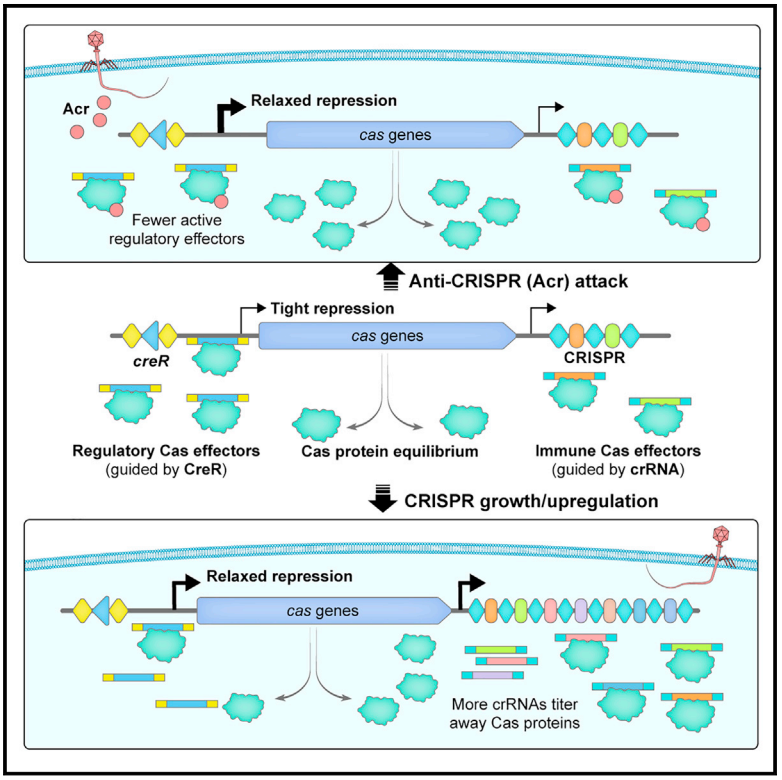
Article

Cell Host & Microbe

# Widespread RNA-based *cas* regulation monitors crRNA abundance and anti-CRISPR proteins

## Graphical abstract




## Highlights

- Identification of a class of *cas*-regulating RNA (CreR) species that resemble crRNAs
- CreR mediates autorepression of type I and V-A Cas effectors to reduce autoimmune risks
- CreR monitors the level of typical crRNAs to proportionately fine-tune *cas* expression
- CreR oversees the inhibition of Cas effectors by phage anti-CRISPR proteins

## Authors

Chao Liu, Rui Wang, Jie Li, ..., Jun Yang, Hua Xiang, Ming Li

## Correspondence

xiangh@im.ac.cn (H.X.),
lim_im@im.ac.cn (M.L.)

## In brief

Liu et al. report a class of crRNA-resembling RNAs that direct the autoregulation of type I and V-A Cas effectors. These *cas*-regulating RNA (CreR) molecules not only reduce the risk of CRISPR-Cas autoimmunity but also monitor the cellular level of typical crRNAs and surveil Cas inhibition by bacteriophage anti-CRISPR proteins.

Liu et al., 2023, Cell Host & Microbe *31*, 1481–1493
September 13, 2023 
https://doi.org/10.1016/j.chom.2023.08.005

CellPress

Cell Host & Microbe

CellPress

Article

# Widespread RNA-based *cas* regulation monitors crRNA abundance and anti-CRISPR proteins

Chao Liu,[1,3,7] Rui Wang,[1,7] Jie Li,[2,7] Feiyue Cheng,[1,7] Xian Shu,[1,3,7] Huiwei Zhao,[1] Qiong Xue,[1] Haiying Yu,[2] Aici Wu,[1,3] Lingyun Wang,[1,4] Sushu Hu,[1] Yihan Zhang,[1,5] Jun Yang,[1,6] Hua Xiang,[2,3,*] and Ming Li[1,3,8,*]

[1]CAS Key Laboratory of Microbial Physiological and Metabolic Engineering, State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China

[2]State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China

[3]College of Life Science, University of Chinese Academy of Sciences, Beijing, China

[4]College of Plant Protection, Shandong Agricultural University, Taian, Shandong, China

[5]School of Life Sciences, Hebei University, Baoding, Hebei, China

[6]Center for Life Science, School of Life Sciences, Yunnan University, Kunming, China

[7]These authors contributed equally

[8]Lead contact

*Correspondence: xiangh@im.ac.cn (H.X.), lim_im@im.ac.cn (M.L.)

https://doi.org/10.1016/j.chom.2023.08.005

## SUMMARY

CRISPR RNAs (crRNAs) and Cas proteins work together to provide prokaryotes with adaptive immunity against genetic invaders like bacteriophages and plasmids. However, the coordination of crRNA production and *cas* expression remains poorly understood. Here, we demonstrate that widespread modulatory mini-CRISPRs encode *cas*-regulating RNAs (CreRs) that mediate autorepression of type I-B, I-E, and V-A Cas proteins, based on their limited complementarity to *cas* promoters. This autorepression not only reduces autoimmune risks but also responds to changes in the abundance of canonical crRNAs that compete with CreR for Cas proteins. Furthermore, the CreR-guided autorepression of Cas proteins can be alleviated or even subverted by diverse bacteriophage anti-CRISPR (Acr) proteins that inhibit Cas effectors, which, in turn, promotes the generation of new Cas proteins. Our findings reveal a general RNA-guided autorepression paradigm for diverse Cas effectors, shedding light on the intricate self-coordination of CRISPR-Cas and its transcriptional counterstrategy against Acr proteins.

## INTRODUCTION

In prokaryotes, CRISPR-Cas systems constitute the adaptive defense line against invading genetic elements, like viruses (phages) and plasmids.[1–5] These systems are highly diversified and currently classified into 2 classes, 6 types, and more than 30 subtypes.[6] The CRISPR arrays contain invader-derived sequences that are inserted between each two direct repeats, known as spacers. These arrays are accompanied by the *cas* (CRISPR-associated) genes that encode a multi-subunit effector complex (class 1) or a single-protein effector (class 2). Mature CRISPR RNAs (crRNAs) guide the Cas effector to precisely recognize and cleave the foreign nucleic acids based on the perfect complementarity between their spacer portion and the target site (known as protospacer), with a conserved protospacer adjacent motif (PAM) playing a critical role during target recognition.[7,8] Typically, the CRISPR-Cas system also encodes Cas proteins (e.g., Cas1, Cas2, and Cas4) that mediate the acquisition of new spacers from the genetic invaders.[9–11] This process, termed CRISPR adaptation, is believed to occur in either of two modes: naive adaptation or primed adaptation, depending on whether a priming spacer that partially matches the invader DNA is required.

The acquisition and loss of spacers in CRISPR arrays during bacteria-phage conflicts result in a constantly evolving volume of CRISPR memory.[12,13] To ensure effective CRISPR immunity, a sufficient supply of Cas proteins is necessary; however, constant high-level Cas production may pose a higher risk of autoimmunity and perhaps other negative effects on the host.[14,15] To achieve a balance between providing sufficient protein effectors for crRNA guides and avoiding negative effects, the regulation of *cas* expression and crRNA production requires precise coordination, yet it remains poorly understood.

Recent studies have shed light on the secondary function of the CRISPR immune effector, revealing its potential as a gene regulator through limited spacer-protospacer complementarity. For instance, the type II effector Cas9 has been found to play a role in transcriptional repression of a virulence-related regulon, a process directed by a noncanonical RNA guide called small CRISPR-associated RNA (scaRNA).[16] In type II and some type V systems, the crRNA maturation process and the interference process essentially involve a *trans*-activating crRNA (tracRNA), which features an anti-repeat that forms a duplex with the repeat portion of crRNA precursors.[17–19] Recently, it has emerged that, in most II-A systems, a long

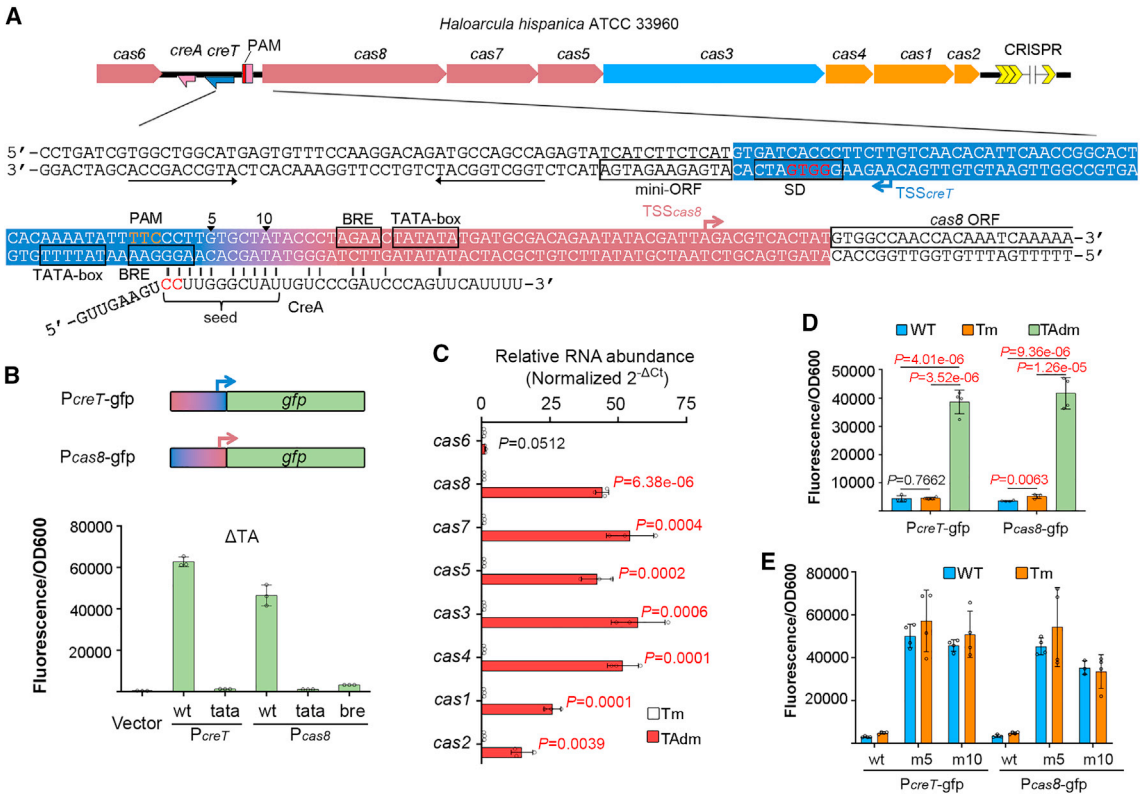



**Figure 1. *H. hispanica* CreA synchronously represses $P_{creT}$ and $P_{cas8}$**

(A) Schematic depiction of the divergent $P_{creT}$ and $P_{cas8}$ and their targeting by CreA. Red nucleotides within the Shine-Dalgarno (SD) sequence of *creT* were mutated to generate the Tm mutant, and then red nucleotides within CreA were further mutated to construct TAdm.

(B) Validation of $P_{cas8}$ using a green fluorescent protein (*gfp*) gene. The TATA box (tata) and BRE (bre) elements were separately mutated.

(C) The relative RNA level of *cas* genes in Tm and TAdm. 7*S* RNA served as the internal control.

(D) Activity of $P_{creT}$ and $P_{cas8}$ in WT, Tm, and TAdm.

(E) Activity of $P_{creT}$ and $P_{cas8}$ in WT and Tm cells when the 5$^{th}$ (m5) or 10$^{th}$ (m10) seed nucleotide within the target site of CreA (labeled in A) was mutated.

Error bars, mean ± SD (n = 3 or 4). p values below 0.01 were highlighted in red (two-tailed Student's t test).

See also Figure S1.

isoform of tracrRNA (tracr-L) additionally contains a spacer-like sequence that partially matches the *cas9* promoter, thus enabling the repurposing of Cas9 for autoregulation.[20] Our recent studies have also unraveled the regulatory role of type I-B CRISPR effector complex.[21–23] We found that a degenerated mini-CRISPR, called CRISPR-resembling antitoxin (*creA*), can reprogram this effector complex to transcriptionally repress a small toxic RNA, CRISPR-repressed toxin (CreT), which acts by sequestering a rare tRNA species in the host cell. This tiny two-RNA toxin-antitoxin (TA) element (considered to represent type VIII TA[24]) makes host cells addicted to active CRISPR effectors because disrupting these effectors would unleash CreT expression and trigger cell death/dormancy. Notably, the spacer portion of CreA share limited complementarity to the promoter DNA of *creT* ($P_{creT}$) (Figure 1A), which causes gene regulation rather than DNA cleavage. Notably, a very recent study has collectively termed these small RNA species as crRNA-like RNAs (crlRNAs).[25] This study not only unraveled their wide presence in a broad variety of CRISPR-Cas loci but also delineated their potential role in regulating *cas* genes and other auxiliary genes.

In this study, we showed that CreA not only directs the type I-B Cas proteins to repress toxin expression but can also mediate their autorepression. We also identified *creA* analogs (or crlRNAs) from type I-E and V-A systems and experimentally validated their regulatory role in directing the autorepression circuit of multi-subunit (I-E) or single-protein (V-A) Cas effector. It is worth noting that these crlRNAs display different characteristics in terms of molecular architecture and physiological function when compared with tracrRNAs and short-complementarity untranslated (scout) RNAs (functional analogs of tracrRNAs in type V-C and V-D[26]), which typically lack spacer-like sequences and participate in crRNA maturation. Therefore, in this study, we propose the term "CreR" to refer specifically to such a *cas*-regulating RNA element. Remarkably, we further showed that the CreR-guided *cas* repression can be fine-tuned by altering the cellular concentration of the canonical crRNAs, as these molecules compete for Cas proteins with CreR RNAs. In addition, we also demonstrated that this autorepression effect can be relieved or even subverted by Acr proteins that inhibit Cas proteins, which illuminates a distinct anti-anti-CRISPR (Acr) strategy that acts on transcription level.

## RESULTS

### *H. hispanica* CreA synchronously represses *creT* and *cas* transcription

In our earlier study, we characterized the Cascade-regulated toxin-antitoxin (CreTA) RNA pair that safeguards the type I-B CRISPR-Cas in the archaeon *Haloarcula hispanica* (*H. hispanica*).[21] As this mini addiction module is situated within the ~300 base pair (bp) intergenic sequence between *cas6* and *cas8* (Figure 1A), we had created a ΔTA mutant by simply deleting this intergenic sequence.[21] In a subsequent study, we observed a significant reduction of 70%–90% of *cas* transcript levels (except that of *cas6*) in the ΔTA mutant.[27] This reduction indicates the presence of an overlooked promoter preceding *cas8*, which we have now identified as $P_{cas8}$. Through reanalyzing the data of a small RNA sequencing assay that was originally aimed to profile *creTA* transcription (Figure S1), we successfully identified the transcription start site (TSS) of *cas8*. This TSS is located 12 bp upstream of the open reading frame (ORF) (Figure 1A), allowing us to predict the corresponding archaeal promoter elements, BRE (TF-IIB recognition element), and TATA box (Figure 1A). To validate our predictions, we utilized the green fluorescent protein (*gfp*) gene as a reporter and discovered that the activity of $P_{cas8}$ was nullified by mutating either BRE or TATA box (Figure 1B). We further showed that, in ΔTA cells, $P_{cas8}$ was remarkably efficient, exhibiting approximately 135 times the activity of *cas6* promoter ($P_{cas6}$) (Figure S1C). And surprisingly, it also outperformed the *creT* promoter ($P_{creT}$) by roughly 1.4 times and a strong constitutive promoter we typically utilized for gene overexpression in haloarchaea ($P_{phaR}$) by approximately 2.0 times.[28]

Notably, $P_{cas8}$ and $P_{creT}$ run divergently and tightly flank the target site of CreA (Figure 1A), suggesting their simultaneous repression by CreA. Because simply mutating CreA would lead to CreT de-repression and hence cell death, we first created a CreT mutant (designated as Tm) by disrupting its Shine-Dalgarno element and then mutated the seed sequence (essential for target recognition) of CreA to generate a CreT/CreA double mutant (TAdm) (Figure 1A). To our surprise, *cas8* transcription was observed to increase by as much as ~44-fold in TAdm when compared with Tm. Additionally, all downstream *cas* genes were observed to be markedly up-regulated, with increases ranging from 15- to 54-fold. By contrast, the expression of *cas6* showed no significant change ($p = 0.0512$; Figure 1C). To confirm this effect, we examined the GFP fluorescence from wild type (WT), Tm, or TAdm cells containing a $P_{creT}$ or $P_{cas8}$-controlled *gfp*. As expected, fluorescence levels were not ($P_{creT}$) or only slightly ($P_{cas8}$) increased by CreT mutation (possibly due to the translation-inhibiting effect of CreT toxin[21]), while for both promoters, fluorescence levels were greatly elevated (by over 10-fold) in TAdm where CreA was further mutated (Figure 1D). We proceeded to introduce single nucleotide substitutions into the target site of CreA. Consequently, both $P_{cas8}$ and $P_{creT}$ were de-repressed, as evidenced in both WT and Tm cells (Figure 1E), which confirmed the synchronization of their downregulation by CreA.

### CreA-guided *cas* repression reduces autoimmune risks

The overexpression of *cas* genes in TAdm implies that this mutant may be more proficient in CRISPR immunity. To test this, we introduced into Tm and TAdm cells a synthetic mini-CRISPR with a 34-bp spacer (namely v10) targeting the *H. hispanica* pleomorphic virus 2 (HHPV-2) virus[10,29] (Figure 2A). Upon subjecting the cells to HHPV-2 infection, we observed equivalent viral immunity levels in both hosts. This indicates that the wild expression level of Cas proteins in Tm was sufficient to provide robust immunity. Viruses can escape CRISPR immunity by mutating the nucleotides within the target site, so we introduced various mutations into the v10 spacer sequence to test whether elevated *cas* expression can confer more efficient immunity against imperfect targets (illustrated in Figure 2A). Our data showed that, when the first, second, or third spacer nucleotide of the seed region (proximal to PAM) was separately mutated, TAdm did exhibit a higher immunity level than Tm ($p = 0.0020$, 0.0059, and 0.0108, respectively; two-tailed Student's t test) (Figure 2A). Similarly, when a 3-bp mismatch was introduced at the PAM-distal end, TAdm also exhibited enhanced immunity ($p = 0.0038$). However, when a more extensive mismatch (6, 9, or 12 bp) was introduced, neither mutant showed virus immunity (Figure 2A). Therefore, TAdm does have a superior CRISPR immunity against some escape mutants of a target virus.

Because the *cas* genes involved in adaptation (*cas1*, *cas2*, and *cas4*) were also up-regulated in TAdm, we inferred this mutant would be more effective at acquiring new spacers. The *H. hispanica* CRISPR array has 13 spacers, and the terminal spacer (s13) shares ~70% sequence identity with HHPV-2, making it a priming candidate for efficient spacer acquisition from this virus.[10] Although we did not detect a marked increase in crRNA abundance in TAdm (Figure S2A), CRISPR array in this strain was still possibly up-regulated. So, we introduced an artificial CRISPR overexpressing the s13 crRNA to ensure equivalent amounts of priming molecules in WT, Tm, and TAdm cells (Figure 2B), which was further verified by northern blotting (Figure S2B). As expected, TAdm incorporated new spacers into the chromosomal CRISPR array more frequently than WT and Tm when infected at low or high multiplicity of infection (MOI, 0.1 or 40) (Figure 2B).

However, our Illumina sequencing data revealed that TAdm mistakenly acquired self-derived spacers at a higher frequency of around 1.69%, compared with Tm which showed a frequency of only 0.08% (Figure 2C). We suspect that many, if not all, of these self-derived spacers were acquired via the naive adaptation, which does not rely on a priming step. Consistently, without virus infections, TAdm cells showed very inefficient acquisition of endogenous DNA as new spacers when incubated at room temperature for approximately 1 month (Figure S3). Because the 13 pre-existing spacers within the chromosomal CRISPR array may have a chance of partially matching endogenous DNA, we further created the TAdm-ΔCRISPR mutant to eliminate the possibility of primed adaptation and then monitored spacer acquisition using an "adaptation" plasmid that carries a CRISPR leader and a single-repeat sequence. Remarkably, this plasmid did acquire new spacers in TAdm-ΔCRISPR cells after they were incubated at room temperature for a couple of days (Figure S3C). Therefore, CRISPR adaptation does tone up in TAdm, but with a higher risk of spawning self-targeting spacers via the naive pathway.

We predicted that the accumulation of self-targeting spacers could potentially compromise the fitness of TAdm cells.

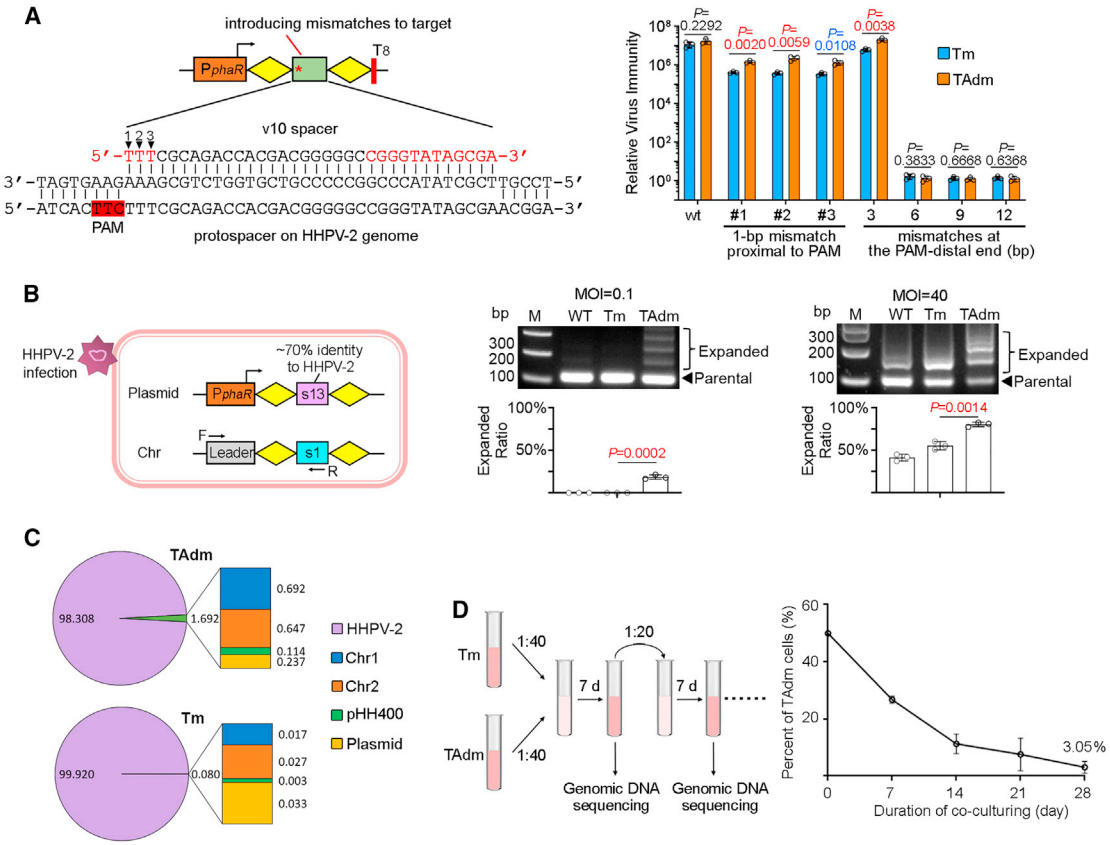



**Figure 2. TAdm showed stronger viral immunity with the risk of autoimmunity**

(A) The relative virus immunity of Tm and TAdm cells expressing a crRNA (v10) that fully or partially targets the HHPV-2 genome, compared with those carry empty vectors (see method details). By mutating the spacer nucleotide(s) (highlighted in red), a 1-bp PAM-proximal mismatch (#1, #2, or #3) was introduced to modify the spacer-protospacer complementarity, or a 3-, 6-, 9-, or 12-bp mismatch was introduced to the PAM-distal end. A constitutive promoter ($P_{phaR}$) was used to express the artificial mini-CRISPR, and a string of eight thymines ($T_8$) was used as a terminator.

(B) CRISPR adaptation to HHPV-2 in Tm or TAdm cells overexpressing the s13-crRNA (partially complementary to the viral DNA). Primers specific to the chromosomal CRISPR were used for PCR amplification, and the expanded PCR products indicated acquisition of new spacers. MOI, multiplicity of infection.

(C) The origin ratio (%) of new spacers. *H. hispanica* genome consists of two chromosomes (Chr1 and Chr2) and one mega-plasmid (pHH400).

(D) Co-cultivation of Tm and TAdm cells. Cell percent was analyzed by high-throughput DNA sequencing.

Error bars, mean ± SD (n = 3). p values below 0.05 or 0.01 were denoted in blue or red, respectively (two-tailed Student's t test).

See also Figures S2–S4.

Consistently, in a long-term co-culture experiment with Tm and TAdm cells, we observed that Tm cells outcompeted TAdm cells over time, with the proportion of TAdm cells dropping to 3.05% after 28 days (Figure 2D). We further repeated this assay using CRISPR-minus strains (Tm-ΔCRISPR versus TAdm-ΔCRISPR) to rule out the effects of autoimmunity. Intriguingly, we observed a similar but slower decrease in the proportion of TAdm-ΔCRISPR cells (Figure S4). This suggests that the growth disadvantage observed for TAdm cells is not solely attributed to higher risks of autoimmunity but also to other unfavorable effects of *cas* overexpression, which warrants further investigation.

## CreA-guided gene repression can monitor crRNA levels

Presuming that crRNA and CreA RNA compete for Cas proteins to form an immune or regulatory effector, we can infer that the CreA-guided gene repression may respond to alterations in crRNA abundance. Because Cascade-CreA possibly not only suppresses the activity of $P_{cas8}$ but also attenuates the potential read-through transcripts that are driven by $P_{cas6}$, we probed the transcription level of *cas8* relative to that of *cas6* in *H. hispanica* WT and ΔCRISPR cells. As expected, we observed a decrease of about 20% in the relative RNA abundance of *cas8* in ΔCRISPR when compared with WT level (Figure 3A). A similar reduction in *cas8* transcription was observed for a CRISPR mutant encoding only one spacer (Δsp2-13 in Figure 3A). Then we introduced the $P_{cas8}$ or $P_{creT}$-controlled *gfp* into these cells and found that fluorescence declined by 10%–20% in Δsp2-13 and ΔCRISPR mutants compared with WT (Figure 3B). This data more directly illustrates that $P_{cas8}$ and $P_{creT}$ were more tightly suppressed by CreA in the case of fewer or no CRISPR spacers.

We then utilized the Tm-ΔCRISPR mutant and supplemented it with a plasmid containing a leader-preceded or a $P_{phaR}$-driven CRISPR array. Cells expressing either CRISPR array produced *cas8* transcripts almost twice as much as cells with an empty vector or those with a leader-less array (Figure 3C), and the northern blotting assay also confirmed that the CRISPR-expressing

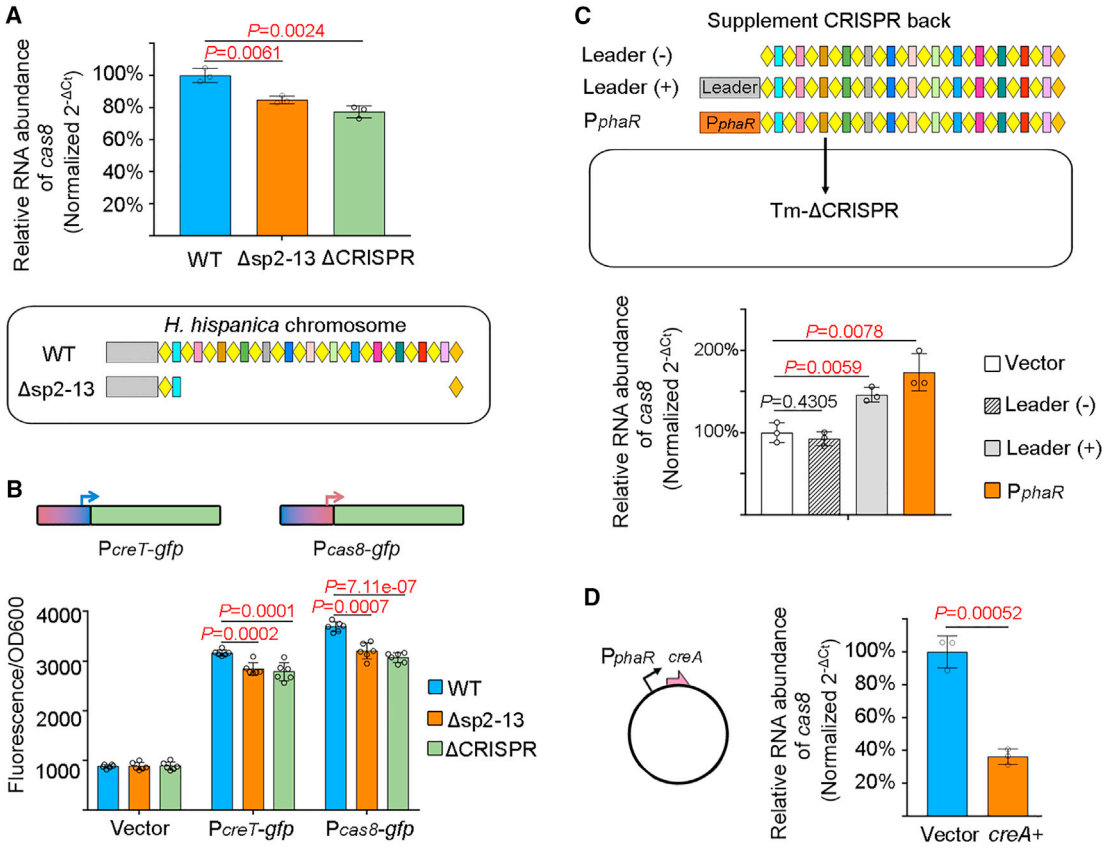



**Figure 3. $P_{creT}$ and $P_{cas8}$ can be fine-tuned by altering the number of CRISPR spacers**

(A) Relative RNA abundance of *cas8* in cells containing 13 (WT), 1 (Δsp2-13), or 0 (ΔCRISPR) CRISPR spacers. RNA of *cas6* served as the internal control. Error bars, mean ± SD (n = 3).

(B) Activity of $P_{creT}$ and $P_{cas8}$ in WT, Δsp2-13, or ΔCRISPR cells. Error bars, mean ± SD (n = 6).

(C) Relative RNA abundance of *cas8* in cells with a plasmid-born leader-less (leader−) or leader-preceded (leader+) CRISPR array, or an array that is driven by $P_{phaR}$. RNA of *cas6* served as the internal control. Tm-ΔCRISPR, a CRISPR-minus mutant constructed based on Tm. Error bars, mean ± SD (n = 3).

(D) Relative RNA abundance of *cas8* in cells containing a $P_{phaR}$-driven *creA* gene or the empty vector. RNA of *cas6* served as the internal control. Error bars, mean ± SD (n = 3).

p values below 0.01 were highlighted in red (two-tailed Student's t test).

See also Figure S2.

cells did produce significantly more crRNAs than the controls (Figure S2C) (note that the leader-less control produced a minimal level of crRNAs possibly due to readthrough transcription events driven by promoters on the plasmid backbone). Therefore, the CreA-guided gene repression can be alleviated by expanding the crRNA pool.

Then we asked whether an elevated level of CreA expression could lead to tighter repression of $P_{cas8}$. To test this possibility, we utilized the strong constitutive promoter $P_{phaR}$ to overexpress CreA and then probed *cas8* transcription. As expected, the relative abundance of *cas8* transcripts decreased by 60% in cells overexpressing CreA in comparison with those with the empty vector ($p = 0.0005$; two-tailed Student's t test) (Figure 3D). Therefore, Cas autorepression can be fine-tuned by altering the relative abundance of the defensive (crRNA) and regulatory (CreA) RNA guides.

## CreA-like CreRs in diverse type I and V-A systems

By revisiting the previously uncovered *creTA* analogs associating with I-B CRISPR systems,[21] we discovered several cases where the target site of CreA locates within or beside the putative promoter of the *cas* operon (see examples in Figure S5), which indicates CreA guides Cas autoregulation in these instances. We then conducted a manual search of the intergenic sequences of more CRISPR-Cas systems to inspect the wider distribution of CreA or CreA-like regulatory crRNAs. We discovered numerous *creA*-like elements in various type I systems (including I-C, I-D, I-E, I-F, and I-U subtypes) and, notably, in dozens of type V-A systems (listed in Table S1). In most cases, CreA targets the predicted promoter of the *cas* genes that encode the multi-subunit type I effector, or that of the *cas12a* gene, which encodes the V-A single-protein effector,[30] as exemplified in Figures S5 and S6, respectively. Notably, many of these *creA*-like elements do not appear to be co-located with a toxic gene, as shown in the example of a *Salmonella enterica* (*S. enterica*) I-E CRISPR-Cas locus (Figure 4A) and a *Moraxella bovoculi* (*M. bovoculi*) V-A locus (Figure 5A). To avoid confusion, we propose referring to these "standalone" *creA* genes as *creR* (*cas*-regulating RNA) until their coupling with toxin genes can be convincingly identified or

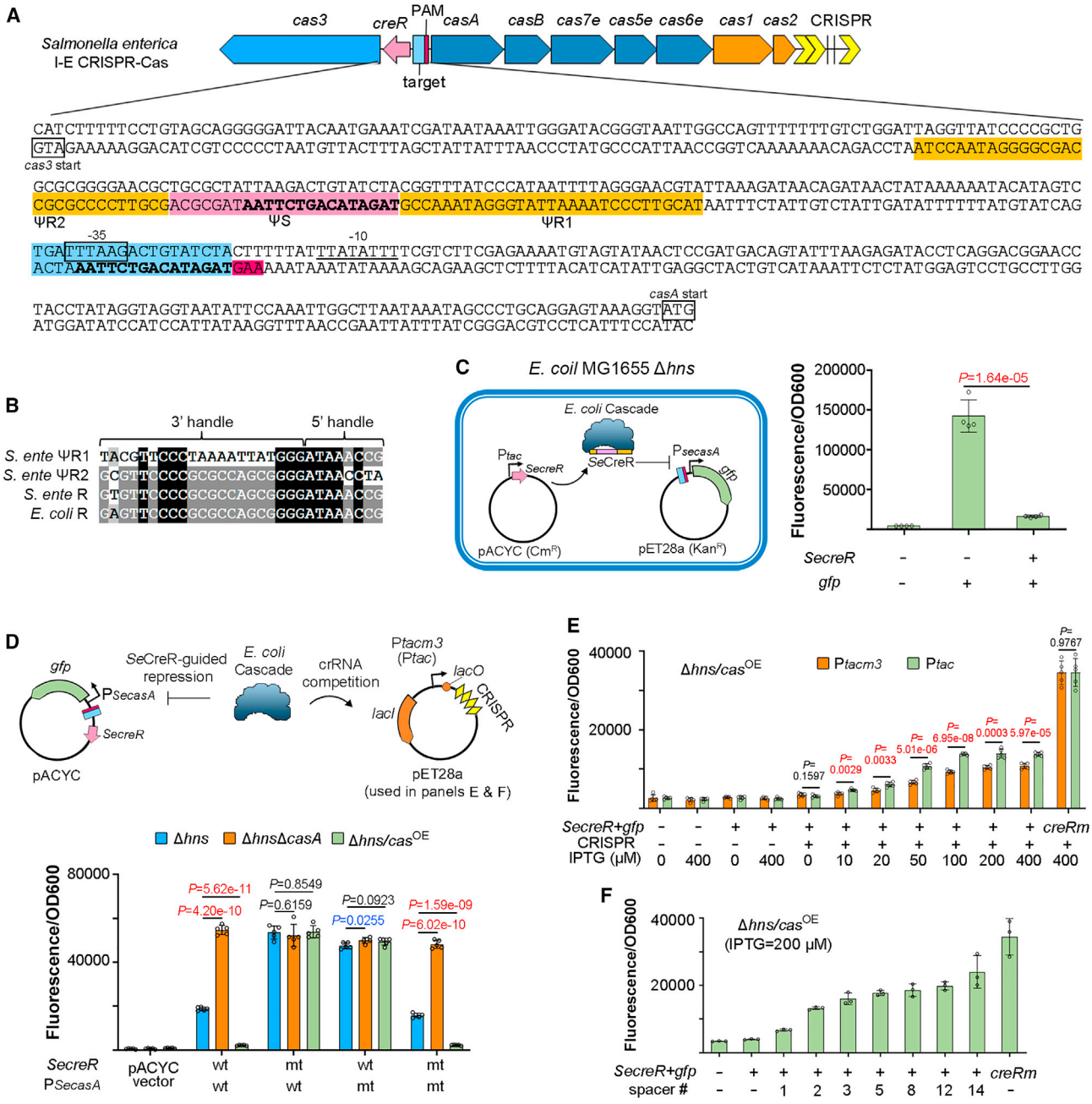



**Figure 4. The *S. enterica* I-E CRISPR-Cas employs CreR to regulate *cas* expression**

(A) Schematic depiction of the CRISPR-Cas architecture and the location and sequence of *creR*. Promoter elements (−35 and −10) of *casA* were predicted. Nucleotides in bold indicate the identical ones shared between the "spacer" (ΨS) of CreR and its corresponding protospacer.

(B) Homology between the CRISPR repeat (R) and the CreR ΨR sequences. The CRISPR repeat of *E. coli* MG1655 is included for comparison.

(C) *S. enterica* CreR (*Se*CreR) reprogramed *E. coli* Cascade to repress its cognate *casA* promoter ($P_{secasA}$). The Δ*hns* mutant was used because the *E. coli* MG1655 CRISPR-Cas system is repressed by H-NS in wild cells. Error bars, mean ± SD (n = 4).

(D) The $P_{secasA}$-repressing effect of *Se*CreR in different *E. coli* mutants. Note that, in the Δ*hns*/*cas*$^{OE}$ mutant, *cas3* was deleted and a J23119 promoter was inserted upstream of *casA* for Cascade overexpression. The complementarity between *Se*CreR and $P_{secasA}$ was disrupted by mutating either of them but remained when they were complementarily mutated. The crRNA-expressing plasmid (pET28a derivate) was used in (E) and (F). Error bars, mean ± SD (n = 5).

(E) The repression effect of *Se*CreR was relieved by overexpressing crRNAs. The CRISPR array of *E. coli* DH5α (containing 14 spacers) was amplified and expressed using an IPTG-inducible $P_{tacm3}$ or $P_{tac}$. *Se*CreR was mutated to provide a control (*creRm*). Error bars, mean ± SD (n = 5).

(F) The repression effect of *Se*CreR in cells containing a CRISPR array with varying numbers of spacers. Error bars, mean ± SD (n = 3).

p values below 0.05 or 0.01 were denoted in blue or red, respectively (two-tailed Student's t test).

See also Figure S5.

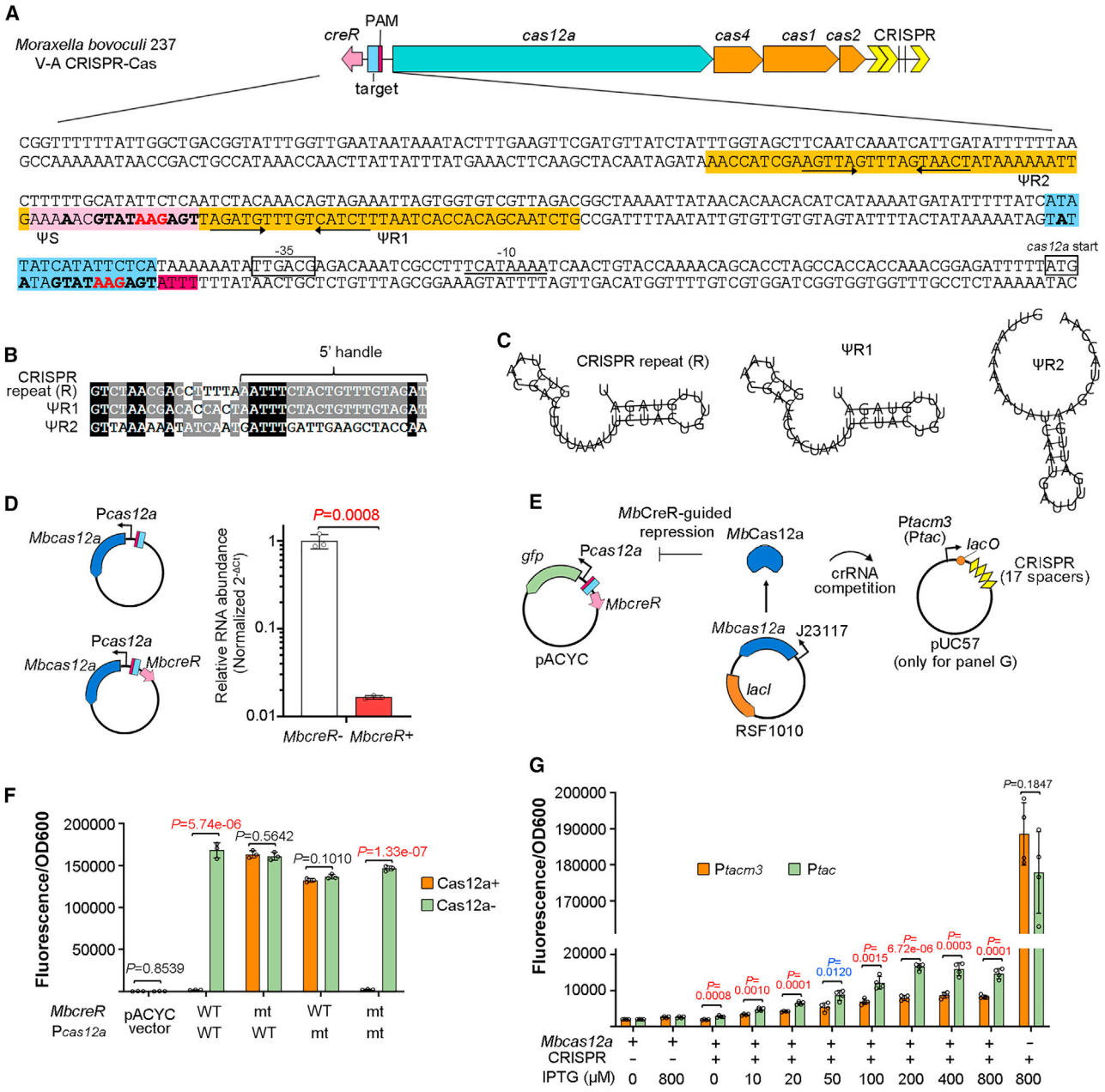



**Figure 5. The *M. bovoculi* V-A CRISPR-Cas employs CreR to regulate *cas* expression**

(A) Schematic depiction of the CRISPR-Cas architecture and the location and sequence of *creR*. Promoter elements (−10 and −35) of *cas12a* were predicted. Nucleotides in bold indicate the identical ones shared between the spacer (ΨS) of CreR and its target site. The convergent black arrows indicate inverted repeats within ΨR sequences.

(B) Homology between the CRISPR repeat (R) and the CreR ΨR sequences. Nucleotides corresponding to the 5′ handle remaining on mature RNA are indicated.

(C) The stem-loop structure predicted for CRISPR repeat and ΨR sequences.

(D) Relative RNA abundance of *Mbcas12a* (driven by its own promoter) in the absence or presence of *Mb*CreR. Error bars, mean ± SD (n = 3).

(E) Schematic illustration of the experimental design in (F) and (G). RSF1010 is a broad host range replication origin. The crRNA-expressing plasmid (pUC57 derivate) was used only in (G). The *M. bovoculi* CRISPR array containing 17 spacers was synthesized.

(F) The $P_{cas12a}$-repressing effect of *Mb*CreR relies on their complementarity. Nucleotides 4–6 in *MbcreR* ΨS sequence (indicated in red in A) and their corresponding nucleotides in the protospacer in $P_{cas12a}$ were separately or complementarily mutated (mt). WT, wild type. Error bars, mean ± SD (n = 3).

(G) The effects of varying crRNA doses on *Mb*CreR-guided gene repression. Error bars, mean ± SD (n = 4).

p values below 0.05 or 0.01 were denoted in blue or red, respectively (two-tailed Student's t test).

See also Figures S6 and S7.

predicted. It appears that Cas autoregulation directed by CreA or CreR is a general mechanism of both class 1 and class 2 CRISPR systems, which consists with the recent study by Shmakov et al.[25] Their study produced a more systematic prediction of diverse crRNA-like RNAs (including tracrRNAs, scout RNAs, scaRNAs, CreA/CreR RNAs, and perhaps RNAs of other functions) in all publicly available CRISPR-Cas loci. In line with our findings, their search revealed that putative *cas*-regulating crlRNAs are broadly dispersed throughout type I and V-A systems.

## Type I-E CreR fine-tunes *cas* expression to match crRNA levels

Due to the high similarity (only differing by one nucleotide) of the CRISPR repeat between the *S. enterica* ATCC 51960 and *Escherichia coli* (*E. coli*) MG1655 CRISPR-Cas systems (Figure 4B), we opted to test the *cas* repression effect of *S. enterica* CreR (*Se*CreR) in the Δ*hns* mutant of MG1655 (*cas* genes are repressed by heat-stable nucleoid structuring (H-NS) protein in the wild strain). Similar to the case of *creA*,[27,31] ΨR1 of *S. enterica creR* gene holds more conservation for the 8 nucleotides that produce a 5′ handle on the mature RNA, while ΨR2 holds more conservation for those corresponding to a 3′ handle (Figure 4B). The spacer-like ΨS sequence shares 15 consecutive nucleotides with its target site, flanked by a 5′-AAG-3′ trinucleotide (the canonical PAM of I-E systems), which is positioned ~140 bp upstream of *S. enterica casA*, the first gene of the *cascade* operon (Figure 4A). Thus, we amplified a long promoter sequence (206 bp) of *S. enterica casA* (referred to as $P_{secasA}$) to include this target site and placed *gfp* under its control. Our findings revealed that fluorescence intensity declined by ~8.6-fold when *Se*CreR was produced in *trans* (driven by the commonly used *tac* promoter) from another plasmid (Figure 4C).

As expected, this repression effect was abolished when *Se*CreR was mutated to disrupt its complementarity to $P_{secasA}$ and then reoccurred by further mutating $P_{secasA}$ to restore their complementarity (Figure 4D). To confirm that the *E. coli* Cascade played an instrumental role during this repression circuit, we constructed the Δ*hns*Δ*casA* mutant (*casA* is the first gene of the *E. coli cascade* operon), where *Se*CreR no longer repressed the activity of $P_{secasA}$ (Figure 4D). According to an early study, we had also created a *cascade* overexpressing mutant (named Δ*hns*/*cas*$^{OE}$) where the *cas3* gene not required for Cascade binding to target DNA was deleted and *cascade* genes were put under the control of the highly active J23119 promoter. This mutant exhibited a significantly heightened *Se*CreR-guided repression of $P_{secasA}$, measuring approximately 7- to 8.6-fold more intense than in Δ*hns* cells (Figure 4D). This confirms that the Cas3 helicase-nuclease is not necessary for the repression effect and also suggests that Cascade overexpression led to more stringent repression. Hence, *Se*CreR utilized *E. coli* Cascade effectively to repress its cognate $P_{casA}$, highlighting the ability of the *S. enterica* Cas homologs to utilize *Se*CreR for autorepression.

We then conducted a comprehensive study utilizing the advanced genetic tools of *E. coli* to investigate and characterize the connection between the activity of *Se*CreR-repressed $P_{secasA}$ and the concentration of crRNA molecules present in the cell (as demonstrated in Figure 4D). We engineered the *S. enterica creR* gene and the $P_{secasA}$-controlled *gfp* into one plasmid and then expressed a 14-spacer CRISPR array (cloned from the *E. coli* DH5α strain) from another plasmid. An isopropyl β-d-1-thiogalactopyranoside (IPTG)-inducible promoter (containing the *lacO* operator) was used to facilitate the induction of crRNA production. It is worth noting that we initially employed a less active mutant of *tac* promoter (*tacm3*)[32] to control CRISPR to avoid excessive basal expression. The exponential Δ*hns*/*cas*$^{OE}$ cells containing both plasmids underwent induction by various doses of IPTG. We observed an increase in the fluorescence intensity by a factor of 1.06, 1.31, 1.89, 2.64, and 2.96 when crRNA production was induced by 10, 20, 50, 100, and 200 μM IPTG, respectively (Figure 4E), which illustrated their positive correlation. Nevertheless, despite an increase of up to 10-fold being observed in a *Se*CreR mutation control, no further increase in fluorescence was observed with higher doses of IPTG (Figure 4E), indicating that the cellular concentration of crRNA or its relieving effect on *Se*CreR repression had reached saturation. Notably, when we changed $P_{tacm3}$ to $P_{tac}$, fluorescence could be further increased by 20%–60% at the same IPTG dose (with p values < 0.01), suggesting that crRNA concentration (or its promoter activity), rather than relieving effect, had reached saturation. We also varied the spacer number of the CRISPR array and further observed a positive correlation between spacer number and GFP fluorescence when a fixed dose of IPTG was used (Figure 4F). These data demonstrate that the cellular concentration of crRNA molecules is capable of fine-tuning the *Se*CreR-mediated Cas autorepression to meet their needs.

## V-A CreR-guided autorepression of Cas12a also monitors the crRNA pool

To validate the function of V-A CreR in *E. coli* cells, we selected the *creR* gene associated with the well-studied *M. bovoculi cas12a* (*Mbcas12a*) (see Figure 5A). This *creR* gene (referred to as *MbcreR*) contains a sequence (ΨR1) that is quite similar to its cognate CRISPR repeat, along with another "repeat" (ΨR2) that is highly degenerated and exhibits very little nucleotide identity to them (Figure 5B). In addition, the RNA of ΨR1 and CRISPR repeats form a similar stem-loop structure (Figure 5C) and share 20 consecutive nucleotides, giving rise to an identical 5′ handle on their mature RNAs (Figure 5B). These findings strongly suggest that mature *Mb*CreR has a crRNA-like architecture. The spacer portion of *Mb*CreR partially complements to a target site that is flanked by a 5′-TTTA-3′ motif (a typical PAM of V-A subtype) and positioned very close to the predicted −35 element of the promoter of *Mbcas12a* (designated as $P_{cas12a}$) (Figure 5A). From a primer extension assay, we determined the TSS of $P_{cas12a}$ (Figure S7) and confirmed the prediction of −35 element.

We initially designed two plasmids: one carried the *Mbcas12a* gene and its promoter, while the other additionally carried the *MbcreR* gene (Figure 5D). In *E. coli* Δ*hns* cells (to avoid potential repression effects on $P_{cas12a}$ from H-NS), the presence of *Mb*CreR led to a significant reduction (by ~98.4%) in *Mbcas12a* transcripts, suggesting that *Mb*CreR tightly repressed $P_{cas12a}$. To confirm this repression, we synthesized a 330-bp DNA construct that included *MbcreR* (and its putative promoter) and $P_{cas12a}$ and put *gfp* under its control (Figure 5E). In *E. coli* Δ*hns* cells, the plasmid carrying this DNA construct produced green fluorescence, which could be very effectively suppressed by ~99-fold (p = 5.74e−06) by the addition of another plasmid that in *trans*-provided *Mb*Cas12a (Figure 5F). This repression effect

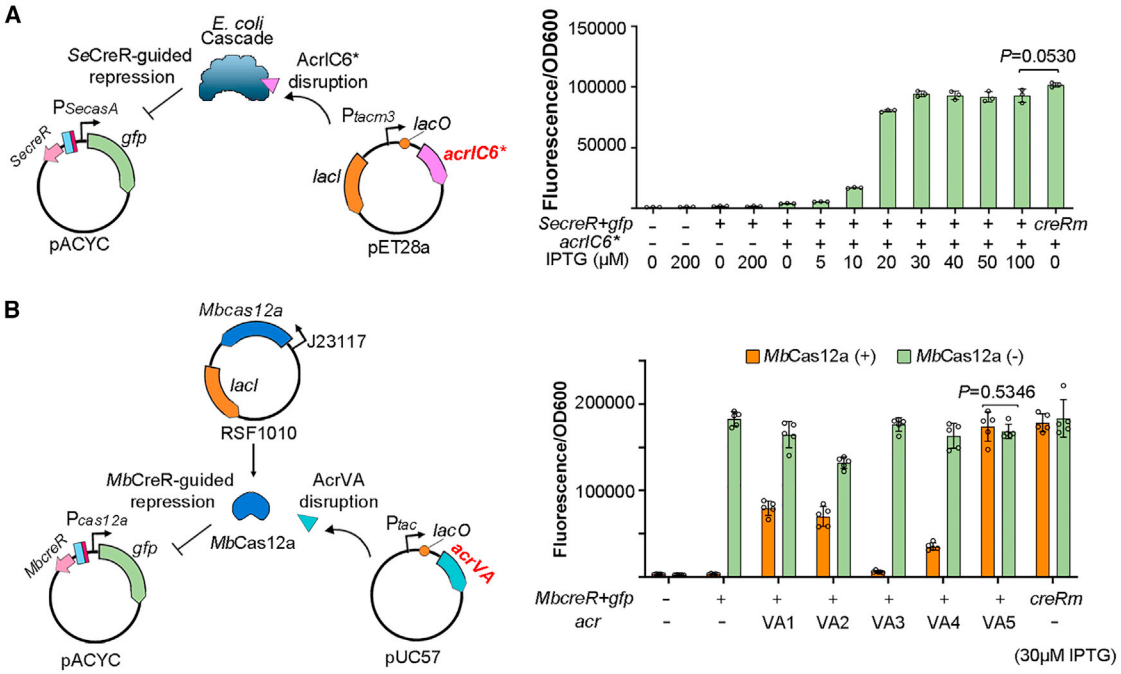



**Figure 6. CreR-guided Cas autorepression can be disrupted by diverse Acr proteins**

(A) Disrupting the autorepression circuit of I-E Cascade by expressing AcrIC6* (a dual CRISPR inhibitor that can inactivate the MG1655 Cascade). Error bars, mean ± SD (n = 3).

(B) Disrupting the autorepression circuit of *Mb*Cas12a by expressing diverse AcrVA proteins. An empty RSF1010-based vector was used for the *Mb*Cas12a (–) assay. Error bars, mean ± SD (n = 5).

p values were obtained through two-tailed Student's t test.

See also Figures S5 and S6.

disappeared when we mutated the spacer portion (ΨS) of *MbcreR* or its target site within $P_{cas12a}$ and, notably, persisted when *MbcreR* and $P_{cas12a}$ were complementarily mutated at the same time (Figure 5F). Therefore, *Mb*CreR guides *Mb*Cas12a to repress its *cas* promoter based on their limited complementarity.

Next, we introduced a third plasmid to express crRNAs to assess their impacts on *Mb*CreR-guided repression (see Figure 5E). The CRISPR array of *M. bovoculi* 237, which features 17 spacers, was synthesized and placed under the control of an IPTG-inducible promoter ($P_{tacm3}$). We then assessed the impact of varying concentrations of IPTG on the fluorescence in exponential *E. coli* Δ*hns* cells containing all these three plasmids (Figure 5G). We observed a positive correlation between increased IPTG concentrations and fluorescence, with saturation nearly being reached after 200 μM. Again, changing $P_{tacm3}$ to $P_{tac}$ further elevated the fluorescence by a factor of 1.5–2.0 at the same IPTG concentration (Figure 5G), illustrating that crRNA concentration, rather than the relieving effect, had reached saturation. Overall, we conclude that the *Mb*CreR-guided autorepression of *Mb*Cas12a has the ability to detect and respond to changes in the cellular concentration of crRNA molecules.

## Acr proteins can relieve or subvert CreR-guided Cas autorepression

To protect themselves against the diverse CRISPR-Cas systems, phages have developed a range of Acr proteins that are at least equally diverse.[33,34] We further investigated how the Cas autorepression responds to the action of Acr proteins. A dual inhibitor, AcrIC6*, which targets both I-E and I-C CRISPR effectors, has been reported for the I-E Cascade of MG1655.[35] So, we synthesized the *acrIC6** gene and expressed it using an IPTG-inducible *tacm3* promoter. The exponential *E. coli* Δ*hns* cells containing the plasmid expressing AcrIC6* and another plasmid carrying the *S. enterica creR* gene and a *gfp* gene controlled by $P_{secasA}$ were treated with different doses of IPTG, and fluorescence intensity was measured (Figure 6A). The results showed that inducing AcrIC6* expression with 5–30 μM IPTG increased fluorescence by a factor of 1.37–24.51, and higher doses did not result in further increases. Notably, the effect of 30–100 μM IPTG induction was comparable to that of *Se*CreR mutation, which completely subverted the repression circuit (Figure 6A). This suggests that even low expression levels of AcrIC6* can effectively relieve or even subvert the repression effect on $P_{secasA}$, indicating that the I-E Cas autorepression circuit can actively respond to the actions of Acr proteins to stimulate mass production of new Cas effectors.

At least five Acr proteins (named AcrVA1-5) have been identified that inhibit the V-A effector Cas12a,[36,37] which allowed us to explore the impact of different Acr proteins on the *Mb*Cas12a autoregulation circuit. Similar to the crRNA induction assay, we employed three plasmids: one plasmid carried *M. bovoculi creR* and $P_{cas12a}$-controlled *gfp*, the second plasmid encoded the *Mb*Cas12a effector, and the third plasmid produced AcrVA proteins under an inducible promoter (Figure 6B). In exponential

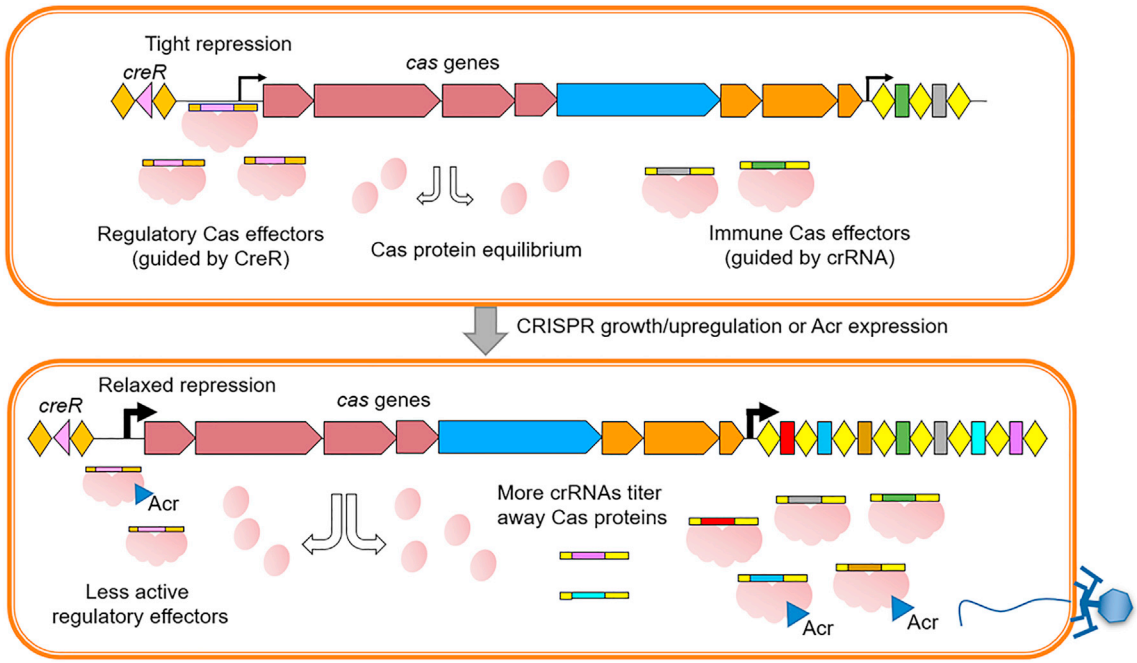



**Figure 7. The model of CreR-guided Cas autorepression**

In cells with a low level of crRNA production, CreR (or CreA) guides Cas effectors to tightly repress the promoter of *cas* genes. When the cellular concentration of crRNA elevates, possibly due to the growth or upregulation of CRISPR arrays, Cas proteins are titrated away from the regulatory circuit, and *cas* repression gets relaxed to promote the recovery of the pool of Cas proteins (until a new equilibrium is achieved). When the Acr proteins from an infecting phage inhibit Cas proteins, the Cas autorepression circuit will be subverted, leading to mass production of new immune effectors.
See also Figures S5 and S6.

*E. coli* Δ*hns* cells containing all these three plasmids, we observed that the *Mb*CreR-repressed fluorescence was markedly relieved by AcrVA1-5 proteins with varying effects (Figure 6B). Upon induction using 30 μM IPTG, AcrVA3 only increased the fluorescence intensity by a factor of 1.84 ($p = 0.0016$), while AcrVA1, AcrVA2, and AcrVA4 increased by a factor of 24.41, 21.48, and 10.77, respectively, with AcrVA5 showing the strongest effect (53.36-fold increase), which was comparable to the effect of disrupting the *Mb*CreR-$P_{cas12a}$ complementarity ($p = 0.5346$). Hence, the autoregulation circuit of *Mb*Cas12a responds differently to various AcrVA proteins, implying distinct Acr mechanisms for each. In fact, it was reported that these AcrVA proteins showed varying inhibiting effects on Cas12a proteins from different *M. bovoculi* strains.[37] To exercise caution, we further omitted the *Mb*Cas12a-expressing plasmid from this assay, and as anticipated, *Mb*CreR no longer suppressed the $P_{cas12a}$-controlled *gfp*, and there was no relieving effect on fluorescence for any AcrVA protein (Figure 6B).

## DISCUSSION

There are mounting evidence supporting that CRISPR-Cas effectors have a secondary physiological role in regulating host genes, in addition to its canonical immune function. For instance, recent studies have shown that the type II effector Cas9 can be reprogrammed by scaRNA to regulate a virulence-related gene that encodes a lipoprotein.[16] Similarly, the discovery of a long isoform of tracRNA (tracr-L) in *Streptococcus pyogenes* has demonstrated its pivotal role in directing the autorepression of Cas9.[20] Our research group has also revealed the regulatory function of Cascade, a type I effector that has been repurposed by a crRNA-resembling antitoxic (CreA) RNA to repress the expression of a toxic RNA (CreT).[21,22,27,31] This CreTA acts as an addiction module that can induce cell dormancy/death when the genes encoding any Cascade subunits are destroyed, thereby enhancing the persistence of CRISPR effectors within bacterial populations. In this study, we further showed that CreA and its CreR analogs (not cooccurring with a toxin gene) widely distribute in type I and V systems (assigned to class 1 and class 2, respectively[6]) and provided experimental evidence that these modulatory RNAs commonly direct the autorepression of both multi-subunit (I-B and I-E) and single-protein (V-A) Cas effectors. The recent study by Shmakov et al. also observed such a Cas autorepression circuit in a type I-F system,[25] and conducted a more comprehensive search for diverse crRNA-like RNAs (discussed below). It appears that in type I and V-A systems, CreR (or sometimes CreA) RNAs play a role akin to that of tracr-L in type II-A systems.[20] Therefore, it seems that repurposing Cas effectors by noncanonical RNA guides to achieve their autoregulation is a general paradigm for both class 1 and class 2 CRISPR-Cas systems.

Note that the genes of these regulatory RNA guides mentioned above typically contain only one or two CRISPR repeats, which are often markedly degenerated. In addition, they feature a spacer-like sequence that only partially complements with the

target promoter. These characteristics have posed challenges in the systematic discovery of such regulatory mini-CRISPRs. Nevertheless, Shmakov et al. recently conducted a comprehensive search for such crRNA-like (crl) RNAs and revealed their possible presence in approximately 15% of archaeal CRISPR-Cas loci and 12% in bacteria.[25] It is worth noting that the crlRNAs surveyed in that study include not only the gene-regulating RNAs but also encompass the tracrRNAs and scout RNAs, which do not regulate gene expression but instead play vital roles in crRNA maturation and target DNA interference. In addition, these two types of crlRNAs lack a spacer-like sequence but contain an anti-repeat that forms a duplex with the repeat portion of crRNAs. Therefore, the widespread crlRNAs display a variety of structural features and physiological functions that are mostly related to the canonical defense mechanism.[38]

We propose that the Cas autoregulation circuit has exquisitely balanced the benefits and downsides of CRISPR-Cas. In order to provide robust immunity, a sufficient level of Cas proteins must be produced to meet the changing needs of crRNA guides during the ongoing battle between bacteria and phages. Furthermore, if phages encoding small Acr proteins that inhibit Cas effectors infect the cell, a rapid increase in the production of new Cas proteins is necessary to quickly restore CRISPR immunity. However, maintaining a constant high expression level of the multi-subunit (class 1) or high molecular weight (class 2) Cas effector could lead to harmful autoimmune events (as demonstrated in Figure 2 when Cas autorepression was lost) and perhaps other negative effects that would compromise the host cell's fitness. The Cas autorepression circuit, which involves Cas proteins per se and a noncanonical guide RNA (CreR or CreA in type I and V-A, while tracr-L in type II-A) that competes with crRNAs, enables *cas* expression to be responsive to Acr elements that inhibit Cas proteins, as well as to alterations in the concentration of crRNAs within the cell (Figure 7). Note that, the repressed *cas* promoters need to be highly effective to facilitate the mass production of Cas proteins when needed. Consistently, in the absence of CreA, the *cas* promoter ($P_{cas8}$) in *H. hispanica* proved to be the most effective promoter among the haloarchaeal promoters we tested (Figure S1C). From the view of arms race, the Cas autorepression circuit may represent a distinct anti-Acr strategy that acts on transcriptional level.

Interestingly, the recent study of Shmakov et al. also revealed the occurrence of virus-encoded crlRNAs (likely CreR RNAs) that have the potential to target the *cas* promoter from their putative bacterial hosts.[25] The authors proposed that these *cas*-regulating crlRNAs may have been explored by viruses to counteract CRISPR immunity. Supporting this claim, we observed that overexpressing CreA RNAs in *H. hispanica* led to a significant reduction (~60%) in *cas* expression (see Figure 3D).

In summary, our data provide substantial experimental evidence for the Cas autoregulation circuit, which is intriguingly directed by modulatory crRNA-like RNAs (specifically, CreR or CreA). Remarkably, this circuit is able to adapt to the changing requirements of the typical defensive crRNAs and is susceptible to the influence of Acr proteins. We surmise that, in combination with the diverse CreT, this circuit might have promoted the wide distribution and stable persistence of CRISPR-Cas in prokaryotes.

## STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Strains and growth conditions
  - Purification and concentration of the virus
- METHOD DETAILS
  - Plasmid construction
  - Transformation
  - Mutant construction and gene knockout
  - Fluorescence measurement
  - RNA extraction
  - RT-qPCR
  - Northern blot analysis
  - Primer extension analysis
  - Competition assay
  - Virus interference assay
  - Spacer acquisition assay
  - Origin analysis of new spacers
  - Bioinformatic analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.chom.2023.08.005.

## ACKNOWLEDGMENTS

We thank Prof. Yanli Wang for sharing with us the plasmid DNA encoding the codon-optimized MbCas12a. This work was supported by the National Natural Science Foundation of China (32150020, 32230061, 32022003, 32200057, 32270092, and 31970544), the Strategic Priority Research Program of the Chinese Academy of Sciences (Precision Seed Design and Breeding) (XDA24000000), the Youth Innovation Promotion Association of CAS (2020090), the China National Postdoctoral Program for Innovative Talents (BX20220331), and the project funded by China Postdoctoral Science Foundation (2022M720160).

## AUTHOR CONTRIBUTIONS

M.L., R.W., C.L., and J.L. designed experiments. C.L., R.W., and J.L. constructed mutant strains with the assistance of F.C. and X.S. C.L. performed the northern blotting assay, spacer acquisition assay, and competition assay with the assistance of S.H. and H.Z. C.L. and R.W. performed the fluorescence analysis with the assistance of J.L., L.W., J.Y., and Y.Z. C.L. and F.C. carried out qPCR and transformation assays with the assistance of Q.X. and A.W. F.C. performed virus interference assay with the assistance of C.L. M.L., J.L., X.S., and H.Y. performed the bioinformatic analyses. M.L. and H.X. analyzed the data and supervised the project. M.L. wrote the manuscript.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: March 23, 2023
Revised: July 4, 2023
Accepted: August 9, 2023
Published: September 1, 2023

## REFERENCES

1. Barrangou, R., and Horvath, P. (2017). A decade of discovery: CRISPR functions and applications. Nat. Microbiol. *2*, 17092. https://doi.org/10.1038/nmicrobiol.2017.92.
2. Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. Science *321*, 960–964. https://doi.org/10.1126/science.1159689.
3. Wiedenheft, B., Sternberg, S.H., and Doudna, J.A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. Nature *482*, 331–338. https://doi.org/10.1038/nature10886.
4. Hille, F., Richter, H., Wong, S.P., Bratovič, M., Ressel, S., and Charpentier, E. (2018). The biology of CRISPR-Cas: backward and forward. Cell *172*, 1239–1259. https://doi.org/10.1016/j.cell.2017.11.032.
5. Nussenzweig, P.M., and Marraffini, L.A. (2020). Molecular mechanisms of CRISPR-Cas immunity in bacteria. Annu. Rev. Genet. *54*, 93–120. https://doi.org/10.1146/annurev-genet-022120-112523.
6. 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.
7. Semenova, E., Jore, M.M., Datsenko, K.A., Semenova, A., Westra, E.R., Wanner, B., van der Oost, J., Brouns, S.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.
8. Wiedenheft, B., van Duijn, E., Bultema, J.B., Waghmare, S.P., Zhou, K., Barendregt, A., Westphal, W., Heck, A.J., 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.
9. Sternberg, S.H., Richter, H., Charpentier, E., and Qimron, U. (2016). Adaptation in CRISPR-Cas systems. Mol. Cell *61*, 797–808. https://doi.org/10.1016/j.molcel.2016.01.030.
10. Li, M., Wang, R., Zhao, D., and Xiang, H. (2014). Adaptation of the *Haloarcula hispanica* CRISPR-Cas system to a purified virus strictly requires a priming process. Nucleic Acids Res. *42*, 2483–2492. https://doi.org/10.1093/nar/gkt1154.
11. Datsenko, K.A., Pougach, K., Tikhonov, A., Wanner, B.L., Severinov, K., and Semenova, E. (2012). Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. Nat. Commun. *3*, 945. https://doi.org/10.1038/ncomms1937.
12. Savitskaya, E., Lopatina, A., Medvedeva, S., Kapustin, M., Shmakov, S., Tikhonov, A., Artamonova, I.I., Logacheva, M., and Severinov, K. (2017). Dynamics of *Escherichia coli* type I-E CRISPR spacers over 42 000 years. Mol. Ecol. *26*, 2019–2026. https://doi.org/10.1111/mec.13961.
13. Levin, B.R., Moineau, S., Bushman, M., and Barrangou, R. (2013). The population and evolutionary dynamics of phage and bacteria with CRISPR-mediated immunity. PLoS Genet. *9*, e1003312. https://doi.org/10.1371/journal.pgen.1003312.
14. Stern, A., Keren, L., Wurtzel, O., Amitai, G., and Sorek, R. (2010). Self-targeting by CRISPR: gene regulation or autoimmunity? Trends Genet. *26*, 335–340. https://doi.org/10.1016/j.tig.2010.05.008.
15. Bikard, D., Hatoum-Aslan, A., Mucida, D., and Marraffini, L.A. (2012). CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. Cell Host Microbe *12*, 177–186. https://doi.org/10.1016/j.chom.2012.06.003.
16. Ratner, H.K., Escalera-Maurer, A., Le Rhun, A., Jaggavarapu, S., Wozniak, J.E., Crispell, E.K., Charpentier, E., and Weiss, D.S. (2019). Catalytically active Cas9 mediates transcriptional interference to facilitate bacterial virulence. Mol. Cell *75*, 498–510.e5. https://doi.org/10.1016/j.molcel.2019.05.029.
17. Chylinski, K., Le Rhun, A., and Charpentier, E. (2013). The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. RNA Biol. *10*, 726–737. https://doi.org/10.4161/rna.24321.
18. Liu, L., Chen, P., Wang, M., Li, X., Wang, J., Yin, M., and Wang, Y. (2017). C2c1-sgRNA complex structure reveals RNA-guided DNA cleavage mechanism. Mol. Cell *65*, 310–322. https://doi.org/10.1016/j.molcel.2016.11.040.
19. Faure, G., Shmakov, S.A., Makarova, K.S., Wolf, Y.I., Crawley, A.B., Barrangou, R., and Koonin, E.V. (2019). Comparative genomics and evolution of trans-activating RNAs in Class 2 CRISPR-Cas systems. RNA Biol. *16*, 435–448. https://doi.org/10.1080/15476286.2018.1493331.
20. Workman, R.E., Pammi, T., Nguyen, B.T.K., Graeff, L.W., Smith, E., Sebald, S.M., Stoltzfus, M.J., Euler, C.W., and Modell, J.W. (2021). A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression. Cell *184*, 675–688.e19. https://doi.org/10.1016/j.cell.2020.12.017.
21. Li, M., Gong, L., Cheng, F., Yu, H., Zhao, D., Wang, R., Wang, T., Zhang, S., Zhou, J., Shmakov, S.A., et al. (2021). Toxin-antitoxin RNA pairs safeguard CRISPR-Cas systems. Science *372*. https://doi.org/10.1126/science.abe5601.
22. Cheng, F., Wang, R., Yu, H., Liu, C., Yang, J., Xiang, H., and Li, M. (2021). Divergent degeneration of *creA* antitoxin genes from minimal CRISPRs and the convergent strategy of tRNA-sequestering CreT toxins. Nucleic Acids Res. *49*, 10677–10688. https://doi.org/10.1093/nar/gkab821.
23. Du, K., Gong, L., Li, M., Yu, H., and Xiang, H. (2022). Reprogramming the endogenous type I CRISPR-Cas system for simultaneous gene regulation and editing in *Haloarcula hispanica*. mLife *1*, 40–50. https://doi.org/10.1002/mlf2.12010.
24. Jurėnas, D., Fraikin, N., Goormaghtigh, F., and Van Melderen, L. (2022). Biology and evolution of bacterial toxin-antitoxin systems. Nat. Rev. Microbiol. *20*, 335–350. https://doi.org/10.1038/s41579-021-00661-1.
25. Shmakov, S.A., Barth, Z.K., Makarova, K.S., Wolf, Y.I., Brover, V., Peters, J.E., and Koonin, E.V. (2023). Widespread CRISPR-derived RNA regulatory elements in CRISPR-Cas systems. Nucleic Acids Res. gkad495. https://doi.org/10.1093/nar/gkad495.
26. Harrington, L.B., Ma, E., Chen, J.S., Witte, I.P., Gertz, D., Paez-Espino, D., Al-Shayeb, B., Kyrpides, N.C., Burstein, D., Banfield, J.F., et al. (2020). A scoutRNA is required for some type V CRISPR-Cas systems. Mol. Cell *79*, 416–424.e5. https://doi.org/10.1016/j.molcel.2020.06.022.
27. Cheng, F., Wu, A., Liu, C., Cao, X., Wang, R., Shu, X., Wang, L., Zhang, Y., Xiang, H., and Li, M. (2022). The toxin-antitoxin RNA guards of CRISPR-Cas evolved high specificity through repeat degeneration. Nucleic Acids Res. *50*, 9442–9452. https://doi.org/10.1093/nar/gkac712.
28. Cai, S., Cai, L., Zhao, D., Liu, G., Han, J., Zhou, J., and Xiang, H. (2015). A novel DNA-binding protein, PhaR, plays a central role in the regulation of polyhydroxyalkanoate accumulation and granule formation in the haloarchaeon *Haloferax mediterranei*. Appl. Environ. Microbiol. *81*, 373–385. https://doi.org/10.1128/AEM.02878-14.
29. Gong, L., Li, M., Cheng, F., Zhao, D., Chen, Y., and Xiang, H. (2019). Primed adaptation tolerates extensive structural and size variations of the CRISPR RNA guide in *Haloarcula hispanica*. Nucleic Acids Res. *47*, 5880–5891. https://doi.org/10.1093/nar/gkz244.
30. Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., et al. (2015). Cpf1 is a single RNA-guided endonuclease of

a class 2 CRISPR-Cas system. Cell *163*, 759–771. https://doi.org/10.1016/j.cell.2015.09.038.

31. Wang, R., Shu, X., Zhao, H., Xue, Q., Liu, C., Wu, A., Cheng, F., Wang, L., Zhang, Y., Feng, J., et al. (2023). Associate toxin-antitoxin with CRISPR-Cas to kill multidrug-resistant pathogens. Nat. Commun. *14*, 2078. https://doi.org/10.1038/s41467-023-37789-y.
32. Zhang, H.M., Chen, S., Shi, H., Ji, W., Zong, Y., Ouyang, Q., and Lou, C. (2016). Measurements of gene expression at steady state improve the predictability of part assembly. ACS Synth. Biol. *5*, 269–273. https://doi.org/10.1021/acssynbio.5b00156.
33. Borges, A.L., Davidson, A.R., and Bondy-Denomy, J. (2017). The discovery, mechanisms, and evolutionary impact of anti-CRISPRs. Annu. Rev. Virol. *4*, 37–59. https://doi.org/10.1146/annurev-virology-101416-041616.
34. Pawluk, A., Davidson, A.R., and Maxwell, K.L. (2018). Anti-CRISPR: discovery, mechanism and function. Nat. Rev. Microbiol. *16*, 12–17. https://doi.org/10.1038/nrmicro.2017.120.
35. León, L.M., Park, A.E., Borges, A.L., Zhang, J.Y., and Bondy-Denomy, J. (2021). Mobile element warfare via CRISPR and anti-CRISPR in *Pseudomonas aeruginosa*. Nucleic Acids Res. *49*, 2114–2125. https://doi.org/10.1093/nar/gkab006.
36. Watters, K.E., Fellmann, C., Bai, H.B., Ren, S.M., and Doudna, J.A. (2018). Systematic discovery of natural CRISPR-Cas12a inhibitors. Science *362*, 236–239. https://doi.org/10.1126/science.aau5138.
37. Marino, N.D., Zhang, J.Y., Borges, A.L., Sousa, A.A., Leon, L.M., Rauch, B.J., Walton, R.T., Berry, J.D., Joung, J.K., Kleinstiver, B.P., et al. (2018). Discovery of widespread type I and type V CRISPR-Cas inhibitors. Science *362*, 240–242. https://doi.org/10.1126/science.aau5174.
38. Koonin, E.V., and Makarova, K.S. (2022). Evolutionary plasticity and functional versatility of CRISPR systems. PLoS Biol. *20*, e3001481. https://doi.org/10.1371/journal.pbio.3001481.
39. Jensen, K.F. (1993). The *Escherichia coli* K-12 "wild types" W3110 and MG1655 have an *rph* frameshift mutation that leads to pyrimidine starvation due to low *pyrE* expression levels. J. Bacteriol. *175*, 3401–3407. https://doi.org/10.1128/jb.175.11.3401-3407.1993.
40. Westra, E.R., Pul, U., Heidrich, N., Jore, M.M., Lundgren, M., Stratmann, T., Wurm, R., Raine, A., Mescher, M., van Heereveld, L., et al. (2010). H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. Mol. Microbiol. *77*, 1380–1393. https://doi.org/10.1111/j.1365-2958.2010.07315.x.
41. Liu, H.L., Han, J., Liu, X.Q., Zhou, J., and Xiang, H. (2011). Development of *pyrF*-based gene knockout systems for genome-wide manipulation of the archaea *Haloferax mediterranei* and *Haloarcula hispanica*. J. Genet. Genomics *38*, 261–269. https://doi.org/10.1016/j.jgg.2011.05.003.
42. Doublet, B., Douard, G., Targant, H., Meunier, D., Madec, J.Y., and Cloeckaert, A. (2008). Antibiotic marker modifications of lambda Red and FLP helper plasmids, pKD46 and pCP20, for inactivation of chromosomal genes using PCR products in multidrug-resistant strains. J. Microbiol. Methods *75*, 359–361. https://doi.org/10.1016/j.mimet.2008.06.010.
43. 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.
44. Reuter, C.J., and Maupin-Furlow, J.A. (2004). Analysis of proteasome-dependent proteolysis in *Haloferax volcanii* cells, using short-lived green fluorescent proteins. Appl. Environ. Microbiol. *70*, 7530–7538. https://doi.org/10.1128/AEM.70.12.7530-7538.2004.
45. Heler, R., Samai, P., Modell, J.W., Weiner, C., Goldberg, G.W., Bikard, D., and Marraffini, L.A. (2015). Cas9 specifies functional viral targets during CRISPR-Cas adaptation. Nature *519*, 199–202. https://doi.org/10.1038/nature14245.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| **Bacterial and virus strains** | | |
| *E. coli* DH5α | Tsingke Biological Technology, Beijing, China | CAT# TSC-C01 |
| *E. coli* MG1655 | Jensen[39] | N/A |
| *E. coli* MG1655 Δ*hns* | Westra et al.[40] | N/A |
| *E. coli* MG1655 Δ*hns*/*cas*$^{OE}$ | This study | N/A |
| *E. coli* MG1655 *Δhns*/*ΔcasA* | This study | N/A |
| *H. hispanica* ATCC 33960 | American Type Culture Collection (ATCC) | ATCC 33960 |
| *H. hispanica* DF60 (*ΔpyrF*) | Liu et al.[41] | N/A |
| *H. hispanica* TAdm | This study | N/A |
| *H. hispanica* Tm | This study | N/A |
| *H. hispanica* ΔCRISPR | Li et al.[10] | N/A |
| *H. hispanica* Tm-ΔCRISPR | This study | N/A |
| *H. hispanica* TAdm-ΔCRISPR | This study | N/A |
| *H. hispanica* ΔTA | Li et al.[21] | N/A |
| *H. hispanica* ΔTAΔ*cas6* | Li et al.[21] | N/A |
| *H. hispanica* Δsp2-13 | Li et al.[10] | N/A |
| *Haloarcula hispanica* pleomorphic virus 2 | Li et al.[10] | N/A |
| **Chemicals** | | |
| Uracil | Aladdin Biochemical Technology, Shanghai, China | Cat#U128370-500g |
| 5-fluoroorotic acid (5-FOA) | Shanghai yuanye Bio-Technology, Shanghai, China | Cat# Y30721-1g |
| Isopropyl β-d-1-thiogalactopyranoside (IPTG) | Inalco, CA, USA | Cat# 1758-1400 |
| Ampicillin | Shanghai Macklin Biochemical Co., Ltd., Shanghai, China | Cat# A830931-25g |
| Apramycin sulfate | Sangon Biotech, Shanghai, China | Cat# A600090 |
| Kanamycin | Shanghai Acmec Biochemical, Shanghai, China | Cat# K12100-5g |
| Chloramphenicol | Shanghai Macklin Biochemical Co., Ltd., Shanghai, China | C804169-25g |
| 2 × Taq Master Mix (Dye Plus) | Vazyme Biotech, Nanjing, China | Cat# P112-01 |
| Phanta Super-Fidelity DNA Polymerase | Vazyme Biotech, Nanjing, China | Cat# P505-d2 |
| T4 DNA ligase | New England Biolabs, MA, USA | Cat#M0202M |
| Random Hexamer Primer | Thermo Fisher Scientific, MA, USA | Cat#SO142 |
| Century-Plus RNA ladder | Thermo Fisher Scientific, MA, USA | Cat# AM7145 |
| Restriction Endonuclease | New England Biolabs, MA, USA | N/A |
| RNA loading dye | Thermo Fisher Scientific, MA, USA | Cat#R0641 |
| Biodyne B nylon membrane | Pall, NY, USA | Cat#60208 |
| DNase I, RNase-free, HC | Thermo Fisher Scientific, MA, USA | Cat# EN0523 |
| TRIzol reagent | Thermo Fisher Scientific, MA, USA | Cat#15596-026 |
| Agarose | Sigma-Aldrich, MO, USA | V900510 |
| **Critical commercial assays** | | |
| Gel Extraction Kit | Omega BIO-TEK, GA, USA | Cat# D2500-02 |
| GeneJET Plasmid Miniprep Kit | Thermo Fisher Scientific, MA, USA | Cat# K0503 |

(*Continued on next page*)

| *Continued* | | |
|---|---|---|
| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
| Century-Plus RNA ladder | Thermo Fisher Scientific, MA, USA | Cat# AM7145 |
| Chemiluminescent Nucleic Acid Detection Module Kit | Thermo Fisher Scientific, MA, USA | Cat#89880 |
| KAPA SYBR® FAST qPCR Kit | Kapa Biosystems, MA, USA | Cat#KK4600 |
| Moloney Murine Leukemia Virus reverse transcriptase kit (M-MLV RT) | Promega, WI, USA | Cat#M1701 |
| Trelief® Seamless Cloning Kit | Tsingke Biological Technology, Beijing, China | Cat# TSV-S3 |
| Deposited data | | |
| Raw Illumina data of Tm-TAdm competition | This study | SRA: PRJNA984824 |
| Raw Illumina data of new spacers | This study | SRA: PRJNA984824 |
| Oligonucleotides | | |
| See "Oligonucleotides" in Table S2 | This study | N/A |
| Synthetic gene | | |
| See "Gene synthesis" in Table S2 | This study | N/A |
| Recombinant DNA | | |
| See "Plasmids" in Table S2 | This study | N/A |
| Software and Algorithms | | |
| GraphPad Prism 9 | Insightful Science | https://www.graphpad.com/scientific-software/prism/ |
| Quantity One software | Bio-Rad, CA, USA | N/A |
| Peak Scanner Software v1.0 | Applied Biosystems™, CA, USA | N/A |
| RNAfold | | http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi |
| BPROM | Softberry | http://www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb |

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ming Li (lim_im@im.ac.cn).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

- Raw Illumina FASTQ data have been deposited at the Sequence Read Archive (SRA) database (BioProject ID PRJNA984824) and are publicly available as of the date of publication. Accession number is also listed in the key resources table.
- This paper does not report original code.
- 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

### Strains and growth conditions

*Haloarcula hispanica* DF60 (an uracil auxotroph mutant of *H. hispanica* ATCC 33960)[41] and its derivatives were cultivated at 37°C in nutrient-rich AS-168 medium, supplemented with uracil at a final concentration of 50 mg/liter. The AS-168 medium contained per liter: 200 g NaCl, 20 g $MgSO_4 \cdot 7H_2O$, 2 g KCl, 3 g trisodium citrate, 1 g sodium glutamate, 50 mg $FeSO_4 \cdot 7H_2O$, 0.36 mg $MnCl_2 \cdot 4H_2O$, 5 g Bacto Casamino Acids, and 5 g yeast extract; pH was adjusted to be around neutral (pH=7.2). For the strains carrying the expression plasmid pWL502 or its derivatives, AS-168 medium without yeast extract was utilized for cultivation. All strains were cultured either on solid agar plates (1.2% agar) or in liquid cultures.

*E. coli* DH5α was utilized for plasmid construction, while *E. coli* MG1655 and its mutants were employed as hosts to investigate the I-E CreR from *S. enterica* ATCC 51960 or the V-A CreR from *M. bovoculi* 237. All bacterial strains were cultured at 37°C in Luria-Bertani (LB) medium, which contained 10 g tryptone, 5 g yeast extract, and 10 g NaCl per liter. Solid plates were prepared with agar at a concentration of 12 g/L. Liquid cultures were agitated at a rate of 200 rpm. When necessary, antibiotics were added at the following final concentrations: apramycin sulfate (50 μg/mL), ampicillin (100 μg/mL), kanamycin (50 μg/mL), or chloramphenicol (25 μg/mL).

### Purification and concentration of the virus

The top agar from a single plaque containing HHPV-2 virions was transferred into an early exponential culture of *H. hispanica* for virus enrichment. The culture was cultivated for 5 days at 37°C with agitation at 200 rpm. After cultivation, the culture was collected and the cells were removed through centrifugation at 9,000 rpm for 15 min at 4°C. The supernatant was purified using the VIVAFLOW 50 system (Sartorius, 50,000 MWCO) for pre-purification and subsequently filtered through 0.22 μm PES filters.

## METHOD DETAILS

### Plasmid construction

The plasmids, oligonucleotides, and synthetic genes used in this study are listed in Table S2. The double-stranded DNA fragments were amplified using Phanta Super-Fidelity DNA polymerase (Vazyme Biotech, Nanjing, China), digested with restriction enzymes (New England Biolabs, MA, USA), and then ligated into the pre-digested vector using T4 DNA ligase (New England Biolabs, MA, USA). Alternatively, they were directly assembled into predigested plasmids through the Gibson assembly strategy using Trelief® Seamless Cloning Kit (Tsingke, Beijing, China). The overlap extension PCR strategy was employed for mutant construction, and the engineered plasmids were confirmed by DNA sequencing.

### Transformation

The transformation of haloarchaeal cells was conducted using the polyethylene glycol-mediated method according to the online Halohandbook (https://haloarchaea.com/wp-content/uploads/2018/10/Halohandbook_2009_v7.3mds.pdf). Transformants were screened on yeast extract-subtracted AS-168 plates and transformation efficiency (colony forming unit per μg plasmid DNA, CFU/μg) was calculated. The results were then log-transformed to obtain the average and standard deviation.

To introduce plasmids into *E. coli* cells, the electro-transformation method was utilized. Bacterial cells were cultured in 3 mL LB broth at 37°C overnight with 200 rpm shaking, and then sub-inoculated into 200 mL fresh medium (1:100 dilution). After the optical density at 600 nm (OD600) reached approximately 0.6, cells were collected by centrifugation at 4°C, washed twice using cold double-distilled water, and finally resuspended in ice-cold 10% glycerol. A total of 50 μL of cells were mixed with plasmid DNA and subjected to electroporation using a Bio-Rad electroporator at 2.5 kV. The shocked cells were subsequently recovered in 1 mL LB medium at 37°C for one hour before being plated on LB agar plates containing appropriate antibiotics. This procedure was repeated to introduce multiple plasmids into *E. coli* cells.

### Mutant construction and gene knockout

Plasmids and oligonucleotides are listed in Table S2. The construction of haloarchaeal mutants followed the method described previously.[41] For instance, to create the TAdm mutant, we initially amplified the *creTA* sequence (NC_015943.1: 145387-145697) and then employed overlap extension PCR to mutate the Shine-Dalgarno (SD) motif of *creT* and the first two nucleotides of the *creA* seed sequence. Subsequently, we separately amplified the upstream ~500 bps and downstream ~500 bps and connected these three fragments using overlap extension PCR. The final DNA products was digested and inserted into the suicide plasmid pHAR,[41] followed by validation through DNA sequencing before introducing the recombinant plasmid into *H. hispanica* ΔTA cells. After screening for single and double cross-over mutants, the TAdm mutant cells were confirmed by colony PCR and subsequent Sanger sequencing.

*E. coli* mutants were constructed based on the Δ*hns* mutant of MG1655 using the λ-Red and FLP/FRT systems, as previously described.[42,43] For example, to create the Δ*hns*/*cas*$^{OE}$ mutant (where the *cas3* gene was replaced with $P_{J23119}$ to overexpress the downstream *cascade* genes), a DNA construct was synthesized that contained the 40 bp upstream of *cas3*, a kanamycin resistance gene (*kan*$^r$) flanked by two FRT sequences, a $P_{J23119}$ sequence, and the 40 bp downstream of $P_{casA}$, according to an earlier study.[43] This construct was then electroporated into Δ*hns* cells that contained pKD46 (which had been pretreated with 1% L-arabinose to induce the expression of the λ-Red system). The kanamycin-resistant colonies were screened and verified using colony PCR. Subsequently, *kan*$^r$ was eliminated by using pCP20, which encodes the flippase, and pCP20 (temperature sensitive) was then cured by growing at 42°C. The resulting mutants were further validated through colony PCR and subsequent Sanger sequencing.

### Fluorescence measurement

To evaluate the activity of haloarchaeal promoters, we employed a modified *gfp* gene that encodes a soluble red-shifted variant of green fluorescence protein.[44] Subsequently, we constructed a DNA sequence comprising the promoter being tested and the modified *gfp* gene, which was then incorporated into the expression vector pWL502. The resulting recombinant plasmid was subsequently introduced into *H. hispanica* cells through transformation for measuring fluorescence intensity.

To evaluate the activity of $P_{secasA}$, we chose the *E. coli* MG1655 Δ*hns* strain or its derivates to avoid potential inhibitory effects of H-NS on the promoter being investigated. The 293-bp DNA segment upstream of the *S. enterica casA* (including *SecreR*) and the *gfp* gene were separately amplified and then assembled into pACYC. Subsequently, the constructed GFP reporter plasmid, or its derivates with mutations in *SecreR* and/or $P_{secasA}$, were introduced into *E. coli* cells for fluorescence measurement.

To evaluate the activity of $P_{cas12a}$, the 330-bp DNA sequence preceding the *M. bovoculi cas12a* (containing its own promoter $P_{cas12a}$ and the *MbcreR* gene) and the *gfp* gene were separately amplified and assembled into the pACYC vector. Then, this plasmid, along with a second plasmid expressing *Mb*Cas12a (driven by $P_{J23117}$), was introduced into the *E. coli* Δ*hns* cells for measuring fluorescence intensity.

To assess the impact of crRNA or Acr on the CreR-repressed promoters, we synthesized the DNA fragment containing $P_{tac}$ or $P_{tacm3}$ and the *lac* operator, and another fragment containing a CRISPR array (or an *acr* gene). These two DNA fragments were then assembled into pET28a or pUC57. The constructed plasmids were introduced into *E. coli* cells containing the GFP reporter plasmid for fluorescence measurement.

For each experimental setting, we randomly selected at least three individual colonies and cultured them in LB medium with appropriate antibiotics (for *E. coli*) or in yeast extract-subtracted AS-168 medium (for *H. hispanica*) until the exponential phase. If necessary, the culture was sub-inoculated into fresh medium (at a 1:100 ratio) with different concentrations of IPTG. When OD600 reached approximately 0.8 (for *E. coli*) or 0.4 (for *H. hispanica*), fluorescence intensity and OD600 were measured simultaneously using the Synergy H4 Hybrid multimode microplate reader (BioTeck, VT, USA). The fluorescence/OD600 ratio was calculated for each biological replicate, and the average and standard deviation were calculated accordingly. For statistical analysis, a two-tailed Student's *t*-test was performed to determine the *P* values.

### RNA extraction

To extract total RNA from *H. hispanica* cells, three independent colonies were randomly selected for each experimental setting. These colonies were then inoculated into 10 mL of AS-168 or yeast extract-subtracted AS-168 medium. When the cultures reached the stationary phase, 100 μL of each culture was sub-inoculated into 10 mL of fresh medium and cultivated for an additional two days. For the extraction of *E. coli* RNA, 30 μL of the late exponential culture was transferred to a 3 mL fresh LB culture containing the appropriate antibiotics. After cultivation for 5 hours, the cells were harvested by centrifugation. Total RNA was then extracted from *H. hispanica* or *E. coli* cells using the TRIzol reagent (Invitrogen, MA, USA), following the standard guidelines. The concentration of RNA was determined using a NanoDrop One spectrophotometer (Thermo Fisher Scientific, MA, USA).

### RT-qPCR

To assess the transcriptional levels of *cas* genes, a total of 20 μg of RNA was treated with 20 U of RNase-free DNase I (Thermo Fisher Scientific, MA, USA) to eliminate any DNA contamination in the samples as instructed by the manufacturer. Subsequently, 3 μg of the pretreated RNA was reverse transcribed into cDNA using the Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT) (Promega, USA). Then, 5 μL of the ten-fold diluted cDNA was utilized as input in a 20 μL reaction volume. The qPCR reaction was carried out using the KAPA SYBR® FAST qPCR Kit (Kapa Biosystems, MA, USA) on an Applied Biosystems ViiA™ 7 Real-Time PCR System. The triplicates were analyzed for each assay based on three independent samples. The primer sequences used for qPCR can be found in Table S2.

### Northern blot analysis

A total of 3 μg of RNA was denatured at 65°C for 10 min with an equal volume of RNA loading dye (New England Biolabs, MA, USA), and then rapidly cooled on ice for 2 min. The RNA samples, the Century-Plus RNA ladder (Thermo Fisher Scientific, MA, USA), and a biotin-labeled 64-nt single-stranded DNA (ssDNA) were loaded onto an 8% polyacrylamide gel (7.6 M urea) and electrophoresed in 1× TBE buffer at 180 V for 45 min. The lane containing the Century-Plus RNA ladder was excised, stained with Ultra GelRed and then imaged using the GenoSens 2000 system (CLiNX, Shanghai, China). Next, the separated RNA samples and the biotin-labeled 64-nt ssDNA were transferred onto a Biodyne B nylon membrane (Pall, NY, USA) using a Mini-Protean Tetra system (Bio-Rad, CA, USA), and then cross-linked with Ultraviolet (UV) light. After pre-hybridization at 42°C, hybridization was performed using biotin-labeled probes (listed in Table S2), and the signal was detected using the Chemiluminescent Nucleic Acid Detection Module Kit (Thermo Fisher Scientific, MA, USA) following the manufacturer's instructions. The membrane was imaged using the Tanon 5200 Multi chemiluminescent imaging system (Tanon Science & Technology, Shanghai, China). 7*S* RNA was used as an internal control.

### Primer extension analysis

The 5′-FAM (6-carboxyfluorescein)-labeled *gfp*-specific primer (refer to Table S2) was ordered from Sangon Biotech Co., Ltd (Shanghai, China). 30 μg of total RNA was digested with 30 U of the DNase I (Thermo Fisher Scientific, MA, USA) for 1h, and re-purified following the phenol: chloroform method. Approximately 5 μg of the total RNA was mixed with 2.5 μg of the labeled primer, and reverse transcription was performed using 200 U of the M-MLV reverse transcriptase (Promega, WI, USA). The extension products were screened using the ABI 3730xl DNA Analyzer (Thermo Fisher Scientific, MA, USA), and the results were visualized using Peak Scanner Software v1.0.

### Competition assay

For each *H. hispanica* mutant (Tm or TAdm), three individual colonies were randomly selected and separately inoculated into fresh AS-168 medium (note that Tm-ΔCRISPR and TAdm-ΔCRISPR mutants were used for the competition assay without CRISPR immunity). Once cultivated to the exponential phase, the cultures were sub-inoculated into fresh medium and allowed to grow until the OD600 reached approximately 1.0. After adjusting the cell concentrations to a similar level, the three Tm and three TAdm cultures were randomly paired and mixed in equal volumes (150 μL each). The resulting three batches of Tm-TAdm mixture were then inoculated into 3 mL of fresh AS-168 medium, cultivated at 37 °C, and passaged every 3 or 7 days at a 1:20 ratio. At various time points, 1 mL of the Tm-TAdm co-culture was sampled from each biological replicate. Cells were collected by centrifugation, stored at -80°C, and genomic DNA was extracted using the phenol: chloroform: isoamyl alcohol (25:24:1, pH=8.0) method. The DNA was quantified using the NanoDrop One spectrophotometer (Thermo Fisher Scientific, MA, USA). To determine the ratio of Tm and TAdm cells, either Illumina sequencing or qPCR assay was then employed.

For Illumina sequencing, the genomic DNA was fragmented by sonication to a size of 350 bp. The fragmented DNA was then used for DNA library construction using the NEB Next® Ultra TM DNA Library Prep Kit (NEB, USA, Catalog#: E7370L), following the manufacturer's instructions. The DNA library was analyzed using the Illumina platforms with the PE150 strategy at Novogene Bioinformatics Technology Co., Ltd (Beijing, China). After removing adapter sequences and low-quality reads, the reads containing the wild sequence (Tm) or the mutated sequence (TAdm) of *creA* were separately retrieved and counted to determine their ratio.

To determine the ratio of Tm and TAdm cells using the qPCR assay, we first generated a standard curve by mixing pure cultures of Tm and TAdm at different ratios (10:0, 9:1, 8:2, 7:3, 6:4, 5:5, 4:6, and 0:10). Each mixed sample was analyzed in triplicate, and 2.5 ng of genomic DNA was used for each reaction. For the qPCR assay, we used the KAPA SYBR® FAST qPCR Kit (Kapa Biosystems, MA, USA) and ran the reactions on an Applied Biosystems ViiA™ 7 Real-Time PCR System following the manufacturer's instruction. We simultaneously probed the DNA sequences of *creA* and *cas8* (used as the internal control) using gene-specific primers (listed in Table S2). It is important to note that, to distinguish the mutated *creA* in TAdm cells and the wide-type *creA* gene in Tm cells, we designed the two 3'-terminal nucleotides of the reverse primer (Ra) to specifically match the mutated sequence in TAdm (see Figure S4A). The 2^(-ΔCt) value (*creA* vs *cas8*) was calculated for each sample, and then plotted against the predetermined percentages of TAdm cells to generate the standard curve. Subsequently, we analyzed the genomic DNA samples from the Tm-TAdm co-cultures using qPCR, along with the DNA sample from pure TAdm culture (the 2^(-ΔCt) value of which was used for normalization). Based on the standard curve, we calculated the percentage of TAdm cells for each sample, using their normalized 2^(-ΔCt) values.

### Virus interference assay

For each mutant of haloarchaea, three individual colonies were randomly selected and separately inoculated into yeast extract-subtracted AS-168 medium. After sub-inoculation and culturing for an additional 2 days, 200 μL of the culture were mixed with 100 μL of HHPV-2 virus at 10-fold serial dilutions and incubated at room temperature for 30 min. The mixture was then mixed with molten 0.7% agar medium that had been pre-incubated at 55°C, and immediately poured onto plates containing 1.2% agar. Once dried, the plates were incubated at 37°C for three days to allow plaque formation. The number of plaque-forming units (PFU) was counted and the ratio of PFUs formed on strain carrying the empty plasmid versus those formed on the strain expressing crRNAs was used to represent the relative virus immunity (RVI). The average RVI values with standard deviations were calculated from three biological replicates.

### Spacer acquisition assay

To monitor the acquisition of spacers from the viral HHPV-2 DNA, the crRNA of spacer13 (which shares approximately 70% sequence identity with HHPV-2 and is a promising candidate for efficient spacer acquisition from this virus) was overexpressed using the strong promoter $P_{phaR}$. This plasmid was then introduced into the WT, Tm, or TAdm mutant of *H. hispanica*. Three individual colonies were randomly selected and separately inoculated into 3 mL of fresh yeast extract-subtracted AS-168 medium, and the cultures were allowed to reach the stationary stage at 37°C. Sub-inoculation was performed at a ratio of 1:100 when OD600 reached approximately 1.0. Then, 100 μL of the exponential cultures were mixed with an equal volume of the virus HHPV-2 at either a low or high MOI (0.1 or 40), and the mixture was inoculated into 3 mL of fresh yeast extract-subtracted AS-168 medium. After 24 hours of culturing, colony PCR was performed for each inoculation. Briefly, 200 μL of the cell culture was centrifuged at 12,000 rpm for 1 min, then the sediment was lysed with 200 μL of distilled water, and 0.5 μL of the lysate was used as the template for PCR reactions. Quantitative analysis of CRISPR expansion was performed using the Quantity One software (Bio-Rad, CA, USA). For each gel lane, appropriate sensitivity was applied to detect parental and amplified bands after background subtraction. Gaussian modelling was then conducted on these bands to determine the number of parental or expanded band(s) present under the fitted curve. For each lane, the intensity of expanded band(s) was divided by that of all band(s) to calculate the percentage of expanded products. At least three biological replicates were examined for each mutant, and the average percentage and standard deviation were calculated.

To facilitate detecting naïve adaptation without virus infections, the empty pWL502 vector was introduced into the WT, Tm, and TAdm cells of *H. hispanica*. After transformation, the colonies were incubated at room temperature for approximately one month prior to colony PCR. The adapt-test_F and adapt-test_R primers (against the CRISPR leader and the first spacer, respectively; listed in Table S2) were used to analyze the expansion of the chromosomal CRISPR array due to new spacer acquisition. The PCR products were then electrophoresed using a 1.5% agarose gel, and the gel was imaged. At least three colonies were tested for each strain, and only representative gel images are presented.

Additionally, we further used the ΔCRISPR, Tm-ΔCRISPR and TAdm-ΔCRISPR mutants to exclude primed adaptation, and, by transformation, introduced an adaptation plasmid carrying the CRISPR leader and one CRISPR repeat sequence to monitor spacer acquisition. After approximately 6 days of incubation until clones formed, the plates were left at room temperature for 2-5 days prior to colony PCR. Three individual colonies were randomly selected for each CRISPR mutant, and the primers adapt-test_F and adapt-test_R' (listed in Table S2) were used. It is worth noting that the terminal nucleotide of the reverse primer adapt-test_R' (refer to Table S2) was specifically designed to mismatch the CRISPR leader sequence in order to enhance the detection efficiency of spacer acquisition events.[45]

### Origin analysis of new spacers

For each mutant, at least three individual colonies were selected for the spacer acquisition assay. The PCR products of three replicates were pooled and electrophoresed on agarose gels. Subsequently, the PCR band(s) corresponding to the 'expanded' CRISPR were purified using the E.Z.N.A.™ Gel Extraction kit (Omega Bio-tek, GA, US), following the manufacturer's instructions. The purified DNA fragments were directly used for DNA library construction using the NEB Next Ultra TM DNA Library Prep Kit (NEB, USA, Catalog#: E7370L), according to the manufacturer's instructions. The DNA library was then analyzed using the Illumina platforms with the PE150 strategy at Novogene Bioinformatics Technology Co., Ltd (Beijing, China). After removing adapter sequences and low-quality reads, the acquired spacers were determined by analyzing the sequence between each two repeats for each read. The acquired spacers were then aligned to the *H. hispanica* genome, the plasmid DNA, or the HHPV-2 genome, to determine the percentage of spacers acquired from each source.

### Bioinformatic analysis

RNA secondary structure was predicted using the RNAfold webserver. Promoter elements were predicted using the BPROM program (Softberry tool).

## QUANTIFICATION AND STATISTICAL ANALYSIS

Microsoft Excel was used to analyze the data, and GraphPad Prism was used to generate the plots. The graphs were then modified in Adobe photoshop to construct the final figures. Quantitative analysis of CRISPR expansion was performed using the Quantity One software. The number of replicates is specified in the associated figure legends. Each replicate represents a biological replicate of the specified experiment. Two-tailed *t* test was performed for statistical analyses. P-values above 0.05 were considered non-significant. Statistical comparisons for the transformation assays relied on log values, which assumes the samples are normally distributed on a log scale.