nature methods

ARTICLES

https://doi.org/10.1038/s41592-020-00980-w



# A compact Cascade–Cas3 system for targeted genome engineering

Bálint Csörgő[1,2,3,9], Lina M. León[1,3,9], Ilea J. Chau-Ly[4], Alejandro Vasquez-Rifo[5], Joel D. Berry[1], Caroline Mahendra[1], Emily D. Crawford[1,6], Jennifer D. Lewis[4,7] and Joseph Bondy-Denomy[1,3,8] ✉

**CRISPR–Cas technologies have enabled programmable gene editing in eukaryotes and prokaryotes. However, the leading Cas9 and Cas12a enzymes are limited in their ability to make large deletions. Here, we used the processive nuclease Cas3, together with a minimal Type I-C Cascade-based system for targeted genome engineering in bacteria. DNA cleavage guided by a single CRISPR RNA generated large deletions (7–424 kilobases) in *Pseudomonas aeruginosa* with near-100% efficiency, while Cas9 yielded small deletions and point mutations. Cas3 generated bidirectional deletions originating from the programmed site, which was exploited to reduce the *P. aeruginosa* genome by 837 kb (13.5%). Large deletion boundaries were efficiently specified by a homology-directed repair template during editing with Cascade–Cas3, but not Cas9. A transferable 'all-in-one' vector was functional in *Escherichia coli*, *Pseudomonas syringae* and *Klebsiella pneumoniae*, and endogenous CRISPR–Cas use was enhanced with an 'anti-anti-CRISPR' strategy. *P. aeruginosa* Type I-C Cascade–Cas3 (*Pae*Cas3c) facilitates rapid strain manipulation with applications in synthetic biology, genome minimization and the removal of large genomic regions.**

CRISPR–Cas systems are a diverse group of RNA-guided nucleases[1] that defend prokaryotes against viral invaders[2,3]. Gene-editing applications have focused on single subunit Class 2 CRISPR systems[4] (for example, Cas9 and Cas12a), but Class 1 systems hold great potential for editing technologies, despite consisting of multi-subunit complexes[5,6]. The signature gene in Class 1 Type I systems is Cas3, a 3′–5′ single-strand DNA helicase-nuclease enzyme that, unlike Cas9 or Cas12a, degrades target DNA processively[7–10]. This property of Cascade (CRISPR-associated complex for antiviral defense)–Cas3 systems raises the possibility of its development as a tool for large genomic deletions, such as the targeted removal of entire genes, gene clusters, islands, prophages or plasmids, a task that Class 2 systems are inefficient at.

Type I systems are the most prevalent CRISPR–Cas systems in nature[1], which has enabled the use of endogenous CRISPR–Cas3 systems for genetic manipulation via self-targeting. This has been accomplished in *Pectobacterium atrosepticum* (Type I-F)[11], *E. coli* (Type I-E)[12,13], *Sulfolobus islandicus* (Type I-A)[14], *Clostridium* species (Type I-B)[15] *Lactobacillus crispatus* (Type I-E)[16], *Serratia* sp. (Type I-F)[17], *Haloarcula hispanica* (Type I-B)[18], *Streptococcus thermophilus* (Type I-E)[19], *P. aeruginosa* (Type I-F)[20] and *Zymomonas mobilis* (Type I-F)[21], most frequently being used to generate small deletions with homologous repair templates. This multitude of different Type I systems have been shown to work to various degrees as editing tools in their native hosts, however, no Type I system has been optimized for efficient heterologous editing in bacteria, beyond the demonstration of the toxic effects of self-targeting[22]. Recent studies have repurposed Type I-E[23–25] systems for DNA cleavage in human cells, and Type I-F, I-E and I-B systems for transcriptional modulation[26–28].

Here, we repurposed and optimized a Type I-C CRISPR system from *P. aeruginosa* (*Pae*Cas3c) for both endogenous and heterologous genome engineering in four microbial species. Compared to other Type I systems, such as the well studied Type I-E system (six different proteins), Type I-C is streamlined, requiring only four proteins. By targeting the genome with a single CRISPR RNA (crRNA) and selecting only for survival after editing, *Pae*Cas3c is a rapid, counter-selection-free approach to programmable large-scale genome engineering and genome minimization. Cascade–Cas3 is capable of efficient genome-scale deletions currently not achievable using other methodologies. It has the potential to serve as a powerful tool for basic research, discovery and strain optimization.

## Results

**Implementation and optimization of genome editing with CRISPR–Cas3.** Type I-C CRISPR–Cas systems use three *cas* genes (*cas5*, *cas8* and *cas7*) to produce the crRNA-guided Cascade surveillance complex[29,30] that can recruit Cas3 (Fig. 1a). A previously constructed[31] *P. aeruginosa* PAO1 strain (PAO1$^{IC}$) with inducible *cas* genes (*cas5-8-7-3*) and plasmid-expressed crRNAs targeting the genome was used to conduct genome manipulation (see Extended Data Fig. 1a for a comparison to other previously identified I-C systems). Introduction of crRNA-expressing plasmids under noninducing conditions was not noticeably toxic, indicating tight regulation of the constructs (see Methods). Induction of genome-targeting crRNAs caused a transient growth delay (Fig. 1b), but survivors were isolated after extended growth. By targeting *phzM*, a gene required for production of a blue-green pigment (pyocyanin), we observed yellow cultures (Fig. 1c) derived from 10/18 and 6/18 surviving colonies, from two independent *phzM*-targeting crRNAs. PCR of genomic DNA confirmed that the yellow cultures had lost this

[1]Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, CA, USA. [2]Genome Biology Unit, European Molecular Biology Laboratory, Heidelberg, Germany. [3]Innovative Genomics Institute, University of California, Berkeley, Berkeley, CA, USA. [4]Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, CA, USA. [5]Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, MA, USA. [6]Chan-Zuckerberg Biohub, San Francisco, CA, USA. [7]Plant Gene Expression Center, United States Department of Agriculture, Albany, CA, USA. [8]Quantitative Biosciences Institute, University of California, San Francisco, San Francisco, CA, USA. [9]These authors contributed equally: Bálint Csörgő, Lina M. León. ✉e-mail: Joseph.Bondy-Denomy@ucsf.edu

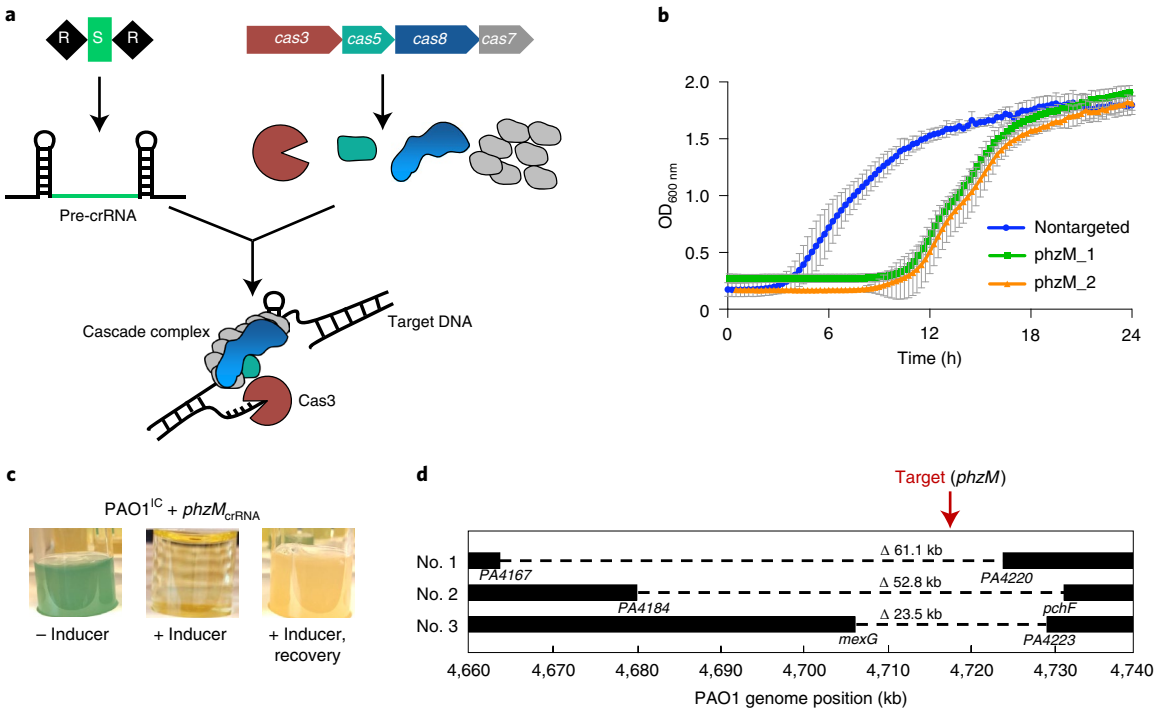



**Fig. 1 | Type I-C CRISPR-mediated self-targeting leads to genomic deletions. a**, A schematic of the Type I-C *cas* gene operon and CRISPR array. The surveillance complex is made up of Cas proteins ($Cas5_1$:$Cas8_1$:$Cas7_7$) and one crRNA, which recruits Cas3 on target DNA recognition. Cas3 then degrades DNA through its dual helicase-nuclease activity. **b**, Growth curves of two $PAO1^{IC}$ strains expressing different crRNAs targeting *phzM* (green and orange) compared to a nontargeting strain (blue). Values are the mean of eight biological replicates each, error bars indicate s.d. values. **c**, Cultures resulting from *phzM* targeting, in the absence of inducer (−ind), presence (+ind) and after recovery. **d**, WGS of three $PAO1^{IC}$ self-targeted survivor strains. Bars indicate boundaries of deletions with open reading frame (ORF) indicated below; red arrow indicates genomic position of targeted sequences.

region, while blue-green survivors maintained it (Extended Data Fig. 1b). Three of these deletion strains were sequenced, revealing deletions of 23.5, 52.8 and 60.1 kb, and each one was bidirectional relative to the crRNA target site (Fig. 1d). This demonstrated the ability of Type I-C Cascade–Cas3 to induce large genomic deletions surrounding a programmed target site.

To determine the in vivo processivity of the Cas3 enzyme, we identified 16 extended nonessential (XNES) regions >100 kb in length (Supplementary Table 1) that lack a known essential gene[32]. Targeting XNES1 and XNES2 (along with additional targeting of *phzM*, which is found in XNES15) with two crRNAs each, led to deletions in 20–40% of the surviving colonies (Fig. 2a). To understand how cells lacking large deletions had survived self-targeting, three possibilities were considered: (1) *cas* gene loss-of-function mutations, (2) PAM or protospacer mutations or (3) mutations in the crRNA-expressing plasmid. Three survivors lacking target deletions from each of the six self-targeting crRNAs were assayed. All had functional *cas* genes when the self-targeting crRNA was replaced with a D3 phage-targeting crRNA (Extended Data Fig. 2a), leading to a reduction in phage replication. Additionally, target sequencing revealed no point mutations. We did, however, observe spacer loss from the crRNA-expressing plasmid, via recombination between the direct repeats (Extended Data Fig. 2b). An additional 17 survivors that lacked target deletions also had plasmids that were missing the spacer (Extended Data Fig. 2c,d),

Spacer excision was successfully prevented by modifying the second repeat, introducing six mutated nucleotides in the stem and three in the loop (Fig. 2b) and thus disrupting homology between the two direct repeats. A phage-targeting crRNA with this new design targeted phage as well or better than the same crRNA with unmodified repeats (Extended Data Fig. 3a). Targeting of the same six genomic sites with modified repeat crRNAs resulted in consistent growth delays (Fig. 2c) and in a robust increase in editing efficiencies to 94–100% (Fig. 2a). Of 216 surviving colonies assayed with deletions generated by the six different crRNAs, 211 (98%) had large deletions (that is, >1 kb), while the remaining 5 had inactive CRISPR–Cas systems when tested with the phage-targeting crRNA (Extended Data Fig. 3b).

When targeting poorly characterized genomes, essential genes may be unknowingly targeted leading to confounding editing outcomes. To assess the phenotype of such an event, we intentionally targeted an essential gene, *rplQ*. Two different modified repeat crRNAs targeting *rplQ* led to a severely extended lag time compared to nonessential gene targeting. Only 8 out of 36 *rplQ*-targeting biological replicates grew after 24 h (Extended Data Fig. 4a). Subsequent analysis of these eight survivor cultures with phage-targeting assays revealed nonfunctional *cas* genes (Extended Data Fig. 4b). No spacer excision events were detected in this experiment, confirming the robustness of the crRNA-engineering and of the deletion method, as the outcome of essential gene versus nonessential gene targeting is noticeably distinct.

**Cas3 generates larger deletions than Cas9 and is recombinogenic.** To determine whether large deletions are a direct consequence of the processive Cas3 enzyme, as opposed to selecting for rare preexisting large genomic deletions[33], we compared Cas3-mediated self-targeting outcomes to those resulting from targeting using *Streptococcus pyogenes* Cas9, which lacks a helicase domain, expressed in an isogenic strain ($PAO1^{IIA}$). If the large deletions we observed in *P. aeruginosa* are preexisting in the population, we would expect these to be selected for at comparable frequencies regardless of the nuclease used to target the genome. A Cas9

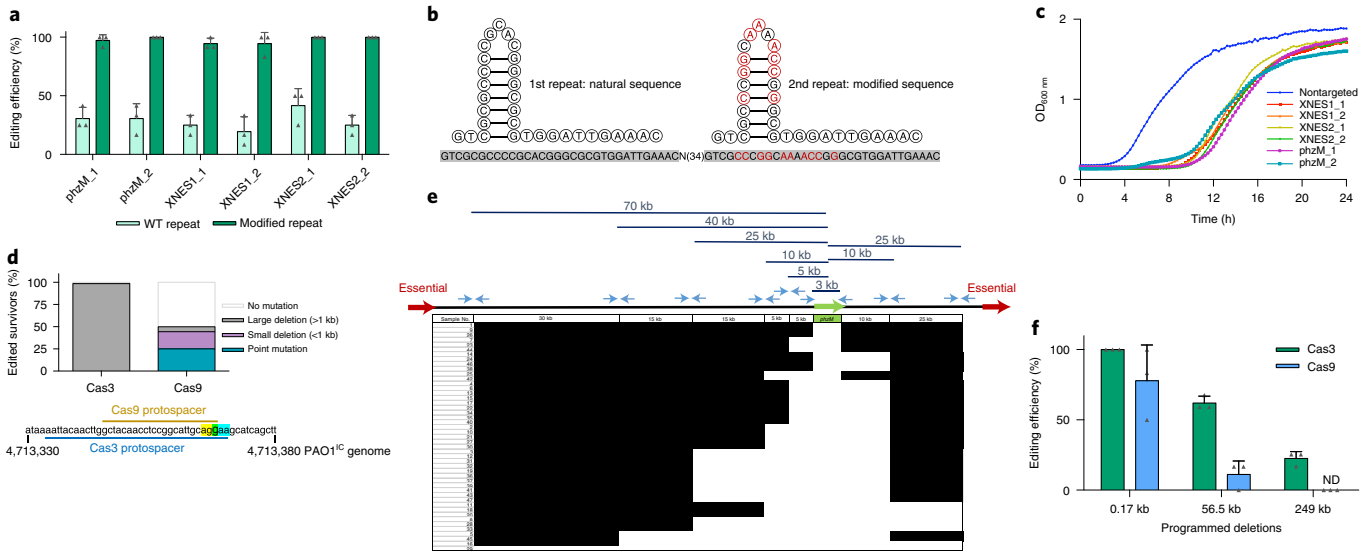



**Fig. 2 | Optimization and characterization of Cascade–Cas3-directed genomic editing. a**, Percentage of survivors with a genomic deletion at the location targeted. Six different crRNA constructs with either wild-type (WT) repeat sequences (light green) or with the second repeat being modified (dark green). Values are means of three biological replicates each, where 12 individual surviving colonies were assayed per replicate, error bars show s.d. values. **b**, Sequence and structure of natural and modified repeat sequences. Specifically engineered modified nucleotides shown in red; repeat sequences highlighted in gray with an arbitrary intervening spacer sequence. **c**, Growth curves of PAO1$^{IC}$ strains expressing distinct self-targeting crRNAs flanked by modified repeats. Nontargeting crRNA-expressing control is marked in blue. Values depicted are averages of four biological replicates each. **d**, Gene-editing outcomes for distinct survivor cells targeted with either a Type II-A SpyCas9 system or a Type I-C Cas3 system ($n$=72 individual colonies isolated from distinct cultures of self-targeting cells for each system). **e**, Deletion size distribution of 47 independently generated *phzM* deletion strains determined using tiling PCR. Black segments indicate the presence of a PCR product for a given sample, while white segments indicate the absence, thus giving an estimate of the size for the generated deletions. **f**, Percentage of analyzed survivors with the specific deletion size present (0.17, 56.5 or 249kb) generated with homologous repair templates, in tandem with the Cascade–Cas3 system (green) or the SpyCas9 system (blue). Values are means of three biological replicates each, where 12 individual surviving colonies were assayed per replicate and have been normalized to the total percentage of edited samples for each enzyme shown in Fig. 2d (98.6% for Cas3, 50% for Cas9); error bars show s.d. values; ND, not detected.

single-guide RNA (sgRNA) was used with a spacer that overlapped with one of the Cas3 crRNAs used to target *phzM* (Fig. 2d, Extended Data Fig. 5). Sequence analysis of these surviving cells revealed that deletions larger than 1 kb were a rare occurrence in the presence of Cas9 (5.6% assayed survivor cells, $n$=72) compared to 98.6% with Cas3 (Fig. 2d). Whole-genome sequencing (WGS) of two large deletion survivors selected by Cas9 showed lesions of 5 and 23 kb around the target site. The more common modes of survival after Cas9 targeting were small deletions between 0.1 and 0.5 kb in length (25% of all survivors), or 1–3 basepair protospacer/PAM deletions/mutations (19.4%), with the remaining 50% of survivors unedited at their target sites. In sum, the apparent shift of deletions toward smaller size resulting from targeting with SpyCas9, compared to a nearly completely distinct set of outcomes when using Cas3, implicates Cas3's enzymatic activity as the cause of large deletions.

To achieve a more granular measurement of the deletion sizes generated by Cas3-mediated editing, we examined 47 individual *phzM*-targeting events. Tiling PCR was used to determine the presence of flanking segments at various intervals spanning a total of 95 kb surrounding *phzM* (Fig. 2e). Of the 47 independent deletion outcomes examined with this method, 44 had deletions of at least 5 kb and 22 had deletions of at least 35 kb in size, with 1 strain having a deletion larger than 95 kb. The average deletion was larger than 26.6 kb and smaller than 48.2 kb, as based on the resolution of the tiling experiment. This comprehensive assessment confirms both the variability and size of Cas3-induced deletion outcomes.

The processive ssDNAse activity of nuclease-helicase Cas3 led us to hypothesize that it may promote recombination. To test this, we provided a double-stranded DNA repair template with 500 bp of the upstream and downstream regions flanking a desired deletion to enable homology-directed repair (HDR) during targeting. We chose 0.17 and 56.5 kb deletions around *phzM*, to model a gene and prophage deletion, respectively, and a large 249 kb deletion within XNES8 for the programmed deletions (see Supplementary Information). The specific editing efficiencies were significantly higher with Cas3 than with Cas9 (Fig. 2f). The 249 kb deletion was incorporated in 22% of the Cas3-generated survivors, compared to 0% using Cas9 ($\chi^2(1, N=72)=9, P=2.7\times10^{-3}$). The 56.5 kb deletion was present in 61 versus 11% ($\chi^2(1, N=72)=25, P=5.73\times10^{-7}$), and the 0.17 kb deletion in 100 versus 78% of survivors, when targeting with Cas3 or Cas9, respectively ($\chi^2(1, N=72)=31.68, P=1.82\times10^{-8}$). During Cas3 targeting, strains that did not incorporate the HDR-programmed deletion had other deletions of random size. Most of the strains that survived Cas9-sgRNA induction without incorporating the HDR template had no change at the target site (84.7% and 80.6% for the 56.5 and 249 kb deletions, respectively), similar to data reported above (Fig. 2e). We presume that mutation or loss of Cas9 occurs more frequently than loss of the Cas3-based system under this experimental set-up. To account for this, normalizing the editing efficiency by roughly twofold (derived from the frequency of unedited clones in Fig. 2d) revealed that whether one considers the absolute percentage of colonies with the desired edits or the normalized value, Cascade–Cas3 targeting outperforms Cas9 for generating large specific deletions (Fig. 2f).

**Rapid genome minimization of *P. aeruginosa* using CRISPR–Cas3 editing**. Large deletions with undefined boundaries provide an unbiased mechanism for genome streamlining, screening and functional genomics. To demonstrate the potential for Cas3, we aimed to minimize the *P. aeruginosa* genome through a series of

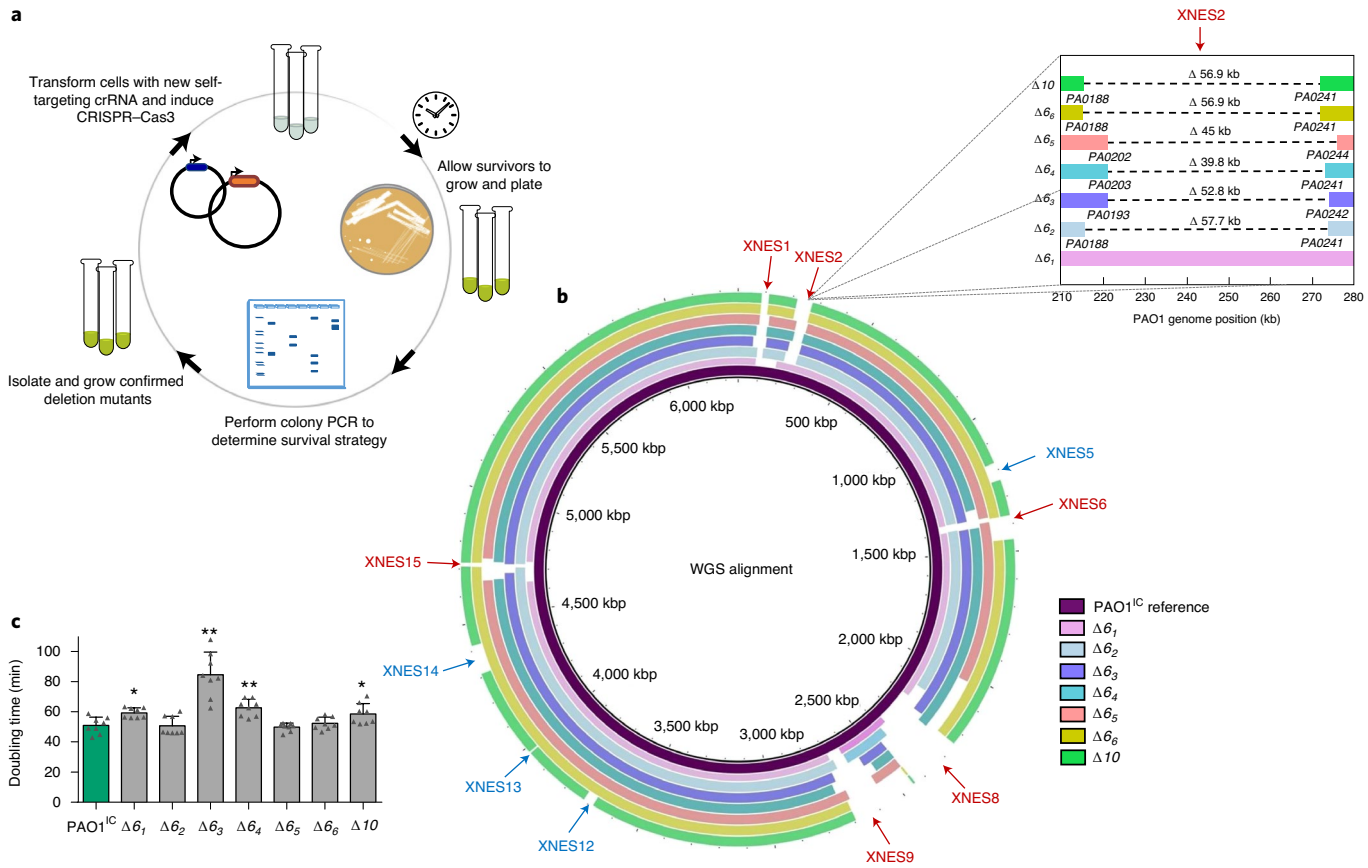



**Fig. 3 | Iterative generation of multiple genomic deletions in *P. aeruginosa*. a**, Schematic overview of the iterative deletion generating process. **b**, Whole-genome sequences of six PAO1$^{IC}$ strains that have been iteratively targeted at six distinct genomic positions and one (derived from strain $\Delta 6_6$) with ten total deletions ($\Delta 10$) aligned to the parental *P. aeruginosa* PAO1$^{IC}$ strain. The first six targeted sites are marked with red arrows and the final four are marked with blue arrows. Inset shows deletion coordinates of XNES2 region of the various strains in finer detail. **c**, Calculated doubling times of the seven genome-reduced strains (strains $\Delta 6_1$–$6_6$ with six deletions, $\Delta 10$ with ten) compared to the parent PAO1$^{IC}$ strain (green). Values are means of eight biological replicates, error bars represent s.d. values, $^*P<0.05$ ($\Delta 6_1$ $P=0.0103$, $\Delta 10$ $P=0.0173$), $^{**}P<0.01$ ($\Delta 6_3$ $P=0.00064$, $\Delta 6_4$ $P=0.00764$), paired two-sided *t*-test compared to PAO1$^{IC}$.

deletions of the XNES regions (Fig. 3a). Six XNES regions were iteratively targeted in six parallel lineages (Fig. 3b), resulting in 35 independent deletions (WGS revealed no deletion at XNES2 in one of the strains). Deletion efficiency remained high (>80%) throughout each round of self-targeting (Extended Data Fig. 6a). WGS of these six multiple deletion strains ($\Delta 6_1$–$\Delta 6_6$) revealed that no two deletions had the exact same coordinates, highlighting the stochastic nature of Cas3. The smallest isolated deletion was 7 kb and the largest 424 kb (mean 92.9 kb, median 58.2 kb). Of note, four genes (PA0123, PA1969, PA2024 and PA2156) previously identified as essential[32] were deleted in at least one of the lineages. Most deletions appeared to be resolved by flanking microhomology regions ranging from 4 to 14 bp in length (Extended Data Fig. 6b and Supplementary Table 2), implicating alternative-end joining[34] as the dominant repair process.

To minimize the genome further, one of the already reduced strains was subjected to four additional rounds of deletions at XNES regions for a total of ten genomic deletions ($\Delta 10$, Fig. 3b). WGS of the $\Delta 10$ strain showed a genome reduction of 849 kb (13.6% of the genome). Generation of large deletions resulted in a growth defect in some cases, with significantly slower growth in three of the six deletion strains ($\Delta 6_1$, $\Delta 6_3$ and $\Delta 6_4$), with the other three growing normally (Fig. 3c). Strain $\Delta 10$ also displayed a slight decrease in fitness, showing a ~15% increase in doubling time compared to the parent strain. Stronger growth defects were likely avoided by the selection of fast-growing colonies at each deletion round. To determine whether any deletions may be preexisting at low frequencies in unedited cells, PCR primers probed for specific deletions at XNES1, 6, 8 and 9, revealing no products (Extended Data Fig. 6c). This again indicates that Cas3 has a direct role in generating large deletions.

**CRISPR–Cas3 editing in distinct bacteria.** To enable expression of this system in other hosts, we constructed an all-in-one vector (pCas3cRh) carrying the I-C specific crRNA with a modified repeat sequence, *cas3*, *cas5*, *cas8* and *cas7*, under rhamnose induction (Extended Data Fig. 7a). As a pilot experiment, we transformed wild-type PAO1 with a nontargeting crRNA and crRNAs targeting *phzM* and *XNES2*. Minimal leaky expression of the system was observed, as transformation efficiency was only slightly reduced comparing targeting and nontargeting constructs. Subsequent induction of the targeting crRNAs resulted in 95–100% of survivors being edited (Extended Data Fig. 7b–d).

Having verified that pCas3cRh was functional, we tested this system in *E. coli* K-12 MG1655. pCas3cRh encoding crRNAs targeting *lacZ* or its vicinity (Fig. 4a) were used to transform cells, which were plated directly on inducing media containing X-gal and scored using blue/white screening. Depending on the crRNA used, directly

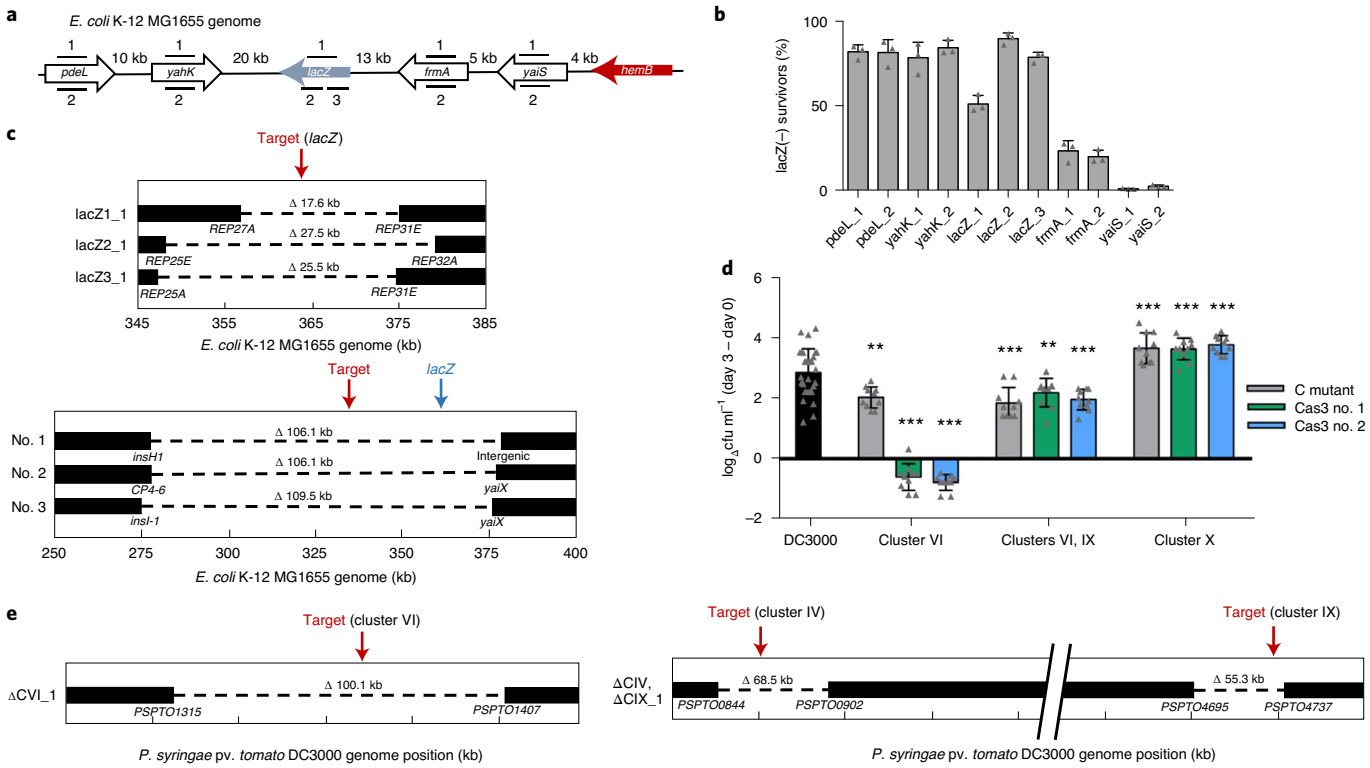
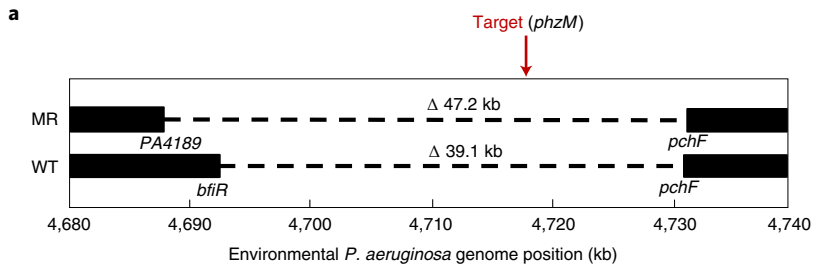
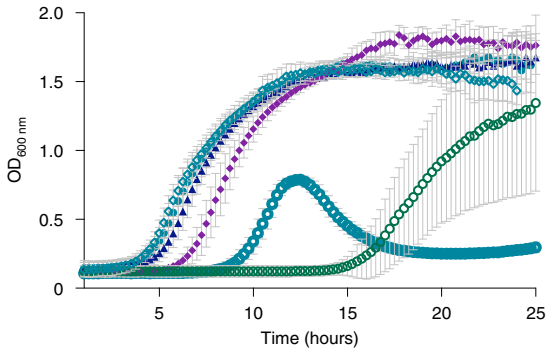



**Fig. 4 | Cascade–Cas3-mediated heterologous editing in various bacteria. a**, Schematic of the crRNA-targeted sites in the *E. coli* MG1655 genome at the *lacZ* locus. **b**, *lacZ* deletion efficiencies using distinct crRNAs targeting the *E. coil* K-12 MG1655 chromosome. Efficiencies calculated based on LacZ activity. Values are averages of three biological replicates, error bars represent standard deviations. **c**, WGS of *E. coli* deletion mutants targeted at *lacZ* and 30kb upstream at *pdeL*. Bars indicate boundaries of deletions with ORF indicated below. **d**, Bacterial growth of deletion mutants in *Arabidopsis thaliana*. Values are differences in colony forming units (cfu)ml$^{-1}$ counted on day 0 of the experiment and day 3, shown on a logarithmic scale. The wild-type DC3000 strain is shown in black, while gray bars represent previously constructed polymutant control (C) strains of the different clusters (labeled at bottom), and green and blue bars show deletion mutants generated using Cas3 (two isolated strains for each targeted cluster, nos. 1 and 2). Values shown are means of ten biological replicates each (30 for DC3000), error bars show s.d. values, $^{**}P<0.01$ (C mutant VI $P=0.00567$, Cas3_1 IV,IX $P=0.00829$) $^{***}P<0.005$, (Cas3_1 VI $P=2.33\times10^{-15}$, Cas3_2 VI $P=9.27\times10^{-18}$, C mutant IV,IX $P=9.95\times10^{-5}$, Cas3_2 IV,IX $P=0.00165$, C mutant X $P=0.00185$, Cas3_1 X $P=0.000858$, Cas3_2 $P=0.000864$), determined based on ANOVA analysis on the day 0 group of values and the day 3 group of values (see Methods). **e**, WGS of *P. syringae* deletion mutants. The left panel shows virulence cluster VI targeting, while the right panel shows virulence cluster IV and IX targeting with a single crRNA, as the clusters share sequence identity. Bars indicate boundaries of deletions with ORF indicated below.

targeting *lacZ* or 30 kb upstream yielded 51–90% and 82–85% LacZ (−) survivors, respectively (Fig. 4b). Of the 96 LacZ (−) survivors, 95 assayed by PCR showed an absence of the *lacZ* region. crRNAs targeting downstream of *lacZ*, however, had reduced efficiency as they approached the essential gene, *hemB*. *frmA* targeting (9.2 kb upstream of *hemB*) resulted in lower editing efficiencies (21–25%) while targeting *yaiS* (4.2 kb upstream of *hemB*) was even lower (2%). This decrease in efficiency was independent of the strand being targeted (and therefore the predicted strand for Cas3 loading and 3′–5′ translocation), confirming the importance of Cas3 bidirectional deletions. WGS of selected Δ*lacZ* cells revealed bidirectional deletions ranging from 17.6 to 109.5 kb encompassing the targeted region (Fig. 4c). Based on these findings, the nearby presence of an essential gene can substantially lower editing efficiency and must be considered when targeting a selected region.

Next, we tested Cas3-mediated editing in the plant pathogen *P. syringae* pv. *tomato* DC3000, which does not naturally encode a CRISPR–Cas system. *P. syringae* encodes many nonessential virulence effector genes whose activities are difficult to disentangle due to their redundancy[35]. We designed crRNAs targeting four chromosomal virulence effector clusters (IV, VI, VIII and IX), or one plasmid cluster (pDC3000, ref. [36]; cluster X). Two clusters (IV and IX) shared identical sequences that could be targeted simultaneously using a single crRNA. Expression of targeting crRNAs caused a growth delay compared to nontargeted controls (Extended Data Fig. 8a) and 67–92% of survivors had deletions (Extended Data Fig. 8b). In planta and in vitro growth assays of three deletion mutants effectively recapitulated the phenotypes of previously described mutants[36] (Fig. 4d and Extended Data Fig. 8c–h). Targeting cluster X cured the 73 kb plasmid (as observed by the absence of plasmid-specific reads in WGS) and simultaneous cluster IV and IX targeting led to dual deletions in eight out of 12 survivors, with a sequenced representative having 68.5 and 55.3 kb deletions, respectively, at the target sites (Fig. 4e). The effector cluster VI Cas3-derived mutant (100.1 kb deletion) had a more severe growth defect in vitro and in planta than the control mutant, likely from a fitness defect owing to the missing genetic material (Fig. 4d,e and Extended Data Fig. 8c,f). In contrast to *P. aeruginosa*, IS elements were present at deletion junctions, suggesting the involvement of homologous recombination between insertion sequence (IS) elements flanking the virulence gene clusters. In such instances, we have not ruled out that the loss of these large regions was not a natural occurrence in the population, as seen in *S. thermophilus*[33]. In two out of three cases, however, the deletions did entail significant fitness costs (Fig. 4d), decreasing this likelihood. Using our portable streamlined system, we achieved three distinct applications in *P. syringae*: plasmid curing (similar to




| | Prophage *acr* | crRNA | *cas*/crRNA induction | *aca1* | Edited survivors |
|---|---|---|---|---|---|
| ○ | AcrIC1 | phzM_1 | +/+ | + | 6/10 |
| ● | AcrIC1 | phzM_1 | +/+ | − | 0/10 |
| ◆ | AcrIC1 | NT | +/+ | − | 0/10 |
| ▲ | AcrIC1 | phzM_1 | ++/++ | − | NA |
| ○ | AcrIIA4 | phzM_1 | +/+ | − | NA |
| ◆ | AcrIIA4 | NT | +/+ | − | NA |

**Fig. 5 | Cascade–Cas3-mediated gene editing in native settings. a**, Schematic of WGS of an environmental isolate of *P. aeruginosa* with an endogenous Type I-C system. Two survivors were isolated post-targeting using either wild-type (WT) direct repeats flanking the spacer or modified repeats (MR). Bars indicate boundaries of deletions with ORF indicated below. **b**, Growth curves of PAO1$^{IC}$ lysogenized by recombinant DMS3m phage expressing *acrIIA4* or *acrIC1* from the native *acr* locus. CRISPR–Cas3 activity is induced with either 0.5mM (+) or 5mM (++) IPTG and 0.1% (+) or 0.3% (++) arabinose. Edited survivors reflect number of isolated survivor colonies missing the targeted gene (*phzM*). NA means that editing was not assessed as no growth defect was seen and NT means a nontargeting crRNA was expressed. Each growth curve is the average of ten biological replicates and error bars represent s.d.

previous observations with a I-E system[37]), single-step deletion of large virulence regions and multiplexed targeting.

Finally, we tested the feasibility of heterologous editing using the I-C system in more distantly related and clinically relevant bacteria *K. pneumoniae*. Using pCas3cRh, *K. pneumoniae* strain KPPR1 was targeted with four distinct crRNAs, with two each targeting *rfaH* and *sacX*, which are flanked by nonessential genes[38]. On induction, all four crRNAs resulted in a substantial growth delay compared to a nontargeted control, indicating functionality (Extended Data Fig. 9a). Individual surviving clones were isolated and 38–63% of survivors had deletions, showing the feasibility of Cas3 editing in *K. pneumoniae* as well ((Extended Data Fig. 9b,c). *rfaH* deletion mutants had a smaller colony morphology, consistent with previous work[38] (Extended Data Fig. 9d). Overall, we have demonstrated portable Type I-C Cascade–Cas3 editing to be a generally applicable tool capable of generating large genomic deletions in four distinct species.

**Repurposing endogenous Cascade–Cas3 systems for gene editing.** Type I CRISPR–Cas3 systems are the most common CRISPR–Cas systems in nature[1]. Therefore, many bacteria have a built-in genome editing tool to be harnessed. We introduced self-targeting *phzM* crRNAs into the environmental isolate (PaLML1) from which our Type I-C system was derived. Genome targeting led to the isolation of 33.7 and 39 kb deletions (Fig. 5a and Extended Data Fig. 10a). Additionally, HDR-based editing with a single construct was again efficacious, with 7/10 survivors acquiring the specific 0.17 kb deletion (Extended Data Fig. 10b).

We next evaluated the feasibility of repurposing other Type I systems, using the naturally active Type I-F systems[39] encoded by laboratory strain *P. aeruginosa* PA14, and the clinical strain *P. aeruginosa* z8. Plasmids with Type I-F specific crRNAs were expressed, targeting various genomic sites for deletion (Supplementary Table 3). HDR templates (600 bp arms on average) were included in the plasmids to generate deletions of defined coordinates ranging from 0.2 to 6.3 kb. Overall, at five different genomic target sites in strain z8 and 2 sites in PA14, we observed desired deletions in 29–100% of analyzed survivor colonies (Supplementary Table 3). Together, these experiments demonstrate the capacity for different forms of high-efficiency genome editing using a single plasmid and an endogenous CRISPR–Cas system.

Finally, one potential impediment to the implementation of any CRISPR–Cas bacterial genome editing tool is the presence of anti-CRISPR (acr) proteins that inactivate CRISPR–Cas activity[40]. In the presence of a prophage expressing AcrIC1 (a Type I-C anti-CRISPR protein[31]) from a native *acr* promoter, targeting was completely inhibited, but not by an isogenic prophage expressing a Cas9 inhibitor AcrIIA4 (ref. [41]) (Fig. 5b). To attempt to overcome this impediment, we expressed *aca1* (anti-CRISPR-associated gene 1), a direct negative regulator of *acr* promoters[42], from the same construct as the crRNA. Using this repression-based 'anti-anti-CRISPR' strategy, CRISPR–Cas function was enabled, allowing the isolation of edited cells despite the presence of *acrIC1* (Fig. 5b and Extended Data Fig. 10c). In contrast, simply increasing *cas* gene and crRNA expression did not overcome AcrIC1-mediated inhibition as assessed by growth kinetics (Fig. 5b). Therefore, using anti-CRISPR repressors to combat anti-CRISPR impediments presents a viable route toward enhanced efficiency of CRISPR–Cas editing and necessitates continued discovery and characterization of anti-CRISPR proteins and their cognate repressors.

## Discussion

By repurposing a minimal Cascade–Cas3 system (referred to as *Pae*Cas3c), large deletions of random or programmed sizes can be obtained with high efficiencies. Using only a single crRNA with modified repeat sequences, we isolated deletions with variable sizes,

one as large as 424 kb, without requiring the insertion of a selectable marker. Notably, the mean (92.9 kb) and median (58.2 kb) deletion sizes are roughly in the range of the average size of *Pseudomonas* bacteriophages (35–100 kb for 92% of sequenced genomes[43]), suggesting that the Cascade–Cas3 machinery can efficiently degrade entire phage genomes. Few studies have directly measured the processivity of the Cas3 enzyme in vivo. Additionally, the I-C system appears to produce bidirectional deletions contrary to unidirectional deletions observed with Type I-E[8,9,23]. Cascade–Cas3 presents a genome editing tool useful for the targeted removal of large elements for genome streamlining. As a long-term goal of microbial gene editing has been genome minimization[44,45], we used our optimized Cascade–Cas3 system to generate ten iterative deletions, achieving >13% genome reduction of the targeted strain. This spanned only 30 d while maintaining editing efficiency, a great improvement over previous genome reduction methods[45]. Some basic microbial applications of Cas3 include studying chromosome biology (for example, replichore asymmetry[46]), pathogenesis, the impact of the mobilome and a better understanding the essential building blocks for life.

An important outcome of this work is the high efficiency of recombination observed at cut sites when comparing Cas3 and Cas9 directly. The potential for Cas3 to be recombinogenic through the generation of exposed ssDNA may be advantageous for both programmed knockouts and knockins. Although knockins were not systematically explored here, a preliminary attempt to affix a chromosomal mCherry tag to *cas3*, using the Type I-C system for selection, was successful. The direct comparison presented here between Cas3 (large deletions) and Cas9 (small deletions), coupled with the high variability of deletions not observed to be preexisting also demonstrates the causality of Cas3 in the deletion outcomes.

Our study revealed benefits and challenges of applying Cascade–Cas3. While the Type I-C system was functional in heterologous hosts, it remains unclear whether the approach will be limited by differences in DNA repair mechanisms. Indeed, in *E. coli* and *P. syringae*, larger regions of homology, such as 34-bp long repetitive extragenic palindromic sequences were observed[47], indicating the role of RecA-mediated homologous recombination in the repair process. Meanwhile in *P. aeruginosa*, the borders of the deletions showed either small (4–14 bp) microhomology or no noticeable sequence homology. The former implies a role for alternative-end joining[34], which has also been observed in *P. atrosepticum*[11], while the latter implicates nonhomologous end joining[48] in the repair process. Downstream studies are required to dissect the roles of each mechanism in the deletion generation process.

CRISPR–Cas3 is an especially promising tool for use in eukaryotic cells as it would facilitate the interrogation of large segments of noncoding DNA, much of which has unknown function. Additionally, it was recently shown that Cas9-generated 'gene knockouts' (that is, small indels causing out-of-frame mutations) frequently encode pseudo-messenger RNAs that may produce protein products, necessitating methods for full gene removal[49,50]. Type I-E CRISPR–Cas systems were recently shown to generate large (up to 100–200 kb) deletions in human cells[23–25], demonstrating the potential wide applicability of Cas3. Overall, the intrinsic properties of Cas3 make it a promising tool to fill a void in current gene-editing capabilities. Using Cas3 to make large genomic deletions will facilitate the manipulation of repetitive and noncoding regions, having a broad impact on genetics research by providing a tool to probe genomes en masse.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41592-020-00980-w.

Received: 1 February 2020; Accepted: 15 September 2020;
Published online: 19 October 2020

## References

1. Makarova, K. S. et al. An updated evolutionary classification of CRISPR–Cas systems. *Nat. Rev. Microbiol.* **13**, 722–736 (2015).
2. Barrangou, R. et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).
3. Garneau, J. E. et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71 (2010).
4. Barrangou, R. & Doudna, J. A. Applications of CRISPR technologies in research and beyond. *Nat. Biotechnol.* **34**, 933–941 (2016).
5. Brouns, S. J. J. et al. Small CRISPR RNAs guide antiviral defense in Prokaryotes. *Science* **321**, 960–964 (2008).
6. Hidalgo-Cantabrana, C. & Barrangou, R. Characterization and applications of Type I CRISPR-Cas systems. *Biochem. Soc. Trans.* **48**, 15–23 (2020).
7. Sinkunas, T. et al. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J.* **30**, 1335–1342 (2011).
8. Sinkunas, T. et al. In vitro reconstitution of Cascade-mediated CRISPR immunity in *Streptococcus thermophilus*. *EMBO J.* **32**, 385–394 (2013).
9. Mulepati, S. & Bailey, S. In vitro reconstitution of an *Escherichia coli* RNA-guided immune system reveals unidirectional, ATP-dependent degradation of DNA target. *J. Biol. Chem.* **288**, 22184–22192 (2013).
10. Redding, S. et al. Surveillance and processing of foreign DNA by the *Escherichia coli* CRISPR-Cas system. *Cell* **163**, 854–865 (2015).
11. Vercoe, R. B. et al. Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. *PLoS Genet.* **9**, e1003454 (2013).
12. Gomaa, A. A. et al. Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. *mBio* **5**, e00928–00913 (2014).
13. Kiro, R., Shitrit, D. & Qimron, U. Efficient engineering of a bacteriophage genome using the type I-E CRISPR-Cas system. *RNA Biol.* **11**, 42–44 (2014).
14. Li, Y. et al. Harnessing Type I and Type III CRISPR-Cas systems for genome editing. *Nucleic Acids Res.* **44**, e34–e34 (2016).
15. Pyne, M. E., Bruder, M. R., Moo-Young, M., Chung, D. A. & Chou, C. P. Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in *Clostridium*. *Sci. Rep.* **6**, 25666 (2016).
16. Hidalgo-Cantabrana, C., Goh, Y. J., Pan, M., Sanozky-Dawes, R. & Barrangou, R. Genome editing using the endogenous type I CRISPR-Cas system in *Lactobacillus crispatus*. *Proc. Natl Acad. Sci. USA* **116**, 15774–15783 (2019).
17. Hampton, H. G. et al. CRISPR-Cas gene-editing reveals RsmA and RsmC act through FlhDC to repress the SdhE flavinylation factor and control motility and prodigiosin production in *Serratia*. *Microbiology* **162**, 1047–1058 (2016).
18. Cheng, F. et al. Harnessing the native type I-B CRISPR-Cas for genome editing in a polyploid archaeon. *J. Genet. Genomics Yi Chuan Xue Bao* **44**, 541–548 (2017).
19. Cañez, C., Selle, K., Goh, Y. J. & Barrangou, R. Outcomes and characterization of chromosomal self-targeting by native CRISPR-Cas systems in *Streptococcus thermophilus*. *FEMS Microbiol. Lett.* **366**, fnz105 (2019).
20. Xu, Z. et al. Native CRISPR-Cas-mediated genome editing enables dissecting and sensitizing clinical multidrug-resistant *P. aeruginosa*. *Cell Rep.* **29**, 1707–1717.e3 (2019).
21. Zheng, Y. et al. Characterization and repurposing of the endogenous Type I-F CRISPR-Cas system of *Zymomonas mobilis* for genome engineering. *Nucleic Acids Res.* **47**, 11461–11475 (2019).
22. Edgar, R. & Qimron, U. The *Escherichia coli* CRISPR system protects from λ Lysogenization, Lysogens, and prophage induction. *J. Bacteriol.* **192**, 6291–6294 (2010).
23. Dolan, A. E. et al. Introducing a spectrum of long-range genomic deletions in human embryonic stem cells using type I CRISPR-Cas. *Mol. Cell* **74**, 936–950.e5 (2019).
24. Morisaka, H. et al. CRISPR–Cas3 induces broad and unidirectional genome editing in human cells. *Nat. Commun.* **10**, 5302 (2019).
25. Cameron, P. et al. Harnessing type I CRISPR–Cas systems for genome engineering in human cells. *Nat. Biotechnol.* **37**, 1471–1477 (2019).
26. Pickar-Oliver, A. et al. Targeted transcriptional modulation with type I CRISPR–Cas systems in human cells. *Nat. Biotechnol.* **37**, 1493–1501 (2019).
27. Chen, Y. et al. Repurposing type I–F CRISPR–Cas system as a transcriptional activation tool in human cells. *Nat. Commun.* **11**, 3136 (2020).
28. Young, J. K. et al. The repurposing of type I-E CRISPR-Cascade for gene activation in plants. *Commun. Biol.* **2**, 383 (2019).
29. Nam, K. H. et al. Cas5d protein processes Pre-crRNA and assembles into a cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system. *Structure* **20**, 1574–1584 (2012).

30. Hochstrasser, M. L., Taylor, D. W., Kornfeld, J. E., Nogales, E. & Doudna, J. A. DNA targeting by a minimal CRISPR RNA-guided cascade. *Mol. Cell* **63**, 840–851 (2016).
31. Marino, N. D. et al. Discovery of widespread type I and type V CRISPR-Cas inhibitors. *Science* **362**, 240–242 (2018).
32. Turner, K. H., Wessel, A. K., Palmer, G. C., Murray, J. L. & Whiteley, M. Essential genome of *Pseudomonas aeruginosa* in cystic fibrosis sputum. *Proc. Natl Acad. Sci. USA* **112**, 4110–4115 (2015).
33. Selle, K., Klaenhammer, T. R. & Barrangou, R. CRISPR-based screening of genomic island excision events in bacteria. *Proc. Natl Acad. Sci. USA* **112**, 8076–8081 (2015).
34. Chayot, R., Montagne, B., Mazel, D. & Ricchetti, M. An end-joining repair mechanism in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **107**, 2141–2146 (2010).
35. Lindeberg, M., Cunnac, S. & Collmer, A. *Pseudomonas syringae* type III effector repertoires: last words in endless arguments. *Trends Microbiol.* **20**, 199–208 (2012).
36. Kvitko, B. H. et al. Deletions in the repertoire of *Pseudomonas syringae* pv. tomato DC3000 type III secretion effector genes reveal functional overlap among effectors. *PLoS Pathog.* **5**, e1000388 (2009).
37. Caliando, B. J. & Voigt, C. A. Targeted DNA degradation using a CRISPR device stably carried in the host genome. *Nat. Commun.* **6**, 6989 (2015).
38. Bachman, M. A. et al. Genome-wide identification of *Klebsiella pneumoniae* fitness genes during lung infection. *mBio* **6**, e00775 (2015).
39. Cady, K. C., Bondy-Denomy, J., Heussler, G. E., Davidson, A. R. & O'Toole, G. A. The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. *J. Bacteriol.* **194**, 5728–5738 (2012).
40. Bondy-Denomy, J., Pawluk, A., Maxwell, K. L. & Davidson, A. R. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **493**, 429–432 (2013).
41. Rauch, B. J. et al. Inhibition of CRISPR-Cas9 with bacteriophage proteins. *Cell* **168**, 150–158.e10 (2017).
42. Stanley, S. Y. et al. Anti-CRISPR-associated proteins are crucial repressors of Anti-CRISPR transcription. *Cell* **178**, 1452–1464.e13 (2019).
43. Ha, A. D. & Denver, D. R. Comparative genomic analysis of 130 bacteriophages infecting bacteria in the genus *Pseudomonas*. *Front. Microbiol.* **9**, 1456 (2018).
44. Pósfai, G. et al. Emergent properties of reduced-genome *Escherichia coli*. *Science* **312**, 1044–1046 (2006).
45. Csörgő, B., Nyerges, Á., Pósfai, G. & Fehér, T. System-level genome editing in microbes. *Curr. Opin. Microbiol.* **33**, 113–122 (2016).
46. Képès, F. et al. The layout of a bacterial genome. *FEBS Lett.* **586**, 2043–2048 (2012).
47. Cui, L. & Bikard, D. Consequences of Cas9 cleavage in the chromosome of *Escherichia coli*.*Nucleic Acids Res.* **44**, 4243–4251 (2016).
48. Bowater, R. & Doherty, A. J. Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining. *PLoS Genet*. **2**, e8 (2006).
49. Tuladhar, R. et al. CRISPR–Cas9-based mutagenesis frequently provokes on-target mRNA misregulation. *Nat. Commun*. **10**, 4056 (2019).
50. Smits, A. H. et al. Biological plasticity rescues target activity in CRISPR knock outs. *Nat. Methods* **16**, 1087–1093 (2019).

## Methods

**Bacterial strains, plasmids, DNA oligonucleotides and media.** A previously described[31] environmental strain of *P. aeruginosa* was used as a template to amplify the four *cas* genes of the Type I-C CRISPR–Cas system genes (*cas3*, *cas5*, *cas7* and *cas8*). The genes were cloned into the pUC18-mini-Tn7T-LAC vector[51] using the SacI-PstI restriction endonuclease cut sites in the order *cas5*, *cas7*, *cas8*, *cas3* to generate the plasmid pJW31 (Addgene number 136423). This vector was introduced into *P. aeruginosa* PAO1 (ref. [52]), inserting the *cas* genes into the chromosome, following previously described methods[53]. Following integration, the excess sequences, including the antibiotic resistance marker, were removed via Flp-mediated excision as described previously[67]. The resulting strain, dubbed PAO1$^{IC}$, allowed inducible expression of the I-C system through induction with isopropylβ-D-1 thiogalactopyranoside (IPTG). This same method was used to integrate the Cas3–Cas8 tether mutant in the order *cas5, cas3, cas8, cas7*. The linker amino acid sequence is RSTNRAKGLEAVS. An isogenic strain carrying Cas9 derived from *S. pyogenes* was constructed in the same fashion, resulting in the strain PAO1$^{IIA}$. For experiments to test the system in *P. syringae*, we used the previously characterized strain DC3000 (ref. [54]). *E. coli* editing experiments were conducted with strain K-12 MG1655 (ref. [55]). Experiments conducted with *K. pneumoniae* were performed using strain KPPR1 (ref. [56]).

To construct the Cas3 helicase and nuclease mutant strains, the PAOI$^{IC}$ system was used to introduce point mutations. crRNAs were designed to target Cas3 along with a HDR template that included the desired mutation, and silent mutations to prevent CRISPR–Cas targeting of the final strain.

To achieve genomic self-targeting of the I-C CRISPR–Cas strains, crRNAs designed to target the genome were expressed from the pHERD20T and pHERD30T shuttle vectors[57]. So-called 'entry vectors' pHERD20T-ICcr and pHERD30T-ICcr were first generated by cloning at the EcoRI and HindIII sites an annealed linear dsDNA template carrying the I-C CRISPR–Cas system repeat sequences flanking two BsaI Type IIS restriction endonuclease recognition sites. Additionally, a preexisting BsaI site in a noncoding site of the pHERD30T and pHERD20T plasmids was mutated using whole-plasmid amplification so it would not interfere with the cloning of the crRNAs[31]. Oligonucleotides with repeat-specific overhangs encoding the various spacer sequences were annealed and phosphorylated using T4 polynucleotide kinase and cloned into the entry vectors using the BsaI sites. For experiments using Cas9, sgRNAs were expressed from the same pHERD30T vector, with the sgRNA construct cloned using the same restriction sites as with the I-C crRNAs.

The all-in-one vector pCas3cRh (Addgene number 133773) is a derivative of the pHERD30T-IC plasmid, with the four I-C system genes cloned downstream of the crRNA site. This was achieved by amplifying the genes *cas3*, *cas5*, *cas8* and *cas7* in two fragments with a junction within *cas8* designed to eliminate an intrinsic BsaI site with a synonymous point mutation. The amplified fragments were cloned into pHERD30T-IC using the Gibson Assembly protocol[58]. Finally, to guard against potential leaky toxic expression, we replaced the *araC*-Para$_{BAD}$ promoter with the rhamnose-inducible *rhaSR*-Prha$_{BAD}$ system[59]. The sequence for *rhaSR*-Prha$_{BAD}$ was amplified from the pJM230 template[59], provided by the laboratory of J.B. Goldberg (Emory University) and cloned into the pHERD30T-IC plasmid to replace *araC*-Para$_{BAD}$ using the Gibson Assembly (New England Biolabs). Without induction, transformation efficiencies of targeting constructs of assembled pCas3cRh were on average 5–10-fold lower when compared to nontargeting controls (Extended Data Fig. 7c), indicating residual leakiness of the I-C system.

The *aca1*-containing vector pICcr-*aca1* is a derivative of the pHERD30T-ICcr plasmid, with *aca1* cloned downstream of the crRNA site under the control of the pBAD promoter. The *aca1* gene was cloned from *P. aeruginosa* phage DMS3m.

All oligonucleotides used in this study were obtained from Integrated DNA Technologies. For a complete list of all DNA oligonucleotides and a short description, see Supplementary Table 4.

*P. aeruginosa*, *E. coli* and *K. pneumoniae* strains were grown in standard lysogeny broth (LB): 10 g of tryptone, 5 g of yeast extract and 10 g of NaCl per 1 l of $dH_2O$. Solid plates were supplemented with 1.5% agar. *P. syringae* was grown in King's medium B: 20 g of Bacto Proteose Peptone No. 3, 1.5 g of $K_2HPO_4$, 1.5 g of $MgSO_4 \bullet 7H_2O$, 10 ml of glycerol per 1 l of $dH_2O$, supplemented with 100 μg ml$^{-1}$ rifampicin. The following antibiotic concentrations were used for selection: 50 μg ml$^{-1}$ gentamicin for *P. aeruginosa*, *P. syringae* and *K. pneumoniae*, 15 μg ml$^{-1}$ for *E. coli*; 50 μg ml$^{-1}$ carbenicillin for all organisms. Inducer concentrations were 0.5 mM IPTG, 0.1% arabinose and 0.1% rhamnose. For transformation protocols, all bacteria were recovered in Super optimal broth with catabolite repression: 20 g of tryptone, 5 g of yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM, $MgSO_4$ and 20 mM glucose in 1 l of $dH_2O$.

**Bacterial transformations.** Transformations of *P. aeruginosa*, *E. coli*, *P. syringae* and *K. pneumoniae* strains were conducted using standard electroporation protocols. Then, 10 ml of overnight cultures were centrifuged and washed twice in an equal volume of 300 mM sucrose (20% glycerol for *E. coli*) and suspended in 1 ml of 300 mM sucrose (20% glycerol for *E. coli*). Next, 100-μl aliquots of the resulting competent cells were electroporated using a Gene Pulser Xcell Electroporation System (BioRad) with 50–200 ng of plasmid with the following settings: 200 Ω, 25 μF, 1.8 kV, using 0.2-mm gap width electroporation cuvettes (BioRad). Electroporated cells were incubated in antibiotic-free Super optimal broth with catabolite repression media for 1 h at 37 °C (28 °C for *P. syringae*), then plated onto LB agar (King's medium B agar for *P. syringae*) with the selecting antibiotic and grown overnight at 37 °C (28 °C for *P. syringae*). Cloning procedures were performed in commercial *E. coli* DH5α cells (New England Biolabs) or *E. coli* XL1-Blue (QB3 Macrolab Berkeley), according to the manufacturer's protocols.

**Construction of recombinant DMS3m acr phages.** The isogenic DMS3m*acrIIA4* and *acrIC1* phages were constructed using previously described methods[60]. A recombination cassette, pJZ01, was constructed with homology to the DMS3m *acr* locus. Using the Gibson Assembly (New England Biolabs), either *acrIC1*or *acrIIA4* were cloned upstream of *aca1*, and the resulting vectors were used to transform PAO1$^{IC}$. The transformed strains were infected with wild-type DMS3m, and recombinant phages were screened for. Phages were stored in SM buffer at 4 °C.

**Isolation of PAO1$^{IC}$ lysogens.** PAO1$^{IC}$ was grown overnight at 37 °C in LB media. Then 150 μl of overnight culture was added to 4 ml of 0.7% LB top agar and spread on 1.5% LB agar plates supplemented with 10 mM $MgSO_4$·5 μl of phage, expressing either *acrIC1* or *acrIIA4* were spotted on the solidified top agar and plates were incubated at 30 °C overnight. Following incubation, bacterial growth within the plaque was isolated and spread on a 1.5% LB agar plate. After an overnight incubation at 37 °C, single colonies were assayed for the prophage. Confirmed lysogens were used for genomic targeting experiments.

**Genomic targeting.** *P. aeruginosa.* Genomic self-targeting of *P. aeruginosa* PAO1$^{IC}$ was achieved by electroporating cells with pHERD30T (or pHERD20T) expressing the self-targeting spacer of choice. Cells were plated onto LB agar plates containing the selective antibiotic, without inducers, and grown overnight. Single colonies were then grown in liquid LB media containing the selective antibiotic, as well as IPTG to induce the genomic expression of the I-C system genes, and arabinose to induce the expression of the crRNA from the plasmid. The *aca1*-containing crRNA plasmids do not need additional inducers, as the pBAD promoter controls *aca1*. Cultures were grown at 37 °C in a shaking incubator overnight to saturation, then plated onto LB agar plates containing the selecting antibiotic, as well as the inducers, and incubated overnight again at 37 °C. The resulting colonies were then analyzed individually using colony PCR for any differences at the targeted genomic site compared to a wild-type cell. gDNA was isolated by resuspending one colony in 20 μl of $H_2O$, followed by incubation at 95 °C for 15 min. Then 1–2 μl of boiled sample was used for PCR. The primers used to assay the targeted sites were designed to amplify genomic regions 1.5–3 kb in size. In the event of a PCR product equal to or smaller than the wild-type fragment (as was often observed when analyzing Cas9-targeted cells), Sanger sequencing (Quintara Biosciences) was used to determine any modifications of the targeted sequences. In some cases, additional analysis of the crRNA-expressing plasmids of the surviving colonies was also performed, by isolating and reintroducing the plasmids into the original I-C CRISPR–Cas strain, where functional self-targeting could be determined based on a substantial increase in the lag time of induced cultures, characteristic of self-targeting events.

In cases where a HDR template was used, homology arms ranging in size of 500–600 bp were cloned using a nested PCR-based approach where the two different arms were stitched together via 25-bp overlaps. These fragments were then cloned into the pHERD30T plasmid expressing self-targeting crRNAs at the NheI restriction sites. Genomic targeting was performed as described above. Surviving cells were analyzed using colony PCRs amplifying the desired deletion junction (verified with Sanger sequencing), as well as the wild-type target site. Editing efficiencies were counted as the number of colonies producing a desired deletion junction fragment from the total number of analyzed colonies.

*P. aeruginosa* PA14 genome editing using the endogenous Type I-F system. A self-targeting crRNA and HDR template encoded in pHERD30T were used to direct genome editing in PA14. For each desired transformation, 2 ml of cultured cells were prepared. PA14 was grown in LB at 37 °C until late log phase. Cells were centrifuged at 16,000*g* for 2 min in 2-ml tubes. The supernatant was removed and cells were resuspended in 300 μl of 300 mM sucrose. Cells were spun down again and resuspended in 100 μl of 300 mM sucrose. Next, 800 ng of plasmid was added to cells before electroporation (200 Ω, 25 μF, 2,500 V). Then 2 ml of LB media was added to the cells. Cells were first incubated at room temperature for 10 min then 37 °C for 50 min and were plated on LB agar plates supplemented with gentamicin (50 μg ml$^{-1}$). The following day, single colonies were resuspended in 1 ml of M9 media. Then 6 μl of this resuspension were further diluted into 1 ml of M9. A 6-μl aliquot of the final resuspension was spread on LB gentamicin plates and these were incubated overnight. Single colonies from the replated bacteria were tested for deletions using colony PCR.

*E. coli.* Genomic self-targeting of *E. coli* was conducted in a similar fashion to *P. aeruginosa*, except using the pCas3cRh all-in-one vector. Electrocompetent *E. coli* cells were transformed with pCas3cRh expressing a crRNA targeting the genome. Individual transformants were selected and grown in liquid LB media containing the selecting antibiotic (gentamicin) overnight without any inducers added. The overnight cultures were then plated in the presence of inducer and

X-gal to screen for functional *lacZ* (LB agar + 15 μg ml$^{-1}$ gentamicin + 0.1% rhamnose + 1 mM IPTG + 20 μg ml$^{-1}$ X-gal) and blue/white colonies were counted the next day.

*P. syringae.* Electrocompetent *P. syringae* cells were also transformed with pCas3cRh plasmids targeting selected genomic sequences. Initial transformants were plated onto King's medium B agar + 100 μg ml$^{-1}$ rifampicin + 50 μg ml$^{-1}$ gentamicin plates and incubated at 28 °C overnight. Single colony transformants were then selected and inoculated in King's medium B liquid media supplemented with rifampicin, gentamicin and 0.1% rhamnose inducer, and grown to saturation in a shaking incubator at 28 °C. Cultures were finally plated onto King's medium B agar plates with rifampicin, gentamicin and rhamnose and incubated at 28 °C. Individual colonies were finally assayed with colony PCR to determine the presence of deletions at the targeted genomic sites.

*K. pneumoniae.* Electrocompetent *K. pneumoniae* cells were transformed with pCas3cRh plasmids targeting selected genomic sequences. Individual transformants were selected and grown in liquid LB media (containing 50 μg ml$^{-1}$ gentamicin, as well as 0.1% rhamnose inducer) overnight. Various dilutions from saturated cultures were then plated the next day onto LB agar plates containing the selective antibiotic (gentamicin 50 μg ml$^{-1}$) and 0.1% rhamnose inducer. Individual colonies were then assayed for deletions using colony PCR.

*Iterative genome minimization.* Iterative targeting to generate multiple deletions in the *P. aeruginosa* PAO1$^{IC}$ strain was carried out by alternating the pHERD30T and pHERD20T plasmids each expressing different crRNAs targeting the genome. Each crRNA designed to target the genome was cloned into both the pHERD30T plasmid, which confers gentamicin resistance, as well as the pHERD20T plasmid, which confers carbenicillin resistance. After first transforming and targeting with a pHERD30T plasmid expressing a specific crRNA, deletion candidate isolates were transformed with a pHERD20T expressing a crRNA targeting a different genomic region. As the two plasmids are identical to the exception of the resistance marker, this eliminated the necessity for curing of the original plasmid to be able to target a different region. For the next targeting event, the pHERD30T plasmid could again be used, this time expressing another crRNA targeting a different genomic region. In this manner, pHERD30T and pHERD20T could be alternated to achieve multiple deletions in a rapid process. At each new transformation step, the cells were checked for any residual resistance to the given antibiotic from a previous cycle. Additionally, functionality of the CRISPR–Cas system of the edited cells could be determined through the introduction of a plasmid expressing crRNA targeting the D3 bacteriophage[40], then performing a phage spotting assay to see if phage targeting was occurring or not.

**Measurement of growth rates.** *P. aeruginosa.* Growth dynamics of various strains were measured using a Synergy 2 automated 96-well plate reader (Biotek Instruments) and the accompanying Gen5 software (Biotek Instruments). Individual colonies were picked and grown overnight in 300 μl volumes of LB in 96-well deep-well plates at 37 °C. The grown cultures were then diluted 100-fold into 100 μl of fresh LB in a 96-well clear microtiter plate (Costar) and sealed with Microplate sealing adhesive (Thermo Scientific). Small holes were punched in the sealing adhesive for each well for increased aeration. Doubling times were calculated as described previously[61].

*P. syringae.* To test bacterial growth in planta, we used the *Arabidopsis thaliana* ecotype *Columbia* (Col-0), which has previously been shown to be susceptible to infection by *P. syringae* DC3000. Plants were grown for 5–6 weeks in 9 h light/15 h darkness and 65% humidity. For each inoculum, we measured bacterial growth in ten individual Col-0 plants. Four leaves from each plant were infiltrated at an optical density ($OD_{600}$) of 0.0002, and cored with a no. 3 borer. The four cores from each plant were then ground, resuspended in 10 mM $MgCl_2$ and plated in a dilution series on selective media for colony counts at both the time of infection and 3 d postinfection.

To test bacterial growth in vitro, we used both King's medium B and plant apoplast mimicking minimal media[62]. Overnight cultures were prepared from single colonies of each strain, washed and diluted to $OD_{600} = 0.1$ in 96-well plates using either King's medium B or minimal media. Plates were incubated with shaking at 28 °C. $OD_{600}$ was measured over the course of 24–25 h using an Infinate 200 Pro automated plate reader (Tecan). Statistical analysis determined significantly different groups based on analysis of variance (ANOVA) analysis on the day 0 group of values and the day 3 group of values. Significant ANOVA results ($P < 0.01$) were further analyzed with a Tukey's honestly significant difference post hoc test to generate adjusted $P$ values for each pairwise comparison. A significance threshold of 0.01 was used to determine which treatment groups were significantly different.

*Bacteriophage plaque (spot) assays.* Bacteriophage plaque assays were performed using 1.5% LB agar plates supplemented with 10 mM $MgSO_4$ and the appropriate antibiotic (gentamicin or carbenicillin, depending on the plasmid used to express the crRNA) and 0.7% LB top agar supplemented with 0.5 mM IPTG and 0.1% arabinose inducers added covering the whole plate. Then, 150 μl of the appropriate overnight cultures was suspended in 4 ml molten top agar poured onto an LB agar plate leading to the growth of a bacterial lawn. After 10–15 min at room temperature, 3 μl of tenfold serial dilutions of bacteriophage was spotted onto the solidified top agar. Plates were incubated overnight at 30 °C and imaged the following day using a Gel Doc EZ Gel Documentation System (BioRad) and Image Lab (BioRad) software. The following bacteriophage were used in this study: bacteriophage JBD30 (ref. [40]), bacteriophage D3 (ref. [63]) and bacteriophage DMS3m[64].

*WGS.* The gDNA for WGS analysis was isolated directly from bacterial colonies using the Nextera DNA Flex Microbial Colony Extraction kit (Illumina) according to the manufacturer's protocol. The gDNA concentration of the samples was determined using a DS-11 Series Spectrophotometer/Fluorometer (DeNovix) and all fell into the range of 200–500 ng μl$^{-1}$. Library preparation for WGS analysis was done using the Nextera DNA Flex Library Prep kit (Illumina) according to the manufacturer's protocol starting from the tagment genomic DNA step. Tagmented DNA was amplified using Nextera DNA CD Indexes (Illumina). Samples were placed overnight at 4 °C following the tagmented DNA amplification step, then continued the next day with the library clean up steps. Quality control of the pooled libraries was performed using a 2100 Bioanalyzer Instrument (Agilent Technologies) with a High Sensitivity DNA Kit (Agilent Technologies). Most samples were sequenced using a MiSeq Reagent Kit v.2 (Illumina) for a 150 bp paired-end sequencing run using the MiSeq sequencer (Illumina). *P. syringae* and Cas9-generated *P. aeruginosa* deletion strains were sequenced using a NextSeq 500 Reagent Kit v.2 (Illumina) for a 150 bp paired-end sequencing run using the NextSeq 500 sequencer (Illumina).

Genome sequence assembly was performed using Geneious Prime software v.2019.1.3. Paired read data sets were trimmed using the BBDuk (decontamination using kmers) plugin using a minimum $Q$ value of 20. The genome for the ancestral PAO1$^{IC}$ strain was de novo assembled using the default automated sensitivity settings offered by the software. The consensus sequence of PAO1$^{IC}$ assembled in this manner was then used as the reference sequence for mapping all of the PAO1$^{IC}$ strains with multiple deletions. As a control, the sequences were also mapped to the reference *P. aeruginosa* PAO1 sequence (NC_002516) to verify deletion border coordinates. Coverage of these sequenced strains ranged from 66- to 143-fold, with an average of 98.3-fold. The sequenced *P. aeruginosa* environmental strains were also mapped to the PAO1 (NC_002516) reference, while the sequenced *E. coli* strains were mapped to the *E. coli* K-12 MG1655 reference sequence (NC_000913). Finally, sequenced *P. syringae* strains were mapped to the *P. syringae* DC3000 (NC_004578) reference sequence, along with the pDC3000A endogenous 73.5 kb plasmid sequence (NC_004633). All of the remaining sequenced strains had >100-fold coverage. All deletion junction sequences were manually verified by the presence of multiple reads spanning the deletions, containing sequences from both end boundaries.

WGS data were visualized using the BLAST Ring Image Generator tool[65] using BLAST+ v.2.9.0. In several cases, short sequences were aligned inside previously determined large deletions at redundant sequences such as transposase genes. Such misrepresentations created by the BLAST Ring Image Generator were manually removed to reflect the actual sequencing data.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Plasmids pJW31 and pCas3cRh are available through Addgene (numbers 136423 and 133773, respectively). Raw WGS data associated with Figs. 1d, 3b, 4a,c,f and 5a) have been uploaded to GenBank (accession numbers CP047061–CP047079) and are also available, along with bacterial strains, upon request from the corresponding author. *P. aeruginosa* strains available for laboratories with BSL-2 clearance. Source data are provided with this paper.

## References


51. Choi, K.-H. et al. A Tn7-based broad-range bacterial cloning and expression system. *Nat. Methods* **2**, 443–448 (2005).
52. Stover, C. K. et al. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* **406**, 959–964 (2000).
53. Choi, K.-H. & Schweizer, H. P. mini-Tn7 insertion in bacteria with single attTn7 sites: example *Pseudomonas aeruginosa*. *Nat. Protoc.* **1**, 153–161 (2006).
54. Buell, C. R. et al. The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. tomato DC3000. *Proc. Natl Acad. Sci. USA* **100**, 10181–10186 (2003).
55. Blattner, F. R. et al. The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1462 (1997).
56. Broberg, C. A., Wu, W., Cavalcoli, J. D., Miller, V. L. & Bachman, M. A. Complete genome sequence of *Klebsiella pneumoniae* strain ATCC 43816 KPPR1, a Rifampin-resistant mutant commonly used in animal, genetic, and molecular biology studies. *Genome Announc.* **2**, e00924–14 (2014).

57. Qiu, D., Damron, F. H., Mima, T., Schweizer, H. P. & Yu, H. D. PBAD-based shuttle vectors for functional analysis of toxic and highly regulated genes in *Pseudomonas* and *Burkholderia* spp. and other bacteria. *Appl. Environ. Microbiol.* **74**, 7422–7426 (2008).
58. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
59. Meisner, J. & Goldberg, J. B. The *Escherichia coli* rhaSR-PrhaBAD inducible promoter system allows tightly controlled gene expression over a wide range in *Pseudomonas aeruginosa*. *Appl. Environ. Microbiol.* **82**, 6715–6727 (2016).
60. Borges, A. L. et al. Bacteriophage cooperation suppresses CRISPR-Cas3 and Cas9 immunity. *Cell* **174**, 917–925.e10 (2018).
61. Nyerges, Á. et al. Directed evolution of multiple genomic loci allows the prediction of antibiotic resistance. *Proc. Natl Acad. Sci. USA* **115**, E5726–E5735 (2018).
62. Huynh, T. V., Dahlbeck, D. & Staskawicz, B. J. Bacterial blight of soybean: regulation of a pathogen gene determining host cultivar specificity. *Science* **245**, 1374–1377 (1989).
63. Kropinski, A. M. Sequence of the genome of the temperate, serotype-converting, *Pseudomonas aeruginosa* bacteriophage D3. *J. Bacteriol.* **182**, 6066–6074 (2000).
64. Budzik, J. M., Rosche, W. A., Rietsch, A. & O'Toole, G. A. Isolation and characterization of a generalized transducing phage for *Pseudomonas aeruginosa* strains PAO1 and PA14. *J. Bacteriol.* **186**, 3270–3273 (2004).
65. Alikhan, N.-F., Petty, N. K., Ben Zakour, N. L. & Beatson, S. A. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* **12**, 402 (2011).

## Acknowledgements

B.C. is supported by the Eötvös National Scholarship of Hungary and a Marie Skłodowska-Curie Actions Individual Global Fellowship ('GenDels', no. 844093) of the Horizon 2020 Research Program of the European Commission. L.M.L. is supported by the HHMI Gilliam Fellowship for Advanced Study and the UCSF Discovery Fellowship. Research on plant immunity in the Lewis laboratory is supported by the USDA grant nos. ARS 2030-21000-046-00D and 2030-21000-050-00D (J.D.L.), and the NSF Directorate for Biological Sciences grant no.IOS-1557661 (J.D.L.). I.J.C.-L. is supported by a Grace Kase fellowship from UC Berkeley and the NSF Graduate Research Fellowship Program. A.V.R. is supported by funding from the Pew Charitable Trusts. E.D.C. is funded by the Chan Zuckerberg Biohub. CRISPR–Cas3 projects in the Bondy-Denomy Laboratory are supported by the UCSF Program for Breakthrough Biomedical Research funded in part by the Sandler Foundation, the Innovative Genomics Institute and an NIH Director's Early Independence Award DP5-OD021344. We thank J.B. Goldberg (Emory University) for providing the plasmid pJM230, and A. Borges (UCSF) for providing pAB01 to clone Type I-F crRNAs. We thank the Bondy-Denomy laboratory for productive conversations pertaining to this project.

## Author contributions

B.C. and L.M.L. participated in designing and performing experiments, analyzing data, acquiring funding for the project and writing the manuscript. I.J.C.-L. performed in vitro and in planta *P. syringae* experiments. A.V.-R. performed and analyzed Type I-F CRISPR–Cas editing experiments. J.D.B. assisted in designing and constructing the all-in-one pCas3cRh vector. C.M. constructed the PAO1$^{IIA}$ strain and Cas9 gRNA expression vector. E.D.C. performed WGS. J.D.L. designed experiments with *P. syringae*. J.B.-D. conceived the study, designed experiments, analyzed data, acquired funding for the project and wrote the manuscript.

## Competing interests

J.B.-D. is a scientific advisory board member of SNIPR Biome and Excision Biotherapeutics and a scientific advisory board member and cofounder of Acrigen Biosciences. J.B.-D., L.M.L. and B.C. have filed a patent application relating to various aspects of Cas3-based genome editing.

## Additional information

**Extended data** is available for this paper at https://doi.org/10.1038/s41592-020-00980-w.

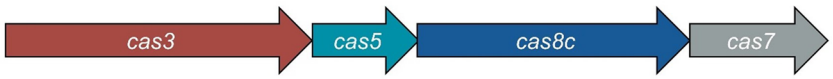
**Supplementary information** is available for this paper at https://doi.org/10.1038/s41592-020-00980-w.

**Correspondence and requests for materials** should be addressed to J.B.-D.

**Peer review information** Lei Tang was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

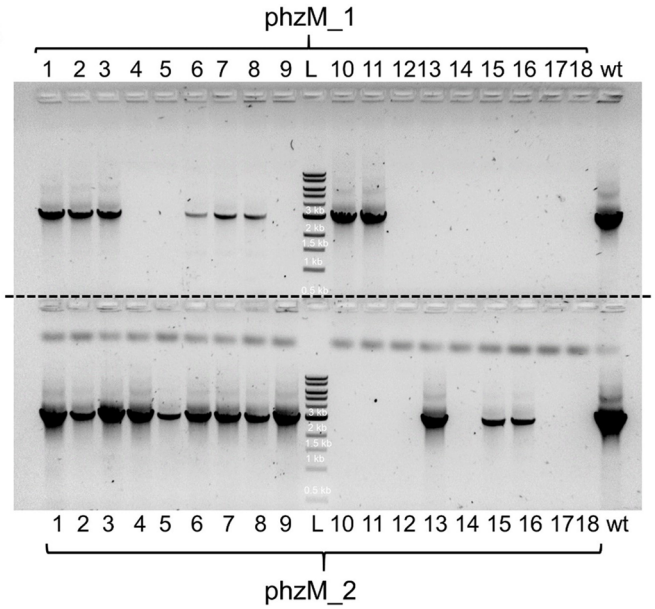
**Reprints and permissions information** is available at www.nature.com/reprints.

A

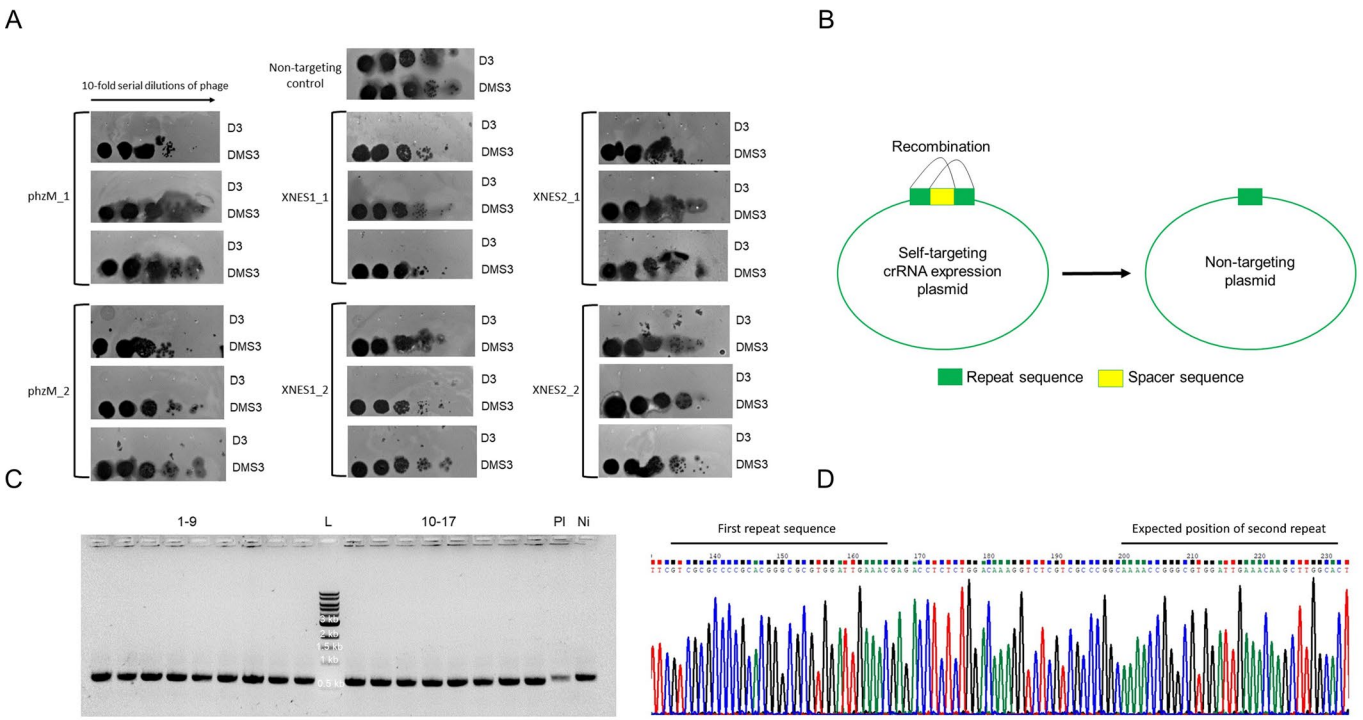



| | Cas3 (%) | Cas5 (%) | Cas8 (%) | Cas7 (%) |
|---|---|---|---|---|
| *Bacillus halodurans** | Query cover.: 98<br>ID: 32 | Query cover.: 93<br>ID: 31 | Query cover.: 96<br>ID: 24 | Query cover.: 94<br>ID: 35 |
| *Bacillus cereus* | Query cover.: 91<br>ID: 31 | Query cover.: 93<br>ID: 29 | Query cover.: 19<br>ID: 29 | Query cover.: 94<br>ID: 33 |
| *Lactobacillus fermentum* | Query cover.: 97<br>ID: 27 | Query cover.: 93<br>ID: 28 | Query cover.: 78<br>ID: 22 | Query cover.: 87<br>ID: 31 |
| *Mycobacterium canettii* | Query cover.: 93<br>ID: 36 | Query cover.: 93<br>ID: 41 | Query cover.: 99<br>ID: 28 | Query cover.: 95<br>ID: 45 |
| *Streptococcus pyogenes* | Query cover.: 92<br>ID: 28 | Query cover.: 91<br>ID: 29 | Query cover.: 23<br>ID: 31 | Query cover.: 84<br>ID: 34 |

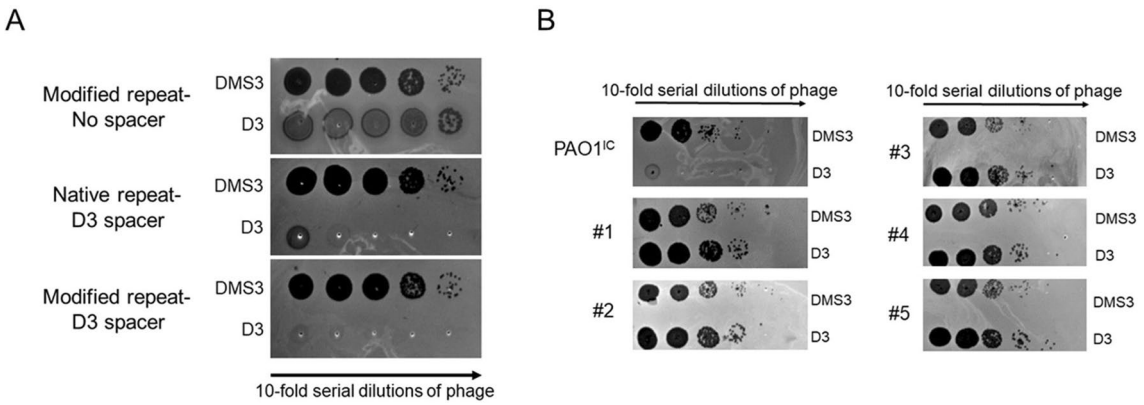
B



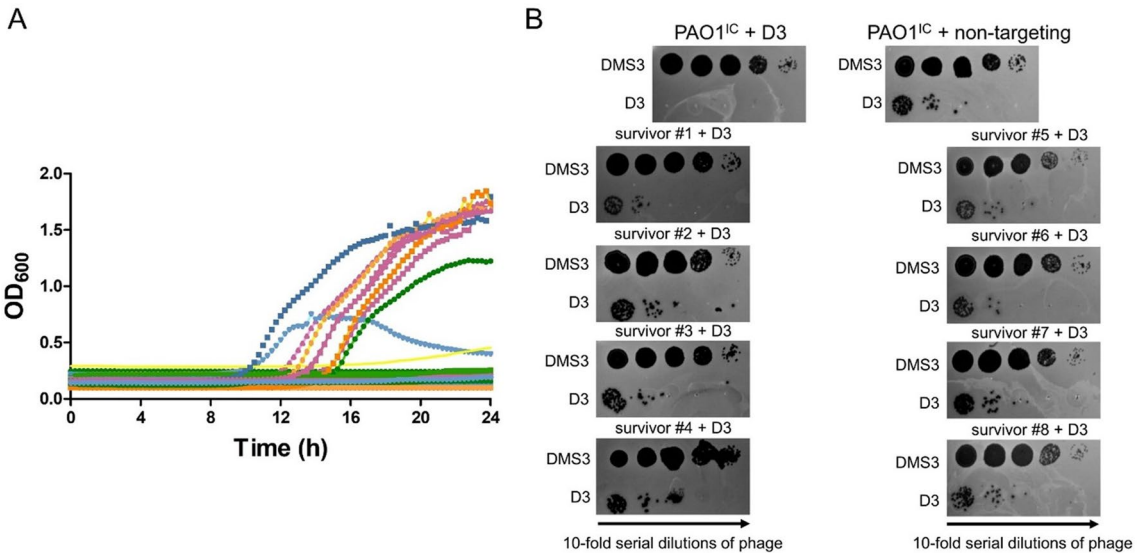

**Extended Data Fig. 1 | Type I-C CRISPR targeting leads to genomic deletions. a**, Comparison of Type I-C CRISPR system from *P. aeruginosa* used in the study, to various other previously identified I-C systems from a range of different bacteria. Values show query coverage and percent identity (ID) percentages comparing the four genes of the *P. aeruginosa* system to each of the other four. * Denotes the reference Type I-C CRISPR system referred to in Ref. [1]. **b**, PCR amplification of a 3 kb genomic fragment flanking the *phzM* gene targeted using two different crRNAs, phzM_1 and phzM_2. Colony PCRs were performed on 18 biological replicates of self-targeted strains for each crRNA. The PAO1$^{IC}$ parental strain is used as a positive control (wt). L indicates a 1 kb DNA ladder.



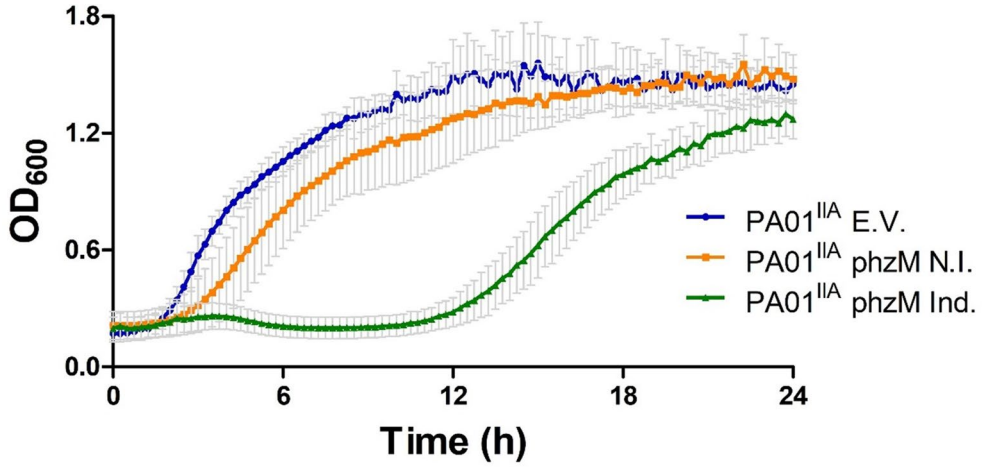

**Extended Data Fig. 2 | Excision of plasmid-encoded spacer sequences. a**, Phage targeting assays with survivors that had no discernable deletion of the crRNA-targeted genomic site. Strains were transformed with a D3 phage-targeting crRNA to assay for IC CRISPR-Cas3 activity. Three unique survivors were isolated from six self-targeting assays for a total of 18 survivors. Control is a non-targeting crRNA. **b**, Schematic of spacer excision events where the two direct repeats recombine, resulting the loss of the targeting spacer. **c**, PCR amplification of the crRNA sequence from plasmids isolated from 17 non-deletion self-targeted survivors (selected from 3 biological replicates of 12 analyzed colonies (see Fig. 2a). Pl indicates the original plasmid as the PCR template, Ni indicates a sample where the crRNA was not induced, L indicates a 1 kb DNA ladder. **d**, Sample chromatogram of a sequenced plasmid with the spacer flipped out. Only one 32 bp repeat sequence remains in the plasmid, the 34 bp spacer sequence and other 32 bp repeat are missing.



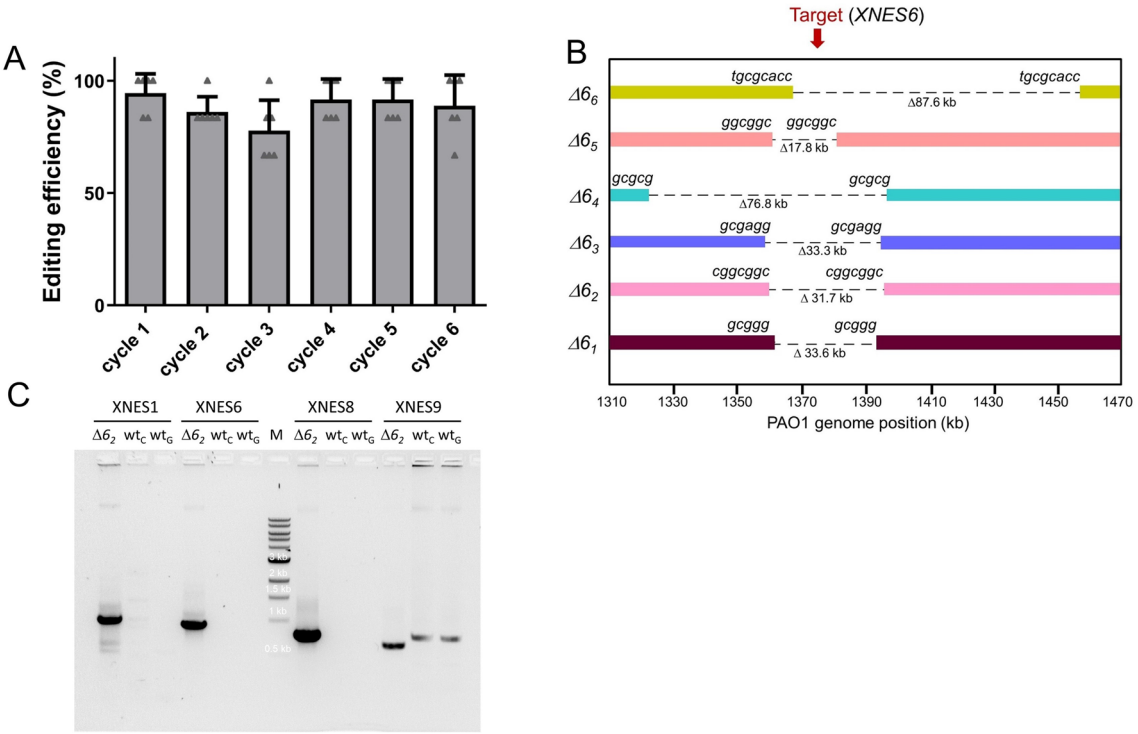

**Extended Data Fig. 3 | Phage-targeting assays to confirm CRISPR-Cas functionality. a**, Phage-targeting assay showing the activity of the modified repeat crRNA constructs. Ten-fold serial dilutions of DMS3 phage and D3 phage were spotted on lawns of PAO1$^{IC}$ expressing either empty vector (top), a crRNA targeting D3 with WT direct repeats (middle), or a crRNA targeting D3 with modified repeats (bottom). **b**, Phage targeting assay of five non-deletion self-targeting survivors expressing a D3 phage targeting crRNA. Unsuccessful targeting of phage indicates a non-functional CRISPR-Cas system in these strains. The parental PAO1$^{IC}$ strain with a functional CRISPR-Cas system was used as a control.



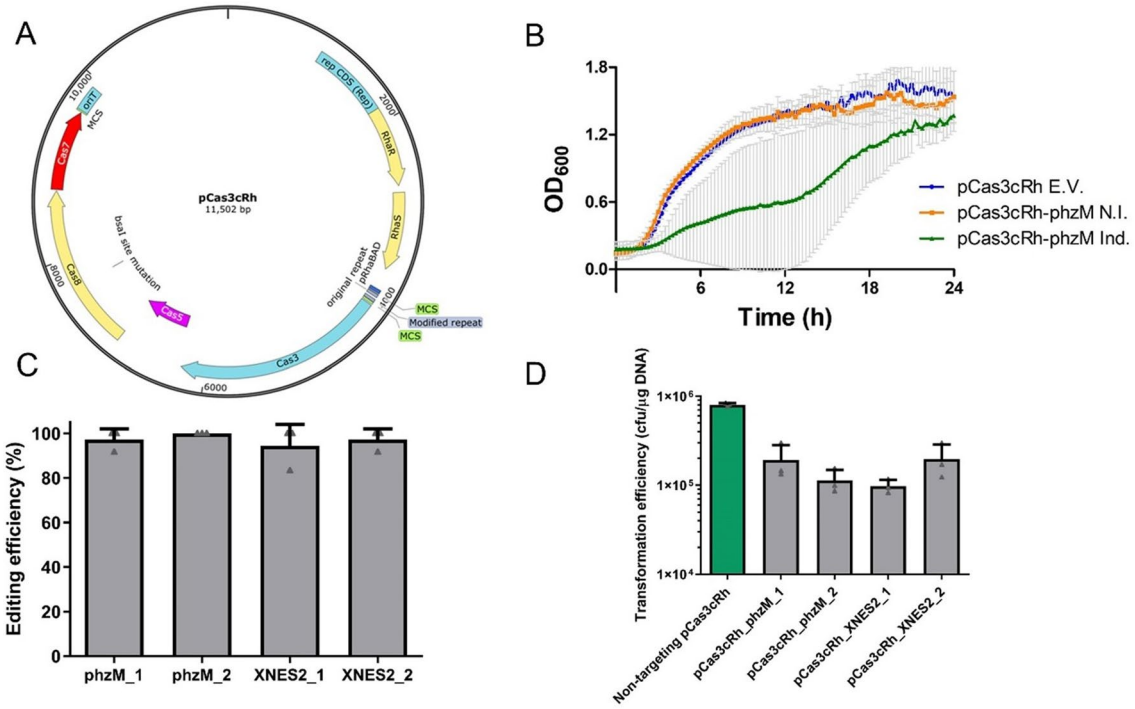

**Extended Data Fig. 4 | Genomic targeting of essential gene *rplQ*. a**, Growth curves of 36 PAO1$^{IC}$ biological replicates targeting the essential gene, *rplQ*, using the MR crRNA plasmid. **b**, Phage targeting assays with eight isolated *rplQ*-targeted survivors to assay for I-C CRISPR-Cas activity. Serial dilutions of DMS3 phage and D3 phage were spotted on lawns of PAO1$^{IC}$ expressing a crRNA targeting phage D3. The parent PAO1$^{IC}$ strain expressing a D3 targeting crRNA (top left) was used as a positive control, while PAO1$^{IC}$ expressing a non-targeting crRNA was used as a negative control.



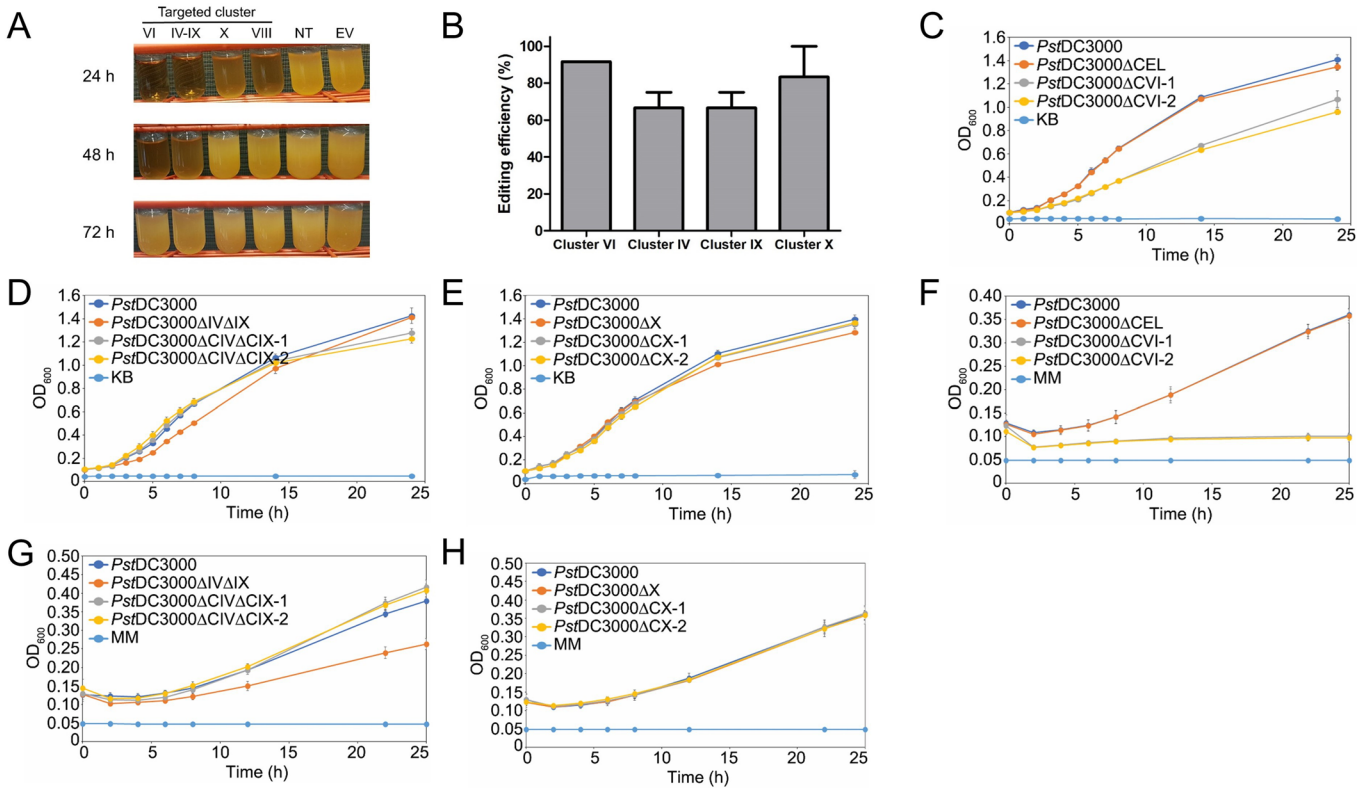

**Extended Data Fig. 5 | Genomic targeting using a Type II-A CRISPR-Cas system.** Growth of self-targeting strains of PAO1$^{IIA}$ expressing a self-targeting gRNA targeting the genome at *phzM* (Ind.). An empty vector (E.V.) and a non-induced *phzM* targeting strain (N.I.) were used as controls. Mean OD values measured at 600 nm are shown for 8 biological replicates each, error bars indicate SD values.



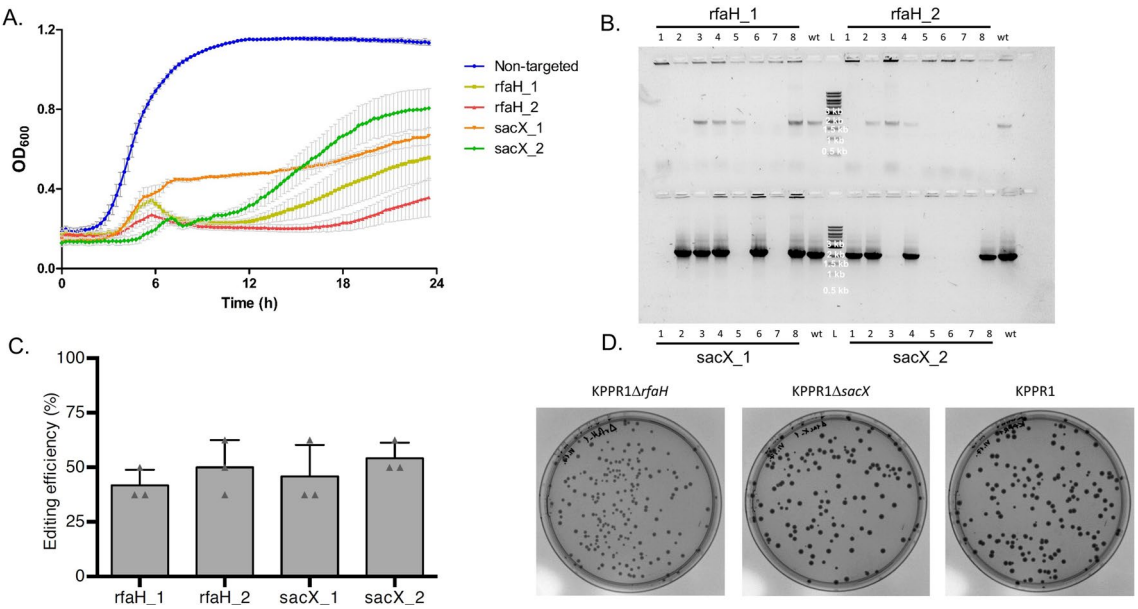

**Extended Data Fig. 6 | Genomic deletions and junction sites. a**, Deletion efficiencies observed over six cycles of iterative self-targeting. Six genomic targets were targeted in six different orders. Six survivors were analyzed using site-specific PCR after each cycle, for a total of 36 analyzed colonies (6*6) after each cycle, error bars represent standard deviations. **b**, Deletion junctions at *XNES6* target site of the 6 PAO1$^{IC}$ strains with 6 iterative targeting events each. Sequences of each specific microhomology for the junctions are shown for each strain above the bars representing the given genomes at both ends, deletion sizes are shown below dashed lines for each strain. **c**, PCR analysis using a representative set of primers amplifying various large deletion junctions (at XNES1, 6, 8, and 9 regions) of the whole-genome sequenced $\Delta 6_2$ strain. $\Delta 6_2$ served as a positive control template, while $wt_C$ represents untargeted $PAO1_{IC}$ cells scraped from a lawn of colonies from a single overnight culture grown on plates serving as templates, and $wt_G$ represents isolated genomic DNA from a different 1.5 ml overnight culture of untargeted $PAO1_{IC}$ used as templates. Bands appearing for the XNES9 deletion junction for the $PAO1_{IC}$ samples were aspecific and when sequenced, did not match any genomic region of the $PAO1_{IC}$ genome. L indicates a 1 kb DNA ladder.



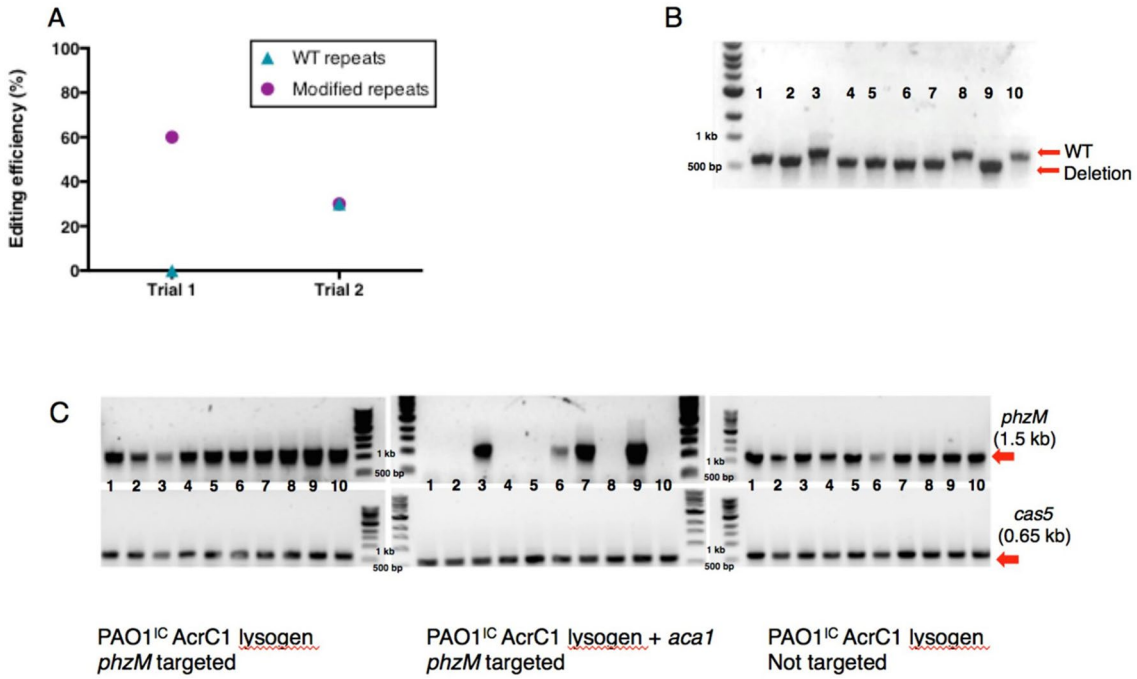

**Extended Data Fig. 7 | Genomic targeting of *P. aeruginosa* PAO1 with all-in-one vector pCas3ch. a**, Map of the I-C CRISPR-Cas all-in-one plasmid pCas3cRh carrying I-C crRNA and genes *cas3*, *cas5*, *cas8*, and *cas7* under the control of the rhamnose-inducible *rhaSR*-Prha$_{BAD}$ system. **b**, Growth curve of PAO1 transformed with the pCas3cRh vector expressing a self-targeting crRNA targeting *phzM* (Ind.). An empty vector (E.V.) and a non-induced *phzM* targeting strain (N.I.) were used as controls. Mean OD values measured at 600 nm are shown for six biological replicates each. **c**, Deletion efficiencies for WT PAO1 using the all-in-one vector pCas3cRh carrying all necessary components of the I-C CRISPR-Cas system. Values are averages of three replicates where 12 individual colonies were analyzed using site-specific PCR. Error bars show standard deviations. **d**, Transformation efficiencies with self-targeting pCas3cRh vectors expressing crRNAs for *phzM* or XNES 2 compared to a non-targeting control (green bar) in PAO1. Values are means of 3 replicates each, error bars represent SD values.




**Extended Data Fig. 8 | Genomic targeting of *Pseudomonas syringae* and growth phenotypes of deletion strains. a**, Growth of *P. syringae* DC3000 strains expressing the I-C system and distinct crRNAs. Constructs VI, IV-IX, and VIII target *P. syringae* DC3000 non-essential chromosomal genes, non-targeting crRNA (NT), empty vector (EV). **b**, Percentage of survivors with targeted deletions in clusters of non-essential virulence effector genes in *P. syringae* pv. *tomato* DC3000. Values are averages of three biological replicates where 12 individual colonies were analyzed using site-specific PCR for each, error bars show standard deviations. **c**, *In vitro* growth of cluster VI deletion strains in King's medium B (KB). ΔCEL is the previously published polymutant, while ΔCVI-1 and ΔCVI-2 are Cas3-generated mutants. Values shown are the means of 4 biological replicates each, error bars represent standard deviations. **d**, *In vitro* growth of cluster IV, cluster IX deletion strains in KB. ΔCEL is the previously published polymutant, while ΔCIVΔCIX-1 and ΔCIVΔCIX-2 are Cas3-generated mutants. Values shown are the means of 4 biological replicates each, error bars represent standard deviations. **e**, *In vitro* growth of cluster X deletion strains in KB. ΔCEL is the previously published polymutant, while ΔCX-1 and ΔCX-2 are Cas3-generated mutants. Values shown are the means of 4 biological replicates each, error bars represent standard deviations. **f**, *In vitro* growth of cluster VI deletion strains in apoplast mimicking minimal media (MM). ΔCEL is the previously published polymutant, while ΔCVI-1 and ΔCVI-2 are Cas3-generated mutants. Values shown are the means of 4 biological replicates each, error bars represent standard deviations. **g**, *In vitro* growth of cluster IV, cluster IX deletion strains in MM. ΔCEL is the previously published polymutant, while ΔCIVΔCIX-1 and ΔCIVΔCIX-2 are Cas3-generated mutants. Values shown are the means of 4 biological replicates each, error bars represent standard deviations. **h**, *In vitro* growth of cluster X deletion strains in MM. ΔCEL is the previously published polymutant, while ΔCX-1 and ΔCX-2 are Cas3-generated mutants. Values shown are the means of 4 biological replicates each, error bars represent standard deviations.




**Extended Data Fig. 9 | CRISPR-Cas3 editing in *Klebsiella pneumoniae*. a**, Growth curves of *K. pneumoniae* strains expressing distinct crRNAs targeting *rfaH* and *sacX* (2 each). Non-targeting crRNA expressing control is marked in blue. Values depicted are averages of 8 biological replicates each. **b**, Representative gel electrophoresis of PCR fragments amplified from 8 total surviving colonies each from the 4 crRNA targeting constructs (representing 1 biological replicate of 3 total). Primer pairs amplified regions flanking the targeted position at *rfaH* and *sacX*. Wild-type KPPR1 (wt) colonies were used as controls, L represents 1 kb DNA marker ladder. **c**, Percentage of survivors with targeted deletions at the targeted genomic positions. Values are averages of three biological replicates where 8 individual colonies were analyzed using site-specific PCR for each, error bars show standard deviations. **d**, Colony morphologies of deletion candidate strains of *rfaH* and *sacX* compared to wild-type *K. pneumoniae* KPPR1.




**Extended Data Fig. 10 | Genomic editing in native host of Type I-C CRISPR-Cas system and effect of I-C specific anti-CRISPR protein on the process.** **a**, Editing efficiencies for the *Pseudomonas aeruginosa* environmental isolate naturally expressing the Type I-C *cas* genes, transformed with a plasmid targeting *phzM* with WT repeats or modified repeats. Each data point represents the fraction of isolates with the deletion out of ten isolates assayed. **b**, Genotyping results for the *Pseudomonas aeruginosa* environmental isolate using the 0.17 kb HDR template. 10 biological replicates were assayed. Larger band corresponds to the WT sequence, smaller band corresponds to a genome reduced by 0.17 kb. **c**, Genotyping results of PAO1$^{IC}$ AcrC1 lysogens after self-targeting induction in the presence or absence of *aca1* and a non-targeted control. Ten biological replicates per strain were assayed. gDNA was extracted from each replicate and PCR analysis for the *phzM* gene (targeted gene, top row of gels) or *cas5* gene (non-targeted gene, bottom row) was conducted. Only cells that co-expressed *aca1* with the crRNA showed loss of the *phzM* band, indicating genome editing. All replicates had a *cas5* band, indicating successful gDNA extraction and target specificity for the *phzM* locus.

nature research

Corresponding author(s): Joseph Bondy-Denomy

Last updated by author(s): Aug 26, 2020

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☒ | ☐ | A description of all covariates tested |
| ☒ | ☐ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | No software was used for data collection. |
| Data analysis | Genome sequence assemblies in the study were performed using Geneious Prime software version 2019.1.3. Gel and plate documentation were performed using ImageLab software version 6.0.1. Automated plate reader experiments were performed using Gen5 software version 3.05. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Raw whole-genome sequencing data associated with Figures 1D, 3B, 4A, 4C, 4E, and 5A) has been uploaded to GenBank (Accession numbers CP047061-CP047079) and is also available, along with bacterial strains, upon request from the corresponding author.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | For gene editing experimetnts, no specific sample size calculation was performed, however genomic self-targeting experiments in Pseudomonas aeruginosa PAO1, Pseudomonas syringae DC3000, Escherichia coli K-12 MG1655, and Klebsiella pneumoniae KPPR1 were all performed in at least 3 biological replicates each with each replicate consisting of the analysis of at least 8 individual colonies. We deemed this to be a sufficient sample size, as multiple genomic target sites were selected for studying each organism (ranging from 4 to 12), providing extensive information regarding gene editing efficiencies. When testing editing capabilities with endogenous systems, less replicates were made as the goal here (as highlighted in the text) was to demonstrate the possibility of genome editing occurring without the emphasis on a detailed characterization in each organism. Whole-genome sequencing was performed on 6 parallel multiple deletion strains (and an additional strain carrying an extra 4 genomic target sites) in the case of Pseudomonas aeruginosa to obtain a detailed as possible picture of deletion generation characterisitics in this organism. 6 parallel multiple deltion lines were the most technically achievable taking into consideration our particular experimental capabilities (i.e. laboratory space, number of incubators, etc.). Less strains were used for whole-genome sequencing from other organisms as the goal in those cases was to demonstrate the capability to generate large deletions for which a smaller number of self-targeted strains was sufficient. |
| Data exclusions | No data were excluded from the analyses. |
| Replication | All genome editing experiments in Pseudomonas aeruginosa PAO1, Pseudomonas syringae DC3000, Escherichia coli K-12 MG1655, and Klebsiella pneumoniae KPPR1 were performed in at least 3 biological replicates each with each replicate consisting of the analysis of at least 8 individual colonies. Additionally, iterative genome editing in P. aeruginosa PAO1 was performed in 6 parallel strains, with large deletions occurring in all of them. All findings were replicated with no major conflicting results observed. The same can be said for the replicates of all other experiments, namely measuring growth rates of strains and induced self-targeting events of interest, as these were measured using at least 8 biological replicates each using automated 96 well microplate readers, with replicates positioned to minimize plate positional effects. Experiments involving phage plaquing were performed as qualitative assays to determine the functionality of the CRISPR systems of selected strains of interest, appropriate controls ensured the validity of these results. |
| Randomization | As the goal of this work was to determine and optimize CRISPR-Cas3 as a genomic editing tool, it was sufficient to always include an appropriate internal control (e.g. expression of a non-targeting crRNA) and compare the effects with the gene-editing samples. In this respect, it was not reasonable to randomize samples as we were investigating the possibility of Cas3-based editing in different bacterial strains all of which had the appropriate internal controls. Additionally, the same strains of bacteria were either induced for self-targeting or had a non-targeting control expressed, it was not technically feasible to individually characterize and separate individual bacterial cells to then randomize which single cells would undergo self-targeting or not. The various species and strains of bacteria undergoing induced self-targeting or expressing a non-targeting crRNA all originated from the same initial bacterial cultures, which were then divided and transformed with the various targeting or non-targeting constructs. Randomizing the expressed self-targeting constructs would have confounded results, maiking it difficult to determine which specific crRNA constructs were responsible for creating large genomic deletions. For growth-rate analysis experiments, it made no sense to randomize any of the samples, as it would have made determining growth phenotypes of specific strains impossible. For whole-genome sequencing, there was no reason to randomize samples, as it would have made determining the genomic effects of specific targeting events difficult to determine. |
| Blinding | Blinding was not relevant in this study, as the effects of genome editing were individually tested for each bacterial organism, with the relevant internal control always present. Blinding the samples when performing gene editing experiments would have prevented the identification of the individual effects of the targeting crRNAs and therefore genetic deletions could not be attributed to the activity of a targeting CRISPR-Cas3 system, raising the possibility that the results could be interpreted as the deletions being present beforehand in the popoulation of cells. Similarly, when measuring growth phenotypes, blinding the samples would have made it difficult to determine the growth rates of individual strains and treatments of interest, as the various samples would have had to be genotypically analyzed after the measurements to identify strain identities. Researchers performing the whole-genome sequencing were unaware to the identities of the different samples. |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☒ | ☐ Animals and other organisms |
| ☒ | ☐ Human research participants |
| ☒ | ☐ Clinical data |

## Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |