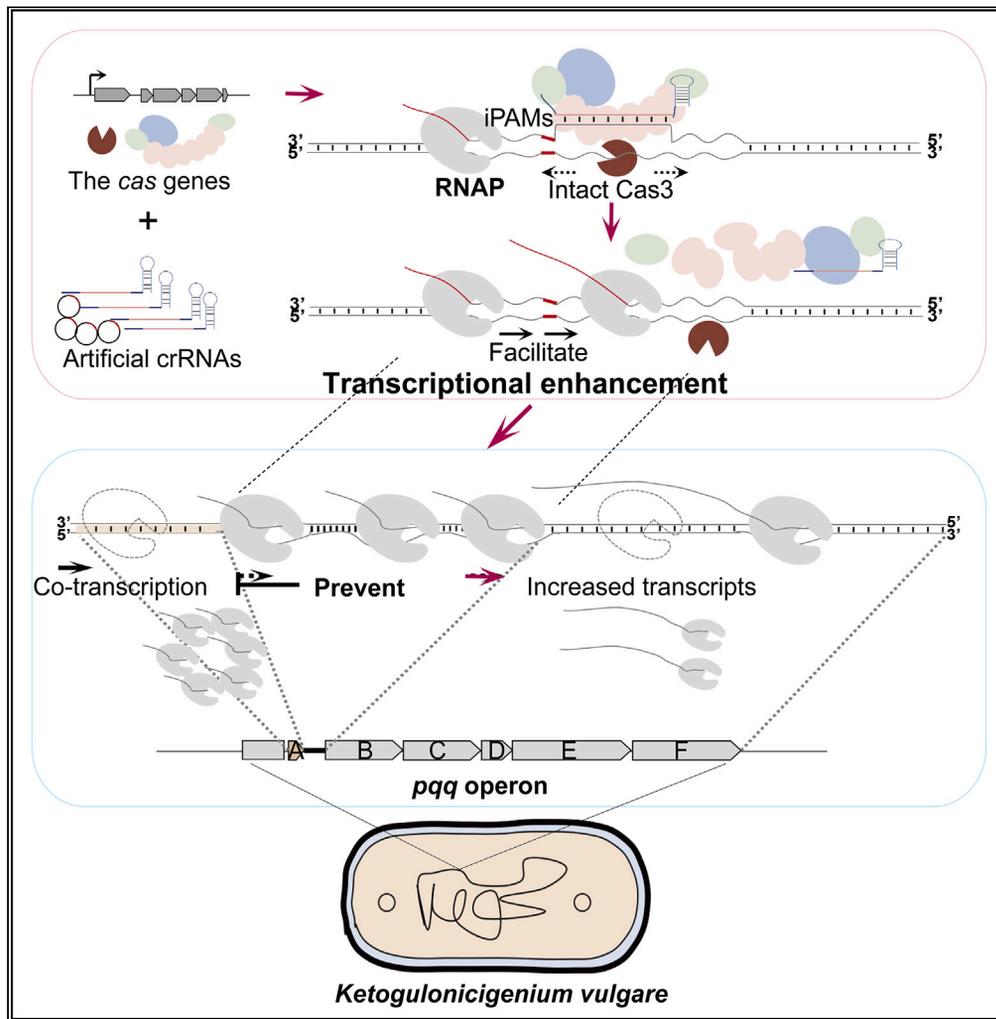


Article

# CRISPRe: An innate transcriptional enhancer for endogenous genes in CRISPR-Cas immunity



Dan Li, Yihong Chen, Fei Huang, Jianmei Wang, Xufeng Li, Yi Yang

xfli@scu.edu.cn (X.L.)  
yangyi528@scu.edu.cn (Y.Y.)

**Highlights**

CRISPR-based transcriptional enhancement is developed in *Ketogulonicigenium vulgare*

CRISPRe promotes gene expression by facilitating transcriptional elongation

An operon regulated by coupling transcriptional-translational elongation is revealed



## Article

CRISPR<sub>e</sub>: An innate transcriptional enhancer for endogenous genes in CRISPR-Cas immunityDan Li,<sup>1,2</sup> Yihong Chen,<sup>1</sup> Fei Huang,<sup>1</sup> Jianmei Wang,<sup>1</sup> Xufeng Li,<sup>1,3,\*</sup> and Yi Yang<sup>1,3,4,\*</sup>

## SUMMARY

CRISPR-Cas system has been repurposed to the promising strategy of CRISPR-based transcriptional interference/activation (CRISPRi/CRISPRa) without eliciting DNA breaks that enables Cas complex a block for transcription initiation or elongation, which greatly expands its application fields and values. However, loss of Cas nuclease ability, especially the endogenous nuclease, may affect genome stability seriously. Here, we found a transcriptional enhancer for genes (CRISPR<sub>e</sub>) in type I-C system of industrial strain *Ketogulonicigenium vulgare* by maintaining the natural activity of Cas3 nuclease and introducing the specific motifs that do not trigger immunity. CRISPR<sub>e</sub> greatly improved the expression of heterologous and endogenous genes and the biosynthesis of products by facilitating transcriptional elongation. Besides, the mechanism for pyrroloquinoline quinone (PQQ) biosynthesis regulated by coupling transcriptional-translational elongation in operon was elucidated. Hence, we enrich the toolbox for CRISPR-Cas system and provide a new framework for gene regulation at transcription.

## INTRODUCTION

CRISPR-Cas systems provide adaptive immunity against invaders in prokaryotes<sup>1–4</sup> and are classified to six distinct types (from I to VI) based on cas genes.<sup>5</sup> In the most familiar type II system, the nuclease Cas9 inactivates foreign genetic elements and involves in RNA-guided targeting as a single-subunit effector.<sup>6</sup> However, for the most prevalent system in nature - type I, Cas3 nuclease and a multi-subunit effector are the main components.<sup>6</sup> Despite different types contain various functional members, they share immunity pathway. In adaptation, spacers with specific lengths from invaders are captured and integrated into CRISPR loci for memory.<sup>7</sup> In expression, CRISPR loci are transcribed and processed to mature RNAs (crRNAs) for target recognition upon the invaders' second irruption.<sup>3</sup> At last stage, crRNAs guide Cas complex to interfere with "protospacer adjacent motifs" (PAMs)-flanking cognate DNA.<sup>8,9</sup> Presence of PAMs implies the accurate cleavage of Cas nuclease at corresponding positions.<sup>9</sup> CRISPR-Cas systems together with PAMs have been repurposed for chromosomal targeting and genomic editing in eukaryotes and prokaryotes.<sup>10</sup> However, to avoid DNA breaks and allow the Cas complex to act as the transcriptional barrier in CRISPR interference (CRISPRi), Cas nuclease has to be artificially deficient.<sup>11</sup> CRISPRi was reported to repress the expression of operons (i.g. *araBAD*, *lacZYA*, *xylAB*, and *rhaBAD*) in the catabolism of sugars.<sup>12</sup> By contrast, fusing an activation domain to "dead" Cas9 nuclease (dCas9) directionally activated  $\sigma$ -dependent promoters to initiate transcription.<sup>13</sup> CRISPRi and CRISPR activation (CRISPRa) have become powerful tools for gene regulation, but the applications are limited due to the severe damage to immunity.<sup>11,14</sup> Apart from direct loss of nuclease activity, we were wondering whether it was feasible to evade nucleases by introducing motifs that do not trigger immunity, not PAMs, in these tools. Recently, crRNAs-guided Cascade complex (CRISPR-associated complex for antiviral defense) has been proved to be capable of binding DNA in PAM-dependent and PAM-independent pathways by three-dimensional diffusion and Watson-Crick pairing,<sup>15</sup> which inspired us to make some nontraditional attempts in CRISPR-derived tools.

Heterologous expression of Cas nuclease in hosts led to off-target mutations, large deletions, and lethal damage, although it enabled programmable genomic editing.<sup>11,16</sup> To develop more effective ways, endogenous type I systems, including I-B, I-E, and I-F, have been harnessed as an alternative strategy to study prokaryotic engineering.<sup>17–19</sup> Type I systems are present in more than 90% of the sequenced bacteria and archaea.<sup>19</sup> Here, we performed the study by industrial strain *Ketogulonicigenium vulgare*,<sup>20–22</sup> which not only has wide applications in L-ascorbic acid synthesis but also harbors a type I-C system, to explore the motifs that do not elicit immunity and addressed the previously described issue. In addition, a constitutively expressed promoter was optimized to detect the consequences of crRNAs-guided Cascade complex in time by the expression of *gfp* gene (green fluorescent protein gene). Surprisingly, we discovered an innate transcriptional enhancer for genes in immunity and defined it as CRISPR<sub>e</sub>, which improved the expression of GFP protein drastically. Pyrroloquinoline quinone (PQQ), the co-enzyme for several dehydrogenases in *K. vulgare*, is important for strains and mammals as a novel redox cofactor after flavin and

<sup>1</sup>Key Laboratory of Bio-resource and Eco-environment of Ministry of Education, College of Life Sciences, State Key Laboratory of Hydraulics and Mountain River Engineering, Sichuan University, Chengdu 610065, China

<sup>2</sup>School of Liquor-making Engineering, Sichuan University Jinjiang College, Meishan 620680, China

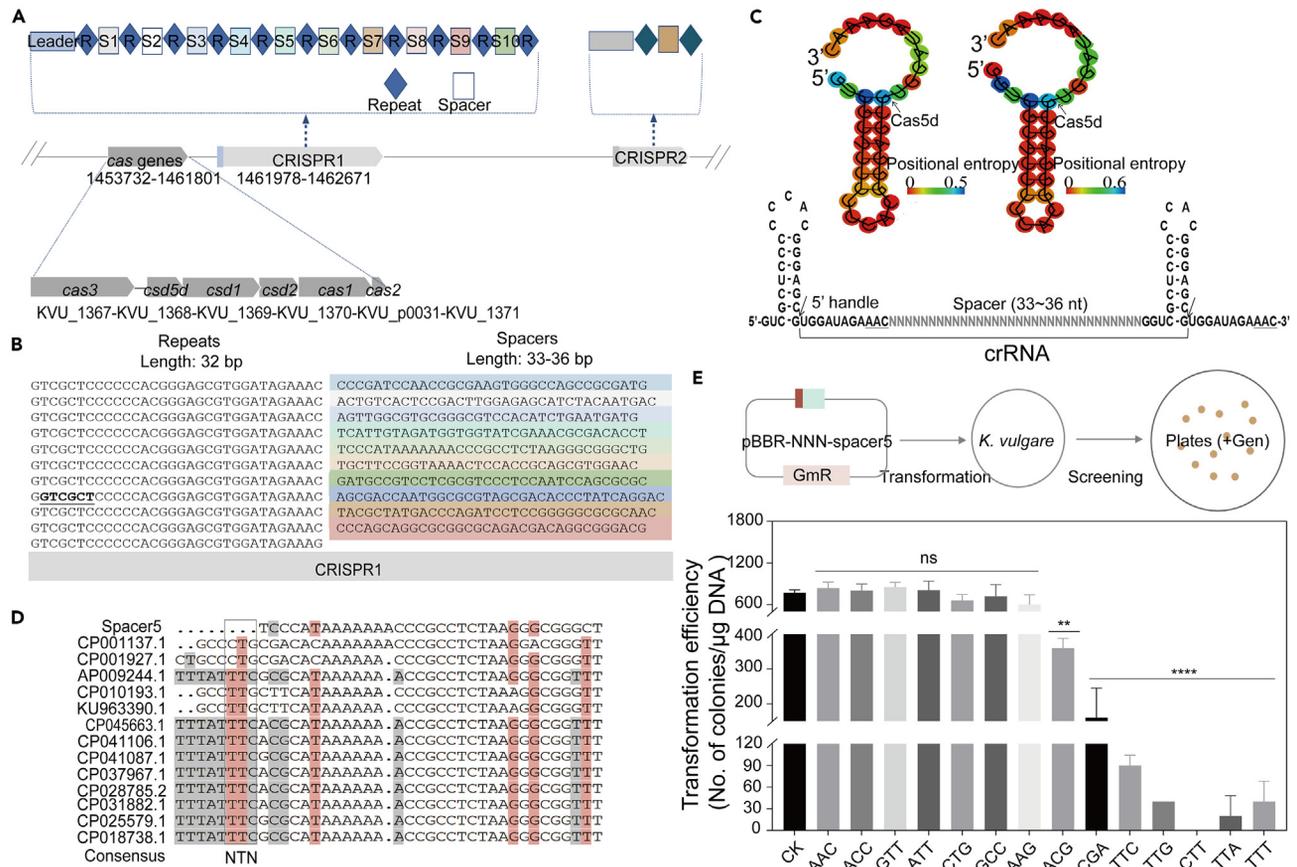
<sup>3</sup>These authors contributed equally

<sup>4</sup>Lead contact

\*Correspondence: xfli@scu.edu.cn (X.L.), yangyi528@scu.edu.cn (Y.Y.)

<https://doi.org/10.1016/j.isci.2023.107814>





**Figure 1. Identification of I-C CRISPR-Cas system in *K. vulgare***

(A) Distribution of cas genes and CRISPR array. Leader: promoter of CRISPR array; R: Repeat; S: Spacer. The positions of cas genes and CRISPR array on genome were identified. See also Figure S1.

(B) Spacers and repeats in CRISPR loci. The distinctive repeat is highlighted in bold.

(C) Repeat structure. Left: most of the repeats; Right: the unique repeat. Pre-crRNA is cleaved by Cas5d<sup>30</sup> at the arrow.

(D) Sequence alignment. The sequences of protospacers from different sources were compared and analyzed. 5'-NTN represents the adjacent conserved motifs immediately upstream the protospacers. Spacer5 indicates "protospacer5" here.

(E) CRISPR interference assay. The conditions that do or do not trigger immunity were determined. Plasmids carrying the potential motifs and protospacer were transformed into *K. vulgare* to observe CRISPR interference. 5'-NNN represents the potential motifs. Spacer5 stands for "protospacer" here. pBBR1MCS-5 was used as the control plasmid (CK). Values are means  $\pm$  s.d. (n = 2 biological replicates). \*\*P < 0.01, \*\*\*P < 0.001, \*\*\*\*P < 0.0001. One-way ANOVA test. ns = non-significant.

nicotinamide.<sup>23–25</sup> Six function-related genes participate in the synthetic pathway and constitute an operon.<sup>26</sup> Due to the differential expression of internal genes within operon ( $\sim$ 1,000-fold), the biosynthesis mechanism is still unknown. We applied CRISPRe to the pathway and found that the synthesis of PQQ is regulated by coupling transcriptional-translational elongation in operon. CRISPRe promoted the expression of genes and the generation of products by facilitating transcriptional elongation in the process. The findings of coupling transcriptional-translational elongation regulation for operons may be helpful to improve Jacob operon theory<sup>26</sup> regulated by co-transcriptional initiation and applicable to a large class of products with significant biological roles. Altogether, we provide a new tool for gene regulation and expand the applications of CRISPR-Cas immunity.

## RESULTS

### Identification of I-C CRISPR-Cas system in *K. vulgare*

To adapt to highly dynamic viral populations,<sup>27</sup> the industrial strain *K. vulgare* WSH-001<sup>28</sup> has evolved its immunity equipped with CRISPR loci and cas genes (Figure 1A). CRISPR1 with 10 spacers was a typical CRISPR array and was immediately downstream of cas genes in the same direction. The cas3 and the other cas genes were presumed to be expressed by two different promoters. Similarly, CRISPR1 had its own leader sequence as a promoter. Spacers of CRISPR1, half of which being in the same size (34 bp), were in lengths of 33–36 bp interrupted by substantially identical

repeats (Figure 1B). The sequence of most repeats was 5'-GTCGCTCCCCACGGGAGCGTGGATAGAAAC-3', whereas the distinctive repeat was 5'-GGTCGCTCCCCACGGGAGCGTGGATAGAAAC-3'. These repeats were analogous to that of the well-studied type I-C system in *Bacillus halodurans*,<sup>18</sup> but differed in three nucleotides on the loop of stem-loop structure predicted by RNAfold WebServer<sup>29</sup> (Figures 1C and S1A). The stem-loop structure here was composed of a 7-bp stem and a 4- or 5-nt loop (Figure 1C). The previous study in *B. halodurans* indicated that pre-crRNAs are processed by Cas5d between G21 and U22 to produce crRNAs.<sup>30</sup> With the roles of Cas5d, the crRNAs in *K. vulgare* were predicted to have the identical 11-nt handle in 5' terminal, although their repeats folded into structure with different loops. Compared with the loop in crRNA, the 5' terminal handle may play an essential role in defining an invading DNA as an attacker.<sup>18</sup> Thus, crRNA-guided Cas complex in *K. vulgare* may provide similar defense mechanism. The cas genes were organized as *cas3-csd5d-csd1-csd2-cas1-cas2* on genome, being architecturally in accord with the reported type I-C system.<sup>31</sup> However, the cas genes did not contain *cas4* that is present in *B. halodurans* (Figure S1B), implying a compact composition in *K. vulgare*. Cas5d, Csd1, and Csd2 proteins have been shown to function as Cascade complex in target identification.<sup>18,29</sup> Conserved Cas1-Cas2 integrase catalyzes new spacer acquisition at the AT-rich leader end of CRISPR loci.<sup>4</sup> And Cas3 is responsible for target cleavage owing to its similar HD nuclease and DEXH/Q-box helicase domains (Figure S1C), even though the amino acid residues among proteins are different.<sup>32,33</sup> Therefore, the type I-C system in *K. vulgare* has typical features in immunity.

### Determination of the motifs that do not trigger immunity

We then attempted to determine the conditions that do not trigger immunity in *K. vulgare*. A 3-bp short sequence flanking protospacer has been verified to the crucial factor in type I-C immunity.<sup>34</sup> 5'-NNN (including NTN) combined with the downstream "protospacer" was regarded as the invading DNA in CRISPR interference assay and spacer5 in CRISPR1 acted as the protospacer (Figures 1D and 1E). Besides "protospacer5," it was not easy to trace the origin of all the highly variable protospacers. 5'-NTN was derived from protospacer5 comparison in GenBank (Figure 1D). CRISPR interference was assessed by transformation efficiency (Figure 1E). The results showed that plasmids carrying the motifs of 5'-TTC/TTG/CTT/TTA/TTT/CGA suffered from CRISPR interference, implying that these motifs are exactly the PAM motifs and cause the neighboring downstream fragments to be targets for attack. PAM was basically represented by 5'-YTN (Y = T/C) and covered 5'-TTC in *B. halodurans*.<sup>35</sup> By contrast, the motifs of 5'-AAC/ACC/GTT/ATT/CTG/GCC/AAG helped the downstream fragments escape from the immunity. Obviously, they were the motifs that we were looking for. To further distinguish from PAM motifs, the motifs that were independent on PAMs and did not trigger CRISPR immunity were referred to as the "iPAM motifs" in this work. Unlike PAMs, we had found no commonalities among iPAMs. However, prior findings revealed that complementarity between CRISPR DNA and crRNA determined CRISPR interference.<sup>36</sup> Although some iPAM motifs enabled base pairing at the indicated positions (-1, -2, and -3), not all iPAM motifs (e.g., 5'-GTT and 5'-CTG) followed the rules (Figure 2A). Similarly, formation of base pairs in PAMs (5'-TTC) did not help escape the fate of being interfered (Figure 2A). Hence, we argued that CRISPR interference might be related to Cas3 nuclease. Precise recognition led to inevitable interference that would otherwise prevent immunity. In addition, we found that 5'-ACC and 5'-AAC happened to be the native motifs located upstream of the spacers in CRISPR loci and formed a part of crRNA in 5' terminal (Figures 1B and 1C), reconfirming the irreplaceable roles of iPAMs in protecting genomic fragments from autoimmunity.

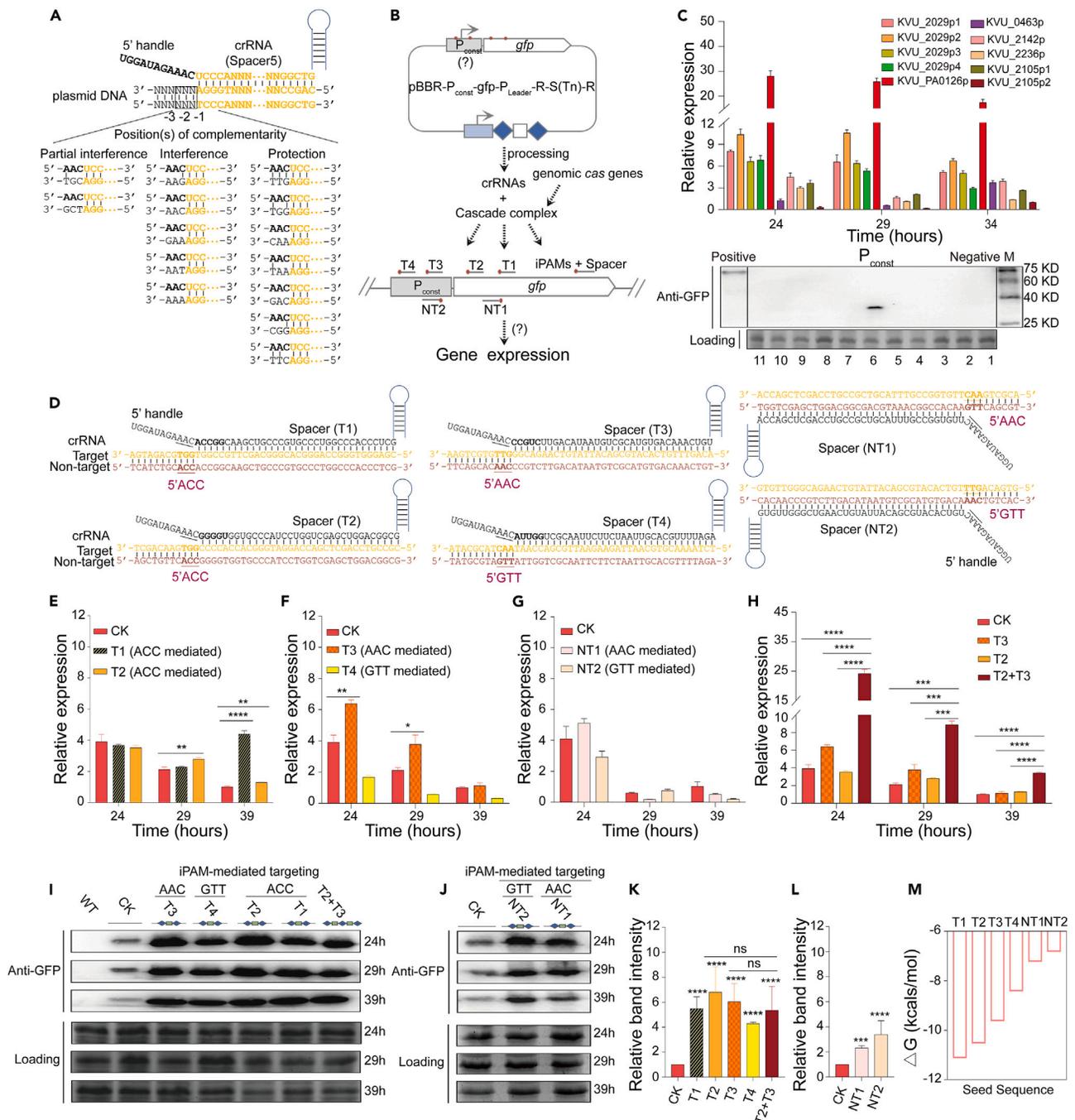
### The new findings of CRISPR in immunity

Next, we explored whether targets successfully avoided the cleavage of Cas3 nuclease by introducing iPAM motifs into CRISPR-Cas-derived tool in *K. vulgare*. Based on the host's endogenous type I-C system, artificial CRISPR loci were recombined to generate specific crRNAs-guided Cascade complex targeting iPAM-flanking targets (Figure 2B). The fragments were placed into a heterologous expressed gene (*gfp*) driven by a constitutive promoter ( $P_{const}$ ) to record the consequences of directional targeting in time through gene expression (Figures 2B and 2C). KUV\_PA0126p was used as the promoter here (Figure 2C). Tn/NTn indicated targeting to different regions of template/non-template strand of DNA, respectively. The iPAM motifs and seed sequence in crRNA were marked in Figure 2D.

As expected, none of iPAM-mediated targeting was subjected to CRISPR interference. However, we also did not observe the inhibition of CRISPR-based system on targets that was originally imagined. More intriguingly, a completely different scene emerged. 5'-ACC-mediated targeting (T1/T2) increased the expression of *gfp* by 6–8 times (Figures 2E, 2I, and 2K). Similarly, 5'-AAC and 5'-GTT showed a 4- to 7-fold enhancement (Figures 2F, 2I, and 2K). Moreover, targeting the non-template strand of DNA exhibited identical results (Figures 2G, 2J, and 2L). The previously described results suggested that both strands of DNA can be targeted to improve expression effectively. Notably, simultaneous targeting led to a more pronounced enhancement on transcriptional level (Figure 2H), which indicated that CRISPR-Cas immunity together with iPAMs plays a role on transcriptional regulation. Collectively, iPAM-mediated targeting was not merely a way to escape CRISPR interference, but a tool to enhance gene expression. The enhancement was found to be governed by the seed sequences in crRNAs (Figure 2M), conforming to the rules that the invaders are interfered with.<sup>37,38</sup> When the seed sequences started with 5'-GGG/CCG and showed a relatively high affinity (T2/T3) (Figure 2M), the effects were particularly prominent (Figure 2I). Furthermore, the enhanced expression was possible by targeting either promoters or coding regions, implying that it may be achieved by promoting transcriptional elongation rather than transcriptional initiation. In addition, the type I-C system in *K. vulgare* remained intact and uncompromised in the process, which suggested that the immunity itself is a natural transcriptional enhancer for genes in the presence of iPAM motifs. To further explore the roles, we defined the transcriptional enhancement in immunity as "CRISPRE" in this study.

### CRISPRE is achieved by facilitating transcriptional elongation

PQQ offers an essential redox cofactor for quinoprotein dehydrogenases that catalyze the oxidation of alcohols and aldose sugars and is beneficial to the growth of strains.<sup>23,24</sup> Besides, PQQ is necessary for the long-term health of mammals as a new redox cofactor after flavin



**Figure 2. The new findings of CRISPR in immunity**

(A) Base pairing between crRNA and target DNA. Base pairing was analyzed at the indicated positions.

(B) Schematic of targeting to the iPAM-flanking fragments on constitutively expressed gene. Artificial CRISPR loci were recombined to produce the corresponding crRNAs. The cas genes in CRISPR-based tools were from genome. Different iPAM motifs were attempted. Tn/NTn represents targeting to different regions of template/non-template strand, respectively.

(C) Promoter screening. The constitutively expressed promoters from genes in *K. vulgare* were identified. The primary screened genes come from mass spectrometry analysis and are represented by their loci on genome. KVVU\_PA0126p drives the stable and efficient expression of *gfp* at transcription and protein (Lane 6) levels and is used as the constitutive promoter (P<sub>const</sub>), which is derived from the gene of amine oxidase flavin-containing protein (gene locus: KVVU\_PA0126).

(D) iPAM-mediated targeting. The iPAM motifs are marked with red bold. Seed sequence in crRNA is marked with black bold.

**Figure 2. Continued**

(E–J) The expression of GFP was measured at different time in strains. WT: wild-type *K. vulgare*, blank control; CK: WT expressed GFP, control; T1/T2/T3/T4/NT1/NT2/T2+T3: CK with the corresponding crRNAs. Values are means  $\pm$  s.d. (n = 3 biological replicates); \*p value < 0.05, \*\*P < 0.01, \*\*\*P < 0.001, \*\*\*\*P < 0.0001. Student's t test. ns = non-significant.

(K–L) Protein abundance. The protein in (I) and (J) was quantitatively analyzed by Image Lab.

(M) Affinity of seed sequences. The affinity of different seed sequences in crRNAs was analyzed.

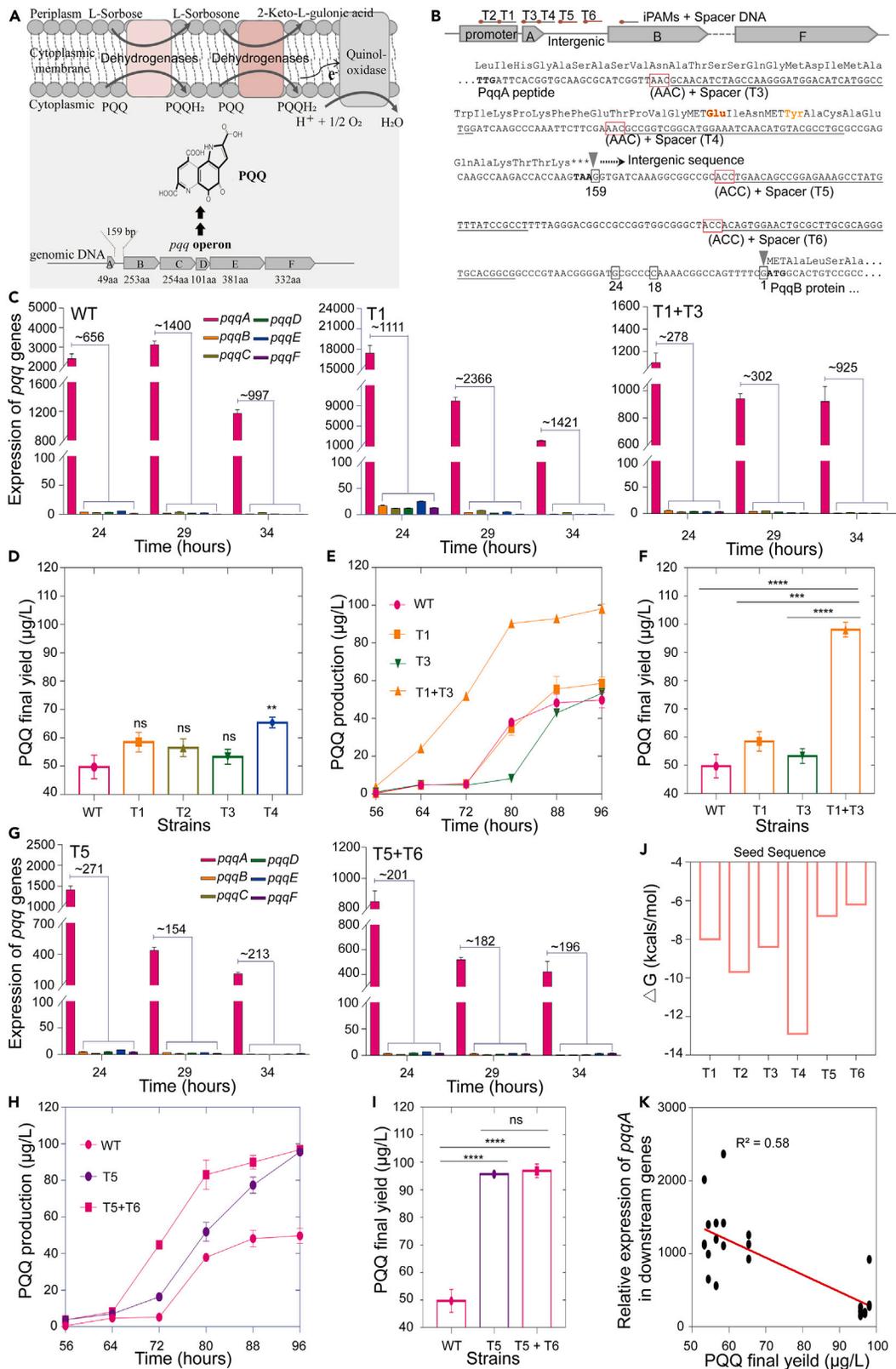
and nicotinamide.<sup>25</sup> The gene cluster responsible for PQQ synthesis in *K. vulgare* involves six genes organized as *pqqA-pqqB-pqqC-pqqD-pqqE-pqqF* (*pqqABCDEF*) on genome (Figure 3A). A 159-bp intergenic region between *pqqA* and *pqqB* divides the cluster into upstream (*pqqA*) and downstream (*pqqB/C/D/E/F*) genes (Figure 3A). Short peptide PqqA was catalyzed and modified to PQQ by proteins encoded by downstream genes.<sup>39–41</sup> RT-PCR, gene deletion, and gene replacement experiments confirmed that the cluster was regulated as an operon<sup>42,43</sup> due to the co-transcription feature even with huge differential expression ( $\sim$ 1,000-fold) (Figures S2A–S2D). In addition, increasing the expression of *pqqA* had no apparent contribution to product synthesis (Figures S3A–S3C), while overexpressing of downstream genes significantly improved the situation (Figure S3D). Hence, we inferred that RNA polymerase (RNAP) might have encountered roadblocks in the operon during the traveling of transcription, thereby limiting the expression of downstream genes and the generation of end products. To assess this point, we applied the newly found transcriptional enhancement tool, CRISPR<sub>e</sub>, to study the operon.

CRISPR<sub>e</sub> targeted different locations (including the sharing promoter, the gene body of *pqqA*, and the intergenic sequence) in the operon and the motifs used were native iPAM motifs (Figures 3B and S4). Under normal condition, the expression of *pqqA* at any time was  $\sim$ 1,000-fold that of downstream genes (Figure 3C). When CRISPR<sub>e</sub> targeted upstream of the operon (T1/T2/T3/T4), the transcriptional difference between *pqqA* and downstream genes was further expanded (Figures 3C and S5A–S5C) and the final yield of PQQ products was only increased by 20%–30% ( $\sim$ 65  $\mu$ g/L) (Figure 3D). It indicated that the targeting locations did not obviously prevent the elongation of RNAP. Interestingly, when single-site targeting by CRISPR<sub>e</sub> appeared at intergenic region, the transcriptional difference between *pqqA* and downstream genes was drastically reduced (Figure 3G) and the generation of products was significantly improved (Figures 3H and 3I), though the seed sequence in crRNA (T5) showed an extremely lower affinity compared with other crRNAs (Figure 3J). The final yield of products was twice that of wild-type strains and was maintained at  $\sim$ 100  $\mu$ g/L. Moreover, simultaneously targeting this region (T5+T6) further improved the synthesis efficiency (Figure 3H). The results suggested that the intergenic region acted as the barriers to hinder the transcription of downstream genes and prevent the elongation of RNAP, while CRISPR<sub>e</sub> promoted the expression of genes by facilitating transcriptional elongation. Besides, targeting two locations upstream of the operon (T1+T3) by CRISPR<sub>e</sub> with higher affinity crRNAs also achieved remarkable results (Figures 3C, 3E, and 3F). Based on the previously described data, the applications of CRISPR<sub>e</sub> (without negative effects to hosts) (Figure S5D) were helpful to find the positive correlation between the expression of downstream genes and the synthesis of products (Figure 3K), further supporting the idea that the *pqq* operon limits the generation of end products by regulating the elongation of RNAP.

Additionally, reducing the barriers on intergenic region also promoted transcriptional elongation (Figures 4A–4G). However, the lower synthesis of PQQ in artificial operon *pqqA(18NNN)BCDEF* than that of *pqqA(24NNN)BCDEF* implied that the operon had evolved complex regulatory means at translational level (Figures 4H and 4I). To prove this point, we generated the artificial units “p/sp+gfp” and “p/sp+pqqA(nNNN)gfp” by replacing the genes in *pqq* operon with *gfp* (Figure 5A). The results showed that the intergenic region further diminished the expression of downstream genes by tens of fold at translational level without great alteration in transcription (Figures 5B and 5C). However, the identical transcription of *gfp* (Figure 5B) indicated that *pqqA* provides a pleasant journey for RNAP due to its enrichment in AT nucleotides (Figures S6A and S6B). Subsequently, we found that multiple Shine-Dalgarno (SD)-like sequences, 5'-GCGCYG (Y=C/T) and 5'-GMCGMCG (M=C/A), were present in intergenic region (Figures 5D–5J, S6C and S6D). Identical gene transcription and similar RNA structure (Figures 5E and 5F) revealed that they were not the factors to inhibit gene translation. Smooth translation of GFP protein in *pqqA(-1NNN)gfp* (Figure 5G) showed that the terminated but not released ribosomes did not separate from transcription products after completing the synthesis of PqqA peptide, but slid forward to reach an adjacent initiation codon and began the translational re-initiation. Hence, the SD-like sequences in intergenic region may delay the translational re-initiation by their high-affinity with ribosomes, thus decreasing the translation of downstream genes. Taken together, the *pqq* operon regulates the expression of internal genes by coupling transcriptional-translational elongation rather than only by co-transcriptional initiation to strictly limit the synthesis of end products (Figure 6A). The regulatory mechanism may be applicable to the natural products of ribosomally synthesized and posttranslationally modified peptides that play significant biological roles.<sup>39</sup> Also, the findings of coupling transcriptional-translational elongation regulation may be helpful to improve Jacob operon theory that a single polycistronic mRNA molecule is produced by co-transcription.<sup>26,42</sup>

**CRISPR<sub>e</sub> is dependent on both iPAMs and active Cas3**

We found the transcriptional enhancement tool by introducing iPAM motifs and maintaining complete Cas members (including Cas3) in immunity. In addition to Cas3 nuclease, Cas5d, Csd1, and Csd2 in type I-C system have been found to constitute Cascade complex that is involved in target recognition under the guidance of crRNA.<sup>18</sup> Cas3 is recruited to perform nuclease activity in the presence of PAM motifs.<sup>18</sup> In *K. vulgare*, 5'-YTN (Y=T/C) was used to elicit the cleavage of Cas3 at targeted positions to obtain mutants required for this study. However, introduction of iPAM motifs flanking targets disrupted the phenomenon, making Cas3 fail to cleave at the predicted positions. Instead, Cas3 may translocate along iPAM-flanking targets to unwind double-stranded DNA under another domain - helicase, thereby facilitating the proceeding for working RNAP (Figure 6B). This is the proposed model for CRISPR<sub>e</sub>. Targeting any strand of DNA leading to efficient



**Figure 3. CRISPR promotes the synthesis of PQQ by enhancing the expression of downstream genes in operon**

(A) Schematic of the roles of PQQ in *K. vulgare*. The uncharacterized gene (locus: KVU\_2026) adjacent to *pqqE* was named as *pqqF*. The long intergenic region divides the operon to up (*pqqA*) and downstream (*pqqB/pqqC/pqqD/pqqE/pqqF*) genes. The number of amino acids encoded by each gene is shown. PQQ provides an essential redox cofactor for dehydrogenases and is ultimately converted into energy through the oxidative pathway in *K. vulgare*<sup>23,39</sup>.

(B) The targets and partial sequences in *pqq* operon. Tn represents the iPAM-flanking targets in operon. The sequences of *pqqA* and long intergenic region in operon are indicated. The targeted fragments located within *pqqA* gene body and intergenic sequence are underlined. The iPAM motifs are in red box. Termination and initiation codons are marked in black bold. Numbers show the distance between the stop and the restart site (in nucleotides). Orange marked amino acid residues are the backbone sources of PQQ products.<sup>39,40</sup> See also Figure S4.

(C–F) Gene expression and product synthesis at different time in strains. The transcriptional differences between upstream and downstream genes in operon are analyzed and shown (C). PQQ synthesis in strains is shown in (D), (E), (F). T1/T2/T3/T4/T1+T3: WT with the corresponding crRNAs. WT was the control. Values are means  $\pm$  s.d. (n = 3 biological replicates); \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ . Student's *t* test. ns = non-significant. See also Figures S5A–S5C.

(G–I) Gene expression and product synthesis in strains carrying crRNA (T5/T5+T6). The data analysis was consistent with (C–F).

(J) Affinity of seed sequences.

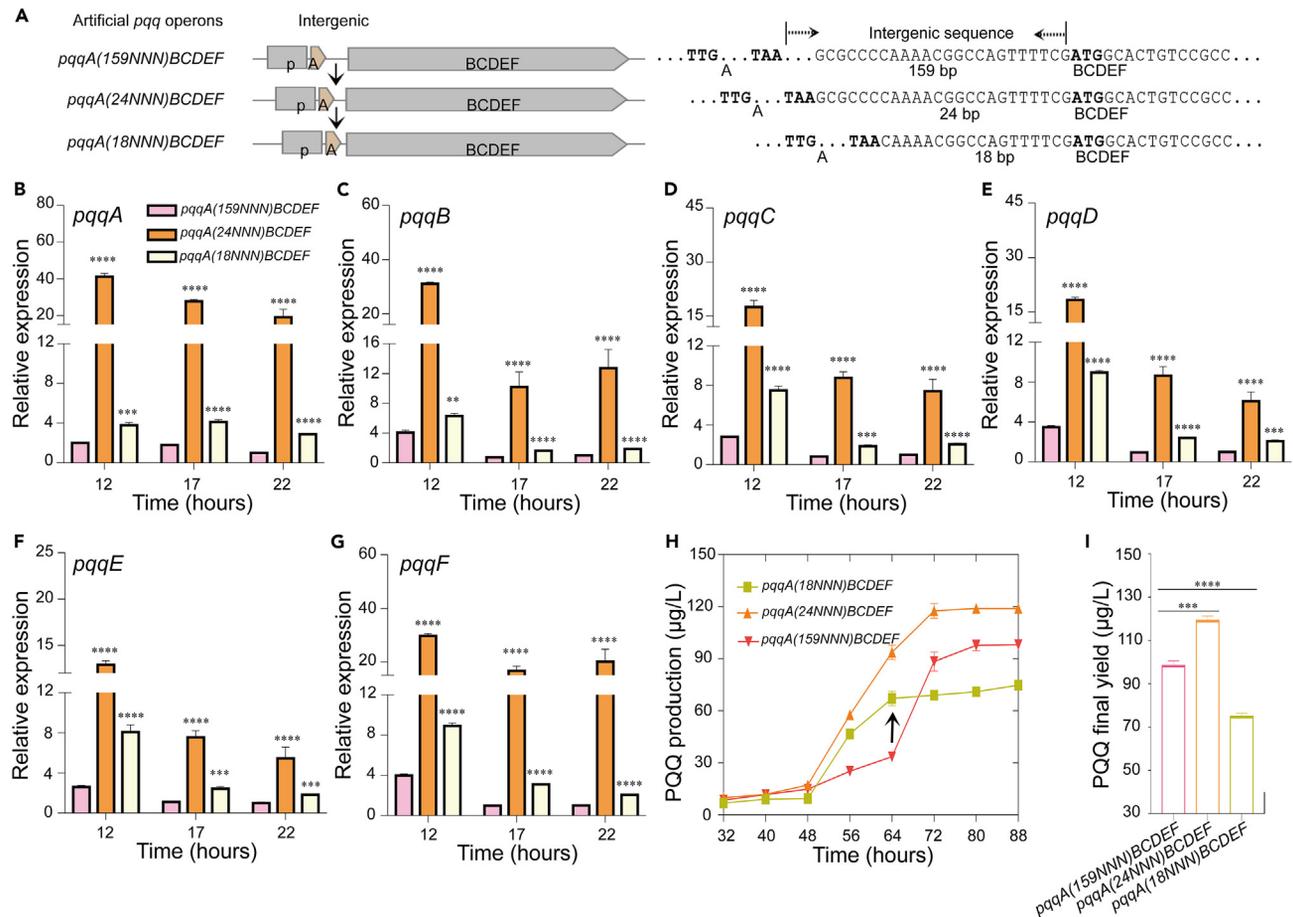
(K) Correlation between gene expression and product synthesis. The correlation between gene expression (relative expression of *pqqA* in downstream genes) and PQQ synthesis in different strains was analyzed under the influence of CRISPRs.

transcriptional enhancement also implies that Cas3 moves in either direction. The proposal is supported by the study of surveillance and processing of foreign DNA in *E. coli* with type I-E system, which has demonstrated that Cas3 is recruited to target DNA in PAM-dependent and PAM-independent pathways and that Cas3 attenuates its nuclease activity and travels along DNA in PAM-independent processing.<sup>15,44</sup> Besides, it is also in accordance with the conclusion that GC-rich sequences are more likely to be the obstacles for RNAP due to the difficulty in unwinding DNA strands (Figures S6A and S6B). It is critical to keep Cas3 nuclease domain intact in CRISPR. Premature termination of Cas3 in *Gluconobacter oxydans* repurposed the endogenous type I-E system to be a tool for gene repression.<sup>29</sup> Deletion/absence of Cas3 in *E. coli* also led to gene repression.<sup>11,12,34</sup> Here, CRISPR does no damage to Cas3 in *K. vulgare* and modulates gene expression in an entirely new pathway. Targeting promoters, open reading frames, or intergenic regions in operon can enhance gene expression by promoting RNAP elongation. Therefore, we discover a tool for gene regulation in this study.

## DISCUSSION

In this work, we tried to make some different attempts in CRISPR-Cas-derived tools. Amazingly, we discovered a transcriptional elongation tool based on the endogenous type I-C immunity in *K. vulgare*, which is of great significance for gene regulation. For years, transcriptional initiation was regarded as the main step of gene regulation, followed by monotonous and uneventful transcriptional elongation. CRISPRa aims to activate promoters to initiate transcription.<sup>14</sup> For the *lac* operon in *E. coli*, complex regulation is also focused on transcriptional initiation to control the synthesis of a single polycistronic mRNA product.<sup>26,42</sup> However, for the *pqq* operon in *K. vulgare*, no complex transcriptional initiation switch has yet been found<sup>45</sup> (Figure 3C), which may be due to the positive effects of PQQ products on strain growth (Figures S5 and S7). The transcriptional elongation in this operon is uneven and perturbed by intrinsic regulatory barriers (GC-rich intergenic region) to synthesize the desired products by restricting the expression of downstream genes (Figure 6A). Thus, gene regulation is far complicated than we have expected. Surprisingly, in addition to activating transcriptional initiation, gene expression can be raised from promoting transcriptional elongation. When employed for *gfp* expression, CRISPRa achieved 2- to 8-fold enhancement (Figure 2I). When applied to PQQ synthesis, CRISPRa shortened the generation time by dozens of hours and increased the yield to ~2-fold (Figures 3E, 3F, 3H, and 3I). The enhanced effects depended on crRNA affinity, targeting location, and targeting way. However, the expression at transcriptional and translational levels was different. The over-enhanced transcription in turn affected protein translation and product synthesis, leading to the limited increase in our study (Figures 2H, 2I, 3H, and 3I). More importantly, CRISPRa helps to reveal a new regulatory model for operons, although how the excessive end products feedback inhibits the expression of genes is unclear. Altogether, CRISPRa can be used in gene expression, product synthesis, and regulatory mechanism analysis.

The exploration of accurate iPAM motifs is crucial for CRISPRa applications. The iPAM motifs may help to initiate an alternative pathway of immunity other than target degradation and enable Cas3 to exhibit helicase function in DNA unwinding. However, whether Cas3 recruitment upon introduction of iPAM motif in *K. vulgare* requires the assistance of Cas1-Cas2 as in *E. coli*<sup>15</sup> and how the activity of Cas3 nuclease is attenuated to avoid DNA breaks at the target DNA are still the issues to be resolved. Nevertheless, we have recorded the effects of Cas3 recruitment in time by using gene expression and product generation and discovered the built-in property of endogenous immunity. Coincidentally, the native iPAM motifs, located at the end of 5' handle in crRNA, are generally conserved, e.g., 5'-AMC (M = A/C) in type I-C (*K. vulgare*), 5'-AAA in type I-F (*Pseudomonas aeruginosa* and *Zymomonas mobilis*),<sup>18,19</sup> and 5'-CCG in type I-E (*E. coli*, *G. oxydans*, and *Thermus thermophilus*).<sup>18,29</sup> And the transcriptional enhancement may be the driving force. The crRNA-guided Cascade complex binding to CRISPR DNA was once regarded as a form of self-protection from autoimmunity - "self-identification".<sup>36,46</sup> This part-time work of crRNA is an intermediate process in innate immunity and is difficult to detect. The findings of transcriptional enhancement have revealed an incredible design in nature immunity. Upstream genes in the co-transcriptional unit always benefiting from CRISPRa due to the natural locations help to explain why newly captured protospacer is integrated at the position of first spacer.<sup>7</sup> Specific crRNAs with a low proportion in total crRNAs were proved to bind to invading DNA and CRISPR DNA equally by base pairing,<sup>2,15</sup> which increases the difficulty of target acquisition and interference. However, iPAM-mediated binding makes us insist that sufficient amounts of crRNAs would be produced in a short time. The remaining crRNAs (except ~50% specific crRNAs) have to bind to iPAM-flanking CRISPR DNA for rapid burst of more crRNAs, thereby



**Figure 4. Reducing the roadblocks on intergenic region improves the expression of downstream genes in operon**

(A) Schematic of artificial *pqq* operon construction. Artificial *pqq* operons driven by its own promoter were recombined by shortening intergenic sequences. Numbers show the length of intergenic region (in nucleotides). The plasmids containing artificial operons were transferred into wild-type strains. The gene expression and product generation in strains were examined, respectively. In *pqqA*(*n*NNN)*BCDEF*, *n*NNN represents the nucleotide number and original nucleotide arrangement in intergenic sequences, respectively.

(B–I) Gene expression and PQQ synthesis in strains harboring artificial *pqq* operons. *pqqA*(*n*NNN)*BCDEF* here indicates the strains harboring operons. *pqqA*(159NNN)*BCDEF* (= *pqqABCDEF*) was the control. Values are means  $\pm$  s.d. ( $n = 3$  biological replicates); \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ . Student's *t* test.

triggering subsequent efficient interference. Thus, besides storing adaptive memory, CRISPR DNA serves as a target for crRNA. In addition, crRNAs may bind competitively to regulate unlimited transcriptional enhancement. Although we lack the evidence described previously for perfect immunity, CRISPRE provides a framework for immediate practical applications and a direction for further study.

### Limitations of the study

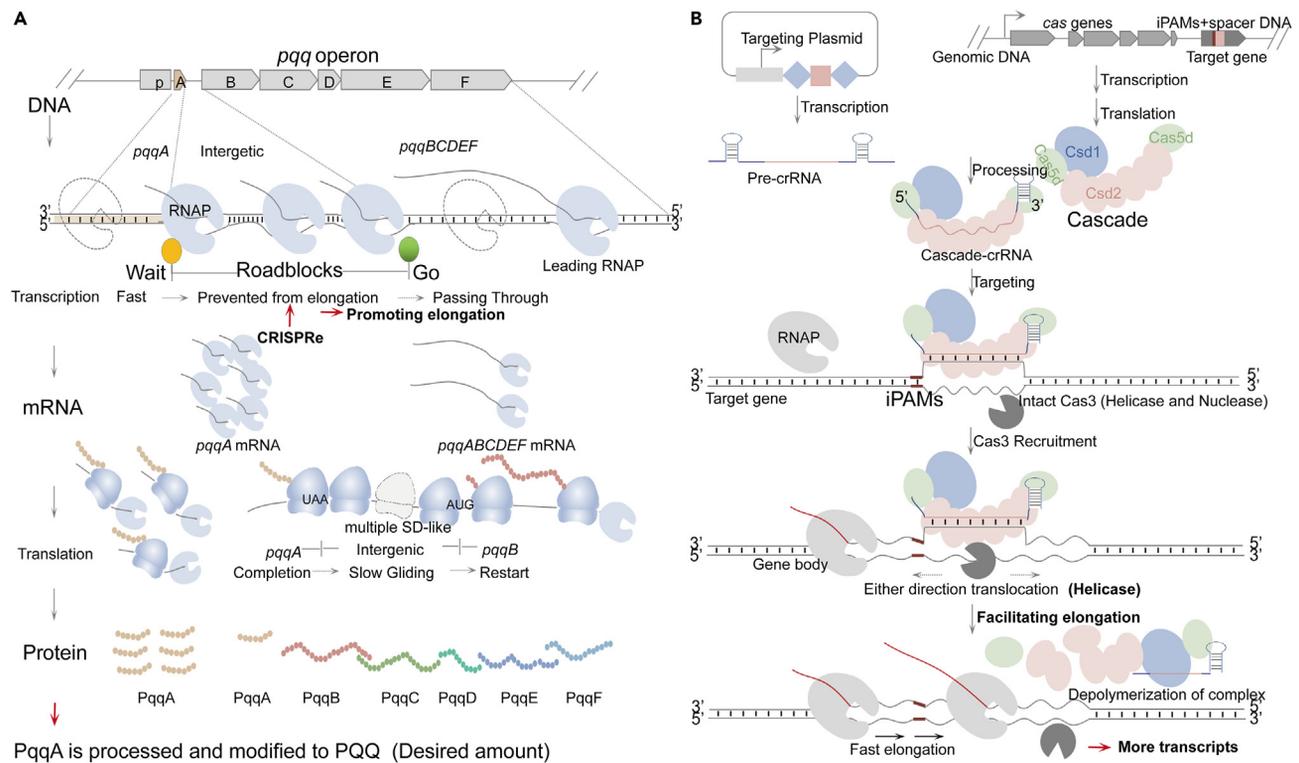
Presence of iPAM motifs avoids DNA breaks at the desired positions and enables Cas3 to translocate away from Cascade complex along targets to unwind double-stranded DNA, which provides possibility for developing a CRISPR-based tool for transcriptional enhancement in *K. vulgare*. However, how the nuclease/helicase domains of Cas3 are coordinated and when the activity of Cas3 nuclease is attenuated remain unknown. In addition, it is unclear whether iPAM-mediated Cas3 recruitment needs the help of other Cas proteins. So, more work for the mechanistic characterization of CRISPRE would be necessary.

### STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY





**Figure 6. CRISPRe works by promoting transcriptional elongation**

(A) CRISPRe facilitates RNAP elongation. Unlike ordinary operons where the complex regulation is focused on transcriptional initiation to activate promoters, the *pqq* operon in *K. vulgare* regulates gene expression by providing an uneven travel for RNAP. Particularly, the GC-rich intergenic region, located immediately downstream AT-rich *pqqA* gene, acts as the roadblocks that prevent the elongation of RNAP and reduces the expression of downstream genes in operon. CRISPRe, however, can significantly promote RNAP elongation and PQQ synthesis by directional targeting. Hence, CRISPRe plays a role by facilitating RNAP elongation. Besides, the *pqq* operon has evolved a ribosomal coupling pathway. Multiple SD-like sequences are placed on intergenic region to slow down the translational re-initiation by their high affinity with ribosomes and further decrease the translation of downstream genes, thus limiting the synthesis of end products.

(B) The proposed model for CRISPRe. Presence of iPAM motifs prevents Cas3 from cleaving at the desired positions and the targets successfully escape interference from CRISPR-Cas immunity. However, Cas3, relying on its helicase domain, can still translocate away from Cascade complex along targets to unwind double-stranded DNA, thereby facilitating RNAP elongation. Hence, more transcripts are generated with the roles of CRISPRe.

● **METHOD DETAILS**

- CRISPR-cas system identification
- CRISPR interference assay
- DNA cloning and plasmid construction
- RT-PCR and quantitative RT-PCR
- Western blot analysis
- PQQ determination assay

● **QUANTIFICATION AND STATISTICAL ANALYSIS**

**SUPPLEMENTAL INFORMATION**

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2023.107814>.

**ACKNOWLEDGMENTS**

We thank Prof. Yan Liu and Dr. Linli Han (Sichuan University) for their technical help. This work was supported by the National Natural Science Foundation of China [31870240] and the Natural Science Foundation of Sichuan Province [2022NSFSC0162].

**AUTHOR CONTRIBUTIONS**

Y.Y., X.L., and D.L. conceived and supervised the project. D.L., Y.C., and F.H. performed the studies. D.L., F.H., J.W., Y.Y., and X.L. analyzed the data. D.L., Y.Y., and X.L. wrote the manuscript. Y.Y. and X.L. revised the manuscript and provided funding.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 26, 2023

Revised: April 20, 2023

Accepted: August 30, 2023

Published: September 1, 2023

## REFERENCES

- Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60, 174–182. <https://doi.org/10.1007/s00239-004-0046-3>.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712. <https://doi.org/10.1126/science.1138140>.
- Marraffini, L.A. (2015). CRISPR-Cas immunity in prokaryotes. *Nature* 526, 55–61. <https://doi.org/10.1038/nature15386>.
- Núñez, J.K., Harrington, L.B., Kranzusch, P.J., Engelman, A.N., and Doudna, J.A. (2015). Foreign DNA capture during CRISPR-Cas adaptive immunity. *Nature* 527, 535–538. <https://doi.org/10.1038/nature15760>.
- Mohanraju, P., Makarova, K.S., Zetsche, B., Zhang, F., Koonin, E.V., and van der Oost, J. (2016). Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* 353, aad5147-10. <https://doi.org/10.1126/science.aad5147>.
- Makarova, K.S., Wolf, Y.I., Iranzo, J., Shmakov, S.A., Alkhnbashi, O.S., Brouns, S.J.J., Charpentier, E., Cheng, D., Haft, D.H., Horvath, P., et al. (2020). Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* 18, 67–83. <https://doi.org/10.1038/s41579-019-0299-x>.
- Núñez, J.K., Lee, A.S.Y., Engelman, A., and Doudna, J.A. (2015). Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature* 519, 193–198. <https://doi.org/10.1038/nature14237>.
- Shimori, M., Garrett, S.C., Chambers, D.P., Glover, C.V.C., 3rd, Graveley, B.R., and Terns, M.P. (2017). Role of free DNA ends and protospacer adjacent motifs for CRISPR DNA uptake in *Pyrococcus furiosus*. *Nucleic Acids Res.* 45, 11281–11294. <https://doi.org/10.1093/nar/gkx839>.
- Hayes, R.P., Xiao, Y., Ding, F., van Erp, P.B.G., Rajashankar, K., Bailey, S., Wiedenheft, B., and Ke, A. (2016). Structural basis for promiscuous PAM recognition in type I-E Cascade from *E. coli*. *Nature* 530, 499–503. <https://doi.org/10.1038/nature16995>.
- Zhang, S., Guo, F., Yan, W., Dai, Z., Dong, W., Zhou, J., Zhang, W., Xin, F., and Jiang, M. (2019). Recent Advances of CRISPR/Cas9-Based Genetic Engineering and Transcriptional Regulation in Industrial Biology. *Front. Bioeng. Biotechnol.* 7, 459–511. <https://doi.org/10.3389/fbioe.2019.00459>.
- Vigouroux, A., and Bikard, D. (2020). CRISPR Tools To Control Gene Expression in Bacteria. *Microbiol Mol Biol R 84*, e00077-19. <https://doi.org/10.1128/MMBR.00077-19>.
- Luo, M.L., Mullis, A.S., Leenay, R.T., and Beisel, C.L. (2015). Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression. *Nucleic Acids Res.* 43, 674–681. <https://doi.org/10.1093/nar/gku971>.
- Fontana, J., Dong, C., Kiattisewee, C., Chavali, V.P., Tickman, B.I., Carothers, J.M., and Zalatan, J.G. (2020). Effective CRISPR-mediated control of gene expression in bacteria must overcome strict target site requirements. *Nat. Commun.* 11, 1618. <https://doi.org/10.1038/s41467-020-15454-y>.
- Liu, Y., Wan, X., and Wang, B. (2019). Engineered CRISPRa enables programmable eukaryote-like gene activation in bacteria. *Nat. Commun.* 10, 3693. <https://doi.org/10.1038/s41467-019-11479-0>.
- Redding, S., Sternberg, S.H., Marshall, M., Gibb, B., Bhat, P., Guegler, C.K., Wiedenheft, B., Doudna, J.A., and Greene, E.C. (2015). Surveillance and Processing of Foreign DNA by the *Escherichia coli* CRISPR-Cas System. *Cell* 163, 854–865. <https://doi.org/10.1016/j.cell.2015.10.003>.
- Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 31, 822–826. <https://doi.org/10.1038/nbt.2623>.
- Hidalgo-Cantabrana, C., Goh, Y.J., Pan, M., Sanozky-Dawes, R., and Barrangou, R. (2019). Genome editing using the endogenous type I CRISPR-Cas system in *Lactobacillus crispatus*. *Proc. Natl. Acad. Sci. USA* 116, 15774–15783. <https://doi.org/10.1073/pnas.1905421116>.
- Zheng, Y., Li, J., Wang, B., Han, J., Hao, Y., Wang, S., Ma, X., Yang, S., Ma, L., Yi, L., et al. (2020). Endogenous Type I CRISPR-Cas: From Foreign DNA Defense to Prokaryotic Engineering. *Front. Bioeng. Biotechnol.* 8, 62. <https://doi.org/10.3389/fbioe.2020.00062>.
- Zheng, Y., Han, J., Wang, B., Hu, X., Li, R., Shen, W., Ma, X., Ma, L., Yi, L., Yang, S., et al. (2019). Characterization and repurposing of the endogenous Type I-F CRISPR-Cas system of *Zymomonas mobilis* for genome engineering. *Nucleic Acids Res.* 47, 11461–11475. <https://doi.org/10.1093/nar/gkz940>.
- Pappenberger, G., and Hohmann, H.P. (2014). Industrial production of L-ascorbic Acid (vitamin C) and D-isoascorbic acid. *Adv. Biochem. Eng. Biotechnol.* 143, 143–188. [https://doi.org/10.1007/10\\_2013\\_243](https://doi.org/10.1007/10_2013_243).
- Takagi, Y., Sugisawa, T., and Hoshino, T. (2010). Continuous 2-Keto-L-gulonol acid fermentation by mixed culture of *Ketogulonicigenium vulgare* DSM 4025 and *Bacillus megaterium* or *Xanthomonas maltophilia*. *Appl. Microbiol. Biotechnol.* 86, 469–480. <https://doi.org/10.1007/s00253-009-2312-1>.
- Zou, W., Liu, L., Zhang, J., Yang, H., Zhou, M., Hua, Q., and Chen, J. (2012). Reconstruction and analysis of a genome-scale metabolic model of the vitamin C producing industrial strain *Ketogulonicigenium vulgare* WSH-001. *J. Biotechnol.* 161, 42–48. <https://doi.org/10.1016/j.jbiotec.2012.05.015>.
- Miyazaki, T., Sugisawa, T., and Hoshino, T. (2006). Pyrroloquinoline quinone-dependent dehydrogenases from *Ketogulonicigenium vulgare* catalyze the direct conversion of L-sorbosone to L-ascorbic acid. *Appl. Environ. Microbiol.* 72, 1487–1495. <https://doi.org/10.1128/AEM.72.2.1487-1495.2006>.
- Salisbury, S.A., Forrest, H.S., Cruse, W.B., and Kennard, O. (1979). A novel coenzyme from bacterial primary alcohol dehydrogenases. *Nature* 280, 843–844. <https://doi.org/10.1038/280843a0>.
- Kasahara, T., and Kato, T. (2003). Nutritional biochemistry: A new redox-cofactor vitamin for mammals. *Nature* 422, 832. <https://doi.org/10.1038/422832a>.
- Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3, 318–356. [https://doi.org/10.1016/s0022-2836\(61\)80072-7](https://doi.org/10.1016/s0022-2836(61)80072-7).
- Bergh, O., Borsheim, K.Y., Bratbak, G., and Haldal, M. (1989). High abundance of viruses found in aquatic environments. *Nature* 340, 467–468. <https://doi.org/10.1038/340467a0>.
- Liu, L., Li, Y., Zhang, J., Zhou, Z., Liu, J., Li, X., Zhou, J., Du, G., Wang, L., and Chen, J. (2011). Complete genome sequence of the industrial strain *Ketogulonicigenium vulgare* WSH-001. *J. Bacteriol.* 193, 6108–6109. <https://doi.org/10.1128/JB.06007-11>.
- Qin, Z., Yang, Y., Yu, S., Liu, L., Chen, Y., Chen, J., and Zhou, J. (2021). Repurposing the Endogenous Type I-E CRISPR/Cas System for Gene Repression in *Gluconobacter oxydans* WSH-003. *ACS Synth. Biol.* 10, 84–93. <https://doi.org/10.1021/acssynbio.0c00456>.
- Nam, K.H., Haitjema, C., Liu, X., Ding, F., Wang, H., DeLisa, M.P., and Ke, A. (2012). Cas5d protein processes pre-crRNA and assembles into a cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system. *Structure* 20, 1574–1584. <https://doi.org/10.1016/j.str.2012.06.016>.
- Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J.M., Wolf, Y.I., Yakunin, A.F., et al. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9, 467–477. <https://doi.org/10.1038/nrmicro2577>.
- Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J.* 30, 1335–1342. <https://doi.org/10.1038/emboj.2011.41>.
- Westra, E.R., van Erp, P.B.G., Künne, T., Wong, S.P., Staals, R.H.J., Seegers, C.L.C., Bollen, S., Jore, M.M., Semenova, E.,

- Severinov, K., et al. (2012). CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol. Cell* 46, 595–605. <https://doi.org/10.1016/j.molcel.2012.03.018>.
34. Rath, D., Amlinger, L., Hoekzema, M., Devulapally, P.R., and Lundgren, M. (2015). Efficient programmable gene silencing by Cascade. *Nucleic Acids Res.* 43, 237–246. <https://doi.org/10.1093/nar/gku1257>.
35. Nimkar, S., and Anand, B. (2020). Cas3/I-C mediated target DNA recognition and cleavage during CRISPR interference are independent of the composition and architecture of Cascade surveillance complex. *Nucleic Acids Res.* 48, 2486–2501. <https://doi.org/10.1093/nar/gkz1218>.
36. Marraffini, L.A., and Sontheimer, E.J. (2010). Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* 463, 568–571. <https://doi.org/10.1038/nature08703>.
37. Wiedenheft, B., van Duijn, E., Bultema, J.B., Waghmare, S.P., Zhou, K., Barendregt, A., Westphal, W., Heck, A.J.R., Boekema, E.J., Dickman, M.J., et al. (2011). RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. USA* 108, 10092–10097. <https://doi.org/10.1073/pnas.1102716108>.
38. Semenova, E., Jore, M.M., Datsenko, K.A., Semenova, A., Westra, E.R., Wanner, B., van der Oost, J., Brouns, S.J.J., and Severinov, K. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci. USA* 108, 10098–10103. <https://doi.org/10.1073/pnas.1104144108>.
39. Zhu, W., and Klinman, J.P. (2020). Biogenesis of the peptide-derived redox cofactor pyrroloquinoline quinone. *Curr. Opin. Chem. Biol.* 59, 93–103. <https://doi.org/10.1016/j.cbpa.2020.05.001>.
40. Hölscher, T., and Görisch, H. (2006). Knockout and Overexpression of Pyrroloquinoline Quinone Biosynthetic Genes in *Gluconobacter oxydans* 621H. *J. Bacteriol.* 188, 7668–7676.
41. Ramamoorthi, R., and Lidstrom, M.E. (1995). Transcriptional analysis of *pqqD* and study of the regulation of pyrroloquinoline quinone biosynthesis in *Methylobacterium extorquens* AM1. *J. Bacteriol.* 177, 206–211. <https://doi.org/10.1128/jb.177.1.206-211.1995>.
42. Jacob, F. (2011). The Birth of the Operon. *Science* 332, 767. <https://doi.org/10.1126/science.1207943>.
43. Eyre-Walker, A. (1995). The distance between *Escherichia coli* genes is related to gene expression levels. *J. Bacteriol.* 177, 5368–5369. <https://doi.org/10.1128/jb.177.18.5368-5369.1995>.
44. Loeff, L., Brouns, S.J.J., and Joo, C. (2018). Repetitive DNA Reeling by the Cascade-Cas3 Complex in Nucleotide Unwinding Steps. *Mol. Cell* 70, 385–394.e3. <https://doi.org/10.1016/j.molcel.2018.03.031>.
45. Ge, X., Wang, W., Du, B., Wang, J., Xiong, X., and Zhang, W. (2015). Multiple *pqqA* genes respond differently to environment and one contributes dominantly to pyrroloquinoline quinone synthesis. *J. Basic Microbiol.* 55, 312–323. <https://doi.org/10.1002/jobm.201300037>.
46. Rollins, M.F., Schuman, J.T., Paulus, K., Bukhari, H.S.T., and Wiedenheft, B. (2015). Mechanism of foreign DNA recognition by a CRISPR RNA-guided surveillance complex from *Pseudomonas aeruginosa*. *Nucleic Acids Res.* 43, 2216–2222. <https://doi.org/10.1093/nar/gkv094>.
47. Liu, L., Li, Y., Zhang, J., Zou, W., Zhou, Z., Liu, J., Li, X., Wang, L., and Chen, J. (2011). Complete Genome Sequence of the Industrial Strain *Bacillus megaterium* WSH-002. *J. Bacteriol.* 193, 6389–6390. <https://doi.org/10.1128/JB.06066-11>.
48. Wang, C.Y., Li, Y., Gao, Z.W., Liu, L.C., Wu, Y.C., Zhang, M.Y., Zhang, T.Y., and Zhang, Y.X. (2018). Reconstruction and analysis of carbon metabolic pathway of *Ketogulonicigenium vulgare* SPU B805 by genome and transcriptome. *Sci. Rep.* 8, 17838–17911. <https://doi.org/10.1038/s41598-018-36038-3>.
49. Cai, L., Yuan, M.Q., Li, Z.J., Chen, J.C., and Chen, G.Q. (2012). Genetic engineering of *Ketogulonicigenium vulgare* for enhanced production of 2-keto-L-gulononic acid. *J. Biotechnol.* 157, 320–325. <https://doi.org/10.1016/j.jbiotec.2011.12.004>.
50. Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M., 2nd, and Peterson, K.M. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* 166, 175–176. [https://doi.org/10.1016/0378-1119\(95\)00584-1](https://doi.org/10.1016/0378-1119(95)00584-1).
51. Hilgarth, R.S., and Lanigan, T.M. (2020). Optimization of overlap extension PCR for efficient transgene construction. *MethodsX* 7, 100759. <https://doi.org/10.1016/j.mex.2019.12.001>.
52. Fu, S., Zhang, W., Guo, A., and Wang, J. (2007). Identification of promoters of two dehydrogenase genes in *Ketogulonicigenium vulgare* DSM 4025 and their strength comparison in *K. vulgare* and *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 75, 1127–1132. <https://doi.org/10.1007/s00253-007-0930-z>.
53. Kato, C., Kawai, E., Shimizu, N., Mikekado, T., Kimura, F., Miyazawa, T., and Nakagawa, K. (2018). Determination of pyrroloquinoline quinone by enzymatic and LC-MS/MS methods to clarify its levels in foods. *PLoS One* 13, e0209700–e0209714. <https://doi.org/10.1371/journal.pone.0209700>.
54. Ameyama, M., Nonobe, M., Shinagawa, E., Matsushita, K., and Adachi, O. (1985). Method of enzymatic determination of pyrroloquinoline quinone. *Anal. Biochem.* 151, 263–267. [https://doi.org/10.1016/0003-2697\(85\)90174-5](https://doi.org/10.1016/0003-2697(85)90174-5).

## STAR★METHODS

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Monoclonal mouse anti-GFP antibody	Abmart	Cat#M20004; RRID:AB_2619674
Goat anti-mouse IgG HRP-linked antibody	Transgen Biotech	Cat#HS201-01; RRID:AB_2629433
<b>Bacterial and virus strains</b>		
<i>E. coli</i> DH5 $\alpha$	Lab stock	N/A
<i>B. megaterium</i>	Lab stock	N/A
<i>K. vulgare</i> WSH-001	Lab stock	N/A
MT <i>K. vulgare</i> ( $\Delta$ ppqA-B)	This study	N/A
<i>K. vulgare</i> WSH-001 derivatives	This study	N/A
<b>Chemicals, peptides, and recombinant proteins</b>		
BamHI	Thermo Fisher Scientific	Cat#FD0054
HindIII	Thermo Fisher Scientific	Cat#FD0504
SacI	Thermo Fisher Scientific	Cat#FD1133
RNAiso Plus reagent	Takara	Cat#D9108A
PQQ-dependent Glucose dehydrogenase	Toyobo	Cat#GLD-321
PQQ standards	Sigma-Aldrich	Cat#80198
2,6-Dichlorophenolindophenol (DCIP)	Sangon Biotech	Cat#A600396
Phenazine methosulfate (PMS)	Sangon Biotech	Cat#A610361
<b>Critical commercial assays</b>		
PrimeScript RT reagent Kit with gDNA Eraser	Takara	Cat#RR047A
SYBR Premix Ex Taq II	Takara	Cat#RR820A
<b>Deposited data</b>		
<i>K. vulgare</i> WSH-001 genome	NCBI Reference Sequence	CP002018
<b>Oligonucleotides</b>		
All oligonucleotides used for plasmid construction and qRT-PCR are listed in supplemental materials.	This study	N/A
<b>Recombinant DNA</b>		
Plasmid: pBBR1MCS-5	Lab stock	N/A
Plasmid: pBBR1MCS-2	Lab stock	N/A
pBBR1MCS-5/-2 derivatives, see <a href="#">Table S8</a>	This study	N/A

## RESOURCE AVAILABILITY

## Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Yi Yang ([yangyi528@scu.edu.cn](mailto:yangyi528@scu.edu.cn)).

## Materials availability

All plasmids generated and or used in this study are listed in [Table S8](#). All plasmids generated in this study are available from the [lead contact](#) upon request.

### Data and code availability

- All data reported in this study will be shared by the lead contact upon request.
- This study does not report original code. The genome information of *K. vulgare* WSH-001 is available on NCBI repository under accession number CP002018.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

*K. vulgare* WSH-001 (PQQ producing strain)<sup>28</sup> and *Bacillus megaterium* (the helper strain)<sup>47</sup> were mainly used in the study. The culture of *B. megaterium* was prepared as follows. The media was modified from Wang et al.<sup>48,49</sup> P media for isolation and purification of strains contained 2% L-Sorbose, 0.02% MgSO<sub>4</sub>, 0.1% KH<sub>2</sub>PO<sub>4</sub>, 0.3% yeast extract, 1% peptone, 0.3% beef extract and 0.1% urea with pH 7.1–7.3. Also, cultivation of *B. megaterium* was completed in P media. While PQQ synthesis in *K. vulgare* was conducted in a mixture composing of P media and culture of *B. megaterium* (50%, v/v) at 28°C–30°C, 180 rpm and ~3% (v/v) inoculation. Culture was prepared by inoculation, cultivation, centrifugation and sterile filtration. When required, media were supplemented with gentamicin at 20 µg/mL or kanamycin at 25 µg/mL for the selection of pBBR1MCS-derived plasmids.<sup>50</sup> The vectors were transformed to *K. vulgare* by Bio-Rad Gene Pulser (Bio-Rad) (0.1 cm-gap cuvettes, 2.2 kV, 200 Ω, and 25 µF). The endogenous CRISPR-Cas system in *K. vulgare* was developed into a genomic editing technology for gene deletion at the indicated positions as previously described.<sup>19</sup> The 400-bp fragment in *pqq* operon comprising of *pqqA*, the intergenic region and part of *pqqB* was deleted and the mutant strain (MT) was acquired.

## METHOD DETAILS

### CRISPR-cas system identification

CIRSPRCasFinder (<https://crisprcas.i2bc.paris-saclay.fr/CrisprCasFinder/Index>) was used to identify CRISPR-Cas system in *K. vulgare* WSH-001. The *cas* genes and CRISPR loci were also detected by CIRSPRCasFinder. RNA structure in this study was predicted on RNAfold WebServer (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>).

### CRISPR interference assay

CRISPR interference plasmids were constructed according to the ref.<sup>19,36</sup> The oligonucleotides bearing Spacer5 and 5'-NNN were synthesized. After annealing, fragments were ligated to the linearized vector pBBR1MCS-5 by HindIII and BamHI digestion. All synthesis was completed by TsingkeBiotechnology (Chengdu, China). The oligonucleotides were listed in [Table S1](#).

### DNA cloning and plasmid construction

*E. coli* DH5α was used for plasmid construction and cultured in Luria-Bertani (LB) media. To screen for the constitutively expressed promoters, the plasmids carrying reporter gene (*gfp*) driven by different promoters were constructed by overlap PCR respectively.<sup>51,52</sup> Genomic DNA from *K. vulgare* was used as the templates for promoters. Plasmid pBI221-GFP provided the template for *gfp* gene. The fused fragments were cloned into pBBR1MCS-5 plasmid by HindIII and BamHI digestion. Related primers were listed in [Table S2](#). CRISPR expression plasmids were constructed as follows. The fragments consisting of the leader of CRISPR1 as a promoter and the repeats interrupted by designed spaces were synthesized and used as the templates for plasmid construction. Products digested by SacI and BamHI were recombined into pBBR1MCS-5 to acquire the crRNAs expressing plasmids. The overexpressed vectors carrying *pqq* genes were constructed by cloning fragments into pBBR1MCS-2. Primers were shown in [Table S3](#). The artificial units *pqqA(nNNN)gfp* and *pqqA(nNNN)BCDEF* were recombined also by overlap PCR. Primers were listed in [Table S4](#).

### RT-PCR and quantitative RT-PCR

Samples were taken at a predetermined time when the growth of strains reached about 0.4 (OD<sub>600</sub>). RNA was extracted with RNAiso Plus reagent (Takara). The cDNA was synthesized using PrimeScript RT reagent Kit with gDNA Eraser (Takara). Quantitative RT-PCR (qRT-PCR) was performed on a Bio-Rad CFX96 Real-Time PCR Detection System using SYBR Premix Ex Taq II (Takara). 16S rDNA was used as an internal control. Primers were listed in [Table S5](#). RT-PCR was performed with cDNA templates and the corresponding primers. Primers were in [Table S6](#).

### Western blot analysis

Western blot (WB) was to detect GFP expression. WB and RNA extraction samples were collected at the same time. The quantified loading protein was separated on a 12% SDS-PAGE gel and then blotted onto a PVDF membrane (Bio-Rad). Primary antibody was Anti-GFP (Abmart, M20004, diluted 5000-fold). Secondary antibody used was Goat Anti-Mouse IgG HRP-linked antibody (Transgen, HS201-01, diluted 5000-fold).

### PQQ determination assay

PQQ products were measured according to the ref.<sup>53,54</sup> PQQ-dependent Glucose dehydrogenase (GLD) (Toyobo) was dissolved into 100 mM phosphate buffer (containing 1 mM ethylenediamine-N,N,N',N'-tetraacetic acid (EDTA) and 2 M KBr; pH 7.3) for dialysis treatment.

The obtained apo-GLD was then dialyzed against 20 mM 3-morpholinopropanesulfonic acid buffer (MOPS) (pH7.0). The resultant solution was measured and diluted to the desired concentration with MOPS buffer (including 0.1% Triton X-100, 1 mM  $\text{CaCl}_2$  and 0.1% BSA) for PQQ detection. PQQ standards (Sigma) were prepared with MOPS buffer. Apo-GLD solution was incubated with standards at 30°C for 35 min to reactivate GLD. MOPS buffer (including 2 mM  $\text{CaCl}_2$ , 400  $\mu\text{M}$  phenazine methosulfate (PMS), 200  $\mu\text{M}$  2, 6-dichloroindophenol (DCIP) sodium salt and 20 mM glucose) was used as the enzymatic activity assay buffer. After incubation, 1 mL assay buffer was added, 200  $\mu\text{L}$  of which was taken out immediately to determine the rate of discoloration of DCIP blue at 600 nm. For the samples, the culture was centrifuged at 12,000 rpm  $\times$  5 min, and the supernatant was taken for PQQ analysis.

### QUANTIFICATION AND STATISTICAL ANALYSIS

All data are shown in figures as mean and represent standard deviation (SD) from the mean. Number of replicates and statistical tests used can be found in figure legends. One-way ANOVA test and Student's *t* test analysis were performed on GraphPad Prism 7 software. Differences were considered statistically significant if the *p* value was less than 0.05. \**P* < 0.05; \*\**P* < 0.01; \*\*\**P* < 0.001; \*\*\*\**P* < 0.0001.