

Repurposing endogenous type I-E CRISPR-Cas systems for natural product discovery in *Streptomyces*

Received: 9 May 2024

Accepted: 1 November 2024

Published online: 13 November 2024

Qun Zhou^{1,2,7}, Yatong Zhao^{1,2,7}, Changqiang Ke^{3,7}, Haojun Wang^{4,7}, Sheng Gao⁴, Hui Li⁴, Yan Zhang ^{1,5}, Yang Ye ³✉ & Yunzi Luo ^{1,2,6}✉

The multifunctional proteins of class 2 CRISPR systems such as Cas9, have been employed to activate cryptic biosynthetic gene clusters (BGCs) in *Streptomyces*, which represent a large and hidden reservoir of natural products. However, such approaches are not applicable to most *Streptomyces* strains with reasons to be comprehended. Inspired by the prevalence of the class 1 subtype especially the type I-E CRISPR system in *Streptomyces*, here we report the development of the type I-E CRISPR system into a series of transcriptional regulation tools. We further demonstrate the effectiveness of such activators in nine phylogenetically distant *Streptomyces* strains. Using these tools, we successfully activate 13 out of 21 BGCs and lead to the identification and characterization of one polyketide, one Ripp and three alkaloid products. Our work is expected to have a profound impact and to facilitate the discovery of numerous structurally diverse compounds from *Streptomyces*.

Microbes are renowned as prolific antibiotic producers, among them, *Streptomyces* are particularly important for producing approximately two-thirds of natural antibiotics¹. However, the pace of discovering new antibiotics from microbes has slowed down since the late twentieth century². With the development of second-generation sequencing technology in the new century, the volume of microbial genomic data has been rapidly expanded. The genomic data disclosed a rich repertoire of biosynthetic gene clusters (BGCs). However, further bioinformatic analysis revealed that, indeed, over 90% of putative microbial metabolites are not produced under normal laboratory culture conditions in the absence of unknown in vivo regulatory signals that trigger their biosynthesis³. Therefore, developing tools to activate cryptic BGCs has become an effective strategy for mining new natural products.

Cas9 of the class 2 CRISPR system and its homologs, such as Cas12a, versatile enzymes that contain multiple domains and exhibit

both nuclease and helicase activities, have been extensively employed for genome editing in various organisms^{4–6}. Their applications in *Streptomyces* have also been reported, which include deletion and insertion of large fragments^{7–9}, and direct capture of large BGCs^{10,11}. The class 2 CRISPR system can also be used for gene regulation. Previously, a promoter knock-in strategy was applied to drive the in situ expression of cryptic BGCs, leading to the activation of four BGCs in five *Streptomyces* strains¹².

Gene expression regulation can also be precisely programmed in a variety of organisms, from bacteria, and yeast to humans, by CRISPRi (CRISPR interference) and CRISPRa (CRISPR activation), which employ dCas proteins (dCas9, dCas12a), mutants of Cas9 or Cas12a lacking nuclease activity and their fusion proteins with an effector attached respectively^{4,13,14}. The CRISPRi system targets the ORF upstream region, blocking the recruitment of RNA polymerase and attenuating its activity¹⁵, resulting in up to 99% decrease in gene expression in

¹Frontiers Science Center for Synthetic Biology and Key Laboratory of Systems Bioengineering (Ministry of Education), School of Chemical Engineering and Technology, Tianjin University, Tianjin, China. ²Haihe Laboratory of Sustainable Chemical Transformations, Tianjin, China. ³State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China. ⁴Department of Gastroenterology, State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu, China. ⁵New Cornerstone Science Laboratory, School of Pharmaceutical Science and Technology, Tianjin University, Tianjin, China. ⁶Georgia Tech Shenzhen Institute, Tianjin University, Tangxing Road 133, Nanshan District, Shenzhen, China.

⁷These authors contributed equally: Qun Zhou, Yatong Zhao, Changqiang Ke, Haojun Wang. ✉e-mail: yye@simmm.ac.cn; yunzi.luo@tju.edu.cn

bacteria¹⁶. This system can be developed to target multiplex genes. Down-regulation of seven genes simultaneously in *Saccharomyces cerevisiae* was recently reported¹⁷. By contrast, the CRISPRa system turns on gene expression, in which the dCas protein is fused to an RNA polymerase activation domain, guiding the recruitment of RNA polymerases to specific genomic loci and triggering the transcription of downstream genes¹⁸. Taken together, the power of CRISPRi and CRISPRa lies in the flexibility of gRNAs designed to target any DNA sequence in close proximity to a three-nucleotide protospacer adjacent motif (PAM), as well as the ability to rapidly and precisely modulate endogenous gene expression without chromosome engineering. Although dCas9-based CRISPRi systems have demonstrated functionality in select *Streptomyces* strains^{4,19–22}, the dCas9-based CRISPRa systems have typically shown restricted efficacy in the majority of *Streptomyces* strains, with a recent report of success only in a specific *Streptomyces* strain²³.

Exploring the diverse CRISPR systems presents a promising avenue for developing bacterial gene regulation tools with unique characteristics. Given that the majority of prokaryotes possess indigenous CRISPR systems, it is feasible to leverage these inherent immune mechanisms for gene regulation purposes. This approach allows for the utilization of nature's own toolkit, potentially leading to more efficient and targeted genetic manipulation strategies within bacterial hosts. A recent in-depth study revealed that most *Streptomyces* strains harbor the class 1 type I-E CRISPR systems²⁴. Unlike the class 2 CRISPR system that consists of a single protein, all class 1 systems contain multiple Cas subunits, which form a complex known as Cascade (the CRISPR-associated complex for antiviral defense)²⁵. A typical type I-E CRISPR system contains three core genes (*cas1*, *cas2*, and *cas3*) and Cascade genes (*casA-casB-cas7-cas5-cas6*)²⁵. Cas1 and Cas2 form an alliance to capture a piece of DNA from the invader during the adaptation process. Cas3 alone is capable of unwinding DNA/DNA and RNA/DNA duplexes and cutting single-stranded DNA²⁶, while the Cascade/Cas3 complex mediates DNA nicking in the interference process. Using the type I-A CRISPR system, different types of mutation were generated, including deletion, insertion, and point mutations²⁷. The employment of type I-B CRISPR systems has illustrated the facile programmability and remarkable efficacy of native CRISPR in facilitating genome engineering within *Clostridium tyrobutyricum*^{28,29}. Similarly, the application of type I-E CRISPR systems has unveiled the capacity to engineer a CRISPRa system endowed with unique and broadened regulatory capabilities in *E. coli*^{30,31}. These discoveries suggested that the *Streptomyces* endogenous class 1 CRISPR systems may be developed into tools for transcriptional regulation.

In this study, we systematically explore the *Streptomyces* type I-E CRISPR Cascade, based on which a series of repressors/activators are designed and demonstrated to be effective in activating cryptic BGCs of the non-model *Streptomyces* strains. Thirteen BGCs responsible for the synthesis of various natural products, including linaridins, NRPS, and polyketides, were activated, allowing the identification and characterization of these compounds. In summary, we develop a robust toolkit that aids in precisely regulating gene expression and effectively mining natural products in *Streptomyces*.

Results

Identification of type I-E CRISPR systems in *Streptomyces*

67 complete genomic sequences of *Streptomyces* strains, all that is available at NCBI, as well as the sequences of seven strains isolated and sequenced in our laboratory, were retrieved for analysis (Supplementary Data 1). Sequence alignment of the most conserved Cas1, followed by PILER-CR analysis to find the striking signature of a CRISPR system in the neighborhood, the repetitive sequences formed by the reciprocal structure of the CRISPR array, uncovered type I-E CRISPR systems in a total of 25 *Streptomyces* strains. Among them, three strains, CFMR-7, Mg1, and SCSIO 03032, possess two systems, while some

other strains harbor variants of the system, with one lacking the *cas2* gene and seven possessing a *casA-cas3* fusion gene. These gene clusters were grouped and displayed based on the phylogenetic distances of their hosts (Fig. 1). Next, we decided to employ a *cas3*-free system for gene regulation, as the *cas3*-free complex is expected to introduce site-specific hindrance of polymerase access while maintaining the integrity of the target DNA³². To avoid complexity, those clusters containing a *casA-cas3* fusion were excluded. Drawing from the protein alignment outcomes of the Cascade (Supplementary Data 2), we chose two type I-E systems that exhibit the highest and lowest homology to the Sav Cascade among the strains in our study to provide an analysis that spans the spectrum of genetic similarity, thereby enhancing our understanding of the functional diversity and potential applications of these CRISPR systems. Thus, three type I-E CRISPR systems from *Streptomyces avermitilis* MA-4680, *Streptomyces* T1-5, and *Streptomyces* A-23, strains available in our laboratory, were proceeded for activity screen.

Developing the CRISPR-based transcriptional repressor 1.0 (CTR 1.0)

To achieve stable regulation in *Streptomyces*, we designed a dual plasmid system to express the Cascade module and spacer plasmid separately. Once the Cascade system is integrated into the genome, subsequent editing of various pathways becomes more manageable and efficient, facilitated by the use of a smaller plasmid. To test these CRISPR systems in *S. coelicolor*, a strain that does not possess its own type I-E CRISPR system, the Cascade expression cassettes derived from *Streptomyces* T1-5, *Streptomyces* A-23 and *Streptomyces avermitilis* MA-4680, with either the *kasO**p promoter and *oop* terminator pairs or the native promoter and terminator pairs (Fig. 2a), were integrated into the Φ C31 *attB* site of *S. coelicolor* genome, generating the strains R1-1 to R1-6 (Supplementary Table 1) presumably expressing the corresponding Cascade proteins at stable levels. Next, a pYL plasmid carrying a crRNA cassette with direct repeat (DR) sequences at both ends of the spacer was conjugated into the six *S. coelicolor* strains, generating the CRISPR-based transcriptional repressor 1.0 (CTR 1.0).

We next decided to screen the CTR 1.0 systems, therefore, the repression of the *actII-ORF4* gene (*scoS085*) using a colorimetric assay was designed. The *actII-ORF4* gene, located in the gene cluster responsible for the biosynthesis of actinorhodin (ACT), a characteristic blue pigment of *S. coelicolor*, encodes a transcription activator³³. Its repression shuts down the entire BGC, leading to the color of the bacterial colony shifting from blue to brown. Five crRNAs (a1–a5), each containing the canonical 5'-AAG-3' PAM and one of the five spacers in the *actII-ORF4* region, including the promoter, the coding strand, and the non-template strand were employed for each of the six R1 strains (Fig. 2b). We observed the expected color change with three crRNAs in the R1-1 strain and four crRNAs in the R1-5 strain (Fig. 2c, S1). Among the six R1 strains, R1-5, that exhibited the best repression effect was chosen for subsequent experiments.

The biosynthesis of undecylprodigiosin (RED), a red-pigmented antibiotic, was then used as another readout of the assay and the proof of concept for CTR 1.0. The *red Q* gene (*scoS887*) in the RED BGC encoding a positive regulatory protein of the SARPs family was chosen as the target. This protein is required for the synthesis of RED³⁴, and its repressed expression has been reported to result in a shift in colony color from red to yellow. Four crRNAs (r1–r4) with two spacers in the promoter region (r1 and r2) and two spacers in the non-template strand (r3 and r4) were designed, aiming to repress *redQ* in the R1-5 strain (Fig. 2d). All four crRNAs efficiently repressed RED synthesis as observed by the colony colors (Fig. 2e), and transcriptional analysis revealed a maximum reduction of 98% of the *red Q* mRNA level (Supplementary Fig. 2a).

Dual-target transcriptional repression was then evaluated for CTR 1.0. Spacer r2a1(r2 and a1) repressed the synthesis of both RED and

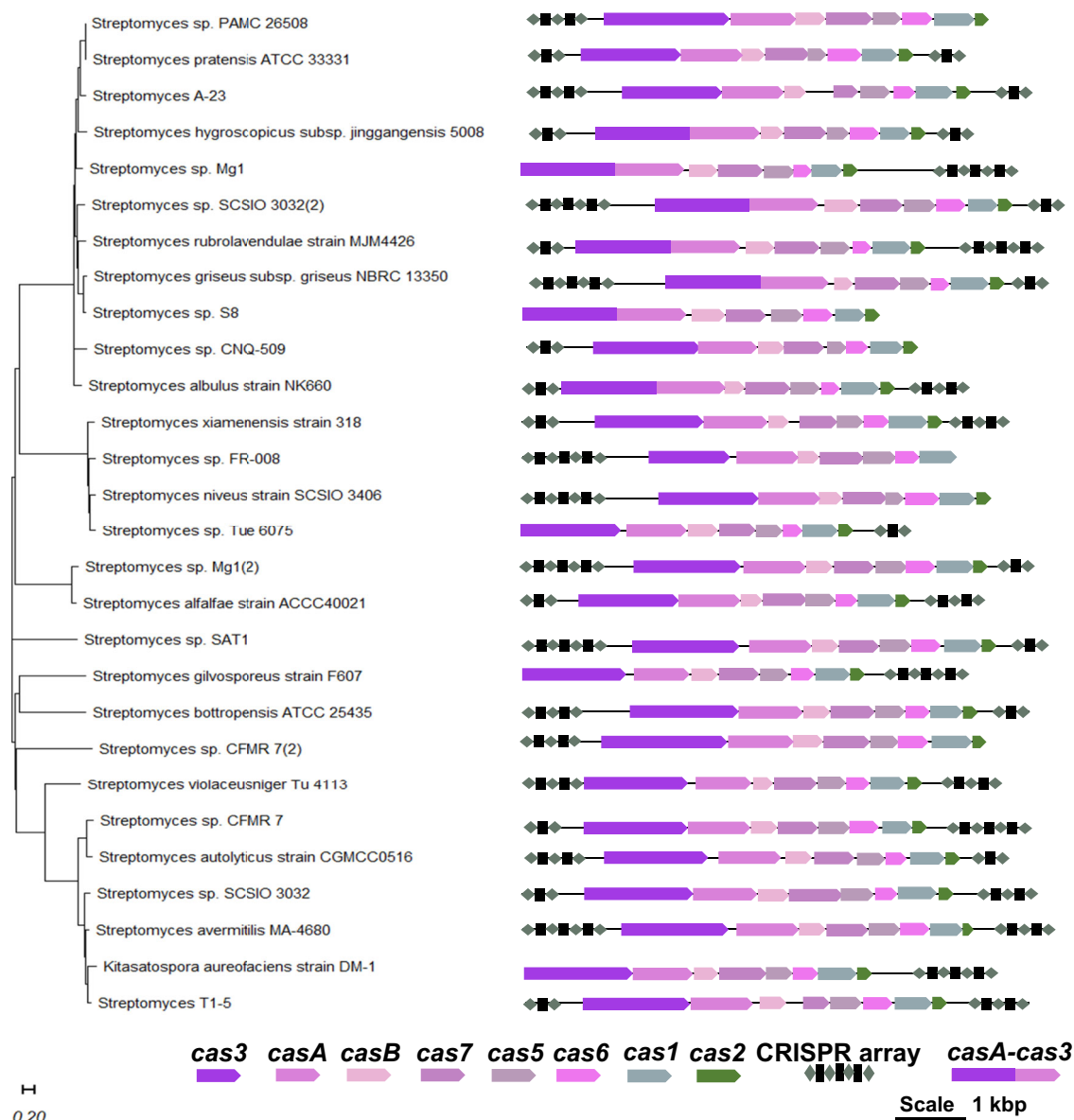


Fig. 1 | Type I-E CRISPR-Cas systems in *Streptomyces*. The phylogenetic tree was built based on the amino acid sequences of Cas1, the most conserved protein in the CRISPR systems. The displayed 28 type I-E CRISPR systems were found from 74 *Streptomyces* strains.

ACT, while spacer r3a4 (r3 and a4) efficiently repressed RED formation, but not as obviously for ACT accumulation (Fig. 2f). The mRNA expression levels of *actII-ORF4* and *red Q* were significantly decreased (Supplementary Fig. 2b).

Creating the CRISPR-based transcriptional activator 1.0 (CTA 1.0)

In pursuit of advancing gene regulatory tools that are particularly amenable to the exploration of natural products in *Streptomyces*, we proceeded to develop the transcriptional activator tool CTA 1.0, building upon the foundation of CTR 1.0 with activation domains fused. Our aim was to exemplify its efficacy in reactivating silenced bacterial gene clusters (BGCs), thereby enhancing the potential for uncovering bioactive compounds. An activation domain fused to the CRISPR system can be directed to the promoter region and activates gene transcription, by recruiting RNA polymerase (RNAP) or stabilizing the RNAP initiation complex³⁵. Two different activation domains were selected: the ω (*rpoZ*) domain, a component of RNA polymerase crucial for antibiotic production in *Streptomyces kasugaensis*³⁶, and the SoxS, a transcriptional regulator belonging to the AraC family and a

component of the essential bacterial *soxRS* redox-sensing system, which is notably upregulated in response to oxidative stress and recruits RNA polymerase to the promoter region³⁷. Both activation domains have been demonstrated effective in *Escherichia coli*³⁸. The crystal structure of the Cascade reveals that CasA caps one end of the helical arrangement³⁹. Therefore, a construct with the CasA N-terminus fused to an activation domain via a linker (two or five small amino acid residues, AA or GGGGA) was designed, in which the Cascade was derived from *Streptomyces avermitilis* MA-4680 (Fig. 3a). The Cascade-activator expression cassette with the *kasO**p promoter and *oop* terminator pairs was integrated into the $\Phi C31$ *attB* site of *Streptomyces* A-14, a non-model *Streptomyces* isolated from the inter-root soil of the traditional Chinese medicine *Ferula officinalis* with rich BGCs and fast growing property (Supplementary Data 3), to form strains A1-1 to A1-4. Subsequently, the pYL-crRNA plasmid was conjugated into the four A1 strains, generating the CRISPR-based transcriptional activator 1.0 (CTA 1.0).

Analysis of the *Streptomyces* A-14 genome sequence by anti-SMASH indicates that cluster 8 is responsible for the biosynthesis of polycyclic tetramate macrolactams (PTMs) that exhibit antitumor

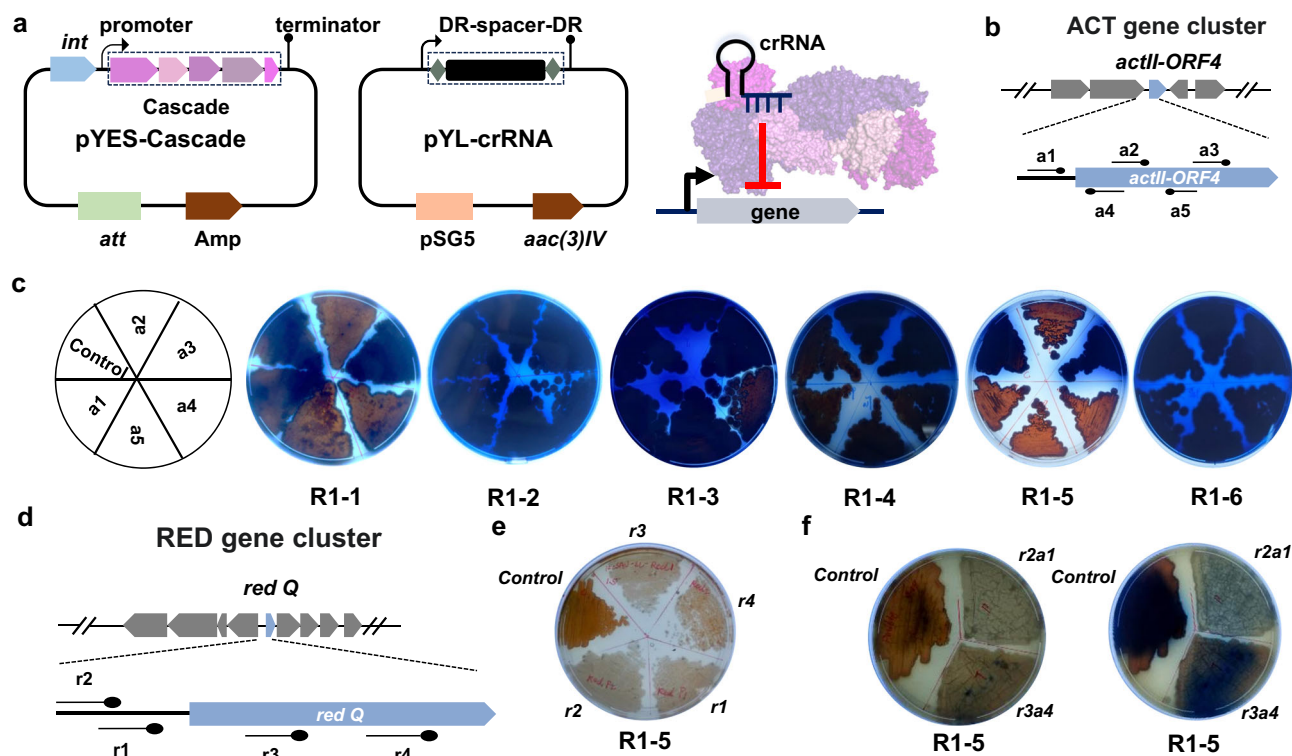


Fig. 2 | Development and characterization of CTR 1.0 in *S. coelicolor*.

a Schematics of CRISPR-based transcriptional repressor 1.0 (CTR 1.0) in this work. The Cascade consists of 5 genes (*casA-casB-cas7-cas5-cas6*). **b** CrRNAs targeting the *act-orf4* coding region and promoter region of the ACT BGC are shown.

c Repression results for targeting the *act-orf4* coding sequence using the CTR 1.0.

d CrRNAs targeting the *red Q* gene region of the RED BGC are shown. **e** Repression results for targeting the *redQ* coding sequence using the CTR 1.0. **f** Repression results for targeting both the *redQ* and *act-orf4* genes using the CTR 1.0. Control (The wild-type strain of *S. coelicolor*).

activities⁴⁰. All six coded gene products are enzymes involved in the biosynthetic pathway and show 100% similarity to previously characterized *S. griseus* PTM BGC. We designed three crRNAs (p1–p3) located at 80–300 bp upstream of the transcription start site (Fig. 3b). All three crRNAs appeared to activate the PTM BGC, as the HPLC profiles of the CTA 1.0 fermentation extracts displayed prominent peaks in comparison with no observable peaks in that of the wild type strain (Fig. 3c, S3). The best hit, A1-3-p3, with Cascade under *kasO**p promoter and *oop* terminator and CasA linked to the SoxS domain via AA, was subjected to further LC/MS analysis. The molecular weights for the products of peak **a** (*m/z*, 510.2730), peak **b** (*m/z*, 508.2430), and peak **c** (*m/z*, 494.2584) were compared and confirmed to be identical to PTMs derived from *S. griseus* (Fig. 3d, S4). A comparative analysis of the linker revealed that AA is superior to GGGGA. Moreover, the activation domain SoxS outperformed ω (*rpoZ*) (Supplementary Fig. 3), which could be attributed to its interaction with the conserved σ -factor region 4 and core RNAP α -subunit⁴¹, and was used for subsequent gene cluster activation.

Development of the CRISPR-based transcriptional repressor/activator 2.0 (CTR/CTA 2.0)

To craft more efficient and versatile gene regulation tools capable of addressing a spectrum of application needs, including the concurrent repression and activation of genes, we have subsequently developed the transcriptional regulation tool CTR/CTA 2.0 employing Scaffold RNAs. Scaffold RNAs (each scRNA is a crRNA followed by the MS2 sequence of a bacteriophage origin) have been used in *E. coli* as gene regulatory tools in concert with the employment of an activation domain fused to the MS2 capsid protein (MCP) that specifically recognizes the MS2 sequence⁴². This design can be used as a transcriptional activator or repressor depending on the positioning of

scRNA. When the scRNA binds to a narrow region upstream of the transcription start site, gene activation occurs. By contrast, if the scRNA binds to the coding sequence, transcriptional repression of the gene occurs³⁵.

We decided to adopt the scRNA method for programmable transcription regulation in *Streptomyces*. The pYL-scRNA plasmid was designed with MS2 at the 3' end of the crRNA and MCP protein fused to the C-terminus of the ω (*rpoZ*) or the SoxS domain via GGGGG linker. The Cascade expression box with the *kasO**p promoter and *oop* terminator (Fig. 4a) was conjugated into *S. coelicolor*, *S. lividans* and *S. A-14*, respectively to form different *Streptomyces* chassis, followed by the introduction of various pYL-scRNA plasmids, resulting in the CRISPR based transcriptional repressor/activator 2.0 (CTR/CTA 2.0) (Fig. 4a and Supplementary Table 1).

To validate CTR 2.0 in *S. coelicolor*, we chose the same reporter genes, *actII-ORF4* and *red Q* involved in the biosynthesis of ACT and RED respectively, and the corresponding spacers used in CTR 1.0, excluding those targeting sequences within the promoter region. Spacers a2, a4, and a5 were found able to abolish or significantly reduce the pigment accumulation likely due to lack of ACT biosynthesis and to reduce the mRNA expression levels of *actII-ORF4* transcripts by 96%, 80% and 93%, respectively, compared to control strain, while spacer a3 showed no significant repression effect (Fig. 4b, S2c). Among the three effective spacers, a4 was the least effective in gene repression, which perhaps is due to its proximity to the promoter region. Spacers r3 and r4 reduced the *RedQ* transcript by 90% and 75%, respectively, resulting in the eyeball-distinguishable decrease of the RED pigmentation (Fig. 4c, S2d). The CTA 2.0 system was then tested in *S. lividans* that harbors the ACT gene cluster with the *actII-ORF4* silent under normal culturing conditions⁴³. The CTA2.0 system employing either the ω (*rpoZ*) or the SoxS domain and the crRNA

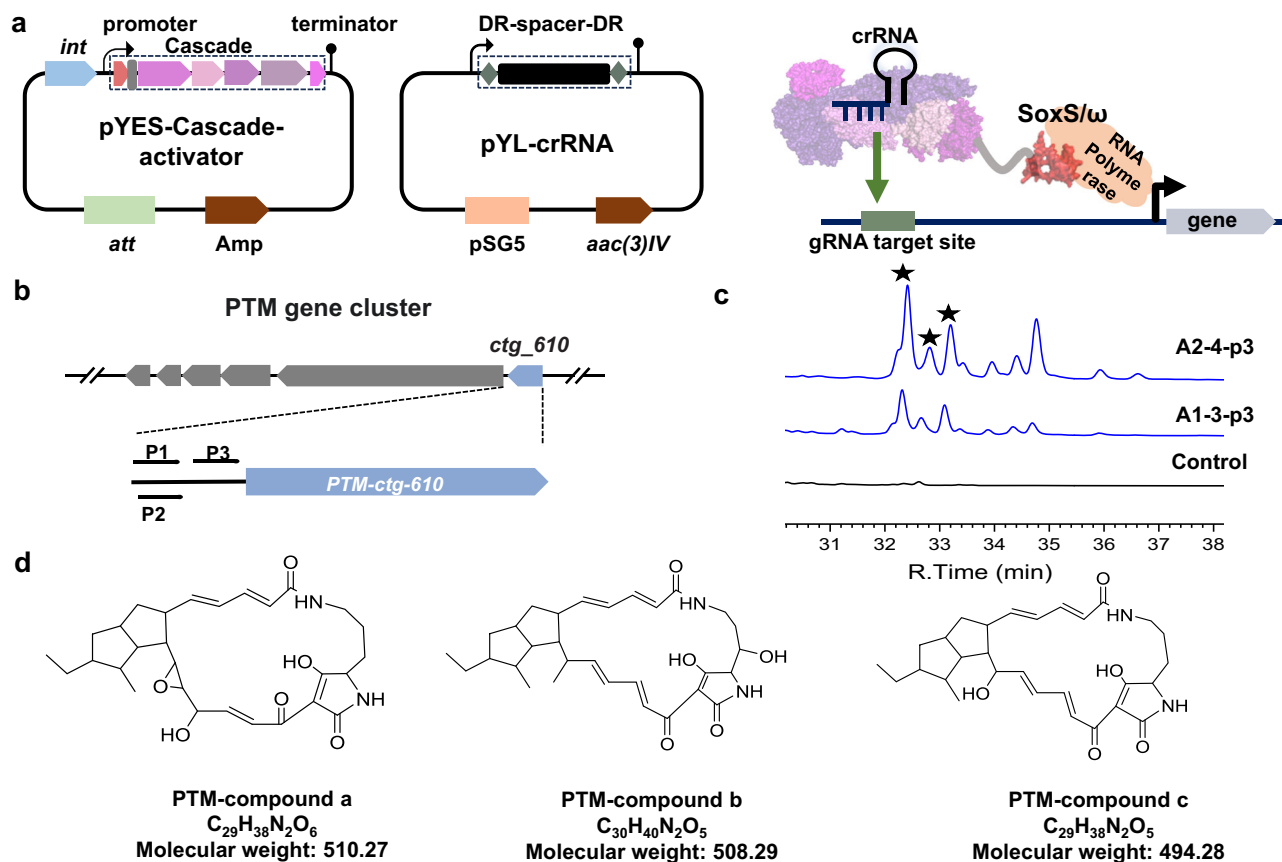


Fig. 3 | Development and characterization of CTA 1.0 in S. A-14. **a** Schematics of CRISPR-based transcriptional activator 1.0 (CTA 1.0) in this work. The activation domain was fused to the Cascade (*casA-casB-cas7-cas5-cas6*) via a linker at the *CasA* N-terminus. **b** CrRNAs targeting the *PTM-ctg-610* promoter region of the PTM BGC

are shown. **c** HPLC analysis for activating PTM BGC using the CTA 1.0 and CTA 2.0 in S. A-14. (★ indicates major products of the PTM BGC). **d** The structures of the compound PTM-a - c. Control (The wild-type strain of S. A-14). Source data are provided as a Source Data file.

targeting 86 bp upstream of the transcription start site of the *actII-ORF4* gene successfully triggered the pigment accumulation of ACT (Fig. 4d).

The BGC of PTM in S. A-14 was found activatable with the CTA 2.0 system. Comparing the two CTA versions in activating the PTM cluster highlighted the efficacy of the 2.0 system. We thus decided to utilize the CTA 2.0 system to activate uncharacterized BGCs in various *Streptomyces* strains.

Activation of cryptic BGCs by the CTA2.0 in S. A-14

31 BGCs in the S. A-14 genome have never been characterized. They are putatively responsible for the biosynthesis of various natural products, including polyketides, NRPs, Ripples, and others. In addition to the PTM gene cluster (C8), 10 of the 31 BGCs were chosen as representatives for the validity test of CTA2.0 with expected natural products ranging from polyketide to Ripp and other uncategorized (Table 1). Among them, cluster 11, expected to produce a polyketide is of particular interest with sequences distinct from those of any cluster responsible for a previously characterized product. Three crRNAs (c1-c3) were designed to target within the region 80–300 bp upstream of the transcription start site of the first polyketides gene, *ctg-949* that encodes the acyl carrier protein (Fig. 5a). LC-MS analysis revealed a prominent peak with a *m/z* of 638.1717 that is absent in both the wild type strain and the activated cluster-deletion strain, Δ C11 (Fig. 5b, S5). The chemical structure of the compound was delineated by nuclear magnetic resonance (NMR) spectroscopy with a scale-up sample from approximately 200 L fermentation broth (0.83 mg/L) (Fig. 5c, Supplementary Fig. 6 and Supplementary Table 2).

Based on the availability of the PAM region in the gap areas and an optimal 80–300 bp distance from the transcription start site of genes, additional spacers were designed to perform in situ activation of other genes in cluster 11. Fermentation followed by HPLC assays of these mutant strains suggested that a gamma-glutamyltransferase, a phosphatase of the PhoX family, another acyl carrier protein and a flavo-protein of the FixA family, encoded by *ctg1_937*, *ctg1_938*, *ctg1_954* and *ctg1_972* respectively may be involved in the biosynthesis (Supplementary Fig. 7 and Supplementary Table 3). Interestingly, a gene cluster responsible for the biosynthesis of secalonic acid and structurally related compounds has been recently identified in the fungi and biochemically characterized^{44,45} (Supplementary Fig. 8 and Supplementary Tables 4, 5). The composition and organization of the S. A-14 gene cluster 11 suggested that this bacterium employs a synthetic route and mechanism distinct from the fungi.

Genomic sequencing and further analysis via PRISM and BAGEL4 on cluster 14 unveiled a linaridin BGC with only 33% similarity to that responsible for the synthesis of legonaridin⁴⁶. It is noteworthy that the precursor peptide is indeed different from legonaridin as a member of the rare linaridin class. Three crRNAs (c4-c6) located at 80–300 bp upstream of the transcription start site of the first gene in the cluster, *C14-ctg-1306* were designed (Fig. 6a). In the HPLC elution profile, a prominent peak was observed and absent in the negative control (Fig. 6b). The MALDI-TOF analysis of supernatant extracts revealed a mass of $[M + Na]^+ = 2197.9540$ Da, consistent with the high-resolution electrospray ionization mass spectrometry (MS), which detected a mass of $[M + H]^+ = 2175.1531$ Da (Supplementary Fig. 9). The discrepancy between the observed mass and the predicted sequence (Fig. 6c) suggested additional modifications.

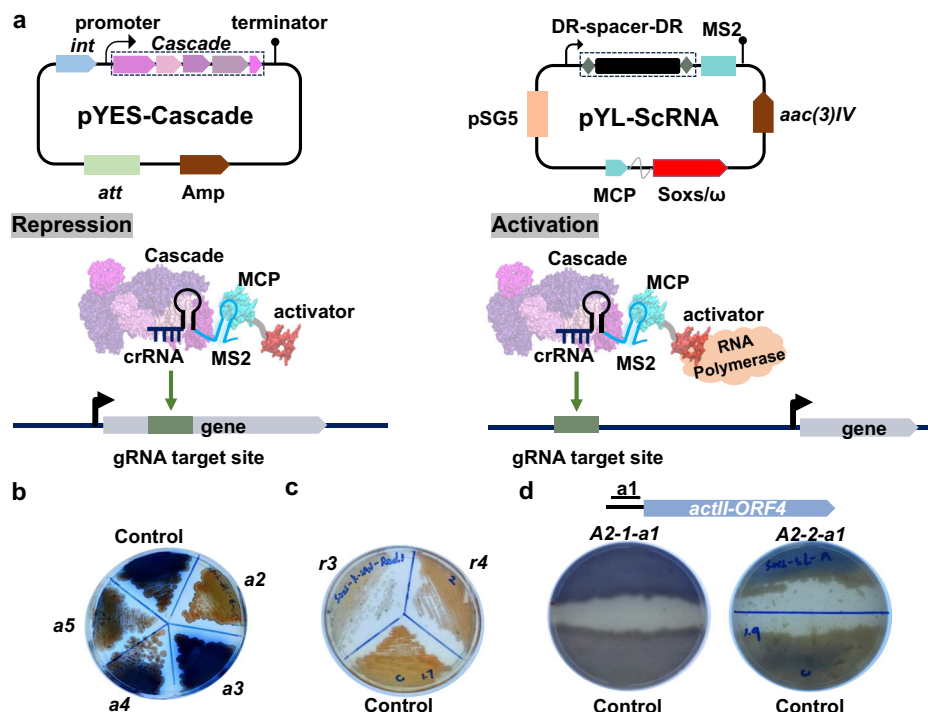


Fig. 4 | Development and characterization of CTR/CTA 2.0. **a** Schematics of CRISPR-based transcriptional repressor/activator 2.0 (CTR/CTA 2.0) in this work. The Cascade consists of 5 genes (*casA-casB-cas7-cas5-cas6*). The scRNA plasmid consists of a scRNA cassette with crRNA coupled to the MS2 sequence, and an activator domain under the control of the *kasO* promoter. **b** Repression results

for targeting the *act-orf4* coding sequence using the CTR 2.0. **c** Repression results for targeting the *redQ* coding sequence using the CTR 2.0. **d** Activation result for targeting the *act-orf4* promoter region using the CTA 2.0. Control (For 4b and 4c, control represents the wild-type strain of *S. coelicolor*; For 4d, control represents the wild-type strain of *S. lividans*).

Strain A2-4 was employed to activate eight unexplored cryptic BGCs, clusters 4, 13, 17, 22, 24, 25, 26, and 28, some of which are highly homologous to previously known BGCs^{47–50}. Six of them, clusters 4, 13, 22, 24, 25, and 28, when activated, resulted in notable peaks in the fermentation extracts (Supplementary Fig. 10a), with some, clusters 4, 22, 24, and 28 finding the expected molecular weights of the natural products by LC-MS (Supplementary Fig. 10b–d).

Activation of multiple gene clusters by CTA2.0 in diverse *Streptomyces*

To demonstrate the general applicability of the CTA2.0 tool to a broad range of *Streptomyces* strains, we attempted to activate nine uncharacterized BGCs of different classes, including four polyketides, two NRPs, one Ripp, and two other uncategorized BGCs. Seven *Streptomyces* strains, namely *S. 539*, *S. T2-10*, *S. T2-4*, *S. A16*, *S. W121*, *S. albus* J1704, and *Nh. DSM 44494* was subjected to the assay. One or more crRNAs located 80–300 bp upstream of the promoter transcription start site of the core gene were designed for each BGC. For six out of the total of nine gene clusters (Supplementary Tables 6–14), C6.1-*S. 539*, C12-*S. A16*, C21-*S. W121*, C24-*S. W121*, C22-*S. albus* J1704 and C11-*Nh. DSM 44494*, we observed remarkable peaks in the fermentation extracts (Supplementary Figs. 11–14a), with most of them, except C21-*S. W121*, matching exactly with the expected molecular weights by MS (Supplementary Figs. 11–14b).

Cluster 11 in *Nocardiopsis halophila* DSM 44494, isolated from extremophilic environments, exhibits a 52% similarity to the pyrrolizidine alkaloids (PAs) BGC (pyrrolizixenamide A) in *Xenorhabdus*⁵¹, was readily activated while the promoter region of the core gene *ctg-16.96* was targeted with CTA2.0 (Fig. 7a). HPLC analysis revealed that the engineered strain produced a prominent peak and a couple of minor peaks that were not present in the fermentation broth of the control strain (Fig. 7b). High-resolution electrospray ionization mass spectrometry detected masses of $[M + H]^+ = 181.0961$, 197.0896 and

320.1593 Da (Supplementary Fig. 15), consistent with the structures of legonmycin C, D and E (0.02 mg/L, 0.04 mg/L and 0.09 mg/L, respectively) as determined by NMR (Fig. 7c, Supplementary Figs. 16–18 and Supplementary Tables 15–17). Taken together, our CTA tools are well suited for in-situ activation of individual genes and/or large BGCs in the model, non-model, and extremophilic *Streptomyces*, and are expected to facilitate the exploration of hidden natural products from this rich source.

Discussion

Streptomyces are ubiquitous Gram-positive bacteria and are known as a treasure trove of structurally diverse natural products, many of which have the potential to be developed into therapeutic compounds. CRISPR systems have been used as valuable tools for genetic manipulation in model *Streptomyces* strains. However, when they are applied to most of the non-model strains, such exogenous systems face limitations due to cellular toxicity and the intricate metabolic milieu of these bacteria. Given the prevalence of type I-E CRISPR systems in *Streptomyces*, our study focused on the development of gene regulatory tools employing such endogenous systems. These endogenous CRISPR systems possess notable advantages over prevalent CRISPR systems like Cas9 or Cas12a, such as enhanced compatibility, extensive distribution, and operational simplicity, given that there is no requirement for the delivery of foreign Cas genes^{52,53}. To ascertain the relative efficacy of the dCas9 and the I-E CRISPR-derived Cascade system, we also assessed the proteotoxicity of Cas9, dCas9 and Cascade within various *Streptomyces* strains (Supplementary Fig. 19), under identical conditions, the conjugative transfer efficiencies followed the order: Cas9 < dCas9 < Cascade. This observation implies that the Cascade system, which originates from *Streptomyces*, may be better adapted for applications within *Streptomyces* species.

Upon systematic mining and comparative analysis, we identified and developed effective gene regulatory tools derived from the type

Table 1 | Summary of the 21 gene clusters activated by the CTA1.0/2.0

Target BGC	Similarity	Predicted BGC type	Construction	Effect
<i>S. lividans</i>	–	–	–	–
C14	100%	T2PKS	Succeed	Activated
<i>S. A-14</i>	–	–	–	–
C4	11%	T3PKS	Succeed	Activated
C8	100%	NRP + PKS	Succeed	Activated
C11	21%	PKS	Succeed	Activated
C13	15%	Class I lanthipeptide	Succeed	Not activated
C14	9%	Linaridin	Succeed	Activated
C17	100%	Lanthipeptide	Failed	–
C22	100%	Class II lasso peptide	Succeed	Activated
C24	80%	NI-siderophore	Succeed	Activated
C25	–	Class II lanthipeptide	Succeed	Not activated
C26	100%	Ectoine	Failed	–
C28	11%	T3PKS	Succeed	Activated
<i>S. 539</i>	–	–	–	–
C6.1	100%	Other	Succeed	Activated
C24	100%	Other	Failed	–
<i>S. T2-10</i>	–	–	–	–
C1.8	96%	T2PKS	Succeed	Not activated
<i>S. T2-4</i>	–	–	–	–
C27	100%	T3PKS	Succeed	Not activated
<i>S. A16</i>	–	–	–	–
C12	80%	NRP + PKS	Succeed	Activated
<i>S. W121</i>	–	–	–	–
C21.1	80%	Lanthipeptide	Succeed	Not activated
C24.1	78%	NRP	Succeed	Activated
<i>S. albus J1704</i>	–	–	–	–
C22	100%	T1PKS	Succeed	Activated
<i>Nh. DSM 44494</i>	–	–	–	–
C11	100%	NRPS	Succeed	Activated

I-E CRISPR system from *Streptomyces avermitilis* MA-4680. These tools were further demonstrated to be effective for the activation of various BGCs in model and non-model strains, including extremophilic *Streptomyces* strains. *Streptomyces* adapt to survival challenges posed by extreme environmental factors by producing distinct profiles of natural products. Our methods advance in overcoming the constraints associated with traditional screening techniques and opening avenues for the discovery of previously untapped bioactive molecules in extreme environments. Despite the fact that the CTR and CTA systems developed in this study utilize two distinct plasmids, the integration of the Cascade system simplifies subsequent gene editing to merely the introduction of the crRNA expression plasmid, thereby enhancing the efficiency of the conjugation process. This refinement allows for a more streamlined approach to genetic manipulation, reducing the complexity and increasing the potential for successful gene editing outcomes. It is widely recognized that a single *Streptomyces* strain typically harbors multiple natural product biosynthetic pathways. Therefore, with the Cascade system integrated just once, subsequent editing of various pathways becomes more manageable and efficient, facilitated by the use of a smaller plasmid. This approach actually

simplifies the genetic manipulation process, making it an advantageous strategy for *Streptomyces* gene regulation.

Activation of Cluster 11 in *Streptomyces* A-14 with our tools leads to the production of secalonic acid, a tetrahydroxanthone dimer, by a bacterial strain. Tetrahydroxanthone dimers were previously isolated from fungi and shown to inhibit the growth of *Helicobacter pylori*, over 30-fold more potent than clinically used metronidazole^{44,45}. Bioinformatics analysis revealed this gene cluster with the core polyketide biosynthesis gene exhibiting only 24% identity to the fungal counterpart⁴⁴, implying a different biosynthetic route and mechanism to be further characterized.

The tools reported in this study are modular, allowing substitution of various components for different purposes, e.g., introducing the gain of function SoxS mutants, R93A and S101A, to boost the activation effect⁵⁰. The increasing number of subclasses within class I CRISPR systems and their wide distribution in different bacterial phyla⁴ may allow quick expansion and advance of the toolbox by adopting new components to the systems.

The distance of the spacer binding site in relation to the promoter is critical for the activation effect, as previously observed for the application of Cas9 in *E. coli*³⁵. Noteworthy, our systems appeared to be less restricted and effective in a broader range of 80–300 bases upstream of the gene transcription start site, compared to 50–90 bases for Cas9 in *E. coli*³⁵. The requirement of specific spatial orientation of gRNA binding was observed for the Cas9 system, working effectively with gRNA binding to the non-template strand in *Streptomyces*²³, whereas our system demonstrated better flexibility, being effective with gRNA binding to both template and non-template strands, perhaps due to a larger complex and a longer interdomain linker employed.

In summary, we demonstrated the versatility of the *Streptomyces* type I-E CRISPR system in gene regulation, encompassing both repression and activation. Robust gene regulation systems for *Streptomyces* successfully constructed in this study led to the discovery of various natural products. Our work rationalizes strategies utilizing endogenous CRISPR systems and promises to catalyze further development of similar strategies in other prokaryotes for various purposes, including but not restricted to the mining of natural products.

Methods

Strains, materials, and reagents

Streptomyces coelicolor A3(2), *Streptomyces lividans* 1326, *E. coli* DH5 α and *E. coli* ET12567/pUZ8002 were gifted from Professor Huimin Zhao (University of Illinois, Urbana, IL). The plasmid pYES and pCRISPOmyces-2 were gifted from Professor Huimin Zhao (University of Illinois, Urbana, IL). The plasmid pWHU2653 was gifted from Professor Yuhui Sun (SIBS, Chinese Academy of Sciences). *S. A-14* and *Nh. DSM44494* were excavated by our laboratory. Reagents and solvents were obtained from Sigma-Aldrich (St. Louis, MO, United States), Oxoid (Basingstoke, Hampshire, United Kingdom), and Sangon Biotech Co., Ltd. (Shanghai, China). All restriction endonucleases, as well as T4 DNA ligase and DNA polymerase, were purchased from New England Biolabs (Beverly, MA, United States). SYBR[®] Green PCR Master Mix was purchased from Applied Biosystems (San Francisco, CA, United States). The DNA isolation and purification kits were purchased from TIANGEN (Beijing, China). All primers were synthesized by Tsingke (Beijing, China). *E. coli* strains were grown in Luria Bertani (LB) broth, containing 100 μ g/mL ampicillin or 50 μ g/mL apramycin as needed. *Streptomyces* strains were grown in MYG liquid medium (10 g/L malt extract, 4 g/L yeast extract, and 4 g/L glucose), or MS agar medium (20 g/L D-mannitol, 20 g/L soybean meal, and 20 g/L agar) at 30 °C unless otherwise indicated. 2 \times YT medium-containing 1% yeast extract, 1% tryptone, and 0.5% NaCl (pH 7.0) was used for strain washing and spore germination before conjugation. M-ISP4 (5 g/L Soybean meal, 5 g/L D-Mannitol, 5 g/L Starch soluble, 2 g/L Tryptone,

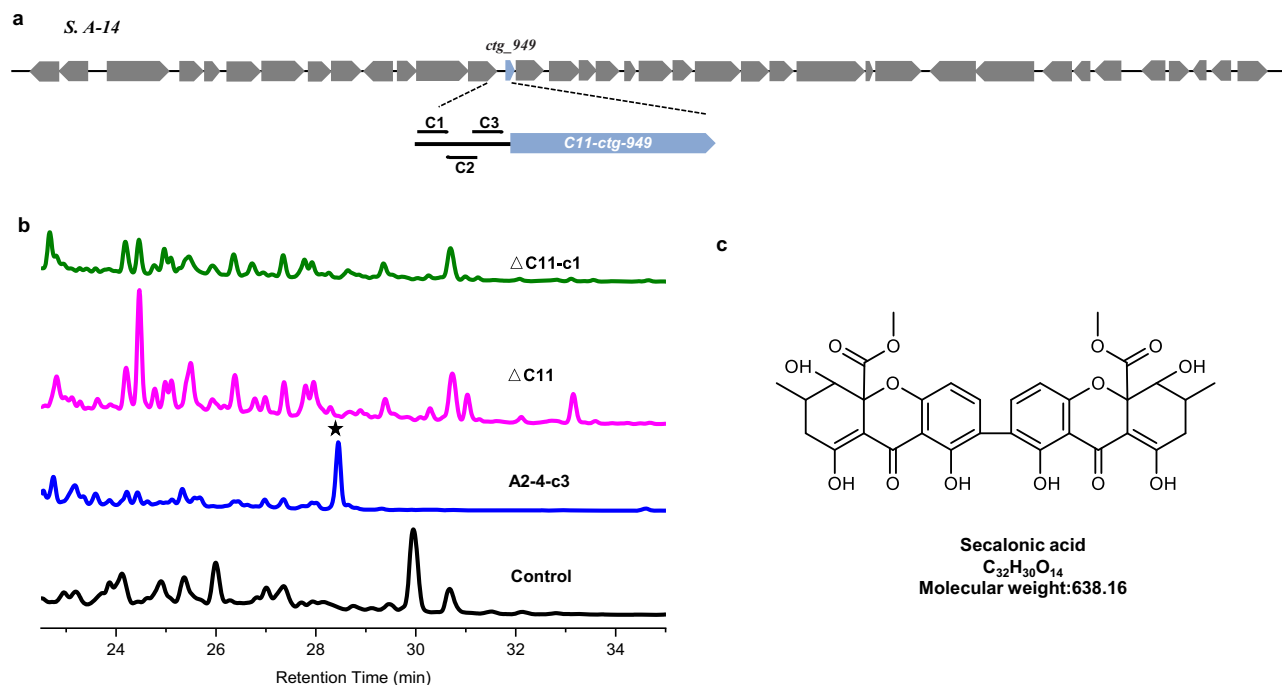


Fig. 5 | Activation of C11 biosynthetic gene cluster in *Streptomyces A-14*.

a CrRNAs targeting the promoter regions of *ctg-949* gene are shown. **b** HPLC analysis for activating C11 BGC using the CTA 2.0 in *S. A-14*. **c** The structure of the compound secalonic acid. A2-4 (Activation strains with *kasO*⁺ driven the expression of the Sav Cascade and MS2-coupled SoxS domain); ΔC11 (Strains with cluster

11 deleted in the chromosome of *S. A-14*, as a control); ΔC11-c1 (The cluster-deletion strain with activation tool applied to target the promoter regions of *ctg-949* gene, as a control); Control (The wild type strain of *S. A-14*). Source data are provided as a Source Data file.

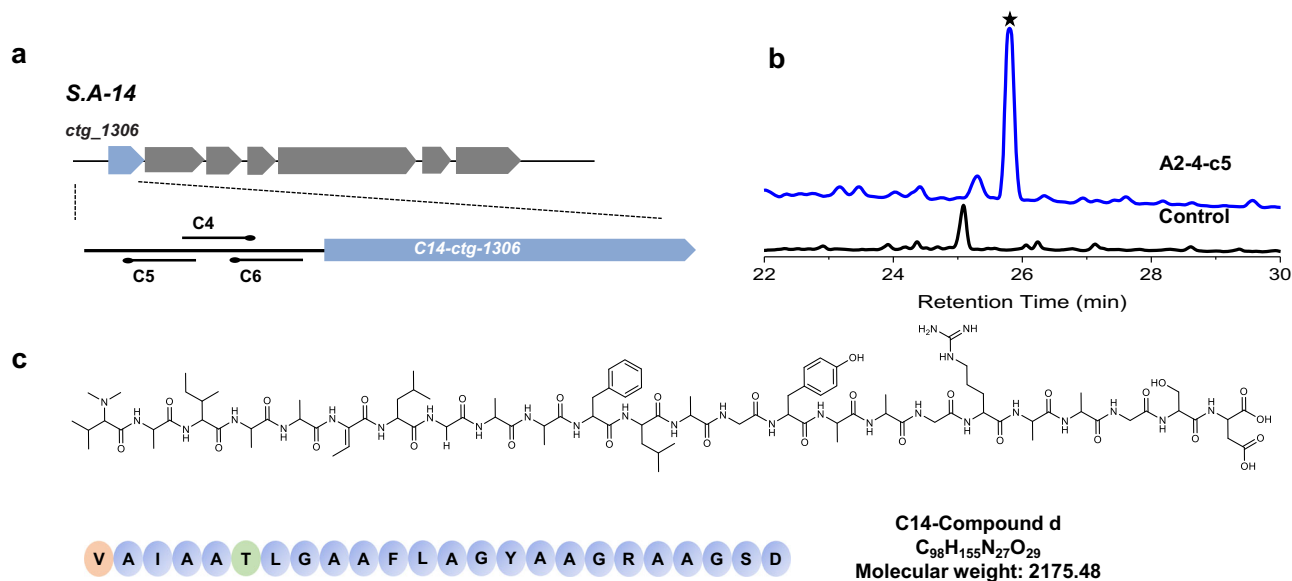


Fig. 6 | Activation of C14 biosynthetic gene cluster in *Streptomyces A-14*.

a CrRNAs targeting the promoter region of the *C14-ctg-1306* gene are shown. **b** HPLC analysis for activating the C14 BGC using the CTA 2.0 in *S. A-14*. **c** The

structure of the compound C14-d. A2-4 (Activation strains with *kasO*⁺ driven the expression of the Sav Cascade and MS2-coupled SoxS domain); Control (The wild type strain of *S. A-14*). Source data are provided as a Source Data file.

1 g/L Yeast extract, 1 g/L NaCl, 2 g/L $(NH_4)_2SO_4$, 1 g/L K_2HPO_4 , 2 g/L $CaCO_3$, 20 g/L Agar, 0.1% Trace element which contain 1 g/L $FeSO_4$, $MnCl_2$ and $ZnSO_4$) was used for conjugation. R2YE agar medium (103 g/L Sucrose, 0.25 g/L K_2SO_4 , 10.12 g/L $MgCl_2 \cdot 6H_2O$, 10 g/L Glucose, 0.1 g/L Casamino acid, 5 g/L yeast extract, 0.5% KH_2PO_4 , 5 M $CaCl_2 \cdot 2H_2O$, 20% L-proline, 5.73% TES buffer, 0.2 ml Trace element solution, and 20 g/L agar) was used for phenotypic analysis.

Plasmid construction and conjugation

The Cascade or Cascade-activator expression cassette was cloned into the vector pYES⁵⁴ and assembled via Gibson assembly⁵⁵. The pYL-crRNA plasmid and pYL-scRNA plasmid with MS2 at the 3' end of the crRNA and MCP protein fused to the C-terminus of the activation domain were constructed based on pYL vector⁹ by Golden Gate assembly.

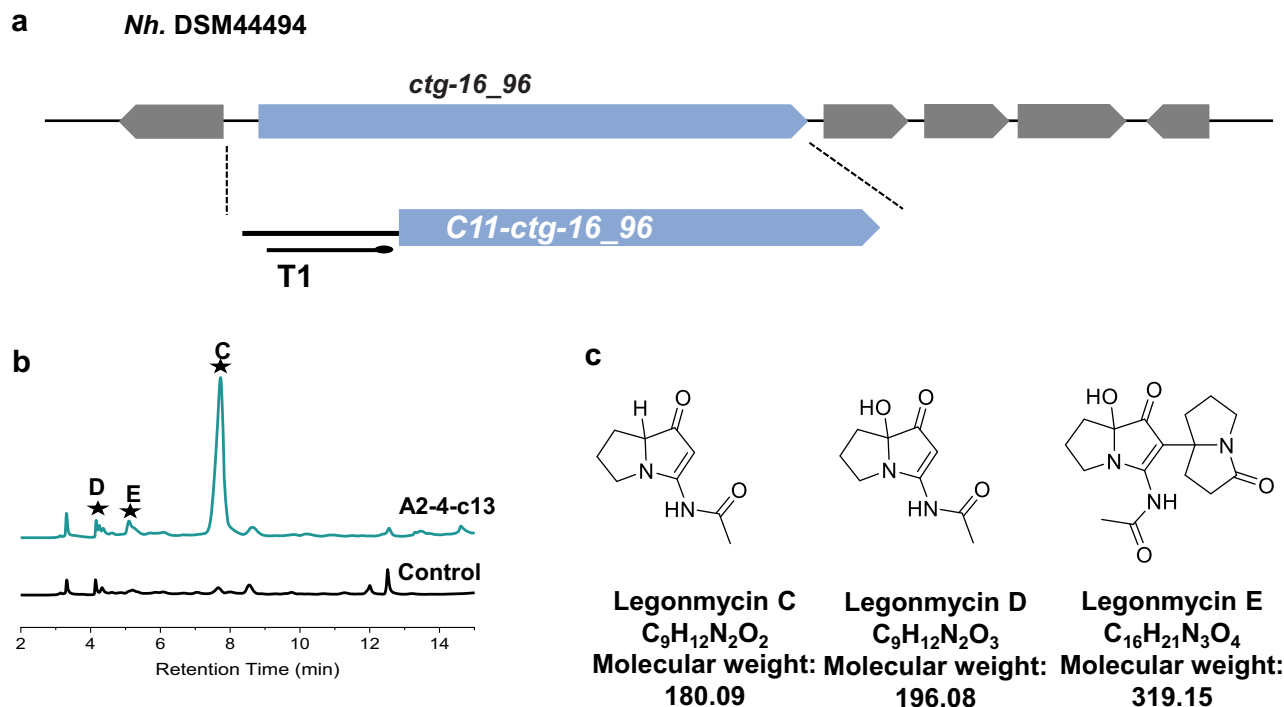


Fig. 7 | Activation of C11 biosynthetic gene cluster in *Nh.* DSM 44494. **a** CrRNA targeting the promoter region of the *C11-ctg1-16_96* gene are shown. **b** HPLC analysis for activating C11 BGC using the CTA 2.0. **c** The structure of the compound C11-

Legonmycin C-E. A2-4 (Activation strains with *kasO*⁺ driven the expression of the Sav Cascade and MS2-coupled SoxS domain); Control (The wild type strain of *Nh.* DSM 44494). Source data are provided as a Source Data file.

Streptomyces codon-optimized MCP was chemically synthesized by Genewiz (Suzhou, Jiangsu, China), and activation domains, including ω and SoxS were amplified from purified genomic DNA of *S. griseus*. The presence of the correct plasmids was confirmed by sequencing (Tsingke, Beijing, China). All primers used in this study are listed in Supplementary Data 4. *E. coli* ET12567/pUZ8002 cells were transformed with the plasmids to be conjugated and selected on LB agar plates supplemented with 50 μ g/mL apramycin, 50 μ g/mL kanamycin and 25 μ g/mL chloramphenicol. These transformants were then used as the donors for conjugative transfer of the assembled plasmids to *Streptomyces* strains. M-ISP4 solid medium containing 25 mM $MgCl_2$ and 2 \times YT liquid medium were used for conjugation. Exconjugants were transferred to M-ISP4 medium containing 25 μ g/mL apramycin and 25 μ g/mL nalidixic acid and were cultured at 30 $^{\circ}C$ ³⁶. *Streptomyces* exconjugants were picked and re-streaked on ISP2 plates supplemented with 50 μ g/mL apramycin and grown for two days.

qPCR analysis

The total RNA of *Streptomyces* cultivations under MYG medium after 72 h was extracted using the BACTERIAL Total RNA Isolation Mini Kit (Bio-Tek, USA). Reverse transcription was carried out using the First-strand cDNA Synthesis Kit (Bio-Rad, Carlsbad, CA, USA). Real-time PCR was performed with SYBR[®] Green PCR Master Mix on the 7900HT Fast Real-Time PCR System (Applied Biosystems, Carlsbad, CA, USA). Primers were designed by the online tool provided by Integrated DNA Technologies (<https://www.idtdna.com/scitools/Applications/RealTimePCR/>). 10 μ L of 2 \times SYBR Green Mix, 1 μ L of cDNA, 1 μ L of each primer at a concentration of 10 pmol/ μ L, and 7 μ L of ddH₂O were mixed gently in each well of the Applied Biosystems[®] MicroAmp[®] Optical 96-Well Reaction Plate. Reactions were performed using the following program: 2 min at 50 $^{\circ}C$, 10 min at 95 $^{\circ}C$ for one cycle, followed by 15 s at 95 $^{\circ}C$, 30 s at 60 $^{\circ}C$ and 30 s at 72 $^{\circ}C$ for 40 cycles with a final cycle of 10 min at 72 $^{\circ}C$. The endogenous gene *hrdB*, encoding RNA polymerase sigma factor, was used as the internal control for gene regulation. The expression levels of

other candidate genes were normalized to the expression of the control. Data were analyzed using SDS2.4 software (Applied Biosystems, Carlsbad, CA, USA).

Fermentation and HPLC-MS analysis

Exconjugants were picked and inoculated in 50 ml of MYG supplemented with apramycin and grown to high density (usually 2-3 days) in 250 ml shake flasks at 30 $^{\circ}C$. For the fermentation of PTM compounds, aliquots of the seed culture were spread for confluence over Petri plates containing approximately 25–30 ml of MYG solid agar medium, and the plates were incubated for 7 days at 30 $^{\circ}C$ ⁴⁰. For other compounds, fermentation cultures were incubated at 30 $^{\circ}C$ with constant shaking (220 rpm) for seven days, with a ratio of 1:25. To isolate PTM compounds, agar, and the adherent cells were stored at 80 $^{\circ}C$ for 24 h, after which the frozen cultures were thawed completely and the released liquids were collected. The liquids were extracted by ethyl acetate with a ratio of 1:1 twice, concentrated 1000-fold, and subjected to HPLC analysis. To isolate potential Ripps compounds, the supernatant of the fermentation was extracted by n-butanol with a ratio of 1:1 twice, concentrated 1000-fold, and subjected to HPLC analysis. To isolate potential other compounds, the supernatant of the fermentation was extracted by ethyl acetate with a ratio of 1:1 twice, concentrated 1000-fold, and subjected to HPLC analysis.

HPLC was performed on an HPLC system (Shimadzu, LC-20a) with a Shimadzu C18 reverse-phase column (4.6 \times 250 mm, 5 μ m) at room temperature. The HPLC parameters were as follows: solvent A, 0.1% trifluoroacetate in water; solvent B, 0.1% trifluoroacetate in acetonitrile; gradient, 10% B for 10 min, to 100% B in 30 min, maintain at 100% B for 10 min, return to 10% B in 1 min and finally maintain at 10% B for 20 min; flow rate 1 mL/min; detection by UV spectroscopy at 280/215 nm. The mass spectrometry system was operated using a drying temperature of 350 $^{\circ}C$, a nebulizer pressure of 35 psi., a drying gas flow of 8.5 L/min, and a capillary voltage of 4500 V.

Structure elucidation of the compounds

Compounds were purified by reverse-phase HPLC using a Shimadzu C18 reverse-phase column (4.6 × 250 mm, 5 μm), on an HPLC system (Shimadzu, LC-20a) with the HPLC conditions described above. The high-resolution MS was performed on a Q Exactive HF (Q Exactive™ HF/Ultime™ 3000 RSLC nano) using electron spray ionization (ESI) in positive or negative mode (Thermo Fisher Scientific, USA). All spectroscopic experiments were carried out on a Bruker AVANCE III 400, 500, or 600 MHz spectrometer (Bruker Biospin AG, Switzerland). The chemical shift values (δ) were given in ppm with the residual signals of CDCl₃ as the internal standard, and coupling constants (J) in Hz. The structures were determined based on the extensive analysis of 1D and 2D-NMR spectroscopic data.

Statistics & reproducibility

Experiments were performed using at least three biological replicates per condition tested unless otherwise indicated. Data for the column chart showed are reported as bars showing mean, with error bars showing standard deviation. Compound production experiments were performed in biological duplicates.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

67 *Streptomyces* genome information were collected online from NCBI, the complete genome sequence of *Streptomyces* A-14 is available with the BioProject ID [PRJNA1083101](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1083101). The complete genome sequence of *Streptomyces* T1-5, *Streptomyces* T2-4, *Streptomyces* T2-10, *Streptomyces* A-7, *Streptomyces* A-16, and *Streptomyces* A-23 have been submitted to the Genome Sequence Archive (GSA) under the accession code [CRA016553](https://www.genome.org/entry/CRA016553). The RNA-seq dataset is available in NCBI under accession code [PRJNA1168667](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1168667). The data needed to evaluate the conclusions in this study have been deposited in Figshare. Sourced data are available in Figure Share: https://figshare.com/articles/online_resource/Sourced_data_in_NCOMMS-24-27794A/27170487. The associated mass spectrometry data are available in Figure share: https://figshare.com/articles/online_resource/MASS_SPECTROMETRY-NCOMMS-24-27794A/27170583. Source data for the main and other Supplementary Figs. are provided in this paper. The data supporting the findings of this study are available within the article, Supplementary, or Source data files. Source data are provided in this paper.

References

- Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **83**, 770–803 (2020).
- Rutledge, P. J. & Challis, G. L. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat. Rev. Microbiol.* **13**, 509–523 (2015).
- Nett, M., Ikeda, H. & Moore, B. S. Genomic basis for natural product biosynthetic diversity in the actinomycetes. *Nat. Prod. Rep.* **26**, 1362–1384 (2009).
- Adli, M. The CRISPR tool kit for genome editing and beyond. *Nat. Commun.* **9**, 1911 (2018).
- Wu, T. et al. An engineered hypercompact CRISPR-Cas12f system with boosted gene-editing activity. *Nat. Chem. Biol.* **19**, 1384–1393 (2023).
- Whitford, C. M. et al. Systems analysis of highly multiplexed CRISPR-base editing in *Streptomyces*. *ACS Synth. Biol.* **12**, 2353–2366 (2023).
- Cobb, R. E., Wang, Y. & Zhao, H. High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS Synth. Biol.* **4**, 723–728 (2015).
- Zeng, H. et al. Highly efficient editing of the actinorhodin polyketide chain length factor gene in *Streptomyces coelicolor* M145 using CRISPR/Cas9-CodA(sm) combined system. *Appl. Microbiol. Biotechnol.* **99**, 10575–10585 (2015).
- Zhang, J. et al. Efficient multiplex genome editing in *Streptomyces* via engineered CRISPR-Cas12a systems. *Front. Bioeng. Biotechnol.* **8**, 726 (2020).
- Jiang, W. et al. Cas9-Assisted targeting of CHromosome segments CATCH enables one-step targeted cloning of large gene clusters. *Nat. Commun.* **6**, 8101 (2015).
- Liang, M. et al. Activating cryptic biosynthetic gene cluster through a CRISPR-Cas12a-mediated direct cloning approach. *Nucleic Acids Res.* **50**, 3581–3592 (2022).
- Zhang, M. M. et al. CRISPR-Cas9 strategy for activation of silent *Streptomyces* biosynthetic gene clusters. *Nat. Chem. Biol.* **13**, 607–609 (2017).
- Alper, H. S. & Beisel, C. L. Advances in CRISPR technologies for microbial strain engineering. *Biotechnol. J.* **13**, 1800460 (2018).
- Ho, H. I., Fang, J. R., Cheung, J. & Wang, H. H. Programmable CRISPR-Cas transcriptional activation in bacteria. *Mol. Syst. Biol.* **16**, e9427 (2020).
- Qi, L. S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).
- Tian, J. et al. 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 (2020).
- Ni, J., Zhang, G., Qin, L., Li, J. & Li, C. Simultaneously down-regulation of multiplex branch pathways using CRISPRi and fermentation optimization for enhancing β -amyrin production in *Saccharomyces cerevisiae*. *Synth. Syst. Biotechnol.* **4**, 79–85 (2019).
- Chavez, A. et al. Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* **12**, 326–328 (2015).
- Zhao, Y. et al. CRISPR/dCas9-mediated multiplex gene repression in *Streptomyces*. *Biotechnol. J.* **13**, 1800121 (2018).
- Zhang, Y. et al. Antisense RNA interference-enhanced CRISPR/Cas9 base editing method for improving base editing efficiency in *Streptomyces lividans* 66. *ACS Synth. Biol.* **10**, 1053–1063 (2021).
- Li, Q. et al. A study of type II ϵ -PL degrading enzyme (pIdII) in *streptomyces albulus* through the CRISPRi system. *Int. J. Mol. Sci.* **23**, <https://doi.org/10.3390/ijms23126691> (2022).
- Tong, Y., Charusanti, P., Zhang, L., Weber, T. & Lee, S. Y. CRISPR-Cas9 based engineering of actinomycetal genomes. *ACS Synth. Biol.* **4**, 1020–1029 (2015).
- Ameruso, A., Villegas Kcam, M. C., Cohen, K. P. & Chappell, J. Activating natural product synthesis using CRISPR interference and activation systems in *Streptomyces*. *Nucleic Acids Res.* **50**, 7751–7760 (2022).
- Qin, Z. Characterization of the multiple CRISPR loci on *Streptomyces* linear plasmid pSHK1. *Acta Biochim. Biophys. Sin.* **43**, 630–639 (2011).
- Tang, L. Exploring class 1 CRISPR systems. *Nat. Methods* **16**, 1079–1079 (2019).
- Sinkunas, T. et al. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J.* **30**, 1335–1342 (2011).
- Li, Y. et al. Harnessing type I and type III CRISPR-Cas systems for genome editing. *Nucleic Acids Res.* **44**, e34–e34 (2016).
- Zhang, J., Zong, W., Hong, W., Zhang, Z.-T. & Wang, Y. Exploiting endogenous CRISPR-Cas system for multiplex genome editing in

- Clostridium tyrobutyricum* and engineer the strain for high-level butanol production. *Metab. Eng.* **47**, 49–59 (2018).
29. Yang, Z. et al. A thermostable type I-B CRISPR-Cas system for orthogonal and multiplexed genetic engineering. *Nat. Commun.* **14**, 6193 (2023).
 30. Villegas Kcam, M. C., Tsong, A. J. & Chappell, J. Uncovering the distinct properties of a bacterial type I-E CRISPR activation system. *ACS Synth. Biol.* **11**, 1000–1003 (2022).
 31. Hidalgo-Cantabrana, C., Goh, Y. J., Pan, M., Sanozky-Dawes, R. & Barrangou, R. Genome editing using the endogenous type I CRISPR-Cas system in *Lactobacillus crispatus*. *Proc. Natl. Acad. Sci. USA* **116**, 15774–15783 (2019).
 32. Devashish et al. Efficient programmable gene silencing by Cascade. *Nucleic Acids Res.* **43**, 237–246 (2015).
 33. Okamoto, S., Taguchi, T., Ochi, K. & Ichinose, K. Biosynthesis of actinorhodin and related antibiotics: discovery of alternative routes for quinone formation encoded in the act gene cluster. *Chem. Biol.* **16**, 226–236 (2009).
 34. Liu, G., Chater Keith, F., Chandra, G., Niu, G. & Tan, H. Molecular regulation of antibiotic biosynthesis in *Streptomyces*. *Microbiol. Mol. Biol. Rev.* **77**, 112–143 (2013).
 35. Dong, C., Fontana, J., Patel, A., Carothers, J. M. & Zalatan, J. G. Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nat. Commun.* **9**, 2489 (2018).
 36. Kojima, I. et al. The *rpoZ* gene, encoding the RNA polymerase Omega subunit, is required for antibiotic production and morphological differentiation in *Streptomyces kasugaensis*. *J. Bacteriol.* **184**, 6417–6423 (2002).
 37. Shi, J. et al. Structural basis of three different transcription activation strategies adopted by a single regulator SoxS. *Nucleic Acids Res.* **50**, 11359–11373 (2022).
 38. Abudayyeh, O. O. et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353**, aaf5573 (2016).
 39. Zhao, H. et al. Crystal structure of the RNA-guided immune surveillance Cascade complex in *Escherichia coli*. *Nature* **515**, 147–150 (2014).
 40. Luo, Y. et al. Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster. *Nat. Commun.* **4**, 2894 (2013).
 41. Zafar, M. A., Shah, I. M. & Wolf, R. E. Protein-protein interactions between $\sigma 70$ region 4 of RNA polymerase and *Escherichia coli* SoxS, a transcription activator that functions by the prerecruitment mechanism: evidence for “Off-DNA” and “On-DNA” interactions. *J. Mol. Biol.* **401**, 13–32 (2010).
 42. Fouts, D. E., True, H. L. & Celander, D. W. Functional recognition of fragmented operator sites by R17/MS2 coat protein, a translational repressor. *Nucleic Acids Res.* **25**, 4464–4473 (1997).
 43. Kim, E.-S., Hong, H.-J., Choi, C.-Y. & Cohen Stanley, N. Modulation of actinorhodin biosynthesis in *Streptomyces lividans* by glucose repression of *afsR2* gene transcription. *J. Bacteriol.* **183**, 2198–2203 (2001).
 44. Greco, C. et al. Structure revision of cryptosporioides and determination of the genetic basis for dimeric xanthone biosynthesis in fungi. *Chem. Sci.* **10**, 2930–2939 (2019).
 45. Neubauer, L., Dopstadt, J., Humpf, H. U. & Tudzynski, P. Identification and characterization of the ergochrome gene cluster in the plant pathogenic fungus *Claviceps purpurea*. *Fungal Biol. Biotechnol.* **3**, 2 (2016).
 46. Rateb, M. E. et al. Legonaridin, a new member of linaridin RiPP from a Ghanaian *Streptomyces* isolate. *Org. Biomol. Chem.* **13**, 9585–9592 (2015).
 47. Funabashi, M., Funa, N. & Horinouchi, S. Phenolic lipids synthesized by type III polyketide synthase confer penicillin resistance on *Streptomyces griseus*. *J. Biol. Chem.* **283**, 13983–13991 (2008).
 48. Ueda, K. et al. AmfS, an extracellular peptidic morphogen in *Streptomyces griseus*. *J. Bacteriol.* **184**, 1488–1492 (2002).
 49. Tietz, J. I. et al. A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat. Chem. Biol.* **13**, 470–478 (2017).
 50. Prabhu, J., Schauwecker, F., Grammel, N., Keller, U. & Bernhard, M. Functional expression of the ectoine hydroxylase gene (*thpD*) from *Streptomyces chrysomallus* in *Halomonas elongata*. *Appl. Environ. Microbiol.* **70**, 3130–3132 (2004).
 51. Schimming, O. et al. Structure, biosynthesis, and occurrence of bacterial pyrrolizidine alkaloids. *Angew. Chem. Int. Ed.* **54**, 12702–12705 (2015).
 52. Wang, S. et al. Unleashing the potential: type I CRISPR-Cas systems in actinomycetes for genome editing. *Nat. Prod. Rep.* **41**, 1441–1455 (2024).
 53. Xu, Z. et al. A transferrable and integrative type I-F Cascade for heterologous genome editing and transcription modulation. *Nucleic Acids Res.* **49**, e94–e94 (2021).
 54. Rajoka, M. I. et al. Cloning and expression of beta-glucosidase genes in *Escherichia coli* and *Saccharomyces cerevisiae* using shuttle vector pYES 2.0. *Folia Microbiol.* **43**, 129–135 (1998).
 55. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
 56. Kieser, T. et al. *Practical Streptomyces Genetics*. John Innes Foundation (2000).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (32471492, 32071426, to Y.L.), the National Key Research and Development Program of China (2018YFA0903300, to Y.L.), the Haihe Laboratory of Sustainable Chemical Transformations for financial support (24HHWCSS00006, to Y.L.), and the Key-Area Research and Development Program of Guangdong Province (2020B0303070002, to Y.L.). We thank Prof. Min Yin (Yunnan University) for providing strain *Nocardiopsis halophila* DSM 44494.

Author contributions

Y.L. designed the experiments. Y.Z., Q.Z., and H.W. performed the experiments and analyzed the data. C.K., S.G., and Y.Y. determined the chemical structures of the compounds. H.L. performed the strain construction. Y.L., Y.Z., and Q.Z. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-54196-z>.

Correspondence and requests for materials should be addressed to Yang Ye or Yunzi Luo.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024