Title

CRAGE-CRISPR facilitates rapid activation of secondary metabolite biosynthetic gene clusters in bacteria

Journal

Cell Chemical Biology, 29(4)

ISSN

2451-9456

Authors

Ke, Jing

Robinson, David

Wu, Zong-Yen

et al.

Publication Date

2022-04-01

DOI

10.1016/j.chembiol.2021.08.009

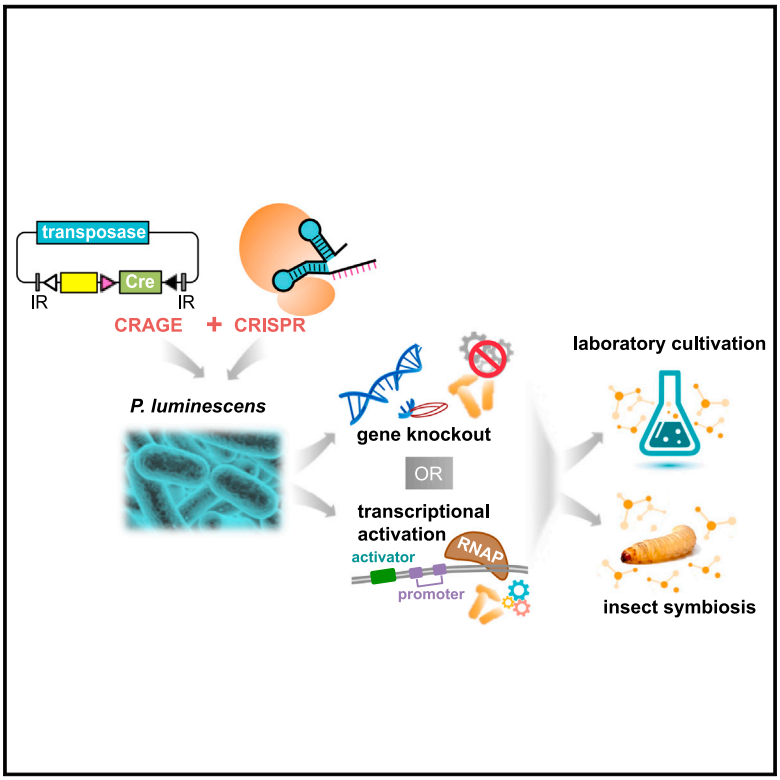
Peer reviewed

Resource

Cell Chemical Biology

CRAGE-CRISPR facilitates rapid activation of secondary metabolite biosynthetic gene clusters in bacteria

Graphical abstract



Authors

Jing Ke, David Robinson, Zong-Yen Wu, ..., Trent Northen, Jan-Fang Cheng, Yasuo Yoshikuni

Correspondence

jfcheng@lbl.gov (J.-F.C.), yyoshikuni@lbl.gov (Y.Y.)

In brief

Ke et al. develop a CRAGE-CRISPR approach, which allows loss- and gain-of-function studies to rapidly confirm activity of six secondary metabolite biosynthetic gene clusters (BGCs). These results prove that CRAGE-CRISPR is a simple but effective strategy to study BGC activities in various microorganisms.

Highlights

- An expansion of the CRISPR toolbox by implementing it on CRAGE
- Gene-to-compound approach in non-model bacteria to characterize BGC and regulator
- CRAGE-CRISPRd/CRISPRa mediate loss- and gain-of-function studies of BGCs
- Target sites for CRISPRa with varied gRNAs show variable levels of BGC activation

Ke et al., 2022, Cell Chemical Biology *29*, 696–710
April 21, 2022
https://doi.org/10.1016/j.chembiol.2021.08.009

CellPress

CellPress
OPEN ACCESS

Cell Chemical Biology

Resource

CRAGE-CRISPR facilitates rapid activation of secondary metabolite biosynthetic gene clusters in bacteria

Jing Ke,[1,6] David Robinson,[1,6] Zong-Yen Wu,[1] Andrea Kuftin,[1] Katherine Louie,[1] Suzanne Kosina,[1] Trent Northen,[1,2] Jan-Fang Cheng,[1,2,*] and Yasuo Yoshikuni[1,2,3,4,5,7,*]

[1]US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[2]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[3]Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[4]Center for Advanced Bioenergy and Bioproducts Innovation, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[5]Global Center for Food, Land, and Water Resources, Hokkaido University, Hokkaido 060-8589, Japan
[6]These authors contributed equally
[7]Lead contact
*Correspondence: jfcheng@lbl.gov (J.-F.C.), yyoshikuni@lbl.gov (Y.Y.)
https://doi.org/10.1016/j.chembiol.2021.08.009

SUMMARY

With the advent of genome sequencing and mining technologies, secondary metabolite biosynthetic gene clusters (BGCs) within bacterial genomes are becoming easier to predict. For subsequent BGC characterization, clustered regularly interspaced short palindromic repeats (CRISPR) has contributed to knocking out target genes and/or modulating their expression; however, CRISPR is limited to strains for which robust genetic tools are available. Here we present a strategy that combines CRISPR with chassis-independent recombinase-assisted genome engineering (CRAGE), which enables CRISPR systems in diverse bacteria. To demonstrate CRAGE-CRISPR, we select 10 polyketide/non-ribosomal peptide BGCs in *Photorhabdus luminescens* as models and create their deletion and activation mutants. Subsequent loss- and gain-of-function studies confirm 22 secondary metabolites associated with the BGCs, including a metabolite from a previously uncharacterized BGC. These results demonstrate that the CRAGE-CRISPR system is a simple yet powerful approach to rapidly perturb expression of defined BGCs and to profile genotype-phenotype relationships in bacteria.

INTRODUCTION

It is generally believed that secondary metabolites (a.k.a. specialized metabolites and natural products) play important roles in mediating microbe-microbe and host-microbe interactions (Huang et al., 2019; Milshteyn et al., 2018; Tobias and Bode, 2019). The first step toward understanding these interactions is to identify and characterize the secondary metabolites and the biosynthetic gene clusters (BGCs) responsible for them. However, BGCs are often highly regulated, and function activation is not trivial in lab environments (Rutledge and Challis, 2015). Modulating the expression of already-defined genes and BGCs is an effective strategy for functional profiling of genotype-phenotype relationships (Deaner and Alper, 2017a; Lian et al., 2019). Several useful technologies, such as promoter replacement and suicide vector systems for gene deletion, have been used (Alper et al., 2005; Hamilton et al., 1989). However, these technologies are generally cumbersome, as they require an extended homologous arm to target the integration location, have low efficiency, and cannot be implemented easily on a large scale. In recent years, a technological breakthrough called clustered regularly interspaced short palindromic repeats (CRISPR), as well as the CRISPR-associated (Pluskal et al., 2010) system, has provided greater precision, ease, and throughput in manipulation of gene regulation (La Russa and Qi, 2015).

In nature, the CRISPR-Cas system exists widely in bacteria and archaea as an RNA-mediated adaptive immune system to eliminate invasive DNA. Repurposed, the CRISPR-Cas system is an effective way to edit genome sequences or perturb gene expression. In this technology, single-guide RNA (sgRNA) is designed to target a specific DNA sequence. Guided by the sgRNA, a Cas9 nuclease introduces a double-strand break at the target site. When researchers are working with bacteria, many of which lack non-homologous end-joining (NHEJ) repair mechanisms, they exogenously introduce a recombinase such as RecET and a template DNA to repair the double-strand break via homologous recombination (Hiom, 2009). Depending on the template DNA, a desired DNA sequence is inserted into or deleted from the target location in the genome (CRISPRd) (Mougiakos et al., 2017). In addition, dCas9, a mutated Cas9 protein without

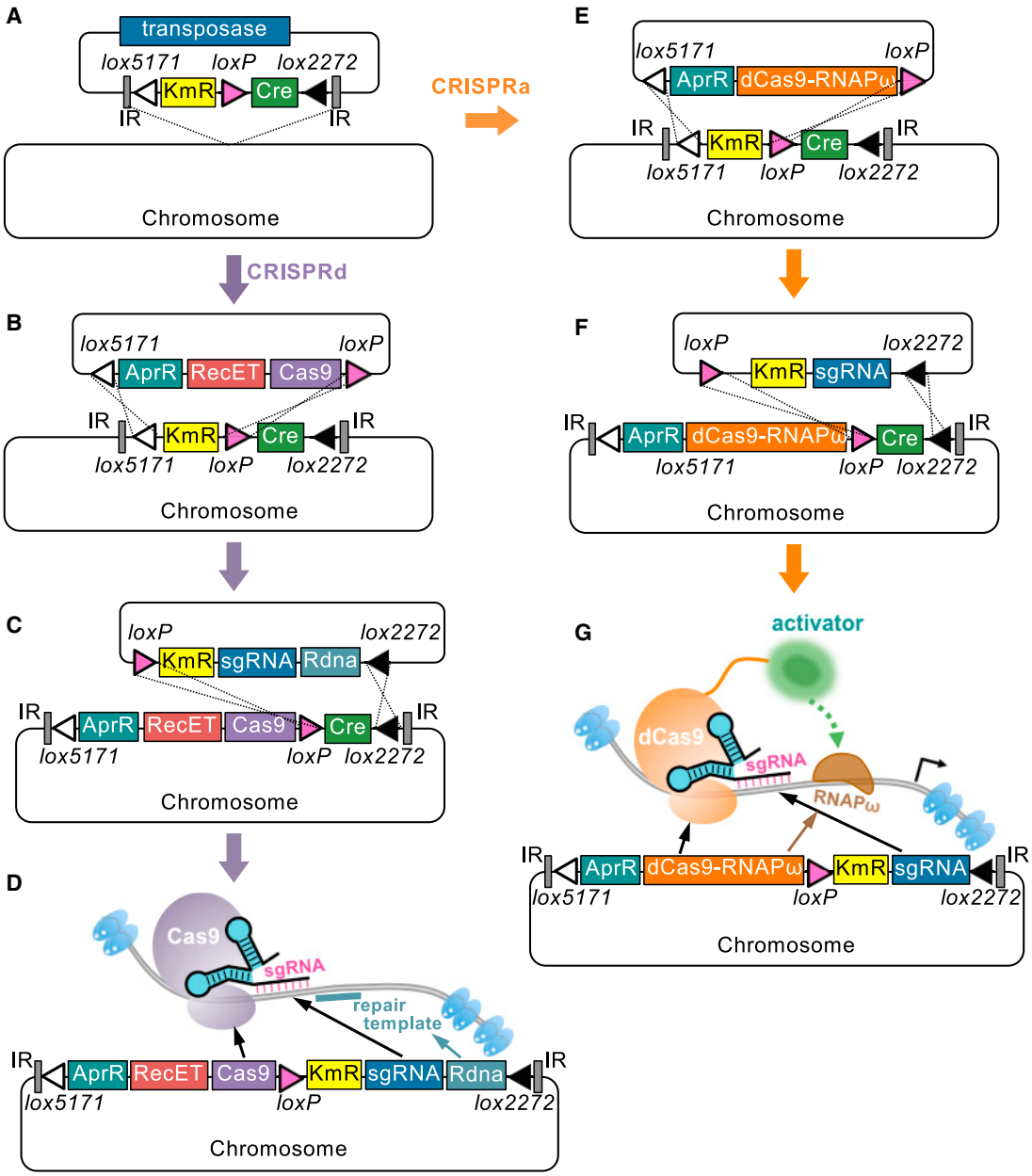


Figure 1. Scheme for the CRAGE-CRISPR system

Step 1. Landing pad (LP) integration (A). A plasmid containing a mariner transposon and transposase was generated. The transposon contained an LP comprising a Cre recombinase gene flanked by two mutually exclusive lox sites (*loxP* and *lox2272*) and a kanamycin-resistance gene, KmR, flanked by *lox5171* and *loxP*. The plasmid was conjugated from donor *E. coli* into the recipient *P. luminescens*, with the LP integrated into the recipient's genome. Step 2. Cre/lox-recombinase-mediated cassette exchange (B and E). A gene encoding Cas9 (B) or dCas9-RNAPω (E) and an apramycin-resistance gene, AprR, was flanked by *lox5171* and *loxP*, and then the Cas9 (B) or dCas9-RNAPω (E) plasmid was conjugated into the recipient cell. Step 3. Cre/lox-recombinase-mediated cassette exchange (C and F). The conjugated plasmid contained an sgRNA carrying a repair DNA (C) or only an sgRNA gene (F) that targeted the promoters or promoter-like motifs of defined BGCs, and a KmR flanked by *loxP* and *lox2272*. Step 4. Targeted gene expression modulation. Guided by sgRNA (encoded by the sgRNA gene in blue), Cas9 mediated the target gene deletion with homology-directed repair (D) or dCas9-RNAPω-recruited RNA polymerase to enhance target gene expression (G).

nuclease activity, can simply bind to the target location and function as a repressor (CRISPRi) (Larson et al., 2013). In contrast, dCas9 fused with a transcription activator can recruit RNA polymerase to activate expression of target genes (CRISPRa) (Dong et al., 2018; La Russa and Qi, 2015). Some studies also suggest that transcription repression and activation are tunable by selecting target locations. Therefore, the CRISPR-Cas system is promising as a way to characterize gene function in various organisms (Gilbert et al., 2013; Peng et al., 2018; Zalatan et al., 2015).

Use of CRISPR for function characterization of BGCs in diverse bacterial species is, however, currently hampered by a

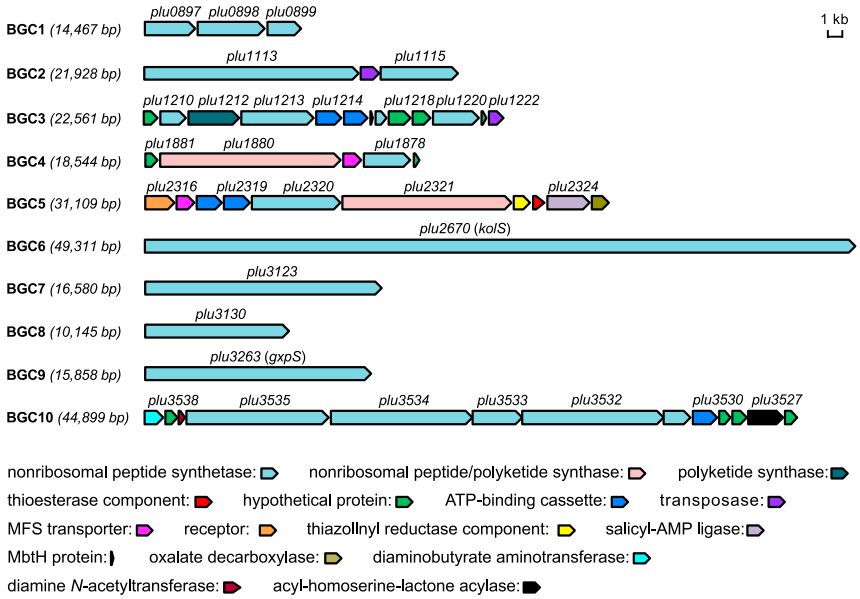


Figure 2. Architectures of model BGCs used in this study

Genes are color coded based on their known or predicted functions as indicated. Homologous genes are presented in the same color. Genes are drawn to scale.

lack of suitable genetic tools. To help remedy this lack, we recently developed chassis-independent recombinase-assisted genome engineering (CRAGE; Wang et al., 2019), which enables single-step integration of large and complex constructs into the chromosomes of diverse bacteria with high accuracy and efficiency. To implement this technology, we first integrate a landing pad (LP) comprising a *cre* recombinase gene flanked by mutually exclusive *lox* sites (Figure 1). Mediated through Cre recombinase, the LP is replaced with the constructs of interest, which are flanked by the same *lox* sites. We have recently upgraded CRAGE to CRAGE-Duet so it allows any applications that dual-plasmid systems can offer at the chromosomal level (Liu et al., 2020; Wang et al., 2020). Using CRAGE and CRAGE-Duet, we have domesticated nearly 60 species of bacteria across multiple phyla, including α-, β-, and γ-Proteobacteria and Actinobacteria. For some species, we recently demonstrated successful implementation of CRISPRd by integrating the Cas9 and RecET genes into the first location and sgRNA and an editing template into the second location (Liu et al., 2020).

In this study, we attempted to expand the utility of CRAGE-CRISPR by implementing CRISPRa on CRAGE, using it to both activate expression of secondary metabolite BGCs and facilitate their function characterization. As models, we selected 10 polyketide synthase and non-ribosomal peptide synthetase (PKS/NRPS) BGCs in the entomopathogenic bacterium *Photorhabdus luminescens* subsp. *laumondii* TT01 (Figure 2) (Fu et al., 2012; Wang et al., 2019). This strain has a unique lifestyle as a symbiont of the nematode *Heterorhabditis* and an entomopathogen with a broad insect host spectrum (Bode, 2009). *P. luminescens* produces various toxins that kill insects and defeat other microbial competitors upon infection (ffrench-Constant et al., 2003; Joyce et al., 2006). Therefore, bacteria in this genus, as well as their relatives in the *Xenorhabdus* genus, are thought to be valuable sources of bioactive secondary metabolites for discovery of new therapeutics and agricultural agents (Antonello et al., 2018; Bode, 2009, 2011; Bode et al., 2015b; Stock et al., 2017; Vizcaino et al., 2014). We previously cloned the same set of 10 BGCs and parallelly expressed them in diverse host strains using CRAGE. This multi-chassis approach enabled identification of 22 products from six BGCs. Three of these BGCs (BGCs 1, 4, and 9) were previously characterized using either heterologous expression in *E. coli* or promoter replacement (Bode et al., 2015a; Fu et al., 2012); three (BGCs 5, 7, and 8) were not previously characterized (Wang et al., 2019). These six BGCs serve as excellent benchmarks for studying the effectiveness of the CRAGE-CRISPR approach.

In brief, we built a series of constructs, using CRAGE-CRISPRd to knock out each BGC and CRAGE-CRISPRa to upregulate expression of each BGC. We tested secondary metabolite production resulting from both culturing in M9-based medium and injecting directly into *Galleria mellonella* larvae. By comparing the production yields of the series of constructs, we confirmed enhanced production of 22 metabolites from six BGCs, expression of which was modulated successfully by CRAGE-CRISPR. Our results demonstrate the combined CRAGE-CRISPR approach as a simple and effective way to study BGC function and thus to accelerate the discovery of bioactive secondary metabolites.

RESULTS

The deletion mutant for each BGC can serve as the loss-of-function or negative control, while the constructs created through CRAGE-CRISPRa with varied gRNA target sites can show various levels of BGC gain of function. CRAGE-CRISPRa was targeted to all the possible gRNA sites within 300 bp upstream of the start codon (Figures 3, 4, 5, 6, 7, S1, and S2). The number of protospacers selected for each BGC depends on the frequency of the protospacer-adjacent motif (PAM) (NGG or CCN) sequences and whether there is a repetitive sequence in those regions. For example, the 300 bp upstream of BGC2 have low (only 25%) GC content, and a repetitive sequence is found 5′ to BGC10; gRNA targets are therefore fewer for these two BGCs than for other BGCs (Figure S2).

Secondary metabolites were produced by cultivating wild-type (WT), BGC-deleted, and BGC-activated constructs in M9-based medium or in larvae. Targeted liquid chromatography-high-resolution mass spectrometry (LC-HRMS)-based metabolite analyses of BGCs 1, 4, 5, 7, and 9 revealed that CRAGE-CRISPRa targeting many (30%–100%) of the sgRNA sites (Figures 3, 4, 5, 6, and 7) activated these BGCs, resulting in the synthesis of many previously described compounds. However, the same analyses of BGCs 6 and 8 failed to detect the associated secondary metabolites. With untargeted metabolite

CellPress
OPEN ACCESS

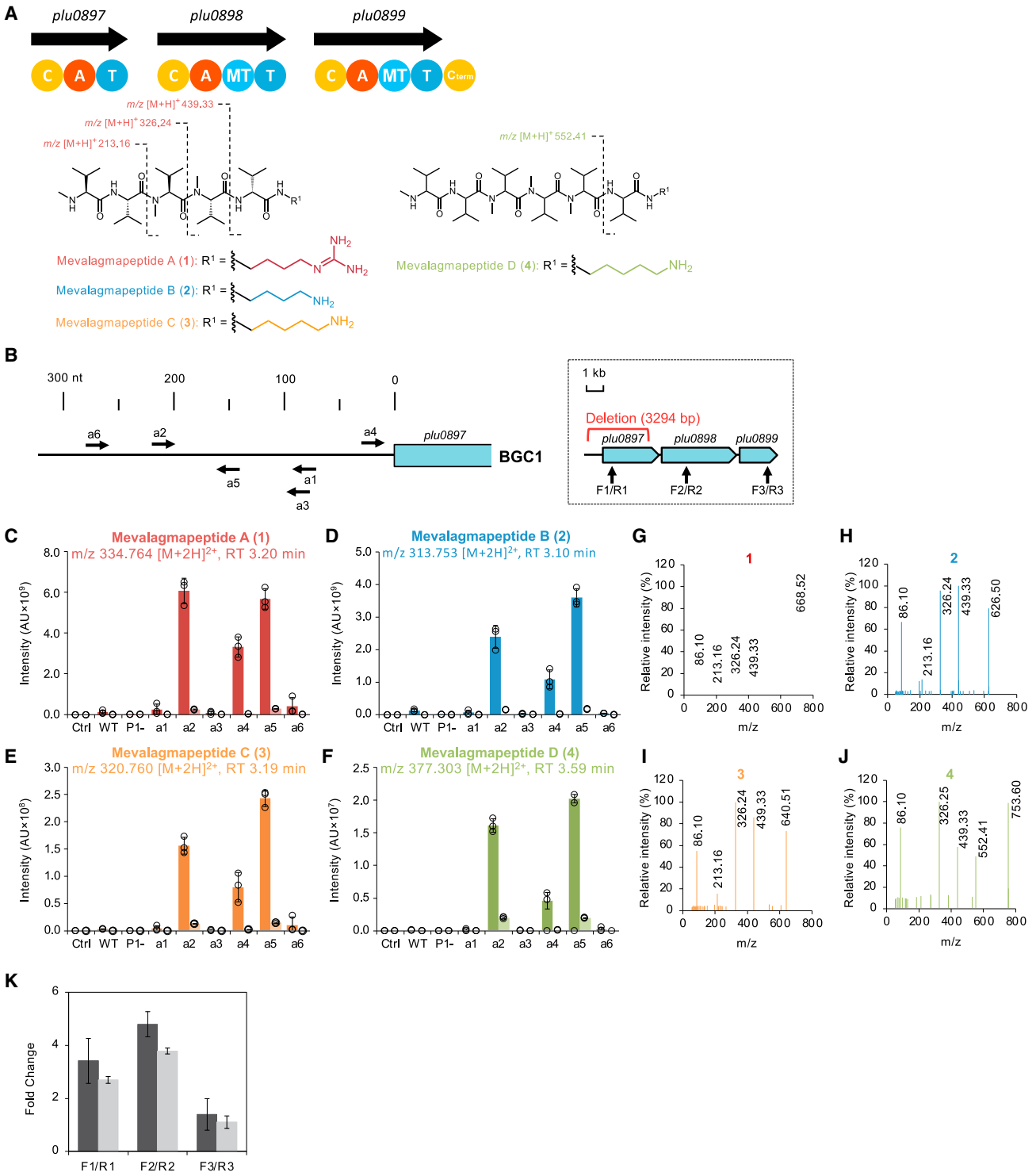


Figure 3. Production of mevalagmapeptides A–D from CRAGE-CRISPRa-modified expression of BGC1 in *P. luminescens*

(A) A proposed scheme of secondary metabolite biosynthesis catalyzed by BGC1 (*plu0897–plu0899*). NRPS domains are shown as circles; letters represent condensation (C), terminal condensation (C_{term}), adenylation (A), thiolation (T), and methylation (MT) domains.

(B) Design of the genome modulation through the CRAGE-CRISPR system for BGC1. The spacer location of the sgRNA before ATG, PAM orientation, and GC content are listed in Table S2. The three pairs of primers for qRT-PCR are labeled F1/R1, F2/R2, and F3/R3.

(C–F) Measured MS1 intensity. Shown are **1–4** (C–F, respectively) from CRAGE-CRISPRa-modulated expression of BGC1 in *P. luminescens*. In these charts, dark and light colors represent measured MS1 intensity for each metabolite extracted from the infected *G. mellonella* larvae and M9-based culture samples,

(legend continued on next page)

analysis for the previously uncharacterized BGC3, we identified a putative secondary metabolite uniquely associated with overexpression of BGC3. Because retrospective analysis of multi-chassis expression of BGC3 also identified the same secondary metabolite (Wang et al., 2019), we report this result in Figure S1.

We could not detect any secondary metabolites associated with BGC 2, 6, 8, or 10 using untargeted analyses. We here describe the results for BGCs 1, 3, 4, 5, 7, and 9, organized by type of activation. The constructs made for BGCs 2, 6, 8, and 10 are described in Figure S2.

CRAGE-CRISPRa can further activate production of secondary metabolites constitutively produced in *P. luminescens*

BGC1 contains three NRPS genes (*plu0897–plu0899*) and is 14.5 kb in size (Figure 2). Previously characterized using promoter replacement (Bode et al., 2015a), BGC1 was found to produce mevalagmapeptides A (**1**) and B (**2**) (Figure 3A) (Bode et al., 2015a). Using a multi-chassis approach, we recently demonstrated that BGC1 could additionally produce mevalagmapeptides C (**3**) and D (**4**) (Figure 3A) (Wang et al., 2019). For BGC1, we built a strain with the *plu0897* gene deleted (BGC1-*Δplu0897*) and strains that expressed one of six sgRNAs for BGC1 activation (BGC1s -a1 through -a6) (Figure 3B). Production of **1–4** (Figures 3G–3J) in the M9-based medium and in the larvae was monitored using LC-HRMS (Figures 3C–3F).

In the M9-based medium, the WT strain produced basal levels of **1** and **2**, but no **3** or **4**. Because insect larval hemocoel is known to activate some BGCs (Nollmann et al., 2015; Tobias et al., 2018; Vizcaino et al., 2014; Wang et al., 2019), the *P. luminescens* strains carrying these constructs were also injected into the larvae. BGC1 in larval hemocoel was significantly activated in the WT strain compared with BGC1 in M9-based medium; production of **1** and **2** increased by 668- and 43-fold, respectively. In addition, the WT strain produced **3** in larvae, suggesting that BGC1 is regulated by factors present in larvae. Deletion of *plu0897* resulted in no production of **1–3**, either in M9-based medium or in larvae, suggesting that this gene and its upstream sequences are required for BGC1 function.

For the CRISPRa constructs with various target sites, we observed varied levels of BGC1 activation (Figures 3C–3E). Compared with **1–3** production by the WT strain, **1–3** production was significantly improved in the three BGC1 strains -a2, -a4, and -a5, in M9-based medium (100- to 1,000-fold increase) and in larvae (10- to 100-fold increase). In our previous study, **4** had not been observed by the isopropyl-β-D-1-thiogalactopyranoside (IPTG)-inducible T7 promoter-driven expression of BGC1 in *P. luminescens* (Wang et al., 2019). Interestingly, however, it was produced by at least three strains engineered with CRAGE-CRISPRa. Various sgRNAs were designed to target different locations and to face both directions toward upstream of *plu0897*. However, we did not see any obvious correlations between sgRNAs and activation of BGC1.

Quantitative reverse-transcriptase PCR (qRT-PCR) indicated that CRISPRa increased the transcription level of BGC1-a5 in M9-based medium at the 3-h cultivation time point 3- to 5-fold (Figure 3K), while the transcription level for the 3′ end of the BGC remained almost the same as for its WT control.

BGC9 contains a single NRPS gene (*plu3263*) and is 15.9 kb in size. One of two BGCs previously characterized using heterologous expression in *E. coli*, BGC9 was found to produce gameXpeptides A (**5**), B (**7**), and E (**12**), as well as luminmides B (**6**) and D–G (**8–11**) (Figures 4A and 4K–4R) (Fu et al., 2012; Wang et al., 2019). For BGC9, we built a strain with the first 5.7 kb of *plu3263* deleted (BGC9-*Δplu3263*) and strains that expressed one of 14 sgRNAs for BGC9 activation (BGC9s -a1 through -a14) (Figure 4B). Production of **5–12** in the M9-based medium and in larvae was monitored using LC-HRMS (Figures 4C–4J).

The WT strain produced high levels of **5–12** both in the M9-based medium and in larvae. Injection of the WT strain into larvae did not significantly increase production of major products **5–9** but did increase production of minor products **10–12**, by more than 10-fold. This result may suggest that available precursors for **10–12** were more abundant in the larvae than in the M9-based medium. Deletion of the first 5.7 kb (the first two NRPS modules) of *plu3263* eliminated production of **5–12**. For four strains (BGC9s -a1, -a3, -a5, and -a11), **5–12** production was significantly improved (up to 20-fold) in both the M9-based medium and the larvae compared with production in the WT strain (Figure 4). For three strains (BGC9s -a2, -a6, and -a10), production of **5–12** remained the same. For the remaining seven strains (BGC9s -a4, -a7 through -a9, and -a12 through -a14), production of **5–12** was inhibited; these strains worked as CRISPRi rather than CRISPRa. One possible interpretation is that these intervening target sequences contain binding sites for endogenous transcriptional regulators. These sites could potentially affect CRISPRa by directly blocking access to an sgRNA target site by blocking the binding of RNA polymerase or by interfering with the ability of a CRISPRa effector protein to engage with RNA polymerase (Fontana et al., 2020). Various sgRNAs were designed to target different locations and to face both directions toward upstream of *plu3236*; however, further study is needed to explore correlations between sgRNA target sites and the activation levels for BGC function.

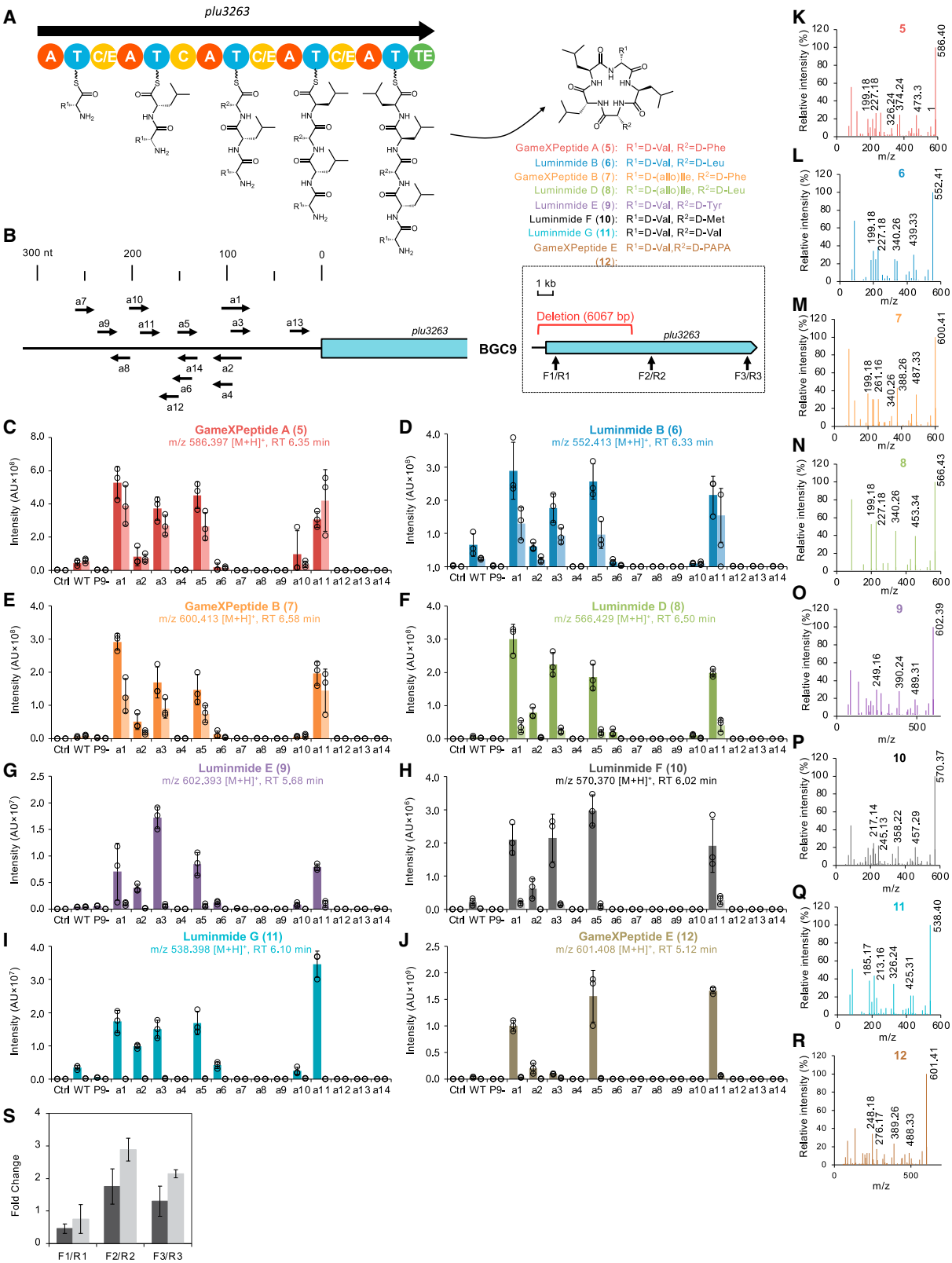
qRT-PCR indicated that the transcription level of BGC9-a1 increased 2- to 3-fold under the M9-based medium cultivation condition at the 3-h time point (Figure 4S), while the transcription level for the 5′ end of the BGC remained almost the same as for its WT control.

CRAGE-CRISPRa can activate silent BGCs in *P. luminescens*

BGC4 (18.5 kb) contains five genes (*plu1881–plu1877*), and encodes a combination of NRPS and NRPS-PKS hybrid (Figure 5A). It is silent in the native *P. luminescens*, and is one of the two

respectively. Ctrl represents larval controls that were injected with 0.9% NaCl without strain inoculation. P1- represents the BGC1 knockout strain. The strains -a1 through -a6 represent constructs with varied sgRNAs. Error bars represent SD; n = 3 technically independent experiments.
(G–J) MS2 spectra for the four metabolites **1–4** (G–J, respectively). The signature fragment ions are marked.
(K) Fold change of the mRNA level of BGC1-a5 determined by qRT-PCR. Relative expression level of BGC1 was normalized using the expression levels of the housekeeping genes GyrB (dark gray) and lpxC (light gray). Error bars represent SD; n = 3 technical replicates.

CellPress
OPEN ACCESS



(legend on next page)

BGCs that were previously characterized via heterologous expression in *E. coli* (Fu et al., 2012). That study showed production of luminmycin A (**13**) and a few intermediates of **13** synthesis. It was previously demonstrated that BGC4 could also produce glidobactin A (**14**) and cepafungin I (**15**) (Figures 5A, 5G, and 5H) (Fu et al., 2012; Stein et al., 2012; Wang et al., 2019). All three products exhibit strong cytotoxic, proteasome inhibitory, and anti-fungal activities (Oka et al., 1988; Terui et al., 1990; Theodore et al., 2012). For BGC4, we built a strain with the *plu1881* and *plu1880* genes deleted (BGC4-*Δplu1881/1880*) as a negative control and strains that expressed one of four sgRNAs for BGC4 activation (BGC4s -a1 through -a4) (Figure 5B). Production of **13–15** in the M9-based medium and in the larvae was monitored using LC-HRMS (Figures 5C–5E).

The WT strain did not produce any **13–15** in either the M9-based medium or the larvae (AU intensity below 1×10^5). In our previous study, no production of **13–15** was observed via the IPTG-inducible T7-promoter-driven expression of BGC4 in *P. luminescens* (Wang et al., 2019), suggesting that BGC4 expression is tightly regulated at multiple levels. As expected, the strain of BGC4-*Δplu1881/1880* did not produce any **13–15**.

For strains BGC4-a1 and -a3, **13–15** were produced at low levels both in the M9-based medium and in larvae (Figures 5C–5H). The sgRNAs for BGC4s -a1 and -a3 were designed to target regions more than 200 bp upstream of *plu1881*, and the PAM sequences are facing the same direction. In contrast, the sgRNAs for BGC4s -a2 and -a4, both facing the opposite of the direction faced by -a1 and -a3, were designed to target regions less than 200 bp upstream of *plu1881.* These results may suggest a potential regulatory element for BGC4 expression at around the 200-bp upstream region of *plu1881*.

qRT-PCR results demonstrated that CRISPRa increased the transcription levels of BGC4 2- to 3-fold in BGC4-a1 in M9-based medium at 3 h (Figure 5I), while the transcription levels for the middle of the BGC remained almost the same as for its WT control.

BGC5 (31.1 kb in size) contains 10 genes (*plu2316–plu2325*), and encodes a combination of NRPS and NRPS-PKS hybrid (Figure 2). The overall architecture of BGC5 is reminiscent of piscibactin BGCs found in Vibrionaceae (Ruiz et al., 2019; Thode et al., 2018). Our multi-chassis approach previously showed that BGC5 could produce 2′-(2-hydroxyphenyl)-4,5-dihydro-2,4′-bi-1,3-thiazole-4-carboxylic acid (HTTPCA) (**16**), prepiscibactin (**17**), and piscibactin-Fe complex (**18**) (Figure 6A) (Wang et al., 2019). For BGC5, we built a strain with five genes deleted (BGC5-*Δplu2316–plu2320*) and strains that expressed 1 of 10 sgRNAs for BGC5 activation (BGC5s -a1 through -a10) (Figure 6B). Production of **16–18** in the M9-based medium and larvae was monitored using LC-HRMS (Figures 6C–6H).

The WT strain did not produce **16–18** in either the M9-based medium or the larvae (Figures 6C–6E). In our previous study, **16–18** production was significantly increased via IPTG-inducible T7 promoter-driven expression of BGC5 in *P. luminescens* (Wang et al., 2019), suggesting that expression of BGC5 is regulated only at the transcriptional level. As expected, the strain with deletion of *plu2316* did not produce **16–18**.

Interestingly, all CRISPRa constructs produced certain levels of **17** and **18** (Figures 6D and 6E); however, only BGC5-a10 produced **16** (Figure 6C), a prematurely terminated intermediate of **17** and **18** (Figure 6E), probably due to metabolic overflow.

Two strains of BGC5s, -a8 and -a10, demonstrated significantly improved production levels, indicating activation of full functionality of the *plu2320*, *plu2323*, and *plu2321* cascade (Figure 6A). Their sgRNAs were designed to target regions of 247 and 360 bp upstream of *plu2316*, respectively, with the same PAM direction. This suggests a potential CRISPRa regulatory element for **16–18** production around the 247- and 360-bp upstream region of *plu2316*.

Up to 2.5-fold increased transcription level of BGC5 was confirmed by qRT-PCR for BGC5-a10 via 3-h M9-based medium cultivation (Figure 6I). The transcription levels for the 5′ end of the BGC were severely reduced and for the 3′ end of the BGC were almost the same as for its WT control. *plu2316–plu2319* were annotated as non-essential to the BGC activity, and this may explain why we saw high **16–18** production despite the reduction in the transcription level at the 5′ end of BGC5.

CRAGE-CRISPRa can conditionally activate BGCs in *P. luminescens*

BGC7 contains a single NRPS gene (*plu3123*) and is 16.6 kb in size. Our previous study characterized BGC7 function using a multi-chassis approach and found that it produces ririwpeptides A–C (**19–21**) (Figure 7A) (Tobias and Bode, 2019; Wang et al., 2019). For BGC7, we built a strain (BGC7*Δplu3123*) with the first 6,034 bp of *plu3123* deleted and strains that expressed 1 of 12 sgRNAs for BGC7 activation (BGC7s -a1 through -a12) (Figure 7B). Production of **19–21** in the M9-based medium and larvae was monitored using LC-HRMS (Figures 7C–7H).

The WT strain did not produce **19–21** in the M9-based medium. Production of these metabolites was activated when the WT strain was injected into larvae, suggesting that BGC7 is regulated by factors present in larvae. As expected, deletion of

Figure 4. Production of gameXpeptides A, B, and E, as well as luminmides B, D, E, F, and G, from CRAGE-CRISPRa-modified expression of BGC9 in *P. luminescens*

(A) A proposed scheme of secondary metabolite biosynthesis catalyzed by BGC9 (*plu3263*)-encoded enzyme. NRPS domains are shown as circles; letters represent adenylation (A), thiolation (T), condensation (C) and epimerization (C/E), and thioesterase (TE) domains.

(B) Design of the genome modulation for CRAGE-CRISPRa and knockout for BGC9. The spacer location of gRNA before ATG, PAM orientation, and GC content are listed in Table S2. The three pairs of primers for qRT-PCR are labeled F1/R1, F2/R2, and F3/R3.

(C–J) Measured MS1 intensity. Shown are **5–12** (C–J, respectively) from CRAGE-CRISPRa-modulated expression of BGC9 in *P. luminescens*. In these charts, dark and light colors represent measured MS1 intensity for each metabolite extracted from the infected larvae and M9-based culture samples, respectively. Ctrl represents larval controls that were injected with 0.9% NaCl without strain inoculation. P9- represents the BGC9 knockout strain. The strains of -a1 through -a14 represent constructs with varied sgRNAs. Error bars represent SD; n = 3 technically independent experiments.

(K–R) MS2 spectra for the eight metabolites **5–12** (K–R, respectively). The signature fragment ions are marked.

(S) Fold change of the mRNA level of BGC9-a1 determined by qRT-PC. Relative expression level of BGC9 was normalized by that of the housekeeping genes GyrB (dark gray) and lpxC (light gray). Error bars represent SD; n = 3 technical replicates.

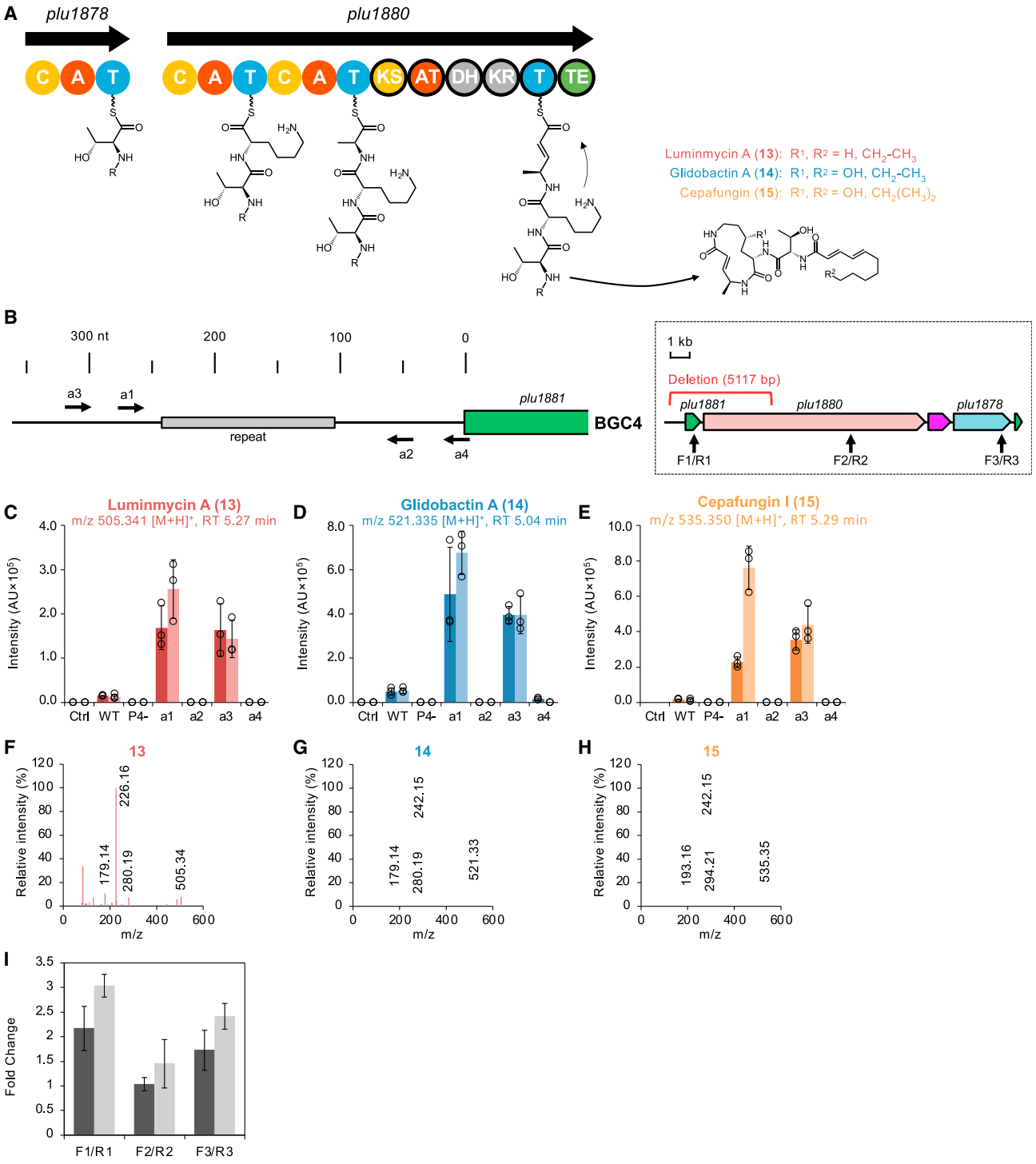


Figure 5. Production of luminmycin A, glidobactin A, and cepafungin I from CRAGE-CRISPRa-modified expression of BGC4 in *P. luminescens*

(A) A proposed scheme of secondary metabolite biosynthesis catalyzed by BGC4 (*plu1877–plu1881*). NRPS and PKS domains are shown as circles (PKS modules with black borders); letters represent condensation (C), adenylation (A), thiolation (T), ketosynthase (KS), acyltransferase (AT), dehydratase (DH), ketoreductase (KR), and thioesterase (TE) domains.

(B) Design of the genome modulation for CRAGE-CRISPRa and knockout for BGC4. The spacer location of gRNA before ATG, PAM orientation, and GC content are listed in Table S2. The three pairs of primers for qRT-PC are labeled F1/R1, F2/R2, and F3/R3.

(C–E) Measured MS1 intensity. Shown are **13–15** (C–E, respectively) from CRAGE-CRISPRa-modulated expression of BGC4 in *P. luminescens*. In these charts, dark and light colors represent measured MS1 intensity for each metabolite extracted from the infected larvae and M9-based culture samples, respectively. Ctrl

(legend continued on next page)

plu3123 completely eliminated BGC7's activity. Twelve sgRNAs were designed to target the sequences upstream of BGC7 related to its activation. For BGC7s -a1, -a3, and -a6, **19–21** production was improved (~5-fold) compared with that of the WT strain (Figures 7C–7E). For five strains (BGC7s -a2, -a5, -a9, -a11, and -a12), production of **19–21** did not change or only mildly increased. Production of **19–21** was induced only when the strains infected larvae, which indicates that BGC7 function is restricted by factors in larvae. It will be difficult to activate the expression of BGCs that are controlled by multiple regulatory factors, as other factors will first need to be explored in order to implement CRAGE-CRISPRa fully. For the remaining four strains (BGC7s -a4, -a7, -a8, and -a10), **19–21** production was inhibited. With varied sgRNAs designed upstream of *plu3236*, no obvious correlation was detected between sgRNAs and their ability to activate BGC7 function. A 4- to 8-fold increase in BGC7 transcription was confirmed for BGC7-a1 in M9-based medium (Figure 7I).

BGC3 contains 13 genes (*plu1210–plu1222*) and is 22.6 kb in size. It encodes a combination of NRPS and NRPS-PKS hybrid (Figure S1A). This BGC was toxic to the heterologous *E. coli* cloning host, which prevented its expression in *E. coli*. For BGC3, we built a strain with four genes deleted (BGC3-*Δplu1210–plu1213*) and strains that expressed one of nine sgRNAs for BGC3 activation (BGC3s -a1 through -a9) (Figure S1B).

The WT strain did not produce **22** in either the M9-based medium or the larval host (Figure S1C). As expected, the strain with deletion of *plu2316* did not produce **22** either. Production of **22** was observed only in larvae infected with the BGC3-CRISPRa strains. Of the nine CRISPRa constructs, all showed **22** production in larvae except BGC3s -a5 and -a9 (Figures S1B–S1D). BGC3s -a2 (−116 to −97 bp) and -a4 (−22 to −3 bp) produce **22** more than the other variants. qRT-PCR demonstrated that BGC3-a2 had a 16-fold enhanced transcription level of BGC3 (Figure S1E). These suggests that BGC3 expression not only is regulated at the transcriptional level, but also requires a trigger(s) and/or substrate/precursor from the insect host.

To further confirm the association of **22** production with BGC3 expression, we re-analyzed the LC-HRMS data for extracts from culturing 27 phylogenetically diverse chassis strains expressing BGC3 (Wang et al., 2019). Interestingly, low levels of **22** were detected in certain strains expressing BGC3, and production levels were correlated with the levels of IPTG induction (Figure S1F). By adding fresh larval tissue extracts, **22** can be produced and secreted in M9-based medium (Figures S1G and S1H), indicating a requirement of a special precursor(s). As fermentation continued, **22** continued accumulating in the cultures of BGC3-a2 and -a4. However, **22** decreased after 120 h, indicating its probable instability or consumption (Figure S1H). We also evaluated **22** production in M9-based medium with different additives (Figure S1G). The fact that **22** was not produced from M1, M2, or M3 indicates that the amino acids proline, aspartic acid, valine, serine, and isoleucine and amino acid dropout mixes are not precursors or they were not enough for biosynthesis of **22** (Asp, Val, Ser, and Ile are present in the predicted product structure of BGC3; Fu et al., 2012). Interestingly, the same amount (wet weight) of freeze-dried or fresh larval body additives resulted in dramatically varied levels of **22** production (Figure S1G), indicating the instability and disruption of a crucial precursor(s) from larval tissues during the freeze-drying process. However, because **22** production was limited, we have not been able to characterize its structure yet. Further optimization of its production is in process.

DISCUSSION

It is noteworthy that, although discovery of natural product BGCs has been largely facilitated through genome-mining efforts in the post-genomic era (Medema and Fischbach, 2015), only a few of them have been functionally characterized so far. To keep pace with the speed of identification of BGCs, robust strategies are needed for modulating the expression of BGCs and characterizing their function.

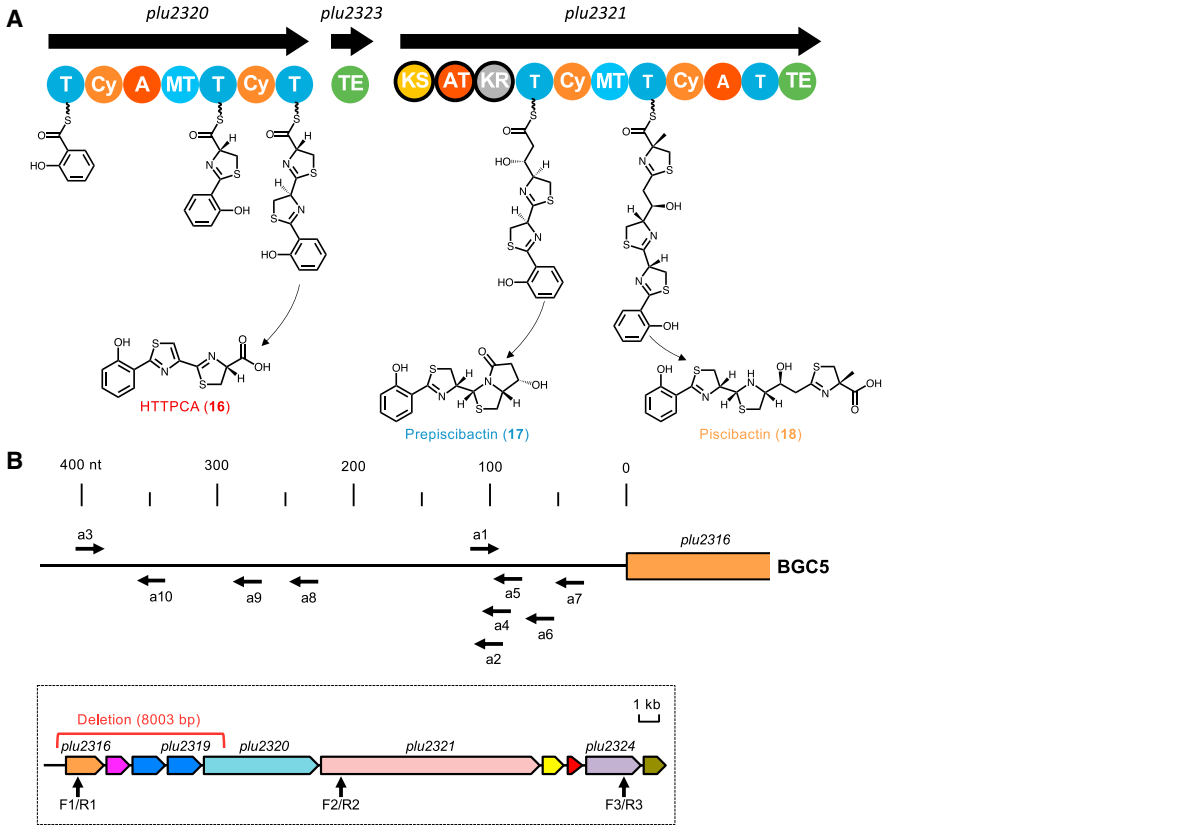
The CRAGE-CRISPR system investigated in this study facilitated activation of diverse BGCs

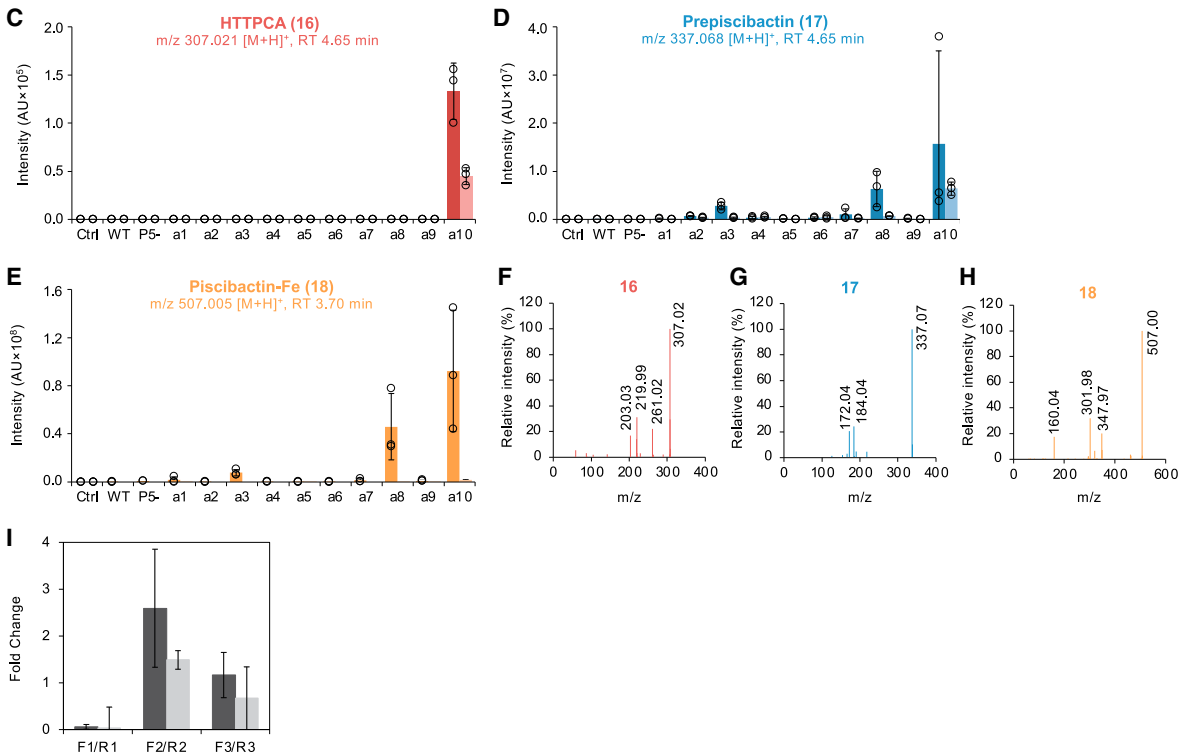
To assess the feasibility of activating PKS/NRPS BGCs with the CRAGE-CRISPRa system, we designed multiple sgRNAs to target the upstream region of each BGC, with the assumption that we could identify predictive rules to efficiently design sgRNAs to activate BGC function. For example, in eukaryotes, target gene expression is increasingly upregulated as an sgRNA is designed to target closer to its start codon (Deaner et al., 2017b), and a broad range of sites upstream of the transcriptional start site are effective (Gilbert et al., 2014). In contrast, our studies revealed no obvious trend in bacteria (Figures 3, 4, 5, 6, and 7).

It is reported that the utility of CRISPRa in bacteria is often limited, as it only moderately promotes gene expression (Wang et al., 2016). Because in our qPCR studies, transcription levels were, in general, improved only 2- to 3-fold, our current study may support this conclusion regarding the gene expression. However, we believe our results clearly suggest that CRISPRa provides sufficient power to effectively increase metabolite production. For example, activation of BGC1 enhanced production of metabolites **1–4** by ~30- to 70-fold in larva and by ~5- to ~300-fold in the M9-based medium. Our previous study suggests that leaky expression may often suffice to activate BGCs integrated into the chromosomes of heterologous hosts (Wang et al., 2019). This would explain why even a low level of expression optimization suffices to activate BGC function.

Chromosomal integration of the CRISPR system might have also increased its stability and consistency among the population. We observed various levels of BGC expression with different sgRNA target sites. This result was reasonable because

represents larval controls that were injected with 0.9% NaCl without strain inoculation. P4- represents the BGC4 knockout strain. The strains of -a1 through -a4 represent constructs with varied sgRNAs. Error bars represent SD; n = 3 technically independent experiments.
(F–H) MS2 spectra for the three metabolites **13–15** (F–H, respectively). The signature fragment ions are marked.
(I) Fold change of the mRNA level of BGC4-a1 determined by qRT-PC. Relative expression level of BGC4 was normalized by that of the housekeeping genes GyrB (dark gray) and lpxC (light gray). Error bars represent SD; n = 3 technical replicates.





(legend on next page)

effective target sites are sensitive to the strength of the target promoter and target site position, even if shifted by a single base (Fontana et al., 2020). Meanwhile, we observed that the forward-oriented sgRNA showed a higher chance of activating BGCs than did the reverse-oriented sgRNA. Among the six activated BGCs, 21 of the 29 forward-oriented sgRNAs (72.4%) achieved increased production of corresponding secondary metabolites, while only 8 of the 26 reverse-oriented sgRNAs (30.8%) achieved increased production of corresponding secondary metabolites at all. This may be explained by a protein structure steric effect on RNAPω according to the crystal structures of Cas9 in complex with sgRNA and target DNA (Jinek et al., 2014; Nishimasu et al., 2014), as the RNAPω was fused to the dCas9 C terminus next to the PAM-interacting (Brown et al., 2019) domain.

A larval hemocoel environment is required for production of some secondary metabolites

BGCs are often regulated by multiple mechanisms. While CRAGE-CRISPRa might activate the promoters of BGCs, another layer of transcriptional and translational regulations may be involved in full BGC activation. In addition, the translated products may need to undergo appropriate post-translational modifications. Unique substrates and co-factors may be required for secondary metabolite production. Because *P. luminescens* is an entomopathogen, production of some secondary metabolites is induced when the strain infects insect hosts to establish and maintain a monoxenic infection (Bode, 2009, 2011; Crawford et al., 2010). Our results for strains injected into insect larvae further demonstrate that the insect hemocoel environment is important for production of some secondary metabolites.

As shown in Figures 7C–7E and S1, BGCs 7 and 3 produced **19–22** in the insect larval hemocoel environment, but stayed silent in the M9-based medium environment. Because **19–22** were produced using heterologous expression in diverse bacteria in the M9-based medium (Figure S1) (Wang et al., 2019), the study shows that **19–22** production does not require metabolites unique to larvae. Instead, expression of BGC7 and BGC3 is likely controlled by an additional layer of regulation. Production of **1–4** from BGC1 was dramatically enhanced after *P. luminescens* was injected into the larvae, suggesting that BGC1 is also subject to another layer of regulatory control (Figures 3C–3F). For BGC9, while production levels of **5–7** were similar in both the M9-based medium and the larvae, **8–12** production was enhanced by 10- to 100-fold or more in larvae (Figures 4C–4J). It has been reported that one of **12's** amino acid constituents is para-aminophenyl alanine (PAPA). The pathway responsible for production of PAPA from chorismate, a metabolic intermediate in a shikimate pathway, was previously reported to be induced in the insect larval hemocoel (Nollmann et al., 2015).

These results suggest that the utility of CRAGE-CRISPRa can be synergistically amplified under specific environmental conditions. As reported in our previous study, BGC8 was characterized as producing *m/z* 357.203 $[M + H]^+$ (Wang et al., 2019). However, this metabolite was not detected in the 15 CRISPRa-engineered strains of the current study (Figure S2C). Neither was any featured metabolite detected from the largest *P. luminescens* NRPS clusters of BGC6 (*plu2670*; ~49 kb) (Figure S2B). In addition, we could not find features unique to BGC 2 (Figure S2A) or 10 (Figure S2D). Further activation, such as that provided by a stronger promoter or some other special trigger(s), might be needed for activation of these BGCs. Meanwhile, the qRT-PCR study confirmed that at least the BGC2 transcription level was significantly increased (Figure S2E), which indicates that some other factors (such as compound stability, metabolite extraction method, and untargeted analysis method for MS data) also need to be considered for identifying metabolites of BGCs. In addition, the utility of CRAGE-CRISPRa will be increased by establishing parameters to best position artificial transcription factors within endogenous promoters for effective upregulation of gene expression (Brown et al., 2019).

Loss-of-function studies using BGC knockouts facilitated identification of secondary metabolites produced from each pathway. CRISPRd technology has contributed to studies of the functions of BGCs. As expected, metabolite production was not detected in *P. luminescens* cultures for BGC1 (Figures 3C–3F), BGC3 (Figure S1C), BGC4 (Figures 5C–5E), BGC5 (Figures 6C–6E), BGC7 (Figures 7C–7E), or BGC9 (Figures 4C–4J), indicating successful deletion of these pathways by CRAGE-CRISPRd. These results demonstrate that CRAGE-CRISPRd is a promising approach for revealing correlations between BGCs and compound production. Our approach combining CRISPRd and CRISPRa helped us identify a metabolite (*m/z* 380.131 $[M + H]^+$) via activation of BGC3, which has not been previously characterized. In addition to helping identify metabolites, CRAGE-CRISPRd can allow comprehensive assessment of the cellular and biosynthetic roles of a given gene in a defined BGC. We found that targeted deletion of a BGC complements the approaches of its activation via CRAGE-CRISPRa in native strains and multi-chassis heterologous expression, allowing researchers to establish the link between a specific BGC and its product. While the pathogenicity of the *P. luminescens* strain

Figure 6. Production of HTTPCA, prepiscibactin, and piscibactin-Fe from CRAGE-CRISPRa-modified expression of BGC5 in *P. luminescens*

(A) A proposed scheme of secondary metabolite biosynthesis catalyzed by BGC5 (*plu2316–plu2324*)-encoded enzymes. NRPS and PKS domains are shown as circles (PKS modules with black borders); letters represent thiolation (T), cyclization (Cy), adenylation (A), methyltransferase (MT), thioesterase (TE), ketosynthase (KS), acyltransferase (AT), and ketoreductase (KR) domains.

(B) Design of the genome modulation for CRAGE-CRISPRa and knockout for BGC5. The spacer location of gRNA before ATG, PAM orientation, and GC content are listed in Table S2. The three pairs of primers for qRT-PC are labeled F1/R1, F2/R2, and F3/R3.

(C–E) Measured MS1 intensity. Shown are **16–18** (C–E, respectively) from CRAGE-CRISPRa-modulated expression of BGC5 in *P. luminescens*. In these charts, dark and light colors represent measured MS1 intensity for each metabolite extracted from the infected larvae and M9-based culture samples, respectively. Ctrl represents larval controls that were injected with 0.9% NaCl without strain inoculation. P5- represents the BGC5 knockout strain. The strains of -a1 through -a10 represent constructs with varied sgRNAs. Error bars represent SD; n = 3 technically independent experiments.

(F–H) MS2 spectra for the three metabolites **16–18** (F–H, respectively). The signature fragment ions are marked.

(I) Fold change of the mRNA level of BGC5-a10 determined by qRT-PC. Relative expression level of BGC5 was normalized to that of the housekeeping genes GyrB (dark gray) and lpxC (light gray). Error bars represent SD; n = 3 technical replicates.

CellPress
OPEN ACCESS

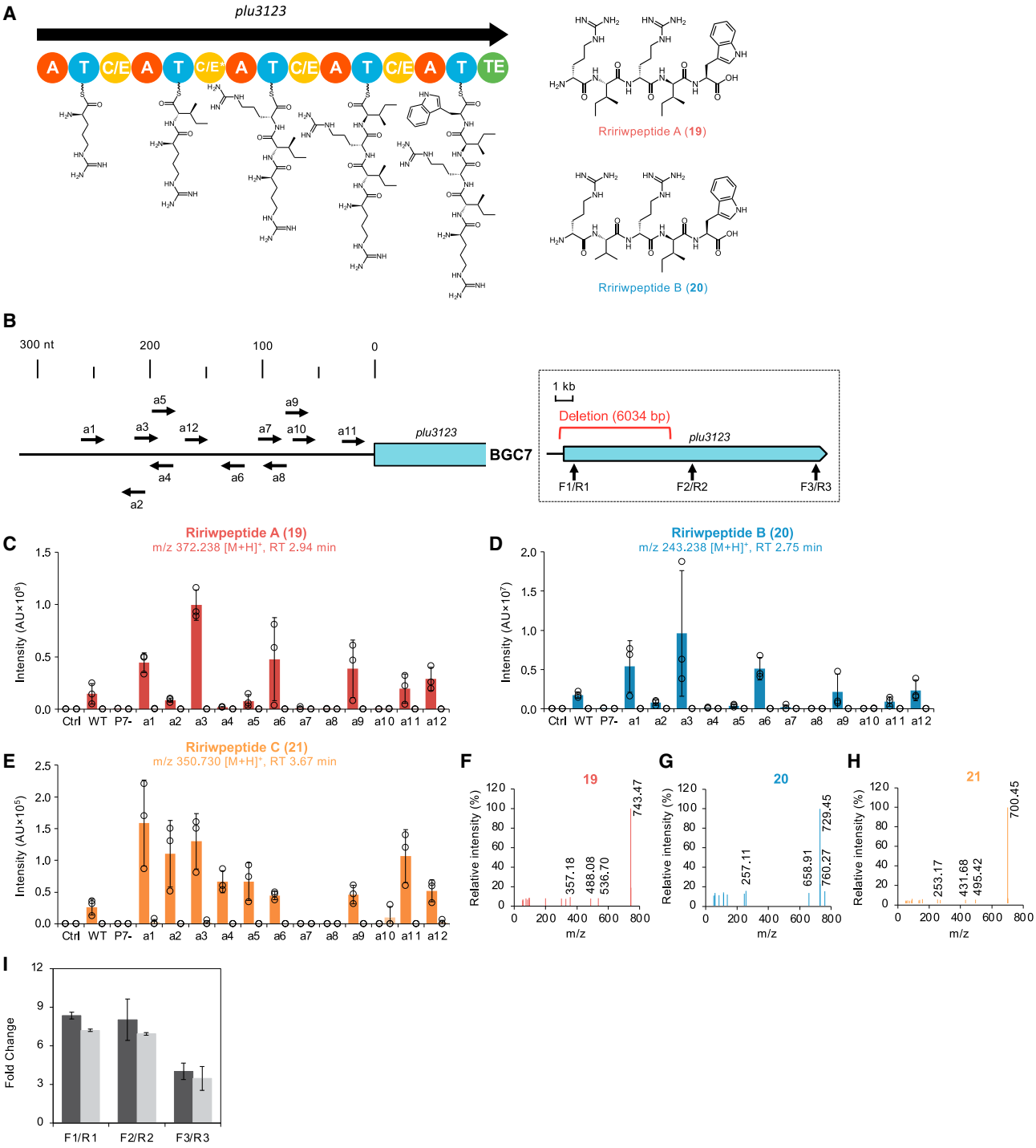


Figure 7. Production of ririwpeptides A–C from CRAGE-CRISPRa-modified expression of BGC7 in *P. luminescens*

(A) A proposed scheme of secondary metabolite biosynthesis catalyzed by BGC7 (*plu3213*)-encoded enzymes. NRPS domains are shown as circles; letters represent adenylation (A), thiolation (T), dual condensation and epimerization (C/E), and thioesterase (TE) domains. C/E* represents dual condensation and epimerization domains with non-functional epimerization domain.

(B) Design of the genome modulation for CRAGE-CRISPRa and knockout for BGC7. The spacer location of gRNA before ATG, PAM orientation, and GC content are listed in Table S2. The three pairs of primers for qRT-PC are labeled F1/R1, F2/R2, and F3/R3.

(C–E) Measured MS1 intensity. Shown are **19–21** (C–E, respectively) from CRAGE-CRISPRa-modulated expression of BGC7 in *P. luminescens*. The secondary metabolites were extracted from cell pellets and analyzed by LC-HRMS. In these charts, dark and light colors represent measured MS1 intensity for each metabolite extracted from the infected larvae and M9-based culture samples, respectively. Ctrl represents larval controls that were injected with 0.9% NaCl and

(legend continued on next page)

to larvae was not affected by deletion of any of the 10 BGCs, we expect that gene deletion by CRAGE-CRISPRd will allow the study of host-microbe interactions mediated through diverse genes more efficiently.

SIGNIFICANCE

In this work, we explored the application of CRAGE-CRISPR-mediated deletion and transcriptional activation of PKS/NRPS BGCs in the native strain of *P. luminescens*. This approach of combining the CRAGE and CRISPR systems enables synergistic exploration of gain- and loss-of-function manipulations, and dramatically expands our ability to regulate BGC expression for functional characterization. Our results also provide evidence that CRAGE-CRISPR could be useful in studies of regulatory controls for natural product BGCs and BGC functions in native environments. Compared with strategies involving heterologous expression, CRAGE-CRISPR may be more useful for characterizing BGCs, especially large ones, because it does not require cumbersome steps for BGC cloning and depends only on the ability of native strains to produce secondary metabolites. While we were unable to find predictive rules for designing effective sgRNA target sites in *P. luminescens*, our analyses suggest that at least one of two or three randomly selected sgRNAs was able to activate BGC function. Therefore, we believe this work represents a valuable expansion of the CRISPR toolbox for functional genomics of non-model microorganisms and provides a foundation for further development of the CRAGE-CRISPRa-mediated transcriptional activator (e.g., multiplexing) as a tool for the gene-to-compound approach for characterization of BGCs and discovery of bioactive secondary metabolites.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Deletion of 10 PKS/NRPS hybrid BGCs in *P. luminescens*
  - Construction of the dCas9-RNAPω plasmid
  - Construction of a series of sgRNA plasmids
  - Activation of 10 PKS/NRPS hybrid BGCs in *P. luminescens*
  - Isolation of total RNA and gene expression analysis
  - Production of secondary metabolites under laboratory cultivation condition
  - Production of secondary metabolites in *G. mellonella* larvae
  - Extraction of secondary metabolites
  - LC-HRMS analysis
  - Untargeted and targeted metabolite analyses
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.chembiol.2021.08.009.

ACKNOWLEDGMENTS

The work conducted by the US Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under contract DE-AC02-05CH11231. We thank A. Wahler for professional editing.

AUTHOR CONTRIBUTIONS

Y.Y., J.F.C., and J.K. conceived the study. Z.Y.W. and D.R. performed the strain engineering. A.K., K.L., S.K., and T.N. ran the LC-HRMS instrument. J.K. conducted the other experiments and data analysis. J.K. and Y.Y. wrote the manuscript.

DECLARATION OF INTERESTS

The Lawrence Berkeley National Laboratory filed a US patent application (20190048354) for CRAGE technology. The application lists Y.Y., J.F.C., and D.R. as inventors. The authors declare no other competing interests.

Received: February 25, 2021
Revised: June 19, 2021
Accepted: August 20, 2021
Published: September 10, 2021

REFERENCES

Alper, H., Fischer, C., Nevoigt, E., and Stephanopoulos, G. (2005). Tuning genetic control through promoter engineering. Proc. Natl. Acad. Sci. U S A *102*, 12678–12683.

Antonello, A.M., Sartori, T., Folmer Correa, A.P., Brandelli, A., Heermann, R., Rodrigues Júnior, L.C., Peres, A., Romão, P.R.T., and Da Silva, O.S. (2018). Entomopathogenic bacteria *Photorhabdus luminescens* as drug source against Leishmania amazonensis. Parasitology *145*, 1065–1074.

Bode, E., Brachmann, A.O., Kegler, C., Simsek, R., Dauth, C., Zhou, Q., Kaiser, M., Klemmt, P., and Bode, H.B. (2015a). Simple "on-demand" production of bioactive natural products. ChemBioChem *16*, 1115–1119.

Bode, H.B. (2009). Entomopathogenic bacteria as a source of secondary metabolites. Curr. Opin. Chem. Biol. *13*, 224–230.

Bode, H.B. (2011). Insect-associated microorganisms as a source for novel secondary metabolites with therapeutic potential. In Insect Biotechnology, A. Vilcinskas, ed. (Springer), pp. 77–93.

Bode, H.B., Brachmann, A.O., Jadhav, K.B., Seyfarth, L., Dauth, C., Fuchs, S.W., Kaiser, M., Waterfield, N.R., Sack, H., Heinemann, S.H., et al. (2015b). Structure elucidation and activity of Kolossin A, the D-/L-Pentadecapeptide

M9-based medium without strain inoculation. P7- represents the BGC7 knockout strain. The strains of -a1 through -a12 represent constructs with varied sgRNAs. Error bars represent SD; n = 3 technically independent experiments.
(F–H) MS2 spectra for the three metabolites **19–21** (F–H, respectively). The signature fragment ions are marked.
(I) Fold change of the mRNA level of BGC7-a3 determined by qRT-PC. Relative expression level of BGC1 was normalized by that of the housekeeping genes GyrB (dark gray) and lpxC (light gray). Error bars represent SD; n = 3 technical replicates.

product of a giant nonribosomal peptide synthetase. Chem. Int. Ed. Engl. *54*, 10352–10355.

Brown, A., Winter, J., Gapinske, M., Tague, N., Woods, W.S., and Perez-Pinera, P. (2019). Multiplexed and tunable transcriptional activation by promoter insertion using nuclease-assisted vector integration. Nucleic Acids Res. *47*, e67.

Clasquin, M.F., Melamud, E., and Rabinowitz, J.D. (2012). LC-MS data processing with MAVEN: a metabolomic analysis and visualization engine. Curr. Protoc. Bioinformatics *14*, Unit14 11.

Crawford, J.M., Kontnik, R., and Clardy, J. (2010). Regulating alternative lifestyles in entomopathogenic bacteria. Curr. Biol. *20*, 69–74.

Deaner, M., and Alper, H.S. (2017a). Systematic testing of enzyme perturbation sensitivities via graded dCas9 modulation in Saccharomyces cerevisiae. Metab. Eng. *40*, 14–22.

Deaner, M., Mejia, J., and Alper, H.S. (2017b). Enabling graded and large-scale multiplex of desired genes using a dual-mode dCas9 activator in Saccharomyces cerevisiae. ACS Synth. Biol. *6*, 1931–1943.

Dong, C., Fontana, J., Patel, A., Carothers, J.M., and Zalatan, J.G. (2018). Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. Nat. Commun. *9*, 2489.

ffrench-Constant, R., Waterfield, N., Daborn, P., Joyce, S., Bennett, H., Au, C., Dowling, A., Boundy, S., Reynolds, S., and Clarke, D. (2003). *Photorhabdus*: towards a functional genomic analysis of a symbiont and pathogen. FEMS Microbiol. Rev. *26*, 433–456.

Fontana, J., Dong, C., Kiattisewee, C., Chavali, V.P., Tickman, B.I., Carothers, J.M., and Zalatan, J.G. (2020). Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. Nat. Commun. *11*, 1618–1678.

Fu, J., Bian, X., Hu, S., Wang, H., Huang, F., Seibert, P.M., Plaza, A., Xia, L., Muller, R., Stewart, A.F., et al. (2012). Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. Nat. Biotechnol. *30*, 440–446.

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., et al. (2014). Genome-scale CRISPR-mediated control of gene repression and activation. Cell *159*, 647–661.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell *154*, 442–451.

Hamilton, C.M., Aldea, M., Washburn, B.K., Babitzke, P., and Kushner, S.R. (1989). New method for generating deletions and gene replacements in Escherichia coli. J. Bacteriol. *171*, 4617–4622.

Hiom, K. (2009). DNA repair: common approaches to fixing double-strand breaks. Curr. Biol. *19*, R523–R525.

Huang, A.C., Jiang, T., Liu, Y.X., Bai, Y.C., Reed, J., Qu, B., Goossens, A., Nutzmann, H.W., Bai, Y., and Osbourn, A. (2019). A specialized metabolic network selectively modulates Arabidopsis root microbiota. Science *364*, eaau6389.

Jinek, M., Jiang, F., Taylor, D.W., Sternberg, S.H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., et al. (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science *343*, 1247997.

Joyce, S.A., Watson, R.J., and Clarke, D.J. (2006). The regulation of pathogenicity and mutualism in Photorhabdus. Curr. Opin. Microbiol. *9*, 127–132.

La Russa, M.F., and Qi, L.S. (2015). The new state of the art: CRISPR for gene activation and repression. Mol. Cell. Biol. *35*, 3800–3809.

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

Lian, J., Schultz, C., Cao, M., HamediRad, M., and Zhao, H. (2019). Multi-functional genome-wide CRISPR system for high throughput genotype-phenotype mapping. Nat. Commun. *10*, 5794.

Liu, H., Robinson, D.S., Wu, Z.Y., Kuo, R., Yoshikuni, Y., Blaby, I.K., and Cheng, J.F. (2020). Bacterial genome editing by coupling Cre-lox and CRISPR-Cas9 systems. PLoS One *15*, e0241867.

Medema, M.H., and Fischbach, M.A. (2015). Computational approaches to natural product discovery. Nat. Chem. Biol. *11*, 639–648.

Milshteyn, A., Colosimo, D.A., and Brady, S.F. (2018). Accessing bioactive natural products from the human microbiome. Cell Host Microbe *23*, 725–736.

Mougiakos, I., Bosma, E.F., Weenink, K., Vossen, E., Goijvaerts, K., van der Oost, J., and van Kranenburg, R. (2017). Efficient genome editing of a facultative thermophile using mesophilic spCas9. ACS Synth. Biol. *6*, 849–861.

Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell *156*, 935–949.

Nollmann, F.I., Dauth, C., Mulley, G., Kegler, C., Kaiser, M., Waterfield, N.R., and Bode, H.B. (2015). Insect-specific production of new GameXPeptides in *Photorhabdus luminescens* TTO1, widespread natural products in entomopathogenic bacteria. ChemBioChem *16*, 205–208.

Oka, M., Nishiyama, Y., Ohta, S., Kamei, H., Konishi, M., Miyaki, T., Oki, T., and Kawaguchi, H. (1988). Glidobactins A, B and C, new antitumor antibiotics. I. Production, isolation, chemical properties and biological activity. J. Antibiot. *41*, 1331–1337.

Peng, R., Wang, Y., Feng, W.W., Yue, X.J., Chen, J.H., Hu, X.Z., Li, Z.F., Sheng, D.H., Zhang, Y.M., and Li, Y.Z. (2018). CRISPR/dCas9-mediated transcriptional improvement of the biosynthetic gene cluster for the epothilone production in Myxococcus xanthus. Microb. Cell Fact. *17*, 15.

Pluskal, T., Castillo, S., Villar-Briones, A., and Oresic, M. (2010). MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics *11*, 395.

Ruiz, P., Balado, M., Fuentes-Monteverde, J.C., Toranzo, A.E., Rodriguez, J., Jimenez, C., Avendano-Herrera, R., and Lemos, M.L. (2019). The fiish pathogen *Vibrio ordalii* under iron deprivation produces the siderophore piscibactin. Microorganisms *7*, 313.

Rutledge, P.J., and Challis, G.L. (2015). Discovery of microbial natural products by activation of silent biosynthetic gene clusters. Nat. Rev. Microbiol. *13*, 509–523.

Stein, M.L., Beck, P., Kaiser, M., Dudler, R., Becker, C.F.W., and Groll, M. (2012). One-shot NMR analysis of microbial secretions identifies highly potent proteasome inhibitor. Proc. Natl. Acad. Sci. U S A *109*, 18367–18371.

Stock, S.P., Kusakabe, A., and Orozco, R.A. (2017). Secondary metabolites produced by Heterorhabditis symbionts and their application in agriculture: what we know and what to do next. J. Nematol *49*, 373–383.

Terui, Y., Nishikawa, J., Hinoo, H., Kato, T., and Shoji, J. (1990). Structures of cepafungins I, II and III. J. Antibiot. *43*, 788–795.

Theodore, C.M., King, J.B., You, J., and Cichewicz, R.H. (2012). Production of cytotoxic glidobactins/luminmycins by *Photorhabdus asymbiotica* in liquid media and live crickets. J. Nat. Prod. *75*, 2007–2011.

Thode, S.K., Rojek, E., Kozlowski, M., Ahmad, R., and Haugen, P. (2018). Distribution of siderophore gene systems on a Vibrionaceae phylogeny: database searches, phylogenetic analyses and evolutionary perspectives. PLoS One *13*, e0191860.

Tobias, N.J., and Bode, H.B. (2019). Heterogeneity in bacterial specialized metabolism. J. Mol. Biol. *431*, 4589–4598.

Tobias, N.J., Shi, Y., and Bode, H.B. (2018). Refining the natural product repertoire in entomopathogenic bacteria. Trends Microbiol. *26*, 833–840.

Vizcaino, M.I., Guo, X., and Crawford, J.M. (2014). Merging chemical ecology with bacterial genome mining for secondary metabolite discovery. J. Ind. Microbiol. Biotechnol. *41*, 285–299.

Wang, B., Zhao, Z., Jabusch, L.K., Chiniquy, D.M., Ono, K., Conway, J.M., Zhang, Z., Wang, G., Robinson, D., Cheng, J.F., et al. (2020). CRAGE-duet

facilitates modular assembly of biological systems for studying plant-microbe interactions. ACS Synth. Biol. *9*, 2610–2615.

Wang, G., Zhao, Z., Ke, J., Engel, Y., Shi, Y.M., Robinson, D., Bingol, K., Zhang, Z., Bowen, B., Louie, K., et al. (2019). CRAGE enables rapid activation of biosynthetic gene clusters in undomesticated bacteria. Nat. Microbiol. *4*, 2498–2510.

Wang, H., La Russa, M., and Qi, L.S. (2016). CRISPR/Cas9 in genome editing and beyond. Annu. Rev. Biochem. *85*, 227–264.

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., et al. (2015). Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. Cell *160*, 339–350.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial strains		
E. coli TransforMax™ EC100D™ pir+	Thermo Fisher	Cat # NC9801351
E. coli BW29427	JGI	# aka WM3064
Chemicals, peptides, and recombinant proteins		
LB agar plate	Teknova	Cat #L1066
LB (Miller's) broth	Growcells	Cat # MBLE-7030
Apramycin sulfate salt	Sigma-Aldrich	Cat # A2024-1G
D-glucose	Sigma-Aldrich	Cat #G8270
Yeast extract	Sigma-Aldrich	Cat #Y1625
M9 minimal salts (2X)	Fisher	Cat # A1374401
1M $MgSO_4$ solution	Sigma-Aldrich	Cat #M3409
1M $CaCl_2$ solution	Sigma-Aldrich	Cat # 21,115
Citric acid monohydrate	Sigma-Aldrich	Cat # 1909
Trace mineral solution	ATCC	Cat # MD-TMS
Vitamin supplement	ATCC	Cat # MD-VS
L-Proline	Sigma-Aldrich	Cat #P0380
L- Aspartic acid	Sigma-Aldrich	Cat # A9256
L-Serine	Sigma-Aldrich	Cat #S4311
L-Valine	Sigma-Aldrich	Cat #V0513
L-Isoleucine	Sigma-Aldrich	Cat #I2752
Amino acids mix	Sunrise Science Products Inc.	Cat # 1001
Critical commercial assays		
PureLink RNA Mini kit	Invitrogen	Cat # 12183018A
ezDNase Enzyme	Invitrogen	Cat # 11,766,051
Quant-it dsDNA HS Assay kit	Invitrogen	Cat #Q32851
SuperScrip IV VIL Master Mix	Invitrogen	Cat # 11,756,050
SYBR Green Supermix	Bio-Rad	Cat # 1,725,270
Experimental models: organisms/strains		
P. luminescens laumondii TT01	DSMZ	Cat # 15,139
G. mellonella larvae	Carolina Biological Supply Company	Cat # 143,928
Software and algorithms		
Xcalibur software	Thermo Fisher	https://www.thermofisher.com/order/catalog/product/OPTON-30965#/OPTON-30965
Maven	Rabinowitz Lab	http://maven.princeton.edu
MZmine	MZmine 2	http://mzmine.github.io/
CFX Maestro	Bio-Rad	https://www.bio-rad.com/en-us/product/cfx-maestro-software-for-cfx-real-time-pcr-instruments?ID=OKZP7E15

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Yasuo Yoshikuni (yyoshikuni@lbl.gov).

Materials availability

All unique/stable reagents generated in this study will be made available from the lead contact on request, upon completion of a Materials Transfer Agreement.

Data and code availability

The raw LCMS data files generated in this study are available upon reasonable request to the lead contact. This study did not generate or analyze any computational datasets or code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

P. luminescens laumondii TT01 was obtained from DSMZ. The native cells were genetically modified and cultured as described in Method Details (below). The last-instar *G. mellonella* larvae weighing between 150 and 250 mg were obtained from Carolina Biological Supply Company, and were used in all experiments within 7 days of shipment from the vendor. Culture conditions and procedures are explained in Method Details.

METHOD DETAILS

The set of plasmids used for this study is listed in Table S1; their sequences are shown as supplementary sequences in GenBank format. All CRISPRa strains built are listed in Table S2. All primers for RT-PCR are summarized in Table S3. The *E. coli* TransforMax EC100D pir^{+} strain was used as a cloning host for constructs containing the R6Kr origin of replication. *E. coli* BW29427 was used as a conjugal donor strain to transfer plasmids to *P. luminescens subsp. laumondii* TTO1. Luria-Bertani (LB) medium was used for cultivation of both *E. coli* and *P. luminescens*. The conjugation was performed as previously described (Wang et al., 2019). In brief, *E. coli* BW29427 as the conjugal donor strain was inoculated in LB medium containing 0.3 mM diaminopimelic acid (DAP) and 50 μg mL^{-1} kanamycin and was grown at 37°C in an incubation shaker at 200 rpm overnight. The recipient *P. luminescens* was inoculated into LB medium and were grown at 28°C in an incubation shaker at 200 rpm until they reached the late log phase. Donor and recipient cells were washed three times with LB medium containing 0.3 mM DAP and were mixed 4:1 by optical density at 600 nm (OD600). This mixture was then pelleted and transferred on a nitrocellulose filter membrane on top of an LB agar plate containing 0.3 mM DAP, and was incubated at 28°C for 5–12 hr. The bacterial mixture grown on the membrane was scraped off, resuspended into the LB liquid medium, and spread on an LB agar plate containing 50 μg mL^{-1} kanamycin.

Deletion of 10 PKS/NRPS hybrid BGCs in *P. luminescens*

We deleted 10 PKS/NRPS hybrid BGCs using CRAGE-CRISPR/Cas9. The CRAGE-Duet system was first implemented in *P. luminescens subsp. laumondii* TTO1 (Liu et al., 2020). Cas9/RecET was integrated into the first LP site, and sgRNA and a repairing arm were integrated into the second LP site to knock out 10 PKS/NRPS hybrid BGCs to yield P1- to P10-. Deletion of these BGCs was confirmed by PCR amplification of the expected DNA band size (Liu et al., 2020).

Construction of the dCas9-RNAPω plasmid

A plasmid, pR6K-2L-dCas9-RNAPω, was built based on existing pR6K-2L-Cas9 plasmids (Liu et al., 2020). The primers used for introduction of point mutations to the Cas9 gene to create a dCas9 gene are listed in Table S1. Subsequently, a synthetic DNA fragment coding for an RNAP ω subunit was inserted into the 3′ end of the dCas9 gene. This plasmid was sequence-verified using a PacBio Sequel II platform.

Construction of a series of sgRNA plasmids

The 20 bp sgRNA target sequences were selected arbitrarily and cloned into accessory vector pR6K-loxWT2272-Plu-sgRNA-BsaI by assembling the pairs of primers listed in Table S1 (Liu et al., 2020). These sgRNA-containing plasmids were sequence-verified by Sanger sequencing.

Activation of 10 PKS/NRPS hybrid BGCs in *P. luminescens*

To activate the 10 BGCs, we first integrated the dCas9-RNAPω gene into the first integration site, flanked by *loxP* and *lox5171*. This strain was then used to integrate a series of sgRNA target constructs into the second integration site, flanked by *lox2272* and *loxP*. Both transformations were mediated by conjugation.

Isolation of total RNA and gene expression analysis

After 3-hr cultivation, total RNA was isolated from 1 mL of culture of WT or BGC1-a5, BGC2-a1, BGC3-a2, BGC4-a1, BGC5-a10, BGC7-a3, or BGC9-a1, respectively, using a PureLink RNA Mini Kit (Invitrogen) and following the manufacturer's instructions. DNA in the RNA sample was removed by digestion using the ezDNaseTM Enzyme Kit (Invitrogen) following the manufacturer's protocol. RNA samples were confirmed to be DNA-free by Quant-it dsDNA HS Assay kit (Invitrogen), and then reverse transcription experiments were performed using the SuperScript IV VILO Master Mix (Invitrogen) following the manufacturer's protocol; the

synthesized cDNAs were stored at −80°C afterward. qPCR of 1:10-diluted cDNA was then conducted using a CFX384 Real-Time PCR Detection System version (Bio-Rad) with the following amplification program: 95°C for 30 s followed by PCR (40 cycles of denaturation at 95°C for 10 s, annealing and extension at 60°C for 30 s). The melt curve analyses were performed from 65°C to 95°C with 0.5°C increments at 5 s per step. Amplification was performed using PowerUp SYBR Green Master Mix (Bio-Rad) and primers listed in Table S3. Three pairs of RT-qPCR primers were designed in-house for each BGC. Relative expression levels of target BGCs were calculated by -$\Delta\Delta C_q$ using Bio-Rad CFX Maestro. All values were normalized using the housekeeping reference expression levels of the *GyrB* and *lpxC* genes. RT-qPCR was carried out in triplicate for each CRISPRa construct.

Production of secondary metabolites under laboratory cultivation condition

Secondary metabolites were produced using a previously described method (Wang et al., 2019). For each strain, a single colony was inoculated into 2 mL of LB media with 50 μg/mL apramycin. The laboratory fermentation was done in triplicate. After overnight growth, an aliquot of this culture was washed with M9-based media (4 g/L glucose, 5 g/L yeast extract, 500 mL/L 2X M9 salts, 3 g/L citric acid monohydrate, 2 mL/L 1M $MgSO_4$ solution, 100 μL/L $CaCl_2$, 2.5 mL/L trace mineral solution, 2.5 mL/L vitamin supplement), and inoculated into 5 mL of the medium with 50 μg/mL apramycin to a final OD_{600nm} of 0.1. The cultures were grown at 28°C in an incubation shaker at 200 rpm for 3 days or other specifically defined time. To prepare the *G. mellonella* instant media, the larvae were frozen at −80°C overnight, freeze-dried or not, and pulverized in a tissue mortar. In 100 mL M9-based media, 8 g of the larvae cadavers were suspended; this was centrifuged at 6000 rpm for 5 min, and the supernatant was filtered through a 0.22-μm PVDF membrane for sterilization.

Production of secondary metabolites in *G. mellonella* larvae

G. mellonella larvae were kept at room temperature in darkness prior to use. Last-instar larvae weighing between 150 and 250 mg were used in all experiments within 7 days of shipment from the vendor (Carolina Biological Supply Company, NC). The strains were grown overnight in 2 mL of LB media at 28°C. Before injection, bacterial suspensions were normalized by 0.9% NaCl to a density of $8x10^7$ colony-forming units (cfu)/mL.

An automated syringe pump was used for intra-hemocoelic injection of the larvae. To control the precise injection volume, the automated syringe pump was set at a volume of 11 μL and an injecting rate of 66 μL/min. A 1 mL sterilized BD syringe was filled with 300 μL bacterial suspension for infection of three larvae. The bubbles were removed by carefully tapping the syringe and injecting the air, and then a 30 gauge 1/2 BD needle was attached to the syringe. After the syringe was placed in the injector, a blank injection of 10 μL in an empty tube was performed to control the injected volume. The strain suspension was injected into the larva hemocoel from the bottom. After injection, larvae were incubated in petri dishes at 25°C in the dark. The mortality was monitored daily, and the larvae were stored at −80°C for 72 hr beyond the time of injection. Larvae were considered dead when they did not move in response to touch. For all experiments, two control groups were used; the first group was larvae inoculated with 0.9% NaCl to monitor for killing caused by physical trauma, and the second was larvae injected with pathway-knockout strains.

Extraction of secondary metabolites

For production in media, 2 mL of each culture was centrifuged at 10,000 rpm for 5 min. The supernatant was transferred to another tube, and the pellet was resuspended in 1 mL acetone for extraction. For production in larval hemocoel, each larva was placed in a 2 mL tube with two of the 5 mm glass beads, homogenized separately by bead beater homogenizer for 30 s, and extracted with 1 mL acetone. Each suspension was sonicated for 15 min, vortexed at 1,500 rpm for 20 min, and centrifuged at 10,000 rpm x *g* at 4°C for 5 min to remove the debris. The supernatant was collected and dried using SAVANT SPD111 SpeedVac Concentrator (Thermo Scientific, Waltham, MA). The debris residue was extracted again with 1 mL ethyl acetate. The collected supernatant was combined with the dried acetone extract and again dried using SAVANT SPD111 SpeedVac Concentrator (Thermo Scientific, Waltham, MA). Lastly, the dried solid was resuspended in 200 μL of methanol containing 1 μg/mL of an internal standard, 2-amino-3-bromo-5-methylbenzoic acid (ABMBA), and was filtered through a 0.22-μm PVDF membrane (Millipore Ultrafree-MC) to prepare a sample for LC-HRMS analysis.

LC-HRMS analysis

Reverse-phase chromatography was performed using an Agilent 1290 LC stack with a C18 column (Agilent ZORBAX Eclipse Plus C18, Rapid Resolution HD, 2.1 × 50 mm, 1.8 μm) at a flow rate of 0.4 mL/min with a 2–3 μL injection volume. To detect compounds, **samples were run** on the C18 column at 60°C and equilibrated with 100% buffer A (100% H_2O with 0.1% formic acid) for 1 min, buffer A was diluted down to 0% with buffer B (100% ACN with 0.1% formic acid) over 8 min, and analytes were removed by isocratic elution in 100% buffer B for 1.5 min. MS1 and MS2 data were collected using a Q Exactive Orbitrap MS (Thermo Scientific, San Jose, CA). Full MS spectra were collected from *m/z* 135 to *m/z* 2000 at a resolution of 70,000, with MS2 fragmentation data acquired using 10, 20, and 30 V collision energies at a resolution of 17,500. Exact mass and retention time (RT) coupled with MS2 fragmentation spectra were used to identify compounds.

Untargeted and targeted metabolite analyses

Untargeted metabolite analysis was performed using MAVEN, as described previously (Clasquin et al., 2012). The minimal peak height was set to 10^7 or 10^6 for MS1, and the minimal ratio of peak height between MS1 for the culture extracts of engineered

and control strains was set to 100, 50, or 10. MAVEN, MZmine, and Thermo Xcalibur software applications were subsequently used to extract peak height for all identified/targeted MS1s and created heat maps. MS1 features were accepted as metabolite ions unique to BGC expression only when they appeared on extracts from strains expressing BGCs except in cases in which the genome analysis suggested that chassis strains contained the BGC homologs. The high resolution MS2 data for the identified MS1s were manually curated using MZmine (Pluskal et al., 2010) and Thermo Xcalibur (Thermo Scientific, San Jose, CA).

QUANTIFICATION AND STATISTICAL ANALYSIS

For LC-HRMS intensity of produced secondary metabolites and RT-qPCR results, each experiment was performed in technically independent triplicates. Details of replicates and data analysis for specific experiments can be found in the figure legends, Figure S1 and S2, Tables S1, S2, and S3, or in the STAR Methods section. The reported mean is equivalent to the average of the values determined for each of the replicates and the reported standard variation is a measure of the variance relative to the determined mean. The software tools used for LC-HRMS are listed on key resources table.