

Internal promoters of the epothilone biosynthetic gene cluster and their activation in *Myxococcus xanthus* by CRISPR-dCas9

Ye Wang

Shandong University

Xin-jing Yue

Shandong University

shu-fei Yuan

Shandong University

Yu Hong

Shandong University

Wei-feng Hu

Shandong University

Yue-zhong Li (✉ lilab@sdu.edu.cn)

Shandong University <https://orcid.org/0000-0001-8336-6638>

Research

Keywords: internal promoters, transcriptional regulation, CRISPR-dCas9 activation, epothilone, biosynthetic gene cluster, *Myxococcus xanthus*

Posted Date: March 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-338190/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

Multiple genes involving in a complex pathway are often clustered into a giant operon with no transcription terminator before the end, and this leads to frangibility of the transcriptional process and arduous engineering work to control the transcription of operon genes. Internal promoters might occur in operon to coordinate the transcription of individual genes, but their effects on the transcription of operon genes have been less investigated.

Results

Epothilones are a kind of polyketides synthesized by seven multifunctional enzymes, which are encoded by a 56-kb operon of the myxobacterium *Sorangium cellulosum*. In this study, we determined that the epothilone operon contained multiple internal promoters. These promoters were activatable by the CRISPRa technique, and the yields of epothilones were accordingly increased. However, the activation efficiencies of promoters in operon and separate forms were greatly different. Further, we found that the transcriptional levels of the epothilone genes were always increased at a greater extent than the epothilone yields, which suggested that the transcriptional activation of single genes probably had a weak effect on the final epothilone yield, and higher yield required an overall transcriptional increase of the multiple operonic genes. Finally, we combined the activation of the starting promoter P_{epoA} together with internal promoters in different epothilone-producing strains, and obtained the highest 15-fold increase of epothilone yield in *Myxococcus xanthus* ZE5.

Conclusions

This is the first time to report the internal promoters in epothilone gene clusters in *Myxococcus xanthus* and the first time to assay the activation effects of these internal promoters by CRISPR-dCas9. Our results highlight that tuning internal promoter activities is critical to control the transcription of operon genes and the production efficiency of microbial secondary metabolites.

Background

Operons, a well-known feature of prokaryotic genomes, are clusters of co-regulated genes with related functions [1]. Big operons, such as pathways for the biosynthesis of secondary metabolites, may contain multiple genes with no transcription terminator before the end. The transcription of multiple genes in operon is initiated by the starting promoter, forming a single polycistronic mRNA. However, the transcription of big mRNA molecules is easily subject to various influences in cells, leading to frangibility of the transcriptional process and arduous engineering work to control the transcription of operon genes. Internal promoters have been observed in operons for many years. According to experiments and genome prediction, internal promoters are universal in operons in different bacteria [2–7]. For example, Ma et al. demonstrated that there are at least three weak internal promoters P2, P3 and P4 in addition to two strong promoters P_L and P1 in the *E. coli rpoBC* operon, which encodes four ribosomal proteins and the β and β' subunits of RNA polymerase [2]. In the 57-kb jamaicamide operon from the marine cyanobacterium *Lyngbya majuscula*, 17 genes are co-transcribed from the starting promoter, and six internal promoters are present in intergenic regions, which potentially assist management of the toxin production in various environments [3].

Similarly, seven internal promoters have also been observed in the microcystin operon of *Microcystis aeruginosa* [4]. However, the effects of internal promoters on the transcription of operon genes have been less investigated.

Epothilones are a kind of antitumor polyketide compounds with the microtubule-stabilizing activity, which were originally discovered in the extracts of some *Sorangium cellulosum* strains [8]. In 2000, Molnár and Julien reported the sequence and construction of the epothilone gene cluster, which is approximately 56 kb, containing seven open reading frames (ORFs) in the same transcriptional direction [9, 10]. These ORFs encode polyketide synthases and a non-ribosomal peptide synthetase with multiple functioning domains for elongation, modification and release of epothilones in a sequential mode [9]. This gene cluster contains no terminator between ORFs before the end, thus forming a giant operon.

In previous work, we have successfully integrated the epothilone biosynthetic gene cluster from *S. cellulosum* So0157-2 [11] into different sites of the *Myxococcus xanthus* genome, and produced dozens of epothilone producers [12]. We found that the transcriptional levels of operon genes greatly varied in either *M. xanthus* or the original *S. cellulosum* producers. We constructed a CRISPRa (CRISPR-dCas9-mediated transcription activation) system in *M. xanthus* and have successfully improved the production of epothilones by activating the starting promoter of the operon [13]. However, the transcriptional levels of operon genes still varied. In fact, uneven gene expression of operon genes was observed many years ago [14], and genome-wide transcriptomics studies have also revealed varied transcriptional levels of the consecutive genes within operons [15], which suggest that the operon genes are complexly regulated, rather than transcribed from the single starting promoter. Internal promoters are suggested to coordinate the transcription of operon genes, but their effects remain mysterious. In this study, we determined the existence of multiple internal promoters in the epothilone biosynthetic gene cluster. We employed the established CRISPRa technique to promote the transcription of internal promoters, and found that the activation efficiency of promoters in operon was distinct from that of them in separate form. We combined the activation of the starting promoter *PepoA* together with internal promoters in different epothilone-producing strains, and obtained the highest 15-fold increase of epothilone yield in *Myxococcus xanthus* ZE5. Our results highlight that coordinating transcriptional activities of internal promoters is critical for the transcription of operon genes and the production efficiency of microbial secondary metabolites..

Results

1. The epothilone gene cluster is a big operon containing multiple internal promoters

The epothilone gene cluster contains seven ORFs (*epoA* to *epoF*) in the same transcriptional direction, which are separated with a short distance (15 ~ 147bp) or shortly overlapped (Fig. 1a). No terminator structure exists in the intergenic regions of these ORFs, which means that the epothilone gene cluster is a huge operon. The continuous transcription of the gene cluster was also proved by co-transcriptional detection of two adjacent ORFs (**Figure S1**). However, these operon ORFs exhibited significant differences in their transcriptional levels. For example, as determined in *M. xanthus* ZE9, an epothilone-producer constructed with the whole gene cluster of *S. cellulosum* So0157-2 [12], the last ORF *epoF* showed the highest expression level by approximately 2.5 times of that of the first ORF *epoA* (Fig. 1b). We suggested that internal promoters might be present in this operon to tune the transcription of individual ORFs.

We predicted the existence of internal promoters in the epothilone operon derived from *S. cellulosum* So0157-2 using the “Neural Network Promoter Prediction” program (https://www.fruitfly.org/seq_tools/promoter.html), and found ten potential promoters (shown as red arrows in Fig. 1a). Prediction with “BSPROM” (<http://www.softberry.com/berry.phtml>) revealed similar results but excavating a new promoter upstream of *epoD* (shown as green arrows in Fig. 1a). Detailed information of the predicted promoters is shown in **Table S1**. These promoters mostly appeared in the junctional regions between ORFs. Up to now, four *S. cellulosum* strains have been reported of their epothilone biosynthetic gene clusters with the GenBank accession numbers of AF210843.1 in SMP44 [9], AF217189.1 in So ce90 [10] and GU063811.1 in KYC3013 [16], in addition to EU414841.1 in So0157-2 [11]. The sequences of these four epothilone gene clusters exhibited more than 98.6% similarity. In the epothilone operons from other three *S. cellulosum* producers, internal promoters were similarly present but with some differences (**Figure S2**; detailed information is provided in **Table S2 ~ S4**).

To analyze activities of internal promoters, we selected six regions that locate in the junction between two ORFs in the So0157-2 epothilone operon. These regions were 1000 bp in length, including the 800-bp fragment of an upstream gene and the 200-bp fragment of a downstream gene, which were amplified from *M. xanthus* ZE9 using the corresponding primer pairs (listed in **Table S5**). Each region, containing at least one predicted internal promoter, was cloned into the pKK232-8 plasmid to control the expression of chloramphenicol acetyltransferase (CAT) (**Figure S3**). The *aphII* promoter was constructed upstream of the *CAT* reporter gene as a positive control. According to the CAT activities, assayed with the CAT ELISA kit, all the six regions exhibited transcriptional activities in *E. coli*: the P_{epoP} promoter showed the highest activity, which was close to that of the *aphII* promoter, while the P_{epoB} and P_{epoE} promoters showed weak activities, slightly higher than that of the negative control (without a promoter before the *CAT* gene) (Fig. 1c).

Simultaneously, we also analyzed promoter activities of the 1 kb sequence upstream of the translation initiation codon (ATG or GTG) of each ORF in *M. xanthus*. The six fragments were constructed respectively in front of the *EGFP* (Enhanced Green Fluorescent Protein) reporter gene in pZJY41, an autonomously-replicating plasmid in *M. xanthus* [17] (**Figure S4**). The identified starting promoter P_{epoA} in front of the epothilone operon [18] was used as the positive control. We introduced these plasmids into *M. xanthus* DZ2 and assayed the green fluorescence values of the mutants after 24 h, 48 h and 72 h of incubation. The results showed that these fragments, even the sequence upstream of *epoP*, where no promoter had been predicted, exhibited promoter activities; P_{epoA} always had the highest activity, and the activity of P_{epoP} is the lowest (Fig. 1d). Notably, the detected promoter activities, either in *E. coli* or *M. xanthus*, were significantly inconsistent with the transcriptional levels of the genes in the epothilone operon. For instance, we did not find a strong promoter upstream of the *epoF*, which, however, displayed the highest transcriptional level among the ORFs in the epothilone operon (Fig. 1b). These results suggested that these internal promoters might be complexly regulated of their transcriptional activities in operon.

2. Operon and separate internal promoters exhibit different transcriptional activities

CRISPR (clustered regularly interspaced short palindromic repeat) Cas system is an RNA-mediated immune system existing in many bacteria and archaea to protect cells from foreign DNA invasion [19–21]. Peng *et al.* constructed a CRISPRa system in *M. xanthus* and successfully activated the P_{epoA} promoter of the epothilone operon, which increased the transcriptional levels of operon genes and improved the yield of epothilones by 230% [13]. The schematic diagram of CRISPRa construction in *M. xanthus* is shown in Fig. 2a.

We performed the CRISPRa-based activation on separate promoters in *E. coli*. To achieve activation, we transferred three plasmids into the HB101 strain: the pSWcuomxdCas9- ω plasmid carrying dCas9 protein and transcription activator Omega(ω) [13], the pZJY41-sgRNA series plasmid carrying sgRNA, and the pkk232 series plasmid carrying promoter P_{epo} and reporter gene *CAT* (**Figure S5**). To construct the pZJY41-sgRNA series plasmids, we employed one spacer for each promoter (refers to Table 1), which were designed using the online software “Cas-OT” [22]. We found that the transcription activities of these separate promoters were all significantly improved: weak promoters were more easily activated by CRISPRa, and the promoters with high transcription activities were also activated but to low extents (Fig. 2b), which is consistent with the previous report [23]. For example, the transcription activity of the weakest promoter P_{epoB} was increased by nearly 33-fold, while the strongest promoter P_{epoA} was activated by approximately 1.6-fold.

Table 1
Information of the spacer sequences

Name	Spacer	Coding strand	Off-target	Distance(bp)	ΔG (kcal/mol)	Hairpin	GC%
CuOm-P2	TCCGGGGGATGATGCTCGAG	+	8	-617	43.6	N	65.0
CuOm-B3	TGAGGAGCCTGTTGCAGAAG	+	14	-376	38.2	N	55.0
CuOm-C2	ACCGTACCGGCAACGCTGTTG	+	4	-165	45.3	N	61.9
CuOm-D2	TGCGGCCGGTATCCTGGACGA	+	3	-117	47.8	N	66.7
CuOm-E1	TGGATGTATCCCAAGGTGCT	+	2	-160	38.1	N	50.0
CuOm-F2	AGCTCTTCTTCCGAAATGCCG	+	2	-193	43.5	N	52.4

We further performed in-situ activation on the internal promoters within the epothilone operon in *M. xanthus* ZE9 by using the same spacer sequences for the activation of each of the six internal promoters (Table 1). The spacer sequences were constructed into the pZJY41 plasmid, respectively (**Figure S6**), producing six series plasmids (**Table S6**), which were separately introduced into the CuOm strain by electro-transformation. The CuOm strain was constructed by introducing the pSWcuomxdCas9- ω plasmid into ZE9 (a diagrammatic sketch for the construction is shown in **Figure S7**). The results showed that the transcriptional levels of most of the activated genes were significantly increased, except that the expression of *epoD* in CuOm-D2 was increased insignificantly (*t*-test, $p > 0.05$) (Fig. 2c). The best activation in *M. xanthus* was achieved on the *epoE* gene, which was activated by about 5-fold. Notably, the results were markedly different from that of the activation on the separate promoters although using the same spacer sequences. Thus, the transcription activities of internal promoters were influenced in operon.

In addition, we found that the epothilone production abilities in these CRISPRa-promoted mutants were significantly increased, even including CuOm-D2, in which the expression of *epoD* was not increased significantly by the CRISPRa technique (Fig. 2d). We found that the transcriptional levels of the epothilone genes were always increased at a greater extent than the epothilone yields. For example, the highest gene expression was found in

CuOm-E1 (increased by 5-fold), but the epothilone yield was only increased by 1.7-fold. We thus checked the transcription of each ORF in the epothilone operon with the activation of single internal promoters. The results showed that the activation of the front promoter P_{epoP} , P_{epoB} or P_{epoC} normally increased the transcription of the front genes *epoA*, *epoP*, *epoB* and *epoC*, but did not change or even decreased the transcription of the hind genes *epoD*, *epoE* and *epoF*, whether the genes were specifically activated or not (**Figure S8**). Similarly, the activation of the hind promoter P_{epoD} , P_{epoE} or P_{epoF} increased the transcription of the hind genes, and sometimes the upstream *epoB* and *epoC* genes, but had no effect on *epoA* or *epoP*. The results suggested that the transcriptional activation of single genes probably had a weak effect on the final epothilone yield, and higher yield required an overall transcriptional increase of the multiple operonic genes.

3. Tuning the activity of multiple promoters to increase the epothilone yield

To improve the activation effects to increase the yield of epothilones, we combined the activation of the starting promoter P_{epoA} together with an internal promoter P_{epoP} , P_{epoB} or P_{epoD} , thus forming the AP, AB or AD combination. We also combined the activation of three promoters, i.e., APB, APD and ABD. The combined sgRNA sequences were cloned into the pZJY41 plasmids, which were introduced into the CuOm strain, respectively (**Figure S9**). After the CRISPRa-based activation, the production of epothilones was increased in each combination, and the highest yields were obtained in the mutants CuOm-AP (11.17 mg/L) and CuOm-ABD (11.69 mg/L), both of which increased approximately 2.4-fold, compared with the 4.95 mg/L in the initial strain ZE9 (Fig. 3a). However, the combined activation of two or three promoters did not lead to an accumulation of the activation effects of single promoters.

We analyzed the transcription of each operon ORF under different activation combinations (Figs. 3b and 3c). The results showed that the transcriptional levels of the front operon genes *epoA*, *epoP* were mostly significantly increased by the combined activation, even when P_{epoP} was not specifically activated (in CuOm-AB, CuOm-AD or CuOm-ABD). However, the transcription of *epoD*, as well as the genes behind, was not increased in CuOm-AD, CuOm-ABD or CuOm-APD, in which P_{epoD} was even specifically activated. Seemingly, the genes close to the starting promoters P_{epoA} were more easily activated than the hind operon genes.

Based on the abovementioned results, we combined CRISPRa-mediated activation on promoters in *M. xanthus* strains with different transcription levels of the epothilone genes. We previously constructed dozens of epothilone-producing *M. xanthus* strains, in which the same epothilone biosynthetic gene cluster was inserted in different sites of the DZ2 genome, resulting in varied production abilities of epothilones [12]. We chose four strains ZE9, ZE5, ZE10 and ZE14 to assay the transcriptional efficiency using the CRISPRa technique. Among these four strains, ZE9 had the highest production ability of epothilones, followed by ZE10, then ZE14, and ZE5 exhibited the lowest epothilone yields (Fig. 4a). Consistently, the transcriptional levels of the seven operon ORFs in ZE9 were higher than that in the other three strains (Fig. 4b). In ZE14, the front operon genes were transcribed at higher levels, but *epoD*, *epoE* and *epoF* were significantly lower than that in ZE10. Similarly, the transcriptional levels of *epoD*, *epoE* and *epoF* were extremely low in ZE5, causing the strain to produce the lowest yield of epothilones among the four strains.

We combined the activations either on the front promoters (P_{epoA} , P_{epoP} and P_{epoB}) or the hind promoters (P_{epoD} , P_{epoE} and P_{epoF}) in these epothilone producing strains. As expected, the epothilone yields were all increased in these strains, and the highest 15-fold increase of epothilone yield was obtained in ZE5 with the DEF promoter

activation (Fig. 4a). Consistent with the yields of epothilones, the highest activation efficiency also occurred in the ZE5 strain. In ZE5-DEF, the transcriptional levels of the three activated genes *epoD*, *epoE* and *epoF* were increased by 9.6, 3.1 and 51.7 times, respectively, and *epoP* and *epoB* were also increased slightly (Fig. 4c). However, the transcriptional changes of operon genes suggested that the interferences between operon promoters were very complex. For example, in the ZE5-APB strain, the transcriptional levels of the three activated genes *epoA*, *epoP* and *epoB* were all increased, but to different extents; the transcriptional levels of the four hind genes (*epoC~epoF*), which were not specially activated, were also mostly increased.

Discussion

Transcription regulation is always a topic of concern. Operons are clusters of co-regulated genes with related functions. Bacteria have established multiple mechanisms to ensure the relative expressional levels of individual genes in operon to meet the requirements of cell and environment [24, 25]. Transcriptional interference between tandem promoters is recognized as a potentially widespread mechanism to regulate gene expression [26, 27]. There are many studies on the regulation of single internal promoters on the expression of operon genes, and a few studies have been performed on transcriptional interferences between multiple operon promoters with no clear conclusions. For example, the 14-kb *CAP1* gene cluster in *Staphylococcus aureus* is transcriptionally controlled by a strong upstream promoter and five weak internal promoters, and the internal promoters showed significant activity only after removing the primary promoter [28]. In the cyanobacterium *Anabaena* sp. strain PCC 7120, a zinc-responsive operon contains 4 distinct promoters, which were induced by metal depletion, and they were constitutively derepressed in a *zur* mutant, despite the two downstream promoters not being direct targets for this regulator [29]. In this study, we demonstrated that the big epothilone gene operon contained multiple internal promoters, and the transcriptional processes of these internal promoters may intricately interfere with each other.

Interference between tandem promoters may be generated by dislodgement of slow-to-assemble pre-initiation complexes and transcription factors, or prolonged occlusion by paused RNA polymerases (RNAPs) [29, 30, 31]. The direct and in cis suppression of one transcriptional process by another transcriptional process is that RNAP transcribing from one promoter may have effects on the supercoiling state of neighbor promoters [26,32.33]. A study in the human and mouse genomes concluded that RNAP collisions were the primary mechanism of interaction between transcripts, based upon decreasing transcript abundance with increasing overlap length, to almost zero when the overlap exceeded 2,000 nucleotides [34]. Mathematical modeling also shows that the probability that an RNAP can avoid colliding with an RNAP from a convergent promoter decreases exponentially with the firing rate of the interfering promoter and with the inter-promoter distance [35]. However, the transcriptional interference between tandem promoters is often not satisfactorily explained by RNAP collisions or occlusion by elongating RNAP [27], and there are other unknown regulatory patterns [29].

Conclusions

The inconsistency in the transcriptional levels of operon genes often limits the yield of secondary metabolites. The inconsistent expression levels of genes in operon observed in different bacterial species are not only challenging the concept of operons [36.37], but also impeding our engineering work to control the transcription of operon genes. Our results present in this study indicated that, multiple internal promoters are present in epothilone gene cluster. Although little is known of the involving mechanism, regulation of operon internal promoters should be crucial for the biosynthetic pathways of secondary metabolites encoded by a big operon. Tuning the

transcriptional activities of operon promoters, such as using CRISPRa technique, can efficiently improve the metabolite yields.

Materials And Methods

Strains and culture conditions

Strains used in this study are listed in **Table S7**.

Escherichia coli DH5 α and HB101 were used for routine transformations and sub-cloning. The *E. coli* strains were grown routinely in Luria Broth (LB) medium (10 g/L peptone, 5 g/L yeast extract, and 5 g/L NaCl, pH 7.2).

Myxococcus xanthus strains were grown in CYE medium [10 g/L casitone, 5 g/L yeast extract, 10 mM 3-(N-morpholino) propanesulfonic acid (MOPS; pH 7.6) and 4 mM MgSO $_4$] or CMO medium [10 g/L casitone, 5 g/L yeast extract, 10 mM MOPS, 4 mM MgSO $_4$, and 7 mL/L methyl oleate, pH 7.6]. When appropriate, kanamycin (40 μ g/mL), tetracycline (10 μ g/mL) and apramycin (25 μ g/mL) were added. The growth temperatures were 37 °C for *E. coli* and 30 °C for *M. xanthus*.

Detection of gene co-transcription of epothilone gene cluster

Total RNA of samples was extracted using BIOZOL kits (Total RNA Extraction Regent, BioFast, China) and then transcribed reversely into cDNA with PrimeScript™ Reagent Kit with DNAase (Takara, Japan). Primers were designed at the junction between two genes. Total RNA without inverse transcription and the plasmid containing the epothilone gene cluster were used as templates in the negative control group and positive control group, respectively. In experimental groups, the cDNAs were used as templates.

Prediction of internal promoters in the epothilone gene cluster

We used the promoter prediction software “Neural Network Promoter Prediction” ([https:// www. fruitfly. org/seq_ tools/ promoter. html](https://www.fruitfly.org/seq_tools/promoter.html)) to predict internal promoters in the epothilone gene clusters derived from *Sorangium cellulosum* strains So ce90, SMP44, So0157-2 and KYC3013. The threshold was set as 0.8. At the same time, we used another online promoter prediction software “BPROM” ([http:// www. softberry. com/ berry. phtml? Topic = bprom & group = programs & subgroup = gfindb](http://www.softberry.com/berry.phtml?Topic=bprom&group=programs&subgroup=gfindb)) to correct predictive results. Bold -35 and -10 binding regions were predicted in comparison to the σ 70 consensus -35 (TTGACA) and -10 (TATAAT) promoter regions of *E. coli*. Then we selected the prediction results with a threshold greater than 0.8.

Construction of plasmids

The plasmids and primers used in this study are provided in **Tables S5** and **S6**.

pkk-232-P_{epoP} ~ pkk-232-P_{epoF} were used as a promoter activity reporter vector in *E. coli*. P_{epoP} ~ P_{epoF} were obtained by PCR with primers P_{epoP}-F/R ~ P_{epoF}-F/R from epothilones gene cluster in So0157-2. pkk-232-P_{epoP} ~ pkk-232-P_{epoF} were constructed by inserting P_{epoP} ~ P_{epoF} into the HindIII/BamHI sites of pkk-232-8. pKK-232-aph were constructed by inserting *aph* into the HindIII/BamHI sites of pkk-232-8.

pZJY41-Ap-EGFP ~ pZJY41-Fp-EGFP were used as a promoter activity reporter vector in *Myxococcus xanthus*. P_{epoA} was obtained by PCR with primers Ap-F/R from So0157-2 genome, while report gene EGFP was amplified

with primers EGFP-F/R, and then overlap PCR was used to obtain P_{epoA} -EGFP with primers P_{epoA} -F and EGFP-R. Whereafter, P_{epoA} -EGFP was inserted into the BamHI/KpnI sites of pZJY41 to construct the reporter vector pZJY41-Ap-EGFP. The other six promoter activity reporter vectors were constructed in the same way.

Plasmid pZJY41-sgRNA was used to express sgRNA. Plasmids pZJY41-sgRNA-P2 ~ pZJY41-sgRNA-F2 were constructed by using p41sg [13] as a template to amplify DNA fragment with primers sgRNA-P2-F ~ sgRNA-F2-F and sgRNA-R, and then linking with T4 DNA ligase to form circular plasmids..

pZJY41-sgRNA-AP was constructed by inserting the sgRNA scaffold with spacer A4 and sgRNA scaffold with spacer P3 into the NdeI site and KpnI/BamHI sites of pZJY41, successively. pZJY41-sgRNA-AB and pZJY41-sgRNA-AD were constructed in the same way. With sgRNA scaffold with spacerA4 inserted into the NdeI site, sgRNA scaffold with spacerP3 inserted into the KpnI/BamHI sites, sgRNA scaffold with spacerB3 inserted into the EcoRI site of pZJY41, pZJY41-sgRNA-APB was constructed. So were the pZJY41-sgRNA-APD and pZJY41-sgRNA-ABD plasmids.

Activity detection of internal promoters in *E. coli*

We constructed plasmids to detect the activity of internal promoters from the epothilone operon in *E. coli*. The promoter activity was characterized by detecting the activity of the report gene chloramphenicol acetyltransferase (*CAT*). The promoter sequences and reporter gene sequence were cloned into plasmid pKK232-8 by digestion with BamHI/HindIII and then ligation with T4 DNA ligase. The activity of reporter gene *CAT* was detected by CAT ELISA Kit. The promoter *aphII* was used as a positive control. The CAT ELISA Kit was purchased from Roche and operated according to the instructions provided ([https:// www. sigmaaldrich. com/ catalog/ product/ roche/ 11363727001? Lang = zh & region = CN](https://www.sigmaaldrich.com/catalog/product/roche/11363727001?Lang=zh®ion=CN)).

Activity detection of internal promoters in *Myxococcus xanthus*

We constructed plasmids to detect the activity of internal promoters from the epothilone operon in *M. xanthus*. The promoter activity was characterized by detecting the fluorescence intensity of the green fluorescence reporter gene *EGFP*. The promoter sequences and the *EGFP* gene sequence were seamlessly connected by fusion PCR, and finally cloned into the plasmid pZJY41 by digestion with BamHI/KpnI and then ligation with T4 DNA ligase. Related primer information was shown in Table S5. The fluorescence intensity of the green reporter gene was detected at 485 nm/528 nm, and three different incubation times (24h, 36h and 72h) were selected for detection.

Construction of CRISPRa-dCas9 system in *E. coli*

We transfected three plasmids into *E. coli* competent cells HB101 at the same time: pSWcuomxdCas9- ω plasmid [13] carrying mxdCas9 protein and transcription activator, Omega(ω); pZJY41-sgRNA series plasmid carrying sgRNA with different spacer sequences; pKK232 series plasmid carrying internal promoter and reporter gene *CAT*. Related sequences and plasmid information were shown in **Tables S5** and **S6**.

Construction of CRISPRa-dCas9 system in *Myxococcus xanthus*

We transfected two plasmids into *Myxococcus xanthus*: pSWcuomxdCas9- ω plasmid [13] carrying mxdCas9 protein and transcription activator Omega(ω); pZJY41-sgRNA series plasmid carrying sgRNA. Related sequences and plasmid information are shown in **Tables S5** and **S6**.

Extraction and detection of epothilones

ZE9 and mutants were grown overnight in 50 mL of CYE medium supplemented with Apra (30 µg/mL). The cultures were inoculated at a ratio of 2:100 into 50 mL of CMO medium containing 2 % of the XAD-16 resin. The resin was harvested with a strainer after 6 days and extracted with 3 mL of methanol by shaking at room temperature overnight. The supernatant was centrifuged for 10 min at 12,000 rpm and filtered with 0.22 µm filter to remove the impurities. 20 µl of the sample was injected into High Performance Liquid Chromatography (HPLC, SHIMADZU, Japan) and analyzed on a Shim-pack MRCODS RP C18 column (4.6 mm × 250 mm, 4.60 µm; Shimadzu, Japan) and monitored at 250 nm, with a mobile phase of 60% of methanol (HPLC grade) and 40% of H₂O at a flow rate of 1.0 mL/min. The yield of epothilone was quantified from the peak area in the UV chromatogram, by reference against a calibration standard.

Transcriptional analysis of epothilone genes with RT-qPCR

We collected samples continuously from the fermentation culture at 48 h of incubation. Then, total RNA of samples was extracted using BIOZOL kits (Total RNA Extraction Regent, Bio Fast, China) and then transcribed reversely into cDNA with PrimeScript™ Reagent Kit with DNAase (Takara, Japan). The *gapA* gene (glyceraldehyde-3-phosphate dehydrogenase gene, MXAN_2815) was chosen as the reference gene for normalization. The transcriptional level of epothilone gene cluster was analyzed by RT-qPCR on LightCycler®480 (Switzerland) with SYBR® Premix Ex Taq™ GC Dye (Takara, Japan). All the primers used in RT-qPCR were listed in TableS5.

Declarations

Author contributions

YW, XJY and YZL designed researches; YW, XJY, SFY, YH and WFH performed researches; YW, XJY and YZL analyzed the data and wrote the paper.

Acknowledgments

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and in its Additional files.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Funding

This work was supported by the National Key Research and Development Program of China (2018YFA0900400 and 2018YFA0901704) to YZL, National Natural Science Foundation of China (31670076) to YZL, Natural Science Foundation of Shandong Province (ZR2019BC041) to XJY and China Postdoctoral Science Foundation (2018M642647) to XJY.

References

1. Ma JC, Newman AJ, Hayward RS. Internal promoters of the *rpoBC* operon of *Escherichia coli*. *Molecular and General Genetics*. 1981; 184(3):548-550.
2. Jones AC, Gerwick L, Gonzalez D, Dorrestein PC, Gerwick WH. Transcriptional analysis of the jamaicamide gene cluster from the marine cyanobacterium *Lyngbya majuscula* and identification of possible regulatory proteins. *BMC Microbiology*. 2009; 9(1):247.
3. Kaebernick M, Dittmann E, Börner T, Neilan BA. Multiple alternate transcripts direct the biosynthesis of microcystin, a cyanobacterial nonribosomal peptide. *Applied and Environmental Microbiology*. 2002; 68(2):449-455.
4. Makita Y, Nakao M, Ogasawara N, Nakai K. DBTBS: database of transcriptional regulation in *Bacillus subtilis* and its contribution to comparative genomics. *Nucleic Acids Research*. 2004; 32(90001):75-7.
5. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiß S, Sittka A, Chabas S, Reiche K, Hackermüller J, Reinhardt R, Stadler PF and Vogel J. The Primary Transcriptome of the Major Human Pathogen *Helicobacter Pylori*. *Nature*. 2010; 464(7286):250-255.
6. Güell M, Noort VV, Yus E, Chen WH, Leigh-Bell J, Michalodimitrakis K, Yamada T, Arumugam M, Doerks T, Kühner S, et al. Transcriptome Complexity in a Genome-Reduced Bacterium. *Science*. 2009; 326(5957):1268-1271
7. Bollag DM, McQueney PA, Zhu J, Hensens O, Koupal L, Liesch J, Goetz M, Lazarides E, Woods CM. Epothilones, a new class of microtubule-stabilizing agents with a taxol-like mechanism of action. *Cancer Research*. 1995; 55(11):2325-33.
8. Julien B, Shah S, Ziermann R, Goldman R, Katz L, Khosla C. Isolation and characterization of the epothilone biosynthetic gene cluster from *Sorangium cellulosum*. *Gene*. 2000; 249(1):153-160.
9. Molnár I, Schupp T, Ono M, Zirkle RE, Milnamow M, Nowak-Thompson B, Engel N, Toupet C, Stratmann A, Cyr DD, Grolach J, Mayo JM, Hu A. The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from *Sorangium cellulosum* So ce90. *Chemistry and Biology*. 2000; 7(2):97-109.
10. Han K, Li ZF, Peng R, Zhu LP, Zhou T, Wang LG, Li SG, Zhang XB, Hu W, Qin N, Li YZ. Extraordinary expansion of a *Sorangium cellulosum* genome from an alkaline milieu. *Scientific Reports*. 2013; 3(1): e1000565-1289.
11. Zhu LP, Yue XJ, Han K, Li ZF, Zheng LS, Yi XN, Wang HL, Zhang YM, Li YZ. Allopatric integrations selectively change host transcriptomes, leading to varied expression efficiencies of exotic genes in *Myxococcus xanthus*. *Microbial Cell Factories*. 2015; 14(1):105.
12. Peng R, Wang Y, Feng WW, Yue XJ, Chen JH, Hu XZ, Li YZ. CRISPR/dCas9-mediated transcriptional improvement of the biosynthetic gene cluster for the epothilone production in *Myxococcus xanthus*. *Microbial Cell Factories*. 2018; 17(1):15.
13. Murakawa GJ, Kwan C, Yamashita J, Nierlich DP. Transcription and decay of the *lac* messenger: role of an intergenic terminator. *Journal of Bacteriology*. 1991; 173(1):28-36.

14. De-Hoon MJL, Makita Y, Nakai K, Miyano S. Prediction of Transcriptional Terminators in *Bacillus subtilis* and Related Species. *PLoS Computational Biology*. 2005; 1(3): e25
15. Hyesook H, Jinwoo C, Kim J, Lee JS, Kwon BM, Son KH, Cho K. Isolation of *Sorangium cellulosum* carrying epothilone gene clusters. *Journal of Microbiology and Biotechnology*. 2008; 18(8):1416-22.
16. Zhao JY, Zhong L, Shen MJ, Xia ZJ, Cheng QX, Sun X, Zhao GP, Li YZ, Qin ZJ. Discovery of the Autonomously Replicating Plasmid pMF1 from *Myxococcus fulvus* and Development of a Gene Cloning System in *Myxococcus xanthus*. *Applied and Environmental Microbiology*. 2008; 74(7):1980-1987.
17. Zhu LP, Li ZF, Sun X, Li SG, Li YZ. Characteristics and activity analysis of epothilone operon promoters from *Sorangium cellulosum* strains in *Escherichia coli*. *Applied Microbiology and Biotechnology*. 2013; 97(15):6857-6866.
18. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*. 2005; 151(8):2551-2561.
19. Pourcel C, Salvignol G, Vergnaud G. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology*. 2005; 151(3):653-663.
20. Marraffini LA, Sontheimer EJ. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nature Reviews Genetics*. 2010; 11(3):181-190.
21. Xie SS, Zhang Y, Zhang LS, Li GL, Zhao CZ, Ni P, Zhao SH. sgRNA design for the CRISPR/Cas9 system and evaluation of its off-target effects. *Hereditas*. 2015; 37(11):1125-1136.
22. Bikard D, Jiang WY, Samai P, Hochschild A, Zhang F, Marraffini LA. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. 2013; 41(15):7429-7437.
23. Huang PM, Pleasance ED, Maydan JS, Hunt-Newbury R, O'Neil NJ, Mah A, Baillie DL, Marra MA. Identification and analysis of internal promoters in *Caenorhabditis elegans* operons. *Genome Research*. 2007; 17(10):1478-1485.
24. Shin JH, Price CW. The SsrA-SmpB ribosome rescue system is important for growth of *Bacillus subtilis* at low and high temperatures. *Journal of bacteriology*. 2007; 189(10):3729-37.
25. Shearwin KE, Callen BP, Egan JB. Transcriptional interference—a crash course. *Trends Genet*. 2005; 21:339-45.
26. Palmer AC, Egan JB, Shearwin KE. Transcriptional interference by RNA polymerase pausing and dislodgement of transcription factors. *Transcription*. 2011; 2(1):9-14.
27. Ouyang S, Lee CY. Transcriptional analysis of type 1 capsule genes in *Staphylococcus aureus*. *Mol. Microbiol*. 1997; 23(3):473-482.
28. Mauro N, Rubio MÁ, Camargo S, Luque L. Regulation of internal promoters in a zinc-responsive operon is influenced by transcription from upstream promoters. *Journal of Bacteriology*. 2013; 195(6):1285-1293.
29. Bird AJ, Gordon M, Eide DJ, Winge DR. Repression of ADH1 and ADH3 during zinc deficiency by Zap1-induced intergenic RNA transcripts. *The EMBO Journal*. 2006; 25(24):5726-5734.
30. Petruk S, Sedkov Y, Riley KM, Hodgson J, Schweisguth F, Hirose S, Jaynes JB, Brock HW, Mazo A. Transcription of bxd Noncoding RNAs Promoted by *Trithorax* Represses Ubx in cis by Transcriptional Interference. *Cell*. 2006; 127(6):1209-1221.
31. Rhee KY, Opel M, Ito E, Hung SP, Arfin SM, Hatfield GW. Transcriptional Coupling between the Divergent Promoters of a Prototypic LysR-Type Regulatory System, the ilvYC Operon of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*. 1999; 96(25):14294-14299.

32. Leng FF, McMacken R. Potent Stimulation of Transcription-Coupled DNA Supercoiling by Sequence-Specific DNA-Binding Proteins. *Proceedings of the National Academy of Sciences of the United States of America*. 2002; 99(14):9139-9144.
33. Sneppen K, Dodd IB, Shearwin KE, Palmer AC, Schubert RA, Callen BP, et al. A mathematical model for transcriptional interference by RNA polymerase traffic in *Escherichia coli*. *J Mol Biol* 2005; 346:399-409.
34. Osato N, Suzuki Y, Ikeo K, Gojobori T. Transcriptional interferences in cis natural antisense transcripts of humans and mice. *Genetics* 2007; 176:1299-306.
35. Kaebernick M, Neilan BA, Börner T, Dittmann E. Light and the transcriptional response of the microcystin biosynthesis gene cluster. *Applied and Environmental Microbiology*. 2000; 66(8):3387-3392.
36. Yoon HS, Golden JW. PatS and products of nitrogen fixation control heterocyst pattern. *Journal of Bacteriology*. 2001; 183(8):2605-2613.

Figures

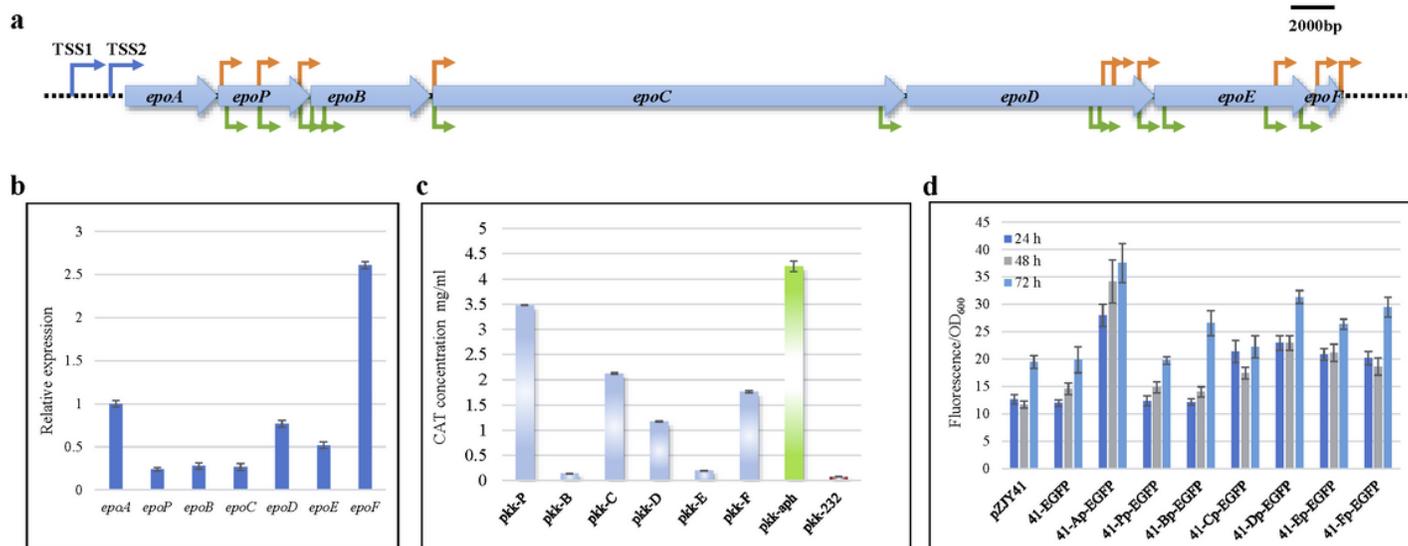


Figure 1

Internal promoters in epothilone operon. (a) Prediction of internal promoters in epothilone operon from So0157-2. The promoter sequence analysis of the epothilone operon was performed using two online promoter prediction softwares. The red arrow is the promoter site predicted by “Neural Network Promoter Prediction”. The green arrow is the promoter site predicted by BPPROM (softberry.com). The threshold was set to 0.8. (b) RT-qPCR analysis of expression levels of the seven genes for epothilone operon in ZE9 at 48h of incubation. The expression of the *epoA* gene was set as 1, and the expressions of the other six genes were shown as the relative expressions to *epoA* gene. (c) The activity of the separate internal promoters in *E. coli*. The activity of the reporter gene chloramphenicol acetyltransferase (CAT) was detected by CAT ELISA kit, with the *aphII* promoter as a positive control, and the original plasmid pKK-232 as a negative control (no promoter upstream of the reporter gene CAT). (d) The activities of the separate internal promoters in *Myxococcus xanthus*. The positive control is the identified promoter *PepoA* upstream of the epothilone operon. The fluorescence intensity of the green reporter gene was detected at

485nm/528nm, and three different incubation times (24h,36h and 72h) were selected for detection. The error bars represent the standard deviation of three independent experiments.

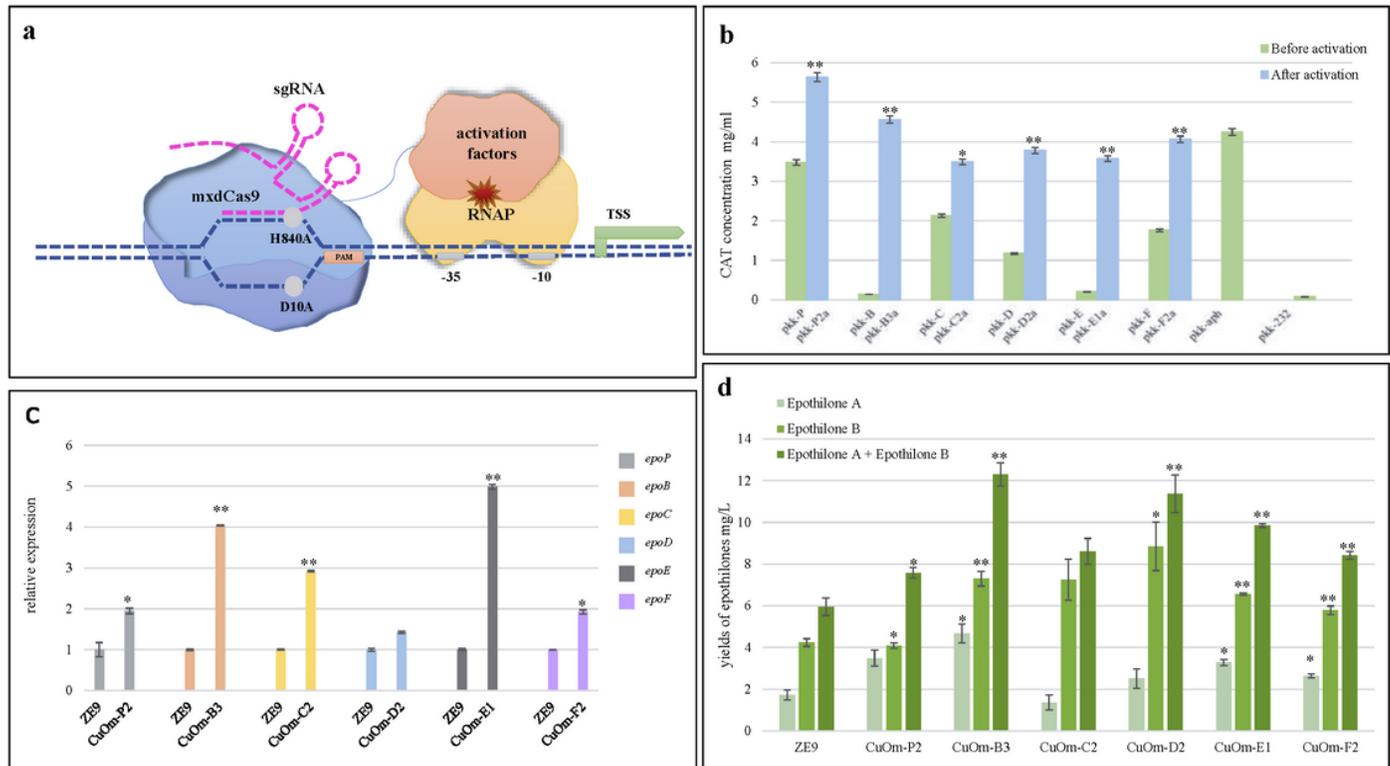


Figure 2

CRISPRa of internal promoters in epothilone operon. (a) The principle of CRISPRa. mxdCas9: codon-optimized inactivated dCas9 protein (H840A, D10A). Activation factors: Transcriptional activation factor, the Omega subunit of RNA polymerase, can play a role in recruiting RNA polymerase. SgrRNA: Single guide RNA sequence, a combination of the CRISPR associated RNA (crRNA) and the trans-activation crRNA (tracrRNA). (b) The activation abilities of the internal promoters in *E. coli*. The activity of the reporter gene CAT was detected by CAT ELISA kit. As the positive control, the aphII promoter was not activated. (c) RT-qPCR analysis of expression levels of the six genes in epothilones operon in different mutants at 48h of incubation. The expressions of the six genes of epothilone operon in *M. xanthus* ZE9 were each set as 1, and the expressions of the genes in mutant strains were shown as the relative expressions to that of their corresponding genes in ZE9. (d) The yield of epothilones A and B and their summation in different activated mutants by CRISPRa and the ancestral strain *M. xanthus* ZE9. The error bars represent the standard deviation of three independent experiments. For statistical analysis between the ancestral strain and mutant strains, the signals of ** and * mean $p < 0.01$ and $p < 0.05$, respectively.

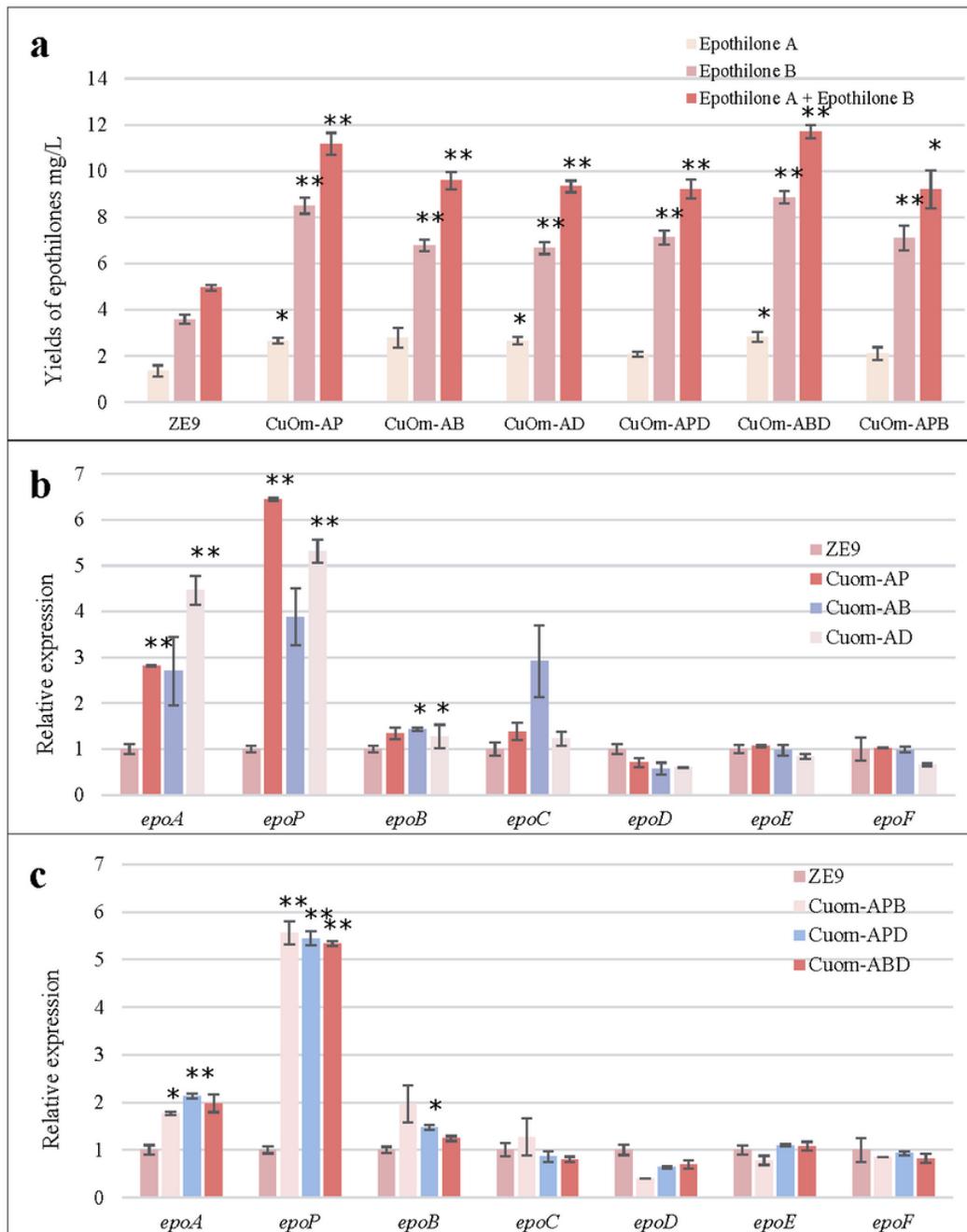


Figure 3

Combinatorial CRISPRa of epothilone operon. (a) The yields of epothilone A and epothilone B in wild strain *M. xanthus* ZE9 and mutants activated by CRISPRa. (b) RT-qPCR analysis of expression levels of the seven epothilone genes in double-activated mutants at 48h of incubation. (c) RT-qPCR analysis of expression levels of the seven epothilone genes in triple-activated mutants at 48h of incubation. The expressions of the seven genes of epothilone operon in *M. xanthus* ZE9 were each set as 1, and the expressions of the genes in mutant strains were shown as the relative expressions to that of their corresponding genes in ZE9. The error bars represent the standard deviation of three independent experiments. For statistical analysis between the ancestral strain and mutant strains, the signals of ** and * mean $p < 0.01$ and $p < 0.05$, respectively.

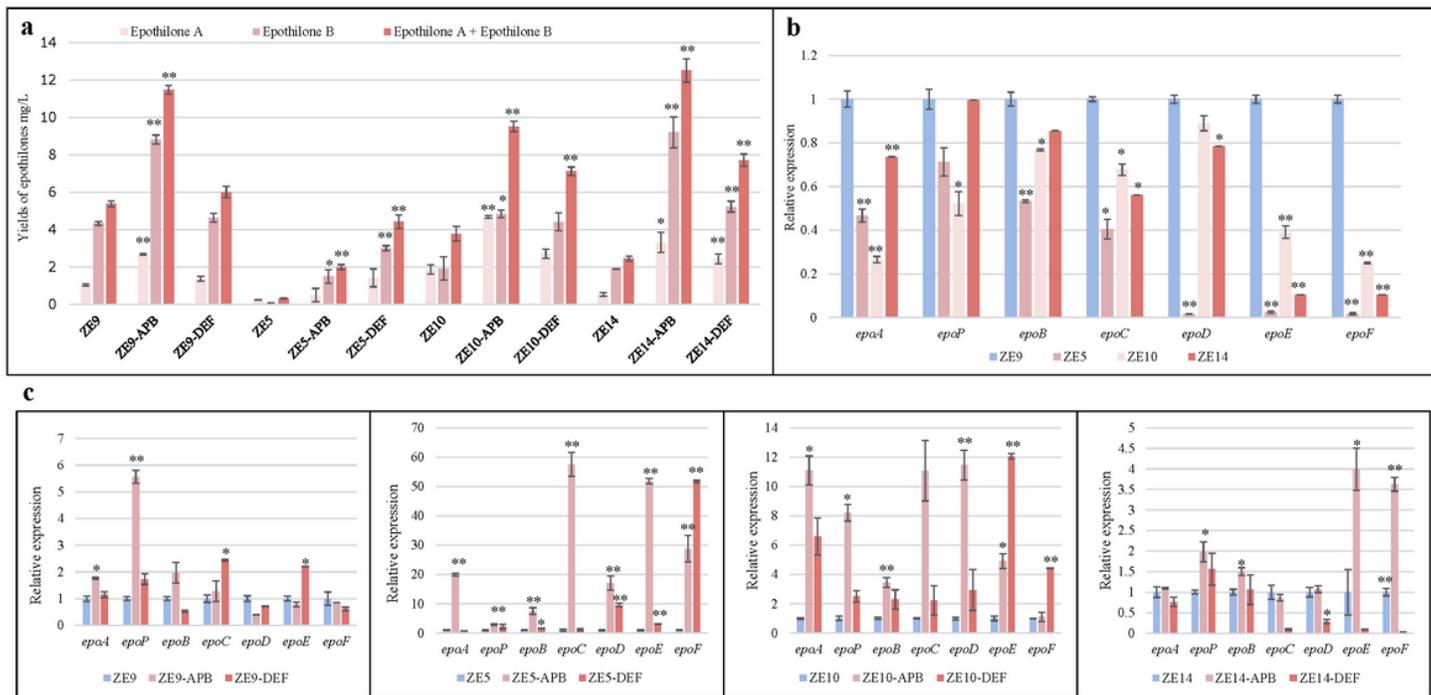


Figure 4

CRISPRa of different epothilone producing strains. (a) The yield of epothilones in wild strains and mutants activated by CRISPRa. (b) RT-qPCR analysis of expression levels of the seven epothilone genes in ZE9, ZE5, ZE10 and ZE14 at 48h of incubation. The expressions of the seven genes of epothilone operon in *M. xanthus* ZE9 were each set as 1, and the expressions of the genes of other strains were shown as the relative expressions to that of their corresponding genes in ZE9. (c) RT-qPCR analysis of expression levels of the seven epothilone genes in activated mutants at 48h of incubation. The expressions of the seven genes of epothilone operon in wild strains were each set as 1, and the expressions of the genes in mutant strains were shown as the relative expressions to that of their corresponding genes in their wild strains, respectively. The error bars represent the standard deviation of three independent experiments. For statistical analysis between the ancestral strain and mutant strains, the signals of ** and * mean $p < 0.01$ and $p < 0.05$, respectively.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1.docx](#)
- [TableS2.docx](#)
- [TableS3.docx](#)
- [TableS4.docx](#)
- [TableS5Primers.docx](#)
- [TableS6Plasmids.docx](#)
- [TableS7Strains.docx](#)

- [figure.S1.jpg](#)
- [figure.S2.jpg](#)
- [figure.S3.jpg](#)
- [figure.S4.jpg](#)
- [figure.S5.jpg](#)
- [figure.S6.jpg](#)
- [figure.S7.jpg](#)
- [figure.S8.jpg](#)
- [figure.S9.jpg](#)