Supplementary Material for

CRISPR repeat sequences and relative spacing specify DNA integration by *Pyrococcus furiosus* Cas1 and Cas2

Julie Grainy, Sandra Garrett, Brenton Graveley, Michael Terns

Supplementary Figure S1.

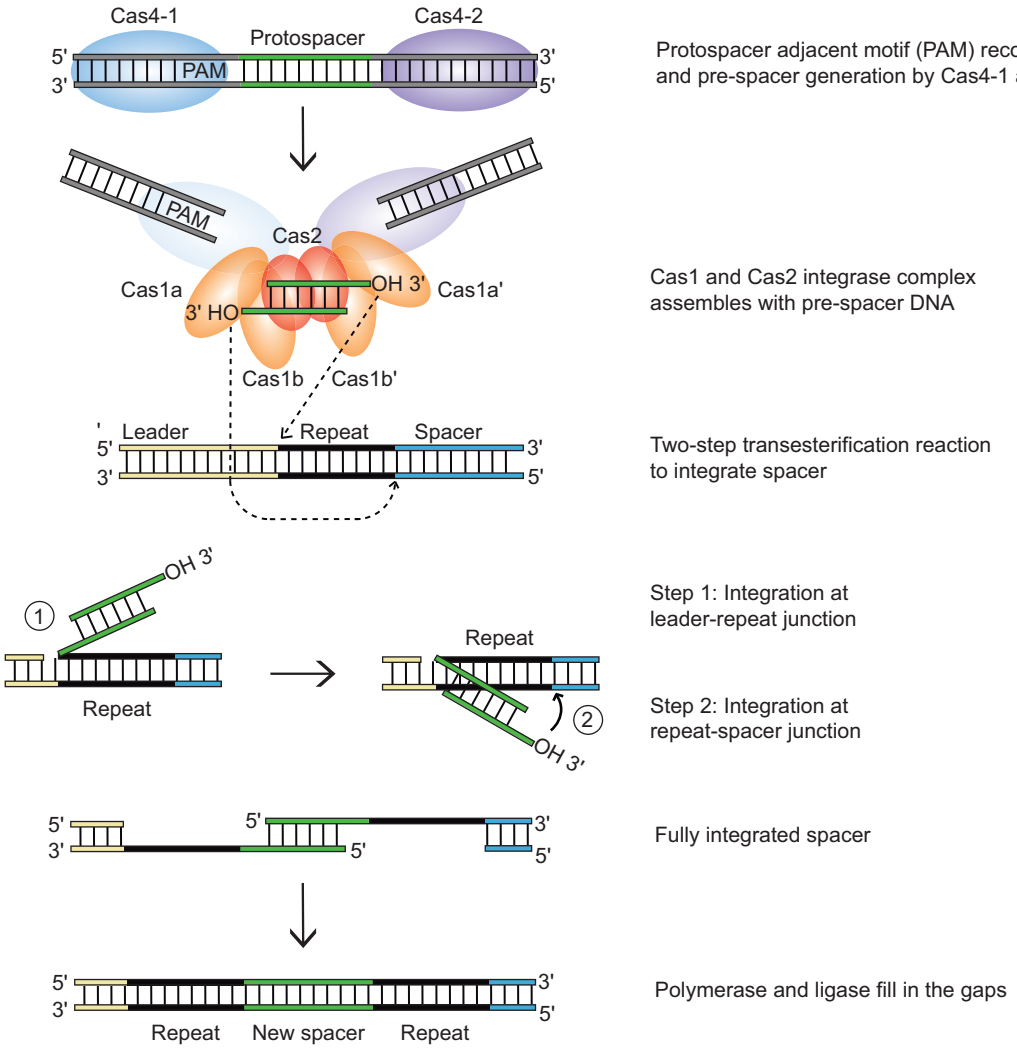


Figure S1. CRISPR adaptation model for *Pyrococcus furiosus*.

Cas4-1 and Cas4-2 generate pre-spacers, which Cas1 and Cas2 then integrate into the CRISPR array by a two-step transesterification reaction at the borders of the leader-proximal repeat. The order of the two integration steps is unknown for this system.

Supplementary Figure S2.

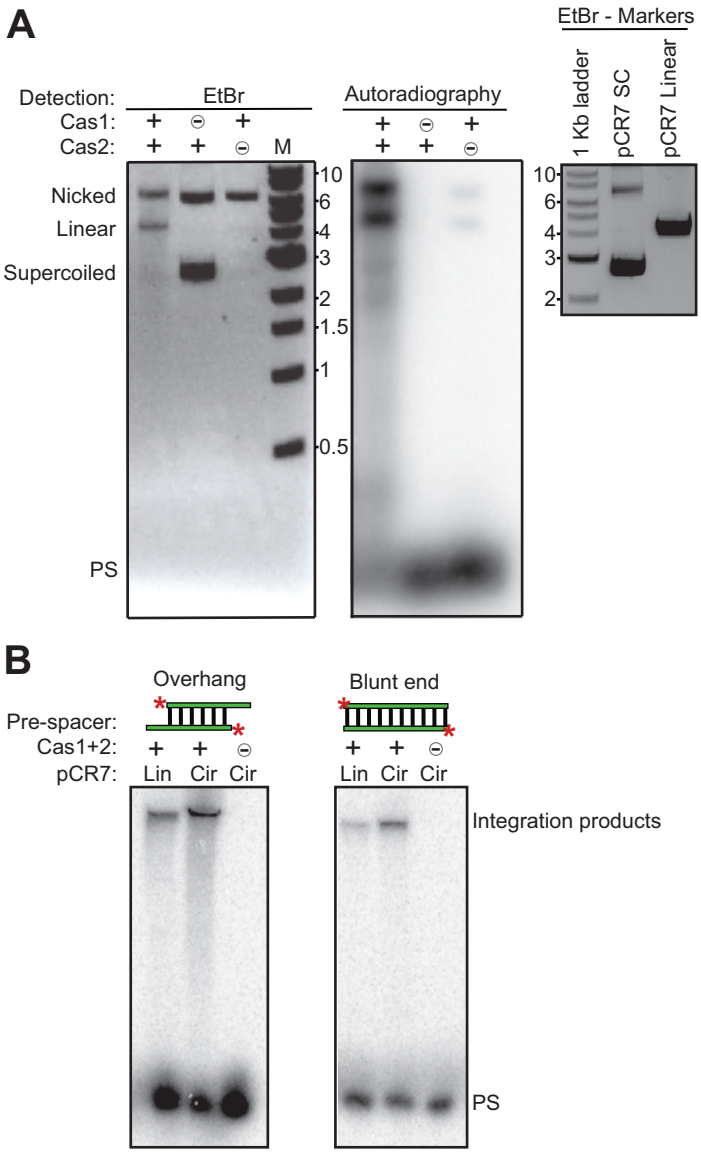


Figure S2. Cas2 is required for robust spacer integration

(**A**) Integration assay comparing Cas1+2, Cas2 alone, and Cas1 alone. Experimental conditions were the same as described for Figure 1. (**B**) Integration assay comparing pre-spacers with five nucleotide long 3′ overhangs and pre-spacers with blunt ends. Linear and circular pCR7 substrates were both tested as the target DNA. Reaction products were separated on 10% denaturing PAGE, followed by autoradiography.

Supplementary Figure S3.

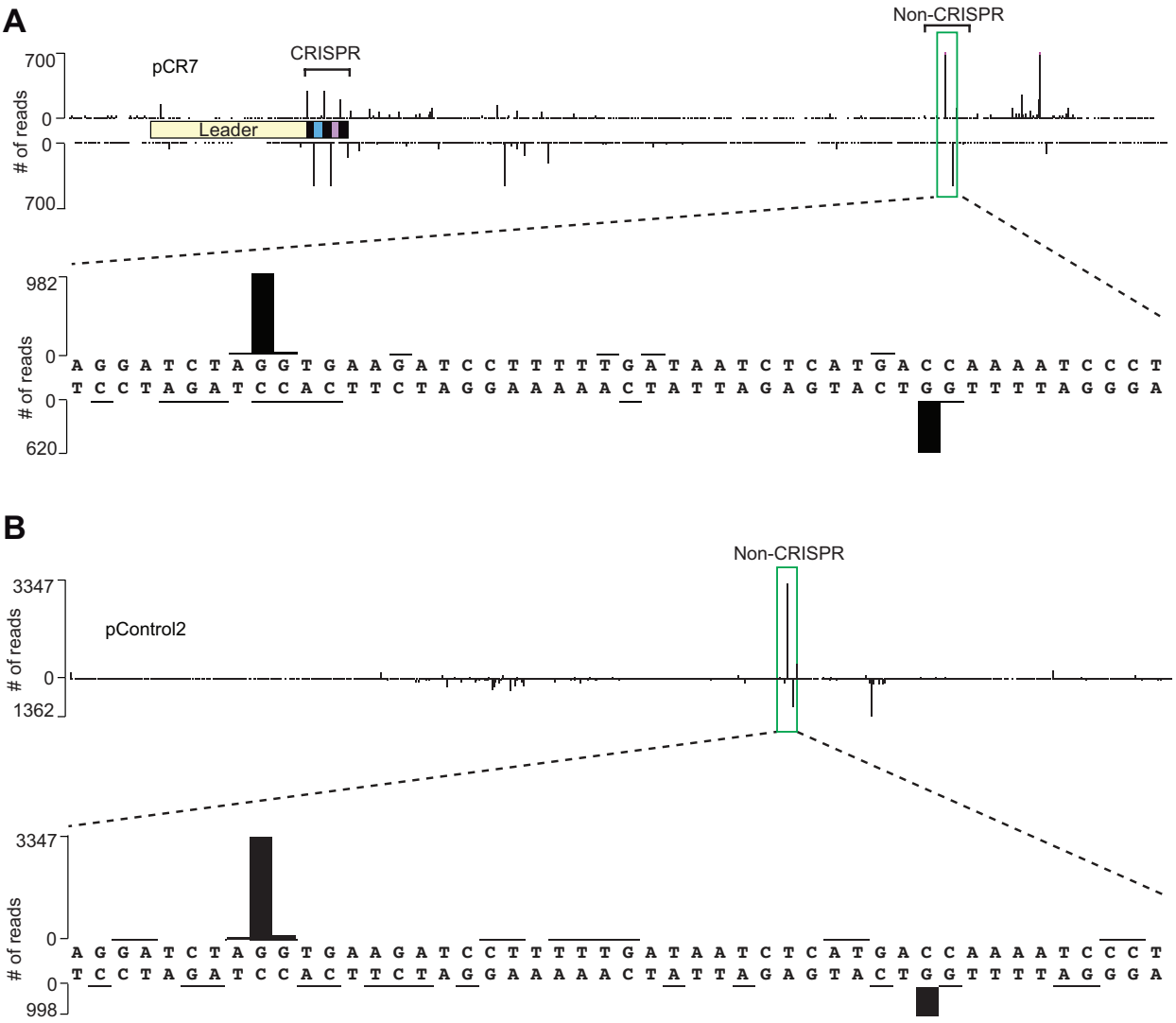


Figure S3. Non-CRISPR integration site

(**A**) Integration locations on pCR7. Increased resolution of the highly preferred non-CRISPR integration site (boxed in green) displayed below. (**B**) Integration locations on pControl2, an alternative plasmid with a different backbone from pCR7. Increased resolution of the same highly preferred non-CRISPR integration site (boxed in green) displayed below. pCR7 and pControl2 share a 120 bp conserved region, in which this preferred site is located, while the rest of the plasmid backbones are not conserved.

Supplementary Figure S4.

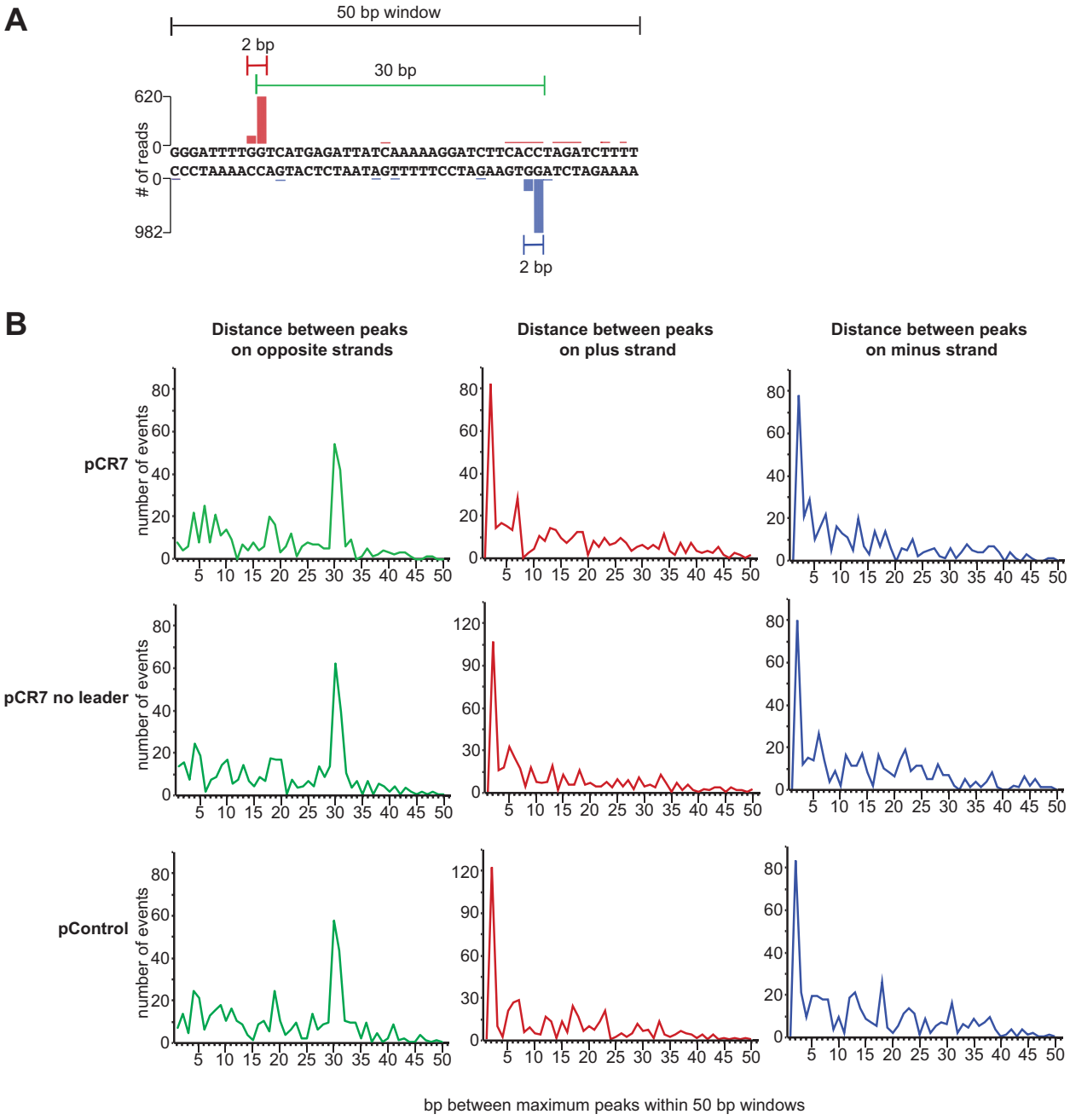


Figure S4. Spacing of integration peaks

(**A**) Spacing between maximum integration peaks within a 50 bp window was measured across the entire plasmid by making 500 random windows. Distance was measured three ways: Highest plus strand peak to highest minus strand peak (green), two highest plus strand peaks (red), and two highest minus strand peaks (blue). (**B**) Peaks on opposite strands (green), plus strand (red), and minus strand (blue) were compared by counting the number of times each bp spacing between highest peaks was found in the 500 random windows.

Supplementary Figure S5.

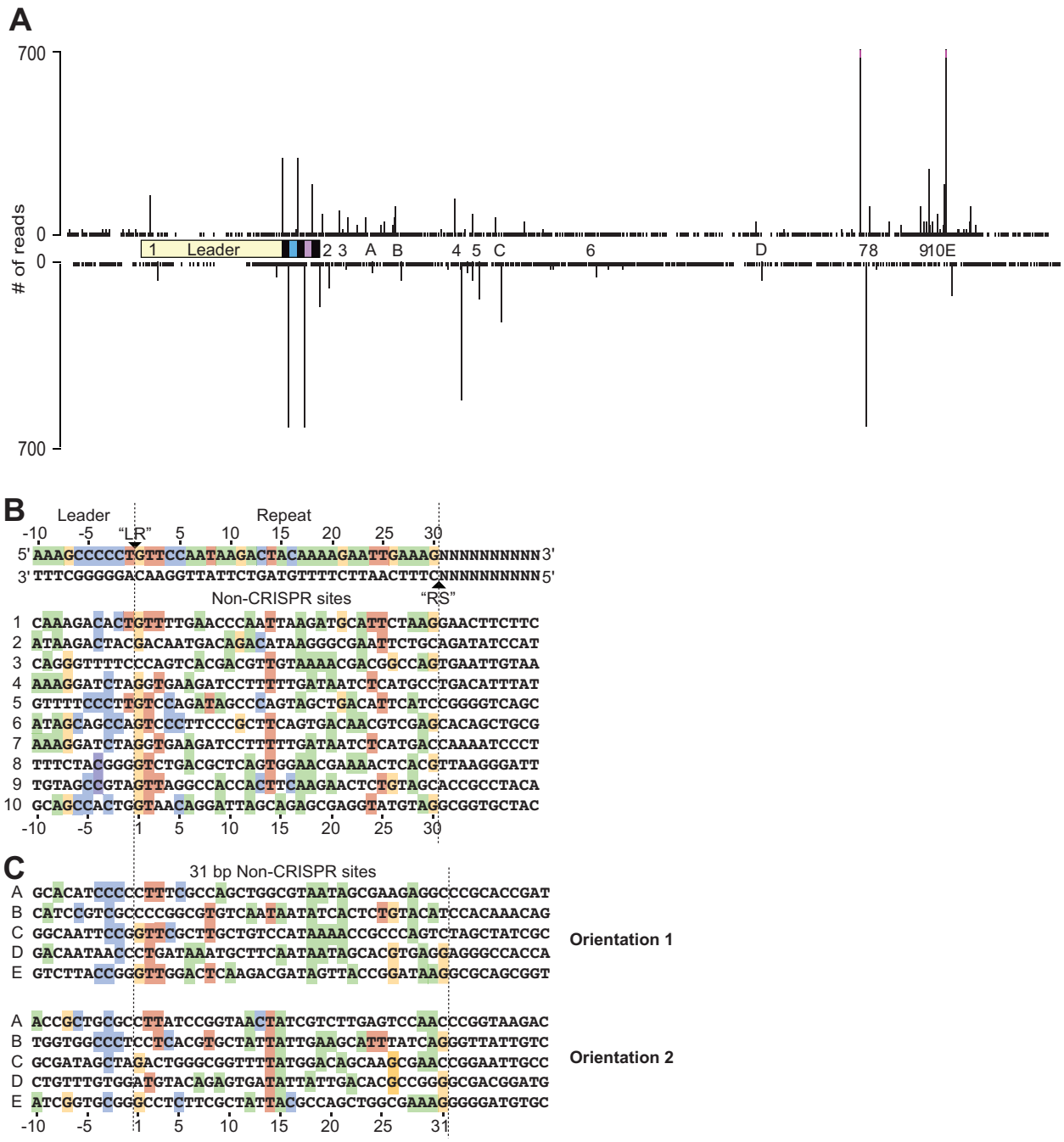
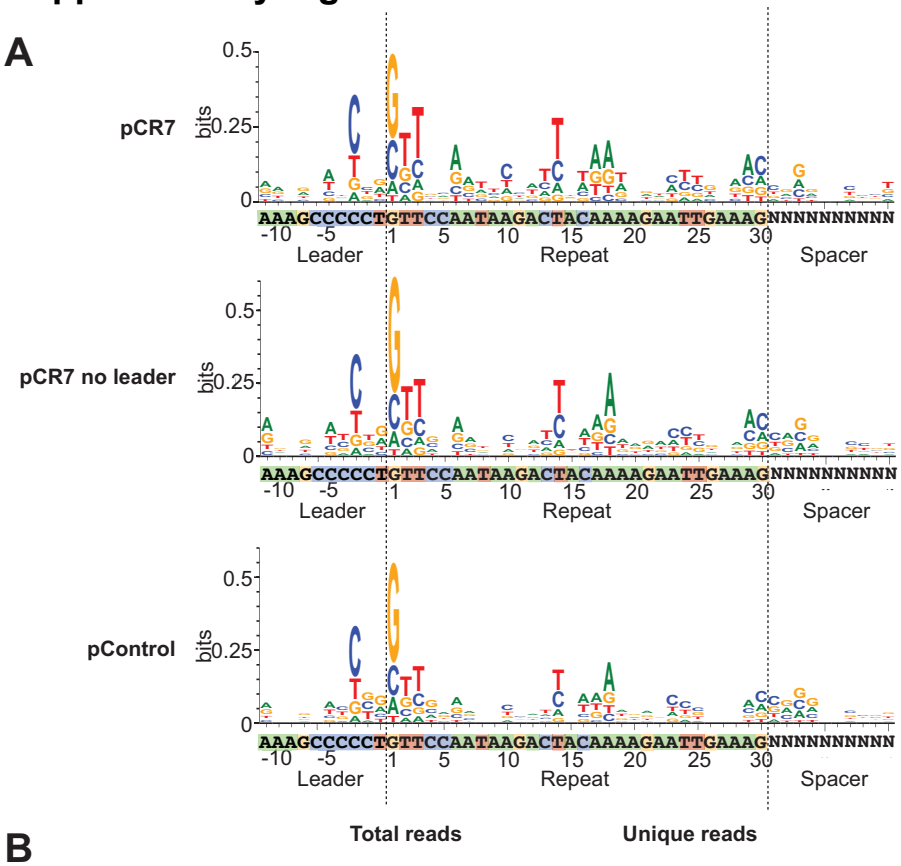


Figure S5. Non-CRISPR integration sites on pCR7

(**A**) Integration sites along pCR7, with top ten 30 bp non-CRISPR sites marked 1 through 10. Highly integrated sites that were 31 bp are marked A through E. (**B, C**) Sequences of 30 bp (B) and 31 bp (C) non-CRISPR sites with corresponding numbers and letters mapping to the track above. Homology with the WT repeat is highlighted for each sequence. The 31 bp sequences are displayed in both orientations, in order to show homology with WT repeat from both sites of integration in the +1 position.

Supplementary Figure S6.



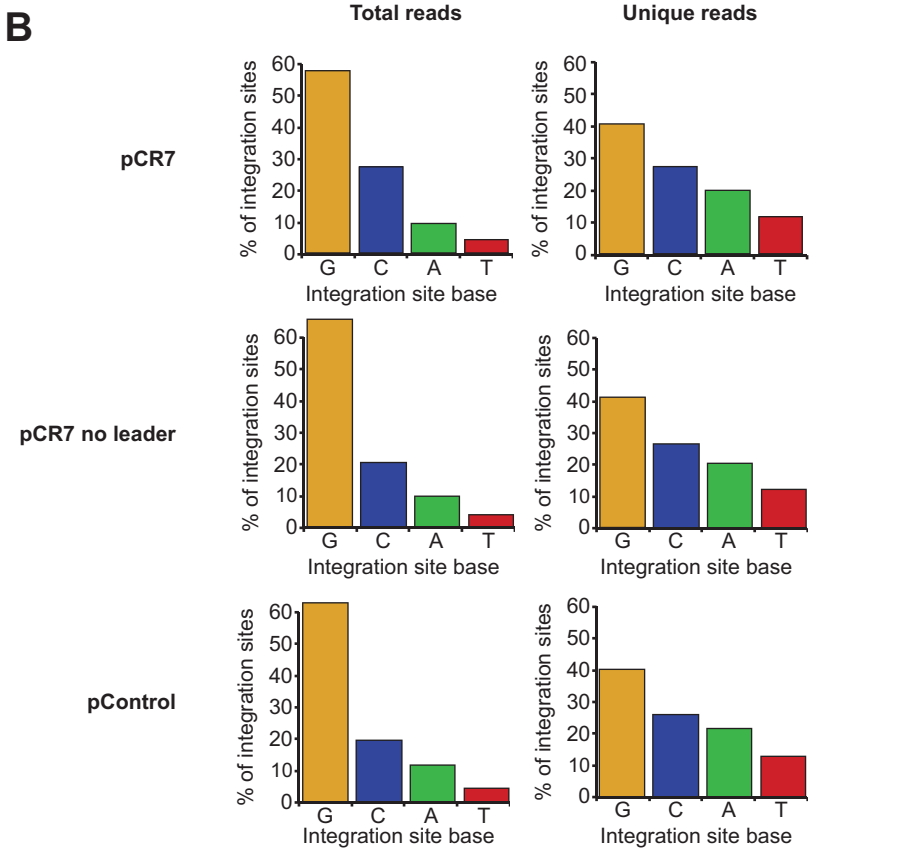


Figure S6. WebLogos and base preferences of all integration sites

(**A**) WebLogos of all integration locations. (**B**) Percent of total or unique integration sites that were located at guanine, cytosine, adenine, and thymine. Many positions in the plasmid had multiple reads showing spacer integration at that point; for “total reads” the base at the integration point was counted once for each supporting read. For “unique reads” the base at a distinct position was counted only once if one or more reads showed integration at that point.

Supplementary Figure S7.

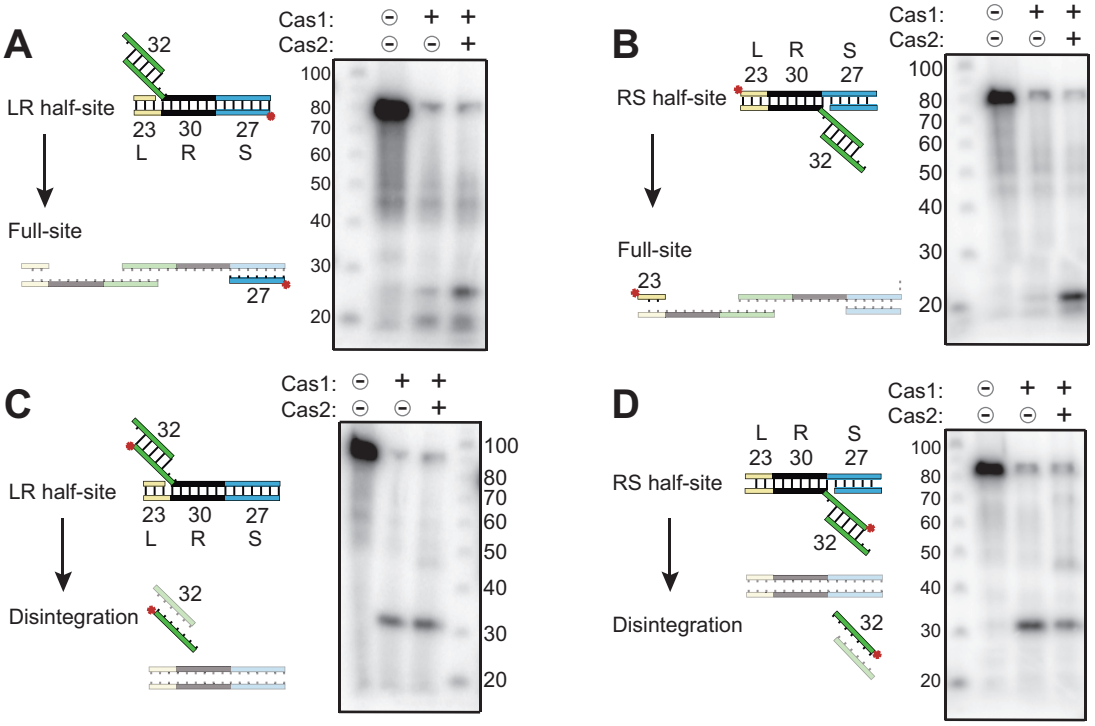


Figure S7. Full-site integration occurs with Cas1 and Cas2, while disintegration is achieved by Cas1 alone. (**A, B**) Continuation of LR half-site (A) or RS half-site (B) substrates to full-site integration detected with radiolabel on bottom strand of the spacer (A) or top strand of leader (B). Reactions contain no protein, Cas1, or Cas1 and Cas2. (**C, D**) Spacer disintegration from LR (C) or RS (D) half-site substrates detected with radiolabel on spacer. Reactions contain no protein, Cas1, or Cas1 and Cas2.

Supplementary Figure S8.

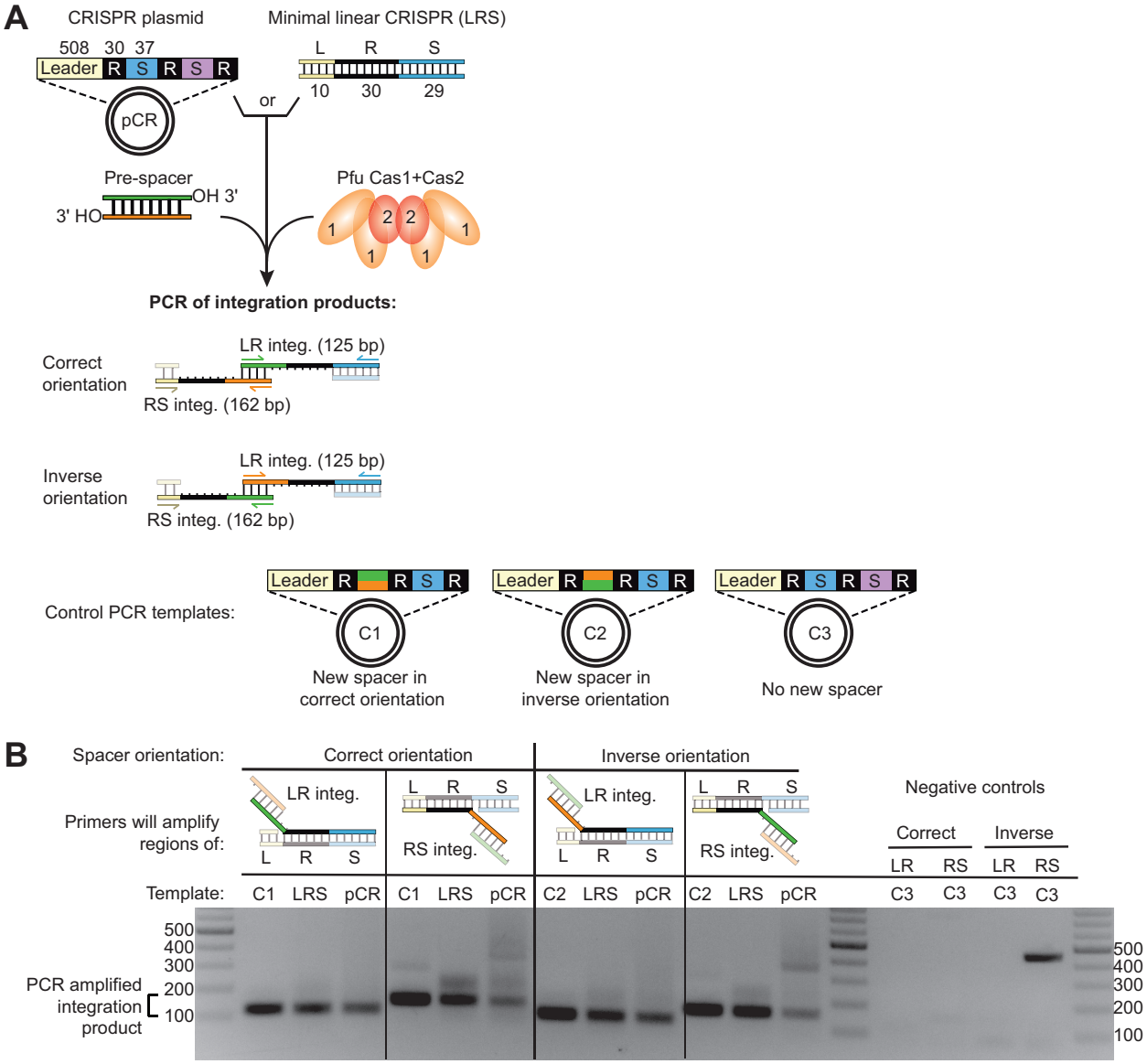


Figure S8. PCR of integration sites determine pre-spacers are integrated in both correct and inverse orientations. (**A**) Spacer integration assays were performed with either CRISPR plasmid (pCR) or minimal linear CRISPRs (LRS). Integration products were used as templates for the following PCR reactions: 1) Spacer with correct orientation integrated at LR, 2) Spacer with correct orientation integrated at RS, 3) Spacer with inverse orientation integrated at LR, 4) Spacer with inverse orientation integrated at RS. Plasmids containing correct (C1) or inverse (C2) orientation spacers were used as positive PCR controls, and pCR with no new spacer (C3) was used in negative PCR controls. (**B**) The PCR amplified integration products were separated on a 1% agarose gel in 1x TAE. Expected PCR products are 125 bp for LR integration and 162 bp for RS integration.

Supplementary Figure S9.

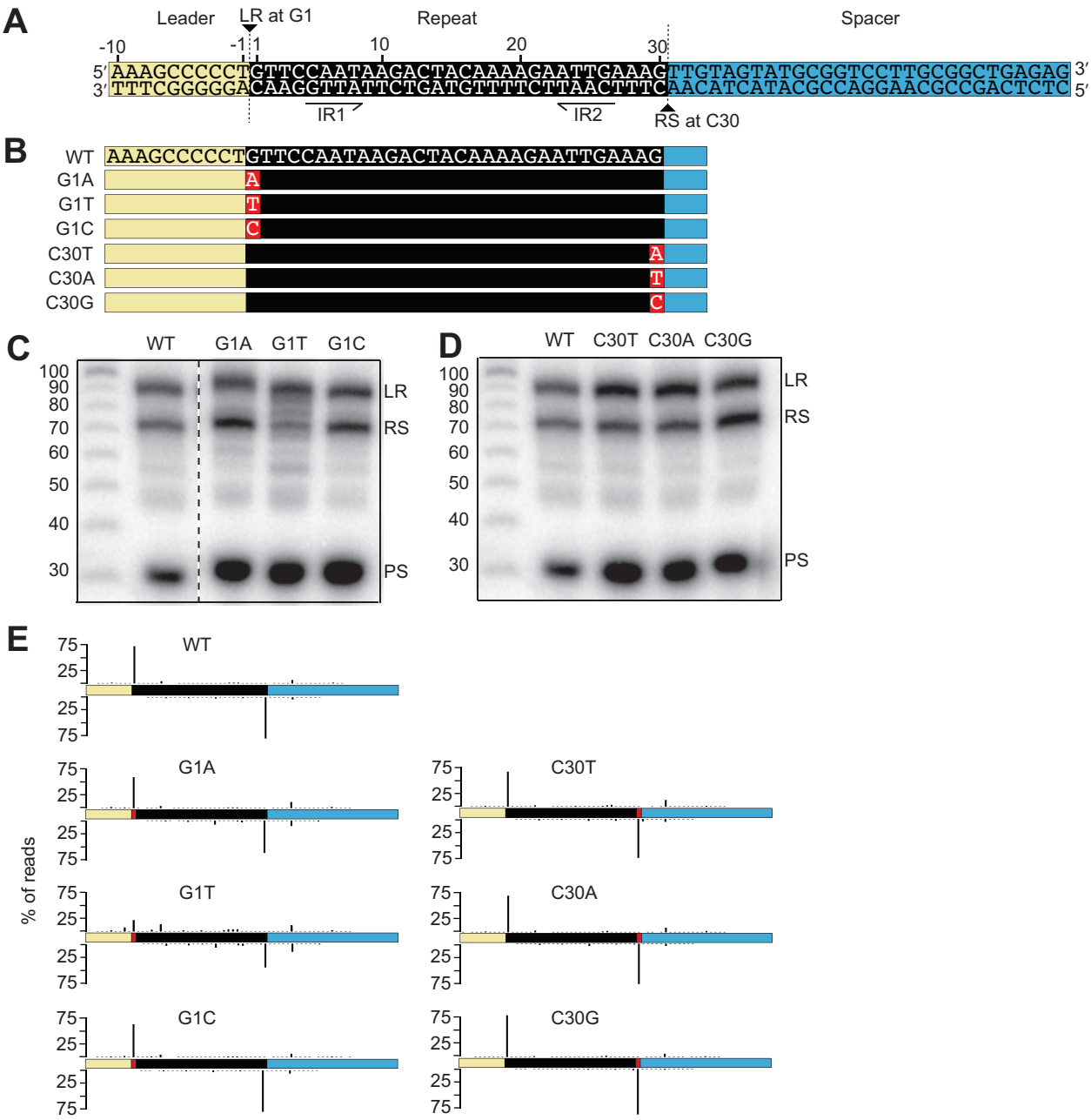


Figure S9. Mutations to the first or last nucleotide of the repeat

(**A**) Annotated sequence of minimal linear CRISPR. (**B**) List of substitution mutations made to the minimal CRISPR. (**C, D**) Integration assay comparing mutations to the (C) first and (D) last nucleotide of the repeat. Noncontiguous lanes from the same gel are indicated by dashed lines. (**E**) Integration sites along the minimal CRISPRs.

Supplementary Figure S10.

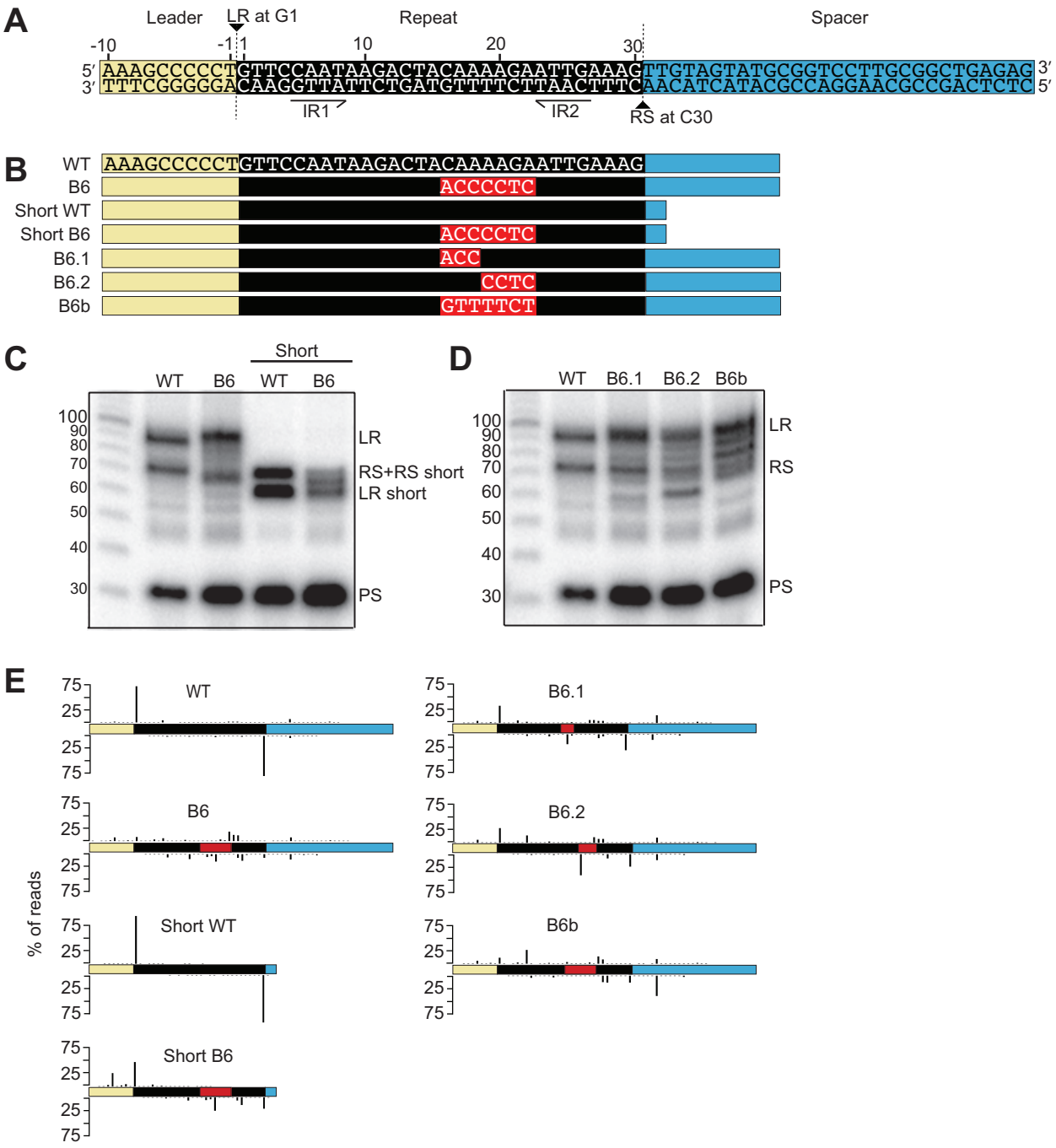


Figure S10. Mutations made to block 6 of the repeat

(**A**) Annotated sequence of minimal linear CRISPR. (**B**) List of substitution mutations made to the minimal CRISPR. (**C**) Integration assay comparing original minimal CRISPR length to a shortened minimal CRISPR to distinguish the integration products of the block 6 mutation. LR integration product of short minimal CRISPR is 64 bp. (**D**) Integration assay comparing alternative mutations to block 6. (**E**) Integration sites along the minimal CRISPRs.

Supplementary Figure S11.

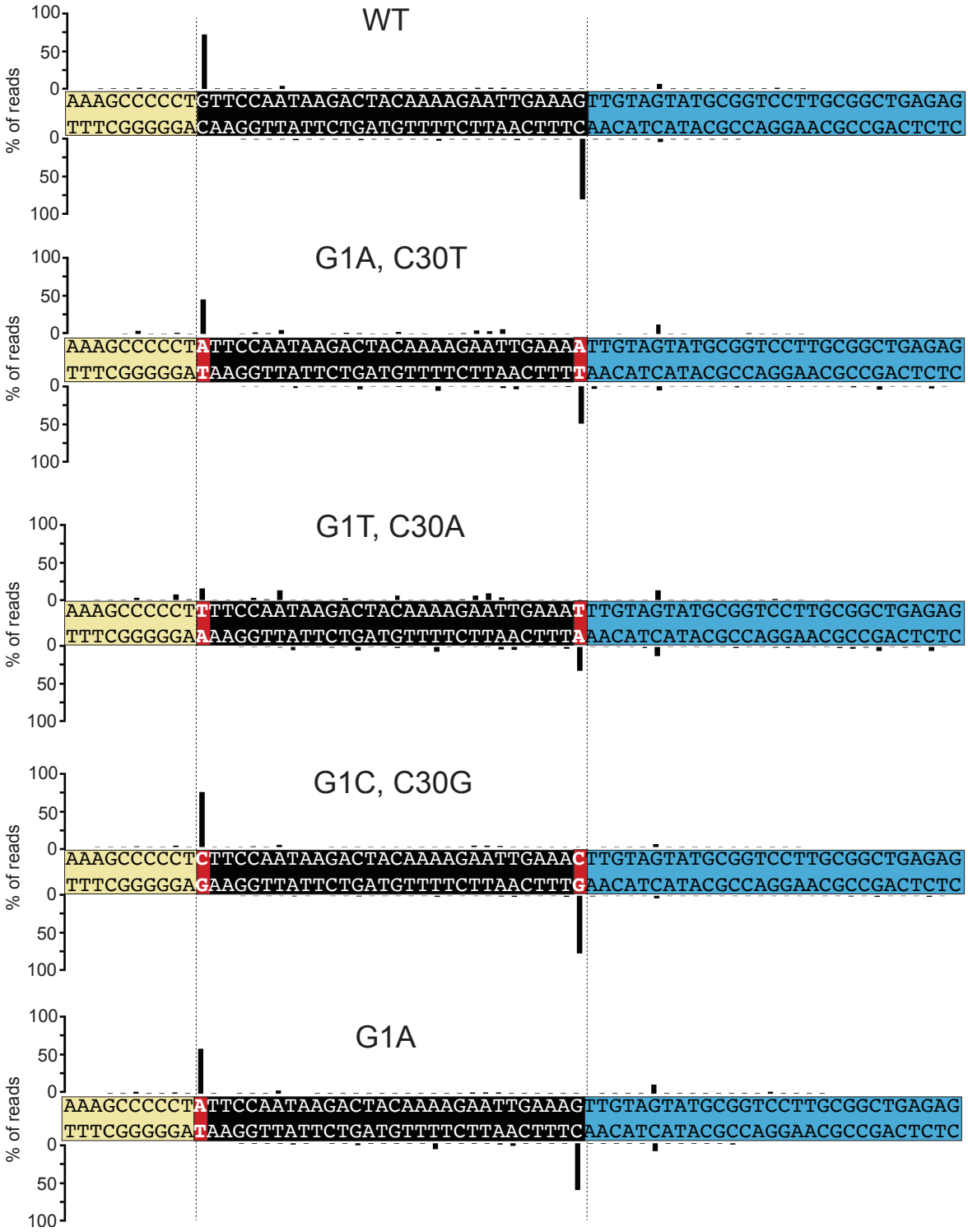
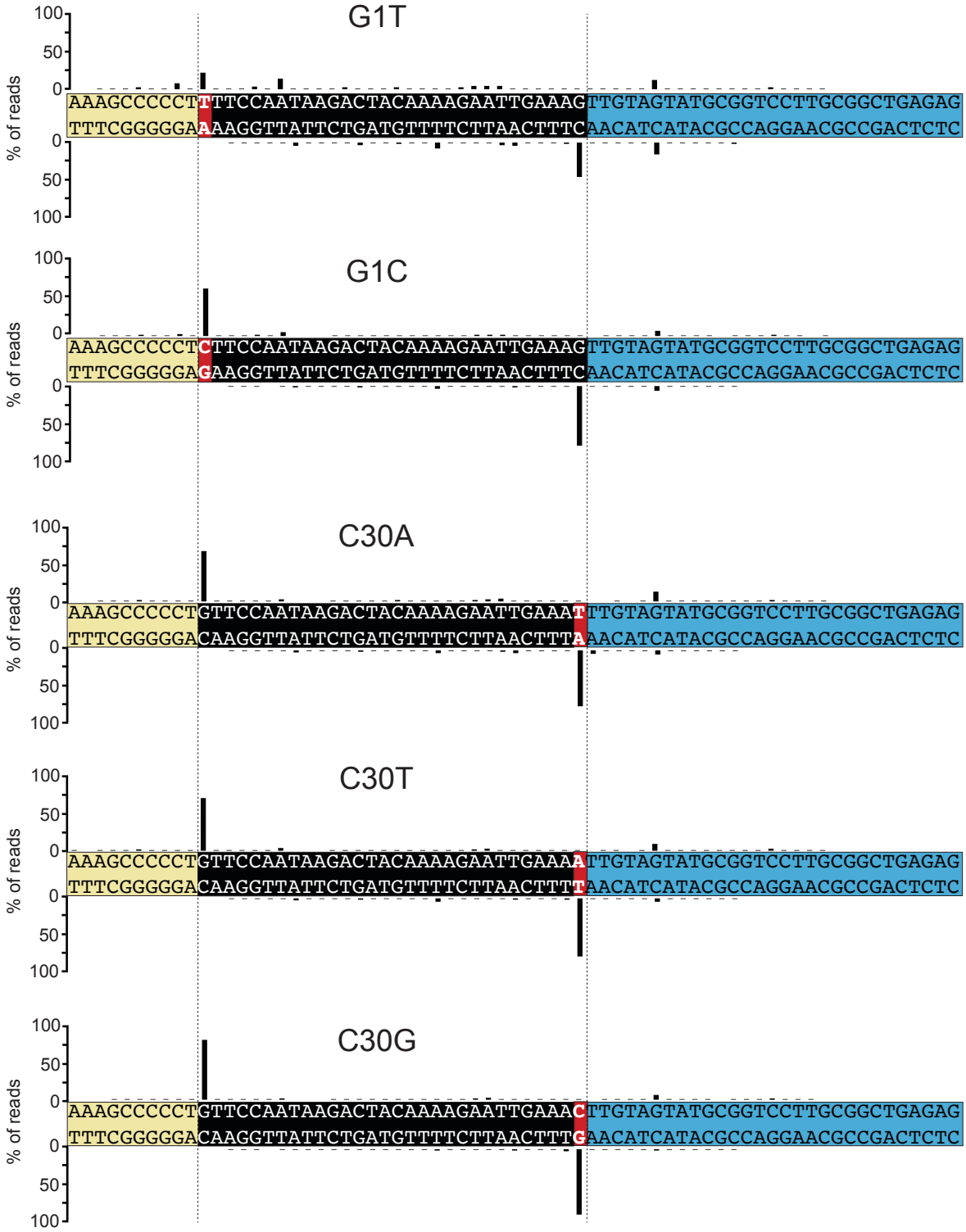
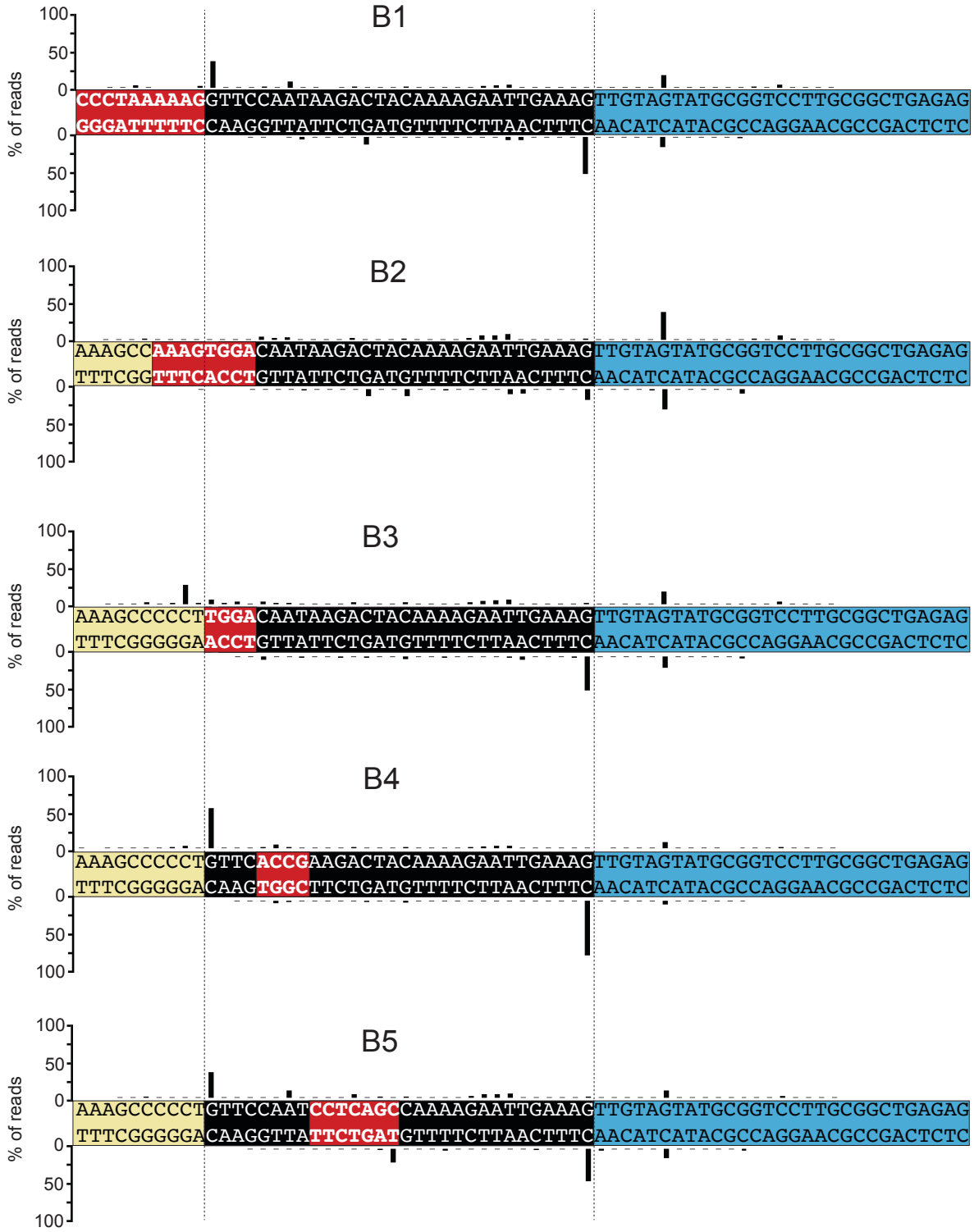
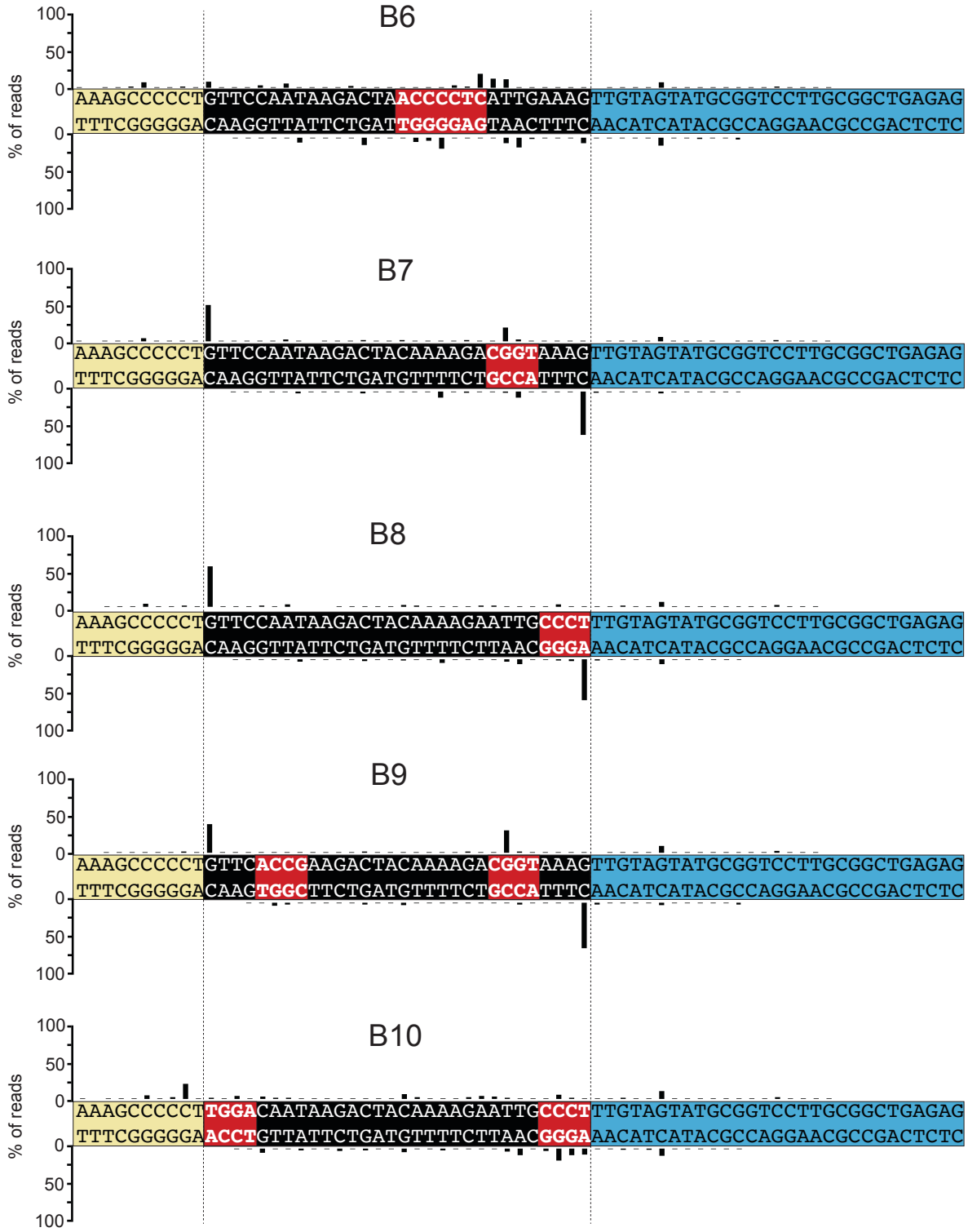
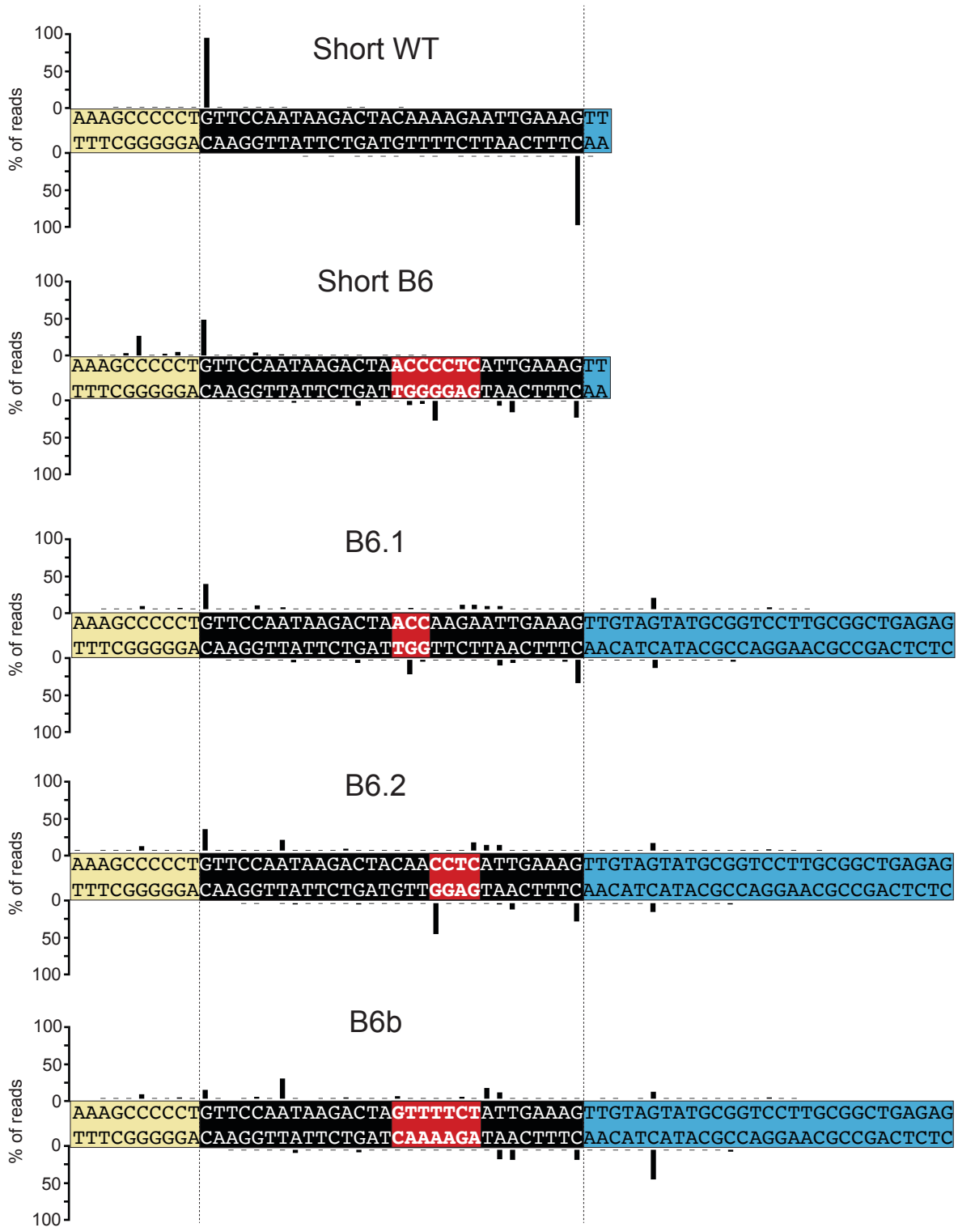


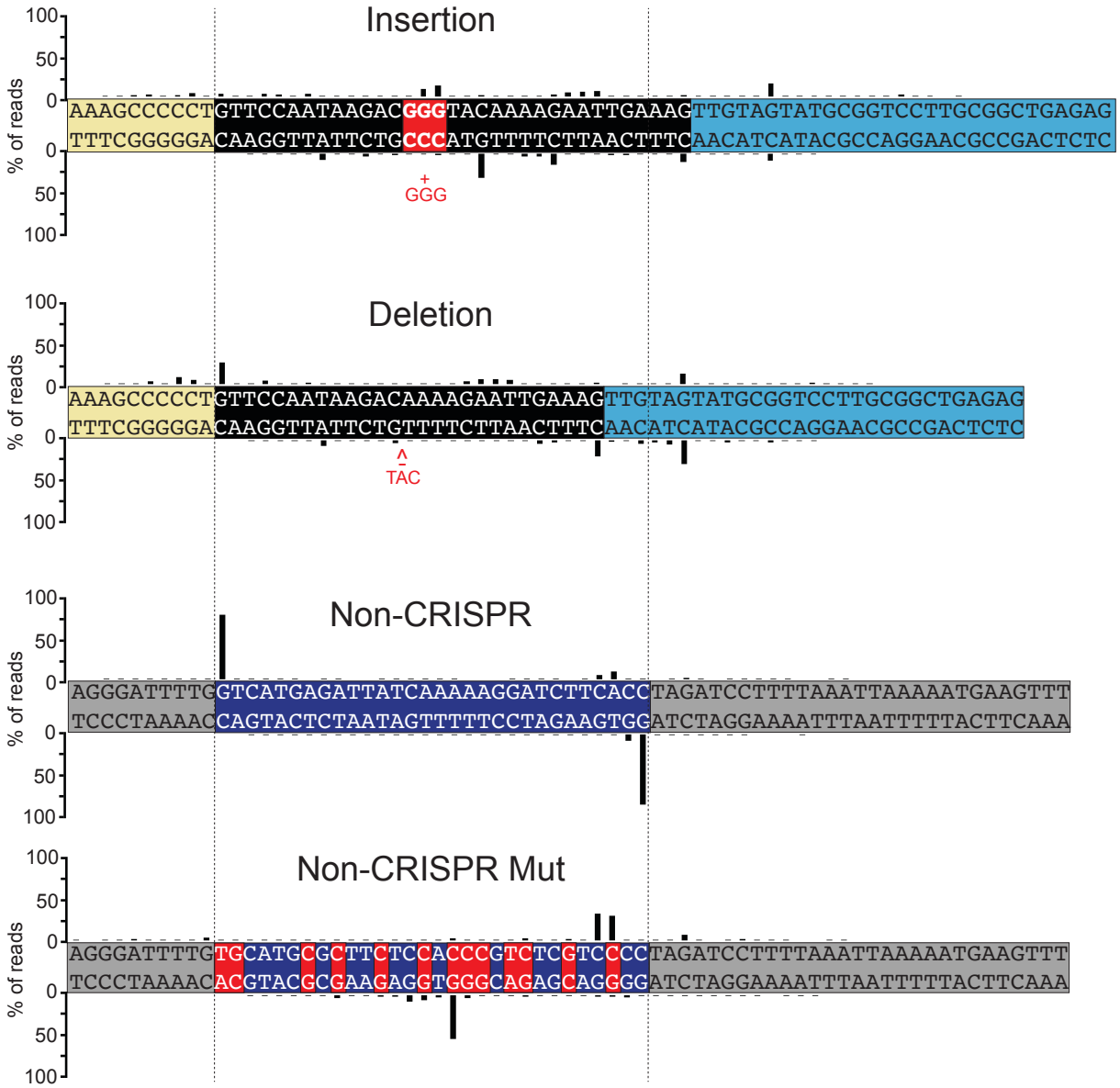
Figure S11. Nucleotide resolution graphs displaying spacer integration sites along minimal CRISPRs used in all experiments. Summary table of % of reads for LR and RS sites can be found in Table S2.


G1T
% of reads
100
50
0
50
100
AAAGCCCCCTTTTCCAATAAGACTACAAAAGAATTGAAAGTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGAAAAGGTTATTCTGATGTTTTCTTAACTTTCAACATCATACGCCAGGAACGCCGACTCTC
G1C
% of reads
100
50
0
50
100
AAAGCCCCCTCTTCCAATAAGACTACAAAAGAATTGAAAGTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGAGAAGGTTATTCTGATGTTTTCTTAACTTTCAACATCATACGCCAGGAACGCCGACTCTC
C30A
% of reads
100
50
0
50
100
AAAGCCCCCTGTTCCAATAAGACTACAAAAGAATTGAAATTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGACAAGGTTATTCTGATGTTTTCTTAACTTTAAACATCATACGCCAGGAACGCCGACTCTC
C30T
% of reads
100
50
0
50
100
AAAGCCCCCTGTTCCAATAAGACTACAAAAGAATTGAAAATTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGACAAGGTTATTCTGATGTTTTCTTAACTTTTAACATCATACGCCAGGAACGCCGACTCTC
C30G
% of reads
100
50
0
50
100
AAAGCCCCCTGTTCCAATAAGACTACAAAAGAATTGAAACTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGACAAGGTTATTCTGATGTTTTCTTAACTTTGAACATCATACGCCAGGAACGCCGACTCTC


B1
B2
B3
B4
B5
% of reads
100
50
0
50
100
CCCTAAAAAGGTTCCAATAAGACTACAAAAGAATTGAAAGTTGTAGTATGCGGTCCTTGCGGCTGAGAG
GGGATTTTTCCAAGGTTATTCTGATGTTTTCTTAACTTTCAACATCATACGCCAGGAACGCCGACTCTC
AAAGCCAAAGTGGACAATAAGACTACAAAAGAATTGAAAGTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGTTTCACCTGTTATTCTGATGTTTTCTTAACTTTCAACATCATACGCCAGGAACGCCGACTCTC
AAAGCCCCCCTTGGACAATAAGACTACAAAAGAATTGAAAGTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGGAACCTGTTATTCTGATGTTTTCTTAACTTTCAACATCATACGCCAGGAACGCCGACTCTC
AAAGCCCCCCTGTTCACCGAAGACTACAAAAGAATTGAAAGTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGGACAAGTGGCTTCTGATGTTTTCTTAACTTTCAACATCATACGCCAGGAACGCCGACTCTC
AAAGCCCCCCTGTTCCAATCCTCAGCCAAAAGAATTGAAAGTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGGACAAGGTTATTCTGATGTTTTCTTAACTTTCAACATCATACGCCAGGAACGCCGACTCTC


B6
% of reads
100
50
0
0
50
100
AAAGCCCCCTGTTCCAATAAGACTAACCCCTCATTGAAAGTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGGACAAGGTTATTCTGATTGGGGAGTAACTTTCAACATCATACGCCAGGAACGCCGACTCTC
B7
% of reads
100
50
0
0
50
100
AAAGCCCCCTGTTCCAATAAGACTACAAAAGACGGTAAAGTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGGACAAGGTTATTCTGATGTTTTCTGCCATTTCAACATCATACGCCAGGAACGCCGACTCTC
B8
% of reads
100
50
0
0
50
100
AAAGCCCCCTGTTCCAATAAGACTACAAAAGAATTGCCCTTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGGACAAGGTTATTCTGATGTTTTCTTAACGGGAAACATCATACGCCAGGAACGCCGACTCTC
B9
% of reads
100
50
0
0
50
100
AAAGCCCCCTGTTCACCGAAGACTACAAAAGACGGTAAAGTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGGACAAGTGGCTTCTGATGTTTTCTGCCATTTCAACATCATACGCCAGGAACGCCGACTCTC
B10
% of reads
100
50
0
0
50
100
AAAGCCCCCTTGGACAATAAGACTACAAAAGAATTGCCCTTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGGAACCTGTTATTCTGATGTTTTCTTAACGGGAAACATCATACGCCAGGAACGCCGACTCTC


Short WT
% of reads
100
50
0
0
50
100
AAAGCCCCCTGTTCCAATAAGACTACAAAAGAATTGAAAGTT
TTTCGGGGGGACAAGGTTATTCTGATGTTTTCTTAACTTTCAA
Short B6
% of reads
AAAGCCCCCTGTTCCAATAAGACTAACCCCTCATTGAAAGTT
TTTCGGGGGGACAAGGTTATTCTGATTGGGGAGTAACTTTCAA
B6.1
% of reads
AAAGCCCCCTGTTCCAATAAGACTAACCAAGAATTGAAAGTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGGACAAGGTTATTCTGATTGGTTCTTAACTTTCAACATCATACGCCAGGAACGCCGACTCTC
B6.2
% of reads
AAAGCCCCCTGTTCCAATAAGACTACAACCTCATTGAAAGTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGGACAAGGTTATTCTGATGTTGGAGTAACTTTCAACATCATACGCCAGGAACGCCGACTCTC
B6b
% of reads
AAAGCCCCCTGTTCCAATAAGACTAGTTTTCTATTGAAAGTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGGACAAGGTTATTCTGATCAAAAGATAACTTTCAACATCATACGCCAGGAACGCCGACTCTC


Insertion
% of reads
100
50
0
50
100
AAAGCCCCCTGTTCCAATAAGACGGGTACAAAAGAATTGAAAGTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGGACAAGGTTATTCTGCCCATGTTTTCTTAACTTTCAACATCATACGCCAGGAACGCCGACTCTC
+
GGG
Deletion
% of reads
100
50
0
50
100
AAAGCCCCCTGTTCCAATAAGACAAAAGAATTGAAAGTTGTAGTATGCGGTCCTTGCGGCTGAGAG
TTTCGGGGGGACAAGGTTATTCTGTTTTCTTAACTTTCAACATCATACGCCAGGAACGCCGACTCTC
^
TAC
Non-CRISPR
% of reads
100
50
0
50
100
AGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTT
TCCCTAAAACCAGTACTCTAATAGTTTTTCCTAGAAGTGGATCTAGGAAAATTTAATTTTTACTTCAAA
Non-CRISPR Mut
% of reads
100
50
0
50
100
AGGGATTTTGTGCATGCGCTTCTCCACCCGTCTCGTCCCCTAGATCCTTTTAAATTAAAAATGAAGTTT
TCCCTAAAACACGTACGCGAAGAGGTGGGCAGAGCAGGGGATCTAGGAAAATTTAATTTTTACTTCAAA

Supplementary Table S1.

Name	Sequence (5'-3')
Pre-spacer top	ATATAAGCAATACGGAACTTATGAGAACTATT
Pre-spacer bottom	TTCTCATAAGTTCCGTATTGCTTATATTTCAA
LRS linear top	AAAGCCCCCTGTTCCAATAAGACTACAAAAGAATTGAAAGTTGTAGTATG CGGTCCTTGCGGCTGAGAG
LRS linear bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTCAATTCTTTTGTAGTCTT ATTGGAACAGGGGGCTTT
LR half-site top	ATATAAGCAATACGGAACTTATGAGAACTATTGTTCCAATAAGACTACAAA AGAATTGAAAGTTGTAGTATGCGGTCCTTGCGGCTGAG
LR half-site leader	GGATTGGGGCAAAAAAGCCCCCT
LR half-site bottom	CTCAGCCGCAAGGACCGCATACTACAACTTTCAATTCTTTTGTAGTCTTAT TGGAACAGGGGGCTTTTTTGCCCCAATCC
LR half-site spacer	TTCTCATAAGTTCCGTATTGCTTATATTTCAA
RS half-site top	GGATTGGGGCAAAAAAGCCCCCTGTTCCAATAAGACTACAAAAGAATTGA AAGTTGTAGTATGCGGTCCTTGCGGCTGAG
RS half-site bottom	TTCTCATAAGTTCCGTATTGCTTATATTTCAACTTTCAATTCTTTTGTAGTC TTATTGGAACAGGGGGCTTTTTTGCCCCAATCC
RS half-site S1	CTCAGCCGCAAGGACCGCATACTACAA
RS half-site spacer	ATATAAGCAATACGGAACTTATGAGAACTATT
LRS spacer hairpin	GCAAAAAAGCCCCCTGTTCCAATAAGACTACAAAAGAATTGAAAGGAGTG TGCGAGGGCTTGGGCCCTGAGGTAACGTGCGATCGCACGTTACCTCAG GGCCCAAGCCCTCGCACACTCCTTTCAATTCTTTTGTAGTCTTATTGGAA CAGGGGGCTTTTTTGC
LRS leader hairpin	GCCCTCGCACACTCCTTTCAATTCTTTTGTAGTCTTATTGGAACAGGGGG CTTTTTTGCCGATAGGCAAAAAAGCCCCCTGTTCCAATAAGACTACAAAA GAATTGAAAGGAGTGTGCGAGGGC
Non-CRISPR hairpin	GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATC CTTTTAAATCGATAGATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAA TCTCATGACCAAAATCCCTTAAC
G1A, C30T top	AAAGCCCCCTATTCCAATAAGACTACAAAAGAATTGAAAATTGTAGTATGC GGTCCTTGCGGCTGAGAG
G1A, C30T bottom	CTCTCAGCCGCAAGGACCGCATACTACAATTTTCAATTCTTTTGTAGTCTT ATTGGAATAGGGGGCTTT
G1T, C30A top	AAAGCCCCCTTTTCCAATAAGACTACAAAAGAATTGAAATTTGTAGTATGC GGTCCTTGCGGCTGAGAG
G1T, C30A bottom	CTCTCAGCCGCAAGGACCGCATACTACAAATTTCAATTCTTTTGTAGTCTT ATTGGAAAAGGGGGCTTT
G1C, C30G top	AAAGCCCCCTCTTCCAATAAGACTACAAAAGAATTGAAACTTGTAGTATG CGGTCCTTGCGGCTGAGAG
G1C, C30G bottom	CTCTCAGCCGCAAGGACCGCATACTACAAGTTTCAATTCTTTTGTAGTCTT ATTGGAAGAGGGGGCTTT

G1A top	AAAGCCCCCTATTCCAATAAGACTACAAAAGAATTGAAAGTTGTAGTATG CGGTCCTTGCGGCTGAGAG
G1A bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTCAATTCTTTTGTAGTCTT ATTGGAATAGGGGGCTTT
G1T top	AAAGCCCCCTTTTCCAATAAGACTACAAAAGAATTGAAAGTTGTAGTATGC GGTCCTTGCGGCTGAGAG
G1T bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTCAATTCTTTTGTAGTCTT ATTGGAAAAGGGGGCTTT
G1C top	AAAGCCCCCTCTTCCAATAAGACTACAAAAGAATTGAAAGTTGTAGTATG CGGTCCTTGCGGCTGAGAG
G1C bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTCAATTCTTTTGTAGTCTT ATTGGAAGAGGGGGCTTT
C30T top	AAAGCCCCCTGTTCCAATAAGACTACAAAAGAATTGAAAATTGTAGTATG CGGTCCTTGCGGCTGAGAG
C30T bottom	CTCTCAGCCGCAAGGACCGCATACTACAATTTTCAATTCTTTTGTAGTCTT ATTGGAACAGGGGGCTTT
C30A top	AAAGCCCCCTGTTCCAATAAGACTACAAAAGAATTGAAATTTGTAGTATGC GGTCCTTGCGGCTGAGAG
C30A bottom	CTCTCAGCCGCAAGGACCGCATACTACAAATTTCAATTCTTTTGTAGTCTT ATTGGAACAGGGGGCTTT
C30G top	AAAGCCCCCTGTTCCAATAAGACTACAAAAGAATTGAAACTTGTAGTATG CGGTCCTTGCGGCTGAGAG
C30G bottom	CTCTCAGCCGCAAGGACCGCATACTACAAGTTTCAATTCTTTTGTAGTCTT ATTGGAACAGGGGGCTTT
Block 1 top	CCCTAAAAAGGTTCCAATAAGACTACAAAAGAATTGAAAGTTGTAGTATG CGGTCCTTGCGGCTGAGAG
Block 1 bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTCAATTCTTTTGTAGTCTT ATTGGAACCTTTTTAGGG
Block 2 top	AAAGCCAAAGTGGACAATAAGACTACAAAAGAATTGAAAGTTGTAGTATG CGGTCCTTGCGGCTGAGAG
Block 2 bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTCAATTCTTTTGTAGTCTT ATTGTCCACTTTGGCTTT
Block 3 top	AAAGCCCCCTTGGACAATAAGACTACAAAAGAATTGAAAGTTGTAGTATG CGGTCCTTGCGGCTGAGAG
Block 3 bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTCAATTCTTTTGTAGTCTT ATTGTCCAAGGGGGCTTT
Block 4 top	AAAGCCCCCTGTTCACCGAAGACTACAAAAGAATTGAAAGTTGTAGTATG CGGTCCTTGCGGCTGAGAG
Block 4 bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTCAATTCTTTTGTAGTCTT CGGTGAACAGGGGGCTTT
Block 5 top	AAAGCCCCCTGTTCCAATCCTCAGCCAAAAGAATTGAAAGTTGTAGTATG CGGTCCTTGCGGCTGAGAG
Block 5 bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTCAATTCTTTTGGCTGAG GATTGGAACAGGGGGCTTT
Block 6 top	AAAGCCCCCTGTTCCAATAAGACTAACCCCTCATTGAAAGTTGTAGTATG CGGTCCTTGCGGCTGAGAG
Block 6 bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTCAATGAGGGGTTAGTCT TATTGGAACAGGGGGCTTT
Block 7 top	AAAGCCCCCTGTTCCAATAAGACTACAAAAGACGGTAAAGTTGTAGTATG CGGTCCTTGCGGCTGAGAG

Block 7 bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTACCGTCTTTTGTAGTCT TATTGGAACAGGGGGCTTT
Block 8 top	AAAGCCCCCTGTTCCAATAAGACTACAAAAGAATTGCCCTTTGTAGTATG CGGTCCTTGCGGCTGAGAG
Block 8 bottom	CTCTCAGCCGCAAGGACCGCATACTACAAAGGGCAATTCTTTTGTAGTCT TATTGGAACAGGGGGCTTT
Block 9 top	AAAGCCCCCTGTTCACCGAAGACTACAAAAGACGGTAAAGTTGTAGTATG CGGTCCTTGCGGCTGAGAG
Block 9 bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTACCGTCTTTTGTAGTCT TCGGTGAACAGGGGGCTTT
Block 10 top	AAAGCCCCCTTGGACAATAAGACTACAAAAGAATTGCCCTTTGTAGTATG CGGTCCTTGCGGCTGAGAG
Block 10 bottom	CTCTCAGCCGCAAGGACCGCATACTACAAAGGGCAATTCTTTTGTAGTCT TATTGTCCAAGGGGGCTTT
Short WT top	AAAGCCCCCTGTTCCAATAAGACTACAAAAGAATTGAAAGTT
Short WT bottom	AACTTTCAATTCTTTTGTAGTCTTATTGGAACAGGGGGCTTT
Short Block 6 top	AAAGCCCCCTGTTCCAATAAGACTAACCCCTCATTGAAAGTT
Short Block 6 bottom	AACTTTCAATGAGGGGTTAGTCTTATTGGAACAGGGGGCTTT
Block 6.1 top	AAAGCCCCCTGTTCCAATAAGACTAACCAAGAATTGAAAGTTGTAGTATG CGGTCCTTGCGGCTGAGAG
Block 6.1 bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTCAATTCTTGGTTAGTCT TATTGGAACAGGGGGCTTT
Block 6.2 top	AAAGCCCCCTGTTCCAATAAGACTACAACCTCATTGAAAGTTGTAGTATG CGGTCCTTGCGGCTGAGAG
Block 6.2 bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTCAATGAGGTTGTAGTCT TATTGGAACAGGGGGCTTT
Block 6b top	AAAGCCCCCTGTTCCAATAAGACTAGTTTTCTATTGAAAGTTGTAGTATGC GGTCCTTGCGGCTGAGAG
Block 6b bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTCAATAGAAAACTAGTCT TATTGGAACAGGGGGCTTT
Insertion top	AAAGCCCCCTGTTCCAATAAGACGGGTACAAAAGAATTGAAAGTTGTAGT ATGCGGTCCTTGCGGCTGAGAG
Insertion bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTCAATTCTTTTGTACCCG TCTTATTGGAACAGGGGGCTTT
Deletion top	AAAGCCCCCTGTTCCAATAAGACAAAAGAATTGAAAGTTGTAGTATGCGG TCCTTGCGGCTGAGAG
Deletion bottom	CTCTCAGCCGCAAGGACCGCATACTACAACTTTCAATTCTTTTGTCTTATT GGAACAGGGGGCTTT
Non-CRISPR top	AGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTT AAATTAAAAATGAAGTTT
Non-CRISPR bottom	AAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATC TCATGACCAAAATCCCT
Non-CRISPR mutant top	AGGGATTTTGTGCATGAGCTTCTCCACCCGTCTCGTCCCCTAGATCCTTT TAAATTAAAAATGAAGTTT
Non-CRISPR mutant bottom	AAACTTCATTTTTAATTTAAAAGGATCTAGGGGACGAGACGGGTGGAGAA GCTCATGCACAAAATCCCT

Supplementary Table S2.

	% total reads relative to WT	
Substrate	**LR**	**RS**
WT	100	100
G1A, C30T	65	70
G1T, C30A	23	46
G1C, C30G	108	112
G1A	84	77
G1T	30	56
G1C	89	95
C30T	87	86
C30A	94	92
C30G	106	103
Block 1	51	62
Block 2	1	18
Block 3	9	56
Block 4	79	95
Block 5	50	55
Block 6	12	9
Block 7	70	76
Block 8	80	72
Block 9	55	80
Block 10	3	11
Short WT	100	100
Short Block 6	51	24
Block 6.1	46	38
Block 6.2	40	30
Block 6b	16	17
Insertion	5	2
Deletion	37	25
Non-CRISPR	110	106
Non-CRISPR mutant	0.01	1

Table S2. Percent of sequencing reads occurring at leader-repeat (LR) and repeat-spacer (RS) junctions of minimal CRISPR sequences, relative to WT normalized to 100%. Substrate 'Short Block 6' is calculated relative to 'Short WT.'