org/10.1002/biot.201800069
Fontana, J., Dong, C., Kiattisewee, C., Chavali, V.P., Tickman, B.I., Carothers, J.M., Zalatan, J.G., **2020a**. Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. *Nat. Commun.* **11**, 1618. https://doi.org/10.1038/s41467-020-15454-y
Fontana, J., Sparkman-Yager, D., Zalatan, J.G., Carothers, J.M., **2020b**. Challenges and opportunities with CRISPR activation in bacteria for data-driven metabolic engineering. *Curr. Opin. Biotechnol.* **64**, 190–198. https://doi.org/10.1016/j.copbio.2020.04.005
Huang, H.-H., Bellato, M., Qian, Y., Cárdenas, P., Pasotti, L., Magni, P., Del Vecchio, D., **2021**. dCas9 regulator to neutralize competition in CRISPRi circuits. *Nat. Commun.* **12**, 1692. https://doi.org/10.1038/s41467-021-21772-6
Hwang, C., Carothers, J.M., **2016**. Label-free selection of RNA aptamers for metabolic engineering. *Vitro Sel. Evol.* **106**, 37–41. https://doi.org/10.1016/j.ymeth.2016.06.016
Intasian, P., Prakinee, K., Phintha, A., Trisrivirat, D., Weeranoppanant, N., Wongnate, T., Chaiyen, P., **2021**. Enzymes, In Vivo Biocatalysis, and Metabolic Engineering for Enabling a Circular Economy and Sustainability. *Chem. Rev.* **121**, 10367–10451. https://doi.org/10.1021/acs.chemrev.1c00121
Juminaga Darmawi, Baidoo Edward E. K., Redding-Johanson Alyssa M., Batth Tanveer S., Burd Helcio, Mukhopadhyay Aindrila, Petzold Christopher J., Keasling Jay D., **2012**. Modular Engineering of l-Tyrosine Production in *Escherichia coli*. *Appl. Environ. Microbiol.* **78**, 89–98. https://doi.org/10.1128/AEM.06017-11
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*. *Metab. Eng.* **66**, 283–295. https://doi.org/10.1016/j.ymben.2021.04.002
Konishi, K., Takaya, N., Masuo, S., Zhou, S., **2018**. 4-amino cinnamic acid production method

using enzyme.
Kuchina, A., Brettner, L.M., Paleologu, L., Roco, C.M., Rosenberg, A.B., Carignano, A., Kibler, R., Hirano, M., DePaolo, R.W., Seelig, G., **2021**. Microbial single-cell RNA sequencing by split-pool barcoding. *Science* **371**, eaba5257. https://doi.org/10.1126/science.aba5257
Lawson, C.E., Martí, J.M., Radivojevic, T., Jonnalagadda, S.V.R., Gentz, R., Hillson, N.J., Peisert, S., Kim, J., Simmons, B.A., Petzold, C.J., Singer, S.W., Mukhopadhyay, A., Tanjore, D., Dunn, J.G., Garcia Martin, H., **2021**. Machine learning for metabolic engineering: A review. *Tools Strateg. Metab. Eng.* **63**, 34–60. https://doi.org/10.1016/j.ymben.2020.10.005
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., **2019a**. A comprehensive metabolic map for production of bio-based chemicals. *Nat. Catal.* **2**, 18–33. https://doi.org/10.1038/s41929-018-0212-4
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., **2019b**. A comprehensive metabolic map for production of bio-based chemicals. *Nat. Catal.* **2**, 18–33. https://doi.org/10.1038/s41929-018-0212-4
Liu, X., Niu, H., Li, Q., Gu, P., **2019**. Metabolic engineering for the production of l-phenylalanine in *Escherichia coli*. *3 Biotech* **9**, 85. https://doi.org/10.1007/s13205-019-1619-6
Luo, Z.W., Lee, S.Y., **2020**. Metabolic engineering of *Escherichia coli* for the production of benzoic acid from glucose. *Metab. Eng.* **62**, 298–311. https://doi.org/10.1016/j.ymben.2020.10.002
Masuo, S., Zhou, S., Kaneko, T., Takaya, N., **2016**. Bacterial fermentation platform for producing artificial aromatic amines. *Sci. Rep.* **6**, 25764. https://doi.org/10.1038/srep25764
Minakawa, H., Masuo, S., Kaneko, T., Takaya, N., **2019**. Fermentation and purification of microbial monomer 4-amminocinnamic acid to produce ultra-high performance bioplastics. *Process Biochem.* **77**, 100–105. https://doi.org/10.1016/j.procbio.2018.11.021
Mohammadi Nargesi, B., Sprenger, G.A., Youn, J.-W., **2019**. Metabolic Engineering of *Escherichia coli* for para-Amino-Phenylethanol and para-Amino-Phenylacetic Acid Biosynthesis. *Front. Bioeng. Biotechnol.* **6**.
Nielsen, J., Keasling, J.D., **2016**. Engineering Cellular Metabolism. *Cell* **164**, 1185–1197. https://doi.org/10.1016/j.cell.2016.02.004
O'Rourke, P.E.F., Eadsforth, T.C., Fyfe, P.K., Shepherd, S.M., Hunter, W.N., **2011**. *Pseudomonas aeruginosa* 4-Amino-4-Deoxychorismate Lyase: Spatial Conservation of an Active Site Tyrosine and Classification of Two Types of Enzyme. *PLOS ONE* **6**, e24158. https://doi.org/10.1371/journal.pone.0024158
Peters, J.M., Colavin, A., Shi, H., Czarny, T.L., Larson, M.H., Wong, S., Hawkins, J.S., Lu, C.H.S., Koo, B.-M., Marta, E., Shiver, A.L., Whitehead, E.H., Weissman, J.S., Brown, E.D., Qi, L.S., Huang, K.C., Gross, C.A., **2016**. A Comprehensive, CRISPR-based Functional Analysis of Essential Genes in Bacteria. *Cell* **165**, 1493–1506. https://doi.org/10.1016/j.cell.2016.05.003
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**, 1173–83. https://doi.org/10.1016/j.cell.2013.02.022
Radivojević, T., Costello, Z., Workman, K., Garcia Martin, H., **2020**. A machine learning Automated Recommendation Tool for synthetic biology. *Nat. Commun.* **11**, 4879. https://doi.org/10.1038/s41467-020-18008-4
Roy, S., Radivojevic, T., Forrer, M., Marti, J.M., Jonnalagadda, V., Backman, T., Morrell, W., Plahar, H., Kim, J., Hillson, N., Garcia Martin, H., **2021**. Multiomics Data Collection, Visualization, and Utilization for Guiding Metabolic Engineering. *Front. Bioeng.*

Biotechnol. **9**.
Santos-Moreno, J., Tasiudi, E., Stelling, J., Schaerli, Y., **2020**. Multistable and dynamic CRISPRi-based synthetic circuits. *Nat. Commun.* **11**, 2746. https://doi.org/10.1038/s41467-020-16574-1
Sheth, R.U., Wang, H.H., **2018**. DNA-based memory devices for recording cellular events. *Nat. Rev. Genet.* **19**, 718–732. https://doi.org/10.1038/s41576-018-0052-8
Shin, J., Porubsky, V., Carothers, J., Sauro, H.M., **2023**. Standards, dissemination, and best practices in systems biology. *Curr. Opin. Biotechnol.* **81**, 102922. https://doi.org/10.1016/j.copbio.2023.102922
Stevens, J.T., Carothers, J.M., **2015**. Designing RNA-Based Genetic Control Systems for Efficient Production from Engineered Metabolic Pathways. *ACS Synth. Biol.* **4**, 107–115. https://doi.org/10.1021/sb400201u
Suvannasara, P., Tateyama, S., Miyasato, A., Matsumura, K., Shimoda, T., Ito, T., Yamagata, Y., Fujita, T., Takaya, N., Kaneko, T., **2014**. Biobased Polyimides from 4-Aminocinnamic Acid Photodimer. *Macromolecules* **47**, 1586–1593. https://doi.org/10.1021/ma402499m
Tickman, B.I., Burbano, D.A., Chavali, V.P., Kiattisewee, C., Fontana, J., Khakimzhan, A., Noireaux, V., Zalatan, J.G., Carothers, J.M., **2021**. Multi-layer CRISPRa/i circuits for dynamic genetic programs in cell-free and bacterial systems. *Cell Syst.* S2405471221004191. https://doi.org/10.1016/j.cels.2021.10.008
Todor, H., Silvis, M.R., Osadnik, H., Gross, C.A., **2021**. Bacterial CRISPR screens for gene function. *Curr. Opin. Microbiol.* **59**, 102–109. https://doi.org/10.1016/j.mib.2020.11.005
Tokic, M., Hatzimanikatis, V., Miskovic, L., **2020**. Large-scale kinetic metabolic models of *Pseudomonas putida* KT2440 for consistent design of metabolic engineering strategies. *Biotechnol. Biofuels* **13**, 33. https://doi.org/10.1186/s13068-020-1665-7
Vigouroux, A., Bikard, D., **2020**. CRISPR Tools To Control Gene Expression in Bacteria. *Microbiol. Mol. Biol. Rev.* **84**, e00077-19. https://doi.org/10.1128/mmbr.00077-19
Wang, G., Zhao, Z., Ke, J., Engel, Y., Shi, Y.-M., Robinson, D., Bingol, K., Zhang, Z., Bowen, B., Louie, K., Wang, B., Evans, R., Miyamoto, Y., Cheng, K., Kosina, S., De Raad, M., Silva, L., Luhrs, A., Lubbe, A., Hoyt, D.W., Francavilla, C., Otani, H., Deutsch, S., Washton, N.M., Rubin, E.M., Mouncey, N.J., Visel, A., Northen, T., Cheng, J.-F., Bode, H.B., Yoshikuni, Y., **2019**. CRAGE enables rapid activation of biosynthetic gene clusters in undomesticated bacteria. *Nat. Microbiol.* **4**, 2498–2510. https://doi.org/10.1038/s41564-019-0573-8
Wang, T., Tague, N., Whelan, S.A., Dunlop, M.J., **2021**. Programmable gene regulation for metabolic engineering using decoy transcription factor binding sites. *Nucleic Acids Res.* **49**, 1163–1172. https://doi.org/10.1093/nar/gkaa1234
Yan, Q., Fong, S.S., **2017**. Challenges and Advances for Genetic Engineering of Non-model Bacteria and Uses in Consolidated Bioprocessing. *Front. Microbiol.* **8**. https://doi.org/10.3389/fmicb.2017.02060
Zaslaver, A., Bren, A., Ronen, M., Itzkovitz, S., Kikoin, I., Shavit, S., Liebermeister, W., Surette, M.G., Alon, U., **2006**. A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nat. Methods* **3**, 623–628. https://doi.org/10.1038/nmeth895
Zhang, S., Voigt, C.A., **2018**. Engineered dCas9 with reduced toxicity in bacteria: implications for genetic circuit design. *Nucleic Acids Res.* **46**, 11115–11125. https://doi.org/10.1093/nar/gky884

Supplementary Information

Supplementary Figures

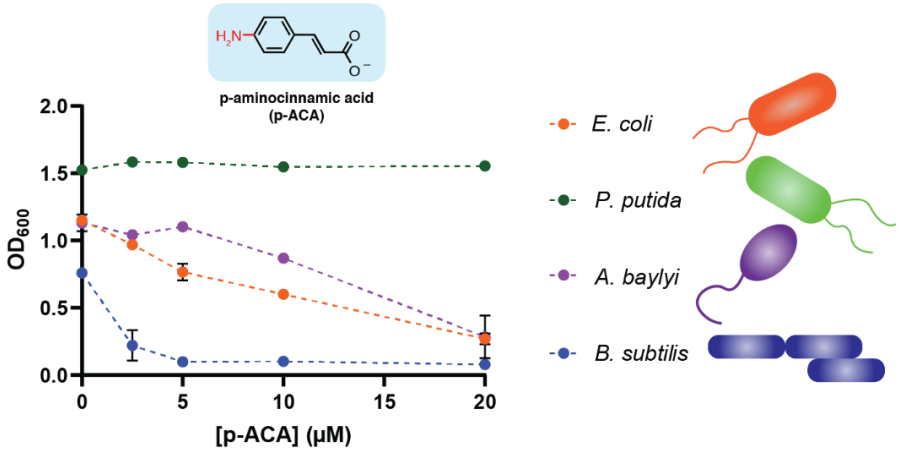


Figure 4.S1: Tolerance of p-ACA of different bacteria

Four selected microbes with industrially-relevant applications — *E. coli* MG1655, *P. putida* KT2440, *A. baylyi* ADP1, and *B. subtilis* strain 168 — were subjected to p-ACA tolerance test in EZ-RDM. It was found that *P. putida* exhibited high tolerance even at the highest concentration tested (20 mM) where *A. baylyi* demonstrated higher tolerance compared to that of *E. coli*.

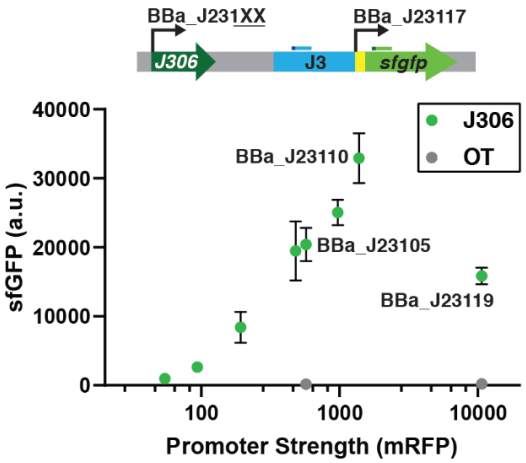
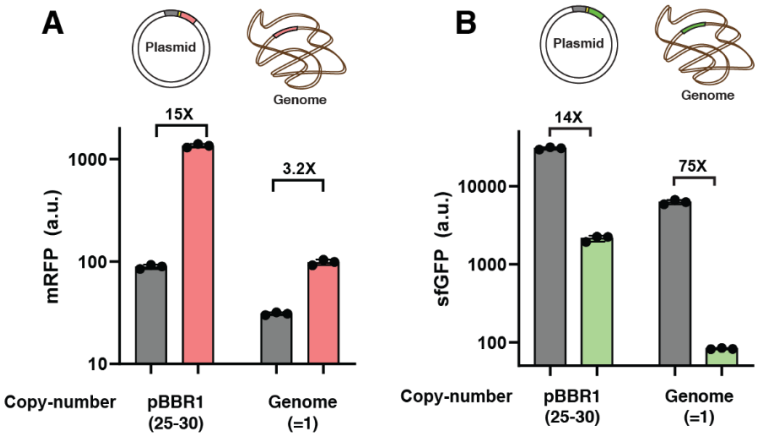


Figure 4.S2: CRISPRa with different scRNA expression level

Different expression levels of scRNA (8 Anderson promoters, BBa_J231XX, http://parts.igem.org) were investigated in *P. putida* for effective CRISPRa where dCas9 and MCP-SoxS were expressed from a single-copy genomic DNA. We found that BBa_J23110 yielded the highest CRISPR-mediated gene expression and fold-activation.



Data retrieved from Kiattisewee et al. 2021 comparing between pCK305 vs CKPP037

Figure 4.S3: CRISPRa and CRISPRi efficiency depends on Copy-Number

CRISPRa and CRISPRi mediated gene activation and repression were compared from the plasmid-bourne expression (pBBR1-GmR, copy-number = 25-30 copies, Cook et al., 2018) and genome-integrated variant (CKPP037). It was found that CRISPRa in *P. putida* performed better in the plasmid system while CRISPRi is more powerful when targeting the single-copy genomic locus. The data were retrieved from Kiattisewee et al.

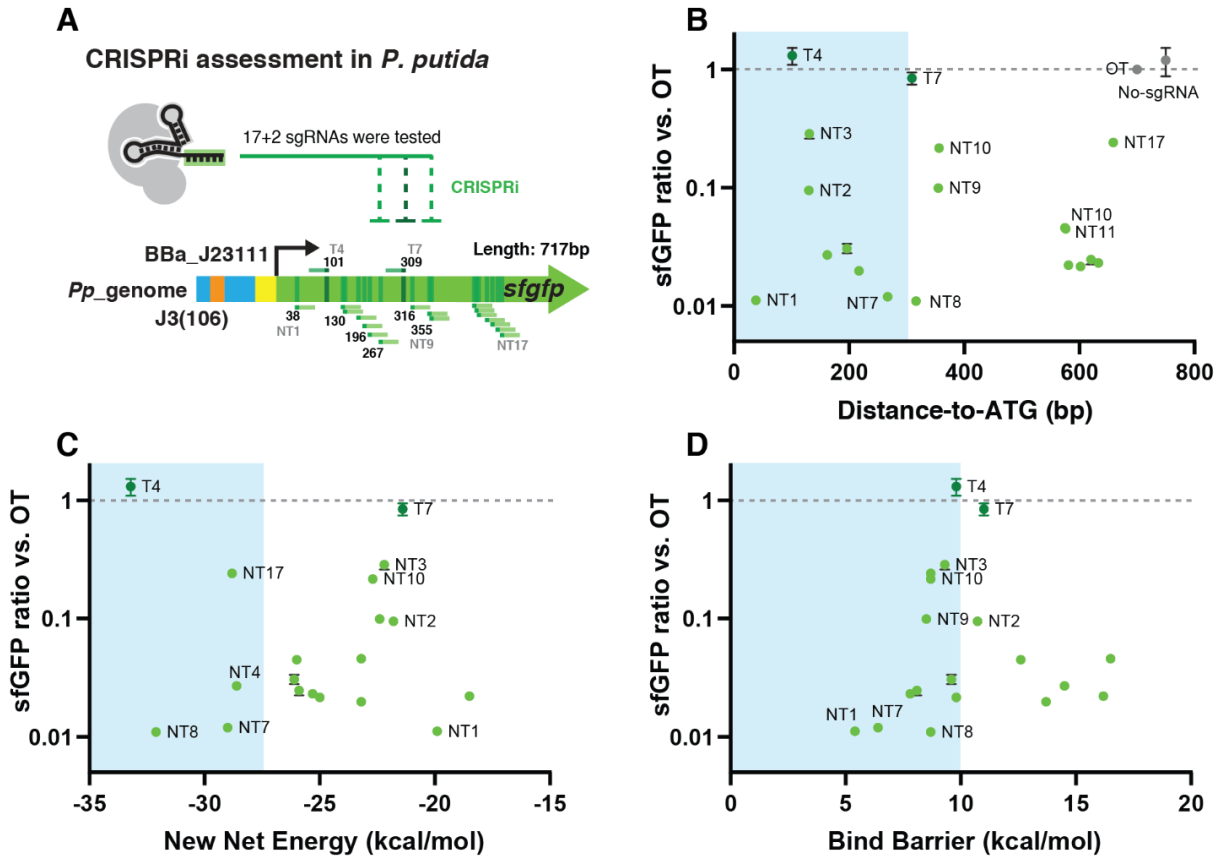


Figure 4.S4: CRISPRi investigation at different distance and folding energetics

(A) CRISPRi was assessed in *P. putida* with 17 different sgRNAs targeting a non-template strand and 2 sgRNAs targeting a template strand. (B) We found that distance can generally suggest high-performance CRISPRi sgRNA where 3 out of 5 highest performance sgRNAs fall into the first 300bp region of ORF. (C & D) We analyzed the RNA folding energetics of all sgRNAs and found that 2 out of 5 of the highest performance sgRNAs have less than -27.5 kcal/mol of New Net Energy. For Bind Barrier, we found that 4 out of 5 highest performance sgRNAs have Bind Barrier lower than 10 kcal/mol suggesting the strong correlation between low Bind Barrier and performance of sgRNA. We further used these three criteria for sgRNA selection of CRISPRi in the genome-scale perturbation characterization.

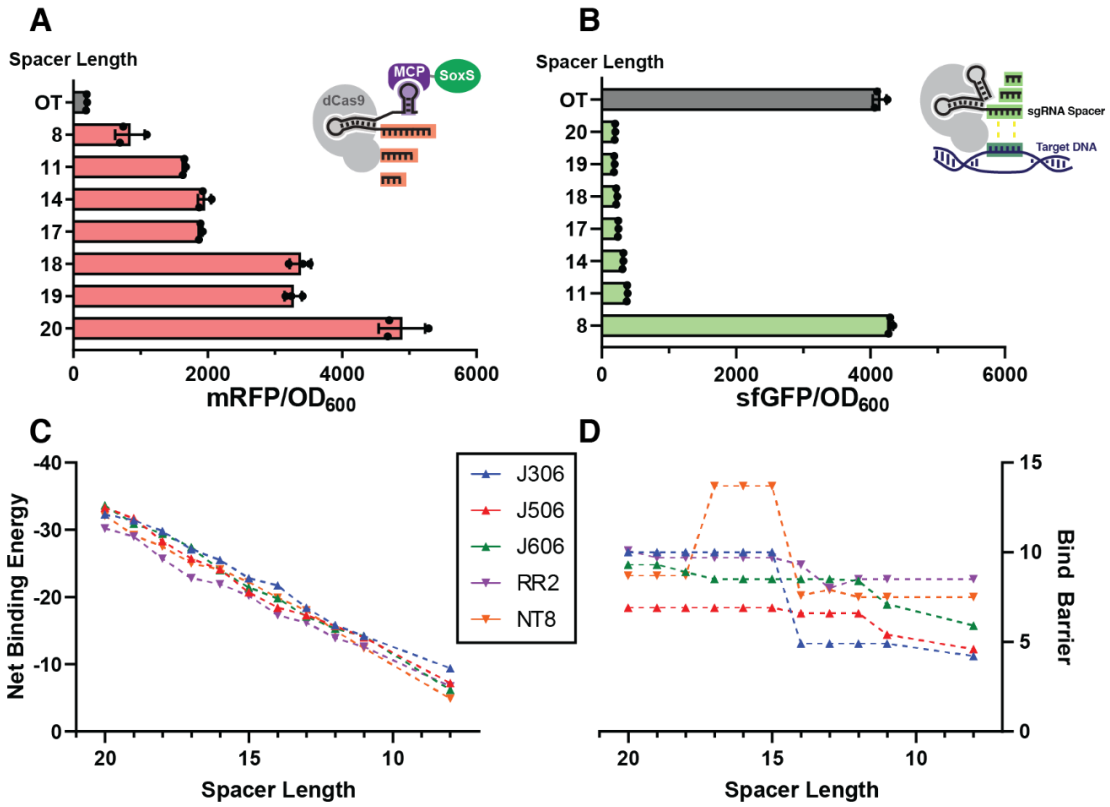


Figure 4.S5: scRNA and sgRNA truncation for CRISPRa/i gene expression tuning

(A & B) gRNA spacers were truncated to reduce the CRISPRa and CRISPRi activities, respectively. For CRISPRa, J306 was chosen for truncation analysis while sfGFP_NT8 was chosen for CRISPRi investigation. (C & D) Truncation of gRNA usually leads to reduced binding energy as shown in the Net Binding Energy plots of truncated spacers analyzed by our proprietary method (Appendix H).

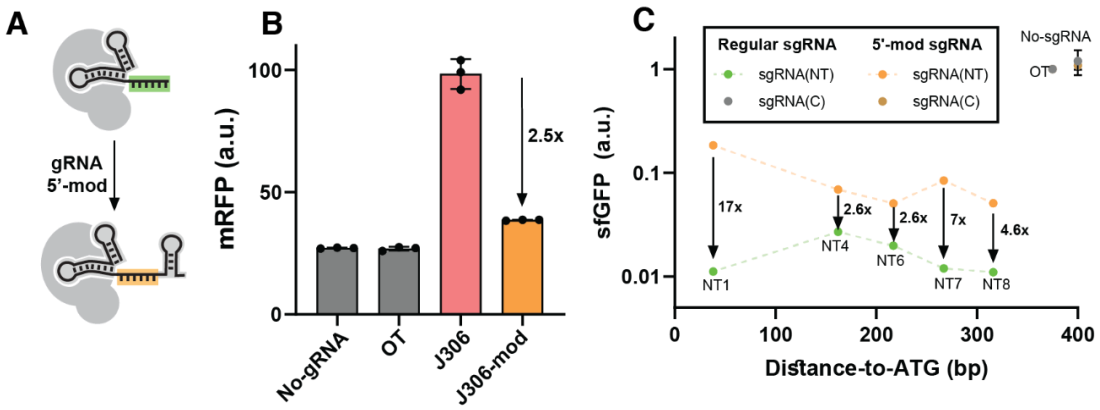


Figure 4.S6: gRNA modification impair CRISPRa and CRISPRi efficiency

(A) gRNA was modified by addition of a hairpin to the 5'-end next to the spacer. (B) The modified J306 scRNA exhibited impaired activity when tested in CKPP037 with J3-BBa_J23117-mRFP integrated into the *P. putida* genome. (C) Five different sgRNAs targeting sfGFP were also investigated in CKPP037 with J3(106)-BBa_J23111-sfGFP and found that CRISPRi-based repression also suffered from gRNA modification. It is possible that additional bases behave similar to mismatched bases and thus impair RNA-DNA binding in the CRISPR-Cas/DNA/RNA ternary complex (Todor et al., 2021).

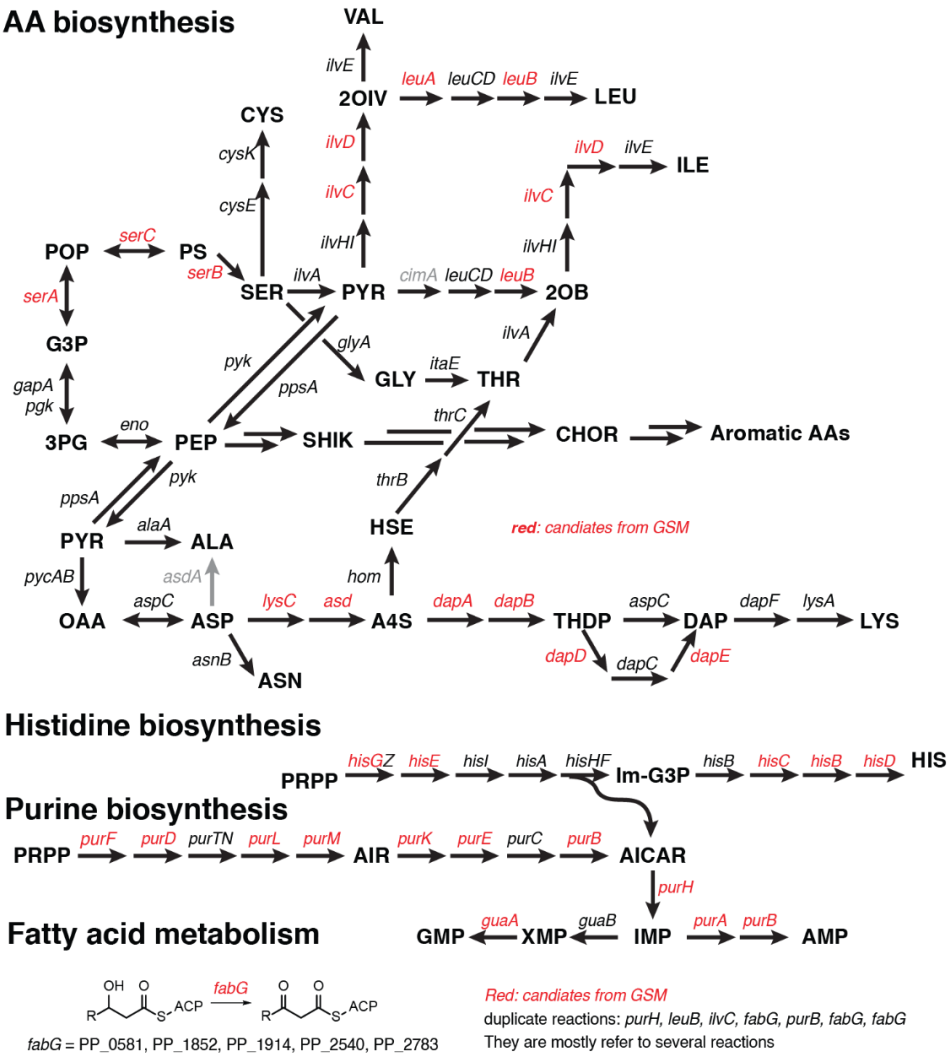


Figure 4.S7: CRISPRi targets for gene repression suggested from metabolic model

Gene candidates for repression as suggested from the metabolic model (Flux-RETAP) were labeled in red. Most candidates are distant to aromatic amino acid biosynthesis (shikimate and chorismate pathways) and fall under the general anabolic pathway of amino acid biosynthesis, nucleic acid biosynthesis, and fatty acid metabolism.

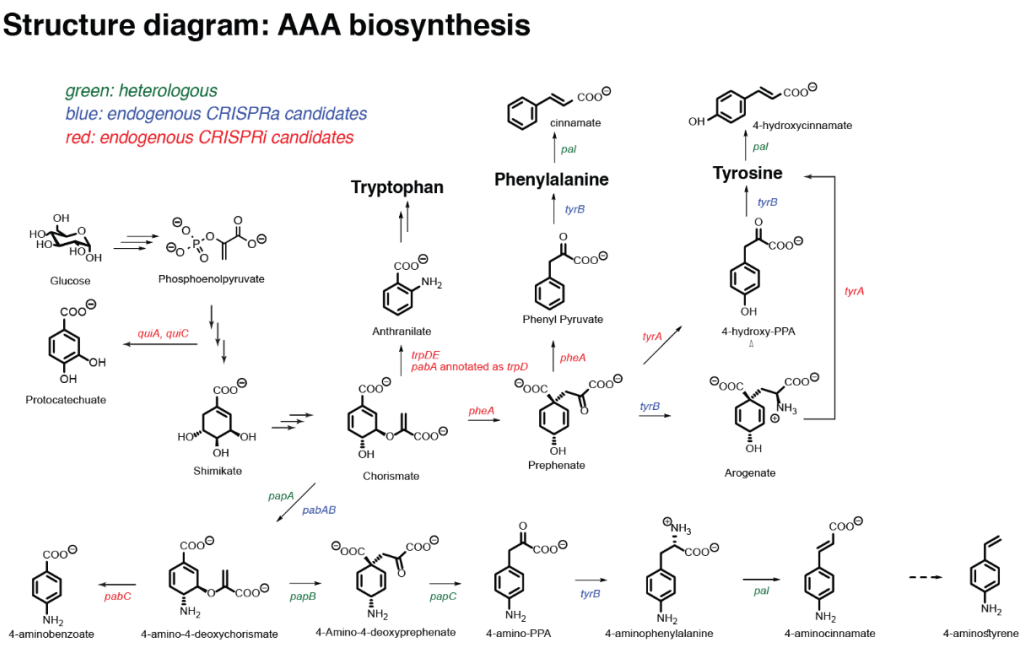


Figure 4.S8: CRISPRi targets suggested from competition in aromatic amino acid biosynthetic pathway

Gene targets previously known to affect shikimate and chorismate pathways were identified and included to the list of CRISPRi gene repression candidates. These gene candidates are likely pulling the flux from the key intermediates including shikimate biosynthesis (*quiA*/*quiC*), tryptophan biosynthesis (*trpDE*/*trpA* from chorismate), phenylalanine and tyrosine biosynthesis (*pheA*/*tyrA* from chorismate), and *pabC* which used an intermediate (4-amino-4-deoxychorismate) from *papA* reaction.

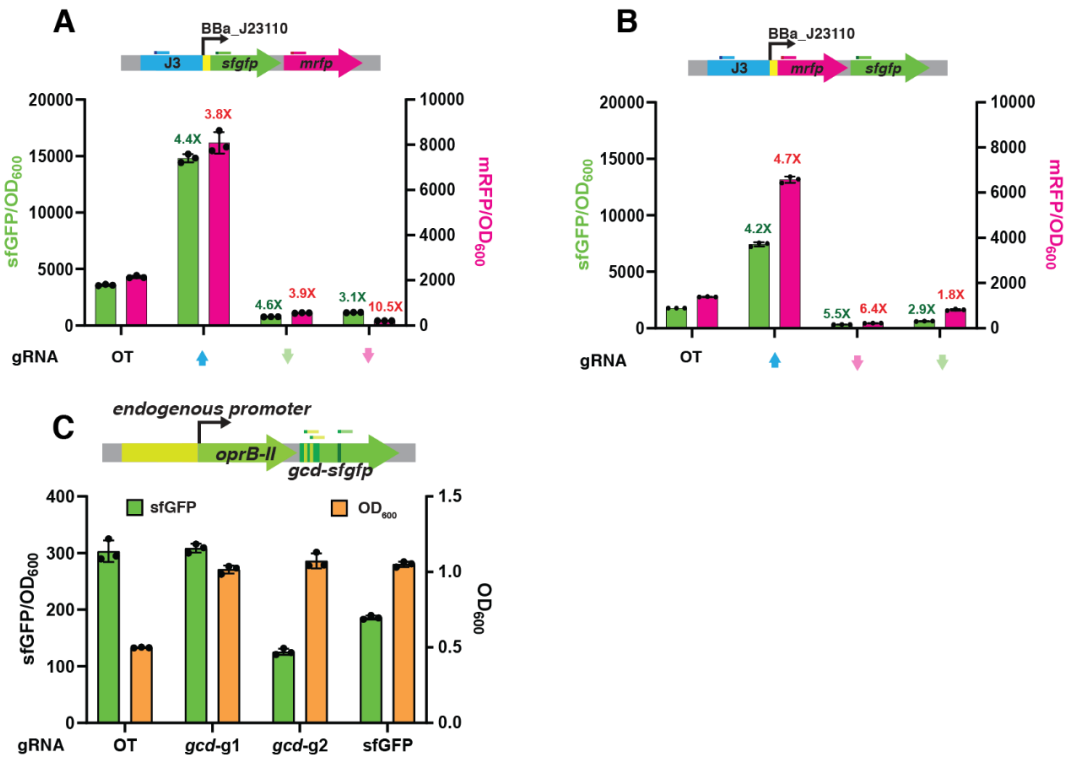


Figure 4.S9: Polar effect on transcriptional operon by CRISPRa/i

(A & B) Polar effect on different genes in the operon were evaluated using BBa_J23110 driving sfGFP-mRFP or mRFP-sfGFP operons under control of J3 synthetic promoters. CRISPRa has similar effects to both genes regardless of the orientation while CRISPRi affects the target gene more than the adjacent gene. (C) Endogenous promoter with *oprB-II-gcd* sfGFP fusion reporter was also evaluated where targeting either gcd or sfGFP ORFs led to alleviation of growth defect and expression of sfGFP.

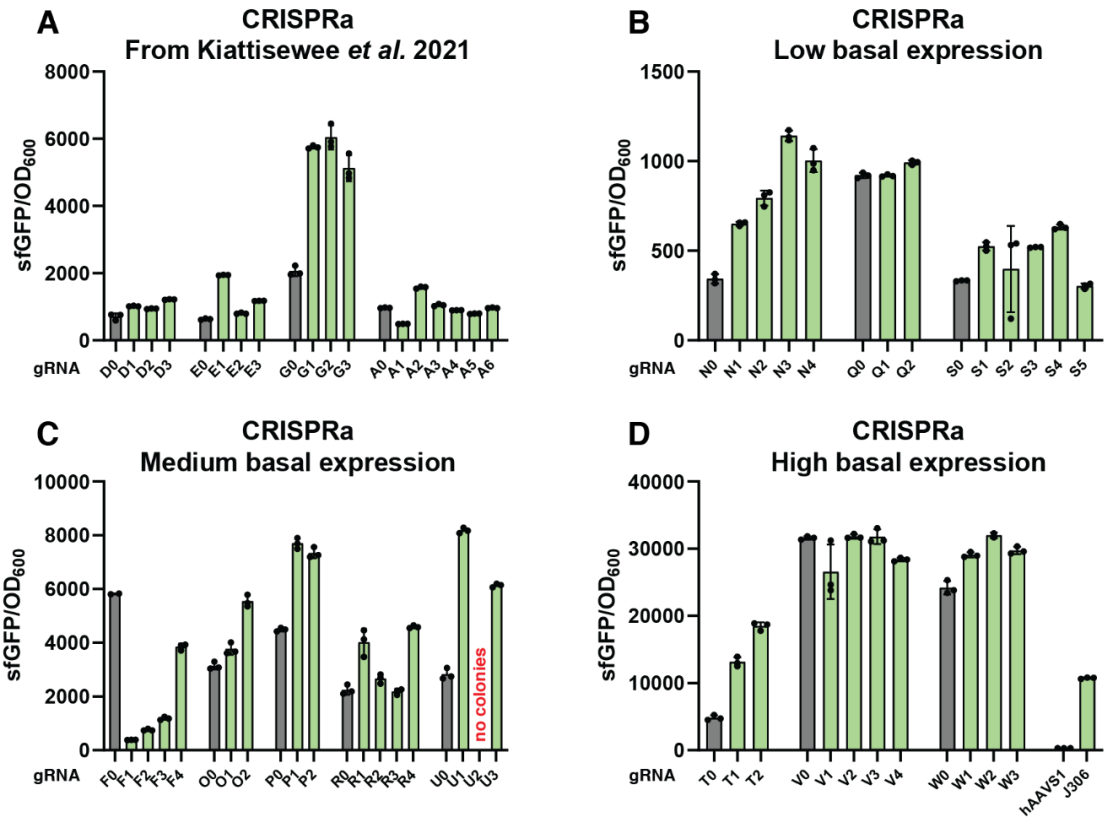


Figure 4.S10: CRISPRa evaluation of endogenous genes

CRISPRa evaluation of endogenous promoters using a synthetic sfGFP fusion strategy. (A) Previous genes studied in the mRFP context (Kiattisewee et al., 2021) were first evaluated and we found similar effects in the sfGFP context. (B-D) One previous promoter and ten additional promoters were evaluated in a similar manner. Plots were grouped by basal expression level — low, medium, and high. J3 synthetic reporter was included in panel D.

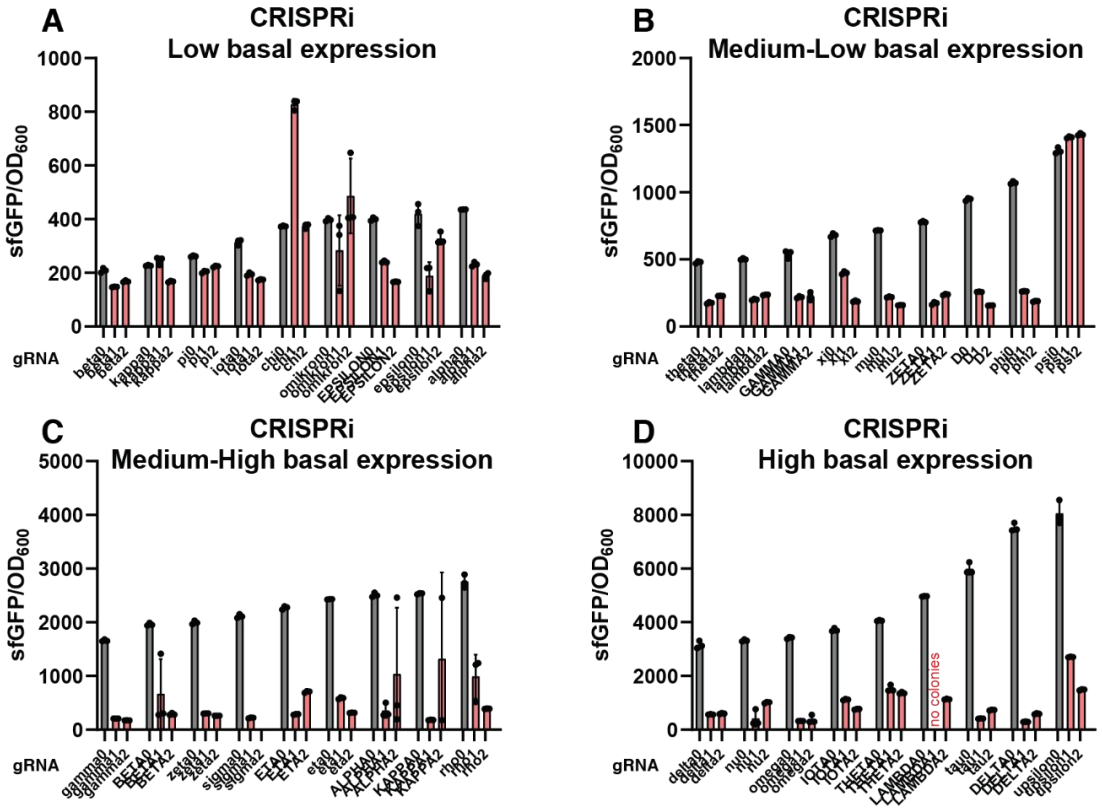


Figure 4.S11: CRISPRi evaluation of endogenous genes

CRISPRi evaluation of endogenous promoters using a synthetic sfGFP fusion strategy. (A-D) 36 promoters were evaluated with 2 sgRNAs for each promoter. Plots were grouped by basal expression level — low, medium-low, medium-high, and high.

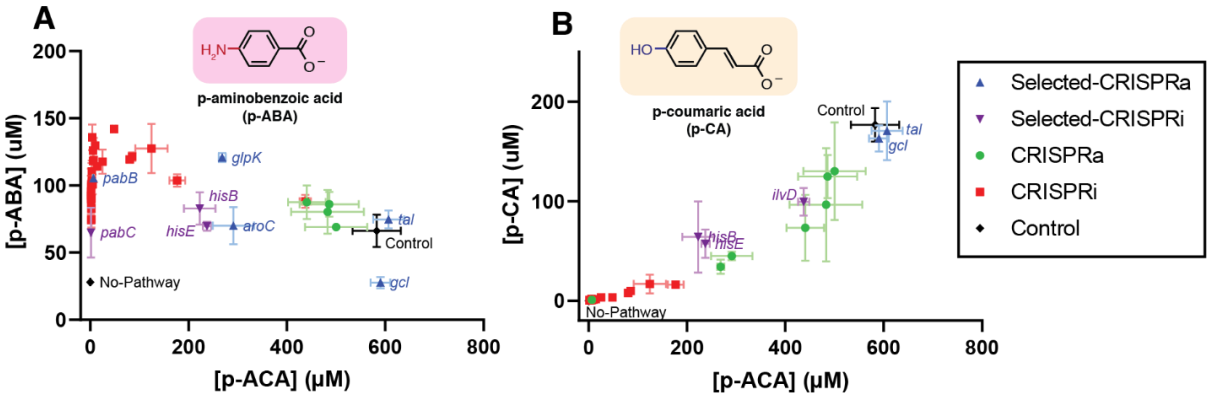


Figure 4.S12: Ratio of p-ACA with other aromatic acids

Other aromatic acids were investigated alongside p-ACA in the supernatants of CRISPRa/i-perturbing strains. (A) It was shown that *gcl* CRISPRa strain led to decrease in p-ABA production close to the no-pathway condition. CIRPSRa of *pabC* significantly impair p-ACA production likely due to pulling the flux to p-ABA synthesis (folate biosynthetic pathway). On the other hand, p-CA which is likely accumulated from the reaction of Pal with tyrosine is linearly correlated with concentration of p-ACA, converted from p-AF. The concentration of p-AF is negligible in these experiments and t-CA concentrations are always high at a similar level (not shown here).

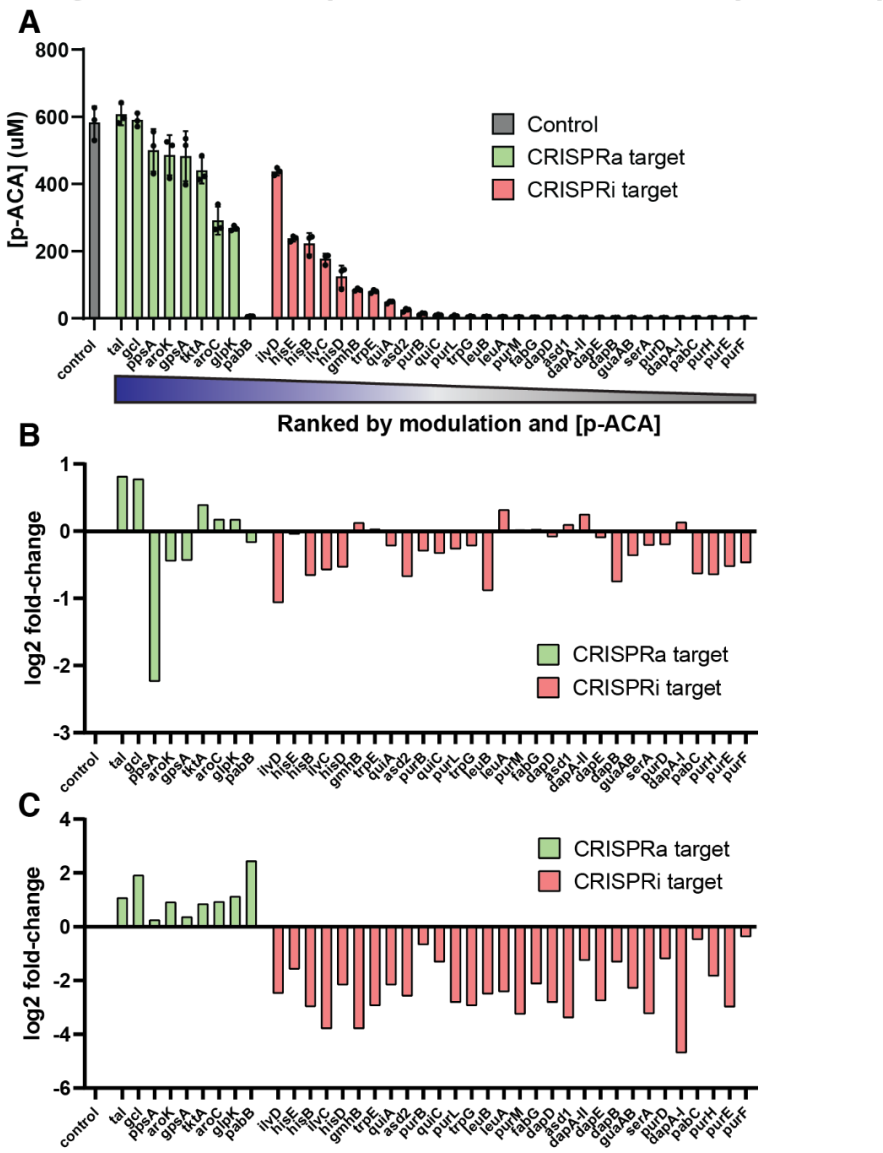


Figure 4.S13: Direct capture of CRISPRa/i effect from RNAseq

According to transcriptomic analysis, the fold-change of target genes in RNA sequencing (B) were compared to that of fold-change from synthetic promoter experiment (C) and ranked by the p-ACA productivity (A). It is clear that a lot of perturbations lead to undesired effects unlike their synthetic reporter counterparts.

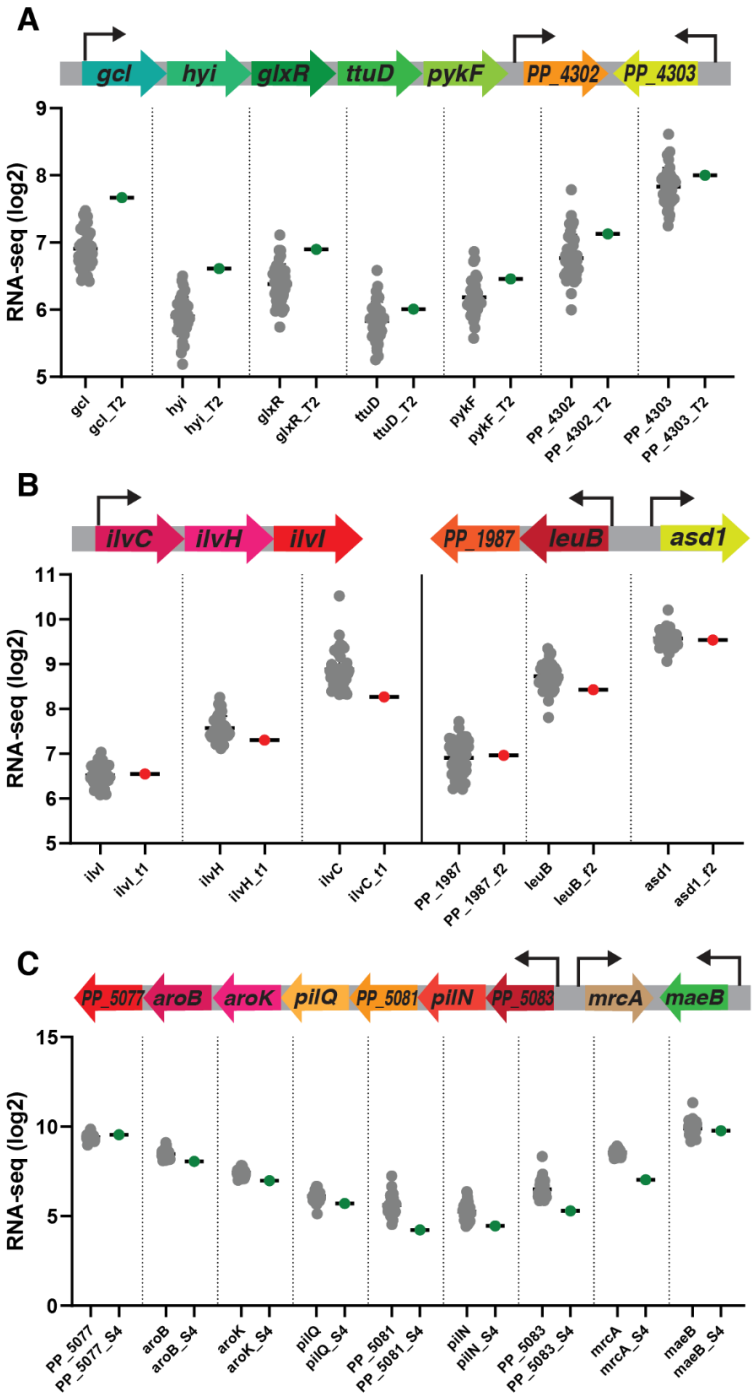


Figure 4.S14: Operon effects from endogenous CRISPRa/i perturbations

Transcriptomic analysis illustrates the polar effect of CRISPRa/i at different endogenous operons. (A) *gcl*-operon was investigated where its downstream genes were upregulated when the *gcl* promoter was upregulated with CRISPRa. The downstream gene (PP_4302), despite having its own promoter, also has upregulation effects. (B) CRISPRi evaluation of *ilvC*-operon (left) and *lueB*-operon (right). It is evident that CRISPRi toward *ilvC* also affects downstream genes in the operon while CRISPRi to the *leuB* operon has no effect on its downstream gene

(PP_1987) and its adjacent gene (*ads1*). (C) *aroKB* repressing CRISPRa is also shown here as an example of the opposite effect. Most genes in the aroKB operon are downregulated as a result of CRISPRa scRNA (S4). The adjacent gene, *mrcA*, is also affected despite facing a different direction suggesting that CRISPRa scRNA might be interfering with the native regulation or cryptic *mrcA* promoter.

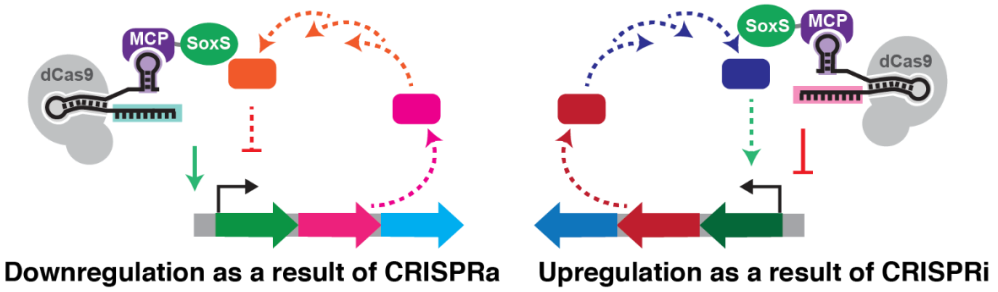


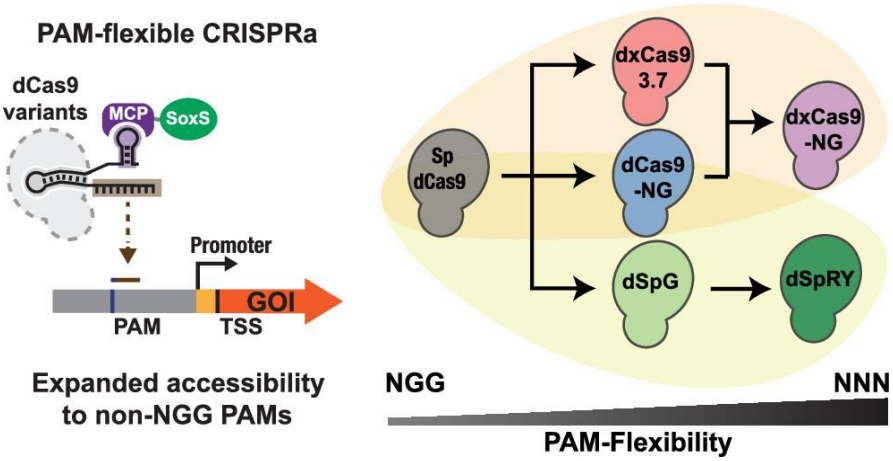
Figure 4.S15: Plausible feedback response of endogenous gene perturbations

The schematic describes plausible feedback responses from CRISPRa (left) or CRISPRi (right). Upregulation of a certain gene could result in a change in concentration of another protein which provides feedback response to the perturbing operon and thus lead to downregulation at the snapshot of mRNA collection instead of initial upregulation. Similar effect could occur in a similar way where cellular response is pushing the expression level back to the normal level after suffering from gene repression in the earlier stages.

Chapter 5:

Expanding the scope of bacterial CRISPR activation with PAM-flexible dCas9 variants

Cholpisit Kiattisewee[+], Ava V. Karanjia[+], Mateusz Legut, Zharko Daniloski, Samantha E. Koplik, Joely Nelson, Benjamin P. Kleinstiver, Neville E. Sanjana, James M. Carothers*, and Jesse G. Zalatan*



Abstract

CRISPR-Cas transcriptional tools have been widely applied for programmable regulation of complex biological networks. In comparison to eukaryotic systems, bacterial CRISPR activation (CRISPRa) has stringent target site requirements for effective gene activation. While genes may not always have an NGG protospacer adjacent motif (PAM) at the appropriate position, PAM-flexible dCas9 variants can expand the range of targetable sites. Here we systematically evaluate a panel of PAM-flexible dCas9 variants for their ability to activate bacterial genes. We observe that dxCas9-NG provides a high dynamic range of gene activation for sites with NGN PAMs while dSpRY permits modest activity across almost any PAM. Similar trends were observed for heterologous and endogenous promoters. For all variants tested, improved PAM-flexibility comes with the trade-off that CRISPRi-mediated gene repression becomes less effective. Weaker CRISPR interference (CRISPRi) gene repression can be partially rescued by expressing multiple sgRNAs to target many sites in the gene of interest. Our work provides a framework to choose the most effective dCas9 variant for a given set of gene targets, which will further expand the utility of CRISPRa/i gene regulation in bacterial systems.

This work is published in ACS Synthetic Biology on November 15th, 2022
DOI: https://doi.org/10.1021/acssynbio.2c00405

5.1 Introduction

CRISPR-Cas transcriptional regulation enables programmable control over gene activation and repression in bacteria (Bikard et al., 2013; Dong et al., 2018; Qi et al., 2013). These tools can be used to regulate endogenous gene networks both to probe biological function and to engineer new behaviors. However, effective CRISPR activation (CRISPRa) is limited by complex and stringent target site requirements (Fontana et al., 2020a; Ho et al., 2020; Liu et al., 2019; Villegas Kcam et al., 2021). Previous work has demonstrated a sharply periodic pattern of effective target sites occurring every 10 bp within a 60-100 bp region upstream of the TSS (Dong et al., 2018; Fontana et al., 2020a, 2020b; Kiattisewee et al., 2021). Targeting these sites with *S. pyogenes* dCas9 requires a compatible protospacer adjacent motif (PAM) at the right position, and endogenous genes that lack an appropriate PAM may be incompatible for activation via Sp-Cas9.

We previously demonstrated that one of the first expanded PAM dCas9 variants, dxCas9(3.7) (Hu et al., 2018), could activate some genes that lacked an appropriately-positioned NGG PAM (Fontana et al., 2020a). Specifically, dxCas9(3.7) produced at least two-fold increases in gene expression at three out of seven endogenous promoters tested (Fontana et al., 2020a). Although these results were encouraging, all seven of the candidate promoters were predicted to encode PAMs compatible with dxCas9(3.7) and all seven were expected to be activated. Since our initial work with dxCas9(3.7), several new expanded PAM dCas9 variants have been described,(Legut et al., 2020; Nishimasu et al., 2018; Walton et al., 2020) raising the possibility that improved variants could further expand the number of targetable genes for CRISPRa in bacteria.

In this work, we demonstrate that PAM-flexible dCas9 variants can improve transcriptional activation at endogenous genes compared to both dCas9 and dxCas9(3.7). We observe activation at previously inaccessible gene targets, and we observe a tradeoff between fold-activation and PAM flexibility. We also demonstrate that expanded PAM dCas9 variants are partially impaired for CRISPR interference (CRISPRi) gene repression. This effect can be mitigated by targeting multiple CRISPR complexes to the desired gene. By systematically characterizing the properties of PAM-flexible dCas9 variants in bacterial CRISPRa/i, we provide a framework to choose the most effective variant for a given gene target or set of targets. This toolbox of dCas9 variants will further expand the utility of CRISPRa/i in bacterial systems for a broad range of applications including metabolic engineering and genome-wide functional screens.

5.2 Results

In-silico PAM availability analysis to predict effective CRISPRa sites

In bacterial systems, effective CRISPRa requires a target site that is precisely positioned upstream of the gene of interest (Fontana et al., 2020a). Consequently, the ability to activate an arbitrary gene is dependent on an appropriately positioned PAM at the desired target site. The widely-used *S. pyogenes* Cas9 (Sp-Cas9) is generally limited to NGG PAMs, but several engineered Cas9 variants have been developed with more flexibility to accommodate various PAMs. We examined two groups of engineered Cas9 variants (Figure 5.1A). The first group includes xCas9-NG, a variant that exhibits high nuclease efficiency and CRISPRa/i performance at NGN PAMs in mammalian systems.(Hu et al., 2018; Legut et al., 2020; Nishimasu et al., 2018) The second group includes SpG and SpRY (Kleinstiver et al., 2016, 2015; Nishimasu et al., 2018; Walton et al., 2021, 2020). The SpRY variant was engineered to be near-PAMless with some preference for NRN PAMs.

We previously identified a set of four precisely-positioned sites upstream of the TSS that are the most effective for CRISPRa with the SoxS activator (positions -70 and -80 for the template strand and -81 and -91 for non-template strand) (Figure 5.1B) (Fontana et al., 2020a). We performed an *in-silico* analysis to identify promoters with accessible PAMs at one or more of the optimal upstream positions in *Escherichia coli* and *Pseudomonas putida*. We restricted the search to promoters with high-confidence TSS positions (Santos-Zavaleta et al., 2019) and with sigma factors previously identified to be effective for CRISPRa with the SoxS activator (Fontana et al., 2020a). Together these criteria were met for 1265 out of 4042 promoters in *E. coli*. PAM compatibility for each Cas9 variant was predicted based on the reported nuclease activity with variable PAM targets (See Methods and Supplementary Methods) (Kim et al., 2020; Legut et al., 2020; Walton et al., 2020). We performed the analysis with 4 deactivated Cas9 variants (dCas9, dxCas9(3.7), dCas9-NG, and dSpRY) of different PAM-flexibility (Table 5.S4). We expect the predictions for dCas9-NG to be representative for dxCas9-NG. Consistent with this expectation, we found that xCas9-NG and Cas9-NG exhibit comparable levels of PAM flexibility in mammalian cell assays for nuclease activity and CRISPRa (Figure 5.S2). The number of promoters with at least one PAM at a suitable position was 47% for dCas9, and increased to almost all promoters for the engineered variants (89% and 93% for dCas9-NG and dSpRY, respectively) (Figure 5.1C). A similar trend was observed in *P. putida* (Figure 5.S3). It is important to note that these predictions are based on nuclease activity in mammalian cells and may not accurately predict bacterial CRISPRa activity, since binding and cleavage determinants may differ.

CRISPRa on non-NGG PAM is improved with engineered dCas9 variants

To test whether PAM-flexible Cas9 variants can increase the number of available CRISPRa target sites, we constructed expression cassettes for the catalytically-inactive versions of each Cas9 variant (dCas9, dxCas9(3.7), dCas9-NG, dxCas9-NG, dSpG, and dSpRY). Our bacterial CRISPRa system uses dCas9 and a modified guide RNA (termed scaffold RNA, scRNA) with an MS2 hairpin to recruit the MCP-SoxS activator (Figure 5,2A) (Dong et al., 2018; Fontana et al., 2020a). We first determined whether each dCas9 variant was effective for CRISPRa at a canonical AGG PAM target site. We used a previously-described reporter gene (J3-BBa_J23117-mRFP) with an AGG PAM positioned at -81 relative to the TSS (Fontana et al., 2020a)). All tested

dCas9 variants exhibited similar activation (~40-50-fold) with the canonical AGG PAM (Figure 5.2). We also evaluated the growth burden associated with each dCas9 variant because dCas9 expression can cause growth defects (Cho et al., 2018; Zhang and Voigt, 2018), and PAM-flexible variants can potentially bind the genome non-specifically at many more sites than the parent dCas9 (Anders et al., 2014; Collias and Beisel, 2021), For each dCas9 variant, we observed similar growth profiles and similar effects on expression capacity (Figure 5.S4), suggesting that PAM flexibility does not produce additional growth burden effects.

To evaluate the effectiveness of each dCas9 variant for recognizing non-canonical PAM sites (non-NGG), we constructed libraries of reporters with varied PAM sequences. The first group of dCas9 variants (dxCas9(3.7), dCas9-NG, and dxCas9-NG) has a preference for NGN PAMs, so we screened three PAM libraries (NGH, NHG, and NHH, where H is not G). This approach follows the strategy previously used to characterize several Cas9 variants in mammalian systems (Legut et al., 2020). Taken together, the data indicate that dxCas9-NG provided the highest fold-activation across the largest number of alternative PAM reporters (Figure 5.2B). dxCas9-NG was effective at all NGH reporters (23 out of 23, or 100% had >10-fold activation) and most NHG reporters (21 out of 24, or 88% had >10-fold activation), but displayed substantially diminished effectiveness at NHH reporters (19 out of 71, or 27% had >10-fold activation). dCas9-NG performed similarly to dxCas9-NG at NGH and NHH PAMs, but notably weaker at NHG PAMs (NGH: 23 out of 23, or 100% had >10-fold activation; NHH: 22 out of 72, or 31% had >10-fold activation; NHG: 13 out of 24, or 54% had >10-fold activation). dCas9 and dxCas9(3.7) led to consistently reduced fold-activation and activated fewer reporters than dxCas9-NG across all three PAM libraries. The broad effectiveness of dxCas9-NG at NGH/NHG PAMs, along with its ability to activate some NHH PAMs, is consistent with prior reports from mammalian systems (Legut et al., 2020).

To evaluate the effectiveness of the dSpRY group of dCas9 variants (Figure 5.1A), we screened NRN and NYN PAM libraries (R = A or G; Y = C or T) following the previously-determined PAM preferences for dSpRY in mammalian cells (Walton et al., 2021, 2020). As expected, we observed the strongest performance across the broadest range of PAM reporters with dSpRY. This variant produced >10-fold CRISPRa at all tested reporters in the NRN library and 29 out of 64 (45%) of the reporters in the NYN library. In both libraries, dSpRY consistently outperformed both dSpG and dCas9 (Figure 5.2C).

We proceeded to directly compare the CRISPRa efficiency of dxCas9-NG and dSpRY at specific non-NGG PAMs. We tested reporters with NGH (GGA, GGT, GGC), NAN (CAA, CAT, CAC), and NYN (ATA) PAMs (Figure 5.2D). We found that dxCas9-NG and dSpRY outperformed both dCas9 and dxCas9(3.7) at all of the non-NGG PAMs. dxCas9-NG exhibited the highest activity at NGH (≥40-fold activation) and moderately weaker activity at NAN and NYN (between 10-fold to 20-fold activation). Compared to dxCas9-NG, dSpRY produced similar or stronger activation at NAN and NYN PAMs (>20-fold activation). These data indicate that different PAM-flexible variants have distinct patterns of optimal PAMs. Together with the PAM library screens, these results suggest a framework for choosing an effective dCas9 variant. For a given target gene of interest, PAMs at the appropriate target site positions should be identified (-70 and -80 for the template strand and -81 and -91 for non-template strand upstream of the TSS). dCas9 should be used for NGG, either dCas9-NG, dxCas9-NG or dSpG should be used for NGH,

dxCas9-NG should be used for NHG, and dSpRY should be used for NHH PAMs. When the four potential target sites include multiple alternative PAMs, they should be prioritized in the order NGG, NGH or NHG, and NHH. Following these guidelines, for subsequent experiments we prioritized dxCas9-NG for NGH/NHG PAMs and dSpRY for NHH PAMs.

CRISPRi efficiency is impaired with PAM-flexible dCas9 variants

CRISPRi acts by physically blocking transcription, and effective repression in bacteria can generally be obtained by targeting within the promoter or near the beginning of the gene on the non-template strand (Peters et al., 2016; Qi et al., 2013). CRISPRi is not subject to the same stringent target site requirements as CRISPRa, and there are often many canonical NGG PAMs available in the promoter or at the beginning of the gene. However, because we desire to express a single dCas9 protein to simultaneously target multiple genes for CRISPRa or CRISPRi, it is important to evaluate the performance of expanded PAM variants for CRISPRi. Previous work in bacterial and eukaryotic systems suggests that expanded PAM variants exhibit impaired CRISPRi function (Legut et al., 2020; Wang et al., 2021). We proceeded to evaluate the PAM-flexible variants dxCas9(3.7), dxCas9-NG, and dSpRY for CRISPRi gene repression. We tested multiple distinct sgRNA target sites, with one site in the promoter region and two sites within the ORF, all with NGG PAMs (Figure 5.3A). At each of these sites, dCas9 produces ~95-fold repression, while the PAM-flexible variants exhibit varying degrees of impaired repression (Figure 5.3B and Figure 5.S6). For target sites within the ORF, the PAM-flexible dCas9 variants repress gene expression by 5 to 30-fold; these effects are significant but substantially weaker than the ~95-fold repression obtained with dCas9 at these sites. For the target site at the promoter, the dxCas9(3.7) performs similarly to dCas9, while dxCas9-NG and dSpRY produce weaker repression effects (28-fold and 14-fold, respectively).

CRISPRi repression with dCas9 can be improved by targeting multiple sgRNAs to the same gene (Qi et al., 2013). We therefore tested whether pairs of sgRNAs could be used to improve CRISPRi repression with PAM-flexible dCas9 variants. In each case, we observed improved repression (Figure 5.3B and 5.3C). Most notably, dSpRY repression can be improved from 14-fold to 26-fold. The dSpRY variants has one of the broadest targeting ranges due to its wide tolerance of PAMs for CRISPRa (Figure 5.2D), and the ability to improve its CRISPRi function via gRNA multiplexing suggests that dSpRY can be used for multi-gene CRISPRa/i programs with a large dynamic range of activation and repression.

PAM-flexible dCas9 variants improve CRISPRa at endogenous promoters

We previously demonstrated that the expanded PAM variant dxCas9(3.7) enables activation of some endogenous genes that are inaccessible to dCas9 (Fontana et al., 2020a). To determine if dxCas9-NG and dSpRY further increase the pool of activatable endogenous genes, we examined four endogenous promoters previously tested with dxCas9(3.7): yajGp, uxuRp, araEp, and ppiDp2. For each endogenous promoter, we tested one NGG and one non-NGG PAM at appropriate positions upstream of the TSS at -70 or -80 (template strand) or at -81 or -91 (non-template strand), following the targeting rules defined previously (Fontana et al., 2020a). For two promoters previously activated by dxCas9(3.7), yajGp and uxuRp, dxCas9-NG and dSpRY produced comparable or improved activation compared to dxCas9(3.7) (Figure 5.4B). We also observed that two promoters that could not be activated by dxCas9(3.7) were slightly activated

with dxCas9-NG and dSpRY: araEp was activated by 1.3-fold by dxCas9-NG while ppiDp2 was activated by 3.1-fold by dSpRY (Figure 5.S7C).

We also tested nine new weakly-expressed endogenous promoters chosen from metabolic pathways related to aromatic amino acid biosynthesis (See Supplementary Methods) (Zaslaver et al., 2006). None of these promoters have NGG PAMs at the ideal distances relative to the TSS for CRISPR activation. For each promoter, we identified two candidate non-NGG targets that we expected to be compatible with dxCas9-NG or dSpRY (see Table 5.S3). These sites were selected with the target site position rules described above (Fontana et al., 2020a). Out of nine new promoters tested, five can be activated by more than 1.5-fold. Activation of aroKp1 by dxCas9-NG and activation of aroLp by dSpRY demonstrated the largest fold-activation values (16-fold and 8-fold, respectively) (Figure 5.4C). We observed modest, ~2-fold activation for pheLp, secBp, and aroFp, and no activation (<1.5-fold) for aroHp1, aroHp2, talAp, and serCp (Figure 5.S7D). For all promoters that were activated >1.5-fold, we observed distinct behaviors with different PAM-flexible variants. For aroKp1, dxCas9-NG significantly outperforms other variants at the GGT PAM. For aroLp, dSpRY outperforms other variants at the TAG PAM. These behaviors are consistent with our results from the mRFP reporter assay, which suggested the use of dxCas9-NG at NGN PAMs and dSpRY at NHH PAMs (Figure 5.2D). Taken together, these results suggest that the newer PAM-flexible variants dxCas9-NG and dSpRY can outperform dCas9 and dxCas9(3.7) for bacterial CRISPRa.

Out of the 13 total endogenous promoters tested, we found that five could not be activated above the 1.5-fold threshold (Figure 5.4B&5.4C, 5.S6C&5.S6D). One possible explanation for the failure of some target promoters to activate is that their basal expression levels are too high or too low. We previously observed that bacterial CRISPRa is sensitive to basal promoter strength, with no activation at the weakest promoters, effective activation at moderately weak promoters, and progressively smaller increases in gene expression as basal expression levels increased (Fontana et al., 2020a). Promoters with innately high expression levels may be inaccessible to high activation (>1.5-fold) due to the metabolic burden associated with increasing protein concentration.(Gyorgy et al., 2015; Wu et al., 2016) However, none of the five inaccessible promoters have unusually high basal expression levels (Figure 5.4D). Four of the promoters with <1.5 fold activation (aroHp1, aroHp2, talAp, and serCp) fall in a basal expression range that is higher than aroLp and lower than aroKp1, the two most highly activated promoters. Other factors beyond basal expression levels may be responsible for the failure of these four promoters to activate. The remaining inaccessible promoter, araEp, has the lowest basal expression level of all promoters tested and may be too weak to be activated with our current CRISPRa system.

5.3 Discussion

In this work, we have demonstrated that PAM-flexible dCas9 variants can improve bacterial CRISPRa in both synthetic and endogenous promoter contexts. Target distance from the TSS has been previously shown to be an important factor for effective bacterial CRISPRa (Fontana et al., 2020a; Ho et al., 2020; Liu et al., 2019; Villegas Kcam et al., 2021). PAM-flexible variants expand the scope of accessible PAM sites and therefore enable targeting at precise, optimal positions upstream of the TSS.

Although PAM-flexible dCas9 variants enable CRISPRa at a majority of previously

inaccessible gene targets, not all endogenous promoters with target sites at the appropriate position were able to be activated (Figure 5.4). Additional native regulatory machinery at endogenous gene targets may be responsible for preventing activation in these cases. We have previously shown that transcription factor binding can interfere with bacterial CRISPRa, presumably by physically obstructing binding of the CRISPRa complex or blocking the SoxS effector protein from engaging with RNA polymerase (Fontana et al., 2020a). We hypothesize that cryptic and unannotated transcription factor binding sites could be responsible for preventing activation at the endogenous genes targeted with PAM-flexible dCas9 variants in this work. Alternatively, we note that our guidelines for effective CRISPRa/i are based on experiments with fluorescent reporter genes, and while many endogenous targets appear to behave according to these rules, some gene sequences could have unexpected, context-specific effects on transcription or translation.

Improved bacterial CRISPRa with PAM-flexible variants comes with tradeoffs. We found that PAM-flexible dCas9 variants exhibited weaker CRISPRi-based gene repression compared to dCas9 (Figure 5.3B). One possible explanation for this behavior follows from the observation that increasing PAM promiscuity reduces affinity towards NGG PAMs (Corsi et al., 2022; Hu et al., 2018; Legut et al., 2020; Nishimasu et al., 2018; Walton et al., 2020). This reduced PAM-binding affinity could allow RNA polymerase to more readily displace the CRISPRi complex. Alternatively, a near-PAMless dCas9 variant would be expected to interrogate almost every DNA sequence in the bacterial genome (Anders et al., 2014; Collias and Beisel, 2021) and increase the time needed to find the correct target site. During every cell division, the CRISPRi complex is displaced from the genome (Qi et al., 2013), and increased time will be needed to bind the target site (Jones et al., 2017), which could allow for increased leaky expression and consequently weaker CRISPRi compared to the parent dCas9.

To implement complex, programmable genetic regulatory networks, both upregulation and downregulation controls with broad dynamic ranges are desirable (Alon, 2007; Brophy and Voigt, 2014; Tickman et al., 2021). Thus, identifying systems that improve CRISPRa while maintaining effective CRISPRi is crucial. Multiplexing the CRISPRi targets with additional sgRNAs led to significantly higher fold-repression, although still weaker than CRISPRi with the parent dCas9 (FIgure 5.3C-5.3D). It remains to be seen whether these improvements are sufficient for genetic circuit applications. If stronger repression is needed, an alternative approach could be to distribute each engineering task to a different subpool of Cas9 proteins (Collias and Beisel, 2021). In this application, a PAM-flexible variant could be used for CRISPRa and an orthogonal Cas protein could be used for CRISPRi.

An additional challenge for multi-gene CRISPRa/i regulation is the possibility that a CRISPRa target site ~80-90 bases upstream of a target gene could have undesired CRISPRi effects on the gene immediately upstream. The average intergenic region upstream of an *E. coli* gene is ~140 bases (Fontana et al., 2018b), so CRISPRa target sites could overlap with promoters or the 3' end of an ORF, depending on the orientation of the preceding gene, and either situation could produce CRISPRi repression (Qi et al., 2013). The use of the PAM-flexible dCas9 variants described here has no impact on this challenge, and any DNA-targeting transcriptional activation system in bacteria is likely to face the same issue. Alternative gene regulation methods, including

mRNA-targeting CRISPR translational activation systems (Otoupal et al., 2022), could be useful in this scenario.

While this manuscript was under revision, another manuscript was published that reported broadly consistent findings with the dSpRY variant (Klanschnig et al., 2022). Our work includes additional comparisons with multiple PAM-flexible dCas9 variants and identifies situations where some of these variants (notably dxCas9-NG) are more effective than dSpRY. We also observed tradeoffs with CRISPRi efficiency that have not been previously described. Overall, the independently-obtained results support the broader idea that PAM-flexible Cas9 variants provide an effective means to overcome challenges associated with bacterial CRISPRa/i gene regulation.

These improvements to bacterial CRISPRa bring us closer to the long-standing goal of performing CRISPRa gain-of-function screens for basic discovery and engineering applications. Previously, some successes have been reported for activating natural product biosynthesis pathways (Ameruoso et al., 2022; Ke et al., 2022) but the ability to perform genome-wide CRISPRa screens in bacteria lags far behind eukaryotic systems (Joung et al., 2017; Kampmann, 2018; Sanson et al., 2018; Schmidt et al., 2022). By unlocking these capabilities in bacteria, CRISPRa screens could enable rapid functional annotation of uncharacterized genes. For bioindustrial applications, we envision identifying genes that can overcome bottlenecks in routing metabolic flux or that confer robustness in harsh, non-native growth conditions. In the long term, combined CRISPRi and CRISPRa screens could provide even more information to map biological functions and deconvolute regulatory networks. In mammalian cells, combining information from single gene CRISPRa/i screens enabled improved chemical genetic profiling to identify drug targets (Jost et al., 2017), and dual-gene activation/repression screens have identified genetic interactions and functional relationships between genes (Boettcher et al., 2018; Najm et al., 2018). Combined CRISPRa/i screens should be possible in bacteria, as CRISPRi loss-of-function screens have been broadly applied (Todor et al., 2021), and it is straightforward to target multiple genes with multiple gRNAs for activation and repression (Dong et al., 2018; Kiattisewee et al., 2021; Wu et al., 2020). Some of these goals are plausibly within reach in bacteria using current CRISPRa tools, and further improvements in CRISPRa systems at endogenous gene targets will enable rapid progress in basic research and bioindustrial applications.

5.4 Materials and methods

Bacterial strains and plasmid constructs

E. coli K-12 substrain MG1655 was used for all CRISPRa tests unless specified. CD38 with highly expressed mRFP was used for CRISPRi experiments (Table S1). Plasmid constructs were cloned using standard molecular biology methods (Fontana et al., 2020a; Kiattisewee et al., 2021). All PCR fragments were amplified with Phusion DNA Polymerase (Thermo-Fisher Scientific) for Infusion Cloning (Takara Bio). Plasmids were transformed into chemically competent NEB Turbo *E. coli* (New England Biolabs) cells, plated on LB-agar, and cultured in LB media supplied with the appropriate antibiotics used in the following concentrations: 100 µg/mL Carbenicillin, 25 µg/mL Chloramphenicol, 30 µg/mL Kanamycin. All plasmid constructs were confirmed by Sanger sequencing (GENEWIZ). Selected plasmids will be available upon request on Addgene (https://www.addgene.org/Jesse_Zalatan/).

PAM-flexible dCas9 variants were cloned from existing dCas9 and dxCas9(3.7) plasmids (pCD442 and pCD564) (Fontana et al., 2020a). dxCas9-NG was cloned by replacing the C-terminus of dxCas9(3.7) with dCas9-NG mutations ordered as a gBlock (IDT). dCas9-NG was cloned by fusing the N-terminus of dCas9 and C-terminus of dxCas9-NG together. dSpG and dSpRY sequences were cloned into a bacterial codon optimized vector (Addgene #101199) and then subcloned into the pCD442 vector. Complete sequences are provided in the supplementary information.

Each dCas9 variant (dCas9, dxCas9, dCas9-NG, dxCas9-NG, dSpG, and dSpRY) was expressed from the endogenous Sp.pCas9 promoter in a p15A vector (Table 5.S2). MCP-SoxS (R93A, S101A) (abbreviated MCP-SoxS) was expressed from the BBa_J23107 promoter (http://parts.igem.org) in the same plasmid with dCas9. The single guide RNAs (sgRNA) or modified scaffold RNAs b2.1xMS2 (scRNAs) were expressed from the strong BBa_J23119 promoter, either in the same plasmid with the dCas9-carrying plasmid or in a separate ColE1 plasmid (Dong et al., 2018; Fontana et al., 2020a). All 20 bp scRNA/sgRNA target sequences are provided in Table 5.S3. The mRFP reporter was expressed from the weak J3-BBa_J23117 promoter on a pSC101** plasmid (Table 5.S2). For CRISPRi experiments, a construct expressing mRFP from the strong BBa_J23119 promoter (strain CD38, Table 5.S1) was integrated into the *E. coli* genome using a previously-described lambda red system (Dong et al., 2018; Fontana et al., 2018a). For endogenous promoter CRISPRa experiments, we used GFPmut2 reporters on pSC101** vectors as described previously (Zaslaver et al., 2006). Reporters were purchased from Horizon Discovery or constructed with the same methodology (Zaslaver et al., 2006).

Plate reader experiments

Single colonies from LB plates were inoculated in 400 µL of EZ-RDM (Teknova) supplemented with the appropriate antibiotics and grown in 96-deep-well plates at 37 °C with shaking overnight 900 RPM on a Heidolph titramax 1000. 150 µL of the overnight culture were transferred into flat, clear-bottomed black 96-well plates (Corning) and the OD_{600} and fluorescence were measured in a Biotek Synergy HTX plate reader. Data were analyzed using the BioTek Gen5 2.07.17 software. For mRFP1 detection, the excitation wavelength was 540 nm and emission wavelength was 600 nm. For GFPmut2 or sfGFP detections, the excitation wavelength was 485 nm and emission wavelength was 528 nm.

Flow cytometry

Single colonies from LB plates were inoculated in 400 µL EZ-RDM (Teknova) supplemented with appropriate antibiotics and grown in 96-deep-well plates at 37 °C, 900 RPM on a Heidolph titramax 1000. Cultures were grown overnight and then diluted in 1:100 in Dulbecco's phosphate-buffered saline (PBS) and analyzed on a MACSQuant VYB flow cytometer with the MACSQuantify 2.8 software (Miltenyi Biotec). To select single cells, we used a previously-described gating procedure (Dong et al., 2018). A side scatter threshold trigger (SSC-H) was applied to select for single cells until 10000 events were collected. FlowJo 10.0.7 software was used to apply a narrow gate along the diagonal line on the SSC-H vs SSC-A plot to exclude the events where multiple cells were grouped together. Within the selected population, events that appeared on the edges of the FSC-A vs. SSC-A plot and the fluorescence histogram were excluded.

E. coli growth profiles and expression capacity

To evaluate *E. coli* expression capacity, we incorporated an sfGFP capacity monitor expressed from a constitutive, medium-strength promoter (BBa_J23110) into the mRFP CRISPRa reporter plasmid. This method follows previously-described approaches to evaluate expression capacity (Ceroni et al., 2015). For measurement of expression burden, CRISPRa plasmids with different dCas9 variants and an scRNA (J306 for CRISPRa and hAAVS1 for an off-target control) were co-transformed with the reporter plasmids (pJF143.J3 or pCK760). Single colonies were inoculated in 400 µL EZ-RDM with appropriate antibiotics and endpoint fluorescent protein levels were measured after 18 hours with a plate reader as described above.

Growth profile time courses (OD_{600} vs. time) were obtained with strains inoculated from overnight stationary cultures with a 1:100 dilution. Growth profiles were also initiated from exponentially-growing cultures by first subculturing overnight cultures 1:100 into 2 mL EZ-RDM in 14 mL culture tubes for 2 hours, then diluting all cultures back to OD_{600} = 0.1. 200 µL of subcultures were transferred into flat, clear-bottomed black 96-well plates (Corning). OD_{600} values were measured in a plate reader for 16 hours at 37 °C.

Pooled PAM library construction and screening

To generate the pooled reporter library with different PAMs at the target site -81 bp from the TSS, we used pJF143.J3 (Supplementary Table 5.S2) as a PCR template with oCK679_NNN (5'-CTCGTCTCCTCACTTTNNNACGGAGCGTTCTGGACACAACG-3') as a forward primer and oCK680 (5'-AAGTGAGGAGACGAGCGAACGC-3') as a reverse primer. Amplified linear fragments were treated with DpnI to remove the parental vector and circularized with Infusion. oCK679_NNN oligos with NGH, NHG, NHH, NRN, and NYN were used to construct each corresponding PAM library. For each screened variant, the number of colonies picked was 2X the number of sequence variants. For example, we picked 24 colonies for NGH (12 possible sequences) and 64 colonies for NRN (32 possible sequences). Fold activation was calculated relative to a strain with pJF143.J3 and an off-target scRNA.

CRISPRa at endogenous promoters

CRISPRa at endogenous promoters was performed with a three plasmid system — dCas9 plasmid, scRNA plasmid, and reporter plasmid. The reporter plasmids were adapted from the *E. coli* promoter collection (Dharmacon), a commercially available library of promoter-GFPmut2 fusions (Zaslaver et al., 2006). We have previously confirmed that CRISPRa effects on these fluorescent protein reporters are also observed by RT-qPCR of the endogenous genomic transcript (Fontana et al., 2020a). Four promoters tested here (yajGp, uxuRp, araEp, and ppiDp2) were evaluated previously with dxCas9(3.7) (Fontana et al., 2020a). These promoters were chosen based on the criteria: 1) genes should not be highly expressed, and 2) genes should be regulated by the sigma70 family (Fontana et al., 2020a). A second set of endogenous promoters (aroKp1, aroLp, pheLp, secBp, aroFp, aroHp1, aroHp2, talAp, and serCp) were selected from genes involved in aromatic amino acid biosynthesis pathways that meet the same criteria described above. scRNAs for each promoter were designed to target the optimal positions (-70 and -80 for the template strand and -81 and -91 for non-template strand relative to the TSS). One scRNA from each strand was chosen, based on which PAM was predicted to be accessible to the

highest number of dCas9 variants. Accessibility was assessed based on the moderate performance threshold cutoff (see Table 5.S4 and Supplementary Methods).

Bioinformatic analysis of targetable genes

Previously reported data for the activity of PAM-flexible dCas9 variants were used to predict the targetable genes for each dCas9 variant. To identify the compatible PAMs for each engineered dCas9 variant, we used data from Cas9 nuclease assays for each variant (Kim et al., 2020; Legut et al., 2020; Walton et al., 2020). Predicted compatible PAMs for each variant are provided in Table 5.S4 (see Supplementary Methods for further details). We then examined *E. coli* and *P. putida* genome sequences for PAM availability. For *E. coli*, 1265 promoters with strong confidence in TSS were retrieved from RegulonDB (Santos-Zavaleta et al., 2019). For *P. putida*, 1104 experimentally-confirmed primary transcriptional units were used for the analysis (D'Arrigo et al., 2016). The PAMs at optimal target site positions were retrieved: -70 and -80 for the template strand and -81 and -91 for non-template strands relative to the TSS. The promoters with at least one compatible PAM out of four target sites were considered targetable. Further information can be found in Figure 5.S3.

Data analysis

Flow cytometry analysis was conducted on FlowJo 10.0.7 software or Python FlowCytometryTools on Jupyter Notebooks and then further processed with Microsoft Excel and Graphpad Prism. Data represents the average and standard deviation of at least three biologically-independent replicates, unless specified. Fold activation and fold repression were calculated by comparing sample fluorescence with a strain expressing off-target scRNA/sgRNA.

Disclosure and competing interests statement

B.P.K. is an inventor on patents and/or patent applications filed by Mass General Brigham that describe genome engineering technologies, including for the development of SpRY. B.P.K. is a consultant for EcoR1 capital and ElevateBio, and is an advisor to Acrigen Biosciences, Life Edit Therapeutics, and Prime Medicine. N.E.S. is an advisor to Vertex and Qiagen. J.M.C. and J.G.Z. are advisors to Wayfinder Biosciences.

Author contributions

C.K., A.V.K., J.M.C., and J.G.Z. developed the initial concept; B.P.K. and N.E.S. provided additional insight to the project; C.K. and J.N. performed bioinformatic analysis; C.K., A.V.K., and S.K. performed experiments in bacteria and analyzed the results; M.L. and Z.D. performed experiments in mammalian cells; C.K. and M.L. analyzed the results from mammalian cells; C.K., A.V.K., J.M.C., and J.G.Z. wrote the manuscript with input from all authors.

Data availability

Screen data generated during this study have been deposited to GEO with an accession number GSE217559.Screen data generated during this study will be deposited to GEO.

Supporting Information

The Supporting Information is available at the end of this chapter. Schematic of mutations in PAM-flexible dCas9 variants; CRISPRko and CRISPRa screens in mammalian systems; PAM

availability and distribution in *E. coli* and *P. putida* promoters; growth burden comparisons; effect of dCas9 variants on basal reporter gene expression; CRISPRi with PAM-flexible dCas9 variants; CRISPRa screening at endogenous promoters; tables of bacterial strains, plasmids, and gRNA target sites; DNA sequences of reporter genes, proteins, and bacterial promoters, and supplementary methods.

Acknowledgements

We thank Benjamin I. Tickman, Ian D. Faulkner, Diego Alba Burbano, Ryan Cardiff, and members of Carothers and Zalatan laboratories for thorough discussion during the experimental planning and data collection/analysis strategy. We also thank Madelynn N. Whittaker for cloning the original bacterial codon optimized dSpG and dSpRY constructs. This work was supported by NSF Award 1817623 (J.M.C, J.G.Z.), DOE Award DE-EE0008927 (J.M.C, J.G.Z.), and NIH/NHGRI DP2HG010099 (N.E.S.). B.P.K. acknowledges funding from an MGH ECOR Howard M. Goodman Award. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-2140004 (A.V.K.). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

5.5 Tables and Figures

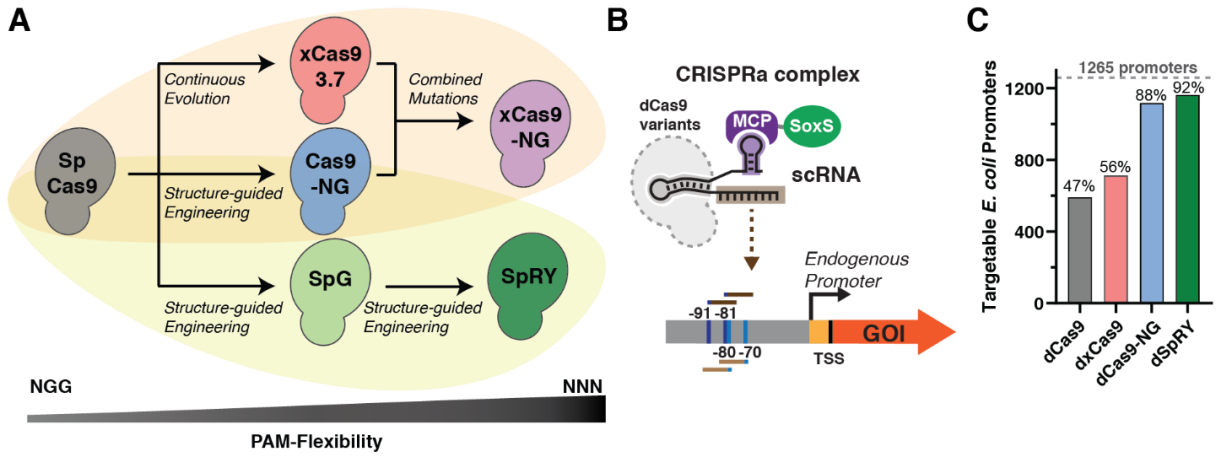


Figure 5.1: Engineered PAM-flexible Cas9 variants and PAM availability analysis

(A) PAM-flexible S. pyogenes Cas9 (Sp-Cas9) have been engineered using various rational design, screening, and directed evolution methodologies (e.g. phage-assisted continuous evolution (PACE), structure-guided engineering monitored via HT-PAMDA) (Hu et al., 2018; Kleinstiver et al., 2016, 2015; Legut et al., 2020; Nishimasu et al., 2018; Walton et al., 2021, 2020). (B) The CRISPRa complex consists of dCas9, an scRNA, and the MCP-SoxS activator (Dong et al., 2018). Previous work suggests that four precisely-positioned sites upstream of the TSS are most effective for CRISPRa (-70 and -80 for the template strand and -81 and -91 for non-template strand relative to the TSS) (Fontana et al., 2020a). (C) 1265 *E. coli* endogenous promoters were analyzed for target site availability with different PAM-flexible dCas9 variants. Promoters were identified as targetable if they have at least one effective target site with a compatible PAM. For each variant, PAM preferences were obtained from previously-reported nuclease screening experiments (see Methods; PAMs with at least 20% of Sp-Cas9 activity at NGG PAMs were considered compatible).

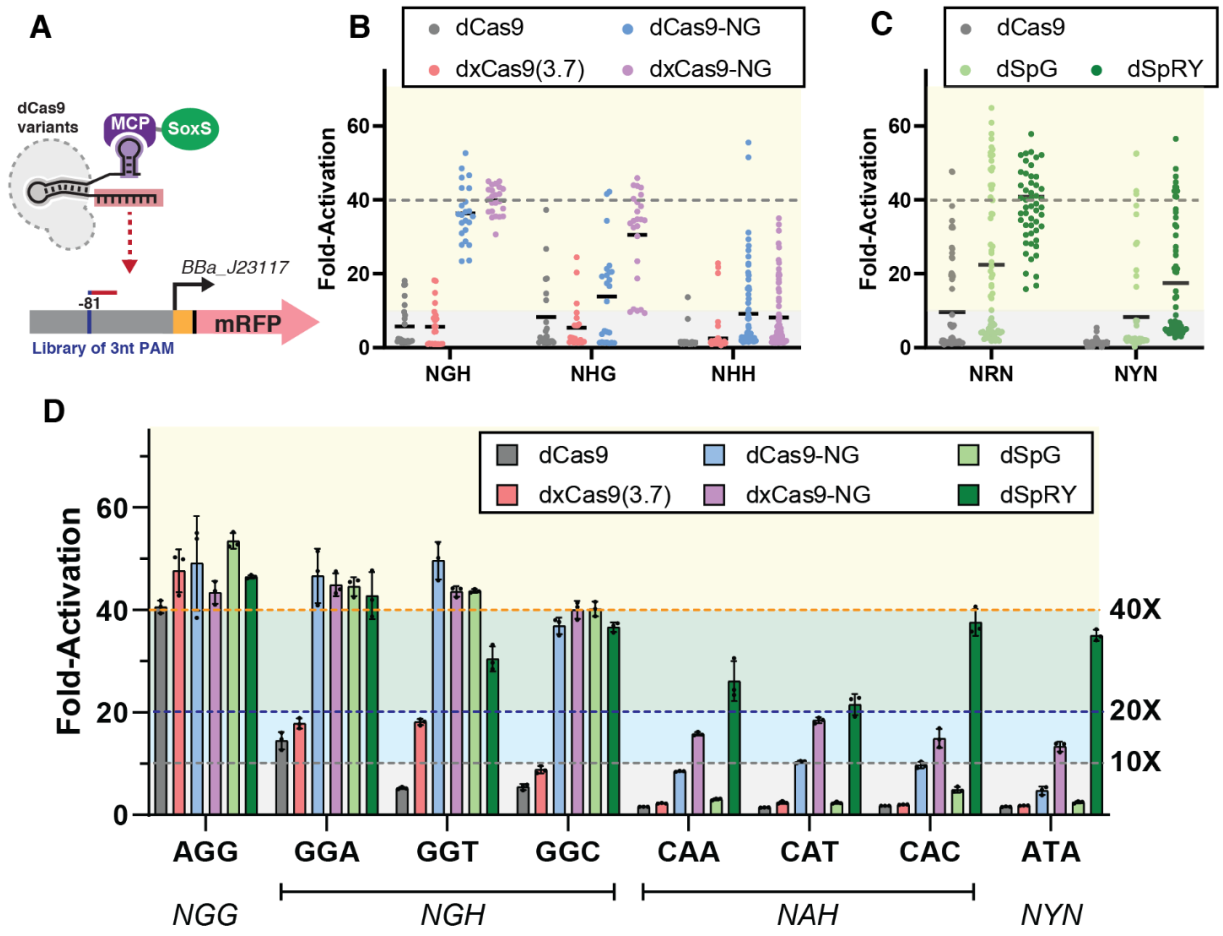


Figure 5.2: CRISPRa with PAM-flexible dCas9 variants

(A) CRISPRa for PAM-flexible dCas9 variants was tested on an mRFP reporter gene with libraries of alternative 3 nucleotide PAMs at the -81 target site. (B) Reporter gene expression with the dxCas9-NG group (Figure 5.1A) at NGH, NHG, and NHH PAM libraries. dxCas9-NG outperforms other variants in every PAM library. (C) Reporter gene expression with the dSpRY group (Figure 5.1A) at NRN and NYN PAM libraries. dSpRY exhibited the highest CRISPRa efficiency. (D) Direct comparisons of reporter gene expression with dCas9, dxCas9(3.7), dCas9-NG, dxCas9-NG, dSpG, and dSpRY at a representative set of PAMs. dCas9-NG, dxCas9-NG, and dSpRY performed best at NGN PAMs while dSpRY outperformed other variants at NAN and NYN PAMs. An scRNA targeting the J306 sequence was used for all library screens and individual PAM assays. To calculate fold-activation, we used an off-target scRNA (hAAVS1) with an AGG PAM reporter to define the basal expression level. Basal reporter expression levels vary <1.5-fold with different dCas9 variants or PAMs (Figure 5.S5). Values in panel D represent the mean ± standard deviation calculated from n = 3.

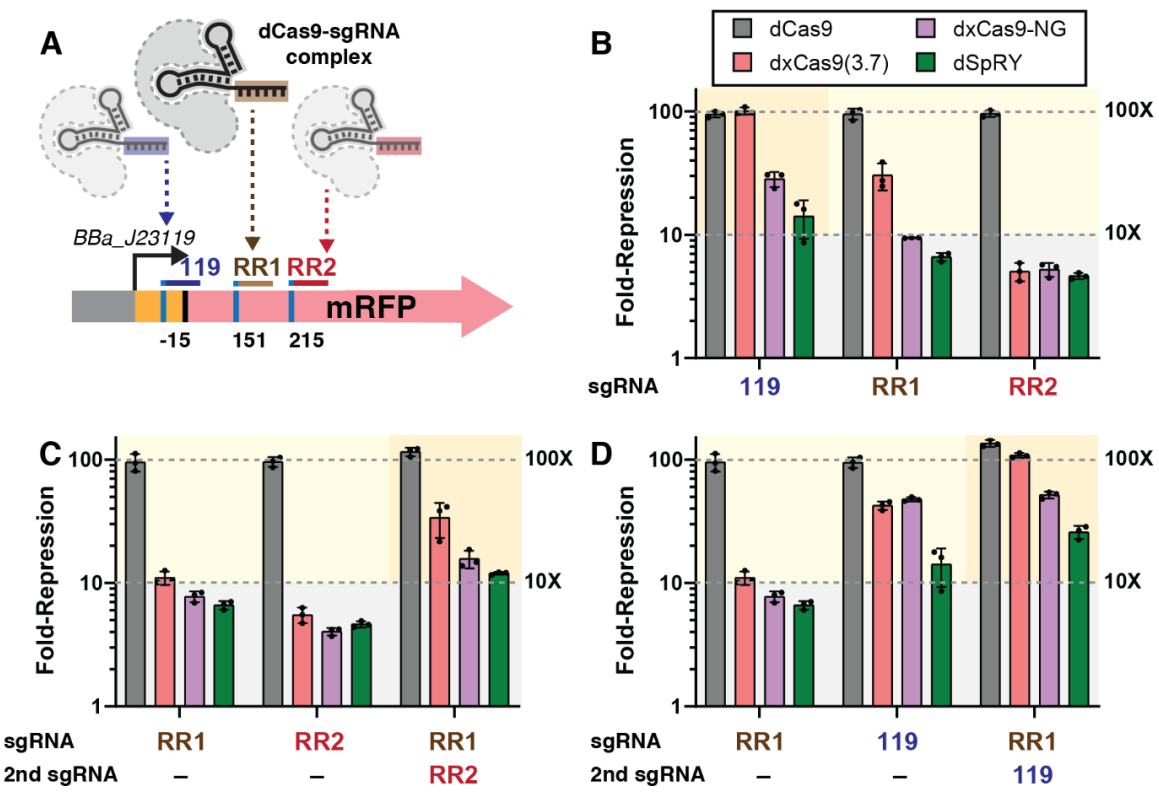


Figure 5.3: CRISPRi with PAM-flexible dCas9 variants

(A) CRISPRi for PAM-flexible dCas9 variants was tested on an mRFP reporter gene with sgRNAs targeting the promoter (119) or coding sequence (RR1 and RR2), where each target site encodes an NGG PAM. (B) Fold-repression of the mRFP reporter gene with a single expressed sgRNA. (C & D) Comparison of fold-repression with one or two sgRNAs expressed. Fold-repression consistently increases when two sgRNAs are expressed. See Figure 5.S6 for a comparison of all dCas9 variants, including dCas9-NG and dSpG, with single and multiple sgRNAs. Values in panel B, C, and D represent the mean ± standard deviation calculated from n = 3.

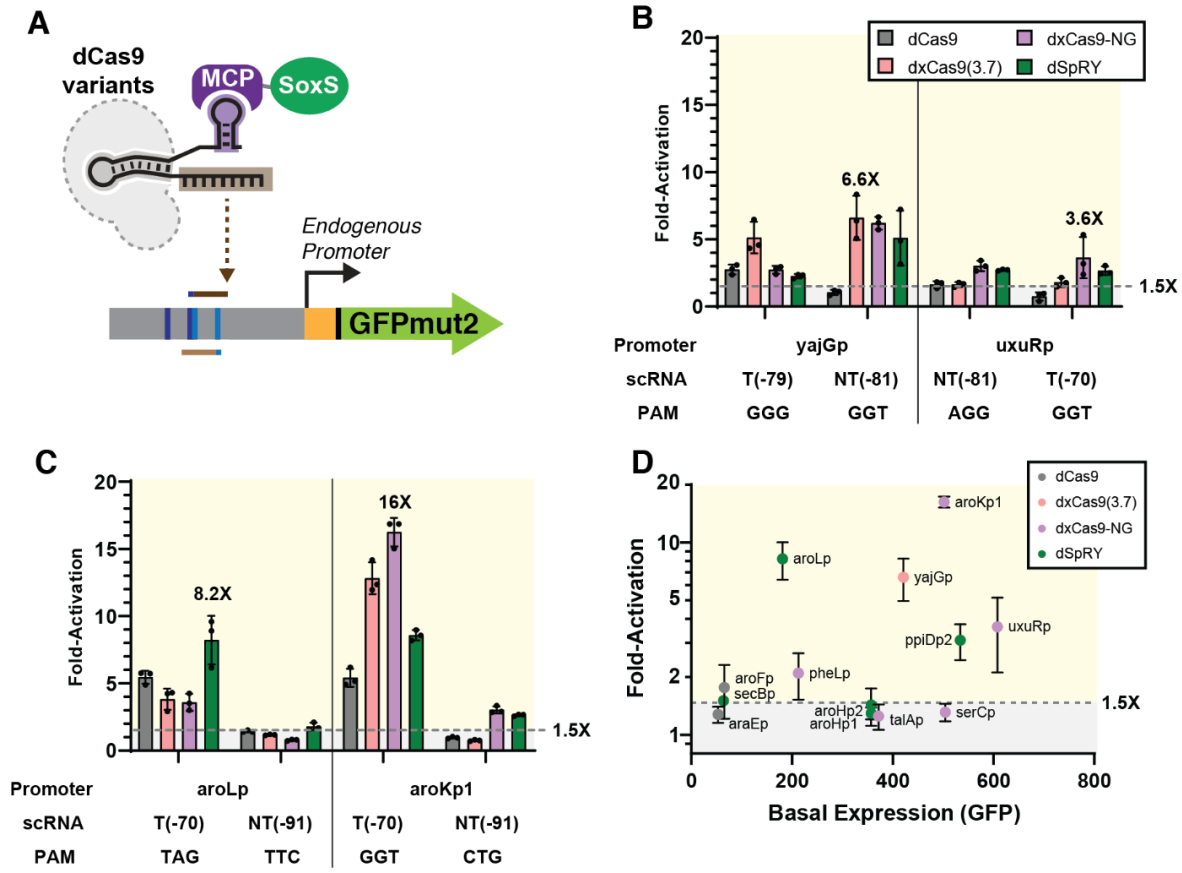


Figure 5.4: CRISPRa at endogenous promoters is enhanced with PAM-flexible dCas9 variants

(A) CRISPRa at endogenous promoters of *E. coli* was tested using two of the four optimal scRNA positions for each promoter (see Supplemental Methods). (B) CRISPRa at endogenous promoters previously tested with dxCas9(3.7) (yajGp and uxuRp) using NGG or non-NGG PAMs. See Figure 5.S7C for araEp and ppiDp2 promoters. (C) CRISPRa at endogenous promoters involved in aromatic amino acid biosynthesis using non-NGG PAMs. See Figure 5.S7D for additional promoters (D) Plot of fold-activation versus basal expression level for all endogenous promoters tested in (B) & (C). See Table 5.S5 for additional details. Data were collected by flow cytometry and fold-activation was calculated relative to a strain expressing the corresponding dCas9 variant and an off-target hAAVS1 scRNA. Values in panel B, C, and D represent the mean ± standard deviation calculated from n = 3.

5.6 References

Alon, U., **2007**. Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* **8**, 450–461. https://doi.org/10.1038/nrg2102

Ameruoso, A., Villegas Kcam, M.C., Cohen, K.P., Chappell, J., **2022**. Activating natural product synthesis using CRISPR interference and activation systems in *Streptomyces*. *Nucleic Acids Res.*

Anders, C., Niewoehner, O., Duerst, A., Jinek, M., **2014**. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513**, 569–573. https://doi.org/10.1038/nature13579

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 Res* **41**, 7429–37. https://doi.org/10.1093/nar/gkt520

Boettcher, M., Tian, R., Blau, J.A., Markegard, E., Wagner, R.T., Wu, D., Mo, X., Biton, A., Zaitlen, N., Fu, H., McCormick, F., Kampmann, M., McManus, M.T., **2018**. Dual gene activation and knockout screen reveals directional dependencies in genetic networks. *Nat. Biotechnol.* **36**, 170–178. https://doi.org/10.1038/nbt.4062

Brophy, J.A.N., Voigt, C.A., **2014**. Principles of genetic circuit design. *Nat. Methods* **11**, 508–520. https://doi.org/10.1038/nmeth.2926

Ceroni, F., Algar, R., Stan, G.-B., Ellis, T., **2015**. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods* **12**, 415–418. https://doi.org/10.1038/nmeth.3339

Cho, S., Choe, D., Lee, E., Kim, S.C., Palsson, B., Cho, B.-K., **2018**. High-Level dCas9 Expression Induces Abnormal Cell Morphology in *Escherichia coli*. *ACS Synth. Biol.* **7**, 1085–1094. https://doi.org/10.1021/acssynbio.7b00462

Collias, D., Beisel, C.L., **2021**. CRISPR technologies and the search for the PAM-free nuclease. *Nat. Commun.* **12**, 555. https://doi.org/10.1038/s41467-020-20633-y

Corsi, G.I., Qu, K., Alkan, F., Pan, X., Luo, Y., Gorodkin, J., **2022**. CRISPR/Cas9 gRNA activity depends on free energy changes and on the target PAM context. *Nat. Commun.* **13**, 3006. https://doi.org/10.1038/s41467-022-30515-0

D'Arrigo, I., Bojanovič, K., Yang, X., Holm Rau, M., Long, K.S., **2016**. Genome-wide mapping of transcription start sites yields novel insights into the primary transcriptome of *Pseudomonas putida*. *Environ. Microbiol.* **18**, 3466–3481. https://doi.org/10.1111/1462-2920.13326

Dong, C., Fontana, J., Patel, A., Carothers, J.M., Zalatan, J.G., **2018**. Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nat Commun* **9**, 2489. https://doi.org/10.1038/s41467-018-04901-6

Fontana, J., Dong, C., Ham, J.Y., Zalatan, J.G., Carothers, J.M., **2018a**. Regulated Expression of sgRNAs Tunes CRISPRi in *E. coli*. *Biotechnol J* **13**, e1800069. https://doi.org/10.1002/biot.201800069

Fontana, J., Dong, C., Kiattisewee, C., Chavali, V.P., Tickman, B.I., Carothers, J.M., Zalatan, J.G., **2020a**. Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. *Nat. Commun.* **11**, 1618. https://doi.org/10.1038/s41467-020-15454-y

Fontana, J., Sparkman-Yager, D., Zalatan, J.G., Carothers, J.M., **2020b**. Challenges and opportunities with CRISPR activation in bacteria for data-driven metabolic engineering. *Curr. Opin. Biotechnol.* **64**, 190–198. https://doi.org/10.1016/j.copbio.2020.04.005

Fontana, J., Voje, W.E., Zalatan, J.G., Carothers, J.M., **2018b**. Prospects for engineering dynamic CRISPR-Cas transcriptional circuits to improve bioproduction. *J Ind Microbiol Biotechnol*. https://doi.org/10.1007/s10295-018-2039-z

Gyorgy, A., Jiménez, J.I., Yazbek, J., Huang, H.-H., Chung, H., Weiss, R., Del Vecchio, D., **2015**. Isocost Lines Describe the Cellular Economy of Genetic Circuits. *Biophys. J.* **109**, 639–646. https://doi.org/10.1016/j.bpj.2015.06.034

Ho, H.-I., Fang, J.R., Cheung, J., Wang, H.H., **2020**. Programmable CRISPR-Cas transcriptional activation in bacteria. *Mol. Syst. Biol.* **16**, e9427. https://doi.org/10.15252/msb.20199427

Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M., Gao, X., Rees, H.A., Lin, Z., Liu, D.R., **2018**. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57–63. https://doi.org/10.1038/nature26155

Jones, D.L., Leroy, P., Unoson, C., Fange, D., Ćurić, V., Lawson, M.J., Elf, J., **2017**. Kinetics of dCas9 target search in *Escherichia coli*. *Science*.

Jost, M., Chen, Y., Gilbert, L.A., Horlbeck, M.A., Krenning, L., Menchon, G., Rai, A., Cho, M.Y., Stern, J.J., Prota, A.E., Kampmann, M., Akhmanova, A., Steinmetz, M.O., Tanenbaum, M.E., Weissman, J.S., **2017**. Combined CRISPRi/a-Based Chemical Genetic Screens Reveal that Rigosertib Is a Microtubule-Destabilizing Agent. *Mol. Cell* **68**, 210-223.e6. https://doi.org/10.1016/j.molcel.2017.09.012

Joung, J., Engreitz, J.M., Konermann, S., Abudayyeh, O.O., Verdine, V.K., Aguet, F., Gootenberg, J.S., Sanjana, N.E., Wright, J.B., Fulco, C.P., Tseng, Y.-Y., Yoon, C.H., Boehm, J.S., Lander, E.S., Zhang, F., **2017**. Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. *Nature* **548**, 343–346. https://doi.org/10.1038/nature23451

Kampmann, M., **2018**. CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine. *ACS Chem. Biol.* **13**, 406–416. https://doi.org/10.1021/acschembio.7b00657

Ke, J., Robinson, D., Wu, Z.-Y., Kuftin, A., Louie, K., Kosina, S., Northen, T., Cheng, J.-F., Yoshikuni, Y., **2022**. CRAGE-CRISPR facilitates rapid activation of secondary metabolite biosynthetic gene clusters in bacteria. *Cell Chem. Biol.* **29**, 696-710.e4. https://doi.org/10.1016/j.chembiol.2021.08.009

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*. *Metab. Eng.* **66**, 283–295. https://doi.org/10.1016/j.ymben.2021.04.002

Kim, N., Kim, H.K., Lee, S., Seo, J.H., Choi, J.W., Park, J., Min, S., Yoon, S., Cho, S.-R., Kim, H.H., **2020**. Prediction of the sequence-specific cleavage activity of Cas9 variants. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-020-0537-9

Klanschnig, M., Cserjan-Puschmann, M., Striedner, G., Grabherr, R., **2022**. CRISPRactivation-SMS, a message for PAM sequence independent gene up-regulation in *Escherichia coli*. *Nucleic Acids Res.* gkac804. https://doi.org/10.1093/nar/gkac804

Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., Joung, J.K., **2016**. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495. https://doi.org/10.1038/nature16526

Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Topkar, V.V., Nguyen, N.T., Zheng, Z., Gonzales, A.P.W., Li, Z., Peterson, R.T., Yeh, J.-R.J., Aryee, M.J., Joung, J.K., **2015**. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485. https://doi.org/10.1038/nature14592

Legut, M., Daniloski, Z., Xue, X., McKenzie, D., Guo, X., Wessels, H.-H., Sanjana, N.E., **2020**. High-Throughput Screens of PAM-Flexible Cas9 Variants for Gene Knockout and Transcriptional Modulation. *Cell Rep.* **30**, 2859-2868.e5. https://doi.org/10.1016/j.celrep.2020.02.010

Liu, Y., Wan, X., Wang, B., **2019**. Engineered CRISPRa enables programmable eukaryote-like gene activation in bacteria. *Nat. Commun.* **10**, 3693. https://doi.org/10.1038/s41467-019-

11479-0
Najm, F.J., Strand, C., Donovan, K.F., Hegde, M., Sanson, K.R., Vaimberg, E.W., Sullender, M.E., Hartenian, E., Kalani, Z., Fusi, N., Listgarten, J., Younger, S.T., Bernstein, B.E., Root, D.E., Doench, J.G., **2018**. Orthologous CRISPR–Cas9 enzymes for combinatorial genetic screens. *Nat. Biotechnol.* **36**, 179–189. https://doi.org/10.1038/nbt.4048
Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T., Abudayyeh, O.O., Gootenberg, J.S., Mori, H., Oura, S., Holmes, B., Tanaka, M., Seki, M., Hirano, H., Aburatani, H., Ishitani, R., Ikawa, M., Yachie, N., Zhang, F., Nureki, O., **2018**. Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* **361**, 1259–1262. https://doi.org/10.1126/science.aas9129
Otoupal, P.B., Cress, B.F., Doudna, J.A., Schoeniger, J.S., **2022**. CRISPR-RNAa: targeted activation of translation using dCas13 fusions to translation initiation factors. *Nucleic Acids Res.* **50**, 8986–8998. https://doi.org/10.1093/nar/gkac680
Peters, J.M., Colavin, A., Shi, H., Czarny, T.L., Larson, M.H., Wong, S., Hawkins, J.S., Lu, C.H.S., Koo, B.-M., Marta, E., Shiver, A.L., Whitehead, E.H., Weissman, J.S., Brown, E.D., Qi, L.S., Huang, K.C., Gross, C.A., **2016**. A Comprehensive, CRISPR-based Functional Analysis of Essential Genes in Bacteria. *Cell* **165**, 1493–1506. https://doi.org/10.1016/j.cell.2016.05.003
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**, 1173–83. https://doi.org/10.1016/j.cell.2013.02.022
Sanson, K.R., Hanna, R.E., Hegde, M., Donovan, K.F., Strand, C., Sullender, M.E., Vaimberg, E.W., Goodale, A., Root, D.E., Piccioni, F., Doench, J.G., **2018**. Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nat. Commun.* **9**, 5416. https://doi.org/10.1038/s41467-018-07901-8
Santos-Zavaleta, A., Salgado, H., Gama-Castro, S., Sánchez-Pérez, M., Gómez-Romero, L., Ledezma-Tejeida, D., García-Sotelo, J.S., Alquicira-Hernández, K., Muñiz-Rascado, L.J., Peña-Loredo, P., Ishida-Gutiérrez, C., Velázquez-Ramírez, D.A., Del Moral-Chávez, V., Bonavides-Martínez, C., Méndez-Cruz, C.-F., Galagan, J., Collado-Vides, J., **2019**. RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in *E. coli* K-12. *Nucleic Acids Res.* **47**, D212–D220. https://doi.org/10.1093/nar/gky1077
Schmidt, R., Steinhart, Z., Layeghi, M., Freimer, J.W., Bueno, R., Nguyen, V.Q., Blaeschke, F., Ye, C.J., Marson, A., **2022**. CRISPR activation and interference screens decode stimulation responses in primary human T cells. *Science*.
Tickman, B.I., Burbano, D.A., Chavali, V.P., Kiattisewee, C., Fontana, J., Khakimzhan, A., Noireaux, V., Zalatan, J.G., Carothers, J.M., **2021**. Multi-layer CRISPRa/i circuits for dynamic genetic programs in cell-free and bacterial systems. *Cell Syst.* S2405471221004191. https://doi.org/10.1016/j.cels.2021.10.008
Todor, H., Silvis, M.R., Osadnik, H., Gross, C.A., **2021**. Bacterial CRISPR screens for gene function. *Host-Microbe Interact. Bact. Viruses* **59**, 102–109. https://doi.org/10.1016/j.mib.2020.11.005
Villegas Kcam, M.C., Tsong, A.J., Chappell, J., **2021**. Rational engineering of a modular bacterial CRISPR–Cas activation platform with expanded target range. *Nucleic Acids Res.* **49**, 4793–4802. https://doi.org/10.1093/nar/gkab211
Walton, R.T., Christie, K.A., Whittaker, M.N., Kleinstiver, B.P., **2020**. Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* eaba8853. https://doi.org/10.1126/science.aba8853
Walton, R.T., Hsu, J.Y., Joung, J.K., Kleinstiver, B.P., **2021**. Scalable characterization of the PAM requirements of CRISPR–Cas enzymes using HT-PAMDA. *Nat. Protoc.* **16**, 1511–1547. https://doi.org/10.1038/s41596-020-00465-2

Wang, J., Teng, Y., Zhang, R., Wu, Y., Lou, L., Zou, Y., Li, M., Xie, Z.-R., Yan, Y., **2021**. Engineering a PAM-flexible SpdCas9 variant as a universal gene repressor. *Nat. Commun.* **12**, 6916. https://doi.org/10.1038/s41467-021-27290-9

Wu, G., Yan, Q., Jones, J.A., Tang, Y.J., Fong, S.S., Koffas, M.A.G., **2016**. Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. *Trends Biotechnol.* **34**, 652–664. https://doi.org/10.1016/j.tibtech.2016.02.010

Wu, Y., Liu, Y., Lv, X., Li, J., Du, G., Liu, L., **2020**. CAMERS-B: CRISPR/Cpf1 assisted multiple-genes editing and regulation system for *Bacillus subtilis*. *Biotechnol Bioeng* **117**, 1817–1825. https://doi.org/10.1002/bit.27322

Zaslaver, A., Bren, A., Ronen, M., Itzkovitz, S., Kikoin, I., Shavit, S., Liebermeister, W., Surette, M.G., Alon, U., **2006**. A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nat. Methods* **3**, 623–628. https://doi.org/10.1038/nmeth895

Zhang, S., Voigt, C.A., **2018**. Engineered dCas9 with reduced toxicity in bacteria: implications for genetic circuit design. *Nucleic Acids Res.* **46**, 11115–11125. https://doi.org/10.1093/nar/gky884

Supplementary Information

Supplementary Figures

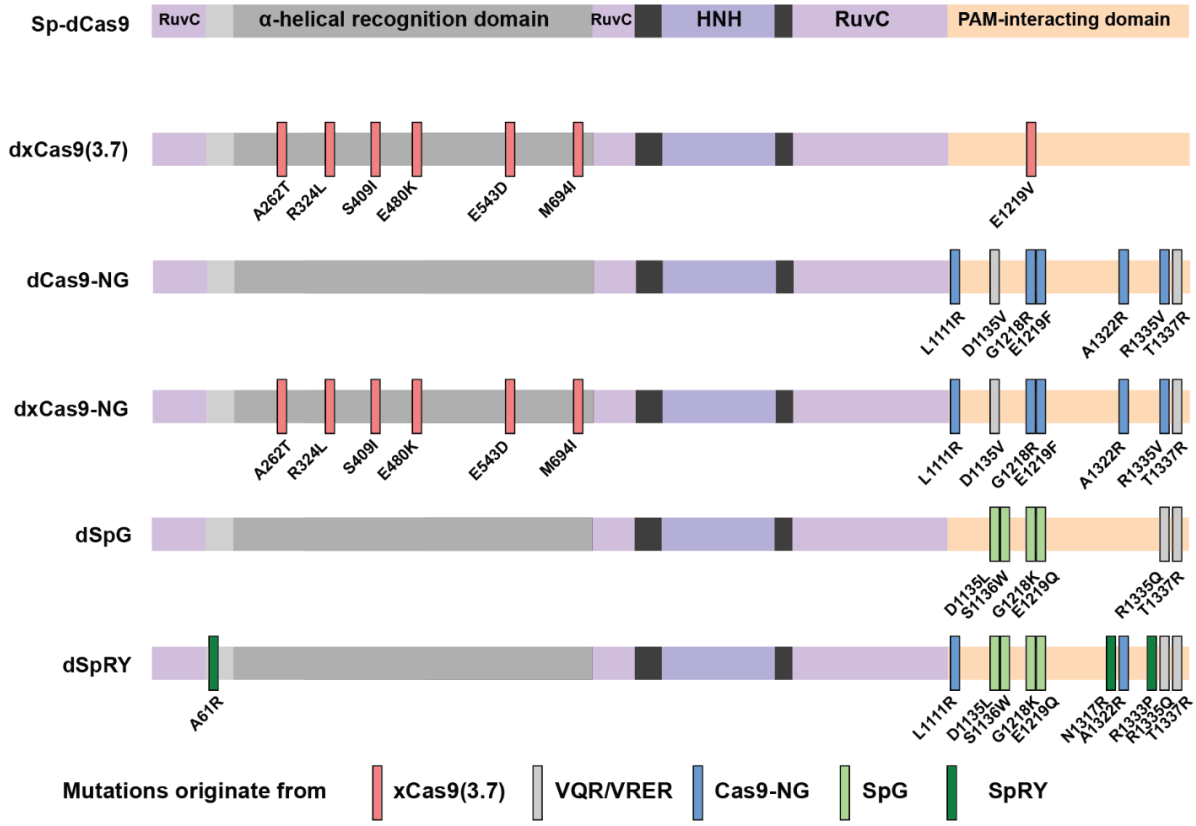


Figure 5.S1: Mutations contributing to PAM-flexibility of engineered dCas9 variants

PAM-flexible dCas9 variants contain differing sets of mutations. All variants share the catalytically-inactivating D10A and H840A mutations of dCas9 (Qi et al., 2013). Mutations of dxCas9(3.7), from phage-assisted continuous evolution (PACE), were mostly at the α-helical recognition domain (Hu et al., 2018). dCas9-NG engineering was guided by the crystal structure with mutations made in the PAM-interacting domain (Nishimasu et al., 2018). dxCas9-NG has combined mutations from dxCas9(3.7) and dCas9-NG (Legut et al., 2020). dSpG and dSpRY were developed through structure-guided engineering with engineering trajectories monitored using a high-throughput PAM determination assay (Walton et al., 2021, 2020). SpG was engineered using mutations from the VRQR variant as a starting point (Kleinstiver et al., 2016, 2015), and SpRY contains mutations from Cas9-NG (Nishimasu et al., 2018). dSpRY is an evolved version of dSpG with a preference for NRN and some NYN PAMs (R is purine nucleotides, Y is pyrimidine nucleotides) (Walton et al., 2020). Mutation colors indicate the origin of each specific mutation. Complete sequences are provided in the DNA sequences section.

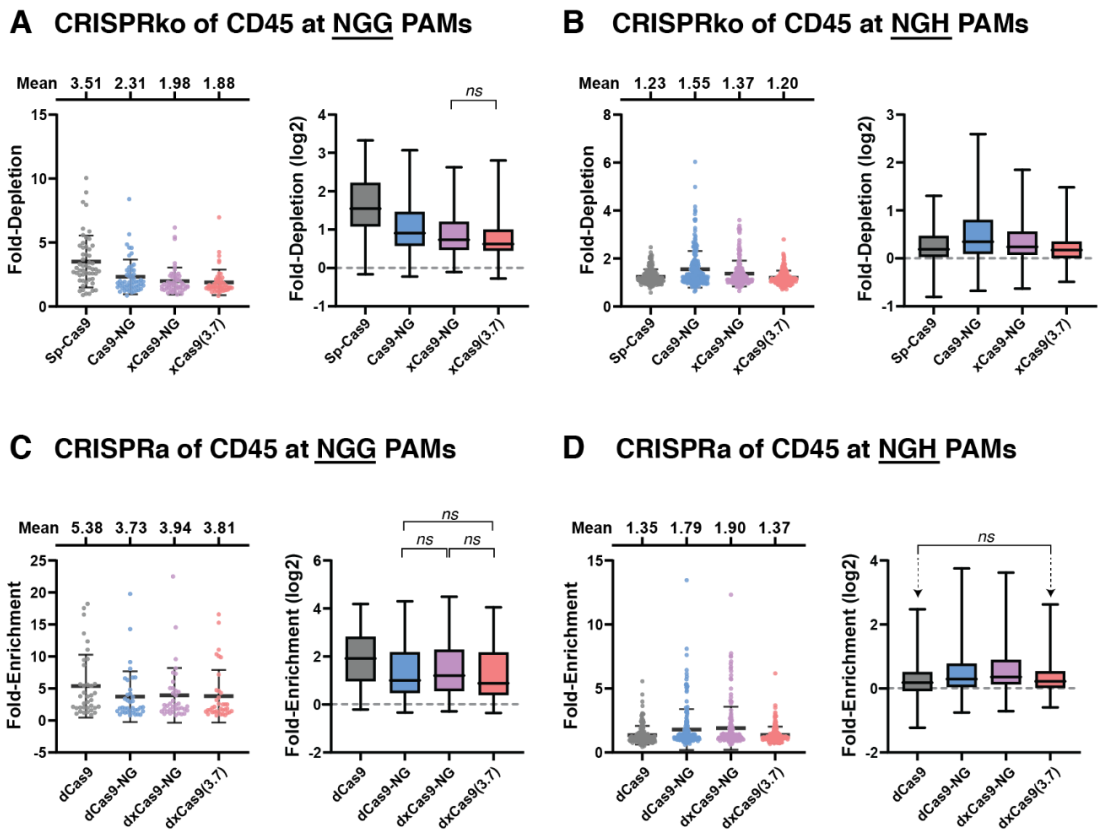


Figure 5.S2: CRISPRko and CRISPRa in mammalian systems with PAM-flexible dCas9 variants

Cas9 variants were evaluated for CRISPRko (knockout) and CRISPRa (activation) at different PAMs. sgRNA-Cas9 effector complexes were targeted to the CD45 gene and changes in expression levels were evaluated by FACS-seq (see Supplemental Methods). xCas9-NG and Cas9-NG exhibit comparable levels of PAM flexibility and both outperform xCas9(3.7). Data were plotted as fold-change scatter plots (left panel) or log2 fold-change box plots (right panel) for (A) CRISPRko at NGG PAMs, (B) CRISPRko at NGH PAMs, (C) CRISPRa at NGG PAMs, and (D) CRISPRa at NGH PAMs. In the scatter plots, bars and whiskers represent mean and standard deviation, respectively. In the box plots, two-tailed unpaired Welch's *t* test for each dCas9 pair were performed. Only non-significant comparisons (*ns*, $p > 0.05$) are indicated; all other differences (between proteins, within modalities) are significant.

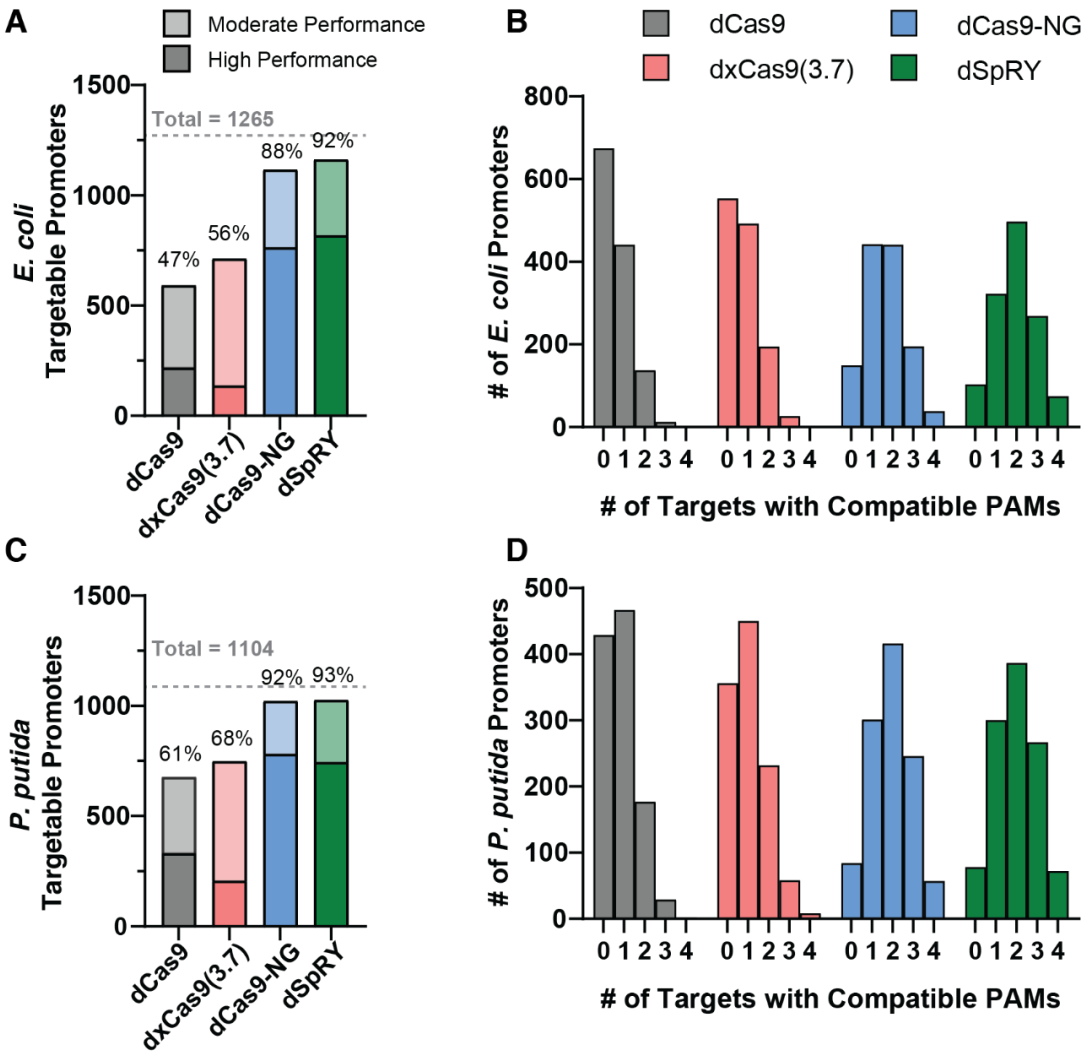


Figure 5.S3: PAM availability and distribution in *E. coli* and *P. putida* promoters

Endogenous *E. coli* and *P. putida* promoters were analyzed for predicted targetable sites with different PAM-flexible dCas9 variants (see Methods section for further details). (A) Number of targetable *E. coli* promoters for different PAM-flexible dCas9 variants. 1265 *E. coli* promoters were analyzed for the presence of a compatible PAM at any of the four effective target site positions upstream of the TSS (see Methods). “High performance” PAMs are classified as PAMs with at least 50% efficiency relative to dCas9 at NGG PAMs. “Moderate performance” PAMs are classified as PAMs with 20-50% efficiency. (B) Distribution of promoters with 0 to 4 compatible PAMs with at least moderate performance in *E. coli*. (C) Number of targetable *P. putida* promoters for different PAM-flexible dCas9 variants. 1104 *P. putida* were analyzed with the same strategy. (D) Distribution of promoters with 0 to 4 compatible PAMs with at least moderate performance in *P. putida*. In both *E. coli* and *P. putida*, dxCas9-NG and dSpRY access more endogenous promoters than dCas9 and dxCas9(3.7). As *P. putida* has a higher GC content than *E. coli* and a correspondingly higher fraction of NGN PAMs, there is an almost identical number of targetable promoters for dCas9-NG and dSpRY.

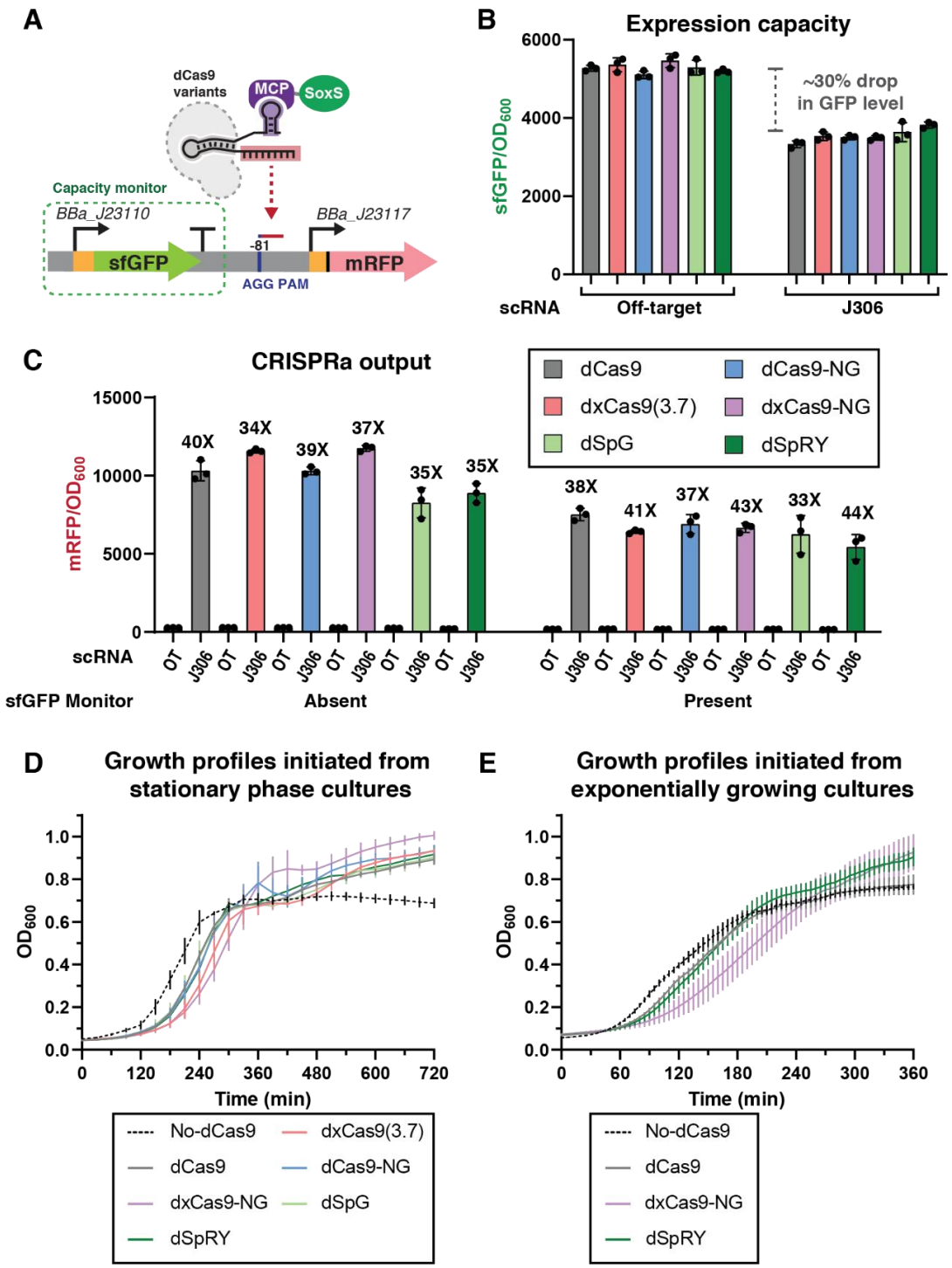


Figure 5.S4: Expression capacity and growth burden associated with dCas9 variants

(A) A constitutive, medium-strength promoter (BBa_J23110) expressing sfGFP acts as a protein expression capacity monitor. (B) CRISPRa upregulation of mRFP (using J306 scRNA) produces a ~30% decrease in the sfGFP expression capacity monitor compared to an off-target scRNA control. Similar effects were observed with all dCas9 variants. (C) The presence of the sfGFP capacity monitor causes a modest decrease in CRISPRa-mediated mRFP expression. The

basal expression levels (off-target controls) also decrease, leading to similar mRFP fold-changes in the presence or absence of the sfGFP capacity monitor. Similar effects were observed with all dCas9 variants. (D&E) Expression of all dCas9 variants caused a modest lag in *E. coli* growth relative to cells that do not express dCas9. In (D), growth profiles were obtained from 1:100 subcultures from overnight, stationary-phase cultures. In (E), growth profiles were obtained by diluting exponentially growing cells back to OD_{600} = 0.1. This experiment was performed to evaluate if the growth lag from dCas9 expression observed in (D) was caused by slow recovery from the stationary phase. We observed a similar growth lag from dCas9 expression in both (D) and (E). A recent paper describes comparable growth experiments with *E. coli* in microbioreactors; these experiments observed similar growth behavior with dCas9, dxCas9, and dSpRY (Klanschnig et al., 2022).

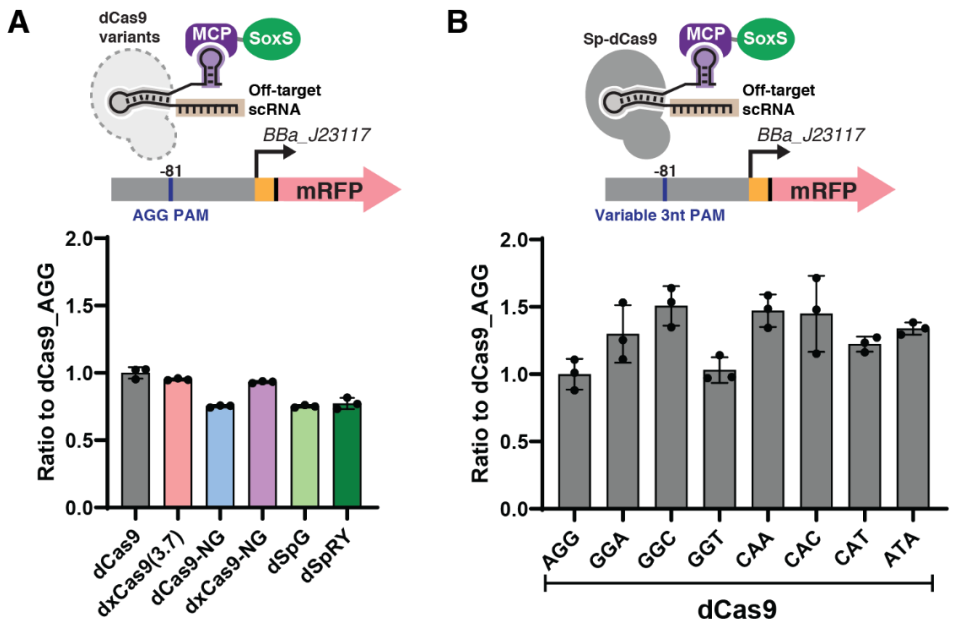


Figure 5.S5: Basal expression of the mRFP reporter gene with different dCas9 variants and PAMs sequences

(A) Expression of different dCas9 variants produces less than 2-fold changes in basal reporter expression level. This experiment was performed with AGG PAM reporter and an off-target scRNA (hAAVS1). These off-target expression levels were used as a basal expression for fold-change calculation with each dCas9 variant in Figure 5.2. (B) Modified mRFP reporter genes with alternative upstream PAM sites produce less than 2-fold changes in basal reporter expression. This experiment was performed with Sp-dCas9 and an off-target scRNA (hAAVS1). Values in panel A and B represent the mean ± standard deviation calculated from n = 3.

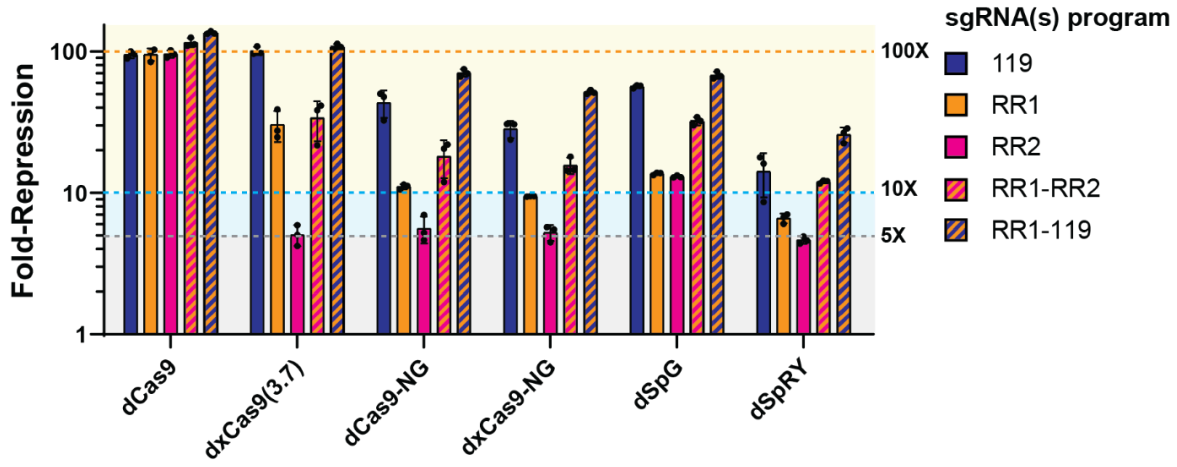


Figure 5.S6: CRISPRi with PAM-flexible dCas9 variants

CRISPRi for PAM-flexible dCas9 variants was tested on an mRFP reporter gene with sgRNAs targeting the promoter (119) or coding sequence (RR1 and RR2) with one or two sgRNAs expressed, each targeting sites harboring NGG PAMs. dCas9 exhibited the highest repression efficiency (>90-fold) among all variants. When only a single sgRNA is expressed, the sgRNA targeting the promoter region (119) produces the largest repression effect with all dCas9 variants. The addition of the second sgRNA (RR1-RR2 or RR1-119) led to significant improvement in repression for all variants. Values represent the mean ± standard deviation calculated from n = 3.

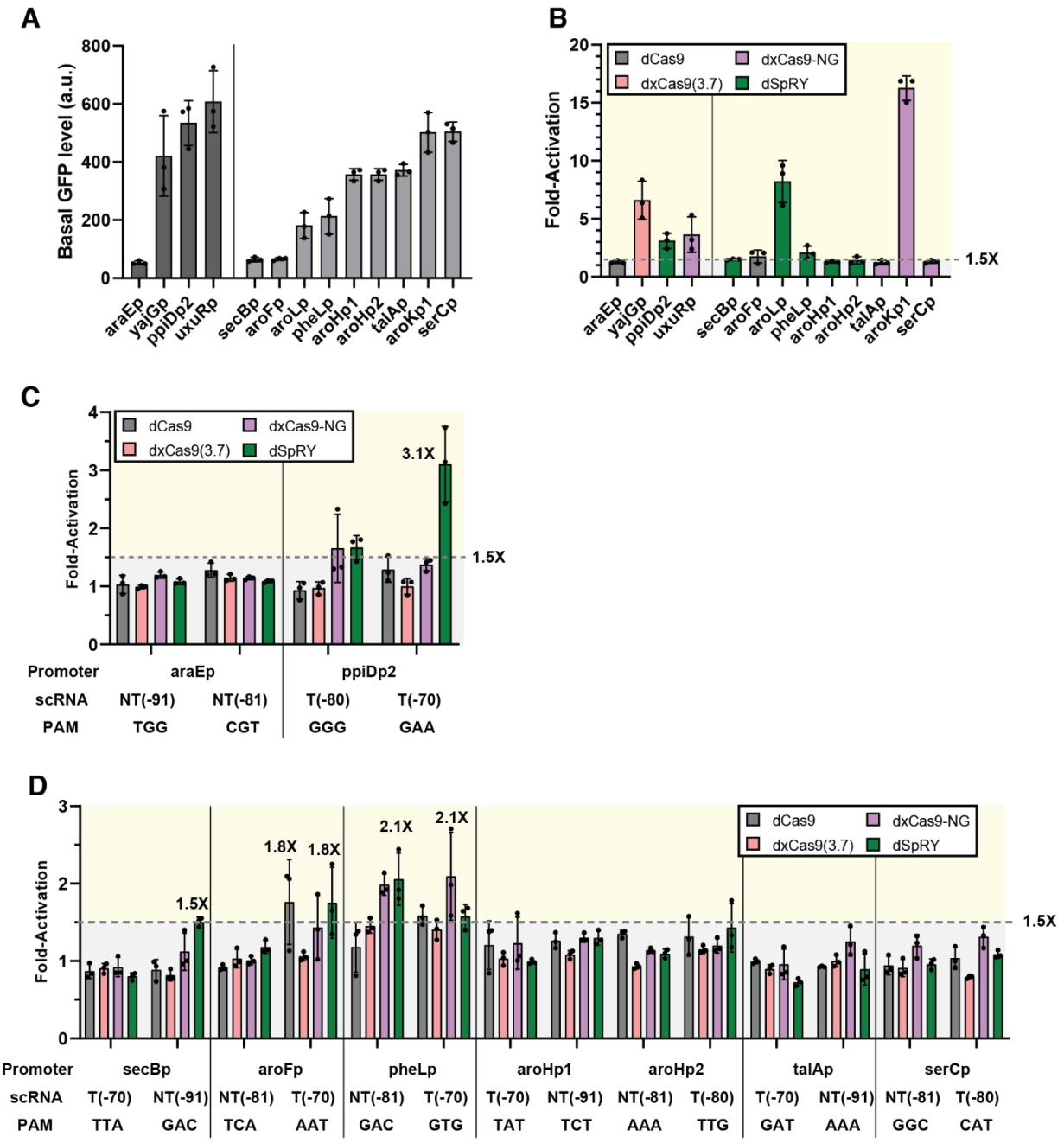


Figure 5.S7: CRISPRa at endogenous promoters with PAM-flexible dCas9 variants

(A) Basal expression levels for endogenous promoters (left: promoters previously tested with dxCas9(3.7) (Fontana et al., 2020a); right: new promoters examined in this work). Basal expression levels were measured in a strain expressing the parent dCas9 and an off-target scRNA (hAAVS1). (B) Maximum fold-activation for each endogenous promoter. Bar color indicates the dCas9 variant with the highest fold-activation for the corresponding endogenous promoter. 8 out of 13 endogenous promoters can be activated (>1.5-fold) with PAM-flexible dCas9 variants. (C) CRISPRa at endogenous promoters previously tested with dxCas9(3.7) (araEp and ppiDp2) using NGG or non-NGG PAMs. See main text Figure 5.4B for the other two promoters

from this set (yajGp and uxuRp). (D) CRISPRa at endogenous promoters involved in aromatic amino acid biosynthesis using non-NGG PAMs (secBp, aroFp, pheLp, aroHp1, aroHp2, talAp, and serCp). See main text Figure 5.4C for the other two promoters from the aromatic amino acid biosynthesis set (aroLp and aroKp1). Data were collected by flow cytometry and fold-activation was calculated compared to the strain expressing the corresponding dCas9 variant and an off-target hAAVS1 scRNA. Values in panels A-D represent the mean ± standard deviation calculated from $n = 3$.

Supplementary Tables

Table 5.S1: *E. coli* strains

Strain	Description	Genotype
MG1655	Wildtype *E. coli strain*	F- λ- ilvG- rfb-50 rph-1
CD38	Integrated BBa_J23119-mRFP used for CRISPRi (adapted from JF01 (Fontana et al., 2018a))	MG1655 *rbsAR*::BBa_J23119-mRFP

Table 5.S2: Selected *E. coli* plasmids

Plasmid	Marker	Replicon	Promoter	Gene	Terminator	Reference
pJF043	CmR	p15A	None	None	None	This work
pCD442	CmR	p15A	1) Sp.pCas9 2) BBa_J23107	1) dCas9 2) MCP-SoxS	1) BBa_B0015 2) BBa_B1002	(Fontana et al., 2020a)
pCD564	CmR	p15A	1) Sp.pCas9 2) BBa_J23107	1) dxCas9(3.7) 2) MCP-SoxS	1) BBa_B0015 2) BBa_B1002	(Fontana et al., 2020a)
pCK668	CmR	p15A	1) Sp.pCas9 2) BBa_J23107	1) dCas9-NG 2) MCP-SoxS	1) BBa_B0015 2) BBa_B1002	This work
pCK669	CmR	p15A	1) Sp.pCas9 2) BBa_J23107	1) dxCas9-NG 2) MCP-SoxS	1) BBa_B0015 2) BBa_B1002	This work
pCK340	CmR	p15A	1) Sp.pCas9 2) BBa_J23107	1) dSpG 2) MCP-SoxS	1) BBa_B0015 2) BBa_B1002	This work
pCK341	CmR	p15A	1) Sp.pCas9 2) BBa_J23107	1) dSpRY 2) MCP-SoxS	1) BBa_B0015 2) BBa_B1002	This work
pCK085.scRNA	CmR	p15A	1) Sp.pCas9 2) BBa_J23107 3) BBa_J23119	1) dCas9 2) MCP-SoxS 3) scRNA	1) BBa_B0015 2) BBa_B1002 3) TrrnB	(Tickman et al., 2021)
pCK281.scRNA	CmR	p15A	1) Sp.pCas9 2) BBa_J23107 3) BBa_J23119	1) dxCas9(3.7) 2) MCP-SoxS 3) scRNA	1) BBa_B0015 2) BBa_B1002 3) TrrnB	This work
pCK670.scRNA	CmR	p15A	1) Sp.pCas9 2) BBa_J23107 3) BBa_J23119	1) dCas9-NG 2) MCP-SoxS 3) scRNA	1) BBa_B0015 2) BBa_B1002 3) TrrnB	This work
pCK671.scRNA	CmR	p15A	1) Sp.pCas9 2) BBa_J23107 3) BBa_J23119	1) dxCas9-NG 2) MCP-SoxS 3) scRNA	1) BBa_B0015 2) BBa_B1002 3) TrrnB	This work

pCK363. scRNA	CmR	p15A	1) Sp.pCas9 2) BBa_J23107 3) BBa_J23119	1) dSpG 2) MCP-SoxS 3) scRNA	1) BBa_B0015 2) BBa_B1002 3) TrrnB	This work
pCK364. scRNA	CmR	p15A	1) Sp.pCas9 2) BBa_J23107 3) BBa_J23119	1) dSpRY 2) MCP-SoxS 3) scRNA	1) BBa_B0015 2) BBa_B1002 3) TrrnB	This work
pJF143. J3	AmpR	pSC101**	J3-BBa_J23117	mRFP	BBa_B0015	(Fontana et al., 2020a)
pCK760	AmpR	pSC101**	1) BBa_J23110 2) J3-BBa_J23117	1) sfGFP 2) mRFP	1) BBa_B0015 2) BBa_B0015	This work
pCK284. NNN	AmpR	pSC101**	J3(NNN-PAM)-BBa_J23117	mRFP	BBa_B0015	This work
pCD443. scRNA	AmpR	ColE1	BBa_J23119	scRNA	TrrnB	This work
pCK411. scRNA/ sgRNA	AmpR	ColE1	BBa_J23119	scRNA/sgRNA	BBa_K2680405	This work
pCK590. XXX	KmR	pSC101	Endogenous (strand +)	GFPmut2		Same as (Zaslaver et al., 2006)
pCK591. XXX	KmR	pSC101	Endogenous (strand –)	GFPmut2		Same as (Zaslaver et al., 2006)

Table 5.S3: scRNA and sgRNA target sites

sc/sgRNA	DNA sequence	Target strand[a]	Distance to TSS[b]	PAM
J306	TTGTGTCCAGAACGCTCCGT	NT	-81	AGG
hAAVS1	GGGGCCACTAGGGACAGGAT	Off-target	NA	–
RR1	AACTTTCAGTTTAGCGGTCT	NT	151[c]	GGG
RR2	TGGAACCGTACTGGAACTGC	NT	215[c]	GGG
119	AATTCAGATCTATTATACCT	NT	-15	AGG
yajGp_Y1	TTGACGAAATAATCGCCCCT	NT	-81	GGT
yajGp_Y2	CATCAGTGTTTCTTTTACCA	T	-79	GGG
uxuRp_U1	TGATTGACCAGTAAGTCTGT	NT	-81	AGG
uxuRp_U2	GATTACCCTACAGACTTACT	T	-70	GGT
ppiDp2_D1	ACTAAGCGTTGTCCCCAGTG	T	-80	GGG
ppiDp2_D2	GTCCCCAGTGGGGATGTGAC	T	-70	GAA
araEp_E1	TGCGACATGTCGTTATGTGA	NT	-91	TGG
araEp_E2	ATTAAATTGCTGCGACATGT	NT	-81	CGT
pheLp_X06	GCGATACACTCAATATAAAG	NT	-81	GAC
pheLp_X07	AGAGTAGTCCTTTATATTGA	T	-70	GTG
secBp_X07	CACCACGGTTCCCCAGATTT	T	-70	TTA
secBp_X08	AATCTGGGGAACCGTGGTGC	NT	-91	GAC
serCp_X06	ACCGTTGAGGGCAAAAATGT	NT	-81	GGC
serCp_X09	CTTTTGTGTGATGCAAGCCA	T	-80	CAT
talAp_X07	GGTAATAATCCTATAACACT	T	-70	GAT
talAp_X08	GTGTTATAGGATTATTACCA	NT	-91	AAA
aroFp_X06	CAGGCAATTTAGTCGCGCTT	NT	-81	TCA
aroFp_X07	AAGGGTTGAAAGCGCGACTA	T	-70	AAT
aroLp_X07	TGGTGGCTGGAAGTGCAACG	T	-70	TAG
aroLp_X08	GTTGCACTTCCAGCCACCAC	NT	-91	TTC
aroHp1_X07	ATTGCCACCAAGATCCTCGA	T	-70	TAT
aroHp1_X08	CGAGGATCTTGGTGGCAATC	NT	-91	TCT
aroHp2_X06	GTGGTTAGCATGATAACAAA	NT	-81	AAA
aroHp2_X09	TAGTGCATTAGCTTATTTTT	T	-80	TTG
aroKp1_X07	GAGTAAACAGCCGTAAAAGC	T	-70	GGT
aroKp1_X08	CTTTTACGGCTGTTTACTCA	NT	-91	CTG

[a] Template strand (T) or non-template strand (NT).

[b] Distance from the 3' end of the guide site (PAM proximal) to the TSS. For synthetic promoters (BBa_J23117 or BBa_J23119, http://parts.igem.org), the TSS is immediately downstream.

[c] For RR1 and RR2 sgRNAs, the positive number refers to the distance downstream of the TSS.

Table 5.S4: Predicted compatible PAMs for each dCas9 variants

PAM	Sp-Cas9	xCas9(3.7)	Cas9-NG	SpG	SpRY
AAA	0.18%	2.75%	19.51%	N.A.	27.97%
AAC	0.06%	1.52%	17.62%	N.A.	69.73%
AAG	46.32%	19.04%	43.10%	N.A.	28.35%
AAT	0.25%	3.80%	28.04%	N.A.	6.58%
ACA	0.61%	0.23%	5.03%	N.A.	19.84%
ACC	-0.50%	-0.15%	2.18%	N.A.	N.A.
ACG	17.85%	7.43%	12.10%	N.A.	N.A.
ACT	-0.45%	0.47%	7.07%	N.A.	N.A.
AGA	27.95%	29.33%	60.81%	65.89%	58.50%
AGC	7.82%	16.70%	52.14%	75.97%	56.19%
AGG	99.80%	62.37%	72.89%	38.75%	21.46%
AGT	1.82%	23.39%	68.90%	78.02%	59.55%
ATA	0.78%	0.21%	12.29%	N.A.	43.63%
ATC	-0.56%	-0.34%	5.07%	N.A.	3.20%
ATG	22.04%	5.45%	28.21%	N.A.	N.A.
ATT	-0.13%	0.55%	10.60%	N.A.	N.A.
CAA	-0.01%	6.21%	27.29%	N.A.	53.17%
CAC	-0.53%	2.69%	26.45%	N.A.	19.04%
CAG	44.62%	25.71%	47.15%	N.A.	18.17%
CAT	-0.63%	6.25%	32.92%	N.A.	22.44%
CCA	-0.13%	-0.21%	2.06%	N.A.	N.A.
CCC	-0.65%	-0.31%	0.89%	N.A.	1.01%

CCG	11.47%	3.99%	6.78%	N.A.	9.06%
CCT	-0.22%	-0.37%	3.09%	N.A.	N.A.
CGA	30.35%	36.77%	59.10%	78.85%	N.A.
CGC	10.29%	24.23%	52.31%	70.58%	N.A.
CGG	102.08%	68.15%	72.25%	48.14%	15.80%
CGT	2.47%	32.69%	69.50%	49.23%	23.77%
CTA	-0.23%	0.64%	9.53%	N.A.	20.11%
CTC	-0.52%	0.19%	4.65%	N.A.	25.22%
CTG	17.31%	3.94%	23.82%	N.A.	19.44%
CTT	-0.44%	-0.06%	9.93%	N.A.	2.30%
GAA	1.02%	11.96%	37.31%	N.A.	100.83%
GAC	-0.13%	4.71%	31.57%	N.A.	86.73%
GAG	53.38%	35.04%	57.91%	N.A.	55.14%
GAT	0.05%	10.29%	42.64%	N.A.	35.73%
GCA	-0.05%	0.76%	7.21%	N.A.	55.08%
GCC	-0.26%	0.28%	4.37%	N.A.	21.67%
GCG	13.01%	1.67%	11.22%	N.A.	N.A.
GCT	-0.58%	0.22%	10.36%	N.A.	43.08%
GGA	30.31%	37.93%	60.78%	62.42%	33.41%
GGC	12.24%	22.95%	53.55%	73.33%	41.75%
GGG	101.50%	70.07%	76.36%	67.74%	54.37%
GGT	7.16%	33.76%	69.40%	76.14%	65.25%
GTA	0.05%	1.79%	20.16%	N.A.	37.98%
GTC	-0.34%	0.56%	11.00%	N.A.	N.A.

GTG	23.36%	6.76%	37.60%	N.A.	50.91%
GTT	-0.15%	0.29%	16.20%	N.A.	4.66%
TAA	0.62%	1.36%	16.05%	N.A.	69.86%
TAC	-0.62%	0.15%	11.41%	N.A.	72.28%
TAG	36.56%	11.45%	34.32%	N.A.	30.23%
TAT	-0.22%	1.52%	19.26%	N.A.	44.40%
TCA	-0.36%	0.01%	1.27%	N.A.	N.A.
TCC	-0.81%	-0.28%	0.72%	N.A.	N.A.
TCG	10.87%	3.10%	5.79%	N.A.	N.A.
TCT	-0.15%	-0.37%	2.47%	N.A.	7.00%
TGA	28.62%	29.68%	61.71%	73.33%	45.34%
TGC	7.91%	15.40%	53.76%	88.36%	65.05%
TGG	96.62%	62.22%	72.08%	95.59%	69.12%
TGT	2.32%	24.78%	70.19%	103.69%	N.A.
TTA	0.06%	0.53%	3.71%	N.A.	N.A.
TTC	-1.16%	-0.09%	0.44%	N.A.	N.A.
TTG	9.19%	2.01%	11.05%	N.A.	0.14%
TTT	-0.46%	-0.14%	1.55%	N.A.	11.70%

The predicted PAM compatibility of each PAM-Cas9 pair was reported as indel frequency (in % units) relative to the benchmark efficiency of Sp-Cas9 at NGG PAMs from each experiment (see Supplementary Methods) (Kim et al., 2020; Walton et al., 2020). Each data point was color-coded based on predicted efficiency: Green (high efficiency, >50% of benchmark), blue (moderate efficiency, 20%–50%), and gray (low efficiency, less than 20%). NGG PAMs are highlighted in yellow. N.A. means data is not available. Some of the reported efficiency values are small negative numbers, likely due to negligible editing frequencies indistinguishable from the background. Comparable data for indel frequencies with xCas9-NG are not available; for this work we assumed Cas9-NG and xCas9-NG were similar based on their comparable performance in nuclease assays (Figure S2).

Table 5.S5: CRISPRa on endogenous *E. coli* promoters

Promoter	Max. FA	dCas9 variant	Position	PAM
yajGp	**6.6** 6.2 5.1	**dxCas9(3.7)** dxCas9-NG dSpRY	NT(-81)	GGT
uxuRp	**3.6** 2.6	**dxCas9-NG** dSpRY	T(-70)	GGT
araEp	1.3	dxCas9-NG	NT(-81)	CGT
ppiDp2	**3.1**	**dSpRY**	T(-70)	GAA
aroKp1	**16.3** 12.8	**dxCas9-NG** dxCas9(3.7)	T(-70)	GGT
aroLp	**8.2** 5.5	**dSpRY** dCas9	T(-70)	TAG
aroF	**1.8** **1.8**	**dCas9** **dSpRY**	T(-70)	AAT
aroHp1	1.3	dxCas9-NG	NT(-91)	TCT
aroHp2	1.4	dSpRY	NT(-81)	AAA
pheLp	**2.1** **2.1**	**dxCas9-NG** **dSpRY**	T(-70) NT(-81)	GTG GAC
secBp	**1.5**	**dSpRY**	NT(-91)	GAC
serCp	1.3	dxCas9-NG	T(-80)	CAT
talAp	1.2	dxCas9-NG	NT(-91)	AAA

The top two conditions were shown for each promoter, if >1.5-fold. The best conditions for each promoter is shown in bold. yajGp has three similar activation levels. Promoters with <1.5-fold activation were shaded in gray.

Supplementary Methods

Evaluating PAM accessibility from CRISPR knockout and CRISPR activation screens in mammalian systems

To predict the PAM accessibility for different Cas9 variants, we used data from CRISPR knockout (CRISPRko) and CRISPR activation (CRISPRa) screens in mammalian cells. Previously, screening data has been reported for three Cas9 variants (Sp-Cas9, xCas9(3.7), and Cas9-NG) (Fontana et al., 2020a). We report here a corresponding screen performed using the same approach with four Cas9 variants: Sp-Cas9, xCas9(3.7), Cas9-NG, and xCas9-NG. Lentiviral sgRNA libraries were constructed to target the CD45 gene, spanning the 3 kb region surrounding the TSS and constitutive CDS exons, using all possible 20-mer sequences upstream of an NG PAM sequence, as described before (Fontana et al., 2020a). In addition to sgRNAs, each vector also contained a Cas9 variant (nuclease for CRISPRko and catalytically dead Cas9 fusion with VPR transcriptional activators for CRISPRa) and a short barcode downstream of the sgRNA to identify the Cas9 variant in the same Illumina read as the sgRNA. sgRNA library cloning, lentivirus production and cell transduction were done separately for each Cas9 variant. The transduced and selected cells were only pooled together prior to FACS sorting based on the CD45 protein expression. CRISPRko was performed in K562 cells and CRISPRa was performed in A375 cells; both cell lines were obtained from ATCC. The presort samples and 10% top/bottom bins were analyzed by high-throughput Illumina sequencing. To calculate fold-depletion for CRISPRko, the relative frequency (normalized to sequencing depth and median frequency of non-targeting sgRNAs for each Cas9 variant) of each sgRNA from the bottom 10% bin (lowest expression level) was divided by its corresponding frequency in the top 10% bin. To calculate fold-enrichment for CRISPRa, the relative frequency of each sgRNA from the top 10% bin (highest expression) was divided by its corresponding frequency in the bottom 10% bin. For CRISPRko, only sgRNAs targeting within the CDS were included in the analysis; for CRISPRa, only sgRNAs targeting within the core promoter region (as determined previously (Legut et al., 2020), chr1:198638250-198639226) were included. The data were then filtered for designated PAMs (NGG or NGH) to visualize the data as shown in Figure 5.S2.

PAM compatibility analysis

For each dCas9 variant, the ability to target each of the 64 possible 3-nucleotide PAMs was predicted based on previously reported CRISPR nuclease (CRISPRko) data in mammalian systems (Kim et al., 2020; Walton et al., 2020). Genome editing efficiencies, measured as indel frequencies, have been reported for Sp-Cas9, xCas9(3.7), and Cas9-NG (Kim et al., 2020) and for Sp-Cas9, SpG, and SpRY (Walton et al., 2020). For each Cas9 dataset, sgRNAs with the same 3-nt PAM were grouped together. We then calculated an average indel frequency for all sgRNAs with the same PAM. We assessed PAM compatibility by comparison to the benchmark indel frequency of Sp-Cas9 at NGG PAMs (Table 5.S4). “High performance” PAMs were those with >50% indel frequency relative to the benchmark. “Moderate performance” PAMs were those with 20-50% efficiency relative to the benchmark. PAMs that were not tested in the previously-reported nuclease assays were assumed to be incompatible (Walton et al., 2020).

Endogenous promoter and scRNA selection strategy

Gene candidates were selected from metabolic pathways related to aromatic amino acid biosynthesis, including the relevant genes in central metabolic pathways. We selected only promoters that are regulated by sigma70 and are available in the previously characterized *E. coli* endogenous promoter library (obtained from Horizon Discovery) (Zaslaver et al., 2006). To identify potentially activatable promoters with moderately-weak basal expression levels, we compared basal expression to yajGp, which was the strongest endogenous promoter that could be activated in previous work (Fontana et al., 2020a). We eliminated any promoters with >10-fold higher expression levels than yajGp, yielding 9 promoters: secBp, aroFp, aroLp, pheLp, aroHp1, aroHp2, talAp, aroKp1, and serCp. Some of these promoters regulate multi-gene operons. Complete sequences of the endogenous promoters used in this study are available in the DNA sequences section below.

Out of selected genes, aroH contains two putative promoters — aroHp1/aroHp2. In Figure 5.S6A, the basal expression levels for aroHp1 and aroHp2 were assumed to be the same value. Expression levels produced by each individual promoter cannot be determined from the fluorescent reporter used in this study because both promoters are upstream of the same fluorescent reporter gene.

In the secBp promoter constructed for the *E. coli* promoter library (Zaslaver et al., 2006), only the sequence to -7 bases from the TSS was included. To include enough sequence for scRNA target sites, we constructed a promoter extending to -137 bases from the TSS.

For each promoter, four scRNAs were identified according to the previously described target site preference. X06 and X08 represent the scRNAs targeting the non-template strand at -81 and -91 positions (Table 5.S3). X07 and X09 represent the scRNAs targeting the template strand at -70 and -80 positions. One target site from the template strand and another from the non-template strand were selected for further analysis based on which PAM was predicted to be accessible to the highest number of dCas9 variants. Accessibility was assessed based on the moderate performance threshold cutoff (Table 5.S4).

DNA sequences

Reporter

>J3(**PAM**)-BBa_J23117-RBS-mRFP

AGCATTTGCGATCATTCACGCAGCGCTTATTCAGTTGCTCACTGCGATGTCATAATCATCGCTACGAGCTGTGAAAG
ATGCATAAAGCTCGTACGACGCGTTCGCTCGTCTCCTCACTTCT**NNN**ACGGAGCGTTCTGGACACAACGTCGTCTTG
AAGTTGCGATTATAGATTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGCGAATTCATTAAAGAGGAGAAAGGTAC
C**ATG**GCGAGTAGCGAAGACGTTATCAAAGAGTTCATGCGTTTCAAAGTTCGTATGGAAGGTTCCGTTAACGGTCACG
AGTTCGAAATCGAAGGTGAAGGTGAAGGTCGTCCGTACGAAGGTACCCAGACCGCTAAACTGAAAGTTACCAAAGGT
GGTCCGCTGCCGTTCGCTTGGGACATCCTGTCCCCGCAGTTCCAGTACGGTTCCAAAGCTTACGTTAAACACCCGGC
TGACATCCCGGACTACCTGAAACTGTCCTTCCCGGAAGGTTTCAAATGGGAACGTGTTATGAACTTCGAAGACGGTG
GTGTTGTTACCGTTACCCAGGACTCCTCCCTGCAAGACGGTGAGTTCATCTACAAAGTTAAACTGCGTGGTACCAAC
TTCCCGTCCGACGGTCCGGTTATGCAGAAAAAAACCATGGGTTGGGAAGCTTCCACCGAACGTATGTACCCGGAAGA
CGGTGCTCTGAAAGGTGAAATCAAAATGCGTCTGAAACTGAAAGACGGTGGTCACTACGACGCTGAAGTTAAAACCA
CCTACATGGCTAAAAAACCGGTTCAGCTGCCGGGTGCTTACAAAACCGACATCAAACTGGACATCACCTCCCACAAC
GAAGACTACACCATCGTTGAACAGTACGAACGTGCTGAAGGTCGTCACTCCACCGGTGCTTAA

PAM location is bolded/underlined above

>BBa_J23110-RBS-sfGFP

TTTACGGCTAGCTCAGTCCTAGGTACAATGCTAGCGAATTCATTAAAGAGGAGAAAGGTACC**ATG**AGCAAAGGAGAA
GAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCCGTGG
AGAGGGTGAAGGTGATGCTACAAACGGAAAACTCACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCGT
GGCCAACACTTGTCACTACTCTGACCTATGGTGTTCAATGCTTTTCCCGTTATCCGGATCACATGAAACGGCATGAC
TTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAACGCACTATATCTTTCAAAGATGACGGGACCTACAAGAC
GCGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATCGTATCGAGTTAAAGGGTATTGATTTTAAAGAAGATG
GAAACATTCTTGGACACAAACTCGAGTACAACTTTAACTCACACAATGTATACATCACGGCAGACAAACAAAAGAAT
GGAATCAAAGCTAACTTCAAAATTCGCCACAACGTTGAAGATGGTTCCGTTCAACTAGCAGACCATTATCAACAAAA
TACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCGACACAATCTGTCCTTTCGAAAGATC
CCAACGAAAAGCGTGACCACATGGTCCTTCTTGAGTTTGTAACTGCTGCTGGGATTACACATGGCATGGATGAGCTC
TACAAATAA

>BBa_J23119-RBS-mRFP

TTGACAGCTAGCTCAGTCCTAGGTATAATAGATCTGAATTCATTAAAGAGGAGAAAGGTACC**ATG**GCGAGTAGCGAA
GACGTTATCAAAGAGTTCATGCGTTTCAAAGTTCGTATGGAAGGTTCCGTTAACGGTCACGAGTTCGAAATCGAAGG
TGAAGGTGAAGGTCGTCCGTACGAAGGTACCCAGACCGCTAAACTGAAAGTTACCAAAGGTGGTCCGCTGCCGTTCG
CTTGGGACATCCTGTCCCCGCAGTTCCAGTACGGTTCCAAAGCTTACGTTAAACACCCGGCTGACATCCCGGACTAC
CTGAAACTGTCCTTCCCGGAAGGTTTCAAATGGGAACGTGTTATGAACTTCGAAGACGGTGGTGTTGTTACCGTTAC
CCAGGACTCCTCCCTGCAAGACGGTGAGTTCATCTACAAAGTTAAACTGCGTGGTACCAACTTCCCGTCCGACGGTC
CGGTTATGCAGAAAAAAACCATGGGTTGGGAAGCTTCCACCGAACGTATGTACCCGGAAGACGGTGCTCTGAAAGGT
GAAATCAAAATGCGTCTGAAACTGAAAGACGGTGGTCACTACGACGCTGAAGTTAAAACCACCTACATGGCTAAAAA
ACCGGTTCAGCTGCCGGGTGCTTACAAAACCGACATCAAACTGGACATCACCTCCCACAACGAAGACTACACCATCG
TTGAACAGTACGAACGTGCTGAAGGTCGTCACTCCACCGGTGCTTAA

119, RR1, and RR2 targets are underlined above

>UP_secBp-UTR-RBS-GFPmut2

GATGGCAACGCCGCCAAGCGTGAAGAGATGATCAAACGCAGCGGTCGCACCACGGTTCCCCAGATTTTTATTGACGC
ACAGCACATTGGCGGCTGTGATGACTTGTATGCATTGGATGCACGTGGTGGACTGGATCCCCTGCTGAAATAACGTG
TGAACGTTGGCATTACATTGCGCAGTATTTAAGGACAACACTTAAGGGTTTTCTACACATGTCAGAACAAAACAACA

CTGAAATGACTTTCCAGATCCAACGTATTTATACCAAGGATATCTCTTTCGAAGCGCCGCTCGAGAGATCCTCTAGA
TTTAAGAAGGAGATATACAT**ATG**AGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGAT
GGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAACTTACCCTTAA
ATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGCGTATGGTCTTCAATGCT
TTGCGAGATACCCAGATCATATGAAACAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGA
ACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAG
AATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCAC
ACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGAT
GGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCA
TTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAA
CAGCTGCTGGGATTACCCATGGTATGGATGAATTGTACAAATAA

Cas9 variants (bolded/underlined mutations)

>*Sp*.pCas9-dCas9-dblTerm

TTACGAAATCATCCTGTGGAGCTTAGTAGGTTTAGCAAGATGGCAGCGCCTAAATGTAGAATGATAAAAGGATTAAG
AGATTAATTTCCCTAAAAATGATAAAACAAGCGTTTTGAAAGCGCTTGTTTTTTTGGTTTGCAGTCAGAGTAGAATA
GAAGTATCAAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATG
CTGTTTTGAATGGTTCCAACAAGATTATTTTATAACTTTTATAACAAATAATCAAGGAGAAATTCAAAGAAATTTAT
CAGCCATAAAACAATACTTAATACTATAGAATGATAACAAAATAAACTACTTTTTAAAAGAATTTTGTGTTATAATC
TATTTATTATTAAGTATTGGGTAATATTTTTTGAAGAGATATTTTGAAAAAGAAAAATTAAAGCATATTAAACTAAT
TTCGGAGGTCATTAAAACTATTATTGAAATCATCAAACTCATTATGGATTTAATTTAAACTTTTTATTTTAGGAGGC
AAAA**ATG**GATAAGAAATACTCAATAGGCTTA**gct**ATCGGCACAAATAGCGTCGGATGGGCGGTGATCACTGATGAAT
ATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATACAGACCGCCACAGTATCAAAAAAAATCTTATAGGGGCT
CTTTTATTTGACAGTGGAGAGACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAA
GAATCGTATTTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTCTTTCATCGACTTG
AAGAGTCTTTTTTGGTGGAAGAAGACAAGAAGCATGAACGTCATCCTATTTTTGGAAATATAGTAGATGAAGTTGCT
TATCATGAGAAATATCCAACTATCTATCATCTGCGAAAAAAATTGGTAGATTCTACTGATAAAGCGGATTTGCGCTT
AATCTATTTGGCCTTAGCGCATATGATTAAGTTTCGTGGTCATTTTTTGATTGAGGGAGATTTAAATCCTGATAATA
GTGATGTGGACAAACTATTTATCCAGTTGGTACAAACCTACAATCAATTATTTGAAGAAAACCCTATTAACGCAAGT
GGAGTAGATGCTAAAGCGATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATCTCATTGCTCAGCTCCC
CGGTGAGAAGAAAAATGGCTTATTTGGGAATCTCATTGCTTTGTCATTGGGTTTGACCCCTAATTTTAAATCAAATT
TTGATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAAGATACTTACGATGATGATTTAGATAATTTATTGGCGCAA
ATTGGAGATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATTTATCAGATGCTATTTTACTTTCAGATATCCTAAG
AGTAAATACTGAAATAACTAAGGCTCCCCTATCAGCTTCAATGATTAAACGCTACGATGAACATCATCAAGACTTGA
CTCTTTTAAAAGCTTTAGTTCGACAACAACTTCCAGAAAAGTATAAAGAAATCTTTTTTGATCAATCAAAAAACGGA
TATGCAGGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTATAAATTTATCAAACCAATTTTAGAAAAAATGGA
TGGTACTGAGGAATTATTGGTGAAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGGACCTTTGACAACGGCTCTA
TTCCCCATCAAATTCACTTGGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAAGAC
AATCGTGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCCATTGGCGCGTGGCAATAGTCG
TTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTT
CAGCTCAATCATTTATTGAACGCATGACAAACTTTGATAAAAATCTTCCAAATGAAAAAGTACTACCAAAACATAGT
TTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTGAAGGAATGCGAAAACCAGC
ATTTCTTTCAGGTGAACAGAAGAAAGCCATTGTTGATTTACTCTTCAAAACAAATCGAAAAGTAACCGTTAAGCAAT
TAAAAGAAGATTATTTCAAAAAAATAGAATGTTTTGATAGTGTTGAAATTTCAGGAGTTGAAGATAGATTTAATGCT
TCATTAGGTACCTACCATGATTTGCTAAAAATTATTAAAGATAAAGATTTTTTGGATAATGAAGAAAATGAAGATAT
CTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGAGATGATTGAGGAAAGACTTAAAACATATGCTC
ACCTCTTTGATGATAAGGTGATGAAACAGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTG
ATTAATGGTATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTGAAATCAGATGGTTTTGCCAATCGCAA
TTTTATGCAGCTGATCCATGATGATAGTTTGACATTTAAAGAAGACATTCAAAAAGCACAAGTGTCTGGACAAGGCG
ATAGTTTACATGAACATATTGCAAATTTAGCTGGTAGCCCTGCTATTAAAAAAGGTATTTTACAGACTGTAAAAGTT
GTTGATGAATTGGTCAAAGTAATGGGGCGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATCAGAC
AACTCAAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGAAGAAGGTATCAAAGAATTAGGAAGTCAGA
TTCTTAAAGAGCATCCTGTTGAAAATACTCAATTGCAAAATGAAAAGCTCTATCTCTATTATCTCCAAAATGGAAGA
GACATGTATGTGGACCAAGAATTAGATATTAATCGTTTAAGTGATTATGATGTCGAT**gcg**ATTGTTCCACAAAGTTT
CCTTAAAGACGATTCAATAGACAATAAGGTCTTAACGCGTTCTGATAAAAATCGTGGTAAATCGGATAACGTTCCAA
GTGAAGAAGTAGTCAAAAAGATGAAAAACTATTGGAGACAACTTCTAAACGCCAAGTTAATCACTCAACGTAAGTTT
GATAATTTAACGAAAGCTGAACGTGGAGGTTTGAGTGAACTTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGA
AACTCGCCAAATCACTAAGCATGTGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGATGAAAATGATAAAC
TTATTCGAGAGGTTAAAGTGATTACCTTAAAATCTAAATTAGTTTCTGACTTCCGAAAAGATTTCCAATTCTATAAA
GTACGTGAGATTAACAATTACCATCATGCCCATGATGCGTATCTAAATGCCGTCGTTGGAACTGCTTTGATTAAGAA
ATATCCAAAACTTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTCGTAAAATGATTGCTAAGTCTG

AGCAAGAAATAGGCAAAGCAACCGCAAAATATTTCTTTTACTCTAATATCATGAACTTCTTCAAAACAGAAATTACA
CTTGCAAATGGAGAGATTCGCAAACGCCCTCTAATCGAAACTAATGGGGAAACTGGAGAAATTGTCTGGGATAAAGG
GCGAGATTTTGCCACAGTGCGCAAAGTATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAGAAGTACAGACAG
GCGGATTCTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTATTGCTCGTAAAAAAGACTGGGATCCA
AAAAAATATGGTGGTTTTGATAGTCCAACGGTAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAATC
GAAGAAGTTAAAATCCGTTAAAGAGTTACTAGGGATCACAATTATGGAAAGAAGTTCCTTTGAAAAAAATCCGATTG
ACTTTTTAGAAGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAATCATTAAACTACCTAAATATAGTCTTTTTGAG
TTAGAAAACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAATTACAAAAAGGAAATGAGCTGGCTCTGCCAAGCAA
ATATGTGAATTTTTTATATTTAGCTAGTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAAT
TGTTTGTGGAGCAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCGTGTTATTTTA
GCAGATGCCAATTTAGATAAAGTTCTTAGTGCATATAACAAACATAGAGACAAACCAATACGTGAACAAGCAGAAAA
TATTATTCATTTATTTACGTTGACGAATCTTGGAGCTCCCGCTGCTTTTAAATATTTTGATACAACAATTGATCGTA
AACGATATACGTCTACAAAAGAAGTTTTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTTATGAAACACGC
ATTGATTTGAGTCAGCTAGGAGGTGAC**TAA**CTCGAGTAAGGATCTCCAGGCATCAAATAAAACGAAAGGCTCAGTCG
AAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGG
TGGGCCTTTCTGCGTTTATA

>dxCas9(3.7)

ATGGACAAGAAGTACTCCATTGGGCTC**gct**ATCGGCACAAACAGCGTCGGCTGGGCCGTCATTACGGACGAGTACAA
GGTGCCGAGCAAAAAATTCAAAGTTCTGGGCAATACCGATCGCCACAGCATAAAGAAGAACCTCATTGGCGCCCTCC
TGTTCGACTCCGGGGAGACGGCCGAAGCCACGCGGCTCAAAAGAACAGCACGGCGCAGATATACCCGCAGAAAGAAT
CGGATCTGCTACCTGCAGGAGATCTTTAGTAATGAGATGGCTAAGGTGGATGACTCTTTCTTCCATAGGCTGGAGGA
GTCCTTTTTGGTGGAGGAGGATAAAAAGCACGAGCGCCACCCAATCTTTGGCAATATCGTGGACGAGGTGGCGTACC
ATGAAAAGTACCCAACCATATATCATCTGAGGAAGAAGCTTGTAGACAGTACTGATAAGGCTGACTTGCGGTTGATC
TATCTCGCGCTGGCGCATATGATCAAATTTCGGGGACACTTCCTCATCGAGGGGGACCTGAACCCAGACAACAGCGA
TGTCGACAAACTCTTTATCCAACTGGTTCAGACTTACAATCAGCTTTTCGAAGAGAACCCGATCAACGCATCCGGAG
TTGACGCCAAAGCAATCCTGAGCGCTAGGCTGTCCAAATCCCGGCGGCTCGAAAACCTCATCGCACAGCTCCCTGGG
GAGAAGAAGAACGGCCTGTTTGGTAATCTTATCGCCCTGTCCCTCGGGCTGACCCCCAACTTTAAATCTAACTTCGA
CCTGGCCGAAGAT**acc**AAGCTTCAACTGAGCAAAGACACCTACGATGATGATCTCGACAATCTGCTGGCCCAGATCG
GCGACCAGTACGCAGACCTTTTTTTGGCGGCAAAGAACCTGTCAGACGCCATTCTGCTGAGTGATATTCTGCGAGTG
AACACGGAGATCACCAAAGCTCCGCTGAGCGCTAGTATGATCAAG**ctc**TATGATGAGCACCACCAAGACTTGACTTT
GCTGAAGGCCCTTGTCAGACAGCAACTGCCTGAGAAGTACAAGGAAATTTTCTTCGATCAGTCTAAAAATGGCTACG
CCGGATACATTGACGGCGGAGCAAGCCAGGAGGAATTTTACAAATTTATTAAGCCCATCTTGGAAAAAATGGACGGC
ACCGAGGAGCTGCTGGTAAAGCTTAACAGAGAAGATCTGTTGCGCAAACAGCGCACTTTCGACAATGGA**atc**ATCCC
CCACCAGATTCACCTGGGCGAACTGCACGCTATCCTCAGGCGGCAAGAGGATTTCTACCCCTTTTTGAAAGATAACA
GGGAAAAGATTGAGAAAATCCTCACATTTCGGATACCCTACTATGTAGGCCCCCTCGCCCGGGGAAATTCCAGATTC
GCGTGGATGACTCGCAAATCAGAAGAGACCATCACTCCCTGGAACTTCGAG**aaa**GTCGTGGATAAGGGGGCCTCTGC
CCAGTCCTTCATCGAAAGGATGACTAACTTTGATAAAAATCTGCCTAACGAAAAGGTGCTTCCTAAACACTCTCTGC
TGTACGAGTACTTCACAGTTTATAACGAGCTCACCAAGGTCAAATACGTCACAGAAGGGATGAGAAAGCCAGCATTC
CTGTCTGGA**gat**CAGAAGAAAGCTATTGTGGACCTCCTCTTCAAGACGAACCGGAAAGTTACCGTGAAACAGCTCAA
AGAAGACTATTTCAAAAAGATTGAATGTTTCGACTCTGTTGAAATCAGCGGAGTGGAGGATCGCTTCAACGCATCCC
TGGGAACGTATCACGATCTCCTGAAAATCATTAAAGACAAGGACTTCCTGGACAATGAGGAGAACGAGGACATTCTT
GAGGACATTGTCCTCACCCTTACGTTGTTTGAAGATAGGGAGATGATTGAAGAACGCTTGAAAACTTACGCTCATCT
CTTCGACGACAAAGTCATGAAGCAGCTCAAGAGGCGCCGATATACAGGATGGGGGCGGCTGTCAAGAAAACTGATCA
ATGGGATCCGAGACAAGCAGAGTGGAAAGACAATCCTGGATTTTCTTAAGTCCGATGGATTTGCCAACCGGAACTTC
attCAGTTGATCCATGATGACTCTCTCACCTTTAAGGAGGACATCCAGAAAGCACAAGTTTCTGGCCAGGGGGACAG
TCTTCACGAGCACATCGCTAATCTTGCAGGTAGCCCAGCTATCAAAAAGGGAATACTGCAGACCGTTAAGGTCGTGG
ATGAACTCGTCAAAGTAATGGGAAGGCATAAGCCCGAGAATATCGTTATCGAGATGGCCCGAGAGAACCAAACCACC
CAGAAGGGACAGAAGAACAGTAGGGAAAGGATGAAGAGGATTGAAGAGGGTATAAAAGAACTGGGGTCCCAAATCCT

TAAGGAACACCCAGTTGAAAACACCCAGCTTCAGAATGAGAAGCTCTACCTGTACTACCTGCAGAACGGCAGGGACA
TGTACGTGGATCAGGAACTGGACATCAATCGGCTCTCCGACTACGACGTGGAC**gct**ATCGTGCCCCAGTCTTTTCTC
AAAGATGATTCTATTGATAATAAAGTGTTGACAAGATCCGATAAAAACAGAGGGAAGAGTGATAACGTCCCCTCAGA
AGAAGTTGTCAAGAAAATGAAAAATTATTGGCGGCAGCTGCTGAACGCCAAACTGATCACACAACGGAAGTTCGATA
ATCTGACTAAGGCTGAACGAGGTGGCCTGTCTGAGTTGGATAAAGCCGGTTTCATCAAAAGGCAGCTTGTTGAGACA
CGCCAGATCACCAAGCACGTGGCCCAAATTCTCGATTCACGCATGAACACCAAGTACGATGAAAATGACAAACTGAT
TCGAGAGGTGAAAGTTATTACTCTGAAGTCTAAGCTGGTCTCAGATTTCAGAAAGGACTTTCAGTTTTATAAGGTGA
GAGAGATCAACAATTACCACCATGCGCATGATGCCTACCTGAATGCAGTGGTAGGCACTGCACTTATCAAAAAATAT
CCCAAGCTTGAATCTGAATTTGTTTACGGAGACTATAAAGTGTACGATGTTAGGAAAATGATCGCAAAGTCTGAGCA
GGAAATAGGCAAGGCCACCGCTAAGTACTTCTTTTACAGCAATATTATGAATTTTTTCAAGACCGAGATTACACTGG
CCAATGGAGAGATTCGGAAGCGACCACTTATCGAAACAAACGGAGAAACAGGAGAAATCGTGTGGGACAAGGGTAGG
GATTTCGCGACAGTCCGGAAGGTCCTGTCCATGCCGCAGGTGAACATCGTTAAAAAGACCGAAGTACAGACCGGAGG
TTTCTCCAAGGAAAGTATCCTCCCGAAAAGGAACAGCGACAAGCTGATCGCACGCAAAAAAGATTGGGACCCCAAGA
AATACGGCGGATTCGATTCTCCTACAGTCGCTTACAGTGTACTGGTTGTGGCCAAAGTGGAGAAAGGGAAGTCTAAA
AAACTCAAAAGCGTCAAGGAACTGCTGGGCATCACAATCATGGAGCGATCAAGCTTCGAAAAAAACCCCATCGACTT
TCTCGAGGCGAAAGGATATAAAGAGGTCAAAAAAGACCTCATCATTAAGCTTCCCAAGTACTCTCTCTTTGAGCTTG
AAAACGGCCGGAAACGAATGCTCGCTAGTGCGGGC**gtg**CTGCAGAAAGGTAACGAGCTGGCACTGCCCTCTAAATAC
GTTAATTTCTTGTATCTGGCCAGCCACTATGAAAAGCTCAAAGGGTCTCCCGAAGATAATGAGCAGAAGCAGCTGTT
CGTGGAACAACACAAACACTACCTTGATGAGATCATCGAGCAAATAAGCGAATTCTCCAAAAGAGTGATCCTCGCCG
ACGCTAACCTCGATAAGGTGCTTTCTGCTTACAATAAGCACAGGGATAAGCCCATCAGGGAGCAGGCAGAAAACATT
ATCCACTTGTTTACTCTGACCAACTTGGGCGCGCCTGCAGCCTTCAAGTACTTCGACACTACCATAGACAGAAAGCG
GTACACCTCTACAAAGGAGGTCCTGGACGCCACACTGATTCATCAGTCAATTACGGGGCTCTATGAAACAAGAATCG
ACCTCTCTCAGCTCGGTGGAGAC**TAA**

>dCas9-NG

ATGGATAAGAAATACTCAATAGGCTTA**gct**ATCGGCACAAATAGCGTCGGATGGGCGGTGATCACTGATGAATATAA
GGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATACAGACCGCCACAGTATCAAAAAAAATCTTATAGGGGCTCTTT
TATTTGACAGTGGAGAGACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAAGAAT
CGTATTTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTCTTTCATCGACTTGAAGA
GTCTTTTTTGGTGGAAGAAGACAAGAAGCATGAACGTCATCCTATTTTTGGAAATATAGTAGATGAAGTTGCTTATC
ATGAGAAATATCCAACTATCTATCATCTGCGAAAAAAATTGGTAGATTCTACTGATAAAGCGGATTTGCGCTTAATC
TATTTGGCCTTAGCGCATATGATTAAGTTTCGTGGTCATTTTTTGATTGAGGGAGATTTAAATCCTGATAATAGTGA
TGTGGACAAACTATTTATCCAGTTGGTACAAACCTACAATCAATTATTTGAAGAAAACCCTATTAACGCAAGTGGAG
TAGATGCTAAAGCGATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATCTCATTGCTCAGCTCCCCGGT
GAGAAGAAAAATGGCTTATTTGGGAATCTCATTGCTTTGTCATTGGGTTTGACCCCTAATTTTAAATCAAATTTTGA
TTTGGCAGAAGATGCTAAATTACAGCTTTCAAAAGATACTTACGATGATGATTTAGATAATTTATTGGCGCAAATTG
GAGATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATTTATCAGATGCTATTTTACTTTCAGATATCCTAAGAGTA
AATACTGAAATAACTAAGGCTCCCCTATCAGCTTCAATGATTAAACGCTACGATGAACATCATCAAGACTTGACTCT
TTTAAAAGCTTTAGTTCGACAACAACTTCCAGAAAAGTATAAAGAAATCTTTTTTGATCAATCAAAAAACGGATATG
CAGGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTATAAATTTATCAAACCAATTTTAGAAAAAATGGATGGT
ACTGAGGAATTATTGGTGAAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGGACCTTTGACAACGGCTCTATTCC
CCATCAAATTCACTTGGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAAGACAATC
GTGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCCATTGGCGCGTGGCAATAGTCGTTTT
GCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGC
TCAATCATTTATTGAACGCATGACAAACTTTGATAAAAATCTTCCAAATGAAAAAGTACTACCAAAACATAGTTTGC
TTTATGAGTATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTGAAGGAATGCGAAAACCAGCATTT
CTTTCAGGTGAACAGAAGAAAGCCATTGTTGATTTACTCTTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAA
AGAAGATTATTTCAAAAAAATAGAATGTTTTGATAGTGTTGAAATTTCAGGAGTTGAAGATAGATTTAATGCTTCAT
TAGGTACCTACCATGATTTGCTAAAAATTATTAAAGATAAAGATTTTTTGGATAATGAAGAAAATGAAGATATCTTA

GAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGAGATGATTGAGGAAAGACTTAAAACATATGCTCACCT
CTTTGATGATAAGGTGATGAAACAGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTA
ATGGTATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTGAAATCAGATGGTTTTGCCAATCGCAATTTT
ATGCAGCTGATCCATGATGATAGTTTGACATTTAAAGAAGACATTCAAAAAGCACAAGTGTCTGGACAAGGCGATAG
TTTACATGAACATATTGCAAATTTAGCTGGTAGCCCTGCTATTAAAAAAGGTATTTTACAGACTGTAAAAGTTGTTG
ATGAATTGGTCAAAGTAATGGGGCGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAGAACCAAACCACC
CAGAAGGGACAGAAGAACAGTAGGGAAAGGATGAAGAGGATTGAAGAGGGTATAAAAGAACTGGGGTCCCAAATCCT
TAAGGAACACCCAGTTGAAAACACCCAGCTTCAGAATGAGAAGCTCTACCTGTACTACCTGCAGAACGGCAGGGACA
TGTACGTGGATCAGGAACTGGACATCAATCGGCTCTCCGACTACGACGTGGAC**gct**ATCGTGCCCCAGTCTTTTCTC
AAAGATGATTCTATTGATAATAAAGTGTTGACAAGATCCGATAAAAACAGAGGGAAGAGTGATAACGTCCCCTCAGA
AGAAGTTGTCAAGAAAATGAAAAATTATTGGCGGCAGCTGCTGAACGCCAAACTGATCACACAACGGAAGTTCGATA
ATCTGACTAAGGCTGAACGAGGTGGCCTGTCTGAGTTGGATAAAGCCGGTTTCATCAAAAGGCAGCTTGTTGAGACA
CGCCAGATCACCAAGCACGTGGCCCAAATTCTCGATTCACGCATGAACACCAAGTACGATGAAAATGACAAACTGAT
TCGAGAGGTGAAAGTTATTACTCTGAAGTCTAAGCTGGTCTCAGATTTCAGAAAGGACTTTCAGTTTTATAAGGTGA
GAGAGATCAACAATTACCACCATGCGCATGATGCCTACCTGAATGCAGTGGTAGGCACTGCACTTATCAAAAAATAT
CCCAAGCTTGAATCTGAATTTGTTTACGGAGACTATAAAGTGTACGATGTTAGGAAAATGATCGCAAAGTCTGAGCA
GGAAATAGGCAAGGCCACCGCTAAGTACTTCTTTTACAGCAATATTATGAATTTTTTCAAGACCGAGATTACACTGG
CCAATGGAGAGATTCGGAAGCGACCACTTATCGAAACAAACGGAGAAACAGGAGAAATCGTGTGGGACAAGGGTAGG
GATTTCGCGACAGTCCGGAAGGTCCTGTCCATGCCGCAGGTGAACATCGTTAAAAAGACCGAAGTACAGACCGGAGG
TTTCTCCAAGGAAAGTATC**cgt**CCGAAAAGGAACAGCGACAAGCTGATCGCACGCAAAAAAGATTGGGACCCCAAGA
AATACGGCGGATTC**gtg**TCTCCTACAGTCGCTTACAGTGTACTGGTTGTGGCCAAAGTGGAGAAAGGGAAGTCTAAA
AAACTCAAAAGCGTCAAGGAACTGCTGGGCATCACAATCATGGAGCGATCAAGCTTCGAAAAAAACCCCATCGACTT
TCTGGAGGCGAAAGGATATAAAGAGGTCAAAAAAGACCTCATCATTAAGCTTCCCAAGTACTCTCTCTTTGAGCTTG
AAAACGGCCGGAAACGAATGCTCGCTAGTGCG**cgtttt**CTGCAGAAAGGTAACGAGCTGGCACTGCCCTCTAAATAC
GTTAATTTCTTGTATCTGGCCAGCCACTATGAAAAGCTCAAAGGGTCTCCCGAAGATAATGAGCAGAAGCAGCTGTT
CGTGGAACAACACAAACACTACCTTGATGAGATCATCGAGCAAATAAGCGAATTCTCCAAAAGAGTGATCCTCGCCG
ACGCTAACCTCGATAAGGTGCTTTCTGCTTACAATAAGCACAGGGATAAGCCCATCAGGGAGCAGGCAGAAAACATT
ATCCACTTGTTTACTCTGACCAACTTGGGCGCGCCT**cgt**GCCTTCAAGTACTTCGACACTACCATAGACAGAAAG**gt**
gTAC**cgc**TCTACAAAGGAGGTCCTGGACGCCACACTGATTCATCAGTCAATTACGGGGCTCTATGAAACAAGAATCG
ACCTCTCTCAGCTCGGTGGAGAC**TAA**

>dxCas9-NG

ATGGACAAGAAGTACTCCATTGGGCTC**gct**ATCGGCACAAACAGCGTCGGCTGGGCCGTCATTACGGACGAGTACAA
GGTGCCGAGCAAAAAATTCAAAGTTCTGGGCAATACCGATCGCCACAGCATAAAGAAGAACCTCATTGGCGCCCTCC
TGTTCGACTCCGGGGAGACGGCCGAAGCCACGCGGCTCAAAAGAACAGCACGGCGCAGATATACCCGCAGAAAGAAT
CGGATCTGCTACCTGCAGGAGATCTTTAGTAATGAGATGGCTAAGGTGGATGACTCTTTCTTCCATAGGCTGGAGGA
GTCCTTTTTGGTGGAGGAGGATAAAAAGCACGAGCGCCACCCAATCTTTGGCAATATCGTGGACGAGGTGGCGTACC
ATGAAAAGTACCCAACCATATATCATCTGAGGAAGAAGCTTGTAGACAGTACTGATAAGGCTGACTTGCGGTTGATC
TATCTCGCGCTGGCGCATATGATCAAATTTCGGGGACACTTCCTCATCGAGGGGGACCTGAACCCAGACAACAGCGA
TGTCGACAAACTCTTTATCCAACTGGTTCAGACTTACAATCAGCTTTTCGAAGAGAACCCGATCAACGCATCCGGAG
TTGACGCCAAAGCAATCCTGAGCGCTAGGCTGTCCAAATCCCGGCGGCTCGAAAACCTCATCGCACAGCTCCCTGGG
GAGAAGAAGAACGGCCTGTTTGGTAATCTTATCGCCCTGTCCCTCGGGCTGACCCCCAACTTTAAATCTAACTTCGA
CCTGGCCGAAGAT**acc**AAGCTTCAACTGAGCAAAGACACCTACGATGATGATCTCGACAATCTGCTGGCCCAGATCG
GCGACCAGTACGCAGACCTTTTTTTGGCGGCAAAGAACCTGTCAGACGCCATTCTGCTGAGTGATATTCTGCGAGTG
AACACGGAGATCACCAAAGCTCCGCTGAGCGCTAGTATGATCAAG**ctc**TATGATGAGCACCACCAAGACTTGACTTT
GCTGAAGGCCCTTGTCAGACAGCAACTGCCTGAGAAGTACAAGGAAATTTTCTTCGATCAGTCTAAAAATGGCTACG
CCGGATACATTGACGGCGGAGCAAGCCAGGAGGAATTTTACAAATTTATTAAGCCCATCTTGGAAAAAATGGACGGC
ACCGAGGAGCTGCTGGTAAAGCTTAACAGAGAAGATCTGTTGCGCAAACAGCGCACTTTCGACAATGGA**atc**ATCCC
CCACCAGATTCACCTGGGCGAACTGCACGCTATCCTCAGGCGGCAAGAGGATTTCTACCCCTTTTTGAAAGATAACA

GGGAAAAGATTGAGAAAATCCTCACATTTCGGATACCCTACTATGTAGGCCCCCTCGCCCGGGGAAATTCCAGATTC
GCGTGGATGACTCGCAAATCAGAAGAGACCATCACTCCCTGGAACTTCGAG**aaa**GTCGTGGATAAGGGGGCCTCTGC
CCAGTCCTTCATCGAAAGGATGACTAACTTTGATAAAAATCTGCCTAACGAAAAGGTGCTTCCTAAACACTCTCTGC
TGTACGAGTACTTCACAGTTTATAACGAGCTCACCAAGGTCAAATACGTCACAGAAGGGATGAGAAAGCCAGCATTC
CTGTCTGGA**gat**CAGAAGAAAGCTATTGTGGACCTCCTCTTCAAGACGAACCGGAAAGTTACCGTGAAACAGCTCAA
AGAAGACTATTTCAAAAAGATTGAATGTTTCGACTCTGTTGAAATCAGCGGAGTGGAGGATCGCTTCAACGCATCCC
TGGGAACGTATCACGATCTCCTGAAAATCATTAAAGACAAGGACTTCCTGGACAATGAGGAGAACGAGGACATTCTT
GAGGACATTGTCCTCACCCTTACGTTGTTTGAAGATAGGGAGATGATTGAAGAACGCTTGAAAACTTACGCTCATCT
CTTCGACGACAAAGTCATGAAGCAGCTCAAGAGGCGCCGATATACAGGATGGGGGCGGCTGTCAAGAAAACTGATCA
ATGGGATCCGAGACAAGCAGAGTGGAAAGACAATCCTGGATTTTCTTAAGTCCGATGGATTTGCCAACCGGAACTTC
attCAGTTGATCCATGATGACTCTCTCACCTTTAAGGAGGACATCCAGAAAGCACAAGTTTCTGGCCAGGGGGACAG
TCTTCACGAGCACATCGCTAATCTTGCAGGTAGCCCAGCTATCAAAAAGGGAATACTGCAGACCGTTAAGGTCGTGG
ATGAACTCGTCAAAGTAATGGGAAGGCATAAGCCCGAGAATATCGTTATCGAGATGGCCCGAGAGAACCAAACCACC
CAGAAGGGACAGAAGAACAGTAGGGAAAGGATGAAGAGGATTGAAGAGGGTATAAAAGAACTGGGGTCCCAAATCCT
TAAGGAACACCCAGTTGAAAACACCCAGCTTCAGAATGAGAAGCTCTACCTGTACTACCTGCAGAACGGCAGGGACA
TGTACGTGGATCAGGAACTGGACATCAATCGGCTCTCCGACTACGACGTGGAC**gct**ATCGTGCCCCAGTCTTTTCTC
AAAGATGATTCTATTGATAATAAAGTGTTGACAAGATCCGATAAAAACAGAGGGAAGAGTGATAACGTCCCCTCAGA
AGAAGTTGTCAAGAAAATGAAAAATTATTGGCGGCAGCTGCTGAACGCCAAACTGATCACACAACGGAAGTTCGATA
ATCTGACTAAGGCTGAACGAGGTGGCCTGTCTGAGTTGGATAAAGCCGGTTTCATCAAAAGGCAGCTTGTTGAGACA
CGCCAGATCACCAAGCACGTGGCCCAAATTCTCGATTCACGCATGAACACCAAGTACGATGAAAATGACAAACTGAT
TCGAGAGGTGAAAGTTATTACTCTGAAGTCTAAGCTGGTCTCAGATTTCAGAAAGGACTTTCAGTTTTATAAGGTGA
GAGAGATCAACAATTACCACCATGCGCATGATGCCTACCTGAATGCAGTGGTAGGCACTGCACTTATCAAAAAATAT
CCCAAGCTTGAATCTGAATTTGTTTACGGAGACTATAAAGTGTACGATGTTAGGAAAATGATCGCAAAGTCTGAGCA
GGAAATAGGCAAGGCCACCGCTAAGTACTTCTTTTACAGCAATATTATGAATTTTTTCAAGACCGAGATTACACTGG
CCAATGGAGAGATTCGGAAGCGACCACTTATCGAAACAAACGGAGAAACAGGAGAAATCGTGTGGGACAAGGGTAGG
GATTTCGCGACAGTCCGGAAGGTCCTGTCCATGCCGCAGGTGAACATCGTTAAAAAGACCGAAGTACAGACCGGAGG
TTTCTCCAAGGAAAGTATC**cgt**CCGAAAAGGAACAGCGACAAGCTGATCGCACGCAAAAAAGATTGGGACCCCAAGA
AATACGGCGGATTC**gtg**TCTCCTACAGTCGCTTACAGTGTACTGGTTGTGGCCAAAGTGGAGAAAGGGAAGTCTAAA
AAACTCAAAAGCGTCAAGGAACTGCTGGGCATCACAATCATGGAGCGATCAAGCTTCGAAAAAAACCCCATCGACTT
TCTGGAGGCGAAAGGATATAAAGAGGTCAAAAAAGACCTCATCATTAAGCTTCCCAAGTACTCTCTCTTTGAGCTTG
AAAACGGCCGGAAACGAATGCTCGCTAGTGCG**cgtttt**CTGCAGAAAGGTAACGAGCTGGCACTGCCCTCTAAATAC
GTTAATTTCTTGTATCTGGCCAGCCACTATGAAAAGCTCAAAGGGTCTCCCGAAGATAATGAGCAGAAGCAGCTGTT
CGTGGAACAACACAAACACTACCTTGATGAGATCATCGAGCAAATAAGCGAATTCTCCAAAAGAGTGATCCTCGCCG
ACGCTAACCTCGATAAGGTGCTTTCTGCTTACAATAAGCACAGGGATAAGCCCATCAGGGAGCAGGCAGAAAACATT
ATCCACTTGTTTACTCTGACCAACTTGGGCGCGCCT**cgt**GCCTTCAAGTACTTCGACACTACCATAGACAGAAAG**gt**
gTAC**cgc**TCTACAAAGGAGGTCCTGGACGCCACACTGATTCATCAGTCAATTACGGGGCTCTATGAAACAAGAATCG
ACCTCTCTCAGCTCGGTGGAGAC**TAA**

>dSpG

ATGGATAAGAAATACTCAATAGGCTTA**gct**ATCGGCACAAATAGCGTCGGATGGGCGGTGATCACTGATGAATATAA
GGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATACAGACCGCCACAGTATCAAAAAAAATCTTATAGGGGCTCTTT
TATTTGACAGTGGAGAGACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAAGAAT
CGTATTTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTCTTTCATCGACTTGAAGA
GTCTTTTTTGGTGGAAGAAGACAAGAAGCATGAACGTCATCCTATTTTTGGAAATATAGTAGATGAAGTTGCTTATC
ATGAGAAATATCCAACTATCTATCATCTGCGAAAAAAATTGGTAGATTCTACTGATAAAGCGGATTTGCGCTTAATC
TATTTGGCCTTAGCGCATATGATTAAGTTTCGTGGTCATTTTTTGATTGAGGGAGATTTAAATCCTGATAATAGTGA
TGTGGACAAACTATTTATCCAGTTGGTACAAACCTACAATCAATTATTTGAAGAAAACCCTATTAACGCAAGTGGAG
TAGATGCTAAAGCGATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATCTCATTGCTCAGCTCCCCGGT
GAGAAGAAAAATGGCTTATTTGGGAATCTCATTGCTTTGTCATTGGGTTTGACCCCTAATTTTAAATCAAATTTTGA

TTTGGCAGAAGATGCTAAATTACAGCTTTCAAAAGATACTTACGATGATGATTTAGATAATTTATTGGCGCAAATTG
GAGATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATTTATCAGATGCTATTTTACTTTCAGATATCCTAAGAGTA
AATACTGAAATAACTAAGGCTCCCCTATCAGCTTCAATGATTAAACGCTACGATGAACATCATCAAGACTTGACTCT
TTTAAAAGCTTTAGTTCGACAACAACTTCCAGAAAAGTATAAAGAAATCTTTTTTGATCAATCAAAAAACGGATATG
CAGGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTATAAATTTATCAAACCAATTTTAGAAAAAATGGATGGT
ACTGAGGAATTATTGGTGAAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGGACCTTTGACAACGGCTCTATTCC
CCATCAAATTCACTTGGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAAGACAATC
GTGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCCATTGGCGCGTGGCAATAGTCGTTTT
GCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGC
TCAATCATTTATTGAACGCATGACAAACTTTGATAAAAATCTTCCAAATGAAAAAGTACTACCAAAACATAGTTTGC
TTTATGAGTATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTGAAGGAATGCGAAAACCAGCATTT
CTTTCAGGTGAACAGAAGAAAGCCATTGTTGATTTACTCTTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAA
AGAAGATTATTTCAAAAAAATAGAATGTTTTGATAGTGTTGAAATTTCAGGAGTTGAAGATAGATTTAATGCTTCAT
TAGGTACCTACCATGATTTGCTAAAAATTATTAAAGATAAAGATTTTTTGGATAATGAAGAAAATGAAGATATCTTA
GAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGAGATGATTGAGGAAAGACTTAAAACATATGCTCACCT
CTTTGATGATAAGGTGATGAAACAGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTA
ATGGTATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTGAAATCAGATGGTTTTGCCAATCGCAATTTT
ATGCAGCTGATCCATGATGATAGTTTGACATTTAAAGAAGACATTCAAAAAGCACAAGTGTCTGGACAAGGCGATAG
TTTACATGAACATATTGCAAATTTAGCTGGTAGCCCTGCTATTAAAAAAGGTATTTTACAGACTGTAAAAGTTGTTG
ATGAATTGGTCAAAGTAATGGGGCGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATCAGACAACT
CAAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGAAGAAGGTATCAAAGAATTAGGAAGTCAGATTCT
TAAAGAGCATCCTGTTGAAAATACTCAATTGCAAAATGAAAAGCTCTATCTCTATTATCTCCAAAATGGAAGAGACA
TGTATGTGGACCAAGAATTAGATATTAATCGTTTAAGTGATTATGATGTCGAT**gcg**ATTGTTCCACAAAGTTTCCTT
AAAGACGATTCAATAGACAATAAGGTCTTAACGCGTTCTGATAAAAATCGTGGTAAATCGGATAACGTTCCAAGTGA
AGAAGTAGTCAAAAAGATGAAAAACTATTGGAGACAACTTCTAAACGCCAAGTTAATCACTCAACGTAAGTTTGATA
ATTTAACGAAAGCTGAACGTGGAGGTTTGAGTGAACTTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACT
CGCCAAATCACTAAGCATGTGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGATGAAAATGATAAACTTAT
TCGAGAGGTTAAAGTGATTACCTTAAAATCTAAATTAGTTTCTGACTTCCGAAAAGATTTCCAATTCTATAAAGTAC
GTGAGATTAACAATTACCATCATGCCCATGATGCGTATCTAAATGCCGTCGTTGGAACTGCTTTGATTAAGAAATAT
CCAAAACTTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTCGTAAAATGATTGCTAAGTCTGAGCA
AGAAATAGGCAAAGCAACCGCAAAATATTTCTTTTACTCTAATATCATGAACTTCTTCAAAACAGAAATTACACTTG
CAAATGGAGAGATTCGCAAACGCCCTCTAATCGAAACTAATGGGGAAACTGGAGAAATTGTCTGGGATAAAGGGCGA
GATTTTGCCACAGTGCGCAAAGTATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAGAAGTACAGACAGGCGG
ATTCTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTATTGCTCGTAAAAAAGACTGGGATCCAAAAA
AATATGGTGGTTTT**ctgtgg**CCAACGGTAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAATCGAAG
AAGTTAAAATCCGTTAAAGAGTTACTAGGGATCACAATTATGGAAAGAAGTTCCTTTGAAAAAAATCCGATTGACTT
TTTAGAAGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAATCATTAAACTACCTAAATATAGTCTTTTTGAGTTAG
AAAACGGTCGTAAACGGATGCTGGCTAGTGCC**aaacag**TTACAAAAAGGAAATGAGCTGGCTCTGCCAAGCAAATAT
GTGAATTTTTTATATTTAGCTAGTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTT
TGTGGAGCAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCGTGTTATTTTAGCAG
ATGCCAATTTAGATAAAGTTCTTAGTGCATATAACAAACATAGAGACAAACCAATACGTGAACAAGCAGAAAATATT
ATTCATTTATTTACGTTGACGAATCTTGGAGCTCCCGCTGCTTTTAAATATTTTGATACAACAATTGATCGTAAA**ca**
gTAT**cgt**TCTACAAAAGAAGTTTTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTTATGAAACACGCATTG
ATTTGAGTCAGCTAGGAGGTGAC**TAA**

>dSpRY

ATGGATAAGAAATACTCAATAGGCTTA**gct**ATCGGCACAAATAGCGTCGGATGGGCGGTGATCACTGATGAATATAA
GGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATACAGACCGCCACAGTATCAAAAAAAATCTTATAGGGGCTCTTT
TATTTGACAGTGGAGAGACAGCGGAA**cgt**ACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAAGAAT

CGTATTTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATGATAGTTTCTTTCATCGACTTGAAGA
GTCTTTTTTGGTGGAAGAAGACAAGAAGCATGAACGTCATCCTATTTTTGGAAATATAGTAGATGAAGTTGCTTATC
ATGAGAAATATCCAACTATCTATCATCTGCGAAAAAAATTGGTAGATTCTACTGATAAAGCGGATTTGCGCTTAATC
TATTTGGCCTTAGCGCATATGATTAAGTTTCGTGGTCATTTTTTGATTGAGGGAGATTTAAATCCTGATAATAGTGA
TGTGGACAAACTATTTATCCAGTTGGTACAAACCTACAATCAATTATTTGAAGAAAACCCTATTAACGCAAGTGGAG
TAGATGCTAAAGCGATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATCTCATTGCTCAGCTCCCCGGT
GAGAAGAAAAATGGCTTATTTGGGAATCTCATTGCTTTGTCATTGGGTTTGACCCCTAATTTTAAATCAAATTTTGA
TTTGGCAGAAGATGCTAAATTACAGCTTTCAAAAGATACTTACGATGATGATTTAGATAATTTATTGGCGCAAATTG
GAGATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATTTATCAGATGCTATTTTACTTTCAGATATCCTAAGAGTA
AATACTGAAATAACTAAGGCTCCCCTATCAGCTTCAATGATTAAACGCTACGATGAACATCATCAAGACTTGACTCT
TTTAAAAGCTTTAGTTCGACAACAACTTCCAGAAAAGTATAAAGAAATCTTTTTTGATCAATCAAAAAACGGATATG
CAGGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTATAAATTTATCAAACCAATTTTAGAAAAAATGGATGGT
ACTGAGGAATTATTGGTGAAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGGACCTTTGACAACGGCTCTATTCC
CCATCAAATTCACTTGGGTGAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAAGACAATC
GTGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCCATTGGCGCGTGGCAATAGTCGTTTT
GCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGC
TCAATCATTTATTGAACGCATGACAAACTTTGATAAAAATCTTCCAAATGAAAAAGTACTACCAAAACATAGTTTGC
TTTATGAGTATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTGAAGGAATGCGAAAACCAGCATTT
CTTTCAGGTGAACAGAAGAAAGCCATTGTTGATTTACTCTTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAA
AGAAGATTATTTCAAAAAAATAGAATGTTTTGATAGTGTTGAAATTTCAGGAGTTGAAGATAGATTTAATGCTTCAT
TAGGTACCTACCATGATTTGCTAAAAATTATTAAAGATAAAGATTTTTTGGATAATGAAGAAAATGAAGATATCTTA
GAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGAGATGATTGAGGAAAGACTTAAAACATATGCTCACCT
CTTTGATGATAAGGTGATGAAACAGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTA
ATGGTATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTGAAATCAGATGGTTTTGCCAATCGCAATTTT
ATGCAGCTGATCCATGATGATAGTTTGACATTTAAAGAAGACATTCAAAAAGCACAAGTGTCTGGACAAGGCGATAG
TTTACATGAACATATTGCAAATTTAGCTGGTAGCCCTGCTATTAAAAAAGGTATTTTACAGACTGTAAAAGTTGTTG
ATGAATTGGTCAAAGTAATGGGGCGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATCAGACAACT
CAAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGAAGAAGGTATCAAAGAATTAGGAAGTCAGATTCT
TAAAGAGCATCCTGTTGAAAATACTCAATTGCAAAATGAAAAGCTCTATCTCTATTATCTCCAAAATGGAAGAGACA
TGTATGTGGACCAAGAATTAGATATTAATCGTTTAAGTGATTATGATGTCGAT**gcg**ATTGTTCCACAAAGTTTCCTT
AAAGACGATTCAATAGACAATAAGGTCTTAACGCGTTCTGATAAAAATCGTGGTAAATCGGATAACGTTCCAAGTGA
AGAAGTAGTCAAAAAGATGAAAAACTATTGGAGACAACTTCTAAACGCCAAGTTAATCACTCAACGTAAGTTTGATA
ATTTAACGAAAGCTGAACGTGGAGGTTTGAGTGAACTTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACT
CGCCAAATCACTAAGCATGTGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGATGAAAATGATAAACTTAT
TCGAGAGGTTAAAGTGATTACCTTAAAATCTAAATTAGTTTCTGACTTCCGAAAAGATTTCCAATTCTATAAAGTAC
GTGAGATTAACAATTACCATCATGCCCATGATGCGTATCTAAATGCCGTCGTTGGAACTGCTTTGATTAAGAAATAT
CCAAAACTTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTCGTAAAATGATTGCTAAGTCTGAGCA
AGAAATAGGCAAAGCAACCGCAAAATATTTCTTTTACTCTAATATCATGAACTTCTTCAAAACAGAAATTACACTTG
CAAATGGAGAGATTCGCAAACGCCCTCTAATCGAAACTAATGGGGAAACTGGAGAAATTGTCTGGGATAAAGGGCGA
GATTTTGCCACAGTGCGCAAAGTATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAGAAGTACAGACAGGCGG
ATTCTCCAAGGAGTCAATT**cgt**CCAAAAAGAAATTCGGACAAGCTTATTGCTCGTAAAAAAGACTGGGATCCAAAAA
AATATGGTGGTTTT**ctgtgg**CCAACGGTAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAATCGAAG
AAGTTAAAATCCGTTAAAGAGTTACTAGGGATCACAATTATGGAAAGAAGTTCCTTTGAAAAAAATCCGATTGACTT
TTTAGAAGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAATCATTAAACTACCTAAATATAGTCTTTTTGAGTTAG
AAAACGGTCGTAAACGGATGCTGGCTAGTGCC**aaacag**TTACAAAAAGGAAATGAGCTGGCTCTGCCAAGCAAATAT
GTGAATTTTTTATATTTAGCTAGTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTT
TGTGGAGCAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCGTGTTATTTTAGCAG
ATGCCAATTTAGATAAAGTTCTTAGTGCATATAACAAACATAGAGACAAACCAATACGTGAACAAGCAGAAAATATT
ATTCATTTATTTACGTTGACG**cgt**CTTGGAGCTCCC**cgt**GCTTTTAAATATTTTGATACAACAATTGAT<u>ccg</u>AAA**ca**

gTAT**cgt**TCTACAAAAGAAGTTTTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTTATGAAACACGCATTG
ATTTGAGTCAGCTAGGAGGTGAC**TAA**

scRNA/sgRNA expression cassettes

>BBa_J23119(SpeI)-scRNA.1xMS2.b2-TrrnB

TTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTNNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAA
GTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACATGAGGATCACCCATGTGCTTTTTTTGAAGCTTG
GGCCCGAACAAAAACTCATCTCAGAAGAGGATCTGAATAGCGCCGTCGACCATCATCATCATCATCATTGAGTTTAA
ACGGTCTCCAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGC
GGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAA
ACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAG
GCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACT

>BBa_J23119(SpeI)-sgRNA-BBa_K2680405

TTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTNNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAA
GTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTAACGCATGAGAAAGCC
CCCGGAAGATCACCTTCCGGGGGCTTTTTTATTGCGC

E. coli endogenous promoters (annotated by RegulonDB) as in pCK590 (+ strand) and pCK591 (– strand). Underlined sequences are 35bp minimal promoters. **Bolded** nucleotides are **TSS** according to RegulonDB (Santos-Zavaleta et al., 2019) or **start codon** of the GFPmut2.

>yajGp
GGATCCGAACGCCGCCCTTAATTTTTCTTCTATCCGCTCACGTATCATCATGAAATATCTCCTCCGACAACGCTGAG
ATGTCATCCATCCCTTTAAATACTAGCCGCTTTTACCGTTTCCATTGCAGCTTAACAACAAATATTTAGGTTTCGTC
CTGGTTTTTCCGTCAAGCGTATGAAGTTTACTGATTTAAAGATTGAAACACAGGCCTCGCATCAGTGTTTCTTTTAC
CAGGGGCGATTATTTCGTCAACATGATGAATTTACATGCCTTACCCACTTCCCCTCGCCGTTGGCAATGTTATGATG
GCGG**G**AATTTTTACCACTCGAGAGATCCTCTAGATTTAAGAAGGAGATATACAT**ATG**

>uxuRp
CTCGAGACTTTGTTGGCGCAGTGACGGCGGCATATCAGCAGCTGTGCGAACGCGGTGCGCGCGAGTGTGTGGCTGCG
CTGTAACTAACTGATTACCCTACAGACTTACTGGTCAATCAAACTGATATTTGGTTGACCAGTTTTCGTTTTTTTGC
CCACCTGTACGTGCCAACTTCCAGT**G**CTAATGGTATAGTTTGAGATTAACGGGGGCCGTAAAATTGCCCGTTGTAGG
CCGGATAAGGCGTTCACGCCGCATCCGGCAAAAATTTGATTAACCGCACCTAACGGACACAACACCATGAAATCTGC
CACCTCTGCGCAAAGACCTTACCAGGAAGTCGGGGCGATGATCCGCGATCTGATCATAAAGACGCCGTACAATCCTG
GGGATCCTCTAGATTTAAGAAGGAGATATACAT**ATG**

>araEp
GGATCCCTGATGGGGCCGATAGTGTTCCTTGCCTCCAGCGCTTCAGATTATGTGAATGGTTATACCATTGCCGTGGA
TGGCGGTTGGCTGGCGCGTTAATTCATTCTTCTTACTTTTATGACCCTGCCGCATGGCAGGGTTTTTTATACCTGTA
GATCATCATAATCCATATCATGGTTATGAAATAATCCATATTAATTATCAATTAATGAACTTTATGAATTTTATCTG
CTGTAAAATTAGGTGGTTAATAATAATCTCAATAATTCAACTTAATTTGAAAATTGGAATATCCATCACATAACGAC
ATGTCGCAGCAATTTAATCCATATTTATGCTGTTTCCGACCTGACACCTGCGTGAGTTGTTCACGTATTTTTTCACT
AT**G**TCTTACTCTCTGCTGGCAGGAAAAAATGGTTACTATCAATACGGAATCTGCTTTAACGCCACGTTCTTTGCGGG
ATACGCGGCGTATGAATATGTTTGTTTCGGTAGCTGCTGCGCTCGAGAGATCCTCTAGATTTAAGAAGGAGATATAC
AT**ATG**

>ppiDp2
CTCGAGCCGCAGACCGGTAAAGAGATCACCATCGCTGCTGCTAAAGTACCGAGCTTCCGTGCAGGTAAAGCACTGAA
AGACGCGGTAAACTAAGCGTTGTCCCCAGTGGGGATGTGACGAAGTTCAAGGGCGCATCTACTGATGTGCCTTTTTT
ATTTGTATTCGGTGACTTTCTGCGTCTTGTGGGC**T**GACAATTGCCCCCGTTTCTTGTCACAATAGGCCTTTGCGCGC
ATCGATACGTTGCGTGAGGTACACAGTCATCTACAGCGGAGTGTTGTTACACCATGATGGACAGCTTACGCACGGCT
GGATCCTCTAGATTTAAGAAGGAGATATACAT**ATG**

>aroKp1 (aroKp2 is not included)
GGATCCTCGGGCAATTATTTCGTCATGACGGAAAAGAAGATGAACGACGCGAGTTAGTGGTGTTTATCACGCCACGA
CTGGTTTCCAGTGAGTAAACAGCCGTAAAAGCGGTAATGTTTTTACGCTGAACGTGTTTCATCTATTTGACGCGCGC
AGGTATTTAGCATACAAGGAGTACC**G**ATTTGAGAGTTGGTGCTCTTCGCTGCCTGCGTTCCATGATGATGATTTATC
ATTCAGGCGGCATTTTGCTGTCTTTTTTACGCTAATCTTACCCGGTGATTTATCGCCAGAGCGGTGGTAGCAAGGCA
GCGCGCTTGCAGCGACCAGATATGCAGAGGGATGGGTGATTTATTCAGTTGCCAAACCCGCTCGAGAGATCCTCTAG
ATTTAAGAAGGAGATATACAT**ATG**

>aroLp
CTCGAGGGCGGACCAGATAGCCTTTCACAACGTGACCGCCAGGCCTTTGCCGCGGAGCTGGAGAAGTGGTGGCTGGA
AGTGCAACGTAGTCGTGGCTAAATGTAATTTATTATTTACACTTCATTCTTGAATATTTATTGGTATAGTAAGGGGT
GT**A**TTGAGATTTTCACTTTAAGTGGAATTTTTTCTTTACAATCGAAATTGTACTAGTTTGATGGTATGATCGCTATT

CTCATGACACCGGCTTTCGCCGCATTGCGACCTATTGGGGAAAACCCACGATGACACAACCTCTTTTTCTGATCGGG
CCTCGGGGCTGTGGTAAAACAACGGTCGGAATGGCCCTTGCCGATTCGCTTAACCGGATCCTCTAGATTTAAGAAGG
AGATATACAT**ATG**

>aroFp
GGATCCCAACAAGGGGGCGATAAACTTTTTCATCATTTCTTTCTCCTTTTTCAAAGCATAGCGGATTGTTTTCAAAG
GGAGTGTAAATTTATCTATACAGAGGTAAGGGTTGAAAGCGCGACTAAATTGCCTGTGTAAATAAAAATGTACGAAA
TATGGATTGAAAACTTTACTTTATGTGTTATCGTTACGTC**A**TCCTCGCTGAGGATCAACTATCGCAAACGAGCATAA
ACAGGATCGCCATCATGCAAAAAGACGCGCTGAATAACGTACATATTACCGACGAACAGGTTTTAATGACTCCGGAA
CAACTGAAGGCCGCTTTTCCATTCTCGAGAGATCCTCTAGATTTAAGAAGGAGATATACAT**ATG**

>aroHp (aroHp1 and aroHp2 are available in this reporter)
CTCGAGTCAGATCCCGTGGATTAACAGTACCAATTATTCGGTAGAAGAGATTGCCACCAAGATCCTCGATATCATGG
GCCTTAGTCGCCGAATGTACTAGAGAACTAGTGCATTAGCTTATTTTTTTGTTATCATGCTA**A**CCACCCGGCGAGGT
GTGACACACCTCGCACTTGAAATCAGCAGCGATTGGTTTATCGTGATGCGC**A**TCACTTCCCGGCAGTCCTGCCGTAG
AAGCAACAAATTTCTGAGACTTGTAATGAACAGAACTGACGAACTCCGTACTGCGCGTATTGAGAGCCTGGTAACGC
CCGCCGAACTCGCGCTACGGTATCCCGTAACGCCTGGGGATCCTCTAGATTTAAGAAGGAGATATACAT**ATG**

>pheLp
CTCGAGACAAAGGCGAAGCACGTCGTGCCGCAACATCGGTGAAAGACGCCAACTTCGTCGAAGAAGTTGAAGAAGAG
TAGTCCTTTATATTGAGTGTATCGCCAACGCGCCTTCGGGCGCGTTTTTTGTTGACAGCGTGAAAACAGTACGGGTA
CTGTACTAA**A**GTCACTTAAGGAAACAAACATGAAACACATACCGTTTTTCTTCGCATTCTTTTTTACCTTCCCCTGA
ATGGGAGGATCCTCTAGATTTAAGAAGGAGATATACAT**ATG**

>secBp
GGATCCGATGGCAACGCCGCCAAGCGTGAAGAGATGATCAAACGCAGCGGTCGCACCACGGTTCCCCAGATTTTTAT
TGACGCACAGCACATTGGCGGCTGTGATGACTTGTATGCATTGGATGCACGTGGTGGACTGGATCC**C**CTGCTGAAAT
AACGTGTGAACGTTGGCATTACATTGCGCAGTATTTAAGGACAACACTTAAGGGTTTTCTACACATGTCAGAACAAA
ACAACACTGAAATGACTTTCCAGATCCAACGTATTTATACCAAGGATATCTCTTTCGAAGCGCCGCTCGAGAGATCC
TCTAGATTTAAGAAGGAGATATACAT**ATG**

>serCp
CTCGAGTCATATGAAAGCGGGGGAAAAACAATTATGTCCGCGCTGTGCAAATCCAGAATGGACGAAGGCAAGTCGGG
CAAAACGGGTGACCTGACAGTAAAAACATCGGCTTTTTGCTAATAATCCGAGAGATTCTTTTGTGTGATGCAAGCCA
CATTTTTGCCCTCAACGGTTTTACTCATTGCGATGTGTGTCACTGAATGATAAAACCGATAGCCACAGGAATAATGT
ATT**A**CCTGTGGTCGCAATCGATTGACCGCGGGTTAATAGCAACGCAACGTGGTGAGGGGAAATGGCTCAAATCTTCA
ATTTTAGTTCTGGTCCGGCAATGCTACCGGCAGAGGTGCTTAAACAGGCTCAACAGGAACTGCGCGACTGGAACGGT
CTTGGGGATCCTCTAGATTTAAGAAGGAGATATACAT**ATG**

>talAp
CTCGAGGGGATCCAATATAGCCTCTGGCAACGCAGCCTCGAGGCTGTGTTGCCAGAGGCTTGGTTGGAGAAACCTGG
ATTTTCCCTGGAACTGGAAATTCATGGAAATCAAGTGCACTTTGTTTTAACTGGTCATCCATTTGGTTGTTCCTTTC
ACGTAACGTTCACAAATAAAGTGTGTGGGCAACAGCCCCTGCCCACAACGTGGCGCACATTATTACCCTGCCGGAGT
CTACAGACTTTGAGCAAGTCCAAACTCTCACCATTAATATAATGTTTTGGTAATAATCCTATAACACTGATGTTACC
TGCTTAATCCAGCAATACCATGCCTGTCTGCTATGCTTTTTTGATGCGTTTAGCGAAATTT**C**TCAGAAGTGTGAATT
AACGCACTCATCTAACACTTTACTTTTCAAGGAGTATTTCCTATGAACGAGTTAGACGGCATCAAACAGTTCACCAC
TGTCGTGGCAGACAGCGGCGATATTGGATCCTCTAGATTTAAGAAGGAGATATACAT**ATG**

Supplemental references

Fontana, J., Dong, C., Ham, J.Y., Zalatan, J.G., Carothers, J.M., **2018**. Regulated Expression of sgRNAs Tunes CRISPRi in *E. coli*. *Biotechnol. J.* **13**, e1800069. https://doi.org/10.1002/biot.201800069

Fontana, J., Dong, C., Kiattisewee, C., Chavali, V.P., Tickman, B.I., Carothers, J.M., Zalatan, J.G., **2020**. Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. *Nat. Commun.* **11**, 1618. https://doi.org/10.1038/s41467-020-15454-y

Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M., Gao, X., Rees, H.A., Lin, Z., Liu, D.R., **2018**. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57–63. https://doi.org/10.1038/nature26155

Kim, N., Kim, H.K., Lee, S., Seo, J.H., Choi, J.W., Park, J., Min, S., Yoon, S., Cho, S.-R., Kim, H.H., **2020**. Prediction of the sequence-specific cleavage activity of Cas9 variants. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-020-0537-9

Klanschnig, M., Cserjan-Puschmann, M., Striedner, G., Grabherr, R., **2022**. CRISPRactivation-SMS, a message for PAM sequence independent gene up-regulation in *Escherichia coli*. *Nucleic Acids Res.* gkac804. https://doi.org/10.1093/nar/gkac804

Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., Joung, J.K., **2016**. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495. https://doi.org/10.1038/nature16526

Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Topkar, V.V., Nguyen, N.T., Zheng, Z., Gonzales, A.P.W., Li, Z., Peterson, R.T., Yeh, J.-R.J., Aryee, M.J., Joung, J.K., **2015**. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485. https://doi.org/10.1038/nature14592

Legut, M., Daniloski, Z., Xue, X., McKenzie, D., Guo, X., Wessels, H.-H., Sanjana, N.E., **2020**. High-Throughput Screens of PAM-Flexible Cas9 Variants for Gene Knockout and Transcriptional Modulation. *Cell Rep.* **30**, 2859-2868.e5. https://doi.org/10.1016/j.celrep.2020.02.010

Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T., Abudayyeh, O.O., Gootenberg, J.S., Mori, H., Oura, S., Holmes, B., Tanaka, M., Seki, M., Hirano, H., Aburatani, H., Ishitani, R., Ikawa, M., Yachie, N., Zhang, F., Nureki, O., **2018**. Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* **361**, 1259–1262. https://doi.org/10.1126/science.aas9129

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**, 1173–83. https://doi.org/10.1016/j.cell.2013.02.022

Santos-Zavaleta, A., Salgado, H., Gama-Castro, S., Sánchez-Pérez, M., Gómez-Romero, L., Ledezma-Tejeida, D., García-Sotelo, J.S., Alquicira-Hernández, K., Muñiz-Rascado, L.J., Peña-Loredo, P., Ishida-Gutiérrez, C., Velázquez-Ramírez, D.A., Del Moral-Chávez, V., Bonavides-Martínez, C., Méndez-Cruz, C.-F., Galagan, J., Collado-Vides, J., **2019**. RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in *E. coli* K-12. *Nucleic Acids Res.* **47**, D212–D220. https://doi.org/10.1093/nar/gky1077

Tickman, B.I., Burbano, D.A., Chavali, V.P., Kiattisewee, C., Fontana, J., Khakimzhan, A., Noireaux, V., Zalatan, J.G., Carothers, J.M., **2021**. Multi-layer CRISPRa/i circuits for dynamic genetic programs in cell-free and bacterial systems. *Cell Syst.* S2405471221004191. https://doi.org/10.1016/j.cels.2021.10.008

Walton, R.T., Christie, K.A., Whittaker, M.N., Kleinstiver, B.P., **2020**. Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* eaba8853. https://doi.org/10.1126/science.aba8853

Walton, R.T., Hsu, J.Y., Joung, J.K., Kleinstiver, B.P., **2021**. Scalable characterization of the PAM requirements of CRISPR–Cas enzymes using HT-PAMDA. *Nat. Protoc.* **16**, 1511–1547. https://doi.org/10.1038/s41596-020-00465-2

Zaslaver, A., Bren, A., Ronen, M., Itzkovitz, S., Kikoin, I., Shavit, S., Liebermeister, W., Surette, M.G., Alon, U., **2006**. A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli. Nat. Methods* **3**, 623–628. https://doi.org/10.1038/nmeth895

Acknowledgement

My Ph.D. study encompasses a significant portion of my life, and I would like to use this opportunity to thank everyone who supported me through this Ph.D. journey. While I may not be able to name everyone who has contributed to this part of my life, if you are reading this, you are undoubtedly one of those people! Thank you!

First and foremost, I would like to thank my Ph.D. advisors, Jesse Zalatan and James Carothers, who provided me freedom to tackle the questions I was eager to answer and guided me through the scientific road. Without your directions, I would likely be lost in the myriad paths I wanted to explore.

My dissertation committees — Carrie Harwood, Jorge Marchand, and Georg Seelig — for all suggestions and support which allowed me to learn through different lenses on conducting scientific research.

Since I joined UW and the labs in 2018 and 2019, respectively, I have been blessed with numerous support from everyone here. Thank you, my mentors, mentees, and colleagues in the labs — including Chen Dong, Jason Fontana, David Sparkman-Yager, Chuhern Hwang, Ryan Kibler, Zibo Chen, Benjamin Tickman, Pramod Chavali, Chenggang Xi, Elizabeth Speltz, Erin Fagan, Robin Kirkpatrick, Daniel Cunningham-Bryan, Joseph Harman, Ian Faulkner, Maire Gavagan, Emily Cliff, Diego Alba, Brianne King, Widianti Sugianto, Ava Karanjia, Ryan Cardiff, Noel Jameson, Janis Shin, Jacob Brandner, Pansa (Fern) Leejaroen, Maggie Cook, Kira Olander, Tommy Primo, Brian Darst, Sam Dyer, Jessica Caruso, Nishi Mehta, Morgan Bean, Sonya Clarkson, Michael Guzman, Semira Beraki, Joely Nelson, Sarah Alvi, Jonathan Zhang, Daniel Ong, Elizabeth Karas, Samantha Boczek, Aleks Grey, Juliana Beall, Marisa Tsunoda, Jolene Nguyen, Luke Zhu, Nicklas Bohmann, Samantha Koplik, Cassandra Maranas, Hinako Kawabe, Anika Gupta — for sharing intriguing discussions both in science and life. You all have made this journey an incredible expedition!

Apart from lab members, I also gained diverse insights from collaborators across institutions from UW, GeorgiaTech, Institute for Systems Biology, PNNL/EMSL, and LBNL/JGI — especially Herbert Sauro, Pamela Peralta-Yahya, Anna Kuchina, Hector Garcia-Martin, Tijana Radivojevic, Yasuo Yoshikuni, Alex Beliaev, Jeremy Zucker, Allan Scott, Nathalie Munoz, Joshua Elmore — along with technical supports during my Ph.D. training.

Outside the lab, the community at UW also encouraged me to be involved with diverse activities filled with knowledgeable and enjoyable events. Especially MolES folks — Philip Leung, Ellie James, Sanaa Mansoor, Jinrong Ma, Stephen Blaskowski, Justin Lee, Steven Hsu, Nicolas Cordozo, Eric Yang, Steven Wu, Nathan Chan, Michael Xie, Will Chen — and friends from ChemE, Chemistry, and other programs at UW.

I also want to thank funding agencies (NSF, DOE, ABF) and the academic associations (EBRC, ACS, AIChE, SIMB, ASBMB, and more), along with their members, for the invaluable knowledge and perspectives they have shared. I am privileged to attend various affiliated conferences, both virtual and in-person, and received a lot of support throughout my Ph.D. from both mentors and trainees in the field.

In addition to academic training, I would not be able to survive in a foreign country without tremendous support from friends and families in Seattle. Thanks to my Thai friends both in Seattle and surrounding Pacific Northwest area — including Apichai (Yao) Yavirach, Sira (Neung) Horradarn, Chonnipa (Amp) Thanarugchok, Pathirat (Pat) Kosakanchit, Najes (Jes) Noparat, Apichaya (Mai) Ponsang, Natthapat (Natt) Sakulborrirug, Sorawis (Bo) Nilparuk, Phanchanok (Nune) Chodsawang, Chard (Aoon) Sujiphinyo, Kulmanee (Makham) Bubphamanee, Kashane (Shane) Ruangdet, Jiraporn (Khaofang) Rattanapapat, Arunrat (Chaonai) Thepna, Supasai (A) Vongkulbhisal, Pariyakorn (Pete) Maneekul, Achalaya (Marine) Thianthong, Panyaporn (Punpun) Chamangwipha, Panutad (Flame) Kuwijitsuwan, Sirapop (Omelete) Theeranantachai, Akachai (Meng) Simasathitchai, Rutchathon (Champ) Chairattana-apirom, Worawin (Tan) Premrasmi, Purim (Purim) Junkham, Nateethorn (Natee) Nuchklang, Nutvara (Goo) Jantarathaneewat, Kittibhum (Bhum) Tasanasuwan, Tanat (Tan) Rungsrithong, Pukjira (Toey) Chaemchuea — I may have missed some names but these folks keep refueling my life with unlimited positive energy. I also want to thank my extended families — Lamoon's family and Iyarin's family — who keep Seattle indistinguishable from my hometown in Thailand.

Since I have been pursuing my Ph.D. during the global pandemic, I have virtually connected with new pals and strengthened long-standing friendships through difficult times. Special thanks to the Biology Beyond Nature team for inviting me to communicate science to a broader audience in Thailand and keeping me involved in the foundation of the Thai Synthetic Biology Consortium. I am grateful toward members of the Association for Thai Democracy (ATD) members for their solidarity in the social movement rally and for providing professional advice over the past few years. These activities expanded my professional perspectives beyond my academic mindset.

As mentioned above, my Ph.D. is accomplishable only with countless support. However, I will not make it to the Ph.D. career without the opportunities provided by my colleagues and mentors since I started my scientific journey. Thank you the VISTEC team for offering me an opportunity to help set up the school of Biomolecular Engineering and Science, Vidyasirimedhi Institute of Science and Technology (VISTEC-BSE) — especially Pimchai (Bee) Chaiyen, Thanyaporn (Tidtee) Wongnate, Nopphon (Nop) Weeranoppanant, Chayasith (Tao) Uttamapinant, Juthamas (Chomphu) Jaroensuk, Litavadee (Tong) Chuaboon, Pornkanok (Cheese) Pongpamorn, Jittima (Pupe) Phonbuppha, Pratchaya (Dew) Watthaisong, Duangrat (Bom) Trisrivirat, Pubthum (Non) Munkajohnpong, Pattarawan (Maprang) Intasian, Supacha (Tam) Buttranon, Pangrum (Pang) Punthong, Kittiya (Art) Sakdaphetsiri, Narongyot (Jeng) Kittipanukul, Jirat (Ik) Chatsirisupachai, Arnut (Jay) Virachotikul, Phattananawee (Tle) Nalaoh. Thank you my friends and coworkers at the Faculty of Science, Mahidol University, for nurturing me in my early scientific career — especially Torsak (King) Luanphaisarnnont, Chutima (Grace) Jiarpinitnun, Nopporn (Nop) Ruangsupapichat, Chanachon (Dream) Thiamsiri, Jaturavit (Jo) Nimnuan, Angkit (Pump) Kaidad, Watcharapon (First) Prasitwatcharakorn, Siriphong (Best) Somprasong, Chomphunut (Chompoo) Thammapichai, Jirawat (Palm) Trakulmututa, Sitthichoke (Chok) Kasemtaweechok, Nisakorn (Tangmo) Yodsanit, Jesarit (Nai) Dithlumdub, Atis (Pleum) Yosprakob, Watsapon (Jay) Wattanasuebsin, Chanachai (Note) Pattanathummasid, Siwat (Chat) Ruangroengkulrith, Wittawin (Mes) Worakitchanon. I also appreciate supervision from

Christopher Uyeda and Ian Powers during my internship at Purdue University, my first research experience outside the country!

Additionally, I would like to thank my lifelong friends from high school, Mahidol Wittayanusorn School, which ignited my interest in scientific research. We are still connected and they will continue to be my safe space forever — including Anan (Pek) Auesuktrakul, Kasit (Aim) Chatsirisupachai, Naruchit (New) Thanuthanakhun, Supanut (Pipe) Thanasilp, Peerapat (Parn) Boonyarittipong, Songyos (Ton) Rajborirug, Toonyawat (Tew) Angkhanawin, Chayatiya (Yok) Leecharoen, Thanapone (Arno) Lohachoonsiri, Warittha (Rung) Panasawatwong, Jatucom (Arm) Rojwatcharapibal, Yossathorn (Mon) Tawabutr, Saksit (Boss) Amornwit, Apiwat (Poom) Jaroonpol, Kanin (Mick) Poolperm, Sutthichai (Hokchai) Chalalaisathaporn, Tanaporn (Baitong) Na Narong, Thanadol (Khunkoei) Sutantiwanichkul, and a lot more who were not mentioned, otherwise this acknowledgement would too long!

Finally, I would like to thank my family for believing in me and raising me to be curious and adaptable in any environment! Thank you to my mom, Sripun (Jaew) Kiatisevi, my first teacher who taught me lessons that any school cannot provide and also for confidence in any decisions I made in life. Thank you to my dad, Kittiwat (Iang) Lertprotsombat, for continuous support without any doubt. Thanks to my big sister, Achiraya (Ink) Praisuwan, who is like my second mom and keeps both eyes on me so that I stay on-track in every stage of life. Thank you to my big brother, Sukrit (Unk) Kiatisevi, who guided me through diverse and unconventional activities. And salute to my second brother, Napaphon (Act) Lertprotsombat, who filled my childhood with joy and challenges. Even though my memory together with Act is less than a decade, and I was still named Elf at that time, all the memories remain vivid like it just happened yesterday. The oldest memory I have is the Act's battle with β-thalassemia major as a child. He is six years older than me but his will to live is never inferior to any adult I have acquainted with. He taught me everything a human being should know starting from how different people can be and to realize that being different is just natural. He passed away when I was turning ten years old (already Ice at that time) and reminded me to live my life earnestly and consciously. He also taught me about science and medicine, both directly and indirectly, with his first-hand experience. Being able to work in CRISPR research has given me such prestige since β-thalassemia is now curable with the CRISPR technology.

Again, I cannot list every name in this section but I am grateful to know and learn from everyone during and before my Ph.D. expedition. Thank you all for your support!

Appendices

Materials provided below are included for reference and as a record of contributions in addition to the main dissertation chapters. Abstracts of manuscripts co-authored by Ice (Cholpisit Kiattisewee) were listed. A short summary of each work was described below.

Appendix A contains the abstract for a manuscript in preparation under the title of "Design rules for multiple bacterial CRISPRa systems enable synergistic gene activation" co-led by Ice, Ava, and Ryan which started during the COVID-19 pandemic. In this work, we explored the design rules of different bacterial CRISPR activation tools to use as a basis to design synergistic CRISPRa systems.

Appendix B contains the abstract for a manuscript in preparation titled "CRISPR-based replication control enables plasmid copy-number programming in bacteria". This work covers my accumulated persistence to investigate the ability to regulate plasmid copy-number both statically (cis-acting) and dynamically (trans-acting) using CRISPRa/i circuit as a tool.

Appendix C is an abstract of the method for building CRISPR genetic circuits "Construction of large-scale genetic circuits using bacterial CRISPRa/i". This work is initiated by Ice and Diego Alba Burbano to carefully design the bacterial genetic circuits with new tools and design rules developed by our team.

Appendix D includes the abstract for a manuscript published in *Nature Communication* "Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements". Ice helped with the study of CRISPRa under control of different σ-factor promoters during early training rotations with Chen and Jason.

Appendix E contains the abstract for a manuscript published in *Cell Systems* with the title of "Multi-layer CRISPRa/i circuits for dynamic genetic programs in cell-free and bacterial systems". During the exploration of this work led by Ben and Diego, Ice was assisting with experimental design and constructed a few tools, together with Pramod, to aid the overall study.

Appendix F includes the abstract for a manuscript published in *Materials Today Bio* entitled "Gene expression dynamics in input-responsive engineered living materials programmed for bioproduction". Ice was helping Widi with the metabolic engineering experiments and LCMS chemical analysis.

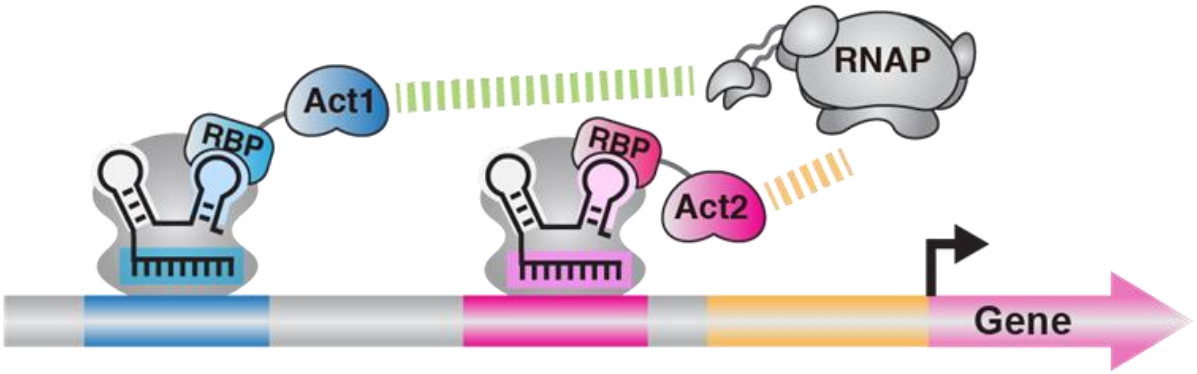
Appendix G contains the abstract for a manuscript published in *Proceedings of the National Academy of Sciences* (*PNAS)* "Engineering activatable promoters for scalable and multi-input CRISPRa/i circuits". During the pilot stages by Diego, Ryan, and Ben, Ice and Cassandra jumped in to assist on the cross-verification experiments.

Appendix H is the abstract for a manuscript in preparation with the working title "Engineering multi-gene CRISPRa programs for combinatorial metabolic expression profiling through guide RNA structure design". Ice helped with the design and characterization of multi-gene CRISPRa during the pioneer stage by Jason and David. Ian is pushing this work further for combinatorial profiling.

Appendix A:

Design rules for multiple bacterial CRISPRa systems enable synergistic gene activation

Cholpisit Kiattisewee[+], Ava V. Karanjia[+], Ryan Cardiff[+], Sarah S. Alvi, James M. Carothers*, Jesse G. Zalatan*



Abstract

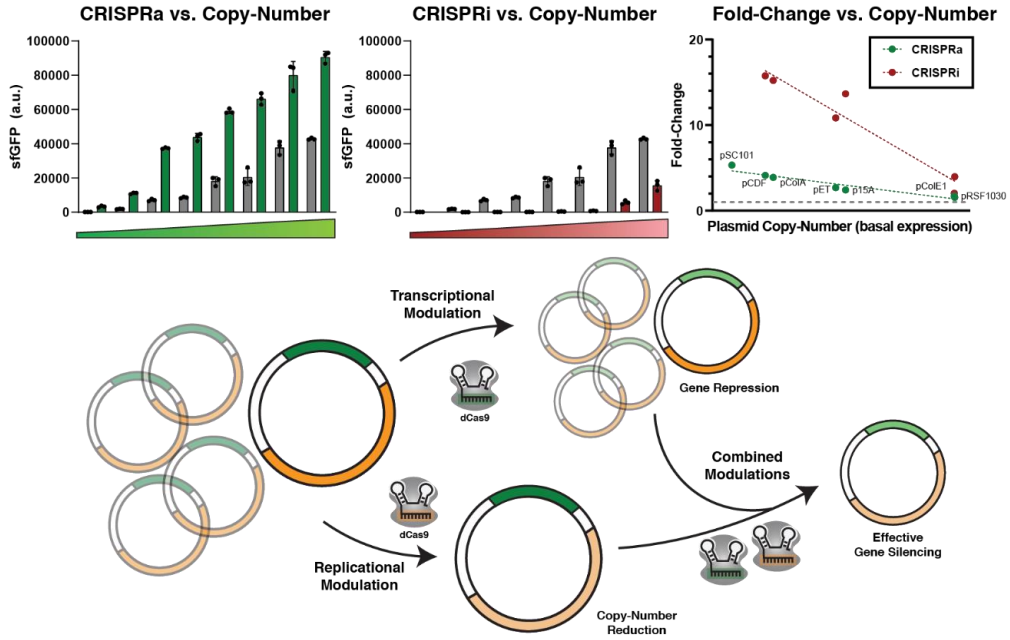
The development of effective CRISPR gene activation (CRISPRa) tools has the potential to dramatically accelerate bacterial strain engineering. Unlike in eukaryotic CRISPRa systems, effective transcriptional activation with bacterial CRISPRa requires precise placement of transcriptional activators relative to a targeted promoter. Different transcriptional activators have distinct mechanisms for recruiting RNA polymerase (RNAP), which impacts the preferred target site positions for effective CRISPRa. Here, we employed a set of engineered, synthetic promoters to investigate distance rules with different bacterial activator proteins. Across the set of CRISPRa effector proteins tested, we found that optimal activator placement relative to the transcription start site (TSS) could vary up to 200 base pairs. Using these guidelines, we tested multiple activators in combination and found that CRISPRa mediated by SoxS and PspF effectors could synergistically activate gene expression. Many other combinations of effectors were antagonistic. This work demonstrates that synergistic CRISPR transcriptional activation can be achieved in bacteria through a systematic investigation of design rules.

This work is in preparation for publication. An abstract of this work was published in a special issue of Journal of Biological Chemistry on March 28th, 2023
DOI: https://doi.org/10.1016/j.jbc.2023.103374

Appendix B:

CRISPR-based replication control enables plasmid copy-number programming in bacteria

Cholpisit Kiattisewee, Pansa Leejaroen, Benjamin I. Tickman, Ian D. Faulkner, James M. Carothers*, Jesse G. Zalatan*



Abstract

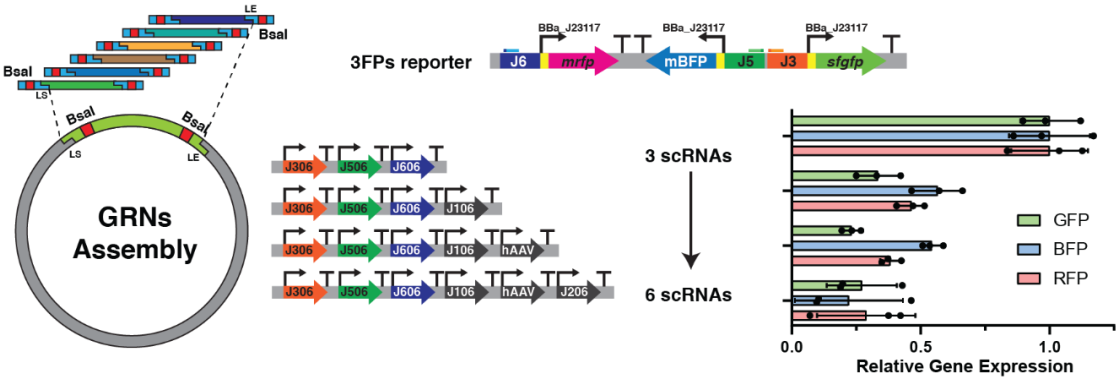
Genetic programming in non-model microorganisms usually relies on a limited set of broad-host-range plasmid replicons with low predictivity on the copy-number. In this study, we repurposed the CRISPR activation (CRISPRa) and CRISPR interference (CRISPRi) tools to modulate the plasmid replication process, resulting in programmable control of gene copy-number. We started by constructing a toolbox of CRISPRa/i-compatible reporters in different *E. coli* plasmid replicons varying in copy-number and host range to demonstrate the accessible dynamic range based on CRISPRa/i input. We observed that a higher copy-number led to a decrease in dynamic range from 100-fold to less than 2-fold. A similar effect was also observed in *Pseudomonas putida* when broad-host-range plasmids were investigated. We found that the expanded plasmid toolkit also enabled CRISPRa/i gene regulation in *Acinetobacter baylyi*. We piloted the CRISPR replication control in *E. coli* using both ColE1-family and REP-based-family where replication initiation depends on antisense RNA and REP protein, respectively. We then investigate replication control strategies in *P. putida* with the broad-host-range REP-dependent plasmids. Finally, we demonstrated that CRISPR-based transcriptional and translational controls can be enacted simultaneously. These combined mechanisms provided improvement in dynamic range compared to a single modulation, thus providing an additional layer to program CRISPR-based genetic circuitry.

This work is in preparation for publication

Appendix C:

Construction of large-scale genetic circuits using bacterial CRISPRa/i

Cholpisit Kiattisewee, Diego Alba Burbano, Semira Beraki, Jesse G. Zalatan*, James M. Carothers*



Abstract

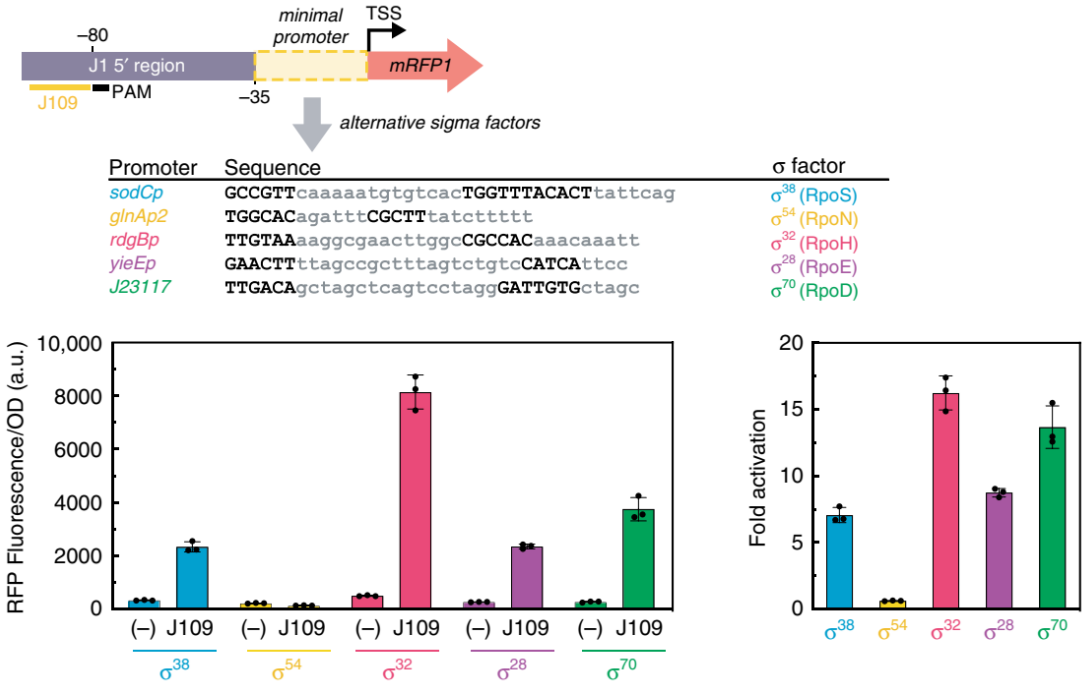
Engineered microbes offer an environmentally friendly route to chemical industries, encompassing areas of fine chemicals and therapeutic productions. To effectively modulate microbial metabolism, genetic circuits can be implemented to program complex gene regulatory networks (GRNs) within microorganisms. CRISPR-mediated gene activation and repression (CRISPRa/i) is an emerging platform suitable for large-scale genetic circuit programming. In CRISPRa/i, a complementary guide RNA (gRNA) was used to program the modulations (activation or repression) and direct CRISPR-dCas9 complex to any DNA target. However, the maximum number of genes that can be simultaneously regulated has not been experimentally investigated. Here, we experimentally investigated the limitation on the number of gRNAs in the chemical bioproduction context. We designed CRISPRa circuits architecture and high-throughput assembly method based on scalable Golden Gate Assembly which can generate at least 6 gRNA nodes at once. CRISPRa circuit performance was then evaluated by simultaneously regulating multiple fluorescent proteins as a proxy for multi-enzyme cascade in biosynthetic pathways. We found impairment of CRISPRa/i activity when the number of gRNAs increased, suggesting a competition of CRISPRa components. We then explore the rapid construction of gRNA-based genetic circuits that could be assembled in a single-step molecular cloning. This method demonstrates the constructions and implementations of multi-gRNA CRISPR circuits in bacteria, and expands the capabilities of CRISPR circuits in a wide range of biosynthetic applications.

This work is in preparation for publication

Appendix D:

Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements

Jason Fontana[+], Chen Dong[+], **Cholpisit Kiattisewee**, Venkata P. Chavali, Benjamin I. Tickman, James M. Carothers*, Jesse G. Zalatan*



Abstract

In bacterial systems, CRISPR-Cas transcriptional activation (CRISPRa) has the potential to dramatically expand our ability to regulate gene expression, but we lack predictive rules for designing effective gRNA target sites. Here, we identify multiple features of bacterial promoters that impose stringent requirements on CRISPRa target sites. Notably, we observe narrow, 2–4 base windows of effective sites with a periodicity corresponding to one helical turn of DNA, spanning ~40 bases and centered ~80 bases upstream of the TSS. However, we also identify two features suggesting the potential for broad scope: CRISPRa is effective at a broad range of σ70-family promoters, and an expanded PAM dCas9 allows the activation of promoters that cannot be activated by S. pyogenes dCas9. These results provide a roadmap for future engineering efforts to further expand and generalize the scope of bacterial CRISPRa.

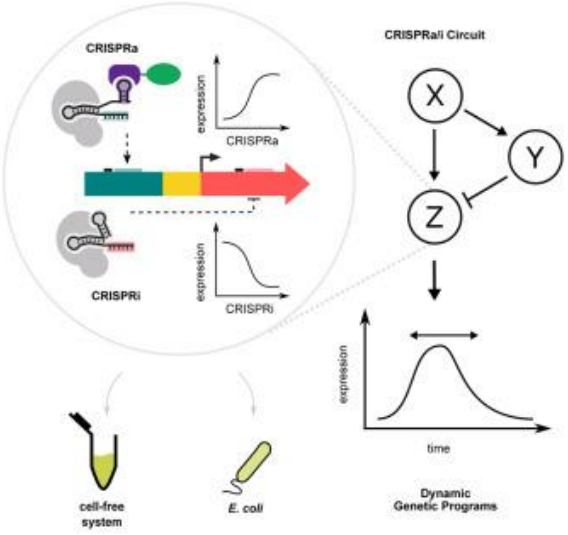
This work is published in Nature Communications on April 1st, 2020

DOI: https://doi.org/10.1038/s41467-020-15454-y

Appendix E:

Multi-layer CRISPRa/i circuits for dynamic genetic programs in cell-free and bacterial systems

Benjamin I. Tickman[+], Diego Alba Burbano[+], Venkata P. Chavali, **Cholpisit Kiattisewee**, Jason Fontana, Aset Khakimzhan, Vincent Noireaux, Jesse G. Zalatan*, James M. Carothers*



Abstract

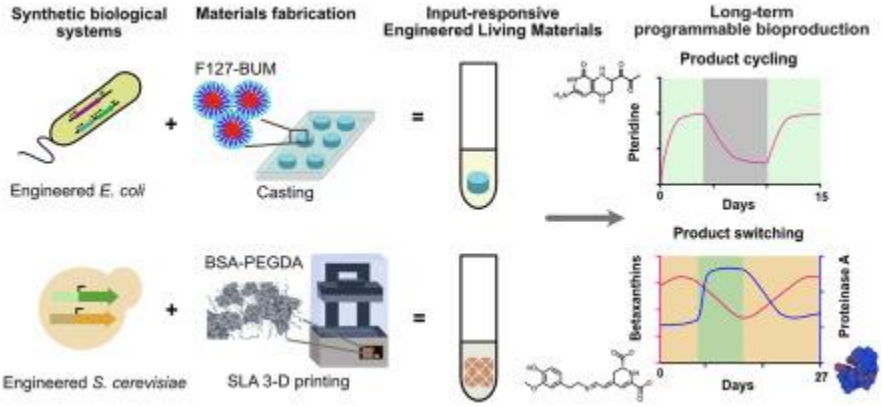
CRISPR-Cas transcriptional circuits hold great promise as platforms for engineering metabolic networks and information processing circuits. Historically, prokaryotic CRISPR control systems have been limited to CRISPRi. Creating approaches to integrate CRISPRa for transcriptional activation with existing CRISPRi-based systems would greatly expand CRISPR circuit design space. Here, we develop design principles for engineering prokaryotic CRISPRa/i genetic circuits with network topologies specified by guide RNAs. We demonstrate that multi-layer CRISPRa/i cascades and feedforward loops can operate through the regulated expression of guide RNAs in cell-free expression systems and *E. coli*. We show that CRISPRa/i circuits can program complex functions by designing type 1 incoherent feedforward loops acting as fold-change detectors and tunable pulse-generators. By investigating how component characteristics relate to network properties such as depth, width, and speed, this work establishes a framework for building scalable CRISPRa/i circuits as regulatory programs in cell-free expression systems and bacterial hosts.

This work is published in Cell Systems on November 19th, 2021
DOI: https://doi.org/10.1016/j.cels.2021.10.008

Appendix F:

Gene expression dynamics in input-responsive engineered living materials programmed for bioproduction

Widianti Sugiant, Gokce Altin-Yavuzarslan, Benjamin I. Tickman, **Cholpisit Kiattisewee**, Shuo-Fu Yuan, Sierra M. Brooks, Jitkanya Wong, Hal S. Alper, Alshakim Nelson*, James M. Carothers*



Abstract

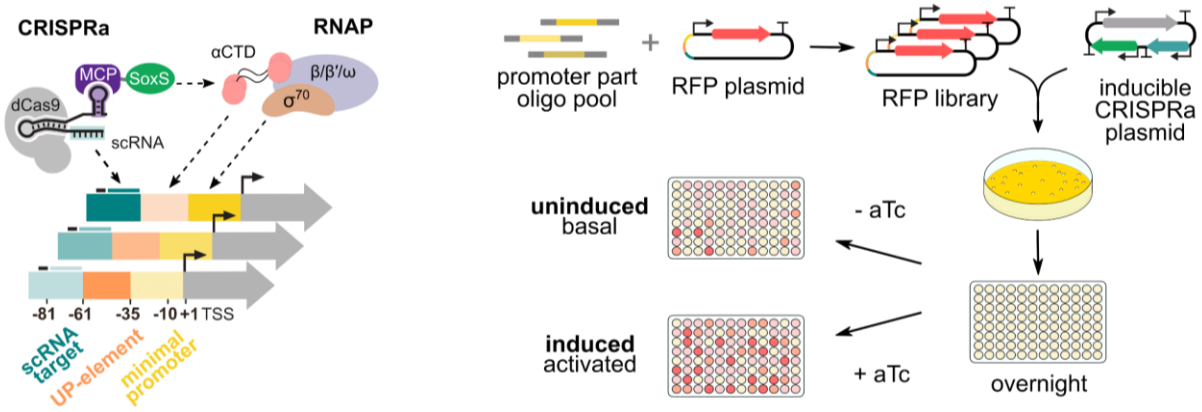
Engineered living materials (ELMs) fabricated by encapsulating microbes in hydrogels have great potential as bioreactors for sustained bioproduction. While long-term metabolic activity has been demonstrated in these systems, the capacity and dynamics of gene expression over time is not well understood. Thus, we investigate the long-term gene expression dynamics in microbial ELMs constructed using different microbes and hydrogel matrices. Through direct gene expression measurements of engineered *E. coli* in F127-bisurethane methacrylate (F127-BUM) hydrogels, we show that inducible, input-responsive genetic programs in ELMs can be activated multiple times and maintained for multiple weeks. Interestingly, the encapsulated bacteria sustain inducible gene expression almost 10 times longer than free-floating, planktonic cells. These ELMs exhibit dynamic responsiveness to repeated induction cycles, with up to 97% of the initial gene expression capacity retained following a subsequent induction event. We demonstrate multi-week bioproduction cycling by implementing inducible CRISPR transcriptional activation (CRISPRa) programs that regulate the expression of enzymes in a pteridine biosynthesis pathway. ELMs fabricated from engineered *S. cerevisiae* in bovine serum albumin (BSA) - polyethylene glycol diacrylate (PEGDA) hydrogels were programmed to express two different proteins, each under the control of a different chemical inducer. We observed scheduled bioproduction switching between betaxanthin pigment molecules and proteinase A in *S. cerevisiae* ELMs over the course of 27 days under continuous cultivation. Overall, these results suggest that the capacity for long-term genetic expression may be a general property of microbial ELMs. This work establishes approaches for implementing dynamic, input-responsive genetic programs to tailor ELM functions for a wide range of advanced applications.

This work is published in Materials Today Bio on May 22nd, 2023
DOI: https://doi.org/10.1016/j.mtbio.2023.100677

Appendix G:

Engineering activatable promoters for scalable and multi-input CRISPRa/i circuits

Diego Alba Burbano[+], Ryan Cardiff[+], Benjamin I. Tickman, **Cholpisit Kiattisewee**, Cassandra Maranas, Jesse G. Zalatan*, James M. Carothers*



Abstract

Dynamic, multi-input gene regulatory networks are ubiquitous in nature. Multi-layer CRISPR-based genetic circuits hold great promise for building gene regulatory networks akin to those found in naturally-occurring biological systems. We develop an approach for creating high-performing activatable promoters that can be assembled into deep, wide, and multi-input CRISPR-activation and -interference (CRISPRa/i) gene regulatory networks. By integrating sequence-based design and in-vivo screening, we engineer activatable promoters that achieve up to 1000-fold dynamic range in an *E. coli*-based cell-free system. These new components enable CRISPRa gene regulatory networks that are six layers deep and four branches wide. We show the generalizability of the promoter engineering workflow by improving the dynamic range of the light-dependent EL222 optogenetic system from 6-fold to 34-fold. Additionally, high dynamic range promoters enable CRISPRa systems mediated by small molecules and protein-protein interactions. We apply these tools to build input-responsive CRISPRa/i gene regulatory networks, including feedback loops, logic gates, multi-layer cascades, and dynamic pulse modulators. Our work provides a generalizable approach for the design of high dynamic range activatable promoters and enables new classes of gene regulatory functions in cell-free systems.

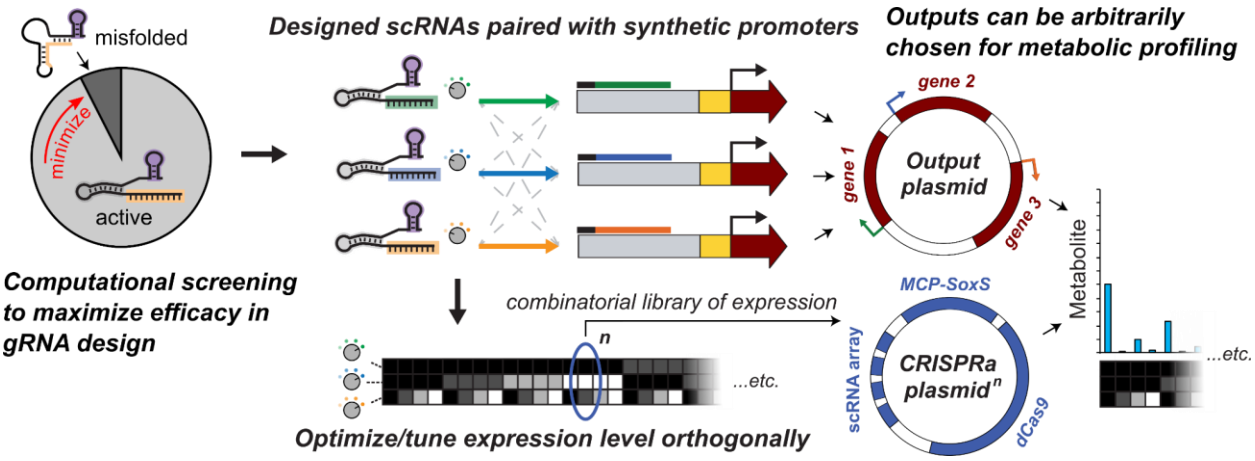
This work is published in PNAS on July 18th, 2023

DOI: https://doi.org/10.1073/pnas.2220358120

Appendix H:

Engineering multi-gene CRISPRa programs for combinatorial metabolic expression profiling through guide RNA structure design

Jason Fontana+, David Sparkman-Yager+, Ian D. Faulkner+, Ryan Cardiff, **Cholpisit Kiattisewee**, Aria Walls, Jesse G. Zalatan*, James M. Carothers*



Abstract

CRISPR gene regulation technologies provided a new landscape for microbial strain engineering to optimize chemical production. Here, we developed a combinatorial library platform of gene expression levels based on bacterial CRISPR transcriptional activation (CRISPRa). Started by developing a computational model to evaluate RNA folding energetics which can increase the chance of functional scRNA from 30% from random spacer to near 100% of computationally pre-screened sequences. Using this RNA folding tool, we constructed additional synthetic promoters and cognate guide RNAs that allowed >40-fold dynamic range total of three orthogonal sets for combinatorial library building. Then, we explored the tunability of CRISPRa-mediated gene expression by individually characterizing each truncated scRNA. We chose four expression levels programmed by scRNA and combined three sets of scRNAs to construct a combinatorial library of 64 variations. We characterized the library with three-fluorescent protein reporters and found that expression of multiple proteins led to decreased expression compared to the single reporter system, likely due to metabolic burden from expressing additional RNA and protein species. Finally, we applied the combinatorial library to the metabolic engineering platform for the biosynthesis of pteridines and human milk oligosaccharides using the architecture of fluorescent reporter. We observed that productions of pteridines and oligosaccharides were acquired from the highest expression level of each gene but through fine-tuned expressions of the whole set. These results suggested that enzyme concentrations should be considered to optimize metabolic flux in each individual reaction together with metabolic burden from overexpressing gene clusters in order to develop a versatile metabolic engineering platform for various biosynthetic hosts of interest.

This work is in preparation for publication