SUPPLEMENT

Supplementary Tables

Table S1. Primers used in this study

pBDPfw	CATCTAGTACTTTCCTGTGTGATTG	primers amplifying pBDP_RFP_GFP plasmid, apart from the *rfp* gene
pBDPrev	TAATAACGCTGATAGTGCTAGTG	primers amplifying pBDP_RFP_GFP plasmid, apart from the *rfp* gene
Km3	CACTAGCACTATCAGCGTTATTAGA AGAACTCGTCAAGAAGGC	primers amplifying *Km* gene of pST76-K
Km5	CAATCACACAGGAAAGTACTAGATG ATTGAACAAGATGGATTG	primers amplifying *Km* gene of pST76-K
cycA1	CTGATGCCGGTAGGTTCT	primers amplifying *cycA* gene
cycA2	GCGCCATCCAGCATGATA	primers amplifying *cycA* gene
bglR1	GTGGCGATGAGCTGGAT	primers for checking size of *bgl* regulatory region
bglR2	CCGACTTCACCAGTATTC	primers for checking size of *bgl* regulatory region
VF2	TGCCACCTGACGTCTAAGAA	primers amplifying the *km-gfp* cassette of pBDP_Km_GFP5
VR	ATTACCGCCTTTGAGTGAGC	primers amplifying the *km-gfp* cassette of pBDP_Km_GFP5
IS1A1	TCGCTGTCGTTCTCA	IS*1* outward primer
IS1A2	AAGCCACTGGAGCAC	IS*1* outward primer
IS1F1	GGTAATGCTGCCAACTTACT	IS*1*F inward primner
UK1R	TCGCAGGCATACCATCAA	IS*2* outward primer
UK2R	CAGACGGGTTAACGGCA	IS*2* outward primer
IS3ki1	CTGCGCCAGTTGTAGAGTTG	IS*3* outward primer
IS3ki2	ATAATGCCTGCGTGGAAAGC	IS*3* outward primer
IS4ki1	GCATGATGTCCAGGCGATTC	IS*4* outward primer
is4ki2	ACAAACCGGCGATAACACTC	IS*4* outward primer
IS5ki1	GACAGTTCGGCTTCGTGA	IS*5* outward primer
IS5ki2	GCTCGATGACTTCCACCA	IS*5* outward primer
IS5Y1	GGAAGGTGCGAATAAGCAGG	IS*5*Y inward primer
IS10ki1	CGTGGCAAGCCAACGTTA	IS*10* outward primer
IS10ki2	TGCGGCATTCTGGCTACA	IS*10* outward primer
IS150ki1	ACGTGCCGAGATGATCCT	IS*150* outward primer
IS150ki2	CAGACCTATATGCCTCGT	IS*150* outward primer
IS150Crisp+	AAACGGGGCTATTCCATTTCATCGT CCAACAAAAGTTTTAGAGCTATGCT GTTTTGAATGGTCCCA	IS*150*-specific spacer, positive strand

IS150Crisp-	GTTTTGGGACCATTCAAAACAGCAT AGCTCTAAAACTTTTGTTGGACGAT GAAATGGAATAGCCCC	IS*150*-specific spacer, negative strand
IS1;IS5;IS3-specific cassette	CATGGTCTCTAAACATGCTGCCAAC TTACTGATTTAGTGTATGAGTTTTAG AGCTATGCTGTTTTGAATGGTCCCA AAACGGCTCCAGATGACAAACATGA TCTCATATCGTTTTAGAGCTATGCTG TTTTGAATGGTCCCAAAACACGCGG CTAAGTGAGTAAACTCTCAGTCAGG TTTTAGAGCTATGCTGTTTTGAATG GTCCCAAAACGGGGCTATTCCATTT CATCGTCCAACAAAAGTTTTAGAGC TATGCTGTTTTGAATGGTCCCAAAA CTGAGACCGTT	Synthetic gene cassette engoding CRSIPR spacers targeting IS*1*, IS*5* and IS*3* for cloning into pCRISPathBrick

Table S2. Genes significantly down-regulated by the presence of pCRIS

	logFC[1]	AveExpr[2]	t	P Value	adjusted P Value	B	fold change	gene name	explanation of downregulation
b0022	-2.14494	5.679112	-30.97	1.4E-10	7.86E-08	15.01269	4.42	*InsA*	**IS1A**
b0259	-4.73737	9.056314	-85.72	1.3E-14	5.84E-11	22.55554	26.67	*InsH*	**IS5**
b0373	-8.50954	-0.27977	-8.684	1.02E-05	0.000712	1.824746	364.44	*InsE*	**IS3**
b0450	-1.01268	4.257646	-5.04	0.000664	0.01572	-0.75015	2.02	*glnK*	
b0553	-1.82978	8.296067	-20.99	4.69E-09	1.62E-06	11.51646	3.55	*nmpC*	IS5 downstream of gene
b0652	-3.16056	4.52066	-23.4	1.77E-09	7.95E-07	12.34931	8.94	*gtlL*	IS5 upstream of *gtl* operon
b0653	-2.86822	3.698821	-22.06	3E-09	1.23E-06	11.57504	7.30	*gtlK*	IS5 upstream of *gtl* operon
b0654	-3.20827	3.96357	-21.22	4.26E-09	1.6E-06	11.34508	9.24	*gtlJ*	IS5 upstream of *gtl* operon
b0655	-4.18404	6.446123	-65.45	1.53E-13	1.53E-10	20.12965	18.18	*gtlI*	IS5 upstream of *gtl* operon
b0987	-1.08532	1.827631	-5.97	0.000196	0.005965	1.110339	2.12	*gfcA*	lying back-to-back with IS1D
b1243	-3.43876	10.08949	-69.72	8.58E-14	1.53E-10	21.46587	10.84	*oppA*	IS5 upstream of opp operon
b1244	-3.60124	6.144254	-59.9	3.44E-13	2.58E-10	19.6422	12.14	*oppB*	IS5 upstream of opp operon
b1245	-3.50465	6.772219	-58.05	4.58E-13	2.94E-10	19.82969	11.35	*oppC*	IS5 upstream of opp operon
b1246	-3.4995	7.152826	-67.14	1.21E-13	1.53E-10	20.87308	11.31	*oppD*	IS5 upstream of opp operon
b1247	-3.38441	7.279867	-64.67	1.7E-13	1.53E-10	20.71187	10.44	*oppF*	IS5 upstream of opp operon
b2028	-1.58354	4.514402	-20.16	6.75E-09	2.17E-06	11.23268	3.00	*ugd*	IS5 upstream
b2029	-1.49571	9.289028	-29.69	2.05E-10	1.02E-07	14.72943	2.82	*gnd*	IS5 upstream
b2725	-1.3485	3.084375	-6.623	8.94E-05	0.003467	1.702269	2.55	*hycA*	
b4294	-3.46957	2.565652	-11.85	7.35E-07	0.00011	6.367982	11.08	*insA*	**IS1F**
b4570	-1.33887	3.727483	-11.28	1.12E-06	0.000152	6.12785	2.53	*lomR*	IS5 interrupting
b4571	-1.54539	6.131028	-14.78	1.07E-07	2.82E-05	8.297627	2.92	*wbbL*	IS5 interrupting

[1] logFC: log2 of fold change in read frequencies
[2] AveExpr: Average of expression levels

Table S3. Genes significantly up-regulated by the presence of pCRIS

	logFC[1]	AveExpr[2]	t	P Value	adjusted P Value	B	fold change	gene name	IS in vicinity
b0297	1.037752	0.886165	4.341705	0.0018	0.030644	-0.96538	2.05303	*eaeH*	directly upstream of IS3
b1319	1.044211	1.168833	4.128482	0.002474	0.037844	-1.31477	2.06224	*ompG*	
b1409	1.064929	0.913986	4.55035	0.001326	0.02486	-0.67252	2.09207	*ynbB*	IS2, IS30 close in Rac prophage
b1503	1.159621	0.612153	4.375767	0.001711	0.029723	-0.89045	2.23399	*ydeR*	
b2055	1.018434	0.444058	3.879706	0.003615	0.047455	-1.58839	2.02572	*wcaE*	26 kbp from IS5
b2357	1.835212	-0.01028	4.89368	0.000814	0.017634	-0.20043	3.56824	*yfdT*	
b3579	1.099776	1.467598	4.643391	0.00116	0.022687	-0.61056	2.14321	*yiaO*	
b3581	1.005112	0.948833	4.204709	0.002206	0.034748	-1.17104	2.0071	*sgbH*	
b4354	1.725097	7.904521	7.834528	2.37E-05	0.00127	2.348302	3.30602	*lgoR*	
b4675	2.316451	-3.36962	4.009705	0.002962	0.041979	-1.73829	4.98105	*yoaJ*	

[1] logFC: log2 of fold change in read frequencies
[2] AveExpr: Average of expression levels

Table S4. Mutations identified in the *km-gfp* cassette of the pBDP_Km_GFP plasmid propagated in *E. coli* JM107MA2

	pCRISPath	pCRIS
point mutation	9	14
insertion	**21**	**13**
deletion	2	5
total	32	32

X^2 test:
p<0.2
p<0.05
p<0.3

IS-specific PCR identified all insertion mutants as IS*2* insertions into the *km-gfp* cassette.

Supplementary Figures

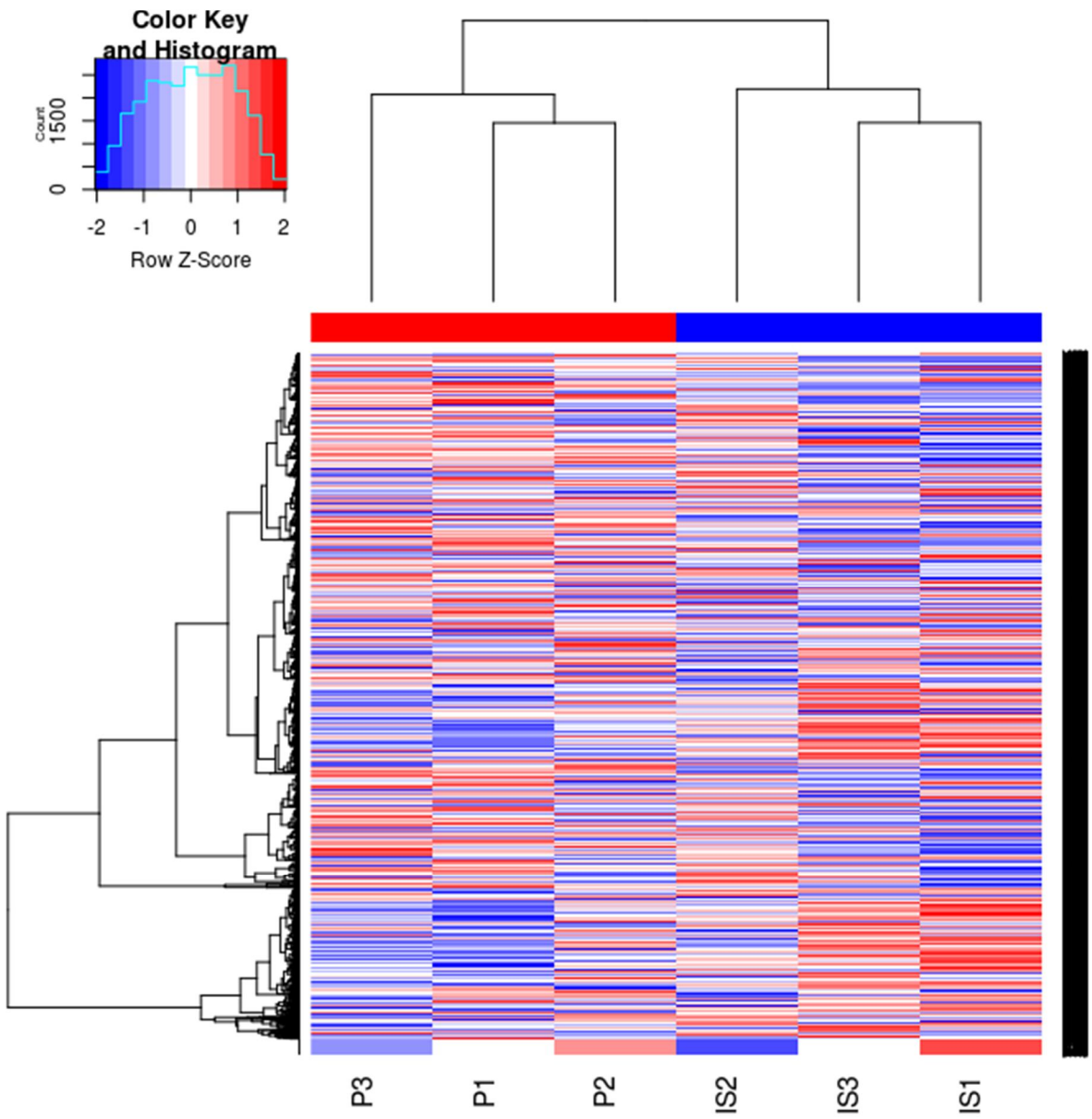

Figure S1. Heat map illustrating the expression level of each annotated gene of *E. coli* MG1655 carrying the control plasmid (samples P1-3) or the IS-silencing plasmid (samples IS1-3).

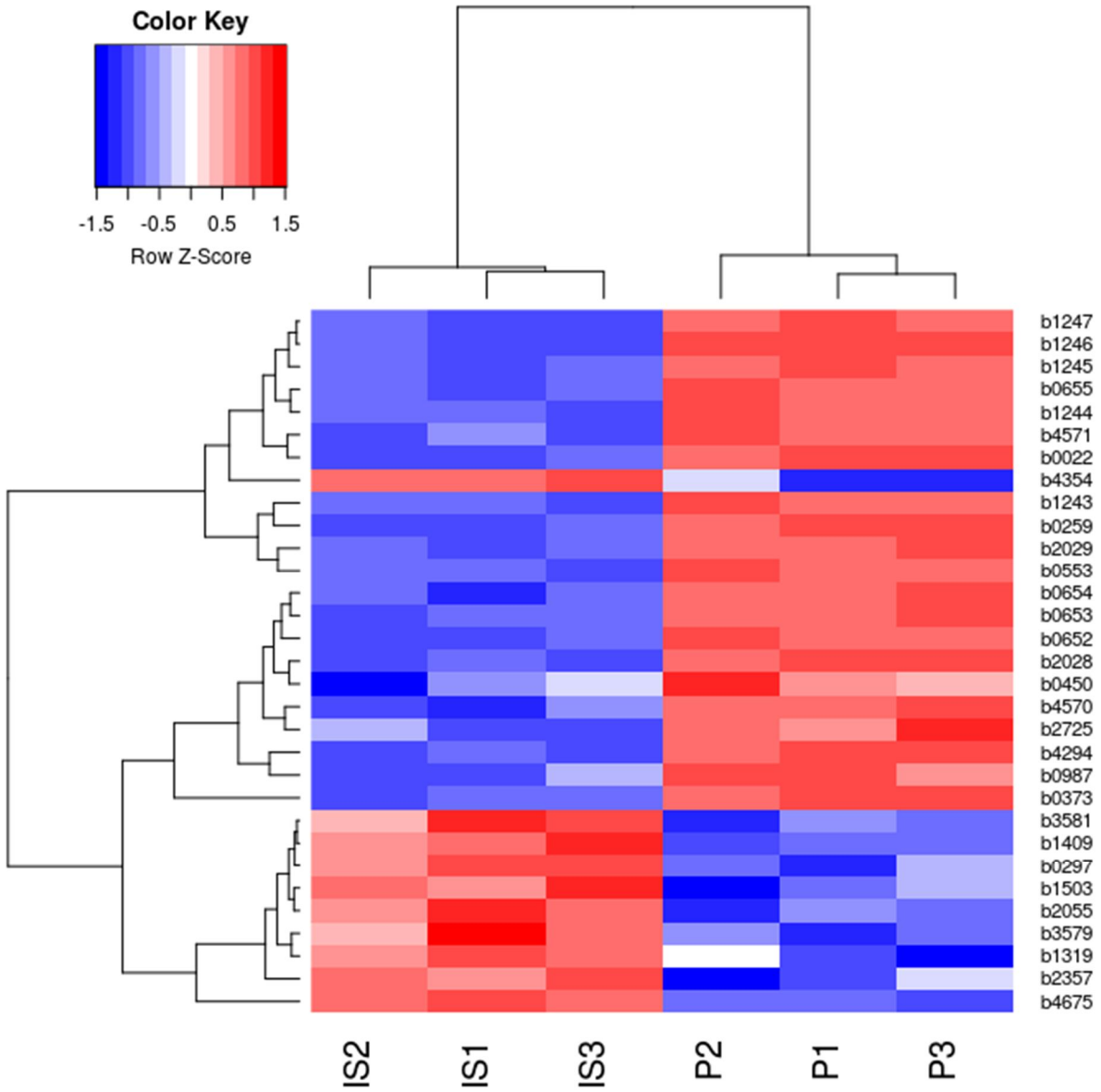

Figure S2. Heat map illustrating the expression level of differentially expressed genes (>2-fold change, p<0.05) of *E. coli* MG1655 carrying the control plasmid (samples P1-3) or the IS-silencing plasmid (samples IS1-3).

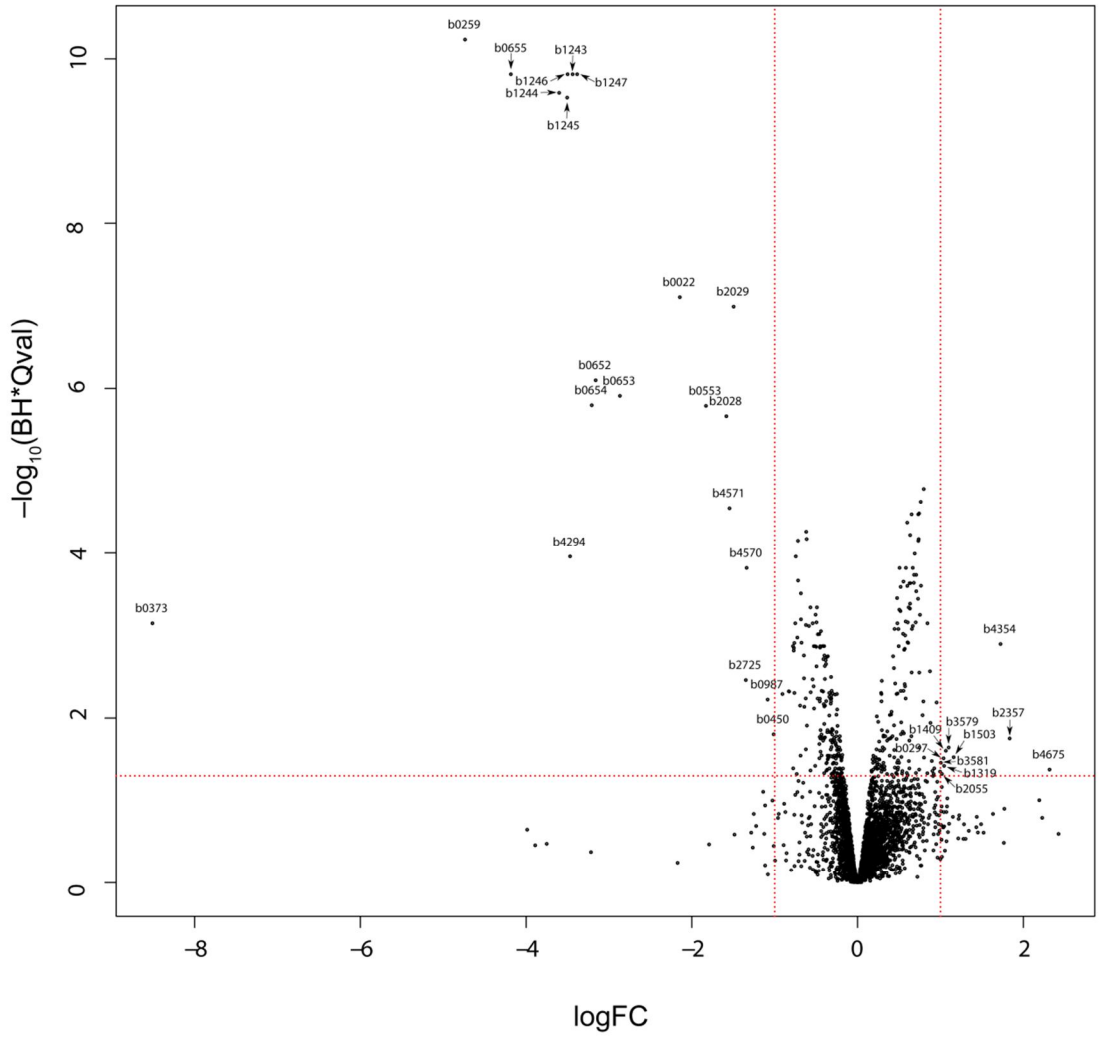

Figure S3. Volcano plot illustrating the differences detected in the expression levels of each annotated gene of *E. coli* MG1655 when exchanging the pCRISPath plasmid for pCRIS. Each dot represents the data for a single gene. The x-axis marks the $\log_2$ values of the fold-changes in read frequencies, the y-axis marks the result of the statistical analysis of the changes seen in expression ($-\log_{10}$(BH*Qval)). Genes significantly downregulated (upper left section, demarked by red dotted lines) or significantly upregulated (upper right section, demarked by red dotted lines) are identified with their respective b-numbers.

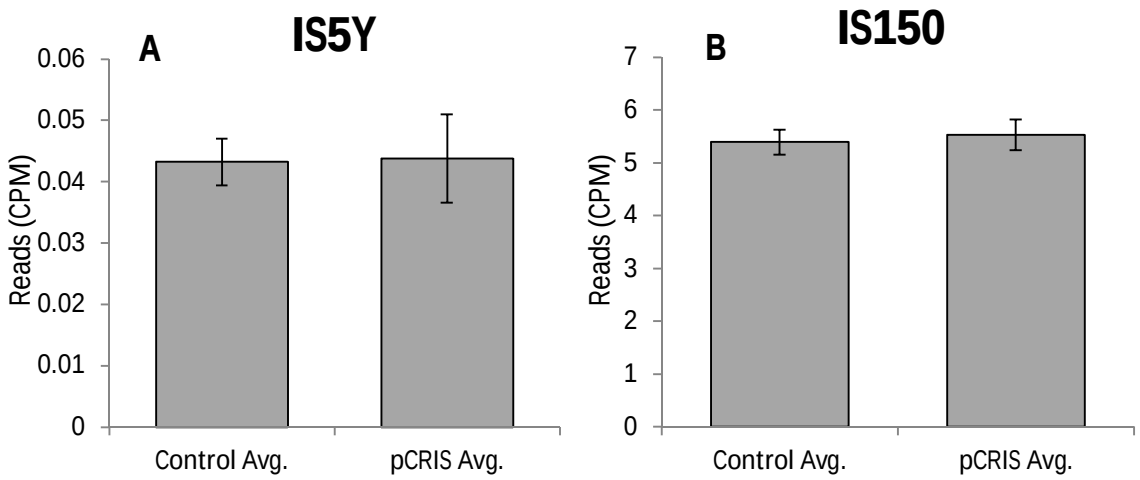

Figure S4. Transcriptional changes of IS*5*Y (A) and IS*150* (B) in *E. coli* K-12 MG1655 caused by the propagation of the pCRIS plasmid. The control was obtained by propagating the pCRISPath plasmid. The figure displays the frequency of Illumina sequencing reads mapping to the given target. Error bars represent SD, *n* = 3. CPM: counts per million reads

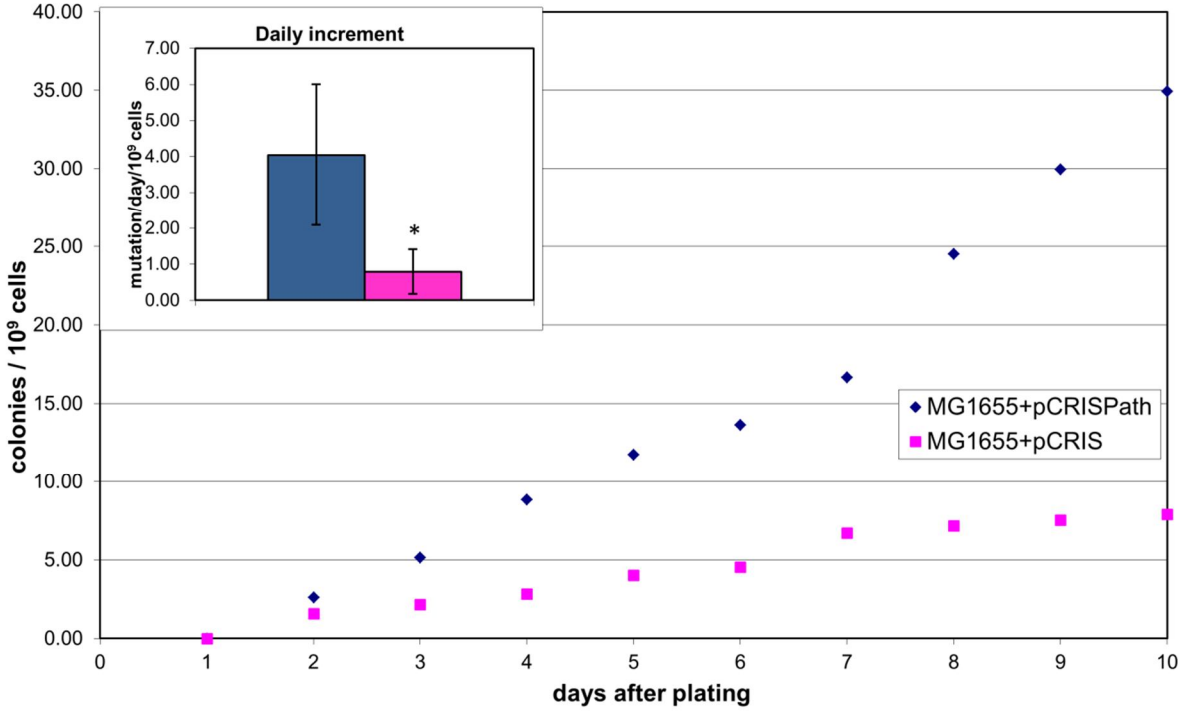

Figure S5. Effect of IS-knockdown on mutations activating the *bgl* operon. The accumulation of salicin-assimilating cells in *E. coli* MG1655 carrying the control plasmid pCRISPath (blue diamonds) or the IS-silencing plasmid pCRISP_ISF (magenta squares) is depicted. The inset shows the mean daily increments of the two strains calculated from the same dataset. * $p<10^{-3}$ using a t-test.

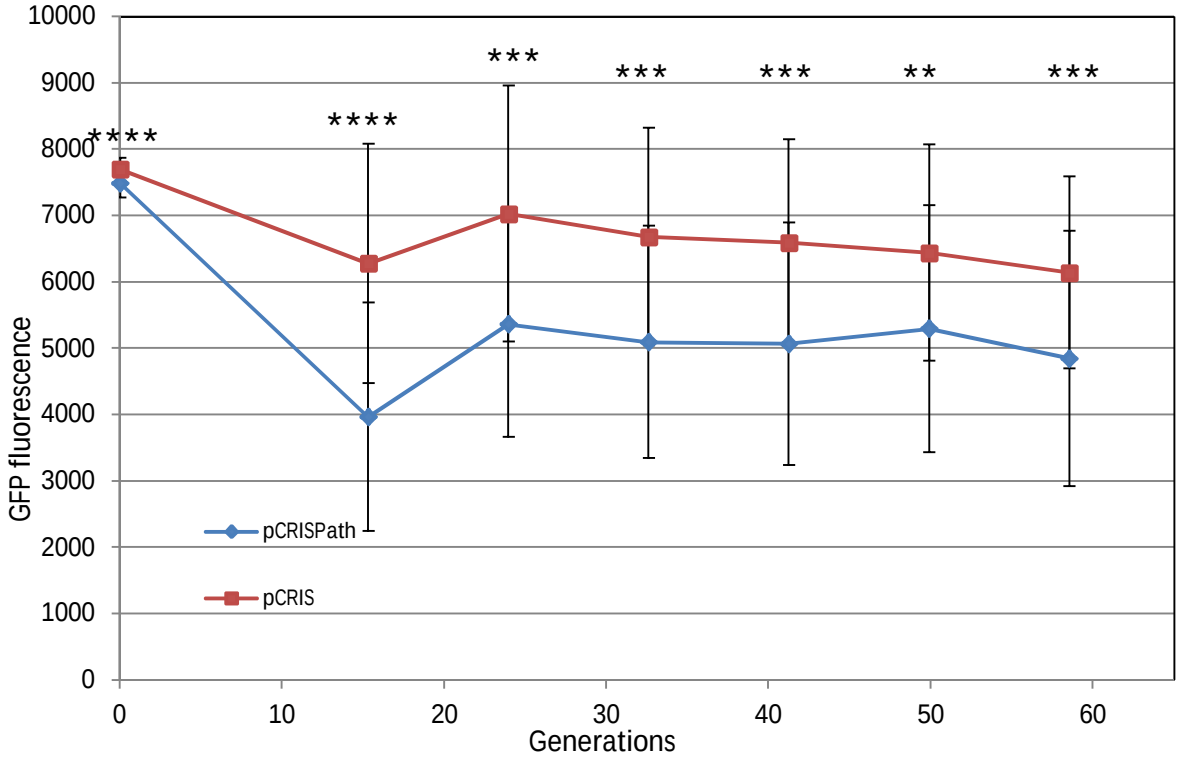

Figure S6. Evolutionary Experiment measuring plasmid stability in *E. coli* DH5αZ1. 48-48 cultures of *E. coli* DH5αZ1 carrying the control plasmid (blue diamonds) or the IS-silencing plasmid (red squares) were grown in a microplate. Green fluorescence was monitored to obtain the peak fluorescence values for each well. The means of the peak fluorescence values were calculated for each cell line and were compared for the two lines using unpaired, two-tailed t-tests. Levels of significance: *$p<5x10^{-2}$, **$p<5x10^{-3}$, ***$p<10^{-3}$, ****$p<10^{-4}$. Error bars represent SD, n=48.

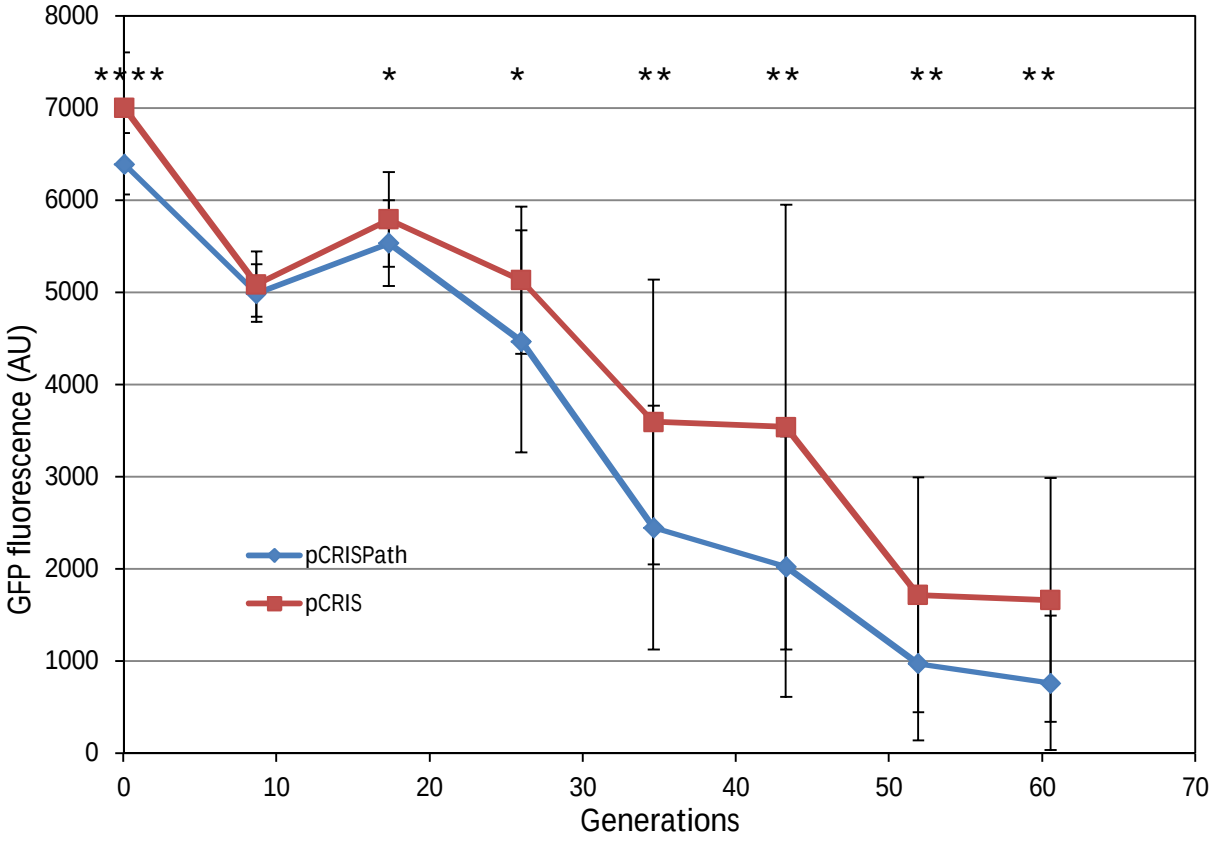

Figure S7. Evolutionary Experiment measuring plasmid stability in *E. coli* JM107MA2. 32-32 cultures of *E. coli* JM107MA2 carrying the control plasmid (blue diamonds) or the IS-silencing plasmid (red squares) were grown in a microplate. Green fluorescence was monitored to obtain the peak fluorescence values for each well. The means of the peak fluorescence values were calculated for each cell line and were compared for the two lines using unpaired, two-tailed t-tests. Levels of significance: *$p<5x10^{-2}$, **$p<10^{-2}$, ****$p<10^{-4}$. Error bars represent SD, n=32.

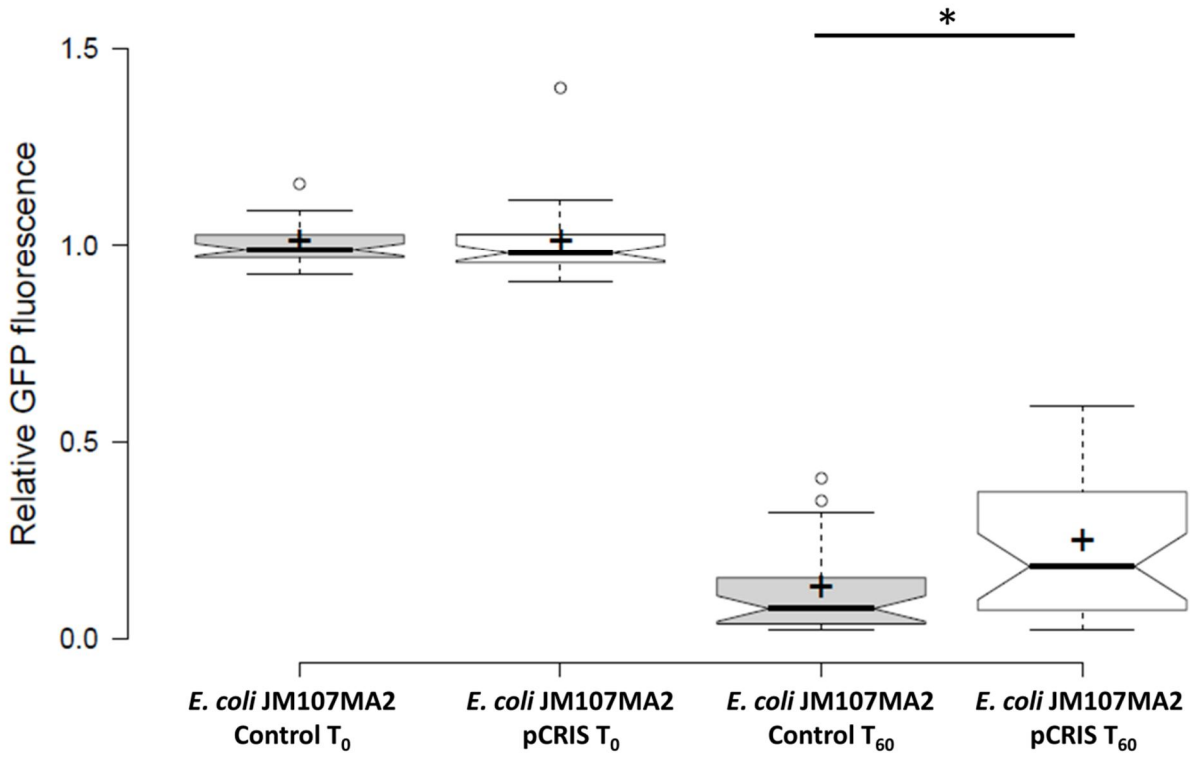

Figure S8. The effect of IS-silencing on plasmid stability in *E. coli* JM107MA2. The peak fluorescence distributions of 32 parallel cultures of *E. coli* JM107MA2 carrying the control plasmid and 32 cultures carrying the IS-knockdown plasmid are compared on day 2 and day 9 of Evolutionary Experiment 3, normalized by the respective mean values of day 2. Center lines show the medians; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, outliers are represented by dots; crosses represent sample means. n=32 sample points. * marks the significant difference of the means ($p<0.05$) using a Mann-Whitney U-test.

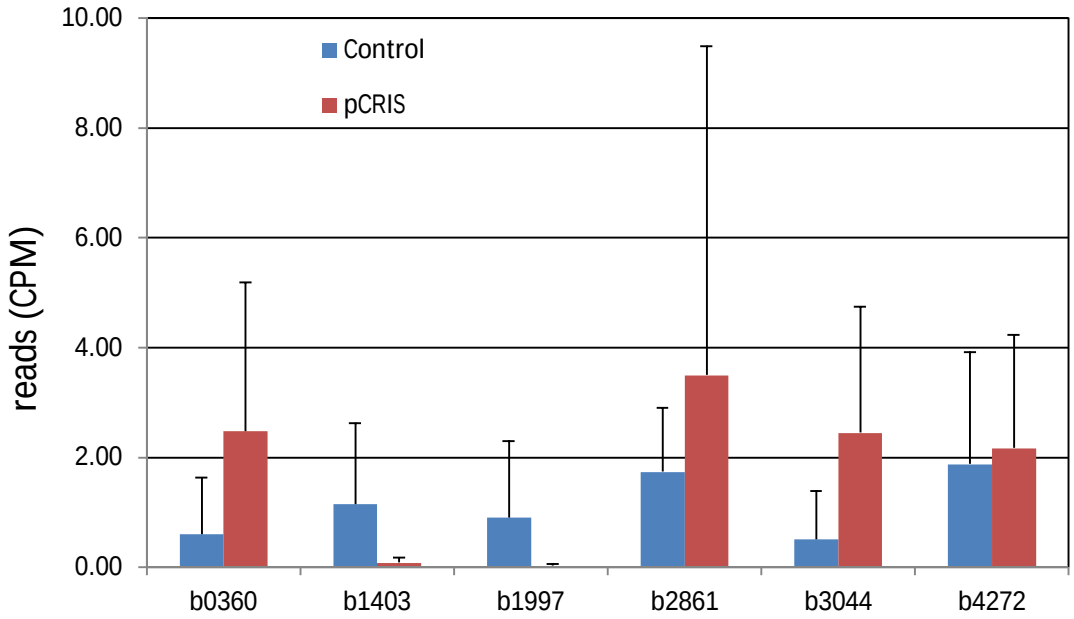

Figure S9. Expression of the *insC* gene of IS*2* elements in *E. coli* MG1655. The x axis shows the *insC* genes of distinct IS*2* elements of *E. coli* MG1655, represented by their b-numbers, the y axis displays their expression in the presence of pCRISPath (control) or the presence of the IS-silencing plasmid (pCRIS). Error bars represent SD values of 3 parallel measurements.

A

Predicted promoter for IS*1*A, IS*1*B, IS*1*D:

GGTGATGCTGCCAACTTACTGATTTAGTGTATGA**TGG**TGT

-35 -10

Predicted promoter for IS*1*F:

GGTAATGCTGCCAACTTACTGATTTAGTGTATGA**TGG**TGA

-35 -10

B

Predicted promoter for IS*5*:

AAGTC**CCT**GATATGAGATCATGTTTGTCATCTGGAGCCATA

-10

Predicted promoter for IS*5*Y:

AAGTT**CCT**GATATGAGATCATCATATTCATCCGGAGCGCAT

-10

Figure S10. Predicted promoters of IS*1* and IS*5* targeted by CRISPRi. (A) The promoter of IS*1*A and IS*1*F, (B) the promoter of IS*5* and IS*5*Y. The -10 and -35 boxes are underlined. Sequence polymorphisms are indicated in red. Green boxes mark the segments targeted by the crRNA, with the PAM sequences indicated in bold.

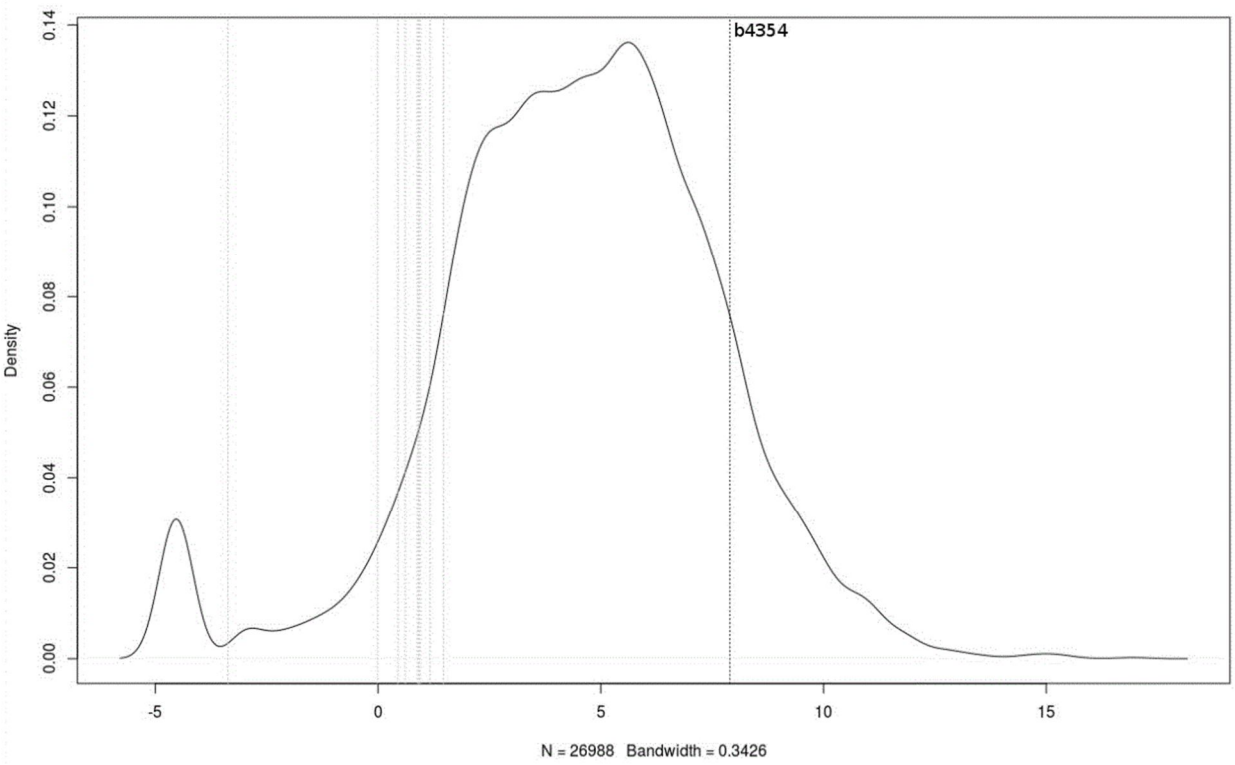

Figure S11. Distribution of expression levels for the gene set of *E. coli* K-12 MG1655, as observed in the RNA-sequencing analysis of this study. Vertical dotted lines mark the genes displaying increased expression in the presence of pCRIS. The expression level of b4354 is specifically marked.

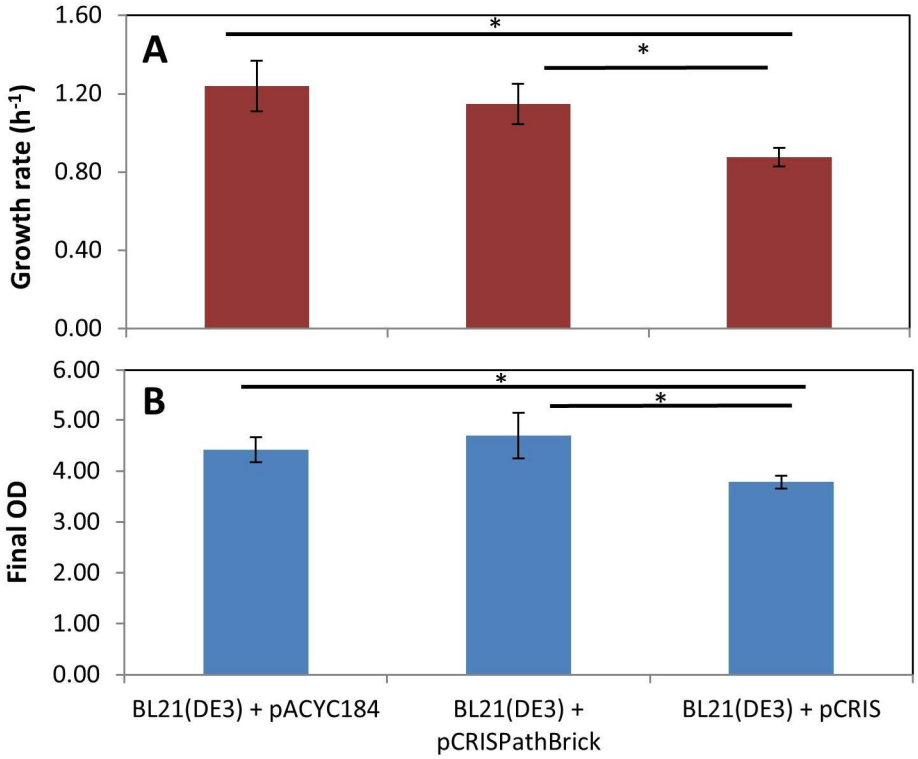

Figure S12. Effect of pCRIS on the growth of *E. coli* in a flask. Growth parameters of *E. coli* BL21(DE3) carrying plasmid pACYC184, pCRISPathBrick or pCRIS, observed in 100 ml of LB+Cm medium at 37 °C, in a flask and shaker system are shown. (A) growth rate; (B) final optical density, measured at 550 nm. Error bars mark SD. * marks the significant difference of the means ($p<0.05$) using an unpaired, two-tailed t-test (n=3).

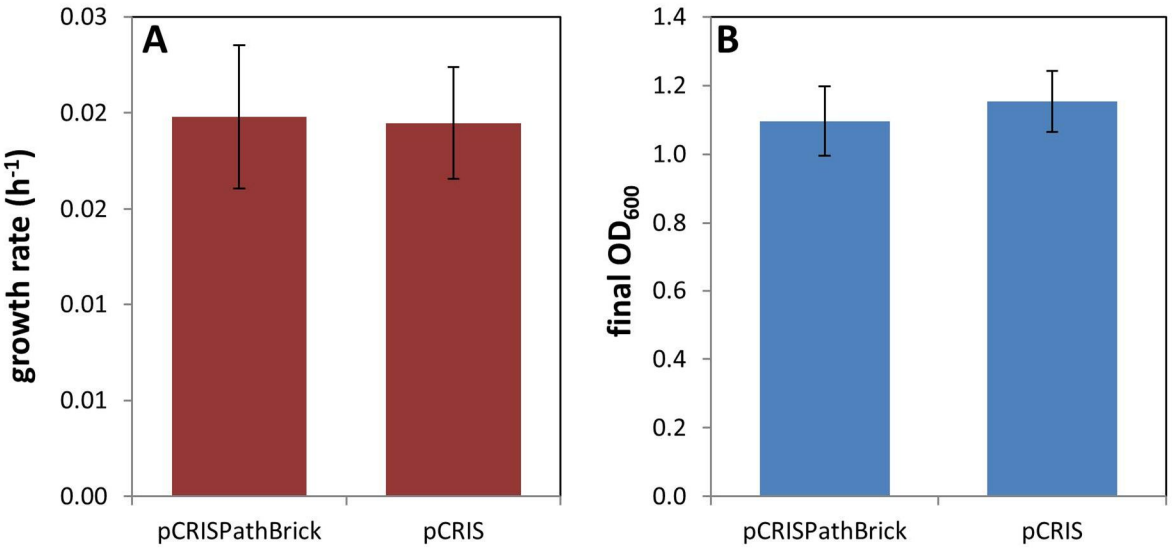

Figure S13. Effect of pCRIS on the growth of *E. coli* in microplate wells. Growth parameters of *E. coli* JM107MA2 carrying pBDP_Km_GFP5 and either the pCRISPathBrick or the pCRIS plasmid are shown, measured in 200 ul of LB+Cm+Km+IPTG medium at 37 ºC, in a microplate shaker/incubator system. (A) growth rate; (B) final optical density, measured at 600 nm. Error bars mark SD (n=32).

Supplementary Note I. Transcriptomics analysis

Regarding the transcriptomics dataset, we were most interested in the specificity of the gene-silencing activity of pCRIS. We divided the specificity issue to two phenomena: transcriptional repression of IS subtypes, and transcriptional modulation of non-IS genes.

A) Transcriptional repression of IS subtypes

The following IS elements were targeted in *E. coli* K-12 MG1655: four copies of IS*1*A (GenBank Acc. No.: X52534), two of IS*1*B (X17345), one each of IS*1*D (X52536) and IS*1*F (X52538), five of IS*3* (X02311), eleven of IS*5* (J01735), and one each of IS*5*Y (HM000058.1) and IS*150* (X07037). A custom-masked *E. coli* K12 MG1655 genome version (based on U00096.3, harbouring only one copy from each relevant IS group: IS*1*A, IS*1*F, IS*3*, IS*5*, IS*5*Y, and IS*150*) was used to map RNA-seq reads.This approach generated a direct readout of these six IS groups as six pools, besides leaving the possibility to investigate the transcription of each further ORF. The active IS*1* copies (IS*1*A, IS*1*B, IS*1*D) were pooled to a single group this way due to their high level of identity (>98 %). The five copies of IS*3* as well as the 11 copies of IS*5* are >99 % identical and binned to the IS*3* and IS*5* groups, respectively. IS*5*Y, however, displays a mere 89.8 % identity with IS*5* and was therefore treated separately. IS*1*F, presumed to be inactive, has only a ~90 % identity with the IS*1*A group, and was also mapped separately. **Figure 1** indicates that the transcription of both the IS*1*A group and IS*1*F were effectively silenced. This was expected despite their mere 90% sequence identity due to the highly conserved promoter. Note that the CRISPR spacer targets completely identical segments of the two elements (**Figure S10A**). IS*5*Y, however, was not silenced (**Figure S4**). The promoter segment of IS*5*Y displayed 6 mismatches with the CRISPR spacer (**Figure S10B**). Albeit these mismatches were at the PAM-distal end of the spacer, they most likely inhibited efficient binding of the dCas machinery to this element. The extremely small transcription level of this IS element (**Figure S4**) impedes the proper assessment of its silencing, and strongly reduces its importance, as well.

The case of IS*150* is more puzzling: although the dramatic reduction was seen in its activity in the presence of pCRIS (see main text), this IS generated an equally high number of transcripts in both the control and the IS-knockdown samples. The unaffected mRNA levels indicate that the P1 promoter of IS*150*, targeted in this work has a negligible role in the transcription of the respective transposase. This may suggest that the P2 promoter, lying downstream of P1 is the functional promoter of IS*150*, despite its overlap with the start codon (1).

B) *Transcriptional modulation of non-IS genes*

The transcriptomics analysis revealed 31 genes that displayed a statistically significant and at least two-fold rate of downregulation in the latter strain (**Table S2, S3, Figure S3**). Four of these genes are the coding regions of IS elements themselves, as indicated in **Table S2**. Of the remaining 17 downregulated genes, 14 lie in the proximity of an IS*5* and one is adjacent to an IS*1*D, their silencing can therefore be regarded as cis-effects on transcription. In two cases, (*gtl* and *opp*) entire operons were downregulated due to the silencing of the upstream IS*5*. The two remaining silenced genes encode a nitrogen-regulatory protein (*glnK*) and a transcription factor (*hycA*). The downregulation of the latter can potentially be explained by a 9 bp segment (AAGTTGGCA) lying 65 bp upstream of its start codon showing identity to the IS*1*-specific spacer, but no such homology could be found in the vicinity of the former. Even the 9 bp match seems to be too short to explain dCas binding, based on earlier studies (2, 3).

In addition to downregulation, pCRIS caused the significant upregulation of ten genes listed in **Table S3**. OmpG is an outer membrane porin, YiaO is a 2,3-diketo-L-gulonate:Na+ symporter - periplasmic binding protein and SgbH is 3-keto-L-gulonate-6-phosphate decarboxylase. The seven remaining genes (*eaeH, ynbB, ydeR, wcaE, yfdT, lgoR, yoaJ*) are assigned with only putative functions yet. The most plausible explanation for gene upregulation by dCas9 expression would be the indirect effect of silencing a repressor, such as *hycA*. None of the upregulated genes however, are known to be repressed by *hycA*. Perhaps more subtle alterations of other transcription factors are responsible for these changes. One also cannot exclude the direct positive effect of dCas binding on gene transcription by modifying the molecular milieu or the accessibility of the promoter in a way that mediates RNA synthesis. Such a cis-effect may come into play for *eaeH*, which lies directly upstream of an IS*3.* The binding of dCas may also influence distal gene expression by trans-regulatory effects resulting from the 3D structural properties of prokaryotic chromosomes, as assumed for numerous transcription factors based on the periodicity of their targets (4). It is also possible, that the binding of a novel proteinaceous factor (dCas) to the genome of *E. coli* at 26 loci initiates a general and yet unknown response, embodied by the induction of some of these genes. Finally, the most plausible cause for the seeming overexpression of these genes is experimental noise. It is apparent from the volcano plot (**Figure S3**) that apart from b4354, all overexpressed genes lie close to the limit of statistical significance. Seven of these barely surpass the thresholds of 2-fold change in expression and 0.05 of FDR (false discovery rate). The corresponding region on the volcano plot among the downregulated genes is practically empty, for those genes mostly display a higher fold change in expression with a higher level of significance. Within this latter

set, the two genes non-proximal to ISes (*glnK* and *hycA*) are nevertheless also among the three least repressed genes with the lowest significance (b0450 and b2725 on the plot). In addition, observing the mean expression level of the upregulated genes (**Figure S11**), it is apparent that these genes are extremely weakly expressed, again with the exception of b4354. Their expression levels fall close to the RNA-Seq detection limit being largely influenced by technical noise. After manually revising the differential gene expression hit lists we consider most of the genes non-proximal to ISes as technical false positive hits.

Overall, out of the 28 ISes targeted among the 4286 genes of *E. coli*, 27 was successfully downregulated (the single copy of IS150 was the exception). Out of the 31 non-IS genes that displayed an altered expression, only 12 could not be easily explained by the cis-effects of dCas-binding. Therefore, less than 0.3% of the *E. coli* gene set seems to have responded in an unexplained manner to pCRIS, which in our interpretation is the characteristic of very high specificity.

Supplementary Note II. Inhibition of IS-transposition

Evaluation of the effects of pCRIS on the mRNA levels and the transposition activity of IS elements shed light on two interesting phenomena. First, it became obvious from our results that transcriptional downregulation is not a prerequisite of IS-silencing. Namely, pCRIS nearly completely eliminated the mobility of IS*150* despite its unchanged mRNA levels. We speculate that the mere steric inhibition of transposase access to the left IR of the IS was sufficient to inhibit transposition. Further experiments (e.g. targeting the right IR of ISes) would be needed to verify this assumption.

Second, our most surprising finding was the significant downregulation of IS*2* transposition to the *km-gfp* cassette of pBDP_Km_GFP5, despite the lack of CRISPR spacers targeting the IS*2* element. We hypothesized that one or more of the crRNA species expressed by pCRIS may bind to IS*2*, and inhibit its expression. This explanation was discarded, since the mRNA levels of the IS*2* transposases displayed no consistent change upon pCRIS transformation (**Figure S9**). We have seen for IS*150* however, that IS mobility can be repressed without trasncriptional silencing, we therefore searched for possible base-pairing between IS*2* and any of the four CRISPR spacers of pCRIS. The longest match found was a 9 bp identity between the IS*1*-specific spacer and the middle (600-608 nt coordinates) of IS*2* (GCTGCCAAC). Based on earlier studies (2, 3), we strongly question whether a spacer of this length, without a proper PAM sequence could explain the potential binding of dCas9 to IS*2*. We

therefore turned to our third hypothesis, attributing our observations to a novel phenomenon we refer to as “transposase cross-reactivity”. Namely, IS*2*, IS*3* and IS*150* all belong to the IS*3* family of ISes. If we assume that their transposases cross-react with each other, the successful transcriptional repression of IS*3* could lead to a lessened mobility of IS*2*, and perhaps IS*150* as well. Further experimental evidence supporting this hypothesis is being collected at this time, and will be published in a separate study.

[1] Schwartz, E., Kroger, M., and Rak, B. (1988) IS150: distribution, nucleotide sequence and phylogenetic relationships of a new *E. coli* insertion element. *Nucleic Acids Res.*, 16, 6789-6802.

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

[3] Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L.A. (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.*, 31, 233-239.

[4] Bouyioukos, C., Elati, M., and Képès, F. (2016) Analysis tools for the interplay between genome layout and regulation. *BMC Bioinformatics*, 17, 191.