Title.

Targeted Transcriptional Activation using a CRISPR-Associated Transposon System

Author List.

Andrea M. Garza Elizondo[1] and James Chappell[1,2]*

[1]Department of Biosciences, Rice University, Houston, TX 77005, USA

[2]Department of Bioengineering, Rice University, Houston, TX 77005, USA

*Corresponding author, email: jc125@rice.edu

Abstract.

Synthetic perturbation of gene expression is central to our ability to reliably uncover genotype-phenotype relationships in microbes. Here, we present a novel gene transcription activation strategy that uses the *Vibrio cholerae* CRISPR-Associated Transposon (CAST) system to selectively insert promoter elements upstream of genes of interest. Through this strategy, we show robust activation of both recombinant and endogenous genes across the *E. coli* chromosome. We then demonstrate precise tuning of expression levels by exchanging the promoter elements being inserted. Finally, we demonstrate that CAST activation can be used to synthetically induce ampicillin-resistant phenotypes in *E. coli*.

Keywords.

CRISPR-Associated Transposon (CAST), transcriptional regulation, gene activation, CRISPR, transposons

Introduction.

Understanding the relationships between genotype and phenotype is one of the grand challenges of biology that has driven the development of molecular strategies for manipulating and deciphering gene function and connectivity. A mainstay approach to studying gene function is the perturbation of endogenous gene expression followed by the observation of resulting phenotypic changes. Perturbations meant to activate transcription or overexpress a gene of interest are known as knock-up studies and their inactivation counterparts as knock-out or knock-down studies. Knock-up studies have been conducted through a variety of methods, including transposon mutagenesis, plasmid-based overexpression, and insertion of promoters through recombination. Through these studies, genes have been identified that confer resistance or tolerance to antibiotics,[1–5] oxidizing agents,[6] metals,[7] and other toxins.[2,8] Likewise, alternate enzymatic activities[9] and the effects of overexpression on growth[10] have been identified through knock-up experiments. In eukaryotic cells, knock-up studies have also been used to uncover the organization and connectivity in regulatory networks.[11] Overall, the activation and overexpression of genes has proven to be a fruitful mechanism by which to study genes.

Recently, CRISPR-based strategies for gene activation have been developed, expanding our ability to achieve targeted overexpression of genes of interest. There are two main strategies for activating transcription in prokaryotes using CRISPR-based editors and regulators. One strategy uses CRISPR-Cas9 DNA editing for the insertion of promoter elements upstream of target genes, also known as promoter knock-in.[12–14] CRISPR-Cas9 knock-in allows for precise, scar-less insertion of promoters. The other strategy, CRISPR activation (CRISPRa), uses a catalytically dead Cas9 (dCas9) to localize activation domains upstream of a target gene's promoter, thereby prompting that gene's expression.[15–19] The same dCas9 used for activation can be repurposed for

CRISPR interference (CRISPRi), which makes this strategy attractive for large-scale screens due to the ability to create libraries of single guide RNAs (sgRNAs) for both CRISPRi and CRISPRa in a single experiment.[16,17,19–21] While both strategies confer advantages unique to their systems, challenges remain for applying them at a scale (e.g. for genome-wide studies). For instance, promoter knock-in by CRISPR-Cas requires a customized DNA donor with large homologous arms specific to each edit site (e.g., ~400 bp).[22] This can increase the complexity of library design as both the sgRNA and the DNA donor must be varied simultaneously. On the other hand, CRISPRa benefits from simplified library design, requiring only a sgRNA for each targeted gene. However, different sets of challenges remain (as reviewed in [23]) that include periodical activation patterns which make it difficult to activate diverse promoters,[16,18,24,25] and the use of activation domains that can require customization for different promoter types[24,26] and hosts.[18,24]

Here, we contribute a novel transcription activation strategy that complements existing CRISPR-based systems, using a recently discovered CRISPR-Associated Transposon (CAST). The *Vibrio cholerae* CAST system is a naturally occurring hybrid of type I-F CRISPR-Cas systems and *E. coli* Tn7-like transposons that has retained the guide RNA-processing and RNA-guided binding properties of CRISPR-Cas systems alongside the transposition mechanism of Tn7.[27,28] The result is the targeted insertion of DNA transposons downstream of sites complementary to the encoded CRISPR RNA (crRNA). This unique mechanism of RNA-guided transposition provides the CAST system with several attractive attributes from a gene perturbation standpoint. For one, the CAST system has already shown great versatility for genome knock-downs[28] and knock-outs in multiple species for creation of auxotrophic strains,[29] for metabolic pathway engineering,[30] and for microbiome programming.[31] Furthermore, the CAST system can leverage universal transposons, analogous to the DNA donor used in CRISPR-Cas9 editing, that are interchangeable between different target sites. This makes the system amenable for library strategies wherein one transposon cargo can be used with different crRNAs and target sites.[29,30,32,33] While CAST has only been used for repression, we reasoned that by encoding outward-facing promoter elements within the transposon

cargo and inserting these cargos upstream of silent or weakly-expressed genes, we could adapt the CAST system to also function for activation.

In this paper, we explore the activation potential of the CAST systems. To demonstrate this novel function, we first use the CAST system to insert a strong constitutive promoter as a transposon cargo upstream of fluorescent reporter genes, leading to robust transcription of the fluorescent gene of interest. We then demonstrate the fine-tuning of transcriptional activation using variable strength and inducible promoter systems as transposon cargos. Furthermore, we demonstrate the utility of this approach in non-synthetic contexts by activating the transcription of endogenous *E. coli* genes. Finally, we demonstrate that CAST can be used to identify endogenous genes that confer antibiotic resistance when overexpressed in *E. coli*.

Results.

Engineering the CAST system for efficient transcriptional activation.

As a starting point, our first goal was to ascertain if CAST can be used to activate the expression of genome-encoded genes through the transposition of outward-facing promoter elements upstream of a targeted gene of interest (GOI) (**Figure 1A**). To test this, we focused on adapting the CAST system derived from the *Vibrio cholerae* Tn6677 transposon, which is composed of a crRNA-encoding CRISPR array, seven proteins, and a transposon containing a DNA cargo flanked by transposon left end (LE) and right end (RE) elements.[33] Prior work has shown that this CAST system can be targeted via the crRNA, resulting in high-efficiency insertion of the transposon ~49 bp downstream of the crRNA binding site.[27,33,34] Interestingly, while the transposon can be inserted in either direction (*i.e.,* RE-LE or LE-RE orientation) it shows a significant bias for RE-LE insertions.[27,33,34] Given these features, we hypothesized that the transposition of promoter elements as the transposon cargo could serve as a reliable strategy for activating gene expression (**Figure 1A**).

To test this, we first created a reporter *E. coli* strain in which fluorescent sfGFP and mRFP1 lacking transcriptional promoters were encoded within the genome in divergent

directions (**Figure 1B**). We reasoned we could use this strain to measure the overall efficiency of editing and directionality of insertion by designing a plasmid construct (called pCAST) containing a CAST system, a transposon cargo containing a strong constitutive promoter facing the LE, and a crRNA to target the intergenic region between the two reporter genes. The crRNA was designed to target the template strand upstream of mRFP1, which we presumed would result in predominantly RE-LE insertions and transcription activation of mRFP1. Although transposon insertion directionality is biased toward the RE-LE direction, we also expected low frequency of insertion in the LE-RE direction, resulting in sfGFP fluorescence, and low frequency of no insertion, resulting in no fluorescence (**Figure 1B**). The pCAST plasmid was transformed into the sfGFP-mRFP1 reporter strain followed by an incubation period of 24 hr at 30°C in liquid culture, a method previously shown to result in high transposition efficiencies.[33] Fluorescence was then quantified using single-cell flow cytometry analysis. The fluorescence of the post-incubation bacterial population showed both high transposition and transcription activation of mRFP1 (**Figure 1C, Supplementary Figure 1**). Specifically, the fraction of the population with observed fluorescence was 74.8%, 96.1% of which were from mRFP1 positive cells for a final mRFP1 activation efficiency of 71.9% (**Figure 1D**). We observed a relatively low frequency of no insertion events (24.6%) and a significantly lower rate of LE-RE insertions (2.95%) in agreement with the prior studies.[33]

Next, we sought to determine the distance between the crRNA binding site and the transposon insertion in transcriptionally activated mRFP1 cells. To do this, we isolated mRFP1 positive cells using fluorescence-activated cell sorting (FACS) and performed next-generation sequencing of the sfGFP-mRFP1 genomic locus (**Supplementary Figure 2**). This analysis revealed that the majority (68.4%) of insertions were located 48 bp downstream of the 3' end of the crRNA binding site, in agreement with insertion distances observed in prior literature.[33,34] Additionally, 26.6% of the remaining insertions occurred at low frequency within a 5 bp window centered on 48 bp.

Having confirmed that our CAST activation strategy yielded high levels of transcriptional activation, we were interested in determining possible effects on long-term activation that resulted from the binding of the CAST proteins. Specifically, prior work has shown that binding by the CAST complex can have repressive effects on nearby genes,[28]

possibly affecting the activation achieved through our strategy if the CAST complex remains in the cell. Likewise, we reasoned that the binding of the TnsA and TnsB proteins to the transposon ends[28] could be blocking polymerase read-through into the gene of interest, reducing activation levels. To investigate this, we transformed an empty vector with no CAST complex, an mRFP1 targeting pCAST plasmid, and a non-targeting pCAST plasmid into sfGFP-mRFP1 reporter strains already containing a RE-LE-oriented transposon (**Supplementary Figure 3**). Fluorescence measurement showed modest reductions in mRFP1 expression levels only in the presence of a non-targeting pCAST, suggesting that the presence of the CAST complex or TnsB was not dramatically reducing transcription readthrough.

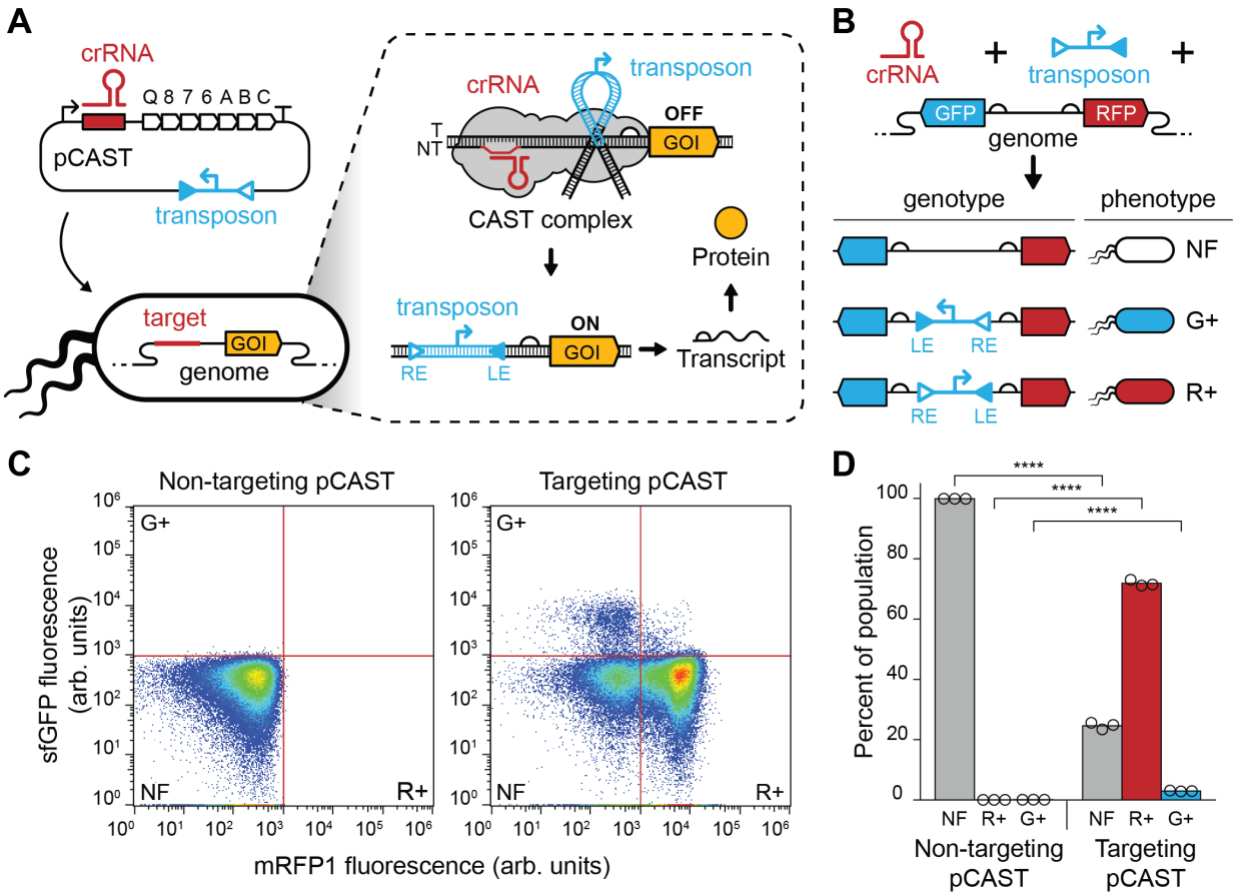

Figure 1. CAST can activate transcription through the insertion of transcriptional promoters. (**A**) Schematic of the mechanism of transcriptional activation by CAST editing through promoter insertion. The pCAST plasmid encodes an operon with the crRNA, the CAST complex, and the transposon with a promoter cargo and the left end (LE) and right end (RE) elements. The crRNA is designed to target the CAST complex

upstream of a gene of interest (GOI). Insertion of the transposon in the right-to-left end (RE-LE) direction results in transcriptional activation of the GOI. (**B**) Schematic of the engineered reporter strain and outcomes of CAST editing. An *E. coli* strain was engineered to contain sfGFP (GFP) and mRFP1 (RFP) reporters lacking promoter elements. CAST editing with a transposon containing a promoter cargo results in cells with mRFP1 fluorescence (R+) if the transposon is inserted in the right-to-left orientation, sfGFP fluorescence (G+) if the transposon is inserted in the left-to-right orientation, and no fluorescence (NF) if editing does not occur. (**C**) Single-cell fluorescence analysis of *E. coli* cells edited with CAST. Flow cytometry data showing the single-cell sfGFP and mRFP1 fluorescence of *E. coli* cells transformed with a non-targeting (left panel) and targeting (right panel) pCAST plasmid. Red lines indicate thresholds used to determine NF, G+, and R+ cell populations. Data are representatives of $n = 3$, with replicates shown in **Supplementary Fig. 1**. (**D**) CAST editing outcomes of *E. coli* cells transformed with a non-targeting and targeting pCAST plasmid. Populations are labeled as in (B), bars represent averages of $n = 3$ replicates. Statistical significance between populations is indicated by asterisks (Student's T-Test, $p < 0.0005$).

Using the CAST system to achieve tunable transcription activation.

After demonstrating that CAST can be used to achieve transcription activation, we next sought to tune the level of activation. To do so, we exchanged the promoter encoded in the transposon cargo for different strength constitutive promoter elements and performed CAST editing to activate mRFP1 expression. As expected, the fluorescence intensity of the mRFP1 positive cells decreased as the strength of the promoter decreased (**Figure 2A, Supplementary Figure 4**). An alternative strategy to achieve tunable control of gene activation is with inducible promoters. To test this, we exchanged the promoter within the transposon for the anhydrotetracycline (aTc)-inducible promoter (PTet) and its cognate repressor, *tetR*. This plasmid was then transformed into the dual reporter strain, and a population containing the transposon inserted in the RE-LE orientation was isolated. Induction of the promoter with increasing concentrations of aTc showed increasing mRFP1 fluorescence intensity, showcasing the ability to tune the level of activation achieved by the promoter (**Figure 2B, Supplementary Figure 5**). Taken together, these results show the potential use of the CAST system as a targeted and tunable transcription activator via the insertion of promoter elements.

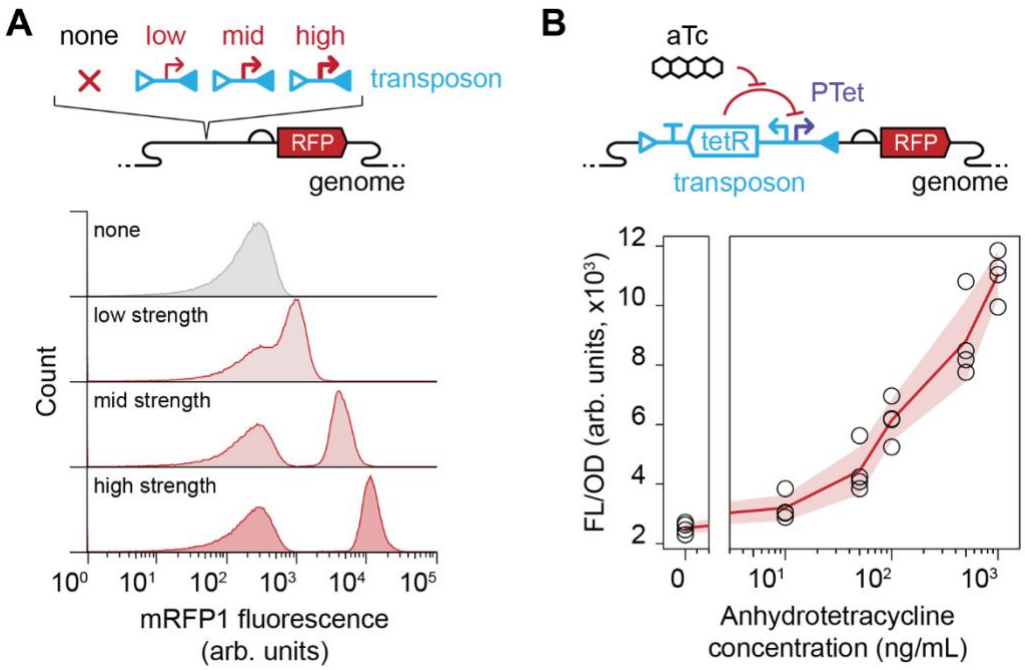

Figure 2. Transcription activation can be tuned by exchanging transposon cargo. (**A**) Schematic of different strength constitutive promoters transposed upstream of mRFP1 using pCAST (top). Flow cytometry data showing single-cell mRFP1 fluorescence of *E. coli* edited with cargos containing high, medium, or low strength constitutive promoters (bottom). Data are representatives of n = 3, with the other replicates shown in **Supplementary Fig. 4**. (**B**) Schematic of the PTet inducible promoter system inserted via CAST (top). Bulk fluorescence analysis of *E. coli* cells edited with a transposon containing an inducible promoter cargo (bottom). Data shows fluorescence characterization (measured in units of fluorescence [FL]/optical density [OD] at 600 nm) of mRFP1 observed at varying anhydrotetracycline (aTc) concentrations (bottom). Red line is the average and shading indicates the standard deviation of n = 4 biological replicates.

Activating transcription of endogenous genes via CAST.

Having shown transcriptional activation in our reporter strain, our next aim was to test the system within a wild-type (WT) strain for activation of endogenous genes. To do so, we selected four genes dispersed across the *E. coli* MG1655 genome with varying genomic contexts and basal expression levels; *ydiN*, a putative transporter, *zraP*, a signaling pathway modulator, *cadB*, a lysine:cadaverine antiporter, and *ffh*, the protein component of a signal recognition particle (**Figure 3A**).[35] A library of pCAST plasmids containing a strong constitutive promoter facing the LE direction as a transposon cargo

and a crRNA targeting the template strand upstream of each gene was designed and transformed into *E. coli* cells. Then, a PCR screen was conducted to isolate edited strains with transposons in the RE-LE direction. To quantify the expression levels of the targeted genes across these strains, reverse transcription quantitative PCR (RT-qPCR) was performed. This analysis showed that, for all four targeted genes, a significant increase in the expression levels between the WT and the CAST-edited strains was observed, suggesting robust activation of endogenous genes can be achieved with our CAST activation strategy (**Figure 3B**). Interestingly, the activated transcription level achieved in the CAST-edited strains was relatively comparable across the different genes, compared to the greater variance seen in their basal expression levels in the WT strain. This leads to lower fold activations in the case of genes with higher basal expression, such as *ffh*, and suggests that the level of expression in CAST-edited strains is primarily determined by the strength of the promoter cargo inserted. Overall, the results demonstrate the ability to robustly activate the transcription of endogenous genes through our CAST-based promoter insertion strategy.

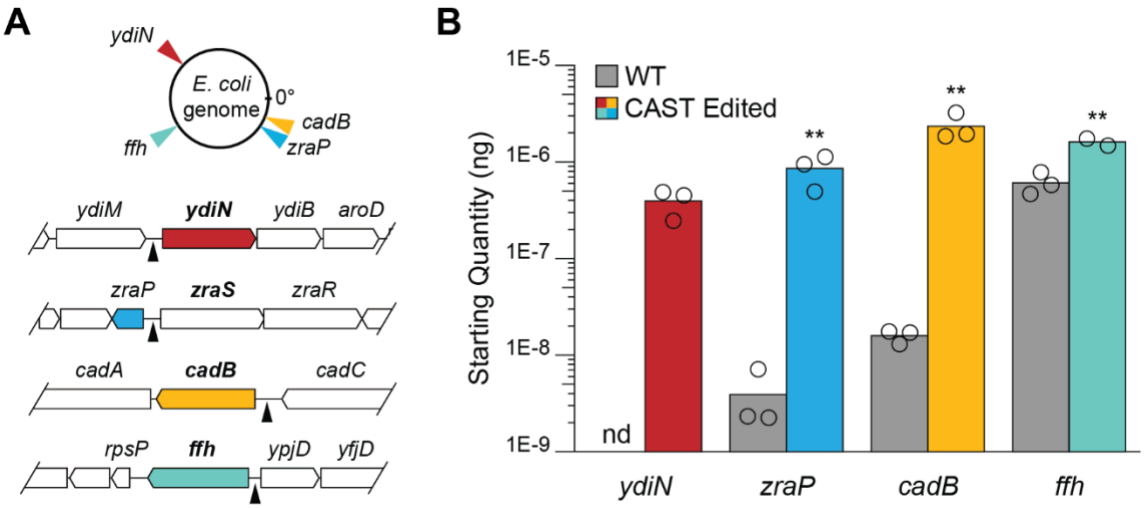

Figure 3. Transcriptional activation of endogenous genes in *E. coli*. (**A**) Location of targeted genes in the *E. coli* K-12 substr. MG1655 genome (top). Genetic context of targeted genes (bottom). Black triangles indicate sites targeted for transposition. (**B**) Starting mRNA quantity of targeted genes with (CAST Edited) and without (WT) insertion of a strong constitutive promoter cargo by the pCAST system, as measured via RT-qPCR. Starting quantities were determined using a relative standard curve. Bars represent averages of $n = 3$ replicates, except for the *ffh* CAST Edited condition that is from $n = 2$ replicates. Statistical significance between conditions is indicated by asterisks (Student's T-Test, $p<0.05$). nd = none detected.

Demonstrating that CAST can selectively activate an ampicillin resistant phenotype.

Having demonstrated activation of endogenous genes, we next sought to investigate if we could selectively induce specific phenotypes using CAST activation. To test this, we sought to activate the *ampC* and *marA* genes that have previously been shown to confer increased resistance to the antibiotic ampicillin when overexpressed.[1,3] As a negative control, we also activated four endogenous genes (*ydiN*, *zraP*, *cadB*, and *ffh)* that have no reported effect on ampicillin resistance. pCAST plasmids were designed with the PTet inducible promoter as a transposon cargo, transformed into WT *E. coli*, and incubated at 30°C in liquid culture for 24 hours to allow for transposition. For the *ampC* and *marA* genes, two different crRNA were designed that targeted distinct positions upstream of each gene. Individual colonies were then grown overnight in liquid culture and sub-cultured the next day, followed by PCR analysis to confirm the presence of the intended RE-LE edits in the population (**Supplementary Figure 6**). We note that some of these populations were also detected to contain LE-RE edits and no edits by PCR analysis, which we reasoned would be a small fraction of the population based on our earlier results (**Figure 1C-D**). Finally, the cells were spotted on plates containing or lacking anhydrotetracycline (aTc) to induce expression of the targeted genes as well as varying concentrations of ampicillin to test for growth. From this analysis, we observed that cells where *ampC* and *marA* were targeted by CAST had an enhanced ability to grow at higher concentrations of ampicillin (8 µg/mL) under the induced condition compared to the negative control genes or the uninduced conditions (**Figure 4**, **Supplementary Figure 7**). This was true for both crRNAs used for *ampC* and *marA*, suggesting that precise targeting upstream of a gene of interest was not necessary for its activation. Overall, the results show that CAST can induce novel phenotypes through activation of endogenous genes.

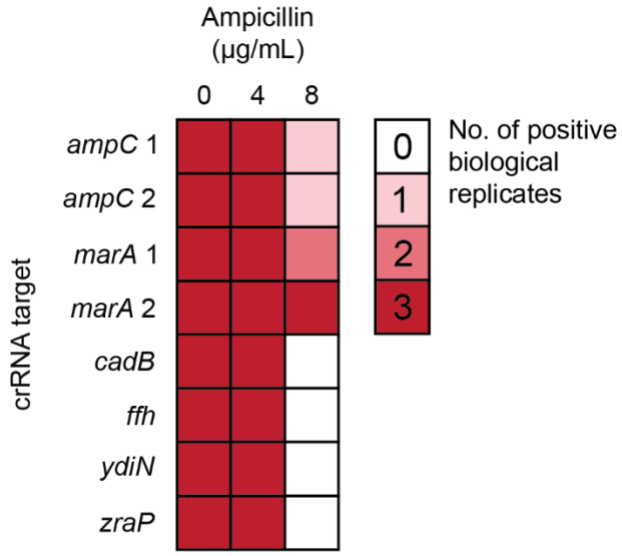

Figure 4. Gain of function via CAST transcriptional activation. The number of positive biological replicates in which growth was observed for different crRNA designs under variable ampicillin concentrations. Full data is shown in **Supplementary Fig. 7.** For each biological replicate, three technical replicates were performed. For a biological replicate to be considered positive, growth had to be observed in >2 technical replicates; otherwise, they were scored as having no growth.

Discussion.

In this work, we demonstrated the use of the *V. cholerae* CAST system for transcriptional activation. We first showed activation of an mRFP1 target gene within a population and then tuned the level of expression using variable strength and inducible promoter elements. We reported high level of editing and activation efficiency, which we believe could be further increased through further optimizing of experimental conditions.[33] Our work also showed robust activation of endogenous *E. coli* genes through the insertion of strong constitutive promoters; Because the inserted promoter was stronger than the native promoters, the genes chosen all showed a significant increase in expression. Finally, we activated the *ampC* and *marA* genes in a WT *E. coli* to induce an ampicillin-resistant phenotype, showcasing the potential for CAST activation to be used as a method for studying genotype-phenotype interactions. Overall, we present a novel CRISPR-based strategy for targeted gene activation.

Within the field of microbiology and synthetic biology, the perturbation of endogenous gene expression is a broadly useful capability. We anticipate that the CAST system will be particularly well-suited for genome-wide expression perturbation studies, as now both knock-outs,[29–31] and knock-ups of gene expression through the CAST system are shown to be possible. Genome-wide, bidirectional studies have been highly successful in eukaryotes for uncovering gene function through both the repression and overexpression of individual genes.[11] The CAST system is well-suited to perform similar screens in bacteria due to several features. For one, the ability to use the same transposon for different insertion sites means only the crRNA needs to be varied, making it more suitable for library generation through pooled DNA synthesis. Furthermore, the *V. cholerae* CAST has a highly promiscuous protospacer-adjacent motif (PAM) preference, with a general bias for a 5'–CN–3' PAM but high flexibility to other sequences, expanding the possible targetable sites across a genome.[36] These characteristics distinguish the CAST system from other CRISPR-based technologies, making it an attractive activation strategy for high-throughput gene editing.

While this work has focused on studies in *E. coli*, we anticipate it is possible to expand this method to other microbes. For example, the VchCAST system has been adapted for use in both gram-negative and -positive species, including biotechnologically relevant species such as *Lactococcus lactis*[37]*, Corynebacterium glutamicum,* and *Bacillus subtilis.*[38] This CAST system has also been used for genome editing of non-model species within soil and gut microbial communities.[31] Additionally, bioinformatic analysis of genomes has uncovered a diversity of CAST systems that could be adapted for our approach. This includes the commonly used *Scytonema hofmanni* type V-K CAST,[39] a large variety of type I-F CAST systems,[34,40,41] as well as some type I-B and I-D CAST systems.[42–44] These and other systems can provide alternatives with different PAM preferences, transposition mechanisms, and insertion behaviors, expanding the possibilities for applying CAST systems as transcriptional perturbators.

In summary, we present a novel approach for gene activation that leverages the unique RNA-guided transposition of CAST systems. We anticipate that our newly described tool is likely to assist efforts to both fundamentally understand and engineer microbes in the coming years.

316 **Materials and Methods.**

317 **Plasmid assembly and strain engineering**

318 All plasmids and strains used in this study can be found in **Supplementary Tables 1** and
319 **2**, respectively, and example annotated plasmids and strains can be found in
320 **Supplementary Table 3**. Plasmids containing the CAST system were derived from
321 pSL1142 (Addgene plasmid #160730).[33] All experiments were performed in wild-type
322 (WT) *E. coli* str. K-12 substr. MG1655 or engineered derivatives. The sfGFP-mRFP1
323 reporter strain (sJEC042), and the versions of it with a promoter insert facing sfGFP
324 (sJEC050) or mRFP1 (sJEC043), were created using CRISPR-Cas9 genome editing.[22]

325

326 **Genome editing using CAST transposition**

327 All crRNAs used in this study can be found in **Supplementary Table 4**. Plasmids were
328 transformed into chemically competent *E. coli* cells, plated on LB-Agar with appropriate
329 antibiotics (LB-agar-Ab), and incubated at 37°C for 24 hours. The plates were then
330 resuspended in 3 mL LB media with antibiotics (LB-Ab) and 100 µL of the suspension
331 was used to inoculate 10 mL of fresh LB-Ab. These cultures were then incubated at 30°C
332 shaking at 200-255 rpm for 24 hours. For experiments using variable strength promoters,
333 the plate suspension was used to inoculate 10 mL of fresh LB-Ab to an optical density at
334 600 nm (OD) of ~0.05 and incubated at 37°C shaking at 200-255 rpm for 24 hours. A
335 similar protocol was followed for creating the sfGFP-mRFP1 reporter strains with the PTet
336 inducible promoter transposon (sJEC057), and the WT strains with the PTet inducible
337 promoter transposon for ampicillin experiments (pJEC1258-65) or the strong constitutive
338 promoter transposon for RT-qPCR analysis (sJEC051-4). The post-incubation
339 populations were plated on LB-agar-Ab and incubated overnight at 37°C to obtain single
340 colonies. For experiments using sJEC051-4 and sJEC057, a clonal strain containing the
341 desired insert in the genome was identified by PCR and confirmed by sequencing, and
342 then used for fluorescence experiments or RT-qPCR analysis. For experiments using
343 pJEC1258-65, colonies were randomly selected and used for the ampicillin resistance
344 experiments.

345

Flow cytometry measurement and analysis

Flow cytometry was performed on the Sony Biotechnology SH800 Cell Sorter. A total of 100,000 events were measured for each biological replicate. FCS files were analyzed using the FlowJo software. Each condition was gated to contain >90% of the measured events, and this sub-population was used for analysis. Fluorescence compensation was performed using samples containing the sfGFP-mRFP1 reporter strain (sJEC042) with a non-targeting plasmid as a negative control and the sfGFP-mRFP1 reporter strain edited to contain the cargo with the constitutive promoter driving either mRFP1 (sJEC043) or sfGFP (sJEC050) expression as positive controls. The quadrant gates used for Fig. 1C-D were manually set at 1,000 arb. units on FlowJo.

Characterization of insertion distance by amplicon sequencing

Fluorescence Assisted Cell Sorting was performed on a Sony Biotechnology SH800 Cell Sorter. A total of 500,000 events per biological replicate were sorted using the gates shown in **Supplementary Figure 2** into 5 mL LB. The cultures were incubated at 37°C shaking at 200-255 rpm for 1 hour, followed by the addition of 15 mL LB-Ab and further incubation at 37°C shaking at 200-255 rpm overnight. The Wizard Genomic DNA Purification Kit (Promega) was used for genome extraction. PCR amplification of the region between the left transposon end and the 5' end of mRFP1 was performed for each of the replicates, column purified, and amplicon sequencing performed (Amplicon-EZ, Genewiz). Primers used can be found in **Supplementary Table 5**. From this data, the distance between the 3' end of the crRNA binding site and the 5' end of the transposon was determined, accounting for the 5 bp duplication that results after insertion.

Bulk fluorescence measurement and analysis

Chemically competent *E. coli* were transformed with plasmids or streaked out from glycerol stocks onto LB-Agar-Ab, and then incubated overnight at 37°C. For each replicate, individual colonies were used to inoculate 300 µL of LB-Ab in a deep 96-well plate and incubated overnight at 37°C shaking at 1,000 rpm. Subsequently, 5 µL of overnight culture was used to inoculate 295 µL of fresh, pre-warmed LB-Ab which was

then incubated at 37°C shaking at 1,000 rpm for 4 hours. For experiments using inducible promoter donors, 5 µL of overnight culture was used to inoculate 290 µL of fresh, pre-warmed LB-Ab which was then incubated at 37°C shaking at 1,000 rpm for 4 hours. The cultures were then induced by adding 5 µL of anhydrotetracycline (Cayman Chem) to final concentrations of 10, 50, 100, 500, and 1,000 ng/mL, and incubated a further 4 hours at 37°C shaking at 1,000 rpm.

For bulk fluorescence measurements, a 96 well plate with 25 µL of culture in 75 µL of phosphate buffer saline (PBS) was used. A media-only control of 25 µL LB and 75 µL PBS was also included. Optical density at 600 nm (OD) was measured, as was fluorescence (FL) for sfGFP (485 excitation, 535 emission) and mRFP1 (560 nm excitation, 590 nm emission). For analysis, the mean OD and FL of the media control was subtracted from each well. The ratio of FL to OD was then measured for each well.

Total RNA extraction

E. coli cells were streaked out from glycerol stocks on LB-agar-Ab and incubated overnight at 37°C. Colonies were used to inoculate LB-Ab and incubated overnight at 37°C shaking at 200-255 rpm. 10 mL of fresh, pre-warmed LB-Ab was inoculated with 100 µL of overnight culture and grown at 37°C shaking at 200-255 rpm to an OD of 0.4-0.6. Total RNA was extracted using the Quick RNA Miniprep Plus Kit (Zymo Research) following the manufacturer's instructions for isolating large RNA species and without in-column DNase treatment. Total RNA was then measured on a Qubit fluorometer (Invitrogen). DNase treatment was performed using Turbo DNase (Invitrogen), treating ~300 ng of total RNA with 4 U of enzyme, and incubating at 37°C for 1 hour. Reaction clean-up was then performed using the Quick RNA Miniprep Plus Kit, again following the manufacturer's instructions and without DNase treatment.

Reverse transcription quantitative PCR (RT-qPCR)

Total RNA samples were measured on a Qubit fluorometer and normalized to a concentration of 0.5 ng/µL. Reverse transcription was performed by mixing 3.5 µL of water, 0.5 µL of 2 µM RT primer, 0.5 µL of 10 mM dNTPs, and 2 µL of 0.5 ng/µL total RNA.

This mixture was incubated at 65°C for 5 min followed by incubation on ice for a further 5 mins. A mixture of 0.5 µL 100 mM Dithiothreitol (DDT), 2 µL 5X First-Strand Buffer (Invitrogen), 0.5 µL 40 U/µL RNase OUT (Invitrogen), 0.5 µL 200 U/µL SuperScript III Reverse Transcriptase (Invitrogen), and the pre-mixture were combined. The reaction was then incubated at 55°C for 1 hour, followed by 70°C for 15 min, and finally stored at -80°C. A CFX Connect Real-Time PCR system (Bio-Rad) was used for qPCR analysis. Reactions were performed using a 96 well plate with an optically clear microseal 'B' film (Bio-Rad). Three technical replicate reactions were performed for each biological replicate. Each reaction contained 0.5 µL of each 2 µM qPCR primer stock, 5 µL of 2X Maxima SYBR Green/ROX qPCR Master Mix (Thermo Scientific), 3 µL of water, and 1 µL of the RT product. To confirm digestion of DNA during the DNase treatment, qPCR analysis was performed directly on the total RNA (i.e., a no-RT control). Additionally, no-template control reaction was also performed using 1 µL of water in place of the RT product. For each qPCR primer set, a standard curve was created using serial dilution of PCR-amplified products. Reaction protocol was as follows: 50°C for 2 min, 95°C for 10 min, and 35 cycles of 95°C for 15 sec, 60°C for 11 min, and measurement of fluorescence. Melting curve analysis was performed after amplification to confirm the absence of primer dimers. RT and qPCR primers used are listed in **Supplementary Table 4**.

For data analysis, CFX Manager Software (Bio-Rad) was used to determine Cq values. To determine the starting quantities (SQ) of each gene of interest, a standard curve was used. In brief, the mean Cq of the technical replicates was determined for each dilution of the gene standard. This mean was fitted to $y = mx + b$, with y being the mean Cq values and x the log SQ. The formula was then used to calculate the SQ of each technical replicate for the test samples, and the technical replicate SQ values were then averaged to obtain a biological replicate SQ.

Ampicillin resistance experiments

For each biological replicate, individual colonies were used to inoculate 300 µL of LB-Ab in a deep 96-well plate and incubated overnight at 37°C shaking at 1,000 rpm. The next day, 5 µL of overnight culture was used to inoculate 295 µL of fresh, pre-warmed LB-Ab

and incubated at 37°C shaking at 1,000 rpm for 4 hours. The OD of each sub-culture was measured as outlined in "bulk fluorescence measurement and analysis," and found to be ~0.6-0.8. A one-way ANOVA was performed on the OD values to confirm equal growth across the conditions. Samples were also taken from each sub-culture for the PCR analysis in **Supplementary Figure 6**. Each biological replicate was then spotted 3 times, for technical replicates, on LB-agar-Ab plates with 0 µg/mL, 4 µg/mL, or 8 µg/mL ampicillin (A9518, Sigma-Aldrich) and with or without 1000 ng/mL anhydrotetracycline, and then incubated for 48 hr at 37°C. For each biological replicate, growth on at least two out of the three technical replicates was counted as a positive hit.

Supporting Information

Additional experimental details and measurements; and plasmids, strains, and primers used in this study (DOCX)

Author Contribution

A.G. and J.C designed the study and experiments, and prepared the manuscript. A.G. collected all experimental data.

Conflict of Interest

The authors declare no competing interests.

Acknowledgements.

The authors acknowledge the Chappell Lab members for helpful discussion. The authors also wish to acknowledge Daniel J. Haller for his assistance in the design and cloning of sJEC057, as well as Biki B. Kundu for his expertise in CRISPR-Cas9 editing that aided the creation of the reporter strains used in this study. This material is based on work supported by the National Science Foundation (grant no. # 2237512).

References.

(1) Palmer, A. C.; Chait, R.; Kishony, R. Nonoptimal Gene Expression Creates Latent Potential for Antibiotic Resistance. *Molecular Biology and Evolution* **2018**. https://doi.org/10.1093/molbev/msy163.

(2) Soo, V. W. C.; Hanson-Manful, P.; Patrick, W. M. Artificial Gene Amplification Reveals an Abundance of Promiscuous Resistance Determinants in Escherichia Coli. *Proceedings of the National Academy of Sciences* **2011**, *108* (4), 1484–1489. https://doi.org/10.1073/pnas.1012108108.

(3) Suarez, S. A.; Martiny, A. C. Gene Amplification Uncovers Large Previously Unrecognized Cryptic Antibiotic Resistance Potential in E. Coli. *Microbiology Spectrum* **2021**. https://doi.org/10.1128/Spectrum.00289-21.

(4) Coward, C.; Dharmalingham, G.; Abdulle, O.; Avis, T.; Beisken, S.; Breidenstein, E.; Carli, N.; Figueiredo, L.; Jones, D.; Khan, N.; Malara, S.; Martins, J.; Nagalingam, N.; Turner, K.; Wain, J.; Williams, D.; Powell, D.; Mason, C. High-Density Transposon Libraries Utilising Outward-Oriented Promoters Identify Mechanisms of Action and Resistance to Antimicrobials. *FEMS Microbiology Letters* **2020**, *367* (22), fnaa185. https://doi.org/10.1093/femsle/fnaa185.

(5) Santiago, M.; Lee, W.; Fayad, A. A.; Coe, K. A.; Rajagopal, M.; Do, T.; Hennessen, F.; Srisuknimit, V.; Müller, R.; Meredith, T. C.; Walker, S. Genome-Wide Mutant Profiling Predicts the Mechanism of a Lipid Ii Binding Antibiotic. *Nat Chem Biol* **2018**, *14* (6), 601–608. https://doi.org/10.1038/s41589-018-0041-4.

(6) Chen, H.; Wilson, J.; Ercanbrack, C.; Smith, H.; Gan, Q.; Fan, C. Genome-Wide Screening of Oxidizing Agent Resistance Genes in Escherichia Coli. *Antioxidants* **2021**, *10* (6), 861. https://doi.org/10.3390/antiox10060861.

(7) Hoegler, K. J.; Hecht, M. H. Artificial Gene Amplification in Escherichia Coli Reveals Numerous Determinants for Resistance to Metal Toxicity. *J Mol Evol* **2018**, *86* (2), 103–110. https://doi.org/10.1007/s00239-018-9830-3.

(8) Mutalik, V. K.; Novichkov, P. S.; Price, M. N.; Owens, T. K.; Callaghan, M.; Carim, S.; Deutschbauer, A. M.; Arkin, A. P. Dual-Barcoded Shotgun Expression Library Sequencing for High-Throughput Characterization of Functional Traits in Bacteria. *Nat Commun* **2019**, *10* (1), 308. https://doi.org/10.1038/s41467-018-08177-8.

(9) Patrick, W. M.; Quandt, E. M.; Swartzlander, D. B.; Matsumura, I. Multicopy Suppression Underpins Metabolic Evolvability. *Molecular Biology and Evolution* **2007**, *24* (12), 2716–2722. https://doi.org/10.1093/molbev/msm204.

(10) Chen, H.; Venkat, S.; Wilson, J.; McGuire, P.; Chang, A. L.; Gan, Q.; Fan, C. Genome-Wide Quantification of the Effect of Gene Overexpression on Escherichia Coli Growth. *Genes* **2018**, *9* (8), 414. https://doi.org/10.3390/genes9080414.

(11) Kampmann, M. CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine. *ACS Chem. Biol.* **2018**, *13* (2), 406–416. https://doi.org/10.1021/acschembio.7b00657.

(12) Li, Y.; Lin, Z.; Huang, C.; Zhang, Y.; Wang, Z.; Tang, Y.; Chen, T.; Zhao, X. Metabolic Engineering of Escherichia Coli Using Crispr–Cas9 Meditated Genome Editing. *Metabolic Engineering* **2015**, *31*, 13–21. https://doi.org/10.1016/j.ymben.2015.06.006.

(13) Zhang, M. M.; Wong, F. T.; Wang, Y.; Luo, S.; Lim, Y. H.; Heng, E.; Yeo, W. L.; Cobb, R. E.; Enghiad, B.; Ang, E. L.; Zhao, H. CRISPR–Cas9 Strategy for Activation of Silent Streptomyces Biosynthetic Gene Clusters. *Nat. Chem. Biol.* **2017**, *13* (6), 607–609. https://doi.org/10.1038/nchembio.2341.

(14) Liu, Y.; Ren, C.-Y.; Wei, W.-P.; You, D.; Yin, B.-C.; Ye, B.-C. A CRISPR-Cas9 Strategy for Activating the Saccharopolyspora Erythraea Erythromycin Biosynthetic Gene Cluster with Knock-in Bidirectional Promoters. *ACS Synth. Biol.* **2019**, *8* (5), 1134–1143. https://doi.org/10.1021/acssynbio.9b00024.

(15) Bikard, D.; Jiang, W.; Samai, P.; Hochschild, A.; Zhang, F.; Marraffini, L. A. Programmable Repression and Activation of Bacterial Gene Expression Using an Engineered CRISPR-Cas System. *Nucleic Acids Research* **2013**, *41* (15), 7429–7437. https://doi.org/10.1093/nar/gkt520.

(16) Dong, C.; Fontana, J.; Patel, A.; Carothers, J. M.; Zalatan, J. G. Synthetic CRISPR-Cas Gene Activators for Transcriptional Reprogramming in Bacteria. *Nat Commun* **2018**, *9* (1), 2489. https://doi.org/10.1038/s41467-018-04901-6.

(17) Ho, H.-I.; Fang, J. R.; Cheung, J.; Wang, H. H. Programmable CRISPR-Cas Transcriptional Activation in Bacteria. *Mol Syst Biol* **2020**, *16* (7), e9427. https://doi.org/10.15252/msb.20199427.
(18) Villegas Kcam, M. C.; Tsong, A. J.; Chappell, J. Rational Engineering of a Modular Bacterial CRISPR–Cas Activation Platform with Expanded Target Range. *Nucleic Acids Research* **2021**, gkab211. https://doi.org/10.1093/nar/gkab211.
(19) Kiattisewee, C.; Karanjia, A. V.; Legut, M.; Daniloski, Z.; Koplik, S. E.; Nelson, J.; Kleinstiver, B. P.; Sanjana, N. E.; Carothers, J. M.; Zalatan, J. G. Expanding the Scope of Bacterial CRISPR Activation with PAM-Flexible dCas9 Variants. *ACS Synthetic Biology* **2022**. https://doi.org/10.1021/acssynbio.2c00405.
(20) Lu, Z.; Yang, S.; Yuan, X.; Shi, Y.; Ouyang, L.; Jiang, S.; Yi, L.; Zhang, G. CRISPR-Assisted Multi-Dimensional Regulation for Fine-Tuning Gene Expression in Bacillus Subtilis. *Nucleic Acids Res* **2019**, *47* (7), e40. https://doi.org/10.1093/nar/gkz072.
(21) Otoupal, P. B.; Erickson, K. E.; Escalas-Bordoy, A.; Chatterjee, A. CRISPR Perturbation of Gene Expression Alters Bacterial Fitness under Stress and Reveals Underlying Epistatic Constraints. *ACS Synth. Biol.* **2017**, *6* (1), 94–107. https://doi.org/10.1021/acssynbio.6b00050.
(22) Jiang, Y.; Chen, B.; Duan, C.; Sun, B.; Yang, J.; Yang, S. Multigene Editing in the Escherichia Coli Genome via the CRISPR-Cas9 System. *Appl. Environ. Microbiol.* **2015**, *81* (7), 2506–2514. https://doi.org/10.1128/AEM.04023-14.
(23) Call, S. N.; Andrews, L. B. CRISPR-Based Approaches for Gene Regulation in Non-Model Bacteria. *Front. Genome Ed.* **2022**, *4*, 892304. https://doi.org/10.3389/fgeed.2022.892304.
(24) Fontana, J.; Dong, C.; Kiattisewee, C.; Chavali, V. P.; Tickman, B. I.; Carothers, J. M.; Zalatan, J. G. Effective CRISPRa-Mediated Control of Gene Expression in Bacteria Must Overcome Strict Target Site Requirements. *Nat Commun* **2020**, *11* (1), 1618. https://doi.org/10.1038/s41467-020-15454-y.
(25) Kiattisewee, C.; Dong, C.; Fontana, J.; Sugianto, W.; Peralta-Yahya, P.; Carothers, J. M.; Zalatan, J. G. Portable Bacterial Crispr Transcriptional Activation Enables Metabolic Engineering in Pseudomonas Putida. *Metabolic Engineering* **2021**, *66*, 283–295. https://doi.org/10.1016/j.ymben.2021.04.002.
(26) Liu, Y.; Wan, X.; Wang, B. Engineered Crispra Enables Programmable Eukaryote-Like Gene Activation in Bacteria. *Nat Commun* **2019**, *10* (1), 3693. https://doi.org/10.1038/s41467-019-11479-0.
(27) Klompe, S. E.; Vo, P. L. H.; Halpin-Healy, T. S.; Sternberg, S. H. Transposon-Encoded CRISPR–Cas Systems Direct RNA-Guided DNA Integration. *Nature* **2019**, *571* (7764), 219–225. https://doi.org/10.1038/s41586-019-1323-z.
(28) Hoffmann, F. T.; Kim, M.; Beh, L. Y.; Wang, J.; Vo, P. L. H.; Gelsinger, D. R.; George, J. T.; Acree, C.; Mohabir, J. T.; Fernández, I. S.; Sternberg, S. H. Selective TnsC Recruitment Enhances the Fidelity of RNA-Guided Transposition. *Nature* **2022**, *609* (7926), 384–393. https://doi.org/10.1038/s41586-022-05059-4.
(29) Aliu, E.; Lee, K.; Wang, K. CRISPR RNA-Guided Integrase Enables High-Efficiency Targeted Genome Engineering in <i>Agrobacterium Tumefaciens</I>. *Plant Biotechnology Journal* **2022**, *20* (10), 1916–1927. https://doi.org/10.1111/pbi.13872.
(30) Zhang, Y.; Yang, J.; Yang, S.; Zhang, J.; Chen, J.; Tao, R.; Jiang, Y.; Yang, J.; Yang, S. Programming Cells by Multicopy Chromosomal Integration Using CRISPR-Associated Transposases. *The CRISPR Journal* **2021**, *4* (3), 350–359. https://doi.org/10.1089/crispr.2021.0018.
(31) Rubin, B. E.; Diamond, S.; Cress, B. F.; Crits-Christoph, A.; Lou, Y. C.; Borges, A. L.; Shivram, H.; He, C.; Xu, M.; Zhou, Z.; Smith, S. J.; Rovinsky, R.; Smock, D. C. J.; Tang, K.; Owens, T. K.; Krishnappa, N.; Sachdeva, R.; Barrangou, R.; Deutschbauer, A. M.; Banfield, J. F.; Doudna, J. A. Species- and Site-Specific Genome Editing in Complex Bacterial Communities. *Nat Microbiol* **2021**, *7* (1), 34–47. https://doi.org/10.1038/s41564-021-01014-7.
(32) Zhang, Y.; Sun, X.; Wang, Q.; Xu, J.; Dong, F.; Yang, S.; Yang, J.; Zhang, Z.; Qian, Y.; Chen, J.; Zhang, J.; Liu, Y.; Tao, R.; Jiang, Y.; Yang, J.; Yang, S. Multicopy Chromosomal Integration Using CRISPR-Associated Transposases. *ACS Synth. Biol.* **2020**, *9* (8), 1998–2008. https://doi.org/10.1021/acssynbio.0c00073.
(33) Vo, P. L. H.; Ronda, C.; Klompe, S. E.; Chen, E. E.; Acree, C.; Wang, H. H.; Sternberg, S. H. CRISPR RNA-Guided Integrases for High-Efficiency, Multiplexed Bacterial Genome Engineering. *Nat Biotechnol* **2020**. https://doi.org/10.1038/s41587-020-00745-y.

(34) Klompe, S. E.; Jaber, N.; Beh, L. Y.; Mohabir, J. T.; Bernheim, A.; Sternberg, S. H. Evolutionary and Mechanistic Diversity of Type I-F Crispr-Associated Transposons. *Molecular Cell* **2022**, *82* (3), 616-628.e5. https://doi.org/10.1016/j.molcel.2021.12.021.

(35) Keseler, I. M.; Gama-Castro, S.; Mackie, A.; Billington, R.; Bonavides-Martínez, C.; Caspi, R.; Kothari, A.; Krummenacker, M.; Midford, P. E.; Muñiz-Rascado, L.; Ong, W. K.; Paley, S.; Santos-Zavaleta, A.; Subhraveti, P.; Tierrafría, V. H.; Wolfe, A. J.; Collado-Vides, J.; Paulsen, I. T.; Karp, P. D. The EcoCyc Database in 2021. *Front Microbiol* **2021**, *12*, 711077. https://doi.org/10.3389/fmicb.2021.711077.

(36) Wimmer, F.; Mougiakos, I.; Englert, F.; Beisel, C. L. Rapid Cell-Free Characterization of Multi-Subunit Crispr Effectors and Transposons. *Molecular Cell* **2022**, *82* (6), 1210-1224.e6. https://doi.org/10.1016/j.molcel.2022.01.026.

(37) Pechenov, P. Y.; Garagulya, D. A.; Stanovov, D. S.; Letarov, A. V. New Effective Method of Lactococcus Genome Editing Using Guide RNA-Directed Transposition. *IJMS* **2022**, *23* (22), 13978. https://doi.org/10.3390/ijms232213978.

(38) Cleto, S.; Jensen, J. V.; Wendisch, V. F.; Lu, T. K. Corynebacterium Glutamicum Metabolic Engineering with CRISPR Interference (CRISPRi). *ACS Synth. Biol.* **2016**, *5* (5), 375–385. https://doi.org/10.1021/acssynbio.5b00216.

(39) Strecker, J.; Ladha, A.; Gardner, Z.; Schmid-Burgk, J. L.; Makarova, K. S.; Koonin, E. V.; Zhang, F. RNA-Guided DNA Insertion with CRISPR-Associated Transposases. *Science* **2019**, *365* (6448), 48–53. https://doi.org/10.1126/science.aax9181.

(40) Roberts, A.; Nethery, M. A.; Barrangou, R. Functional Characterization of Diverse Type I-F Crispr-Associated Transposons. *Nucleic Acids Research* **2022**, *50* (20), 11670–11681. https://doi.org/10.1093/nar/gkac985.

(41) Yang, S.; Zhang, Y.; Xu, J.; Zhang, J.; Zhang, J.; Yang, J.; Jiang, Y.; Yang, S. Orthogonal CRISPR-Associated Transposases for Parallel and Multiplexed Chromosomal Integration. *Nucleic Acids Research* **2021**, No. gkab752. https://doi.org/10.1093/nar/gkab752.

(42) Wang, S.; Gabel, C.; Siddique, R.; Klose, T.; Chang, L. Molecular Mechanism for Tn7-Like Transposon Recruitment by a Type I-B Crispr Effector. *Cell* **2023**. https://doi.org/10.1016/j.cell.2023.07.010.

(43) Hsieh, S.-C.; Peters, J. E. Discovery and Characterization of Novel Type I-D CRISPR-Guided Transposons Identified Among Diverse Tn7-Like Elements in Cyanobacteria. *Nucleic Acids Research* **2023**, *51* (2), 765–782. https://doi.org/10.1093/nar/gkac1216.

(44) Faure, G.; Saito, M.; Benler, S.; Peng, I.; Wolf, Y. I.; Strecker, J.; Altae-Tran, H.; Neumann, E.; Li, D.; Makarova, K. S.; Macrae, R. K.; Koonin, E. V.; Zhang, F. Modularity and Diversity of Target Selectors in Tn7 Transposons. *Molecular Cell* **2023**, *83* (12), 2122-2136.e10. https://doi.org/10.1016/j.molcel.2023.05.013.