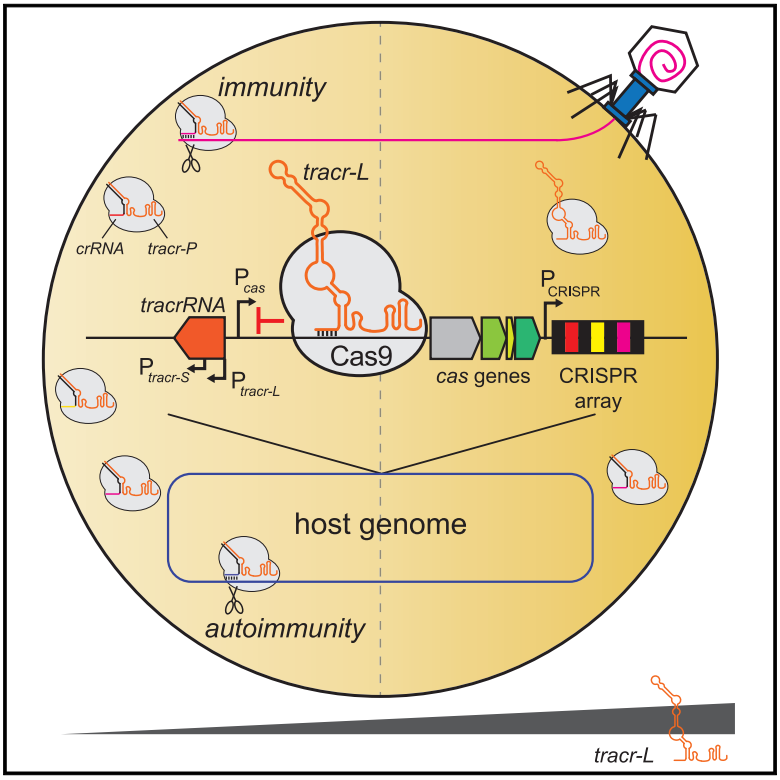
Cell

A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression

Graphical abstract



Authors

Rachael E. Workman, Teja Pammi, Binh T.K. Nguyen, ..., Marie J. Stoltzfus, Chad W. Euler, Joshua W. Modell

Correspondence

jmodell1@jhmi.edu

In Brief

Workman et al. discover a natural single-guide RNA for *S. pyogenes* Cas9, called *tracr-L*, that directs Cas9 to transcriptionally repress its own promoter. This *tracrRNA*-mediated autoregulatory mechanism is a conserved example of intrinsic regulation within bacterial type II CRISPR-Cas systems, which can be exploited for the development of novel Cas9 tools.

Highlights

- A long-form *tracrRNA* (*tracr-L*) is a natural single-guide RNA for *S. pyogenes* Cas9
- *tracr-L* directs Cas9 to repress CRISPR-Cas expression and mitigate autoimmunity
- De-repression of *tracr-L* enhances immunity to bacteriophage infection by 3,000-fold
- *tracr-L* is conserved in many type II-A CRISPR-Cas systems

Workman et al., 2021, Cell *184*, 1–14
February 4, 2021
https://doi.org/10.1016/j.cell.2020.12.017

CellPress

Please cite this article in press as: Workman et al., A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression, Cell (2021), https://doi.org/10.1016/j.cell.2020.12.017

Cell

CellPress

Article

A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression

Rachael E. Workman,[1,5] Teja Pammi,[1,5] Binh T.K. Nguyen,[1] Leonardo W. Graeff,[1] Erika Smith,[2] Suzanne M. Sebald,[1] Marie J. Stoltzfus,[1] Chad W. Euler,[3,4] and Joshua W. Modell[1,6,*]

[1]Department of Molecular Biology & Genetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA
[2]Department of Biological Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA
[3]Department of Medical Laboratory Sciences, Hunter College, CUNY, New York, NY 10065, USA
[4]Department of Microbiology and Immunology, Weill Cornell Medicine, New York, NY 10065, USA
[5]These authors contributed equally
[6]Lead contact
*Correspondence: jmodell1@jhmi.edu
https://doi.org/10.1016/j.cell.2020.12.017

SUMMARY

CRISPR-Cas systems provide prokaryotes with acquired immunity against viruses and plasmids, but how these systems are regulated to prevent autoimmunity is poorly understood. Here, we show that in the *S. pyogenes* CRISPR-Cas system, a long-form transactivating CRISPR RNA *(tracr-L)* folds into a natural single guide that directs Cas9 to transcriptionally repress its own promoter (P_{cas}). Further, we demonstrate that P_{cas} serves as a critical regulatory node. De-repression causes a dramatic 3,000-fold increase in immunization rates against viruses; however, heightened immunity comes at the cost of increased autoimmune toxicity. Using bioinformatic analyses, we provide evidence that *tracrRNA*-mediated autoregulation is widespread in type II-A CRISPR-Cas systems. Collectively, we unveil a new paradigm for the intrinsic regulation of CRISPR-Cas systems by natural single guides, which may facilitate the frequent horizontal transfer of these systems into new hosts that have not yet evolved their own regulatory strategies.

INTRODUCTION

All immunity systems must distinguish ''self'' from ''non-self'' in order to avoid autoimmunity while still providing potent activity against diverse pathogen-associated antigens. The vertebrate adaptive immune system is a tightly regulated cellular network that undergoes spatiotemporal activation and expansion in response to novel or remembered antigens (Gasteiger and Rudensky, 2014). CRISPR-Cas systems provide bacteria and archaea with immunological memory of foreign nucleic acids, thereby protecting their prokaryotic hosts from invasive agents like viruses (Barrangou et al., 2007), plasmids (Marraffini and Sontheimer, 2008), and integrative and conjugative elements (Zhang et al., 2013). How CRISPR-Cas immune systems are regulated within their single-celled hosts is not well understood.

Immunological memory is encoded within CRISPR loci as short, roughly 30 bp DNA ''spacers,'' which are derived from segments of foreign nucleic acids (Barrangou et al., 2007). These spacers are located within CRISPR arrays between similarly sized repeating sequences or ''repeats'' (Mojica and Rodriguez-Valera, 2016). CRISPR immunity occurs in three stages. During ''adaptation,'' spacers are acquired from foreign agents and incorporated into the CRISPR array at the end proximal to a conserved ''leader'' sequence, causing the duplication of the leader-proximal repeat (Barrangou et al., 2007). During ''biogenesis,'' the CRISPR array is transcribed as one long precursor CRISPR RNA (*pre-crRNA*), which is cleaved within repeats to produce individual CRISPR RNAs (*crRNAs*), each containing a single spacer (Brouns et al., 2008; Carte et al., 2008; Deltcheva et al., 2011). Finally, during ''interference,'' *crRNAs* direct effector Cas proteins to matching foreign targets (Gasiunas et al., 2012; Jinek et al., 2012; Jore et al., 2011), where they perform distinct catalytic activities depending on the CRISPR-Cas family and sub-type (Plagens et al., 2015). The type II-A CRISPR-Cas system from *Streptococcus pyogenes* encodes four *cas* genes, expressed as a single operon (Figure 1A). While all Cas proteins are required for adaptation (Heler et al., 2015), Cas9 alone performs interference by introducing double-strand breaks into *crRNA*-specified DNA targets or ''protospacers''(Garneau et al., 2010; Sapranauskas et al., 2011; Jinek et al., 2012). Cas9 targeting requires a 5′-NGG-3′ protospacer-adjacent motif (PAM) (Deveau et al., 2008; Horvath et al., 2008), which is absent in the spacer-adjacent repeats within the CRISPR array, thus preventing the CRISPR system from recognizing its own spacers as foreign targets.

In addition to the *pre-crRNA*, the *S. pyogenes* CRISPR-Cas system harbors a second non-coding RNA, the *tracrRNA*, which functions during all three stages of CRISPR immunity (Deltcheva et al., 2011; Heler et al., 2015). While its role during adaptation is unclear, during biogenesis, *tracrRNA* base pairs

Please cite this article in press as: Workman et al., A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression, Cell (2021), https://doi.org/10.1016/j.cell.2020.12.017

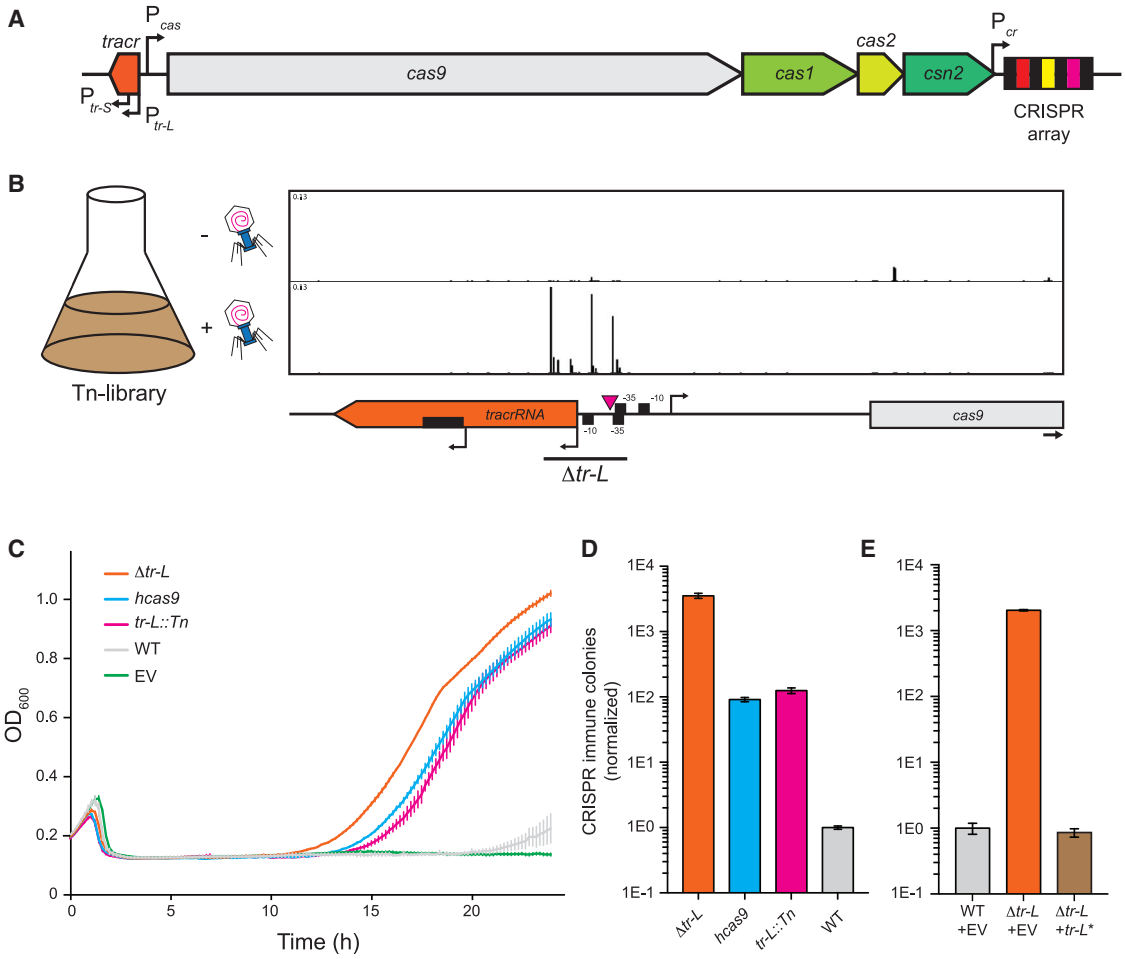


Figure 1. Tn-seq reveals a *tracrRNA* mutant with enhanced CRISPR immunity

(A) Schematic of the *S. pyogenes* type II-A CRISPR-Cas system. Arrows indicate promoters. Black boxes, CRISPR array repeats; colored boxes, spacers.

(B) Tn-seq NGS reads are shown as the fraction of total reads for each experiment. Data are representative of biological replicates. Top panel, – phage; bottom panel, + phage ϕNM4γ4; Δ*tr-L*, *tracr-L* deletion; magenta triangle, transposon insertion site in *tracr-L::Tn*; black rectangle within *tracrRNA*, segment complementary to *pre-crRNA*.

(C) Cells harboring a naive CRISPR-Cas system on a plasmid with the indicated mutations were infected with ϕNM4γ4 at multiplicity of infection (MOI) = 10 and cell densities (OD_{600}) were measured every 10 min. Plasmids here and throughout are medium-copy unless otherwise specified. *hcas9*, hyper-Cas9; *tracr-L::Tn*, transposon insertion in $P_{tracr\text{-}L}$; EV, empty vector. Error bars throughout represent mean ± SEM.

(D and E) Naive CRISPR-Cas systems with the indicated mutations were infected with ϕNM4γ4 at MOI = 25 in top agar, and surviving colonies with expanded CRISPR arrays were quantified by PCR. (E) An empty vector or a plasmid expressing long-form *tracrRNA* (*tr-L**) was introduced to the indicated strains from (D). See also Figures S1 and S7.

with repeat-derived sequences within the *pre-crRNA*, and the *tracrRNA:pre-crRNA* duplex is then cleaved by the host factor RNaseIII within each repeat, producing individual *crRNAs* and a processed *tracrRNA* (*tracr-P*) (Deltcheva et al., 2011). During interference, the mature targeting complex (Cas9:*tracr-P:crRNA*) facilitates target recognition and cleavage. In *S. pyogenes*, the *tracrRNA* is transcribed from two promoters producing a long (*tracr-L*) and short (*tracr-S*) form, both of which contain the RNaseIII processing site (Figures 1A and 1B) (Deltcheva et al., 2011). Because *tracr-S* alone can mediate *crRNA* biogenesis and interference *in vivo* and *in vitro*, it is unclear what role *tracr-L* plays or why two *tracrRNA* forms are produced.

Several studies have demonstrated that CRISPR-Cas expression can be regulated in response to extracellular and intracellular cues (Patterson et al., 2017), including phage infection (Agari et al., 2010; Young et al., 2012; Quax et al., 2013; Fusco et al., 2015; He et al., 2017), quorum sensing (Patterson et al., 2016; Høyland-Kroghsbo et al., 2017), membrane stress (Ratner et al., 2015; Perez-Rodriguez et al., 2011; Yosef et al., 2011; Majsec et al., 2016), metabolic status (Agari et al., 2010; Shinkai et al., 2007; Patterson et al., 2015), and surface association (Borges et al., 2020). In most of these cases, transcription factors residing outside the CRISPR-Cas locus interact with CRISPR-Cas promoters (Medina-Aparicio et al., 2011; Westra et al.,

Please cite this article in press as: Workman et al., A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression, Cell (2021), https://doi.org/10.1016/j.cell.2020.12.017

2010; Shinkai et al., 2007; Patterson et al., 2015). In contrast, the archaeal type I-A system can be intrinsically controlled by a dedicated Cas-encoded transcription factor (Liu et al., 2015; Lintner et al., 2011; Haft et al., 2005; Viswanathan et al., 2007; He et al., 2017). Given that CRISPR systems are frequently horizontally transferred in the wild (Takeuchi et al., 2012; Chakraborty et al., 2010; Godde and Bickerton, 2006), such intrinsic controllers would seem to be essential to prevent autoimmune toxicity (Vale et al., 2015; Vercoe et al., 2013; Jiang et al., 2013; Heler et al., 2017) in a new host. Paradoxically, however, few CRISPR loci—and no type II loci—encode dedicated transcription factors.

Here, we show that *tracr-L* folds into a natural single-guide RNA that directs Cas9 to transcriptionally silence CRISPR-Cas expression. Repression prevents autoimmunity but at the cost of reduced immunity to phage, demonstrating the critical importance of *tracr-L* regulation. Our work demonstrates a new mechanism of CRISPR-Cas autoregulation and suggests that transcriptional regulation by Cas9 is more prevalent than previously thought.

RESULTS

Tn-seq reveals a *tracrRNA* mutant with enhanced CRISPR immunity

We initially sought to identify bacterial genes involved in CRISPR-Cas immunity through an unbiased genetic screen. In our model system, *Staphylococcus aureus* cells that do not encode a native CRISPR-Cas system harbor the *S. pyogenes* type II CRISPR-Cas system on a medium or low-copy plasmid (Heler et al., 2015). The *S. aureus* host provides genetic tractability, an assortment of well-characterized phages, and a compatible cellular environment for the *S. pyogenes* CRISPR-Cas system. Within this heterologous system, we performed Tn-seq (van Opijnen et al., 2009) to identify transposon disruptions that enhanced CRISPR-Cas immunity. Cells harboring a "naive" CRISPR-Cas system, containing a single repeat and no spacers, were mutagenized with Tn917 (Tomich et al., 1979) and infected with bacteriophage ϕNM4γ4 (Heler et al., 2015). To our surprise, the transposon insertions that most strongly promoted survival were within the CRISPR locus itself, spanning the 5′ end and promoter of *tracr-L* (Figures 1B and S1A; Table S1).

To determine whether disruption of *tracr-L* enhances CRISPR immunity, we constructed a naive CRISPR-Cas plasmid with a 52 bp deletion spanning the promoter and 19 nucleotides at the 5′ end of *tracr-L*, generating Δ*tracr-L* (Figure 1B). Strikingly, Δ*tracr-L* significantly enhances immunity to ϕNM4γ4 by roughly 30-fold compared to a representative transposon insertion (*tracr-L::Tn*) or a previously identified hyper-adapting mutant (*hcas9*) (Heler et al., 2017) and by over 3,000-fold relative to the wild-type system (Figures 1C, 1D, and S1B). We recapitulated the Δ*tracr-L* immunity phenotype with an inactivating single nucleotide mutation in the *tracr-L* promoter (Figures S1C and S1D), and we confirmed that enhanced immunity was not due to a change in ϕNM4γ4 infectivity in Δ*tracr-L* cells (Figure S1E). Next-generation sequencing (NGS) of the CRISPR array in ϕNM4γ4-infected cells revealed that Δ*tracr-L* increased the abundance of cells with newly acquired spacers but not the distribution of those spacers within the ϕNM4γ4 genome (Figures S1F–S1H). Finally, we confirmed that Δ*tracr-L* enhanced immunity to another phage, ϕNM2γ1 (Figure S1I).

To test whether expression of *tracr-L* in *trans* can complement the Δ*tracr-L* phenotype, we cloned the *tracrRNA* locus onto a second plasmid and introduced two T > C mutations in nucleotides 71 and 76 of *tracr-L*, which constitute critical residues in the −10 promoter element of $P_{tracr\text{-}S}$ (Figure S1D), effectively eliminating *tracr-S* expression (Figure S1J). Expression of the double mutant, hereafter referred to as *tracr-L**, restored low levels of immunity to Δ*tracr-L* cells (Figures 1E and S1K), indicating that (1) Δ*tracr-L* is a loss-of-function allele and (2) *tracr-L* inhibits CRISPR immunity.

tracr-L inhibits multiple stages of CRISPR immunity

We next sought to determine whether the enhanced immunity of Δ*tracr-L* was due to increased rates of CRISPR adaptation, biogenesis, and/or interference. To test adaptation independently of subsequent steps, we performed two distinct spacer acquisition assays in the absence of phage: (1) we measured spacer acquisition of a defined dsDNA substrate that was electroporated into cells (Figure 2A) and (2) we overexpressed the adaptation genes and measured the abundance of spacers acquired from the chromosome and resident plasmid (Figure S2A). In each case, Δ*tracr-L* cells showed elevated adaptation rates compared to the wild type (Figures 2B and S2B), together indicating that *tracr-L* impairs spacer acquisition.

To determine whether *tracr-L* also inhibits CRISPR biogenesis and/or interference, we programmed wild-type or Δ*tracr-L* CRISPR loci with a spacer targeting ϕNM4γ4 and tested phage defense. On a medium-copy plasmid, the wild-type and Δ*tracr-L* systems both provide robust defense (Figures 2C, 2D, and S2C). However, on a low-copy plasmid, Δ*tracr-L* cells significantly outperformed the wild type (Figures 2C, 2E, and S2D). Collectively, these results demonstrate that *tracr-L* negatively regulates CRISPR-Cas adaptation as well as biogenesis and/or interference.

tracr-L causes widespread downregulation of all CRISPR-Cas RNAs

tracr-L could inhibit CRISPR immunity by affecting the levels or activity of one or more CRISPR-Cas components. Indeed, in the absence of *tracr-L*, the levels of *tracr-S*, *tracr-P,* and processed *crRNAs* were all dramatically enhanced (Figure 3A). In addition, *cas* gene mRNAs were induced by 30- to 50-fold (Figures 3B and S2E), and Cas9 protein levels showed a corresponding increase by western blot (Figure 3C). Induction of CRISPR-Cas components in Δ*tracr-L* cells was observed whether the CRISPR array contained a single repeat or the native six spacer CRISPR array from *S. pyogenes* strain SF370 (Figure S2F). Complementation of Δ*tracr-L* by expression of *tracr-L* from a second plasmid restored low levels of CRISPR-Cas components (Figures 3A–3C). In the *tracr-L::Tn* mutant, *tracr-L* and Cas9 levels were intermediate between the wild type and Δ*tracr-L* (Figure S2G), indicating that the transposon insertion disrupted but did not eliminate $P_{tracr\text{-}L}$ activity. Collectively, these

Please cite this article in press as: Workman et al., A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression, Cell (2021), https://doi.org/10.1016/j.cell.2020.12.017

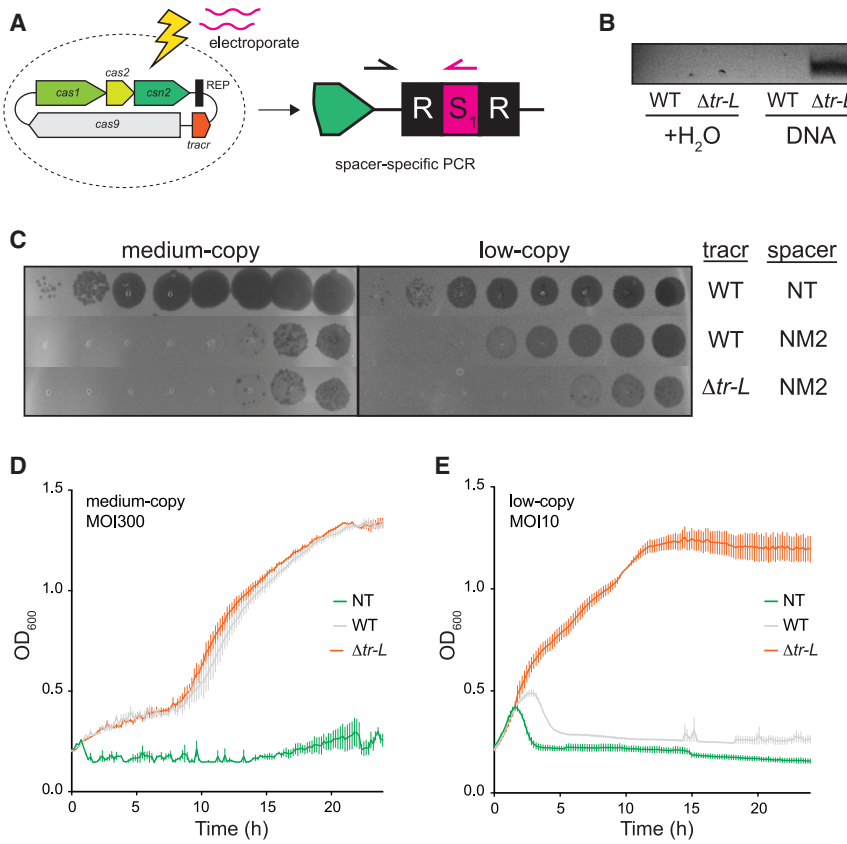


Figure 2. Long-form *tracrRNA* (*tracr-L*) inhibits both spacer acquisition and interference

(A) Schematic of spacer acquisition assay.

(B) Water or a 60 bp dsDNA amplicon with a single PAM was electroporated into cells with a naive WT or Δ*tr-L* CRISPR system, and adaptation was monitored with a spacer-specific PCR.

(C) 10-fold dilutions of ϕNM4γ4 were plaqued on top agar lawns containing cells with medium or low-copy CRISPR plasmids harboring the indicated spacers. NT, non-targeting spacer; NM2, spacer targeting ϕNM4γ4.

(D and E) The medium (D) and low-copy (E) strains from (C) were infected with ϕNM4γ4 in liquid culture at MOI = 300 (D) or MOI = 10 (E). Cell densities (OD_{600}) were measured every 10 min.

See also Figures S2 and S7.

results demonstrate that *tracr-L* inhibits the expression of all CRISPR-Cas coding and non-coding RNAs.

tracr-L directs Cas9 to transcriptionally repress the *cas* gene operon

We next asked whether *tracr-L* downregulates CRISPR-Cas RNA levels by transcriptionally repressing any of the system's four promoters (Figure 1A). Cells expressing *cas9* and wild-type *tracrRNA* but no CRISPR array or other *cas* genes were transformed with a second plasmid harboring a transcriptional GFP fusion to each CRISPR-Cas promoter. Compared to an empty vector, cells expressing *cas9* and *tracrRNA* inhibited P_{cas} activity by roughly 50-fold but had no effect on the other promoters (Figure 4A). Inhibition of P_{cas}-GFP was not observed in cells expressing *cas9* and *tracr-S* (Figure S3A) or *cas9* alone (Figure S3B) indicating that *tracr-L* is specifically required for P_{cas} repression.

To confirm that repression of P_{cas} requires Cas9, we combined a plasmid expressing *tracrRNA* and a *cas9* nonsense mutant with the P_{cas}-GFP reporter. Indeed, P_{cas} repression required Cas9 (Figure S3C) although *tracrRNAs* were undetectable in *cas9*-null mutants (Figure S3D), likely because they are protected from degradation by Cas9 protein. To confirm that Cas9 and *tracr-L* are sufficient to form a repressive complex on P_{cas}, we performed an EMSA with purified Cas9, *in vitro*-transcribed *tracr-L*, and P_{cas} target DNA. Indeed, Cas9:*tracr-L* bound P_{cas} with an affinity of ~1.6 nM, considerably tighter than the nonspecific affinity of Cas9 alone for P_{cas} or Cas9:*tracr-L* for a non-targeted control dsDNA (Figures 4B, S4A, and S4B).

Strikingly, inspection of P_{cas} revealed an 11-nucleotide match between the 5′ end of *tracr-L* and a region 2 bp downstream of the P_{cas} transcriptional start site (TSS) with a 5′-NGG-3′ PAM (Figures 4C, S3E, and S3F). Furthermore, while Cas9 cleaves DNA targets that match the 20 spacer-derived nucleotides with a given processed *crRNA* (Figure 4D), shorter match lengths of up to 16 base pairs allow stable target binding but prevent cleavage (Bikard et al., 2013; Sternberg et al., 2015; Ratner et al., 2019). To test whether this sequence match contributes to repression, we individually mutated bases within the putative target site on the P_{cas}-GFP reporter plasmid. Repression of P_{cas} was abolished in PAM mutants and greatly reduced in mutants of the 9 bases at the 3′ end of the 11 bp target site (Figures 4E and S3G). Together, our results are consistent with a model in which the 5′ end of *tracr-L* directs Cas9 to bind its own promoter and repress transcription of the *cas* gene operon.

tracr-L is a natural single-guide RNA

Key functional domains of *tracr-P* have been well defined, including two segments that base pair with the *crRNA* to form the upper and lower stems of the guide RNA duplex (Figure 4D) (Briner et al., 2014). By contrast, nothing is known about domains that are specific to *tracr-L*. To probe the determinants of *tracr-L* repression, we generated a tiled set of 5 nt mutations in the *tracr-L*-specific 5′ region as well as targeted mutations in the functional domains of *tracr-P*. As expected, mutations in the nexus and bulge, *tracrRNA* regions critical for Cas9 activity, alleviated repression of P_{cas} (Figure 4F) and destabilized *tracr-L* (Figure S3H), likely due to an inability to bind Cas9. Mutations in the 11 nt P_{cas}-matching sequence similarly blocked repression, consistent with our hypothesis that this region guides *tracr-L* to its target. Curiously, mutations in the *crRNA*-binding upper and

Please cite this article in press as: Workman et al., A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression, Cell (2021), https://doi.org/10.1016/j.cell.2020.12.017

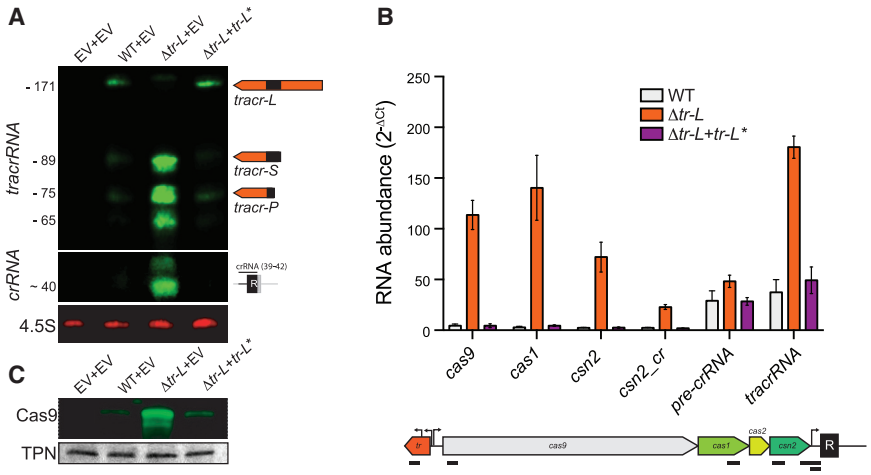


Figure 3. *tracr-L* downregulates all CRISPR-Cas system components

(A) Northern blot of cells harboring a plasmid expressing a naive WT or Δ*tr-L* CRISPR-Cas system and a second plasmid expressing the long-form *tracrRNA* (*tr-L**) or an EV. Membranes were probed with oligos matching the 3′ end of *tracrRNA* (top), the *crRNA* repeat (middle), or the 4.5S RNA (bottom, loading control).

(B) qRT-PCR of the indicated strains using primers within the coding region of *cas9*, *cas1*, or *csn2*, on either side of the crRNA promoter (csn2_cr), within the *pre-crRNA* leader sequence or within the processed *tracrRNA* (*tracr-P*). RNA abundance was expressed as $2^{-\Delta Ct}$ for each indicated primer pair relative to a *rho* control primer pair. Bottom panel shows the locations of the amplicons.

(C) Infrared western blot of the strains in (A) using an α-Cas9 antibody. Membranes were stained with Ponceau S, and the prominent band is shown as a loading control (total protein, TPN).

See also Figures S2, S3, and S7.

lower stems also inhibited *tracr-L* repression and destabilized the *tracrRNA* (Figures 4F and S3H). This was surprising because *tracr-L*-mediated repression of P_{cas} occurred in the absence of *crRNA* (Figure 4A), indicating that the upper and lower stems of *tracr-L* are not occupied by their cognate *crRNA* residues in the Cas9:*tracr-L* repressor complex. We noticed, however, that the 12 nucleotides just downstream of the P_{cas}-targeting region of *tracr-L* almost perfectly mimic those at the 5′ end of the *crRNA* repeat, suggesting that *tracr-L* could fold in on itself to reconstitute the upper and lower stems (Figure 4C). Consistent with this hypothesis, the *crRNA* mimicking residues were also essential for repression and stability (Figures 4F and S3H). These results suggested that *tracr-L* could form a natural single-guide RNA in which the P_{cas}-targeting region takes the place of the crRNA spacer and the upper and lower stems are formed intramolecularly.

In the putative single-guide *tracr-L* structure, residues 14–19 and 113–118 base pair to form the lower stem, and residues 22–25 and 105–108 form the upper stem. Consistent with this model, and the importance of these stems for Cas9 activity, 5 nt mutations in any of these four segments that would disrupt stem formation also disrupt *tracr-L* repression (Figure 4F). We generated a construct in which both strands of the putative lower stem are simultaneously mutated to their complement, which should restore base pairing potential. Indeed, high levels of repression were observed in the double mutant as well as a compensatory mutant in both upper and lower stems (Briner et al., 2014) (Figures S3I–S3K), supporting our model of intramolecular stem formation. Another feature of the natural single guide is the large 79 nt "upper stem extension," which takes the place of an 8 nt duplex or 5′-GAAA-3′ tetraloop in the *crRNA* and synthetic *sgRNA*, respectively (Figures S7A–S7D). We replaced the entire 79 nt segment with a 5′-GAAA-3′ tetraloop and found that repression of P_{cas} was maintained (Figure S3I), suggesting that this region is dispensable for *tracr-L* repression.

To confirm that the *tracr-L* target is specified by the 11 nt P_{cas}-matching region, we mutated this region to match 11 bp sites within artificial promoters unrelated to P_{cas}. Whether these targets were upstream (Figure S4C) or downstream (Figure 5A) from the transcriptional start site, promoter activity was significantly reduced in the presence of the reprogrammed constructs but not the native *tracr-L*. We next varied the *tracr-L* match length and found that increasing complementarity from 11 to 13 or 15 bp further reduced promoter activity (Figure 5B). Compared to nuclease-dead Cas9 (dCas9) programmed with a canonical *sgRNA* (containing a 20 bp spacer) specifying the same target site, repression with the 11 bp or 15 bp *tracr-L* was 4- or 1.7-fold weaker, respectively (Figure S4D). Notably, we could not readily transform cells harboring *tracr-L* constructs with GFP reporter plasmids containing target matches longer than 16 bp (Figure S4E) due to Cas9 cleavage above this threshold (Figure 5C), consistent with previous studies using other guides (Ratner et al., 2019; Bikard et al., 2013; Sternberg et al., 2015). Collectively, our results suggest that *tracr-L* evolved as a natural single guide that directs Cas9 to bind but not cleave P_{cas} and thereby maintain low levels of Cas operon expression.

tracr-L controls a system-wide switch that is responsive to crRNA expression

We next investigated whether the immunosuppressive effects of *tracr-L* were solely due to repression of P_{cas}. We constructed a plasmid encoding a naive CRISPR-Cas system with a PAM mutation (P_{cas}^{NGC}) that renders P_{cas} insensitive to *tracr-L* repression (Figure 4E). The P_{cas}^{NGC} strain displayed high levels of phage resistance comparable to the Δ*tracr-L* mutant (Figure 6A), indicating that transcriptional repression is the primary if not sole inhibitory function of *tracr-L*. As expected, despite the presence of *tracr-L*, P_{cas}^{NGC} cells overexpress Cas9 at Δ*tracr-L* levels (Figure S5A). Interestingly, this construct also phenocopies Δ*tracr-L* with high levels of *tracr-S*, *tracr-P*, and *crRNA*, none of which are transcribed from P_{cas} (Figure S5A). To ask whether *tracrRNAs* and *crRNAs* are stabilized by higher levels of Cas9 in the P_{cas}^{NGC} strain, we overexpressed Cas9 in cells harboring a wild-type

Please cite this article in press as: Workman et al., A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression, Cell (2021), https://doi.org/10.1016/j.cell.2020.12.017

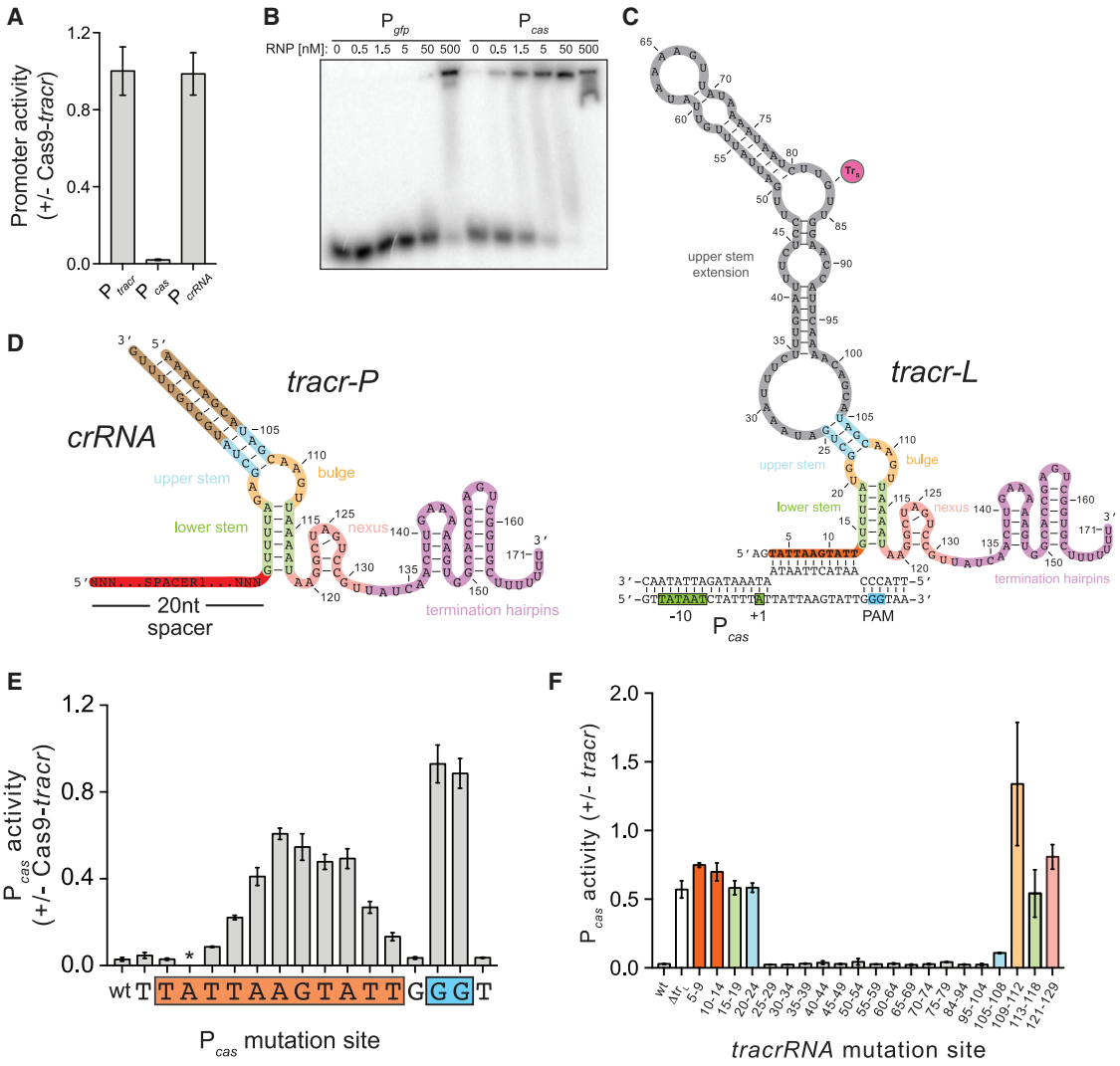


Figure 4. *tracr-L* is a natural *sgRNA* that directs Cas9 to repress its own promoter

(A) Promoter activity was measured (fluorescence/OD_{600}) in cells harboring a plasmid expressing GFP from the indicated promoters, and a second plasmid encoding *cas9* and the full *tracrRNA* locus or an empty vector. The ratios of promoter activity in *cas9+tracrRNA*: EV-containing cells are shown. P_{tracr} reports on the sum of $P_{tr\text{-}S}$ and $P_{tr\text{-}L}$.

(B) Electromobility shift assay (EMSA). 50 pM of a radiolabeled 55 bp dsDNA derived from P_{gfp} or P_{cas} was mixed with Cas9:*tracr-L* RNPs at the indicated concentrations.

(C) Putative structure of *tracr-L* in the repressor conformation. Orange region, P_{cas}-targeting; green, lower stem; yellow, bulge; blue, upper stem; gray, upper stem extension; pink, nexus; purple, termination hairpins; magenta circle, *tracr-S* start site. The upper stem extension structure is based on an RNA fold prediction (Gruber et al., 2008).

(D) Structure of the *tracr-P:crRNA* hybrid (adapted from Briner et al., 2014).

(E) Promoter activity was measured as in (A) with P_{cas}-GFP reporter constructs harboring single mutations to the complementary base at the indicated positions. The asterisk indicates that the mutation at the −10 site significantly reduced basal P_{cas} expression (Figure S3G), rendering the ± Cas9-tracr comparison inconclusive.

(F) Promoter activity was measured in cells harboring a plasmid expressing *cas9* from the constitutive promoter P_{sparc3} and P_{cas}-GFP and a second plasmid expressing the indicated *tracrRNA* mutants or an empty vector. The ratios of promoter activity in *tracrRNA*: EV-containing cells are shown. Sets of mutations were introduced in the *tracrRNA* plasmid at the indicated positions (relative to the TSS of *tracr-L*) to their complementary bases.

See also Figures S4 and S7.

CRISPR system and found that *tracr-S*, *tracr-P*, and *crRNAs* were indeed all upregulated (Figure S5B). These results suggest that *tracr-L* serves as a master regulator for the entire CRISPR-Cas system by (1) directly controlling the protein-coding *cas* operon through P_{cas} and (2) indirectly controlling *tracrRNA* and *crRNA* levels by regulating Cas9 abundance.

We next asked whether *tracr-L* is a specialized transcriptional repressor or whether, like *tracr-S*, it can also bind *crRNAs*. Cells

Please cite this article in press as: Workman et al., A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression, Cell (2021), https://doi.org/10.1016/j.cell.2020.12.017

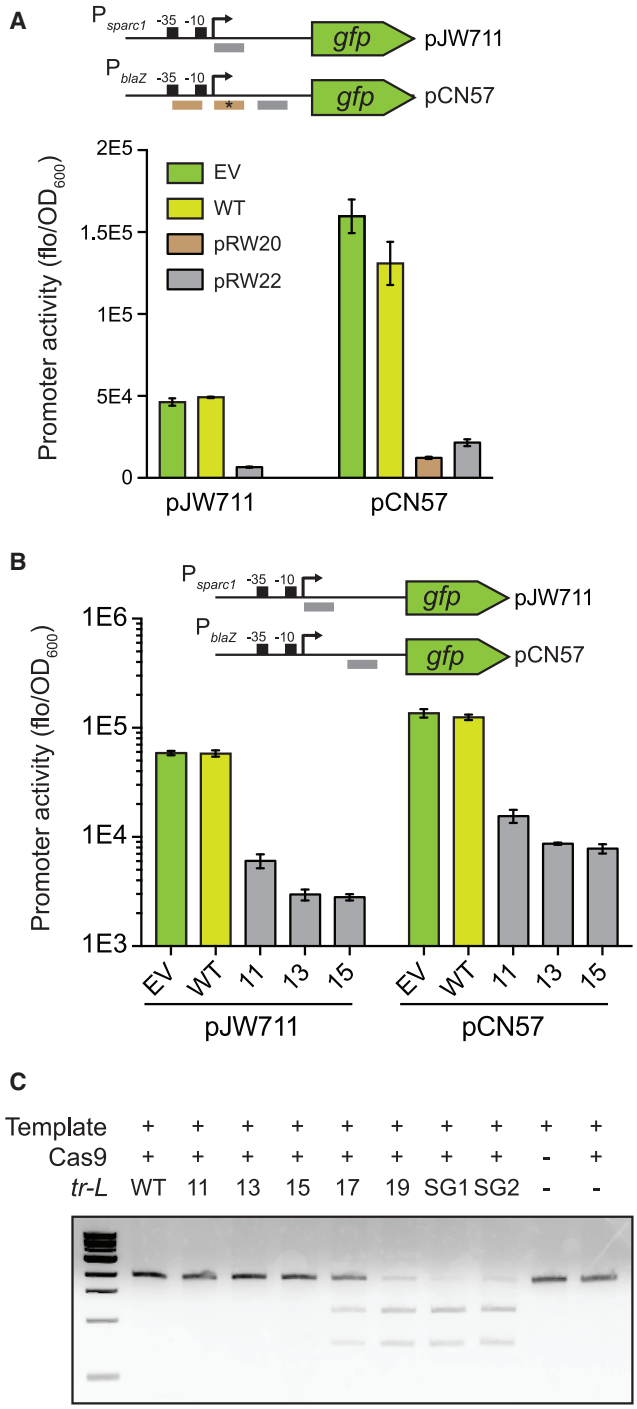


Figure 5. *tracr-L* can be reprogrammed to direct Cas9 repression or cleavage of novel targets

(A) Promoter activity was measured in cells harboring the GFP reporter plasmid pJW711 or pCN57 and a second plasmid expressing *cas9* and *tracr-L*. WT, wild-type *tracr-L*; pRW20 and pRW22, *tracr-L* reprogrammed to target the 11 bp regions indicated in brown or gray bars, respectively. The asterisk indicates a second pRW20 binding site within pCN57 containing a single mismatch at the −9 seed position relative to the PAM.

(B) Promoter activity was measured as in (A) but with *tracr-L* variants with target matches of increasing lengths (11, 13, 15 bp).

(C) *In vitro* cleavage assay. Cas9 and *tracr-L* were incubated with a 1 kb amplicon from pCN57 and cleavage products were separated on an agarose gel. WT, *tracr-L*; 11-19, *tracr-L* variants with matches of the indicated length to the pCN57 amplicon; SG1-2, replicates of an *sgRNA* with a 20 nt pCN57 targeting sequence.

See also Figures S4 and S7.

expressing *tracr-L* as the only *tracrRNA* were capable of performing interference (Figures 6B, S5C, and S5D), albeit less well than those expressing *tracr-S* (Figures S5E and S5F). Nonetheless, these results indicate that *tracr-L* can hybridize with *crRNAs* in order to interfere with viral targets specified by the *crRNA* spacer.

Once bound, *crRNAs* occupy the upper and lower stems of *tracr-L* preventing formation of the natural single guide (Figure S7E). We therefore wondered whether *crRNA* expression interferes with *tracr-L*-mediated repression of P_{cas}. To explore this possibility, we measured Cas9 levels in cells harboring a CRISPR-Cas system with no CRISPR array, a single repeat, or two repeats surrounding one of the six spacers from the natural *S. pyogenes* SF370 CRISPR array. We observed low Cas9 levels in cells without a CRISPR array, and the presence of a single repeat did not enhance Cas9 expression, indicating that a naive *pre-crRNA* did not appreciably interfere with *tracr-L* repression; however, Cas9 levels increased in the presence of any single spacer and in the 6-spacer *S. pyogenes* CRISPR array (Figure S5G; Figure S2F). *crRNAs* could also influence P_{cas} expression by hybridizing with *tracr-S* and sequestering Cas9 away from *tracr-L*. Consistent with this possibility, Cas9 levels in a naive CRISPR-Cas system increased after the introduction of a non-targeting *sgRNA* on a second plasmid (Figure S5H). These results suggest the presence of a regulatory circuit in which *crRNAs* provide feedback through *tracr-L* to affect P_{cas} expression.

CRISPR-Cas repression by *tracr-L* inhibits autoimmunity

Our results indicate that, in wild-type cells, *tracr-L* maintains the CRISPR-Cas system in a lowly active state. We wondered whether constitutive expression of CRISPR-Cas components could cause autoimmunity, stemming from adaptation and subsequent interference against the bacterial chromosome or resident plasmids. Indeed, in the absence of phage, we observed more self-targeting spacers in Δ*tracr-L* cells relative to the wild type and *hcas9*, with many derived from the chromosome terminus, a known hotspot for spacer acquisition in type II CRISPR-Cas systems (Modell et al., 2017) (Figures S6A–S6C). To gauge whether constitutive CRISPR-Cas expression causes a growth defect, we competed wild-type and Δ*tracr-L* cells and found that after 1 and 2 days, the number of Δ*tracr-L* cells dropped to 33% and 3% of their original number, respectively (Figure 6C). Together, these results suggest that *tracr-L* could repress the CRISPR system to avoid autoimmunity while allowing enough expression for some level of viral surveillance (Figures 2B–2E).

tracr-L regulates CRISPR-Cas expression in *S. pyogenes*

We next explored whether *tracr-L* represses P_{cas} in its native *S. pyogenes* host through two parallel approaches: (1) we deleted *tracr-L* in the *S. pyogenes* chromosome, and (2) we depleted

Please cite this article in press as: Workman et al., A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression, Cell (2021), https://doi.org/10.1016/j.cell.2020.12.017

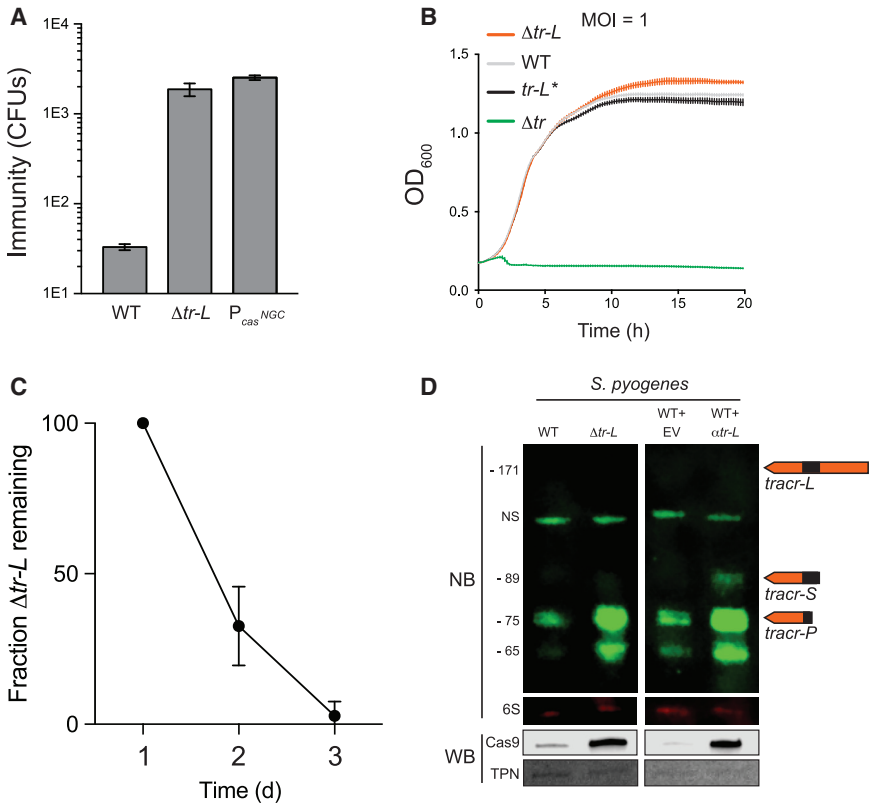


Figure 6. *tracr-L* is a master switch that controls autoimmunity and responds to crRNA spacer identity

(A) *S. aureus* cells harboring a plasmid expressing a naive CRISPR-Cas system were infected with ϕNM4γ4 at MOI = 25 in top agar, and surviving colonies were quantified. P_{cas}^{NGC}, PAM mutation in the *tracr-L* target site within P_{cas}.

(B) Interference assay with ϕNM4γ4 at MOI = 1 with cells harboring a plasmid expressing the ϕNM4γ4-targeting spacer NM2, the indicated *tracrRNA* alleles, and *cas9* from P_{cas}^{NGC} to normalize *cas9* expression across experiments. Δ*tr*, full deletion of the *tracrRNA* locus. The *tr-L** strain harbored a second plasmid expressing additional *tracr-L** from P_{sparc2} in order to ensure levels of *tracr-L* expression comparable to *tracr-S* in the Δ*tr-L* strain. All other strains harbored a second empty vector. IPTG was added at 1 mM to all strains to induce expression from P_{sparc2}.

(C) Cells harboring a plasmid expressing a naive wild-type or Δ*tr-L* CRISPR system were co-cultured in a competition assay. Cells were grown to stationary phase and diluted into logarithmic phase twice per day, and each morning cultures were plated and genotyped by PCR of the *tracrRNA* locus.

(D) Northern and western blots on overnight cultures of *S. pyogenes* cells. α*tr-L*, plasmid expressing an antisense RNA targeting *tracr-L*. A nonspecific band (NS) and *6S* RNA were used as loading controls.

See also Figures S5, S6, and S7.

tracr-L by expressing an antisense RNA. In both cases, *tracr-S*, *tracr-P*, and Cas9 protein were significantly elevated in the absence of *tracr-L* (Figures 6D and S6D) indicating that *tracr-L* controls CRISPR-Cas expression in *S. pyogenes*. Next, we asked whether P_{cas} could be de-repressed to induce CRISPR-Cas expression when needed, for instance, during a phage infection. We infected *S. pyogenes* cells with phage A25 but found no reproducible changes in Cas9 protein levels at any multiplicity of infection (MOI) tested (Figure S6E). Alternatively, we hypothesized that cells could induce CRISPR-Cas expression in response to environmental changes that forebode a phage attack. Growth at 25°C mildly enhanced both Cas9 and *tracr-L* expression relative to 37°C (Figure S6F), similar to what we observed for *tracr-L*-independent mechanisms of Cas9 induction (Figures S5A and S5B). Conversely, increases in Cas9 levels were accompanied by decreases in *tracr-L* during growth into late stationary phase (Figure S6F), suggesting that, in this case, Cas9 induction could be driven by de-repression of P_{cas} by *tracr-L*.

tracrRNA regulation is dynamic on evolutionary timescales

We next investigated whether *tracr-L*-mediated repression of P_{cas} is found in other type II-A systems. We examined 80 CRISPR-Cas loci from *Streptococci*, *Listeria*, and *Lactobacilli* (Faure et al., 2019) and found that 34 (43%) contain a putative *tracr-L* (*tracr-L*$^+$) (Figure 7A; Table S2). To determine whether these *tracr-L*$^+$ loci encode bona fide repressors, we cloned *S. agalactiae* strain A909 *tracrRNA* and *cas9* onto an *S. aureus* plasmid and measured *cas9* mRNA levels by qPCR. Deletion of the putative *tracr-L* or introduction of a P_{cas} PAM mutation enhanced *cas9* expression by 20- to 30-fold, suggesting that *S. agalactiae tracr-L* is a conserved P_{cas} repressor (Figure 7B).

tracr-L$^+$ genomes are distributed intermittently throughout the II-A tree, suggesting frequent loss or gain events. Curiously, while the lower stems and PAMs showed near perfect conservation among *tracr-L*$^+$ loci, the identity of the matching sequence varied and its length ranged from 11 to 15 bp, indicating covariation over evolutionary timescales between the *tracr-L* and P_{cas} targeting determinants (Figure 7A, Table S5). Within *tracr-L*$^-$ loci, the intergenic space between *tracr-S* and *cas9* was on average roughly 200 bp shorter than in *tracr-L*$^+$ loci (Figure 7C) owing to deletions that removed the targeting and/or targeted sites (Table S5). In some cases, we observed *tracr-L*$^-$ loci with intergenic lengths comparable to *tracr-L*$^+$ that contained one or more mismatches within the seed but retained perfect PAMs and lower stems (Table S5). Our data suggest that similar seed mutants can retain intermediate levels of repression (Figure 4E), while deletions of targeting determinants likely result in complete de-repression. Collectively, our results demonstrate that *tracr-L*-mediated regulation of P_{cas} is common in type II-A CRISPR-Cas systems.

DISCUSSION

tracr-L buffers cells against CRISPR-Cas autoimmunity

Since the earliest characterization of the *S. pyogenes* type II-A CRISPR-Cas system (Deltcheva et al., 2011), it has remained a

Please cite this article in press as: Workman et al., A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression, Cell (2021), https://doi.org/10.1016/j.cell.2020.12.017

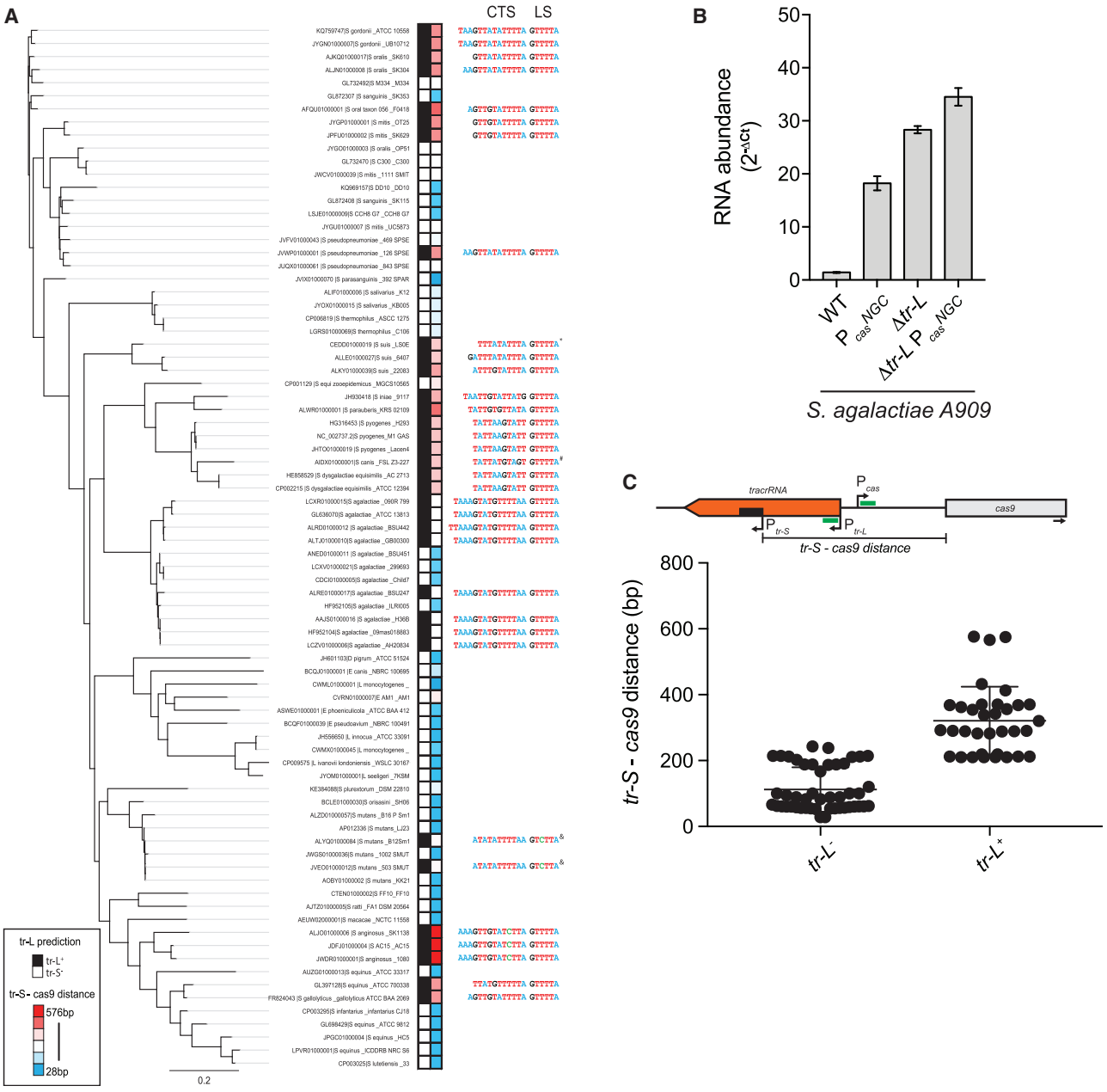


Figure 7. *tracr-L* is a conserved and dynamic feature of type II-A CRISPR-Cas systems

(A) A phylogenetic tree was constructed using Cas9 protein sequences from representative type II-A "branch-1" genomes (Faure et al., 2019). Each CRISPR-Cas locus was queried for the presence of *tracr-L*-mediated regulation of P_{cas} (black boxes, *tracr-L*$^+$; white boxes, *tracr-L*$^-$) by looking for an 11 bp candidate targeting sequence (CTS), a 5′-GTTTTA-3′ lower stem (LS), and target site 5′-NGG-3′ PAM. The intergenic distance between the translation start site of *cas9* and the transcriptional start site of the short-form *tracrRNA* (*tracr-S* - cas9) was measured and plotted as a heatmap. Red boxes, longer distances; blue boxes, shorter distances. The CTS and lower stem LS are listed for each *tracr-L*$^+$ locus, with nucleotides colored by identity. *, 10 bp CTS; #, 5′-NAG-3′ PAM; &, 5′-GTCTTA-3′ LS; scale bar, average amino acid substitutions per site. Full metadata are available in Table S2.

(B) qRT-PCR of *S. aureus* cells expressing the indicated *S. agalactiae* A909 *cas9* and *tracrRNA* alleles from a plasmid. The *tracrRNA* and P_{cas} loci in A909 are identical to those in strain "09mas018883" in (A).

(C) Schematic showing the intergenic distance between *tracr-S* and Cas9 (black bar) with putative targeting determinants (green bars). Below, intergenic distances were plotted for *tracr-L*$^+$ and *tracr-L*$^-$ loci.

See also Figure S7.

Please cite this article in press as: Workman et al., A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression, Cell (2021), https://doi.org/10.1016/j.cell.2020.12.017

mystery why two *tracrRNA* isoforms are expressed given that the short form is sufficient to mediate *pre-crRNA* processing and Cas9 targeting and cleavage (Jinek et al., 2012). Here, we report that long-form *tracrRNAs* form natural single guides that direct Cas9 to bind and transcriptionally repress the *cas* operon promoter (Figure 4C; Figure S7). The *S. pyogenes* Cas9:*tracr-L* complex binds P_{cas} with a K_d of ~1.6 nM (Figure S4B), an affinity that is slightly lower than published values for dCas9:*sgRNA* (Sternberg et al., 2014; Jinek et al., 2012; Mekler et al., 2017), perhaps owing to the shorter targeting sequence (11 versus 20 bp).

The wild-type CRISPR-Cas system is thus maintained in a low activity state that is sufficient to surveil and destroy remembered threats (Figures 2C–2E) but adapts poorly to unrecognized phages (Figures 1C and 2B). De-repression of the system results in a roughly 50-fold increase in CRISPR-Cas components (Figure 3) and a 3,000-fold increase in immunization rates against an unrecognized phage (Figures 1C–1E). As with many immune systems, this hyperactivity comes at a price; cells with de-repressed systems acquire many more autoimmune spacers (Figures S1F and S1G), and the substrate preference is altered with a higher percentage of new spacers coming from the bacterial chromosome relative to plasmid and phage sources (Figures S6A–S6C). We propose that autoimmune spacer acquisition contributes to the growth defect in cells lacking *tracr-L* (Figure 6C), which could be further exacerbated by the metabolic burden of constitutive CRISPR-Cas expression and off-target cleavage by Cas9 at high concentrations. *S. pyogenes* therefore joins a small but growing list of prokaryotes that employ an off or dimmer switch to buffer against the deleterious effects of CRISPR-Cas overexpression (Patterson et al., 2017).

Are CRISPR-Cas systems dynamically regulated?

Can physiological stimuli trigger CRISPR-Cas induction as needed by the bacterial host? One logical candidate stimulus is the phage infection itself. The archaeal *S. solfataricus* type I-A CRISPR-Cas locus encodes the only other known CRISPR-Cas autoregulator, a Cas-encoded transcription factor Csa3b that binds and represses P_{cas} (Liu et al., 2015; Lintner et al., 2011; Haft et al., 2005). When the Cas effector in this system is directed to a non-lytic viral target, it brings Csa3b with it resulting in de-repression of P_{cas} (He et al., 2017). We did not observe a similar induction of the *S. pyogenes* type II-A Cas operon following an A25 phage infection at a range of MOIs (Figure S6E). While other phages may provoke different outcomes, it is unclear whether a transcriptional response can keep pace with the rapid life cycle of a lytic bacteriophage compared to a non-lytic archaeal virus.

Another possibility is that the CRISPR-Cas system is induced by conditions or environments that might presage a phage threat. We observed a modest increase in Cas9 expression during late stationary phase (Figure S6F), when dense bacterial populations could provide ample targets for phage predation. In this case, we also observed a decrease in *tracr-L* suggesting that Cas9 induction could be driven by de-repression of P_{cas}. In contrast, during growth at 25°C, Cas9 expression increased in tandem with *tracr-L*, similar to what we observed for modes of Cas9 induction that do not involve loss of *tracr-L* (Figures S5A and S5B). Increasing levels of Cas9:*tracr-L* could provide negative feedback to P_{cas}, thereby maintaining Cas9 expression within an acceptable window. We note that many bacterial ribonucleases similarly autoregulate their expression, often by cleaving their own mRNA products (Bechhofer and Deutscher, 2019). *tracr-L* hybridizes with the P_{cas} template strand so regulation of the mRNA is unlikely. Regardless, a critical role of *tracr-L* may be to maintain Cas9 homeostasis and thereby buffer cells against stochastic or regulated changes in Cas9 expression. Further studies will be required to identify the full set of signals and mechanisms that regulate Cas9, either via or independently of *tracr-L*.

Genetic control of *tracr-L*

De-repression of *tracr-L* could also occur through genetic changes within the targeting elements. Single mutations in *tracr-L* (targeting seed, lower and upper stems) or P_{cas} (target seed, PAM), encompassing roughly 30 nucleotides in total, partially, or fully relieve repression (Figures 4E and 4F). Given bacterial mutation rates (Chevallereau et al., 2019), as many as one in a million cells are expected to be CRISPR-Cas overproducers, or 10–100 per colony. In a bacterial population, wild-type cells could be spared from autoimmunity, while rare CRISPR-Cas overexpressing mutants could serve as sentinels in case of a phage attack. Seed mutations in the *tracr-L* targeting sequence could function as evolutionary stepping stones, where a second compensatory mutation in P_{cas}, or vice versa, could restore repression as needed. This might provide an explanation for the covariation we observe between the targeting and targeted sites throughout the type II-A *tracr-L* tree.

tracr-L autoregulators are located throughout the type II-A CRISPR-Cas tree in 43% of representative members (Figure 7A). Its intermittent absence suggests that individual strains can use or lose *tracr-L* according to their unique needs. In strains without *tracr-L*, constitutive CRISPR-Cas expression could enable survival in phage-rich environments. Alternatively, other non-canonical guide RNAs could regulate P_{cas} in the absence of *tracr-L*. In the *F. novicida* type II-B CRISPR system, tracrRNA hybridizes with a "*scaRNA*" that, like *tracr-L,* reconstitutes dsRNA stem structures normally formed by the *crRNA:tracrRNA* duplex (Sampson et al., 2013; Ratner et al., 2019). The *scaRNA* also utilizes an 11 bp targeting sequence; however, it targets the promoter of an immunostimulatory lipoprotein in the bacterial genome. Whether *scaRNAs* can target promoters within the CRISPR-Cas system and conversely whether *tracr-L* can target promoters elsewhere in the bacterial chromosome are open questions.

The composition of the CRISPR array can also affect Cas9 expression (Figure S5G). We hypothesize that *crRNAs* affect the repressive potential of *tracr-L* by (1) binding to *tracr-L* and preventing formation of the single-guide (Figure S7E) and/or (2) binding to *tracr-S* and sequestering Cas9 away from *tracr-L* (Figure S5H). *crRNAs* with spacers that bind the *tracrRNAs* better or worse could therefore differentially regulate P_{cas}. Notably, the CRISPR array is transcribed from its own promoter providing another entry point for system-wide regulation. Furthermore, larger CRISPR arrays with more repeats provide more binding sites for the sequestration and/or processing of *tracr-L*. Longer arrays should therefore enhance CRISPR-Cas expression,

supported by the higher baseline levels of Cas expression we observe in the native *S. pyogenes* 6-spacer array relative to a single repeat (Figure S2F). Bacterial populations acquire a wide variety of spacers during a phage infection (Heler et al., 2015; Paez-Espino et al., 2013), and the most recently acquired spacers are the most highly expressed (McGinn and Marraffini, 2016; Deltcheva et al., 2011). In addition to cleavage efficiency (Xu et al., 2015) and target location (Modell et al., 2017; Strotskaya et al., 2017), our results unveil CRISPR-Cas expression as a new criterion for spacer selection and maintenance.

Future perspectives

The emerging literature on CRISPR-Cas regulation in bacteria focuses on type I systems, which are controlled by transcription factors or chaperones encoded outside the CRISPR-Cas locus (Patterson et al., 2017). Our results provide the first example of an intrinsic CRISPR-Cas regulator within a bacterial host. We believe that intrinsic regulators like *tracr-L* could facilitate the rampant horizontal transfer of CRISPR-Cas systems in nature by dampening autoimmunity upon delivery and allowing the new bacterial host to tune CRISPR-Cas9 expression to meet its needs, through transient induction or mutation of the control elements.

Our results will also inform new generations of Cas9 tools. *tracr-L* can be reprogrammed to target promoters of interest and multiplexed with canonical *sgRNAs* in order to simultaneously cleave one site while transcriptionally repressing another. Alternatively, *tracr-L* autoregulation of P_{cas} could mitigate off-target effects by buffering Cas9 expression levels. Furthermore, a better understanding of how *tracr-L* is naturally regulated will facilitate the development of inducible and reversible single guides. As the list of sequenced microbial genomes grows, so does the catalog of novel Cas9 orthologs, each with the possibility of new PAMs and unique cleavage activities. Our results suggest that many of these genomes might be worth a closer look for the blueprints of natural single guides.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Microbes
  - Phages
- METHOD DETAILS
  - Plasmid construction
  - Gibson assembly
  - Oligo cloning
  - Allelic exchange in Streptococcus pyogenes
  - Genomic DNA extraction from Streptococcus pyogenes
  - Transformation in Streptococcus pyogenes
  - Tn-seq screen
  - S. aureus miniprep protocol
  - Electroporation adaptation assay
  - Overexpression adaptation assay
  - Enrichment PCR assay
  - Liquid growth interference assay
  - Top agar interference assay
  - Promoter activity fluorescence assays
  - Liquid growth immunity assay
  - Top agar immunity Assay
  - PCR conditions
  - RNA extraction
  - Infrared Northern (irNorthern)
  - Western blot
  - RNA-seq library preparation and sequencing
  - qPCR
  - Transformation assay
  - *In vitro* transcription (IVT)
  - Electrophoretic mobility shift assay (EMSA)
  - *In vitro* cleavage assay
  - Competition assay
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Tn-seq screen analysis
  - Promoter activity fluorescence assays
  - Electrophoretic mobility shift assay (EMSA) dissociation constant extraction
  - RNA sequencing analyses
  - Spacer sequencing analyses
  - Evolutionary analyses

Supplemental information

Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2020.12.017.

ACKNOWLEDGMENTS

We would like to thank Michael Laub, Luciano Marraffini, Geraldine Seydoux, and Samuel Sternberg for their comments on the manuscript. We thank Nicholas Keith for lending bioinformatic expertise. We thank Karole D'Orazio and members of Rachel Green's and Scott Bailey's labs for sharing technical assistance and reagents for *in vitro* experiments. We thank Jeremy Nathans' and Carol Greider's lab for sharing reagents and equipment. We thank David Mohr and the GRCF High Throughput Sequencing Center for assistance with NGS experiments. Funding was provided by a startup package from the Johns Hopkins School of Medicine.

AUTHOR CONTRIBUTIONS

R.E.W., T.P., B.T.K.N., L.W.G., and J.W.M. designed and executed the research studies. E.S., S.M.S., and M.J.S. assisted with plasmid construction. C.W.E. constructed the *S. pyogenes* Δ*tr-L* strain and assisted in *Streptococcal* experimental design. R.E.W. and J.W.M. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 9, 2020
Revised: September 24, 2020
Accepted: December 9, 2020
Published: January 8, 2021

SUPPORTING CITATIONS

The following references appear in the supplemental information: Charpentier et al. (2004); Devos (1986); Goldberg et al. (2014); Horinouchi and Weisblum (1982a); Horinouchi and Weisblum (1982b); Khan and Novick (1983); Youngman et al., (1983).

REFERENCES

Agari, Y., Sakamoto, K., Tamakoshi, M., Oshima, T., Kuramitsu, S., and Shinkai, A. (2010). Transcription profile of Thermus thermophilus CRISPR systems after phage infection. J. Mol. Biol. *395*, 270–281.

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.

Bechhofer, D.H., and Deutscher, M.P. (2019). Bacterial ribonucleases and their roles in RNA metabolism. Crit. Rev. Biochem. Mol. Biol. *54*, 242–300.

Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., and Marraffini, L.A. (2013). Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res. *41*, 7429–7437.

Borges, A.L., Castro, B., Govindarajan, S., Solvik, T., Escalante, V., and Bondy-Denomy, J. (2020). Bacterial alginate regulators and phage homologs repress CRISPR-Cas immunity. Nat. Microbiol. *5*, 679–687.

Briner, A.E., Donohoue, P.D., Gomaa, A.A., Selle, K., Slorach, E.M., Nye, C.H., Haurwitz, R.E., Beisel, C.L., May, A.P., and Barrangou, R. (2014). Guide RNA functional modules direct Cas9 activity and orthogonality. Mol. Cell *56*, 333–339.

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.

Carte, J., Wang, R., Li, H., Terns, R.M., and Terns, M.P. (2008). Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. Genes Dev. *22*, 3489–3496.

Chakraborty, S., Snijders, A.P., Chakravorty, R., Ahmed, M., Tarek, A.M., and Hossain, M.A. (2010). Comparative network clustering of direct repeats (DRs) and cas genes confirms the possibility of the horizontal transfer of CRISPR locus among bacteria. Mol. Phylogenet. Evol. *56*, 878–887.

Charpentier, E., Anton, A.I., Barry, P., Alfonso, B., Fang, Y., and Novick, R.P. (2004). Novel cassette-based shuttle vector system for gram-positive bacteria. Appl. Environ. Microbiol. *70*, 6076–6085.

Chevallereau, A., Meaden, S., van Houte, S., Westra, E.R., and Rollie, C. (2019). The effect of bacterial mutation rate on the evolution of CRISPR-Cas adaptive immunity. Philos. Trans. R. Soc. Lond. B Biol. Sci. *374*, 20180094.

Culviner, P.H., Guegler, C.K., and Laub, M.T. (2020). A Simple, Cost-Effective, and Robust Method for rRNA Depletion in RNA-Sequencing Studies. MBio *11*. Published online April 21, 2020. https://doi.org/10.1128/mBio.00010-20.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature *471*, 602–607.

Deveau, H., Barrangou, R., Garneau, J.E., Labonté, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P., and Moineau, S. (2008). Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. J. Bacteriol. *190*, 1390–1400.

Devos, W.M. (1986). Gene Cloning in Lactic Streptococci. Neth. Milk Dairy J. *40*, 141–154.

Edgar, R.C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics *5*, 113.

Euler, C.W., Juncosa, B., Ryan, P.A., Deutsch, D.R., McShan, W.M., and Fischetti, V.A. (2016). Targeted Curing of All Lysogenic Bacteriophage from Streptococcus pyogenes Using a Novel Counter-selection Technique. PLoS ONE *11*, e0146408.

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.

Ferretti, J.J., McShan, W.M., Ajdic, D., Savic, D.J., Savic, G., Lyon, K., Primeaux, C., Sezate, S., Suvorov, A.N., Kenton, S., et al. (2001). Complete genome sequence of an M1 strain of Streptococcus pyogenes. Proc. Natl. Acad. Sci. USA *98*, 4658–4663.

Fields, C., Sheng, P., Miller, B., Wei, T., and Xie, M. (2019). Northern Blot with IR Fluorescent Probes: Strategies for Probe Preparation. Biol. Protoc. *9*, e3219.

Fusco, S., Liguori, R., Limauro, D., Bartolucci, S., She, Q., and Contursi, P. (2015). Transcriptome analysis of Sulfolobus solfataricus infected with two related fuselloviruses reveals novel insights into the regulation of CRISPR-Cas system. Biochimie *118*, 322–332.

Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature *468*, 67–71.

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. USA *109*, E2579–E2586.

Gasteiger, G., and Rudensky, A.Y. (2014). Interactions between innate and adaptive lymphocytes. Nat. Rev. Immunol. *14*, 631–639.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

Godde, J.S., and Bickerton, A. (2006). The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. J. Mol. Evol. *62*, 718–729.

Goldberg, G.W., Jiang, W., Bikard, D., and Marraffini, L.A. (2014). Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. Nature *514*, 633–637.

Gruber, A.R., Lorenz, R., Bernhart, S.H., Neuböck, R., and Hofacker, I.L. (2008). The Vienna RNA websuite. Nucleic Acids Res. *36*, W70–W74.

Haft, D.H., Selengut, J., Mongodin, E.F., and Nelson, K.E. (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. PLoS Comput. Biol. *1*, e60.

He, F., Vestergaard, G., Peng, W., She, Q., and Peng, X. (2017). CRISPR-Cas type I-A Cascade complex couples viral infection surveillance to host transcriptional regulation in the dependence of Csa3b. Nucleic Acids Res. *45*, 1902–1913.

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.

Heler, R., Wright, A.V., Vucelja, M., Bikard, D., Doudna, J.A., and Marraffini, L.A. (2017). Mutations in Cas9 Enhance the Rate of Acquisition of Viral Spacer Sequences during the CRISPR-Cas Immune Response. Mol. Cell *65*, 168–175.

Horinouchi, S., and Weisblum, B. (1982a). Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. J. Bacteriol. *150*, 815–825.

Horinouchi, S., and Weisblum, B. (1982b). Nucleotide sequence and functional map of pE194, a plasmid that specifies inducible resistance to macrolide, lincosamide, and streptogramin type B antibodies. J. Bacteriol. *150*, 804–814.

Horvath, P., Romero, D.A., Coûté-Monvoisin, A.C., Richards, M., Deveau, H., Moineau, S., Boyaval, P., Fremaux, C., and Barrangou, R. (2008). Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. J. Bacteriol. *190*, 1401–1412.

Høyland-Kroghsbo, N.M., Paczkowski, J., Mukherjee, S., Broniewski, J., Westra, E., Bondy-Denomy, J., and Bassler, B.L. (2017). Quorum sensing controls the Pseudomonas aeruginosa CRISPR-Cas adaptive immune system. Proc. Natl. Acad. Sci. USA *114*, 131–135.

Jiang, W., Maniv, I., Arain, F., Wang, Y., Levin, B.R., and Marraffini, L.A. (2013). Dealing with the evolutionary downside of CRISPR immunity: bacteria and beneficial plasmids. PLoS Genet. *9*, e1003844.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.

Jore, M.M., Lundgren, M., van Duijn, E., Bultema, J.B., Westra, E.R., Waghmare, S.P., Wiedenheft, B., Pul, U., Wurm, R., Wagner, R., et al. (2011). Structural basis for CRISPR RNA-guided DNA recognition by Cascade. Nat. Struct. Mol. Biol. *18*, 529–536.

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. *30*, 3059–3066.

Khan, S.A., and Novick, R.P. (1983). Complete nucleotide sequence of pT181, a tetracycline-resistance plasmid from Staphylococcus aureus. Plasmid *10*, 251–259.

Kreiswirth, B.N., Löfdahl, S., Betley, M.J., O'Reilly, M., Schlievert, P.M., Bergdoll, M.S., and Novick, R.P. (1983). The toxic shock syndrome exotoxin structural gene is not detectably transmitted by a prophage. Nature *305*, 709–712.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

Lintner, N.G., Frankel, K.A., Tsutakawa, S.E., Alsbury, D.L., Copié, V., Young, M.J., Tainer, J.A., and Lawrence, C.M. (2011). The structure of the CRISPR-associated protein Csa3 provides insight into the regulation of the CRISPR/Cas system. J. Mol. Biol. *405*, 939–955.

Liu, T., Li, Y., Wang, X., Ye, Q., Li, H., Liang, Y., She, Q., and Peng, N. (2015). Transcriptional regulator-mediated activation of adaptation genes triggers CRISPR de novo spacer acquisition. Nucleic Acids Res. *43*, 1044–1055.

Majsec, K., Bolt, E.L., and Ivančić-Baće, I. (2016). Cas3 is a limiting factor for CRISPR-Cas immunity in Escherichia coli cells lacking H-NS. BMC Microbiol. *16*, 28.

Marraffini, L.A., and Sontheimer, E.J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science *322*, 1843–1845.

Maxted, W.R. (1952). Enhancement of streptococcal bacteriophage lysis by hyaluronidase. Nature *170*, 1020–1021.

McGinn, J., and Marraffini, L.A. (2016). CRISPR-Cas Systems Optimize Their Immune Response by Specifying the Site of Spacer Integration. Mol. Cell *64*, 616–623.

Medina-Aparicio, L., Rebollar-Flores, J.E., Gallego-Hernández, A.L., Vázquez, A., Olvera, L., Gutiérrez-Ríos, R.M., Calva, E., and Hernández-Lucas, I. (2011). The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and leucine-responsive regulatory protein in Salmonella enterica serovar Typhi. J. Bacteriol. *193*, 2396–2407.

Mekler, V., Minakhin, L., and Severinov, K. (2017). Mechanism of duplex DNA destabilization by RNA-guided Cas9 nuclease during target interrogation. Proc. Natl. Acad. Sci. USA *114*, 5443–5448.

Modell, J.W., Jiang, W., and Marraffini, L.A. (2017). CRISPR-Cas systems exploit viral DNA injection to establish and maintain adaptive immunity. Nature *544*, 101–104.

Mojica, F.J., and Rodriguez-Valera, F. (2016). The discovery of CRISPR in archaea and bacteria. FEBS J. *283*, 3162–3169.

Paez-Espino, D., Morovic, W., Sun, C.L., Thomas, B.C., Ueda, K., Stahl, B., Barrangou, R., and Banfield, J.F. (2013). Strong bias in the bacterial CRISPR elements that confer immunity to phage. Nat. Commun. *4*, 1430.

Patterson, A.G., Chang, J.T., Taylor, C., and Fineran, P.C. (2015). Regulation of the Type I-F CRISPR-Cas system by CRP-cAMP and GalM controls spacer acquisition and interference. Nucleic Acids Res. *43*, 6038–6048.

Patterson, A.G., Jackson, S.A., Taylor, C., Evans, G.B., Salmond, G.P.C., Przybilski, R., Staals, R.H.J., and Fineran, P.C. (2016). Quorum Sensing Controls Adaptive Immunity through the Regulation of Multiple CRISPR-Cas Systems. Mol. Cell *64*, 1102–1108.

Patterson, A.G., Yevstigneyeva, M.S., and Fineran, P.C. (2017). Regulation of CRISPR-Cas adaptive immune systems. Curr. Opin. Microbiol. *37*, 1–7.

Perez-Rodriguez, R., Haitjema, C., Huang, Q., Nam, K.H., Bernardis, S., Ke, A., and DeLisa, M.P. (2011). Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in Escherichia coli. Mol. Microbiol. *79*, 584–599.

Plagens, A., Richter, H., Charpentier, E., and Randau, L. (2015). DNA and RNA interference mechanisms by CRISPR-Cas surveillance complexes. FEMS Microbiol. Rev. *39*, 442–463.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. PLoS ONE *5*, e9490.

Quax, T.E., Voet, M., Sismeiro, O., Dillies, M.A., Jagla, B., Coppée, J.Y., Sezonov, G., Forterre, P., van der Oost, J., Lavigne, R., and Prangishvili, D. (2013). Massive activation of archaeal defense genes during viral infection. J. Virol. *87*, 8419–8428.

Ratner, H.K., Sampson, T.R., and Weiss, D.S. (2015). I can see CRISPR now, even when phage are gone: a view on alternative CRISPR-Cas functions from the prokaryotic envelope. Curr. Opin. Infect. Dis. *28*, 267–274.

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.

Rizk, G., Lavenier, D., and Chikhi, R. (2013). DSK: k-mer counting with very low memory usage. Bioinformatics *29*, 652–653.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.

Sampson, T.R., Saroj, S.D., Llewellyn, A.C., Tzeng, Y.L., and Weiss, D.S. (2013). A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. Nature *497*, 254–257.

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. Nucleic Acids Res. *39*, 9275–9282.

Shinkai, A., Kira, S., Nakagawa, N., Kashihara, A., Kuramitsu, S., and Yokoyama, S. (2007). Transcription activation mediated by a cyclic AMP receptor protein from Thermus thermophilus HB8. J. Bacteriol. *189*, 3891–3901.

Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature *507*, 62–67.

Sternberg, S.H., LaFrance, B., Kaplan, M., and Doudna, J.A. (2015). Conformational control of DNA target cleavage by CRISPR-Cas9. Nature *527*, 110–113.

Strotskaya, A., Savitskaya, E., Metlitskaya, A., Morozova, N., Datsenko, K.A., Semenova, E., and Severinov, K. (2017). The action of Escherichia coli CRISPR-Cas system on lytic bacteriophages with different lifestyles and development strategies. Nucleic Acids Res. *45*, 1946–1957.

Takeuchi, N., Wolf, Y.I., Makarova, K.S., and Koonin, E.V. (2012). Nature and intensity of selection pressure on CRISPR-associated genes. J. Bacteriol. *194*, 1216–1225.

Tomich, P.K., An, F.Y., and Clewell, D.B. (1979). A transposon (Tn917) in Streptococcus faecalis that exhibits enhanced transposition during induction of drug resistance. Cold Spring Harb. Symp. Quant. Biol. *43*, 1217–1221.

Vale, P.F., Lafforgue, G., Gatchitch, F., Gardan, R., Moineau, S., and Gandon, S. (2015). Costs of CRISPR-Cas-mediated resistance in Streptococcus thermophilus. Proc. Biol. Sci. *282*, 20151270.

van Opijnen, T., Bodi, K.L., and Camilli, A. (2009). Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. Nat. Methods *6*, 767–772.

Vercoe, R.B., Chang, J.T., Dy, R.L., Taylor, C., Gristwood, T., Clulow, J.S., Richter, C., Przybilski, R., Pitman, A.R., and Fineran, P.C. (2013). Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. PLoS Genet. *9*, e1003454.

Viswanathan, P., Murphy, K., Julien, B., Garza, A.G., and Kroos, L. (2007). Regulation of dev, an operon that includes genes essential for Myxococcus xanthus development and CRISPR-associated genes and repeats. J. Bacteriol. *189*, 3738–3750.

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.

Woodcock, D.M., Crowther, P.J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S.S., Michael, M.Z., and Graham, M.W. (1989). Quantitative evaluation of Escherichia coli host strains for tolerance to cytosine methylation in plasmid and phage recombinants. Nucleic Acids Res. *17*, 3469–3478.

Xu, H., Xiao, T., Chen, C.H., Li, W., Meyer, C.A., Wu, Q., Wu, D., Cong, L., Zhang, F., Liu, J.S., et al. (2015). Sequence determinants of improved CRISPR sgRNA design. Genome Res. *25*, 1147–1157.

Yosef, I., Goren, M.G., Kiro, R., Edgar, R., and Qimron, U. (2011). High-temperature protein G is essential for activity of the Escherichia coli clustered regularly interspaced short palindromic repeats (CRISPR)/Cas system. Proc. Natl. Acad. Sci. USA *108*, 20136–20141.

Young, J.C., Dill, B.D., Pan, C., Hettich, R.L., Banfield, J.F., Shah, M., Fremaux, C., Horvath, P., Barrangou, R., and Verberkmoes, N.C. (2012). Phage-induced expression of CRISPR-associated proteins is revealed by shotgun proteomics in Streptococcus thermophilus. PLoS ONE *7*, e38077.

Youngman, P.J., Perkins, J.B., and Losick, R. (1983). Genetic transposition and insertional mutagenesis in Bacillus subtilis with Streptococcus faecalis transposon Tn917. Proc. Natl. Acad. Sci. USA *80*, 2305–2309.

Zhang, Q., Rho, M., Tang, H., Doak, T.G., and Ye, Y. (2013). CRISPR-Cas systems target a diverse collection of invasive mobile genetic elements in human microbiomes. Genome Biol. *14*, R40.

Please cite this article in press as: Workman et al., A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression, Cell (2021), https://doi.org/10.1016/j.cell.2020.12.017

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Cas9 Mouse IgG1 mAb	Cell Signaling	Cat #7A9-3A3; RRID: AB_2750916
Goat anti-Mouse IgG Secondary Antibody (infrared)	LICOR	Cat # 925-32210; RRID:AB_2687825
Goat anti-Mouse IgG1 Secondary Antibody (HRP-conjugated)	Pierce	Cat # PA174421; RRID:AB_10988195
Bacterial and virus strains		
Staphylococcus aureus strain RN4220 (WT)	PMID: 6226876 (Kreiswirth et al., 1983)	Refseq: NC_048107
Streptococcus pyogenes strain SF370	PMID: 11296296 (Ferretti et al., 2001)	Refseq: NC_002737
Streptococcus pyogenes strain SF370Δtracr-L	This paper	N/A
Streptococcus pyogenes strain C13	Gift of Andrew Varble	N/A
Streptococcus agalactiae strain A909	ATCC	BAA-1138
Escherichia coli DH5a	PMID: 2657660 (Woodcock et al., 1989)	N/A
Staphylococcus phage ϕNM4γ4	PMID: 25707807 (Heler et al., 2015)	N/A
Staphylococcus phage ϕNM2γ1	Gift of Marraffini Lab	N/A
Streptococcus phage A25	PMID: 13013300 (Maxted, 1952)	Refseq: NC_028697
Chemicals, peptides, and recombinant proteins		
Lysostaphin	Ambi Products LLC	Cat # LSPN-50
Tetracycline	Sigma	Cat # T7660-25G
Chloramphenicol	Sigma	Cat # C0378-25G
Erythromycin	Sigma	Cat # E5389-5G
Spectinomycin	Sigma	Cat # S4014-25G
IPTG	Sigma	Cat # I6758-5G
Anhydrotetracycline (aTc)	Sigma	Cat #37919-100MG-R
Q5 High-Fidelity DNA polymerase	NEB	Cat # M0491S
Phusion High-Fidelity DNA polymerase	ThermoFisher	Cat # F530L
T4 Polynucleotide Kinase	NEB	Cat # M0201S
T4 DNA Ligase	NEB	Cat # M0202S
NEBNext dsDNA Fragmentase	NEB	Cat # M0348
Proteinase K	NEB	Cat # P8107S
Cas9 nuclease	NEB	Cat # M0386S
PlyC	Gift from Euler lab	N/A
Novex sample buffer	ThermoFisher	Cat # LC6876
Critical commercial assays		
QIAGEN Spin MiniPrep kit	QIAGEN	Cat # 27104
Direct-zol RNA MiniPrep kit	Zymo	Cat # R2071
Wizard Genomic DNA purification kit	Promega	Cat # A1120
QIAquick PCR Purification kit	QIAGEN	Cat # 28104
NEBNext DNA Library Reagent Set for Illumina	NEB	Cat # E6000S
QIAquick Gel Extraction kit	QIAGEN	Cat #28704
NEBNext Multiplex Oligos for Illumina kit	NEB	Cat # E7335S
ProbeQuant G50 spin columns	GE Healthcare	Cat # GE28-9034-08
Trilink small RNA library prep kit	Trilink	Cat # L-3206-24

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Spacer sequencing and RNA sequencing files (Illumina NGS)	Sequence read archives (SRA)	Bioproject ID PRJNA679244
Software and algorithms		
Python v3.8	Python Software Foundation	https://www.python.org/downloads/release/python-380/
Burrows-Wheeler Aligner	(Li and Durbin, 2009)	http://bio-bwa.sourceforge.net/
Integrative Genomics Viewer	Broad Institute	http://software.broadinstitute.org/software/igv/
MAFFT10	Geneious plugin	https://www.geneious.com/plugins/mafft-plugin/
MUSCLE11	Geneious plugin	https://www.geneious.com/features/sequence-alignment/
FastTree	Geneious plugin	https://www.geneious.com/plugins/fasttree-plugin/
Prime 2020.1.1	Geneious	https://www.geneious.com/
Prism 7.4	GraphPad	https://www.graphpad.com/scientific-software/prism/
Custom code used to analyze RNA sequencing and spacer sequencing Illumina data	This study	https://github.com/modelllab/Workman_etal_2020

RESOURCE AVAILABILITY

Lead contact

For further information and requests for resources and reagents, please contact Joshua Modell (jmodell1@jhmi.edu), Department of Molecular Biology and Genetics, Johns Hopkins School of Medicine, Baltimore, MD.

Materials availability

All materials generated for this study are available upon request and without restrictions from the Lead Contact, Joshua Modell.

Data and code availability

- Data and code are publicly available
- Code for all analyses performed publicly available at Github: https://github.com/modelllab/Workman_etal_2020
- Raw Illumina fastq data for RNA sequencing experiments and spacer acquisition sequencing available on SRA: Bioproject ID PRJNA679244

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Microbes

Staphylococcus aureus cells were grown at 37°C, unless otherwise indicated, in Bacto Brain-Heart infusion (BHI) broth with shaking at 220 RPM. During outgrowths from stationary phase preceding phage treatments, BHI was supplemented with calcium chloride at 5 mM to allow phage adsorption and with 1 mM IPTG to allow expression from P_{sparc2} when necessary. Antibiotics were used at the following concentrations for strain construction and plasmid maintenance in *S. aureus*: tetracycline, 5 μg/mL; chloramphenicol, 10 μg/mL; erythromycin, 10 μg/mL; spectinomycin, 250 μg/mL. *Streptococcus pyogenes* cells were grown at 37°C, unless otherwise indicated, in Bacto Brain-Heart infusion (BHI) broth without shaking. During outgrowths from stationary phase preceding phage treatments, BHI was supplemented with calcium chloride at 5 mM to allow phage adsorption. Antibiotics were used at the following concentrations for strain construction and plasmid maintenance in *S. pyogenes:* chloramphenicol, 3 μg/mL.

Phages

Staphylococcus aureus phages (ϕNM4γ4, ϕNM2γ1) were amplified on RN4220 and stored in BHI at 4°C. *Streptococcus pyogenes* phage A25 was amplified on strain C13, a derivative of SF370 prophage-cured strain CEM1ΔΦ (Euler et al., 2016) and a kind gift of Andrew Varble, which has a spontaneous deletion of spacers 2-5 in its CRISPR array, including deletion of the natural spacer targeting A25, and stored in BHI at 4°C.

Please cite this article in press as: Workman et al., A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression, Cell (2021), https://doi.org/10.1016/j.cell.2020.12.017

METHOD DETAILS

Plasmid construction

See Table S4 for strains, plasmids, cloning notes, and oligos used in this study.

Gibson assembly

Gibson assemblies were performed as described (Gibson et al., 2009). Briefly, 100 ng of the largest dsDNA fragment to be assembled was combined with equimolar volumes of the smaller fragment(s) and brought to 5 μL total in dH2O on ice. Samples were added to 15 μL of Gibson Assembly master mix, mixed by pipetting and incubated at 50°C for 1 hour. Samples were drop dialyzed in dH2O for 30 minutes to 1 hour, and 5 μL were electroporated into 50 μL electrocompetent RN4220 *S. aureus* cells.

Oligo cloning

To create a repeat-spacer-repeat CRISPR array with a defined spacer, we used a restriction digest-based cloning approach. Parent plasmids contain two CRISPR repeats flanking a 30bp sequence housing two BsaI restriction sites. 400-800 ng of the plasmid was mixed with the BsaI-HFv2 restriction enzyme (NEB, R3733S) in a 10 μL reaction volume (1 μL BsaI-HFv2 enzyme, 1 μL CutSmart buffer, plasmid + nuclease-free water to 10 μL) and incubated at 37°C for ~8 hours. Two IDT oligos, a "top" strand with sequence 5′-AAAC-(30bp spacer)-G-3′, and a "bottom" strand with sequence 5′-AAAAC-(30bp spacer reverse complement)-3′ were phosphorylated with PNK (NEB, M0201S) in a 50 μL reaction volume (1.5 μL 100 μM top oligo, 1.5 μL 100 μM bottom oligo, 41 μL nuclease free water, 5 μL T4 PNK 5x reaction buffer, 1 μL T4 PNK) at 37°C for 30 minutes to 1 hour. After phosphorylation, oligos were annealed by adding 2.5 μL of 1M NaCl to the 50 μL reaction and incubating for 5 minutes at 98°C, then allowing the reaction to cool to room temperature (1-2 hours). The phosphorylated, annealed oligos were diluted 1:10 in nuclease-free water and ligated to the digested plasmid in a 20 μL reaction (10 μL digested plasmid, 6 μL water, 1 μL 1:10 diluted oligos, 2 μL T4 ligase buffer, 1 μL T4 DNA ligase enzyme (NEB, M0202S)) at room temperature overnight. Reactions were drop dialyzed for 1 hour in dH2O and 5 μL were transformed into electrocompetent RN4220 *S. aureus* cells.

Allelic exchange in Streptococcus pyogenes

To create the Δ*tracr-L* mutation in the native *Streptococcus pyogenes* host chromosome, we performed allelic exchange. The allelic exchange vector pJW854 was constructed by amplification of the pWV01-based shuttle vector pFW13 (a generous gift of Andrew Varble) using primers oJW2637/oJW2638, and insertion of left and right homology arms flanking the chloramphenicol resistance gene *cat194* [Cm^r] by Gibson assembly. The left homology arm was amplified from purified *Streptococcus pyogenes* SF370 genomic DNA (DNeasy Blood and Tissue kit, QIAGEN, modified protocol described below) using primers oJW2641/oJW2642, *cat194* was amplified from pAV259 (a generous gift of Andrew Varble) using primers oJW2640/oJW2639, and the right homology arm containing the 52-bp Δ*tracr-L* deletion mutation was amplified from pGG32_Dtr-L using primers oJW2643/oJW2644. pJW854 was first transformed into *E. coli* Dh5α, then isolated from 150 mL overnight culture using the GeneJET Plasmid Maxiprep kit (Thermo Fisher) according to manufacturer's instructions. The purified plasmid was concentrated using ethanol precipitation and transformed into *Streptococcus pyogenes* strain SF370 using the electroporation protocol described below and plated onto BHI plates with chloramphenicol. After transformation, colonies were restreaked on chloramphenicol plates to reduce background and confirm antibiotic resistance. Restruck colonies were picked, lysed in 100 ul 1X PBS and PlyC (1 μg/ml final concentration), then checked for successful double recombination using PCR, primers oJW2923/oJW2924 and oJW2922/oJW2927.

Genomic DNA extraction from Streptococcus pyogenes

1 mL of *S. pyogenes* overnight culture was centrifuged for 1 minute at 6000 RPM, resuspended in 1X PBS and PlyC (1 μg/ml final concentration) and incubated for 10 minutes at room temperature. After lysis, the genomic DNA extraction was performed using the DNeasy Blood and Tissue kit according to the manufacturer's protocol (QIAGEN).

Transformation in Streptococcus pyogenes

S. pyogenes cells were made electrocompetent by diluting an overnight culture 1:20 in 150 mL and growing cells to an OD of ~0.3, followed by centrifugation at 4000 x g at 4°C for 20 minutes. The supernatant was decanted and the cell pellet was washed with 150 mL cold 10% glycerol and pelleted at 4000 x g at 4°C for 15 minutes. This wash step was repeated once in 150 mL, then twice more in 30 mL in a 50 mL Falcon tube. The washed cell pellet was resuspended in 150 μl 10% glycerol, then separated into 50 μl aliquots. 5 μl of 500 ng/μl plasmid DNA was mixed with the electrocompetent cells, then pipetted into a pre-chilled 0.1cm cuvette (Biorad, 165-2089). The cuvette was dried with a Kimwipe then pulsed at 2.5kV/cm, 200 ohms, and 25 μF. 950 μl of pre-warmed BHI was immediately added to the cuvette and incubated at 37°C without shaking for 3 hours. After incubation, the cells were transferred into an Eppendorf tube, centrifuged at 6000 rpm for 1 min, then resuspended in 300 μl BHI and struck out onto BHI plates with the appropriate antibiotic.

Please cite this article in press as: Workman et al., A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression, Cell (2021), https://doi.org/10.1016/j.cell.2020.12.017

Tn-seq screen

To construct a transposon library in cells lacking a CRISPR system, *Staphylococcus aureus* RN4220 cells harboring pTV1, a temperature-sensitive plasmid containing transposon Tn917 from *Streptococcus faecalis* (Tomich et al., 1979), were grown overnight at the permissive temperature (30°C) in BHI supplemented with chloramphenicol. The next morning, 500 μL cells were washed 1x in plain BHI and then diluted into 500 mL BHI prewarmed to the restrictive temperature (42°C) and supplemented with erythromycin, without chloramphenicol. Cells were grown at 42°C with shaking for ~8 hours and diluted again 500 μL into 500 mL BHI/erythromycin. After an overnight growth at 42°C, cells were plated onto BHI agar plates supplemented with erythromycin and incubated overnight at 42°C. The next day, 10,000 colonies were scraped with 3 mL BHI into a Falcon tube and resuspended by pipetting and vortexing. DMSO was added to 10% and library aliquots were stored at −80°C. A second transposon library was constructed in cells harboring the CRISPR-Cas system on plasmid pJW92 as above, with tetracycline added during each step for plasmid maintenance.

Transposon library aliquots were thawed and diluted to OD = 0.05 in 20 mL (- phage experiments) or 500 mL (+ phage experiments) BHI supplemented with calcium chloride. After ~2 hours of growth at 37°C, culture ODs were roughly 0.4-0.5, and ϕNM4γ4 was added to the + phage experiments in duplicate (CRISPR- library) or triplicate (CRISPR+ library) at MOI = 1. After 24 hours of growth, genomic DNA was prepared using the Wizard Genomic DNA purification kit (Promega) and culture aliquots were frozen at −80°C for storage in 10% DMSO. As above, tetracycline was added in all steps for cells harboring pGG32_tetK.

To prepare NGS libraries, genomic DNA was digested with NEBNext dsDNA Fragmentase (M0348) for 9 minutes, and the reaction was stopped by addition of 0.1 M EDTA (final concentration). The fragmented DNA, centered at 700bp, was then purified using the QIAquick PCR Purification kit and end repaired using the NEBNext DNA Library Reagent Set for Illumina (E6000S). After another QIAquick PCR purification, NEBNext adapters were ligated onto the fragments and DNA in the 0.2 – 1 kb range was purified from an agarose gel with the QIAquick Gel Extraction kit. Next, for each sample, a PCR was performed with a unique oligo from the NEBNext Multiplex Oligos for Illumina kit (E7335S, Index Primers Set 1) and a universal, transposon-specific primer (oJW451). Amplicons in the 0.2 – 0.5 kb range were gel purified as before and sequenced on an Illumina HiSeq using the sequencing primer oJW436. NGS analysis details are given in the Quantification and Statistical Analysis section below.

S. aureus miniprep protocol

1-1.5 mL of an overnight culture, unless otherwise indicated, was pelleted and resuspended in 250 μL Buffer P1. 10-20 μL Lysostaphin (Ambi Products LLC, LSPN-50, 100 μg/mL final) was added and the cells were incubated without shaking at 37°C for ~20 minutes. Following lysis, plasmids were isolated using the QIAGEN Spin Miniprep kit according to the manufacturer's protocol, beginning with addition of P2. DNA was eluted from each column in 30 μL RNase-free water.

Electroporation adaptation assay

Electrocompetent *S. aureus* cells were made by washing overnight cultures twice in full-volumes of dH2O and once in a half-volume of dH2O before resuspending in 1/100 the original volume in 10% glycerol. A 60 bp amplicon from phage ϕNM4γ4 was generated using primers oJW1744 and oJW1745. This amplicon contains a single candidate protospacer with a 5′-NGG-3′ PAM, which Cas9 recognizes during spacer selection (Heler et al., 2015). The amplicon was dialyzed and 50 ng were added to 50 μL of electrocompetent cells in an Eppendorf tube, mixed by pipetting and left at room temperature for 10 minutes. Cells were transferred to a 0.2 cm electroporation cuvette (Bio-Rad, 1652086) and electroporated at 1.8 kV using an Eporator (Eppendorf). Following electroporation, 1 mL BHI broth was added and the contents of the cuvette were mixed by pipetting and then transferred to an Eppendorf tube and grown for 1 hour with shaking at 220 RPM at 37°C. Plasmids were miniprepped, and spacer acquisition was monitored by a spacer-specific PCR reaction with primers oJW1833 and oJW1834. PCR products were visualized on a 2% agarose gel run for 25 minutes at 130 V.

Overexpression adaptation assay

In this assay, overexpression of the adaptation genes *cas1*, *cas2* and *csn2* enables spacer acquisition from the host genome and resident plasmids. Overnight cultures of *S. aureus* were diluted to OD = 0.04 in BHI broth supplemented with antibiotics and grown for 1 hour with shaking at 37°C. Cultures were then supplemented with anhydrotetracycline (ATc) at 0 μg/mL or 0.5 μg/mL (final concentrations). Cultures were incubated with shaking at 37°C for 2 hours and then miniprepped with a lysis step including 15 μL of Lysostaphin incubated at 37°C for 5 minutes. An enrichment PCR was performed as described below and PCR products were visualized on a 2% agarose gel run for 27 minutes at 130 V.

Enrichment PCR assay

For overexpression adaptation and spacer sequencing analysis we enriched for adapted CRISPR loci using an established enrichment PCR protocol (Modell et al., 2017). Briefly, CRISPR plasmids were harvested from *S. aureus* cells with a modified QIAprep Spin Miniprep Kit protocol: bacterial cell pellets were resuspended in 250 μl P1 buffer supplemented with Lysostaphin (100 ug/ml final concentration) and incubated at 37 °C for 20 minutes followed by the standard QIAprep protocol. We used 100ng of plasmid as input for the enrichment PCR of the CRISPR locus using Phusion DNA Polymerase (Thermo) with the following primer mix: three parts oJW8 or oJW419 and one part each oJW3, oJW4 and oJW5. For spacer NGS sequencing, variants of the primer oJW8 with 3–8-bp bar codes at the 5′ end were used to distinguish experiments from each other during multiplexed high-throughput sequencing.

Please cite this article in press as: Workman et al., A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression, Cell (2021), https://doi.org/10.1016/j.cell.2020.12.017

PCRs with conventional primers were performed similarly using primers oJW1131 and oL401 with 5′ bar codes for multiplexing. Amplicons were purified with Ampure XP (Beckman Coulter) and subjected to Illumina high-throughput sequencing with the Miseq platform.

Liquid growth interference assay

Overnight cultures of *S. aureus* were diluted to OD = 0.025 in BHI broth supplemented with calcium chloride and antibiotics and grown for 1 hour and 15 minutes shaking at 37°C. Cultures were normalized to the OD of the lowest culture and phage ϕNM4γ4 was added to the appropriate MOI. After inverting to mix, 150 μL of each culture were added to a flat-bottom 96-well plate (Grenier 655180), and the plate was incubated at 37°C with shaking in a TECAN Infinite F Nano+ with OD600 measurements recorded every 10 minutes for 24 hours.

Top agar interference assay

100 μL of *S. aureus* overnight cultures were added to a 50 mL Falcon tube. 6 mL of BHI top agar (0.75% agar) supplemented with calcium chloride (5mM final concentration) was added to each tube. After swirling to mix, the cells and top agar were poured onto a BHI 1.5% agar plate and rocked gently to create a bacterial lawn. The plate was incubated for 15 - 30 minutes at RT. 3.5 μL of 8 10-fold serial dilutions of phage ϕNM4γ4 in BHI broth were spotted on top of the bacterial lawn using a multichannel pipette. After a 30 minute incubation at room temperature, the plates were moved to a 37°C incubator overnight.

Promoter activity fluorescence assays

200 μL of overnight cultures were spun down in a 1.5 mL Eppendorf tube at 6,000 rpm for one minute. Cell pellets were resuspended in 1 mL of 1X PBS and 150 μL were transferred into a clear, flat-bottomed 96-well plate (Grenier 655180). Measurements for absorbance (at 600 nm) and fluorescence (excitation wavelength = 485 nm; emission wavelength = 535 nm) were recorded (details on calculations in Quantification and Statistical Analysis).

Liquid growth immunity assay

Overnight cultures of *Staphylococcus aureus* harboring naive CRISPR systems with a single repeat were diluted to OD = 0.1 in BHI supplemented with calcium chloride and relevant antibiotics and grown for 1.5 hours. Cultures were normalized to the OD of the lowest culture, and phage ϕNM4γ4 was added at the specified MOIs. 150 μL of each culture was added to a clear, flat-bottom 96-well plate (Grenier 655180), and the plate was incubated at 37°C with shaking in a TECAN Infinite F Nano+. OD600 measurements were recorded every 10 minutes for 24 hours.

Top agar immunity Assay

100 μL of overnight cultures of *Staphylococcus aureus* harboring naive CRISPR systems were mixed with ϕNM4γ4 at MOI = 25 and 6 mL of BHI top agar (0.75% agar) supplemented with calcium chloride and poured onto BHI agar (1.5%) plates. After 15 minutes at room temperature, plates were moved to a 37°C incubator overnight, and surviving colonies were counted. When indicated, colony PCR was performed with primers oJW154 and oJW355 to verify expansion of the CRISPR array due to spacer integration.

PCR conditions

PCR was performed with Phusion HF DNA polymerase using 5X Phusion Green Reaction Buffer (Thermo). Each reaction contained 10 μL buffer, 4 μL dNTPs, 0.5 μL each of 100 μM forward and reverse primers, 10-50 ng template, 0.5 μL polymerase and nuclease-free water to 50 μL. Three-step cycling was performed under the following conditions: 98°C for 30 s, 34 cycles of [98°C 5 s, 45-72°C 15 s, 72°C for 30 s/kb], 72°C 10 minutes, hold at 10°C.

RNA extraction

To extract *S. aureus* RNA for Northern blot, RT-qPCR, and RNA-seq analysis, 7.5E8 cells were spun down, resuspended in 150 μL 1X PBS (10X stock, Corning, 46-013-CM) and Lysostaphin (60 μg/ml final concentration), and incubated at 37°C for 5 minutes. Cells were processed from overnight cultures unless otherwise specified. To the whole cell lysate, 450 μl Trizol (Zymo, R2071) and 600 μL 200 proof ethanol were added, samples were vortexed and RNA was extracted using the Direct-Zol Miniprep Plus spin column according to the manufacturer's protocol (Zymo, R2071). Samples were eluted in 50 μL Ambion RNase-free water (ThermoFisher, AM9937). To extract *S. pyogenes* RNA for Northern blot and RT-qPCR analyses, 7.5E8 cells from an overnight culture were spun down, resuspended in 150 μl 1X PBS and PlyC (1 μg/ml final concentration), and incubated at room temperature for 10 minutes. After lysis, the RNA extraction protocol was followed as detailed above.

Infrared Northern (irNorthern)

Total RNA (3-10 μg) was mixed 1:1 with 2X Novex sample buffer (Thermo, LC6876), boiled at 94°C for 3 minutes and placed on ice for 3-5 minutes. Samples were loaded onto a 15% TBE-Urea gel (MINI-Protean, Bio-rad, 4566053) and run at 150 V for 2-2.5 hours. A Hybond N+ membrane (GE lifesciences, 45000854) and 6 sheets of 3 mm Whatman cellulose paper (Sigma Aldrich, WHA3030861) were pre-soaked for 5 minutes in room temperature 0.5X TBE, then assembled into a sandwich of: 3 layers Whatman paper, Hybond

membrane, TBE-Urea gel, and 3 more layers of Whatman paper. Blotting was performed using a Trans-blot Turbo (Bio-Rad) at 200 mA for 30 minutes. The membrane was then pre-hybridized in 10 mL ExpressHyb (Clontech, NC9747391) at 44°C in a rotating oven for 1 hour, and probed overnight at 44°C, with rotation, using probes conjugated to irDyes (LICOR; 680CW 929-50010, 800CW 929-50000), made following the protocol detailed at (https://bio-protocol.org/e3219) (Fields et al., 2019). The membrane was washed once with 2X SSC/0.1% SDS, and once with 1X SSC/0.1% SDS each for 10 minutes at RT, then visualized on the Odyssey Fc (LICOR). *4.5S* RNA was used as a loading control (stability of *4.5S* across genetic backgrounds was verified by qPCR, data not shown). For Northern blots performed using *S. pyogenes* cells using the *tracrRNA* probe, a nonspecific band (NS) was used as a loading control. Sequences for oligos used to probe *crRNA*, *tracrRNA* and *4.5S* RNA (oJW2313, oJW1991, oJW2172) are listed in Table S4.

Western blot

1.8E8 *S. aureus* cells were resuspended in 1X PBS supplemented with Lysostaphin (60 μg/ml final concentration) and incubated at 37°C for 20 minutes. Cells were processed from overnight cultures unless otherwise specified. Whole cell lysate was mixed 1:1 with 2x Laemmli solution (Bio-rad, 1610737) supplemented with B-mercaptoethanol (55 mM stock, ThermoFisher, 21985023) at a final concentration of 1.3 mM, and boiled at 98°C for 10 minutes. Samples were loaded onto a 4%–20% Tris-glycine gel (MINI-Protean TGX Pre-cast, Bio-rad, 4561095) and run at 200 V for 15 min-1 hour. A PVDF membrane (Bio-rad) was hydrated with methanol for 15-30 s and pre-wet alongside a stack of blotting paper for 3-5 minutes in 1X Transfer buffer. A blotting sandwich was assembled consisting of six layers (one stack) of filter paper, the PVDF membrane, Tris-glycine gel, and another stack of filter paper. Samples were transferred onto a nitrocellulose membrane with the Trans-blot turbo (Bio-Rad, 1704150), with high molecular weight transfer settings (1.6 A for 10 minutes), and the membrane was stained with Ponceau for 5 minutes to perform total protein normalization. In western blot figures, the prominent Ponceau-stained band is shown as a loading control (total protein, TPN). The membrane was blocked with 5% nonfat dry milk for 1 hour, then probed with a 1:1000 dilution of Cas9 monoclonal antibody (Cell Signaling, 7A9-3A3) for 2 hours at RT, or at 4°C overnight. The membrane was washed with 1X TBST buffer 3x for 10 minutes at RT, then probed with 1:15,000 dilution of infrared (LICOR, 925-32210) or 1:10,000 dilution of HRP (Pierce, PA174421) secondary antibody for 1 hour at room temperature. The membrane was washed as before, then visualized on the Odyssey Fc (LICOR). To analyze *S. pyogenes* Cas9 levels, ~1.8E8 cells were resuspended in 1X PBS supplemented with PlyC (1 μg/ml final concentration) and incubated at room temperature for 10 minutes. After lysis, the Western protocol was followed as detailed above.

RNA-seq library preparation and sequencing

To prepare RNA libraries for next-generation sequencing, we first depleted ribosomal RNA from the total purified RNA using reverse capture of biotinylated probes annealing to rRNA with streptavidin beads, following the protocol detailed at (https://mbio.asm.org/content/11/2/e00010-20) (Culviner et al., 2020). rRNA depleted samples were prepared for sequencing using the Cleantag Small RNA library prep kit (Trilink, L-3206-24) according to the manufacturer's instructions, with 10 ng RNA input and 18 PCR cycles. Sequencing was performed on an Illumina MiSeq with a v3 2x75 cycle kit at the Johns Hopkins Genome Resources Core Facility (GRCF).

qPCR

Purified total RNA (1-4 μg) was treated with TURBO DNA-free (ThermoFisher, AM1907) and reverse transcribed with Superscript IV (ThermoFisher, 18090050) according to the manufacturer's instructions. cDNA was diluted to 0.5-1 ng/μL and 4-8 ng was used as input in an 8 μL volume. 1 μL of 10 μM forward/reverse primers were used, and 10 μL of 2X PowerUp SYBR mastermix (ThermoFisher, A25742). qPCR was performed with cycling conditions: 50°C for 2 minutes, 95°C for 2 minutes, 39 cycles of [60°C 1 minute], followed by a melt curve: 65°C to 95°C, incrementing 0.5°C every 5 s. *rho* was used as a loading control and was amplified using primers oJW2003/2004. *cas9* was amplified with primers oW262/oJW1986. *cas1* was amplified using primers oJW150/oL432. *csn2* was amplified using primers oJW2139/2140. The region spanning *csn2* and *crRNA* (*csn2_cr*) was amplified using oJW2141/2142. The *pre-crRNA* was amplified using oJW2142/2143. All forms of *tracrRNA* (*tracr-L*, *tracr-S*, and *tracr-P*) were amplified using oJW2012/2013. *Streptococcus agalactiae cas9* was amplified using primers oJW2849/2850. In the native *Streptococcus pyogenes* system, gyrase A (*gyrA)* was used as a loading control and was amplified with primers oJW2595/2596. qPCR primer sequences are provided in Table S4.

Transformation assay

10 mL overnight cultures of cells harboring derivatives of pTP16, which encodes *cas9* and *tracrRNA* variants with increasing *tracr-L* match lengths to the GFP reporter plasmids pJW711 and pCN57 were diluted to OD = 0.1 and outgrown in BHI supplemented with chloramphenicol for 1-2 hours until the OD was between 0.8-1. The cultures were centrifuged at 4200 RPM for 10 minutes and washed twice with 1 mL of ice-cold deionized water in a 1.5 mL Eppendorf tube with 6000 RPM 1 minute spins between washes. Cells were resuspended in 150 μL 10% glycerol and stored at −80°C if not used immediately. 50 ng of pJW711, pCN57, or an empty vector pE194 was electroporated into 50 μl of competent cells and outgrown in 300 μl BHI at 37°C with shaking for 3 hours. 100 μL of each outgrowth was plated onto BHI agar plates supplemented with chloramphenicol and erythromycin and total transformants were counted the following day.

Please cite this article in press as: Workman et al., A natural single-guide RNA repurposes Cas9 to autoregulate CRISPR-Cas expression, Cell (2021), https://doi.org/10.1016/j.cell.2020.12.017

In vitro transcription (IVT)

IVT templates were generated using pTP16 and pRW22-26, the forward primer oJW2267 and reverse primers oJW2268-2274, which begin with the T7 promoter sequence (5′- GAAATTAATACGACTCACTATAGG-3′). To generate *tracr-L*, pTP16 was amplified with oJW2267 and oJW2268. To generate the GFP-targeting *tracr-L* with an 11-bp match, pRW22 was amplified with oJW2267 and oJW2269. To generate the GFP-targeting *tracr-L* with a 13-bp match, pRW23 was amplified with oJW2267 and oJW2270. To generate the GFP-targeting *tracr-L* with a 15-bp match, pRW24 was amplified with oJW2267 and oJW2271. To generate the GFP-targeting *tracr-L* with a 17-bp match, pRW25 was amplified with oJW2267 and oJW2272. To generate the GFP-targeting *tracr-L* with a 19-bp match, pRW26 was amplified with oJW2267 and oJW2273. To generate the sgRNA version of the GFP targeting *tracr-L* with a 20-bp match, pRW27 was amplified with oJW2267 and oJW2274. *tracr-L* (wild-type and variants) was in-vitro transcribed from 1 μg of these templates using the HiScribe T7 High Yield RNA synthesis kit (NEB, E2040S). Reactions were incubated at 37°C for 4 hours and run on a 6% TBE-Urea gel (Invitrogen, EC6265BOX). Bands of the appropriate size were excised, added to a sterile Eppendorf tube in 450 μL gel extraction buffer (1X: 0.3 M NaOAc, 10 mM Tris-Cl pH 7.5, 1 mM EDTA), and incubated overnight at 4°C in an end-over-end rotator. The buffer with dissolved gel slice was added to 1 mL ice-cold 100% ethanol and 2 μL of glycogen (20 mg/ml, ThermoFisher, R0561) and incubated at −20°C for at least one hour. Samples were centrifuged at max speed at 4°C for 30 minutes, the supernatant was decanted, and the pellet was washed with 1 mL cold 70% ethanol. After another centrifugation step at max speed for 10 minutes at 4°C, the pellets were either dried in a vacuum centrifuge or let air dry for 10 minutes and eluted in Ambion RNase-free water.

Electrophoretic mobility shift assay (EMSA)

DsDNA target substrates were prepared by mixing 20 μL of 20 μM 55-nt top and bottom strand oligos (P_{cas}: oJW2507/2508, P_{gfp}: oJW2509/2510, P_{cas}^{NGC}:oJW2816/2817) in annealing buffer (10 mM Tris pH 7.5, 50 mM NaCl, 1 mM EDTA), heating to 95°C for 5 minutes, then slowly cooling the mixture to room temperature for 1-2 hours. Annealed oligos were mixed with 4 μL 6X loading dye (NEB, B7024S), loaded onto a lab-made 8% TBE gel (Acrylamide/Bis solution 37.5:1, Bio-rad, 1610148), run at 200V for 1 hour, then stained with SYBR Gold (ThermoFisher, S11494) for 10 minutes still, and 10 minutes shaking. Bands of the appropriate size were excised and added to a sterile Eppendorf tube in 450 μL gel extraction buffer (1X: 0.3 M NaOAc, 10 mM Tris-Cl pH 7.5, 1 mM EDTA), and incubated overnight at room temperature in a shaker. Gel extraction buffer was added to 1 mL ice-cold 100% ethanol and 2 μL of glycogen (20 mg/mL, ThermoFisher) and incubated at −20C for at least one hour. Samples were centrifuged at max speed at 4°C for 30 minutes, and the pellet was washed with 1 mL cold 70% ethanol. Following another centrifugation step at max speed for 10 minutes at 4°C, pellets were either dried in a vacuum centrifuge or let air dry for 10 minutes, then eluted in RNase-free water (Ambion). Purified dsDNAs were diluted to 200 nM with binding buffer (20 mM HEPES pH 7.5, 250 mM KCl, 2 mM MgCl2, 0.01% Triton X-100, 0.1 mg/mL BSA, 10% glycerol), then diluted again 1:10 in binding buffer to make a 20 nM stock.

After dsDNA purification, the substrate was radiolabeled with P32 (1 μL 20 nM oligo duplex, 5 μL T4 PNK buffer, 10 mM radioactive ATP, 2 μL T4 PNK in a 50 μL reaction). This reaction was incubated at 37°C for 1 hour and cleaned up with ProbeQuant G50 spin columns (GE Healthcare, GE28-9034-08) as per manufacturer's instructions.

Wild-type *tracr-L* was prepared by performing *in vitro* transcription as described in *in vitro* transcription methods above. *Tracr-L* was diluted to 8 μM in RNA hybridization buffer (20 mM Tris-HCl, pH 7.5, 100 mM KCl, 5 mM MgCl2), then heated at 95°C for 30 s and slow cooled to room temperature (1-2 hours). Cas9 (10 mg/mL, generously gifted by the Seydoux lab) was diluted to 8 μM in storage buffer (20 mM HEPES pH 7.5, 500 mM KCl). Cas9:*tracr-L* RNPs were pre-formed by mixing 10 μL 8 μM Cas9 with 10 μL 8 μM *tracr-L* and incubating at RT for 10 minutes. RNPs were diluted to 500 pM - 1 μM in a 1:1 mix of RNA hybridization buffer and Cas9 storage buffer, then brought to volume with binding buffer. For the Cas9-only control, RNA hybridization buffer was substituted for *tracr-L*. 2 μl of radiolabeled substrate (final concentration 50 pM) was added to each reaction and mixed gently by pipetting. The binding reaction was incubated at 37°C for 1 hour, then quenched with 2 μL 6X purple loading dye (NEB) at RT. 15 μL of each reaction was immediately loaded onto an 8% TBE gel (Acrylamide/Bis solution 37.5:1, Bio-rad) supplemented with 2 mM MgCl2, and run at 4°C at 15W in 1X TBE buffer + 5 mM MgCl2 for 75-90 minutes. The gel was then removed from its casing, wrapped in plastic wrap and exposed to a phosphor screen overnight. The phosphor screen was scanned and imaged using the Typhoon FLA9500 (GE Healthcare).

In vitro cleavage assay

20 μL nuclease-free water, 3 μL NEBuffer 3.1, 3 μL of *in vitro* transcribed 300 nM *sgRNA* or *tracr-L* (produced as described in *in vitro* transcription protocol above) and 1 μL of 1 μM Cas9 nuclease (NEB, M0386S) were mixed and incubated for 10 minutes at 25°C to allow for RNP formation. Next, 3 μL of a 30 nM dsDNA substrate (a 1kb pCN57 amplicon, amplified from pCN57 with oJW134 and oJW1869) was added, mixed thoroughly, pulsed in a microcentrifuge, and incubated at 37°C for 15 minutes. Final concentrations of the reaction components were 30 nM *sgRNA* or *tracr-L*, 33.3 nM purified Cas9 and 3 nM dsDNA. 1 μL of proteinase K (20 mg/mL, QIAGEN) was added and samples were incubated at RT for 10 minutes. 6X loading dye was added, and the reactions were run on a 1.5% agarose gel.

Competition assay

To investigate the fitness costs of Δ*tracr-L*, cells harboring a naive wild-type or Δ*tracr-L* CRISPR system were co-cultured in a long-term competition assay. Wild-type and Δ*tracr-L* cells were inoculated in BHI in triplicate and grown overnight at 37°C, diluted to log

phase OD = 0.1, and replicate pairs were mixed 1:1. 50 μL of a 1:100 dilution of this mixture was immediately plated, and 900 μL was frozen at −80°C in 10% DMSO. The remainder of the mixture was grown to stationary phase in the evening, then diluted back 1:1000 and grown overnight. This continued for a total of 5 days, with frozen stocks taken of the cultures each morning before dilution. To determine the ratio of wild-type:Δ*tracr-L* cells over time, 5 μL of frozen cultures were diluted 1:5000 and plated, and colonies were picked for colony lysis and PCR of the *tracrRNA* locus (oJW2012/oJW1447). Amplicon length (314 bp for wild-type and 262 bp for Δ*tracr-L*) allowed us to estimate the relative abundance of each strain in the mixed culture over time.

QUANTIFICATION AND STATISTICAL ANALYSIS

Tn-seq screen analysis

Sequencing reads were processed with custom Python scripts and aligned to the NCTC8325 genome using the Burrows-Wheeler Aligner (Li and Durbin, 2009). Tn-seq output was visualized with the Integrative Genomics Viewer3 (Robinson et al., 2011).

Promoter activity fluorescence assays

Measurements for absorbance (at 600 nm) and fluorescence (excitation wavelength = 485 nm; emission wavelength = 535 nm) were taken using a TECAN Infinite F Nano+. For each experimental strain, promoter activity was measured as (Fe)/(Ae) - (Fc)/(Ac) where F = fluorescence, A = absorbance, e = experimental strain and c = non-fluorescent control strain.

Electrophoretic mobility shift assay (EMSA) dissociation constant extraction

Bound and unbound intensities for EMSA blots were quantified using ImageJ, and fraction bound was plotted against RNP concentration using GraphPad for Mac version Prism 7.4, GraphPad Software, San Diego, California USA, https://www.graphpad.com:443/. A nonlinear regression curve fit was applied to extract the K_d using Prism 7.4, and was performed separately for two independent analyses.

RNA sequencing analyses

Sequencing reads were aligned to the NCTC8325 genome and CRISPR plasmid pCR using the Burrows-Wheeler Aligner. Binary alignment files (BAM) were converted to bedgraph format and scaled to reads per million (RPM) using bedtools genomecov. Bedgraph pile-ups were visualized with the Integrative Genomics Viewer.

Spacer sequencing analyses

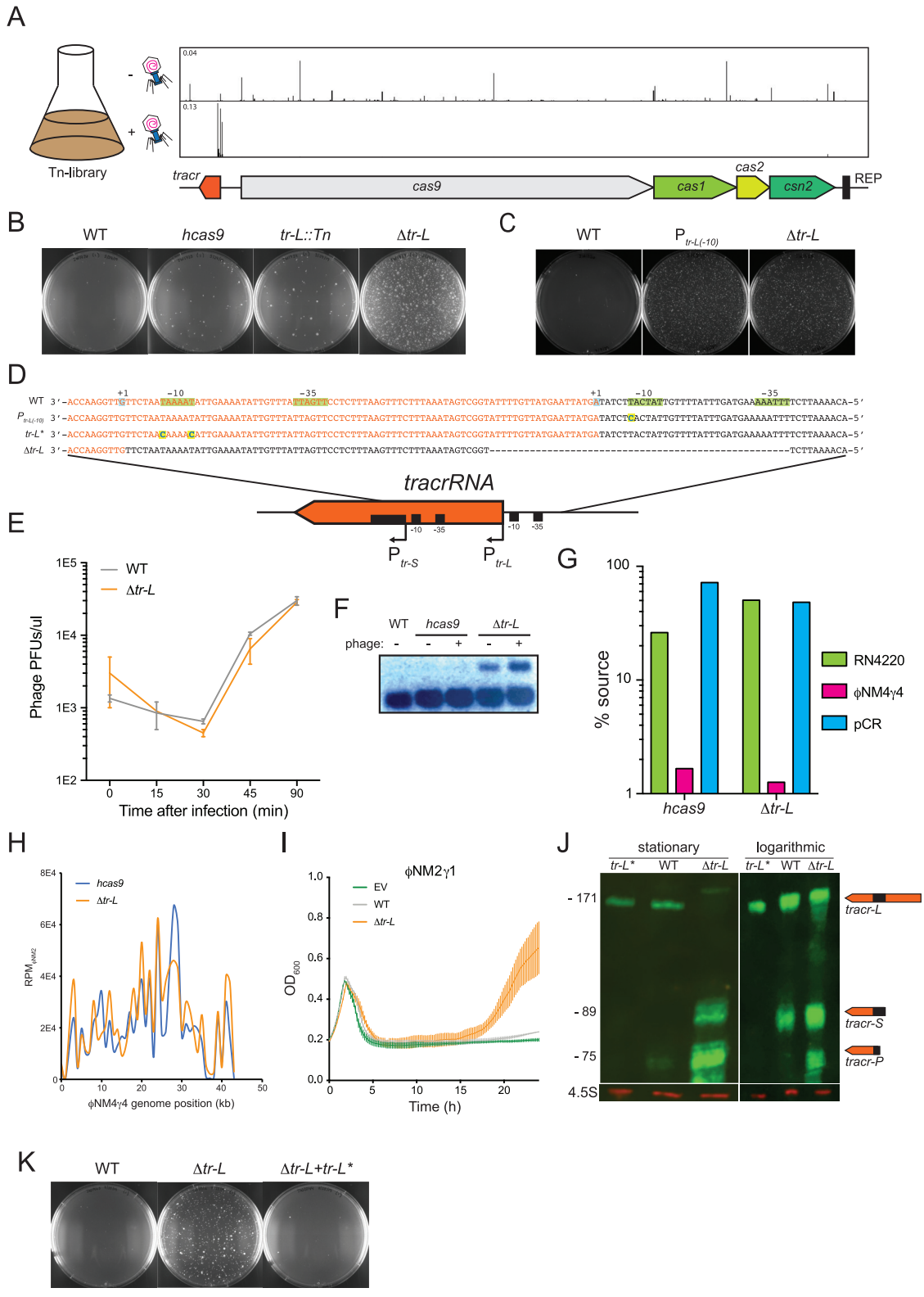
Sequencing reads were processed with custom Python scripts and aligned to the NCTC8325 genome, CRISPR plasmid pCR (shorthand for pGG32, pRH163, or pGG32_Δ*tr-L* depending on the sample), or bacteriophage ϕNM4γ4 reference sequences using the Burrows-Wheeler Aligner. Binary alignment files (BAM) were converted to bed format. Spacers with no mismatches were aggregated by genomic position, then multiplied by a correction factor accounting for the last base of the spacer (A: 1.812265748, C: 0.267975994, T: 5.640697298, G: 2.467159171), in order to normalize values to the expected frequencies. This normalization was applied as the enrichment PCR process intentionally avoids priming against the single repeat in non-adapted plasmids by using a mix of primers complementary to the last base of a newly acquired spacer (3′ ending in T, A, or G) and omitting the reverse primer ending in C. Additionally, because of the increased bond strength of G-C basepairs, spacers ending in C are overrepresented compared to those ending in A or T. To visualize spacer distributions from a given source, reads per million (RPM) were calculated as follows: the number of normalized reads at a given location were divided by the total number of reads aligned to that source for that experiment and multiplied by one million (RPM_{chr} for spacers matching the RN4220 chromosome; $RPM_{\phi NM2}$ for spacers matching ϕNM4γ4). RPM were aggregated in 10 kb bins for the RN4220 chromosome and 1 kb bins for ϕNM4γ4.

Evolutionary analyses

Bacterial strains harboring type II-A CRISPR-Cas systems with previously annotated short-form *tracrRNAs* (Faure et al., 2019) were investigated for evidence of *tracr-L* repression. Accession numbers for species investigated are listed in Table S2 along with all metadata compiled in our analyses. The genomic locus between the annotated 3′ end of *tracrRNA* and the translational start site of Cas9 was interrogated using the DSK k-mer counting software (Rizk et al., 2013) with k-mer length set to 11 bp and a minimal match frequency of 2. We classified strains as "*tracr-L+*" if one k-mer instance was on the same strand as *tracr-S* followed immediately by a 5′-GTTTTA-3′ sequence allowing for 1 mismatch (representing the targeting site within *tracr-L*) and the other instance was upstream of site 1, immediately followed by a 5′-NGG-3′ sequence allowing for 1 mismatch (representing the targeted site within P_{cas}). We manually inspected the 16 *tracr-L* loci with intergenic lengths above 167 bp for candidate *tracr-L* targeting sites with a single seed mismatch but perfect lower stems and PAMs, and found one additional *tracr-L* sequence with a 10 bp match.

We aligned *tracrRNA* sequences using MAFFT (Katoh et al., 2002) (default parameters). Cas9 amino acid sequences for these strains were aligned using MUSCLE (Edgar, 2004), and a phylogenetic tree was constructed using FastTree (Price et al., 2010), with the WAG evolutionary model. All multiple sequence alignments and tree generation were performed in Geneious Prime 2020.1.1.

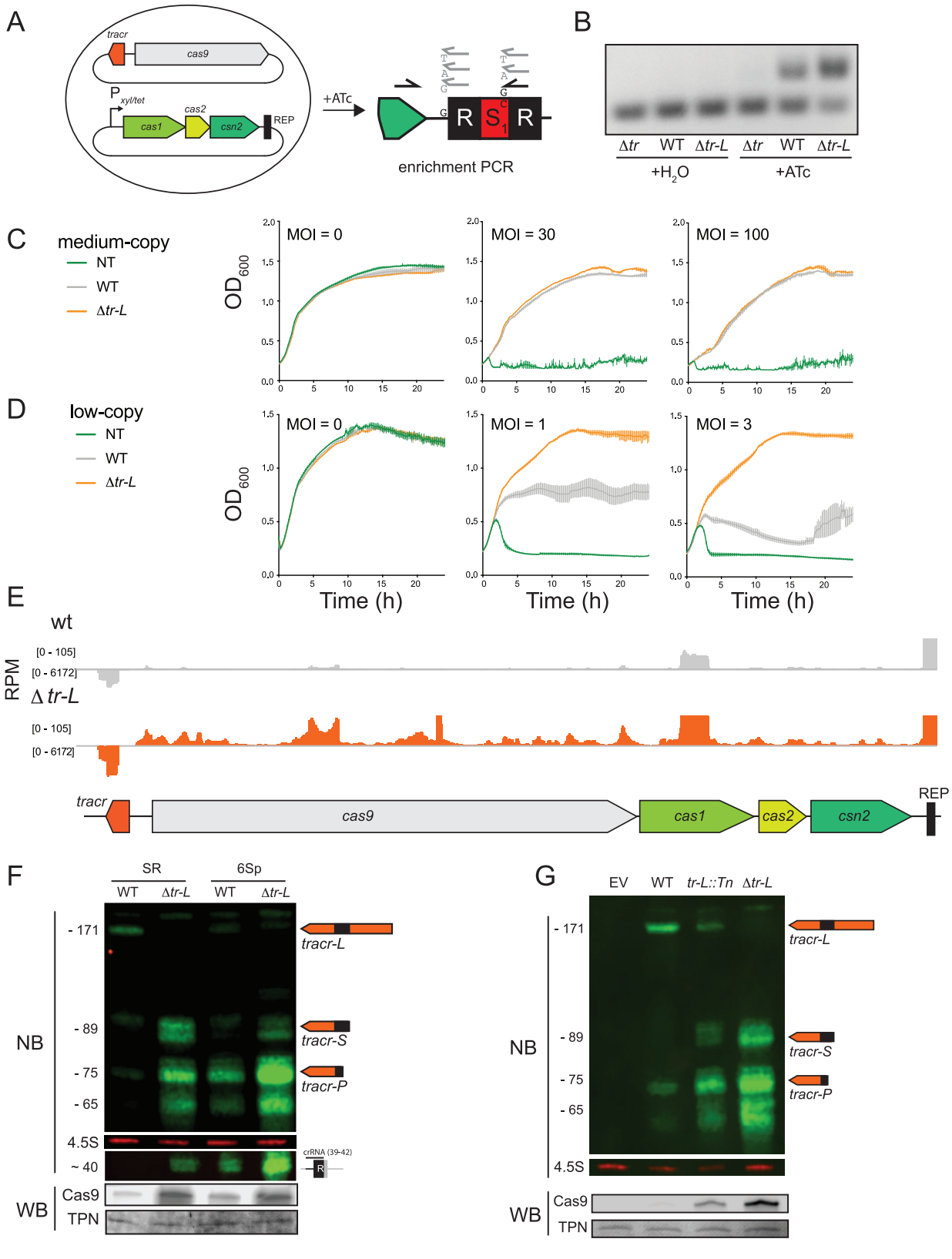
Supplemental Figures



(legend on next page)

Figure S1. Characterization of *tracrRNA* mutants, related to Figure 1

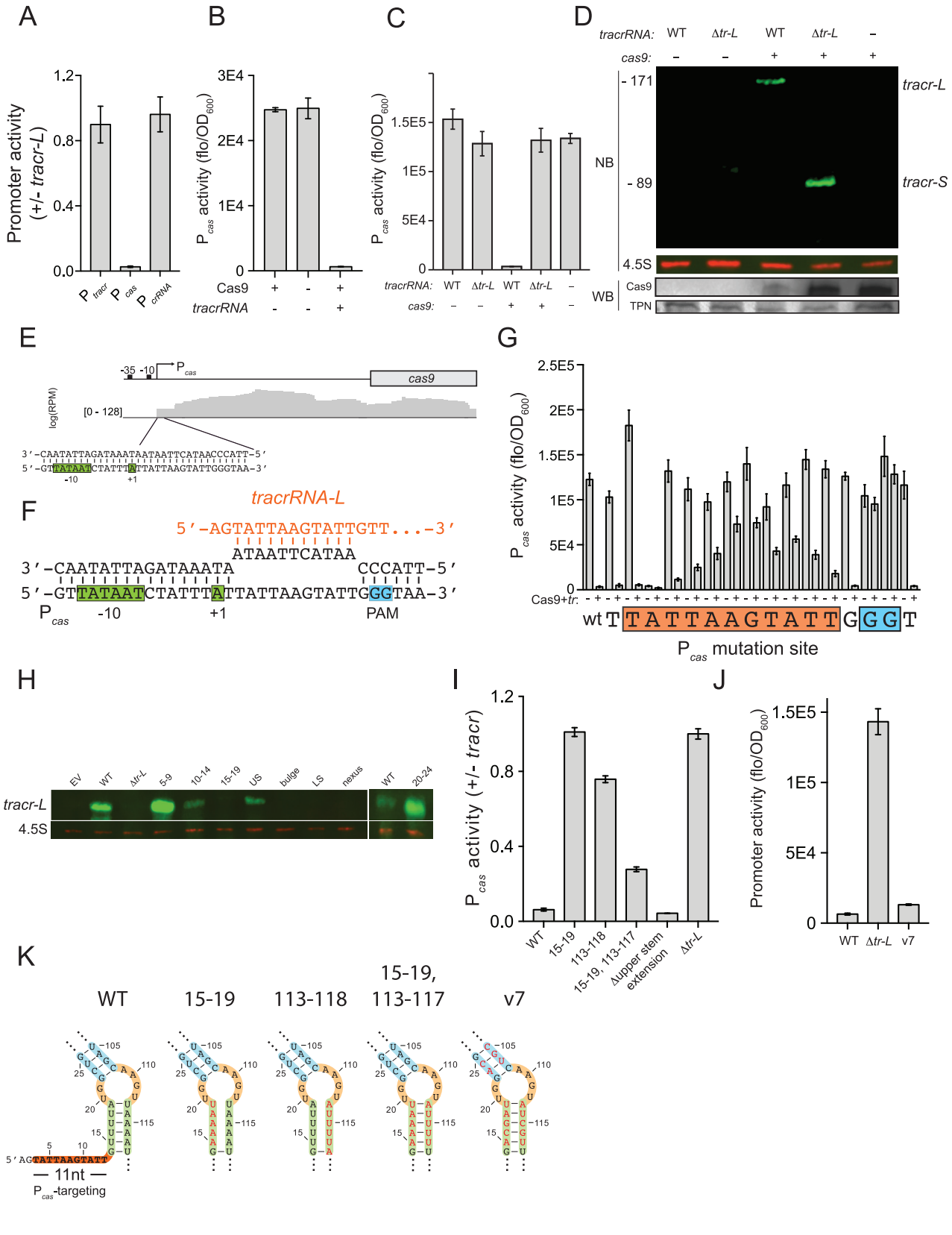
(A) An extended view of Tn-seq reads in the CRISPR-Cas region from Figure 1B for - phage (top panel) and + phage (bottom panel) experiments. The + phage panel is a representative experiment of biological triplicates. The location of each hit corresponds to the gene annotations shown in cartoon format below the Tn-seq reads. (B) Representative plates are shown from the top agar immunity assays performed in Figure 1D. (C) Top agar immunity assay including a single inactivating mutation in the $P_{tr\text{-}L}$ −10 element ($P_{tr\text{-}L(-10)}$) (D) Schematic of the *tracrRNA* promoter region. The wild-type sequence is shown in the top row with transcribed sequences shown in orange. $P_{tracr\text{-}L}$ and $P_{tracr\text{-}S}$ are each indicated with a transcriptional start site (+1) and the −10 and −35 promoter elements that contact RNA polymerase. In $P_{tr\text{-}L(-10)}$, a T > C mutation in the −10 element was introduced to inhibit *tracr-L* transcription. In *tracr-L** (middle row), two T > C mutations in the tracr-S −10 element were introduced to block *tracr-S* transcription. In Δ*tr-L* (bottom row), the *tracr-L* promoter and 19 5′ nucleotides have been deleted. (E) Supernatant PFUs were counted following an infection of *S. aureus* cells harboring a plasmid expressing a naive CRISPR-Cas system with ϕNM4γ4 at the indicated time points. (F) Enrichment PCR (Modell et al., 2017) of the CRISPR array in cells harboring a naive CRISPR-Cas system on a medium-copy plasmid, with the indicated *cas9* or *tracrRNA* alleles, treated with or without ϕNM4γ4 at MOI = 10 for 30 min. Bottom band, un-adapted array; top band, adapted array with a newly acquired spacer. (G and H) Newly acquired spacers from the + phage experiments in (F) were analyzed by NGS, and the percentage of normalized spacers derived from each source (G) and the distribution of spacers across the ϕNM4γ4 genome (H) are shown. (I) Cells harboring a plasmid expressing a naive CRISPR-Cas system were infected with ϕNM2γ1 at MOI = 3 and cell densities (OD_{600}) were measured every 10 minutes in a 96-well plate reader. (J) Northern blot on logarithmic or stationary phase cells harboring a plasmid expressing a naive wild-type, Δ*tr-L* or *tracr-L** CRISPR-Cas system. In *tracr-L**, *tracr-S* transcription is below the limit of detection and *tracr-P* levels are greatly reduced, likely owing to poor Cas9-binding and/or processing of *tracr-L* compared to *tracr-S*. (K) Representative plates are shown from the top agar immunity assays performed in Figure 1E.



(legend on next page)

Figure S2. Δ*tracr-L* cells exhibit enhanced CRISPR adaptation and interference, related to Figures 2 and 3

(A) Schematic of spacer acquisition assay on cells harboring a plasmid expressing *tracrRNA* and *cas9* and a second plasmid expressing the adaptation-specific *cas* genes and CRISPR repeat from the anhydrotetracycline(ATc)-inducible promoter $P_{xyl/tet}$. ATc was added at 0.5 ug/mL for 2 h to induce expression of the adaptation cassette, and spacer acquisition was monitored by an "enrichment PCR" assay (Modell et al., 2017). (B) PCR products were separated on a 2% agarose gel, and the presence of a larger band represents a single newly acquired spacer. (C and D) Interference assays on cells expressing CRISPR systems with the ϕNM4γ4-targeting spacer NM2 on medium (B) and low-copy (C) plasmids. Cells were treated with ϕNM4γ4 at the indicated MOIs and cell densities (OD_{600}) were measured every 10 minutes in a 96-well plate reader. (E) RNaseq from cells with a plasmid expressing a wild-type or Δ*tr-L* CRISPR-Cas system, outgrown from stationary to mid-logarithmic phase for 4 h. (F) Northern and western blots on the indicated strains. SR, single repeat; 6Sp, the six-spacer *S. pyogenes* CRISPR array. (G) Blots were performed as in (B) on the indicated strains from Figures 1C and 1D.



(legend on next page)

Figure S3. Extended characterization of the determinants of *tracr-L*:P_{cas} repression, related to Figure 4

(A) Promoter activity was measured (fluorescence/OD_{600}) in cells harboring a plasmid expressing GFP from the indicated promoters and a second plasmid expressing *cas9* and either the full *tracrRNA* locus or *tracr-S*. The ratio of promoter activity in *tracrRNA-: tracr-S*-containing cells is shown. (B) Promoter activity was measured (fluorescence/OD_{600}) in cells harboring a P_{cas}-*GFP* reporter plasmid and a second plasmid expressing *cas9* and tracrRNA as indicated. (C) Promoter activity was measured (fluorescence/OD_{600}) in cells harboring a P_{cas}-GFP reporter plasmid and a second plasmid expressing *tracrRNA* and *cas9* or the indicated mutants. In *cas9* null mutants (-), two stop codons were inserted after the 15th codon of *cas9*. (D) Northern and western blot on the strains in (C) grown to stationary phase. (E) RNaseq from cells harboring a Δ*tr-L* CRISPR system outgrown from stationary to mid-logarithmic phase for 4 h. A window is shown focused on the P_{cas} region. (F) Base-pairing between the 5′ end of *tracr-L* and the P_{cas} region just downstream of the transcriptional start site (+1). −10 and +1 promoter elements are shown in green, and the PAM downstream of the *tracr-L* targeted site is shown in cyan. (G) P_{cas} promoter activity was measured as in Figure 4D with the raw values for both ± Cas9+*tracrRNA* conditions shown. (H) Northern blot on a subset of strains shown in Figure 4F. Experiments in different panels were done on different days. (I and J) Promoter activity was measured (fluorescence/OD_{600}) in cells harboring a plasmid expressing cas9 from the constitutive promoter P_{sparc3} and P_{cas}-GFP and a second plasmid expressing the indicated *tracrRNA* mutants or an empty vector. The ratio of activities for the indicated *tracrRNA* mutants relative to the empty vector are shown. Label numbers indicate the positions of nucleotides mutated to their complementary base relative to the 5′ end of *tracr-L*. Δupper stem extension, nucleotides 26 – 104 of *tracr-L* were replaced by a 5′-GAAA-3′ tetraloop. (K) Schematic showing the base-pairing potential of the *tracr-L* mutants from (I and J).

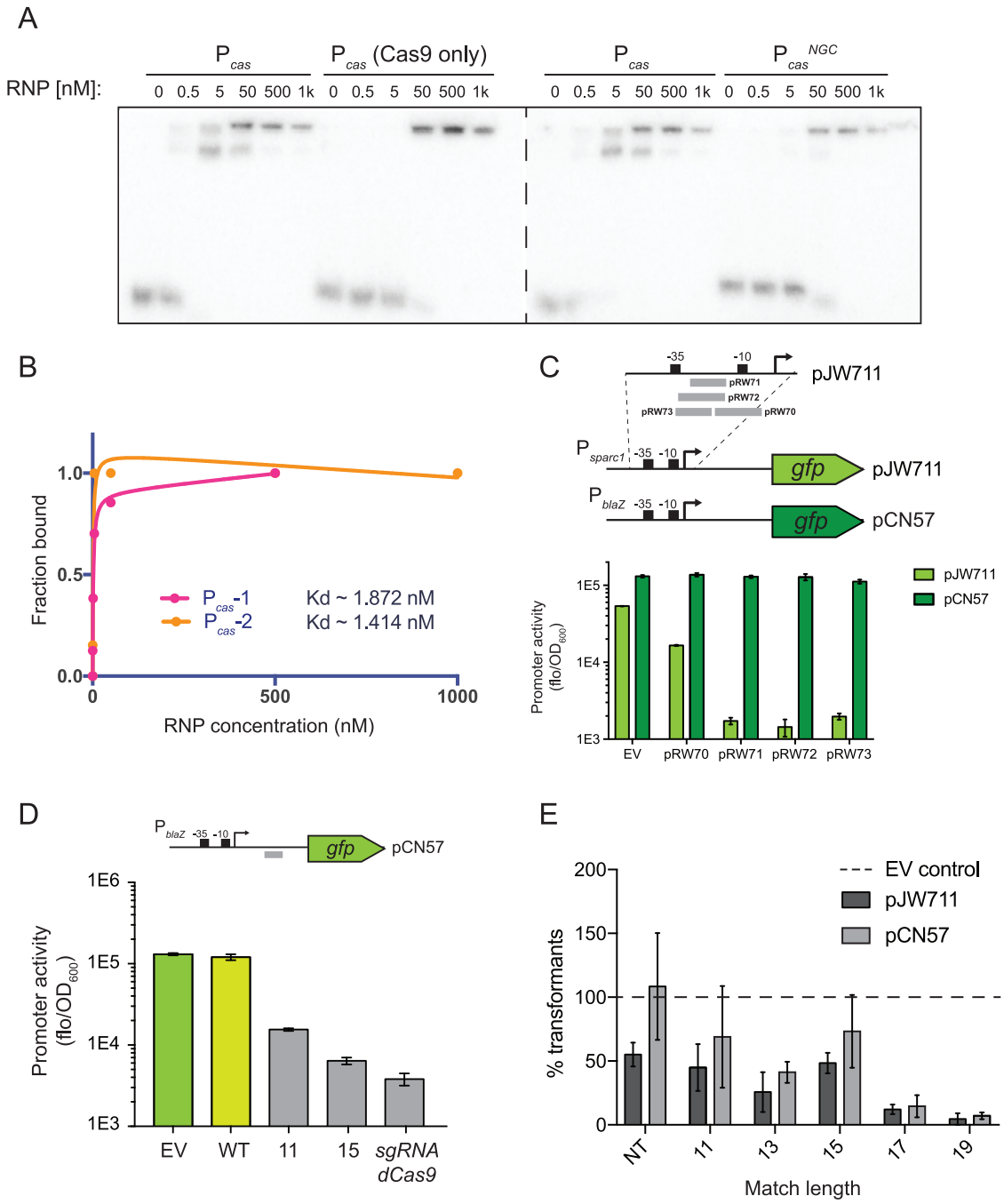


Figure S4. Extended *tracr-L* match lengths lead to tighter repression and cleavage, related to Figures 4 and 5

(A) Electrophoretic mobility shift assay (EMSA) with increasing concentrations of Cas9:*tracr-L* RNPs bound to radiolabeled 55 bp oligos derived from P_{cas} or the P_{cas} PAM mutant (P_{cas}^{NGC}). Also shown is a Cas9-only control in which *tracr-L* was omitted. The presence of a second, more shifted band in the bound fraction is likely due to nonspecific DNA binding by apo-Cas9 in the RNP mixture. (B) The data from each EMSA biological replicate (Figure 4C - P_{cas}-1, S4A - P_{cas}-2) were fit with a standard binding isotherm (solid lines) and the K_d was extracted. (C) Promoter activity was measured (fluorescence/OD_{600}) in cells harboring the indicated GFP reporter plasmid and a second plasmid expressing *cas9* and *tracr-L* reprogrammed to target the sites within P_{sparc1} indicated by gray bars in the top panel. Note that all targets are absent in pCN57 and upstream of the TSS in pJW711. pRW70 and pRW72 target 11 nt sites while pRW71 and pRW73 target 15 nt sites. (D) Promoter activity was measured (fluorescence/OD_{600}) in cells harboring variants of pRW22 (Figure 5A) with target site matches of the indicated lengths and a second GFP reporter plasmid, pJW711 or pCN57. Grey bars in the schematic, targeting site; WT, wild-type *tracr-L* targeting P_{cas}. (E) Cells harboring pRW22 variants with the indicated target site match lengths were transformed with a second plasmid, EV, pJW711 or pCN57. Transformants were plated on selective media with antibiotics to maintain both plasmids, and surviving colonies were counted. Data are presented as a ratio of pJW711 or pCN57 colonies relative to the EV control.

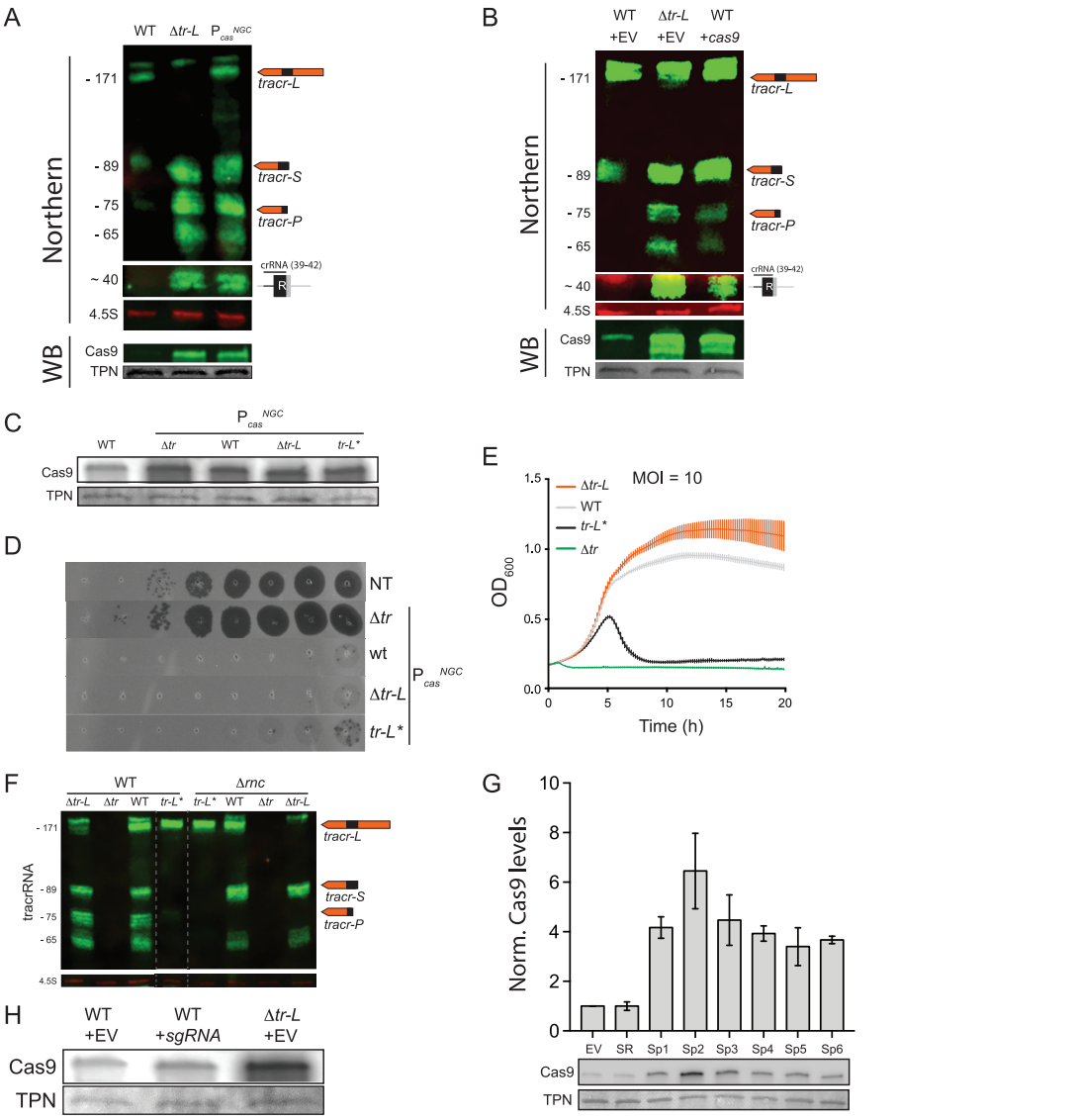


Figure S5. *crRNAs* interact with *tracr-L* and regulate CRISPR-Cas expression, related to Figure 6

(A and B) Northern and western blots of (A) strains from Figure 6A or (B) cells harboring a plasmid with a naive wild-type or Δ*tr-L* CRISPR system and a second empty vector or a plasmid constitutively expressing *cas9* from P_{spac} (+*cas9*). (C) Western blot on cells harboring a plasmid expressing *cas9* from the P_{cas}^{NGC} PAM mutant promoter and the indicated *tracrRNA* mutants. Δ*tr*, deletion of the *tracrRNA* locus. (D) Top agar interference assay with cells harboring a plasmid expressing Cas9, the indicated *tracrRNA* mutants and the ϕNM4γ4-targeting spacer NM2. 10-fold dilutions of ϕNM4γ4 were plated on the indicated bacterial lawns. (E and F) Interference assay with ϕNM4γ4 at MOI = 10 (E) and Northern blot (F) using the strains from Figure 6B which harbor two plasmids. These results show that *tracr-L** can mediate interference at low MOIs (Figure 6B) but not at high MOIs (E). Overnight cultures were diluted into logarithmic phase in the presence of 1mM IPTG for 1:15 h to maximally induce *tracr-L** expression. The Δ*rnc* experiments demonstrate that *tracr-L* and *tracr-S* are expressed at comparable levels in the absence of processing in tr_L* and Δ*tr-L* respectively. Reduced levels of *tracr-P* in the *tracr-L** strain relative to Δ*tr-L* could be due to decreased binding and/or processing of *tracr-L* relative to *tracr-S*. Grey lines indicate where a single gel image was cropped to juxtapose relevant strains. (G) Western blot on logarithmic phase cells harboring a plasmid expressing the wild-type CRISPR system with the CRISPR array deleted and a second plasmid expressing no CRISPR array (EV) a single repeat (SR) or a single spacer from the endogenous *S. pyogenes* CRISPR array (Sp1-6). Quantification of western blot experiments were performed in biological triplicate. (H) Western blot on cells harboring a plasmid expressing the indicated *cas9* and *tracrRNA* alleles and a second plasmid expressing a non-targeting *sgRNA* or an empty vector.

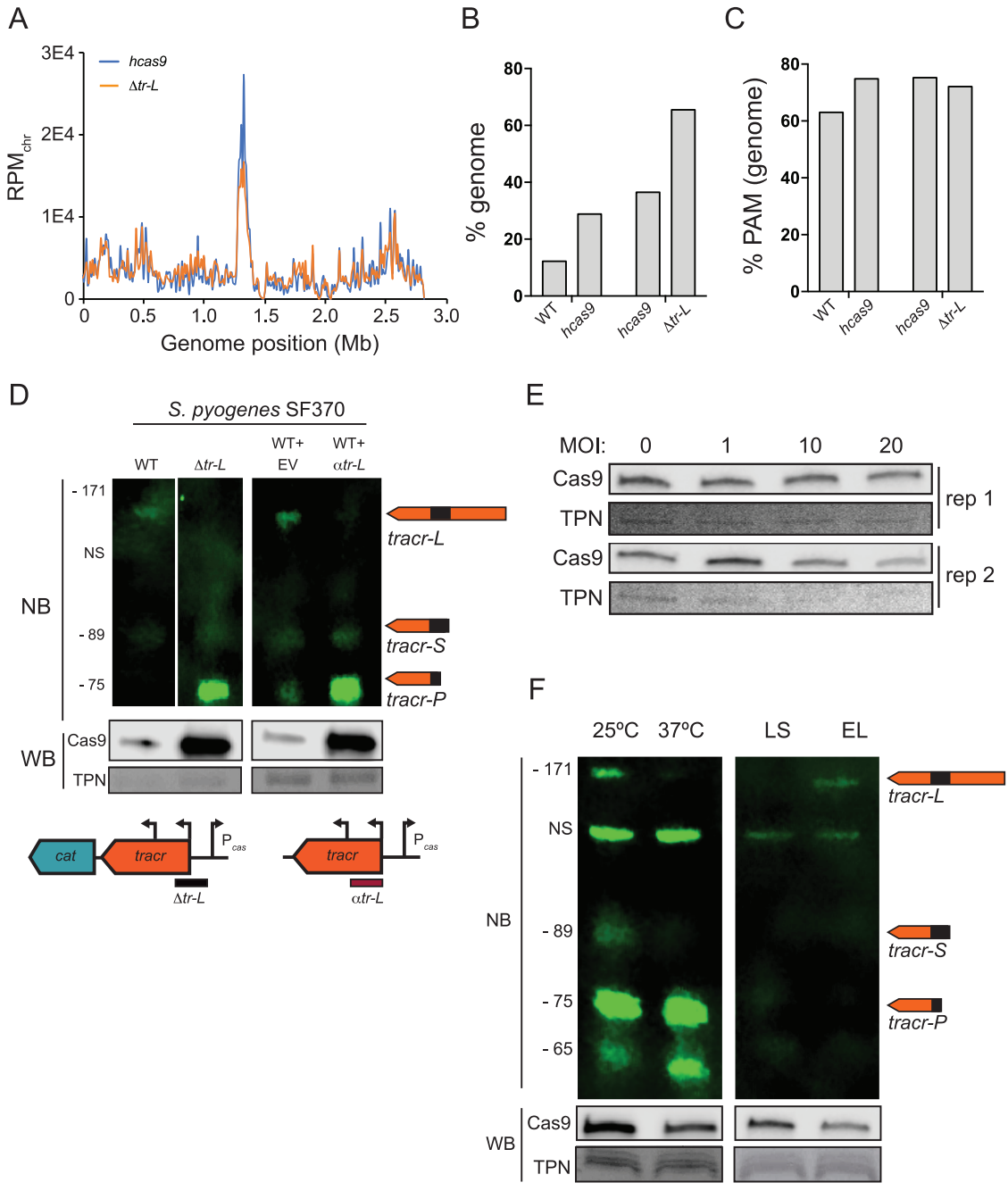
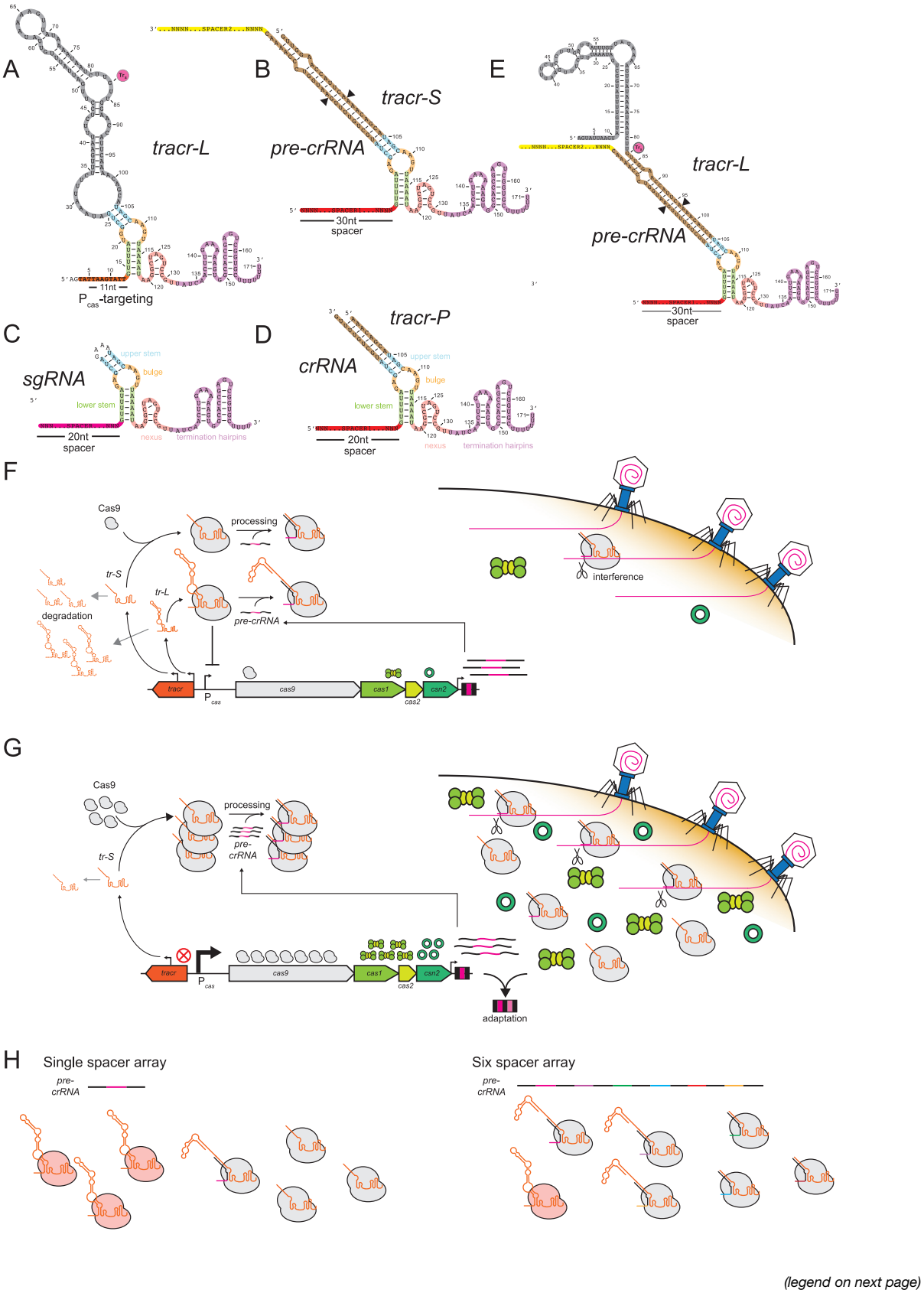


Figure S6. *tracr-L* impacts autoimmunity and is dynamic in response to physiological stresses, related to Figure 6

(A–C) Additional data from the (- phage) NGS analysis performed on the samples shown in Figures S1F and S1G. (A) The distributions of spacers derived from the RN4220 *S. aureus* chromosome are shown. (B and -C) The percentage of chromosome-derived spacers for each *cas9* or *tracrRNA* allele (B) and the percentage of those spacers with correct PAMs (C) are shown. In (B-C), data for the left two bars (WT and *hcas9*) were collected for a previous manuscript (Modell et al., 2017). Note that in Δ*tr-L* cells, more total spacers are acquired in the absence of phage (Figure S1F), and of those, more are derived from the host genome (B). Most of these spacers have correct PAMs (C) indicating that they are not an accumulation of non-functional spacers. Together, these results indicate that autoimmune spacer acquisition occurs more frequently in Δ*tr-L* cells. (D) Northern and western blots on the S. pyogenes strains from Figure 6D grown into mid-logarithmic phase. Cartoons illustrate the location of the chloramphenicol (*cat*) gene in the Δ*tr-L* allelic exchange construct, as well as the location of α*tr-L* antisense RNA target. (E) Western blot of *S. pyogenes* cells infected with lytic phage A25 at MOIs ranging from 0-20 for 45 min. Two replicates representing independent experiments are shown. (F) Northern and western blots on *S. pyogenes* cells grown overnight at either 25°C or 37°C (E), or grown to late stationary (LS) or early logarithmic (EL) phase at 37°C (F). Data are representative of biological triplicates. EL cells were diluted 1:1000 from an overnight culture and grown to OD_{600} = 0.2.



(legend on next page)

Figure S7. Structures of CRISPR-Cas RNAs and model of *tracr-L* regulation, related to Figures 1, 2, 3, 4, 5, 6, and 7

Schematics are shown for the putative *tracr-L* natural *sgRNA* (A), the *tracr-S:pre-crRNA* duplex (B), the processed *tracr-P:crRNA* duplex (C), the artificial *sgRNA* commonly used in CRISPR editing technologies (D), and the *tracr-L:pre-crRNA* duplex (E). Green, lower stem; dark yellow, bulge; cyan, upper stem; pink, nexus; purple, termination hairpins; brown, region of *crRNA:tracrRNA* complementarity outside the stems; gray, putative structures of *tracr-L* regions of unknown function. (F) In wild-type cells, *tracr-L* represses P_{cas} resulting in low levels of *cas* gene expression. Low Cas9 levels destabilize *tracrRNAs* and lead to low levels of *tracrRNA:crRNA* processing. In this state, the CRISPR-Cas system can interfere against viruses at low MOIs but cannot effectively acquire new spacers. *crRNAs* partially relieve repression by binding to *tracr-L* and preventing formation of the natural single guide. (G) In the absence of *tracr-L*, *cas* gene expression is induced, *tracr-S* is stabilized by Cas9 and *tracrRNA:crRNA* processing is enhanced. As a result, interference can occur at higher MOIs and spacer acquisition is stimulated. (H) In cells with a six-spacer array, each *pre-crRNA* provides more repeat-derived binding sites for *tracr-L* relative to a single spacer array, resulting in fewer Cas9:*tracr-L* repressive complexes (shaded in light red).