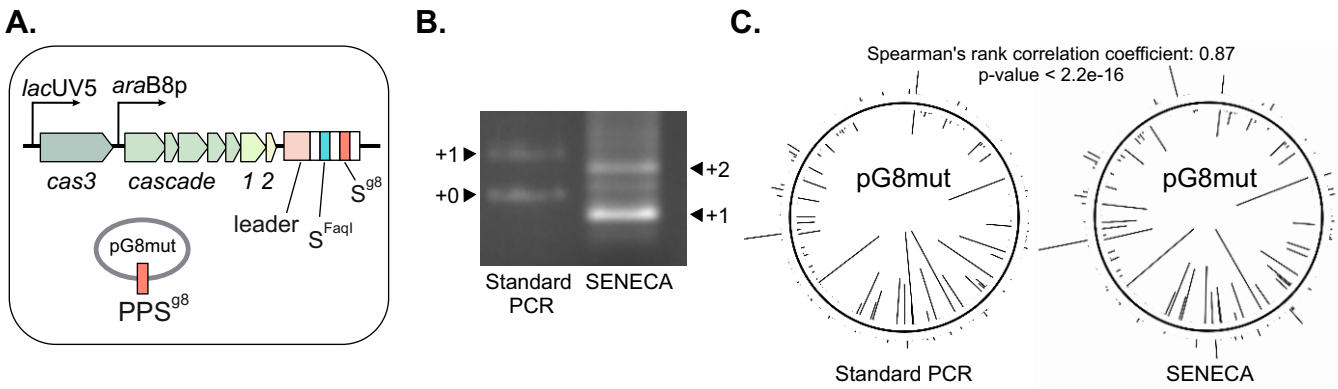
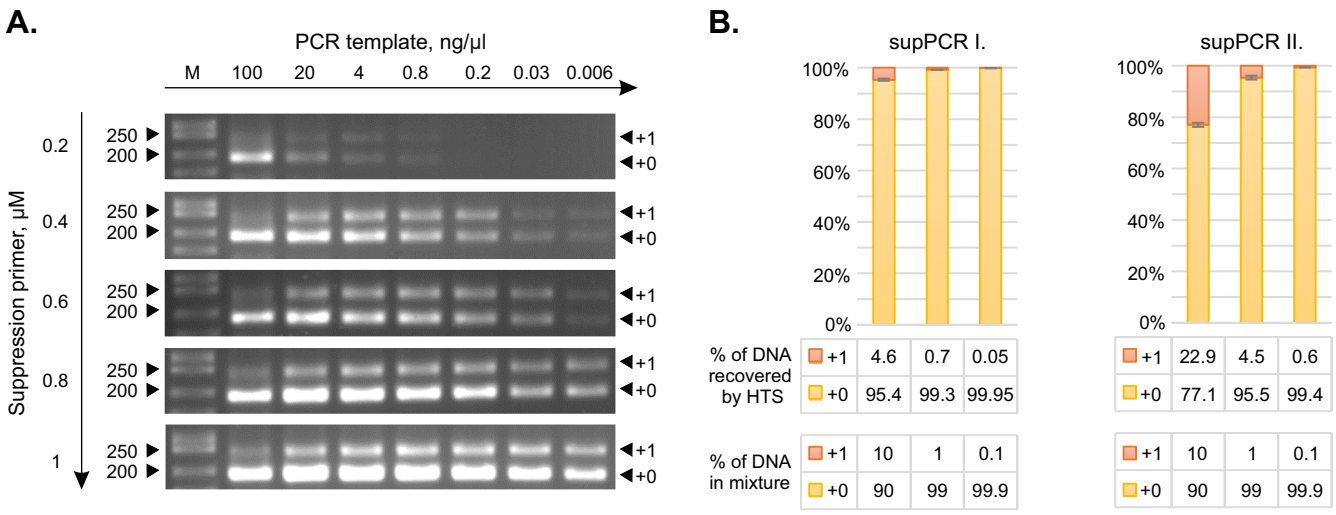


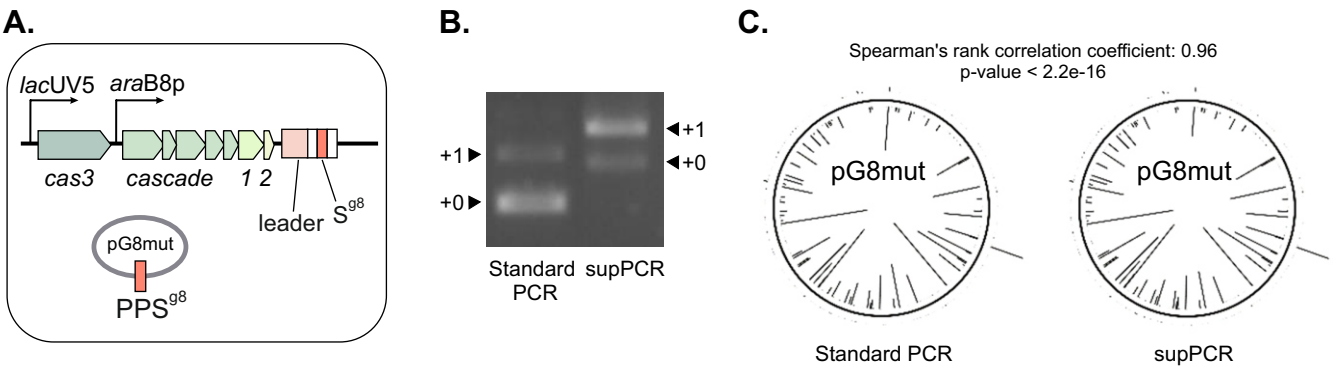
Supplementary Figure 1. Analysis of primed adaptation by SENECA reveals no biases in amplification of individual spacers compared with standard PCR.

A. Schematic representation of an *E. coli* strain KD263_FaqI cell used to monitor primed adaptation by SENECA. The *lac*UV5 and *ara*B8p promoters direct inducible transcription of *cas* genes (shown as block arrows). The engineered CRISPR array contains the leader (light peach rectangle), three CRISPR repeats (white rectangles), spacer S^{FaqI} with an FaqI recognition site (turquoise rectangle), and spacer S^{g8} (dark peach rectangle) matching priming protospacer PPS^{g8} (dark peach rectangle) of plasmid pG8mut. **B.** Agarose gel electrophoresis analysis of primed adaptation in induced KD263_FaqI culture by standard PCR and SENECA. Amplicons corresponding to unexpanded and expanded arrays revealed by standard PCR (left lane) are labeled +0 and +1, respectively. Amplicons corresponding to arrays expanded by one or two spacers amplified by SENECA (right lane) are labeled +1 and +2. **C.** HTS of acquired spacers revealed by standard PCR and SENECA. +1 amplicons from panel **B** were subjected to HTS. Spacers were extracted from reads and mapped to the pG8mut plasmid. Each bar represents an individual protospacer. Bars oriented inside or outside the plasmid circles represent protospacers located on different DNA strands. The bar heights are proportional to the number of HTS reads containing corresponding spacers. Spearman's rank correlation coefficient between the numbers of spacers obtained by standard protocol and SENECA is shown above.



Supplementary Figure 2. Optimizing supPCR conditions to obtain increased sensitivity of CRISPR adaptation detection.

A. Genomic DNA purified from an *E. coli* strain with one additional spacer (+1) was mixed with DNA from a parental strain with unexpanded (+0) array in the ratio of 10% to 90% and used as a template for the first stage of supPCR. The products were diluted to concentrations indicated at the top and used as a template for the second stage supPCR in the presence of a suppression primer at concentrations indicated on the left. Products of the second stage supPCR corresponding to unexpanded (+0) and expanded (+1) arrays separated by agarose gel electrophoresis are shown. M, molecular-weight size marker. **B.** Genomic DNA purified from an *E. coli* strain with one additional spacer (+1) was mixed with DNA from a strain with unexpanded (+0) array in proportions indicated at the bottom. The mixtures were used as a template for the first and second stages of supPCR ("supPCR I" and "supPCR II") followed by high-throughput sequencing. Percentages of recovered reads corresponding to unexpanded and expanded CRISPR arrays are indicated in bar plots and tables immediately below.



Supplementary Figure 3. Analysis of primed adaptation by supPCR reveals no biases in amplification of individual spacers compared with standard PCR.

A. Schematic representation of an *E. coli* strain KD263 cell used to monitor primed adaptation by supPCR. The *lac*UV5 and *ara*B8p promoters direct inducible transcription of *cas* genes (shown as block arrows). The engineered CRISPR array contains the leader (light peach rectangle), two CRISPR repeats (white rectangles), and spacer S^{g8} (dark peach rectangle) matching priming protospacer PPS^{g8} (dark peach rectangle) of plasmid pG8mut. **B.** Agarose gel electrophoresis analysis of primed adaptation in induced KD263 culture by standard PCR and supPCR. Amplicons corresponding to unexpanded and expanded arrays are labeled +0 and +1, respectively. **C.** HTS of acquired spacers revealed by standard PCR and supPCR. The +1 amplicons from panel **B** were subjected to HTS. See Supplementary Figure 1 panel C legend for details.

Supplementary Methods

Bacterial strains and plasmids

The *E. coli* strains used in this study are listed in Supplementary Table 1. The KD263_FaqI strain is a derivative of KD263 strain [1] with an additional spacer S^{FaqI} containing FaqI recognition site (Fig. 3A, Supplementary Fig. 1A). To obtain KD263_FaqI, primed adaptation was induced in KD263 transformed with a pG8mut_FaqI plasmid described below (Fig. 3A). After induction of primed adaptation, individual colonies that lost Amp-resistance were selected and tested by PCR with primers Ec_LDR_F and M13_G8 (Supplementary Table 2). Amplicons corresponding to arrays expanded by a single spacer-repeat unit were subjected to Sanger sequencing. Screening of ~100 clones allowed to identify several colonies with the S^{FaqI} insertion.

Plasmid pG8mut_FaqI is a derivative of pG8mut [2] with a modified hot protospacer HS1 (GCTTTCCCTATAGTGAGTCGTATTAGAGCTTGG) [3]. Modified HS1, PS^{FaqI} (GTCCCCCCTATAGTGAGTCGTATTAGAGCTTGG), contains the FaqI recognition site (underlined). To construct pG8mut_FaqI, plasmid pG8mut was amplified with primers Faq_HS1_plF and Faq_HS1_plR (Supplementary Table 2) using iProof High-Fidelity DNA Polymerase (Bio-Rad) according to the manufacturer's instructions. The resulting PCR product was purified and treated with 0.05 U/µl DpnI (Thermo Fisher Scientific). After 5'-phosphorylation with T4 polynucleotide kinase (Thermo Fisher Scientific) DNA was precipitated, diluted in 5 µl of sterile Milli-Q water and circularized with Gibson Assembly Master Mix (NEB) following manufacturer's instructions. The ligation mixture was transformed into competent *E. coli* DH5α cells and plasmids with desired sequence were selected by Sanger sequencing.

Primed adaptation assay

The same protocol was used to induce primed adaptation in KD263 transformed with pG8mut_FaqI to obtain the KD263_FaqI strain, in KD263_FaqI transformed with pG8mut to monitor spacer acquisition by SENECA, or in KD263 transformed with pG8mut to monitor spacer acquisition by supPCR.

Overnight cultures obtained from individual transformant colonies were grown in liquid LB supplemented with 100 µg/ml ampicillin, diluted 100-fold in LB without antibiotics and grown at 37 °C until culture OD_{600} reached 0.3-0.4. CRISPR interference/primed adaptation was induced by the addition of 1 mM IPTG and 1 mM L-(+)-arabinose as described [4]. After 6

hours growth cells were either plated to select the KD263_FaqI strain (above) or processed for total genomic DNA purification by GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific) and subsequent analysis of spacer acquisition by standard PCR, SENECA, or supPCR.

Detection of spacer acquisition by standard PCR

CRISPR arrays were amplified by PCR in a 25-µl reaction containing 30 ng of total genomic DNA, 2.5 U of Taq DNA Polymerase (αFerment), 1× Taq buffer (αFerment), 0.2 mM dNTPs (Thermo Fisher Scientific), 0.5 µM primer annealing to the leader sequence (Ec_LDR_F, Supplementary Table 2), and 0.5 µM primer annealing to the preexisting spacer (M13_G8 for KD263 strain or HS1_qPCR_adaptation for KD263_FaqI strain, Supplementary Table 2). The PCR cycling conditions were as follows: 3 min at 95 °C followed by 32 cycles of 15 sec at 95 °C, 30 sec at 55 °C, and 30 sec at 72 °C with a final extension of 5 min at 72 °C.

Detection of spacer acquisition by SENECA

For analysis of adaptation by SENECA, CRISPR arrays were first amplified by PCR in a 25-µl reaction containing 30 ng of total genomic DNA, 2.5 U of Taq DNA Polymerase (αFerment), 1× Taq buffer (αFerment), 0.2 mM dNTPs (Thermo Fisher Scientific), 0.5 µM Ec_LDR_F primer (anneals to the leader sequence), and 0.5 µM HS1_qPCR_adaptation primer annealing to S^{FaqI} spacer (Supplementary Table 2). The PCR cycling conditions were as follows: 3 min at 95 °C followed by 35 cycles of 15 sec at 95 °C, 30 sec at 50 °C, and 30 sec at 72 °C with a final extension of 5 min at 72 °C.

The PCR product was purified with the GeneJET PCR Purification Kit (Thermo Fisher Scientific), treated with FaqI restriction endonuclease and ligated to Illumina adapter according to [5]. Oligonucleotides FaqI_adapter_F and FaqI_adapter_R (Supplementary Table 2) were annealed by heating at 95 ˚C for 5 min and cooling the mixture to 25 ˚C to form a double-stranded Illumina adapter (FaqI_adapter) with an overhang mimicking the sticky end generated by FaqI. The cleavage of CRISPR array amplicons by FaqI and simultaneous ligation of the FaqI_adapter was performed in a 50-µl reaction containing 20 – 200 ng of DNA, 0.01 µM FaqI_adapter, 1 mM ATP (NEB), 1 mM DTT (VWR Life Science AMRESCO), 0.2 U of T4 DNA Ligase (Thermo Fisher Scientific), 1 µL of FastDigest FaqI restriction endonuclease (Thermo Fisher Scientific), and 1× FastDigest buffer (Thermo Fisher Scientific). 99 cycles of 3 min at 37 °C and 3 min at 20 °C were performed. Reaction products were purified with the GeneJET PCR Purification Kit (Thermo Fisher Scientific), resuspended in 15 µl of MQ. 3 µl

aliquots (3ng/µl) were used for amplification of CRISPR arrays ligated to the FaqI_adapter in a 100-µl reaction with 2.5 U of Taq DNA Polymerase (αFerment), 1× Taq buffer (αFerment), 0.2 mM dNTPs (Thermo Fisher Scientific), 0.5 µM primer FaqI_adapter_PRIMER annealing to the Illumina adapter, and 0.5 µM primer CRISPR_array_for annealing to CRISPR repeat not cleavable by the FaqI (Figure 2D, oligonucleotide sequences are available in Supplementary Table 2). The PCR cycling conditions were as follows: 3 min at 95 °C followed by 35 cycles of 15 sec at 95 °C, 30 sec at 52 °C, and 30 sec at 72 °C with a final extension of 5 min at 72 °C.

Detection of spacer acquisition by supPCR

For analysis of adaptation by supPCR, two consecutive PCR reactions were performed. At the first stage, CRISPR arrays were amplified by standard PCR in a 25-µl reaction containing 20 ng of total genomic DNA, 2.5 U of Taq DNA Polymerase (αFerment), 1× Taq buffer (αFerment), 0.2 mM dNTPs (Thermo Fisher Scientific), and 0.5 µM primers containing the ITR sequence on 5' ends and annealing to the leader sequence (KD263_Ist_F, Supplementary Table 2) and preexisting spacer (KD263_Ist_R, Supplementary Table 2). The PCR cycling conditions were as follows: 3 min at 95 °C followed by 20 cycles of 15 sec at 95 °C, 30 sec at 55 °C, and 30 sec at 72 °C with a final extension of 5 min at 72 °C.

The amplified DNA was diluted 100-fold and 1 µl of the mixture was used as a template for the second stage PCR reaction performed in a total volume of 20 µl and containing 1× Q5 high-fidelity polymerase master mix (NEB) and 0.4 µM suppression primer (ITR_primer, Supplementary Table 2). The PCR cycling conditions were as follows: 30 sec at 98 °C followed by 40 cycles of 30 sec at 98 °C, 30 sec at 57.3 °C, and 30 sec at 72 °C with a final extension of 2 min at 72 °C.

High-throughput sequencing of CRISPR arrays

HTS libraries were generated using NEBNext Ultra II DNA Library Prep Kit (NEB) and sequenced at the MiSeq Illumina platform in a pair-end 150-bp read mode. Raw sequencing data were analyzed using ShortRead and BioStrings packages [6,7]. Sequencing reads were filtered for quality scores of ≥20 and reads containing two repeats (with up to two mismatches) were selected. Reads that contained 33-bp sequences between two parts of CRISPR repeats (AGCGGGGATAAACCG and GTGTTCCCCGCGCC) were next selected. The 33-bp segments were considered as spacers. For reads corresponding to multiple spacer acquisition events, only spacers that were acquired first were selected for further analysis. Spacers were

next mapped onto the pG8mut plasmid with no mismatches allowed. Read mapping and spacer statistics analysis was performed with R [1]. Graphical representation was carried out using the EasyVisio tool developed by Ekaterina Rubtsova.

Supplementary Table 1. Strains used in this study

Name	Description	Source
KD263	K-12 F^+, *lacUV5-cas3 araBp8-cse1*, CRISPR I: repeat-S^{g8}-repeat, CRISPR II deleted. S^{g8} (CTGTCTTTCGCTGCTGAGGGTGACGATCCCGC) targets g8 gene of M13 phage.	Shmakov *et al.*[1]
KD263_FaqI	A derivative of KD263 with an additional spacer S^{FaqI} in CRISPR I: repeat-S^{FaqI}-repeat-S^{g8}-repeat. S^{FaqI} (GTCCCCCCTATAGTGAGTCGTATTAGAGCTTGG) contains the FaqI recognition site (underlined).	This study

Supplementary Table 2. Oligonucleotides used in this study

Name	Sequence (5' to 3')	Description
Faq_HS1_plF	CCTGCAGGCATGCAAGTCCCCCCTATAGTGAGTCGTATTA	pG8mut_FaqI construction
Faq_HS1_plR	TAATACGACTCACTATAGGGGGGACTTGCATGCCTGCAGG	
Ec_LDR_F	AAGGTTGGTGGGTTGTTTTTATGG	Anneals to the leader sequence (standard PCR; SENECA 1st stage PCR)
M13_G8	GGATCGTCACCCTCAGCAGCG	Anneals to S^{g8} spacer (standard PCR)
HS1_qPCR_adaptation	CCAAGCTCTAATACGAC	Anneals to S^{FaqI} spacer (standard PCR; SENECA 1st stage PCR)
Faq_adapter_F	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	Two strands of FaqI_adapter for SENECA
Faq_adapter_R	AGCGAGATCGGAAGAGCACACGTCTGAACTCCAGTCAC	
Faq_adapter_PRIMER	GTGACTGGAGTTCAGACG	Anneals to FaqI_adapter (SENECA 2nd stage PCR)
CRISPR_array_for	AGCGGGGATAAACCG	Anneals to CRISPR repeat (SENECA 2nd stage PCR)
KD263_Ist_F	GTAATACGACTCACTATAGGGCACGGTTGCGGAAATGTTACATTAAGGTTGG	The underlined region anneals to the leader sequence (supPCR 1st stage)
KD263_Ist_R	GTAATACGACTCACTATAGGGCACGGTTGCCCCTCAGCAGCGAAAGACAG	The underlined region anneals to S^{g8} spacer (supPCR 1st stage)
ITR_primer	GTAATACGACTCACTATAGGGC	Suppression primer (supPCR 2nd stage)

Supplementary References

1. Shmakov, S., Savitskaya, E., Semenova, E., Logacheva, M.D., Datsenko, K.A. and Severinov, K. (2014) Pervasive generation of oppositely oriented spacers during CRISPR adaptation. *Nucleic Acids Res.* **42(9)**, 5907–5916
2. Datsenko, K.A., Pougach, K., Tikhonov, A., Wanner, B.L., Severinov, K. and Semenova, E. (2012) Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat. Commun.* **3**, 945
3. Musharova, O., Sitnik, V., Vlot, M., Savitskaya, E., Datsenko, K.A., Krivoy, A. et al. (2019) Systematic analysis of Type I-E Escherichia coli CRISPR-Cas PAM sequences ability to promote interference and primed adaptation. *Mol. Microbiol.* **111(6)**, 1558–1570
4. Semenova, E., Savitskaya, E., Musharova, O., Strotskaya, A., Vorontsova, D., Datsenko, K.A. et al. (2016) Highly efficient primed spacer acquisition from targets destroyed by the Escherichia coli type I-E CRISPR-Cas interfering complex. *Proc. Natl. Acad. Sci. U.S.A.* **113(27)**, 7626–7631
5. Schmidt, F., Cherepkova, M.Y., Platt, R.J. (2018) Transcriptional recording by CRISPR spacer acquisition from RNA. *Nature.* **562(7727)**, 380–385
6. Morgan, M., Anders, S., Lawrence, M., Aboyoun, P., Pagès, H., Gentleman, R. (2009) ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics.* **25(19)**, 2607–2608
7. Pagès, H., Aboyoun, P., Gentleman, R., DebRoy, S. (2017) Biostrings: Efficient manipulation of biological strings. R package version 2.46.0.