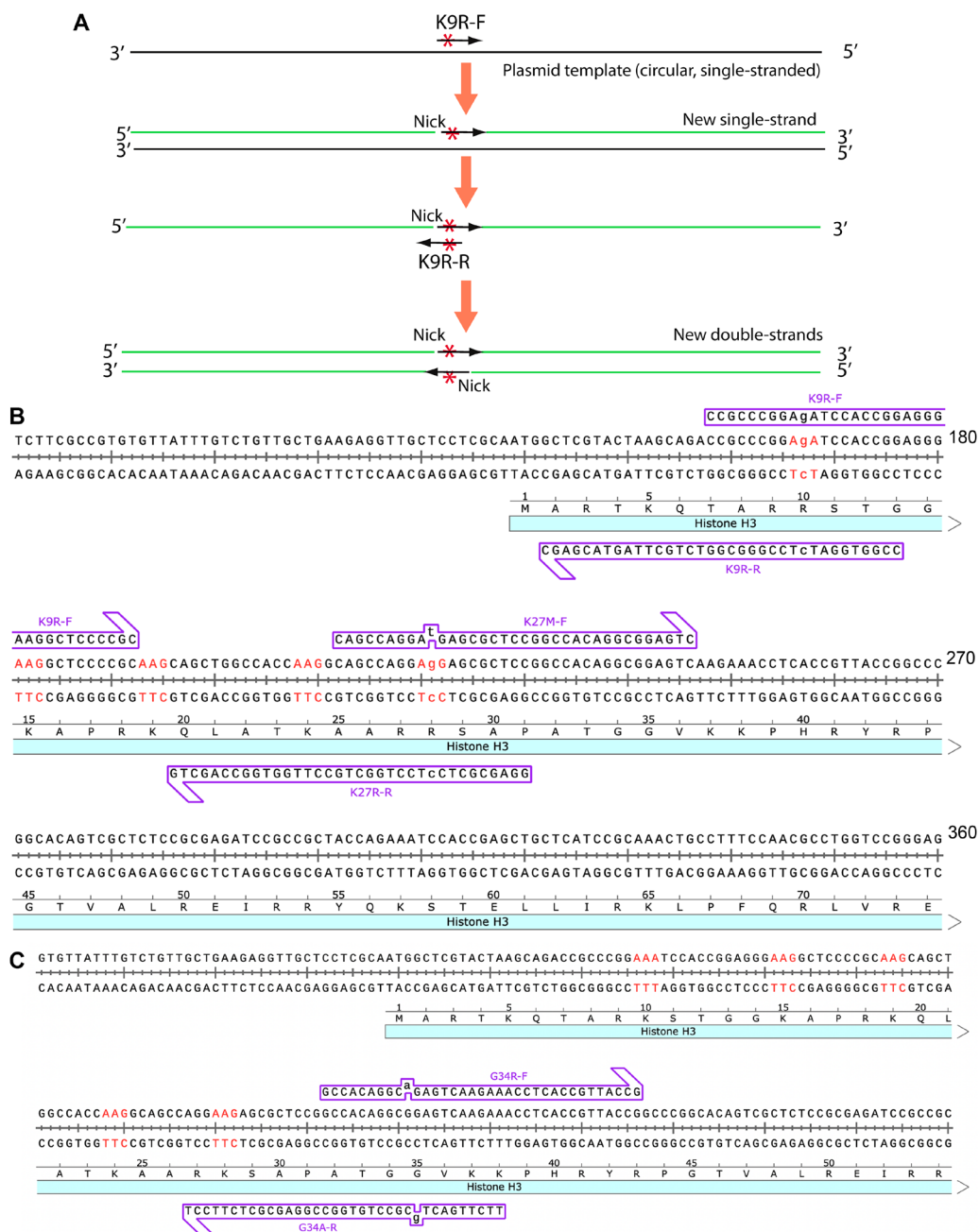
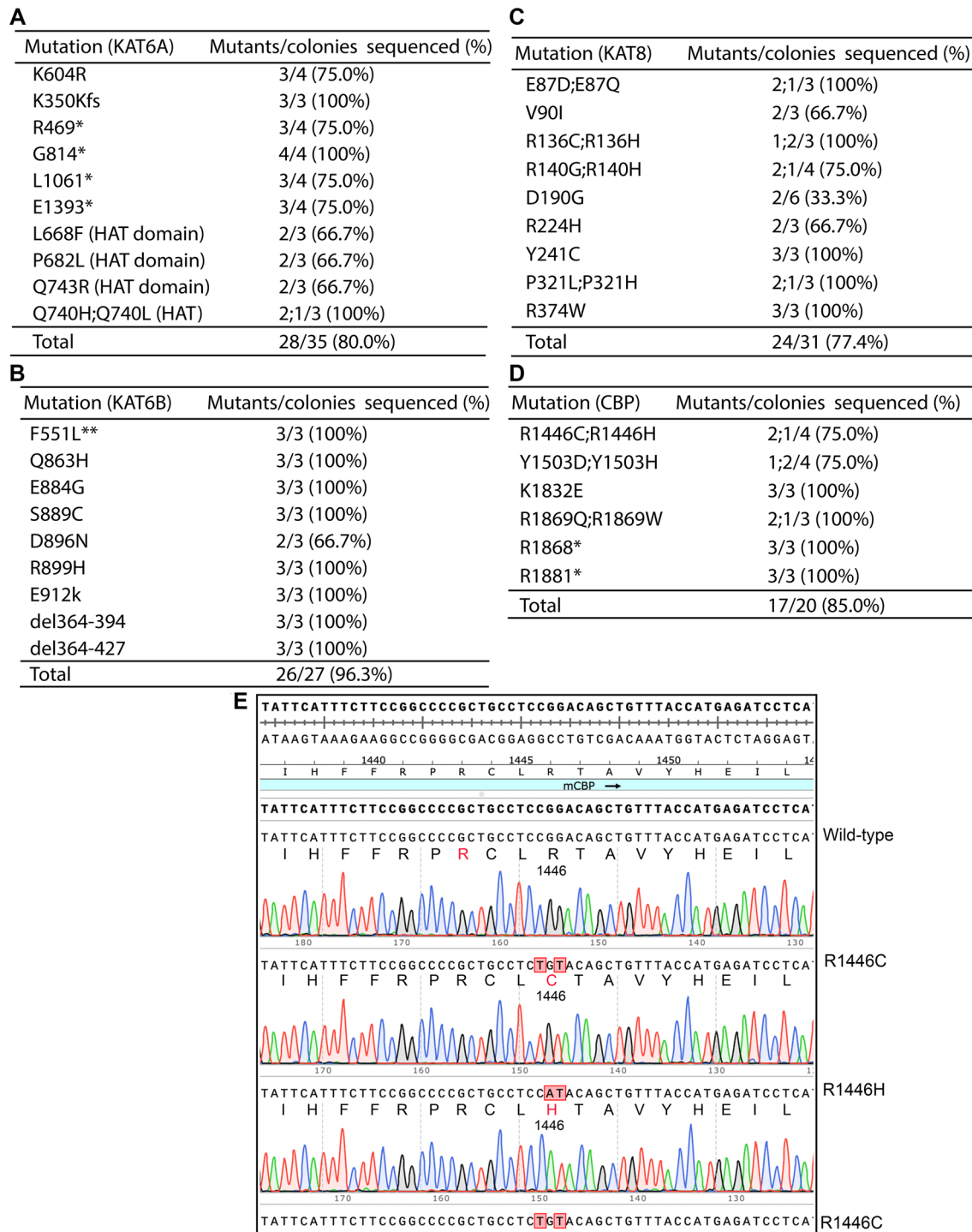


P3a site-specific and cassette mutagenesis for seamless protein, RNA and plasmid engineering

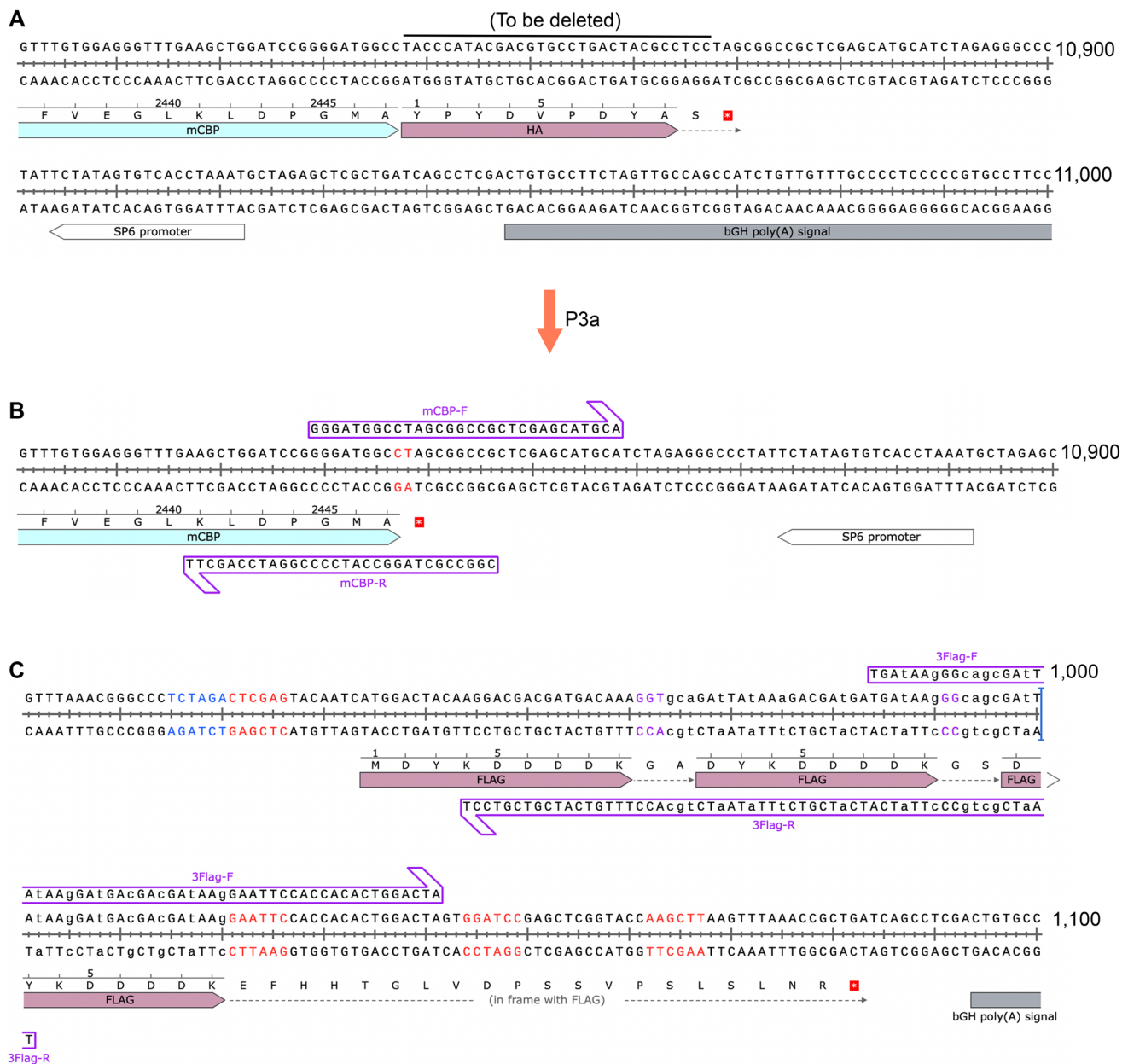
SUPPLEMENTARY MATERIALS



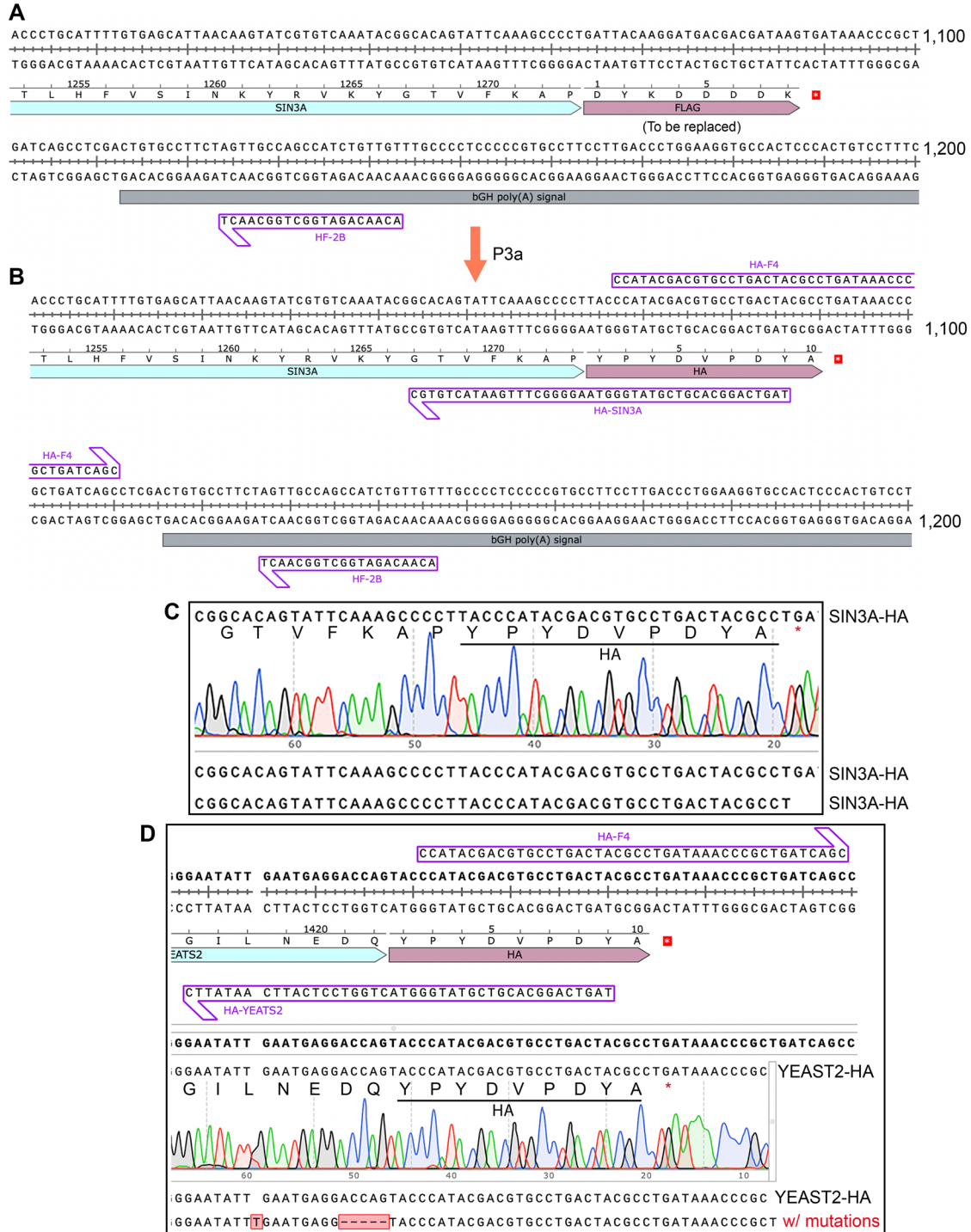
Supplementary Figure 1: Characteristics of mutagenic primer pairs with 3'-overhangs. (A) Schematic illustrating two partially complementary primers with 3'-overhangs (Figure 1B) directing synthesis from nicked newly synthesized DNA strands. In stark contrast, completely complementary primers (Figure 1A) cannot do so. One primer pair (K9R-F and K9R-R) with 3'-overhangs was used to engineer the K9R mutation in histone H3. (B) Sequences showing how two primer pairs were designed to engineer K9R and K27R/M. The codons of K9 and K27 have already been replaced with those for R9 and R27, respectively. Conventionally, for histone proteins, amino acid residue positions are numbered starting from the second residue (alanine), instead of the first methionine. As a result, K9 and K27 are encoded by the 10th and 28th codons, respectively. (C) Sequences showing how a single primer pair was designed to engineer G34A and G34R point mutations.



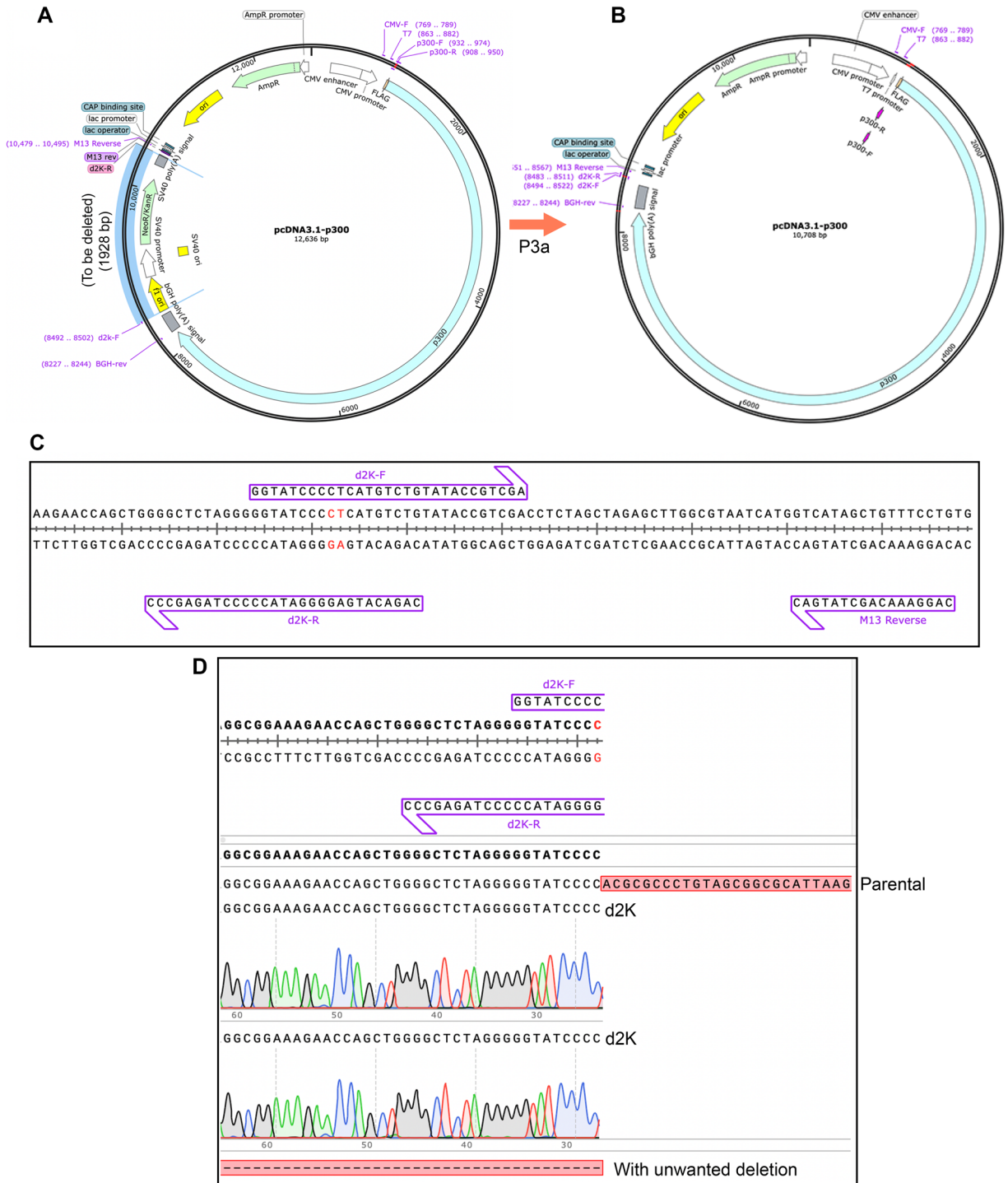
Supplementary Figure 2: P3a mutagenesis to engineer missense or nonsense mutants of epigenetic regulators. (A) Efficiency of engineering six KAT6A mutants and five mutants of its lysine acetyltransferase (HAT) domain. Asterisks in three C-terminal truncation mutants, such as G814*, refer to stop codons. (B) Efficiency of engineering seven mutants and two deletion mutants of KAT6B. A vector encoding residues 364-810 of KAT6B (GenBank accession, AF113514) was used for the experiments. The protein encoded by AF113514 carries an F551L substitution, possibly due to a sequencing error during cDNA cloning and analysis. F551 is invariant among different KAT6A and KAT6B proteins, so the mutant F551L was tested to assess the importance of F551, as denoted by two asterisks (**). (C) Efficiency in engineering 13 KAT8 mutants. The notation 'R136C;R136H' refers to generating two mutants R136C and R136H, with a single pair of primers (see Figure 1C). (D) Efficiency in engineering mouse CBP mutants. Asterisks in two C-terminal truncation mutants such as R1868* refer to stop codons. (E) Sequence chromatograms of four representative plasmids sequenced for engineering the CBP mutants R1446C and R1446H. The top chromatogram is the wild-type and the remaining three are mutants.



Supplementary Figure 3: Deleting or converting an epitope tag by P3a cassette mutagenesis. (A) Sequence of the region encoding the HA tag fused to the C-terminus of mouse CBP. The HA coding sequence to be deleted is indicated with a horizontal solid line. This expression vector is 13.4 kb in size. (B) Sequence of the region after deletion and the two partially complementary primers (mCBP-F and mCBP-R) that were used for deletion by P3a cassette mutagenesis. (C) Sequence of two partially complementary primers (3Flag-F and 3Flag-R) used to convert the coding sequence for a FLAG tag on pAW48 (a pcDNA3.1 derivative) to that for a 3xFLAG tag by P3a site-directed mutagenesis.



Supplementary Figure 4: Replacing a FLAG tag with an HA tag by P3a cassette mutagenesis. (A) Sequence of the region encoding the FLAG tag fused to the C-terminus of human SIN3A. The tag needs to be replaced with an HA tag for subsequent co-immunoprecipitation with partners to be expressed as FLAG fusion proteins. (B) Sequence of the resulting mutant and the two partially complementary primers (HA-F4 and HA-SIN3A) that were used for P3a cassette mutagenesis. (C) Sequence chromatograms of three candidate plasmids sequenced for engineering the replacement of the FLAG tag with an HA tag in the SIN3A-FLAG mammalian expression vector. The results indicate that all three are correct. (D) Sequence chromatograms of three candidate plasmids sequenced for engineering the replacement of the FLAG tag with an HA tag in the YEATS2-FLAG mammalian expression vector. The results indicate that two are correct and the third possesses a T insertion and 5-bp deletion just upstream from the coding sequence for the HA tag.

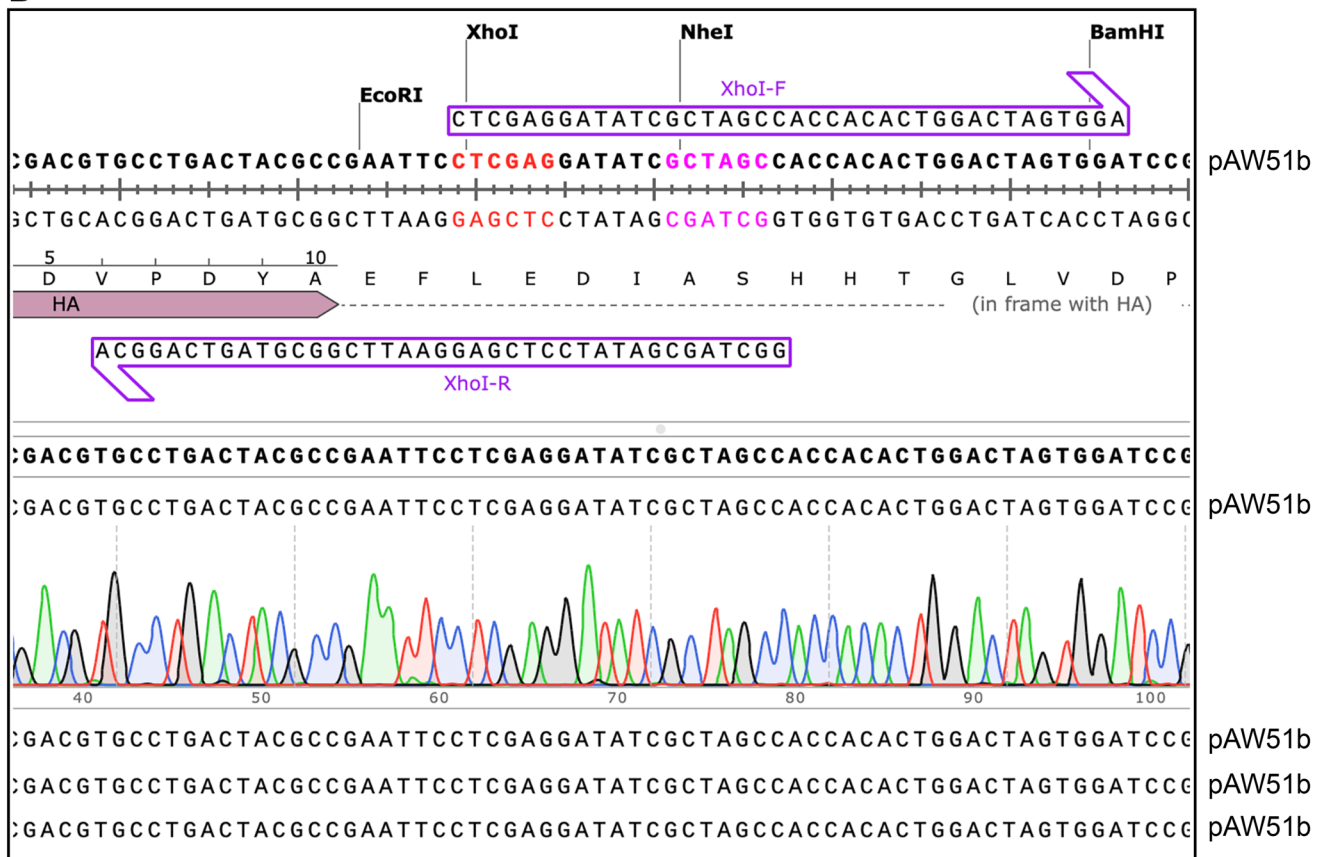


Supplementary Figure 5: Optimization of a 12.8-kb expression vector by deletion of a 1.9-kb fragment. (A) Map of the 12.6-kb mammalian expression vector for FLAG-p300. The 1.9-kb fragment to be removed by P3a mutagenesis is highlighted in blue. The original p300 expression vector from Addgene (Cat. 23252) is ~12.8 kb, whereas the one shown here results from introduction of the FLAG coding sequence, deletion of a 92-bp fragment and removal of the 6xHis coding sequence (see Figure 5B). (B) Map of the mutated 10.7-kb expression vector for FLAG-p300. d2k, deletion of ~2 kb. (C) Sequence of the resulting deletion mutant and the two partially complementary primers (d2k-F and d2K-R) that were used for P3a cassette mutagenesis. (D) Sequence chromatograms of four candidate plasmids sequenced for engineering the ~2-kb deletion. The top one is the wild-type, containing the ~2-kb fragment, the middle two are the mutant with the fragment deleted and the fourth possesses a large unwanted deletion.

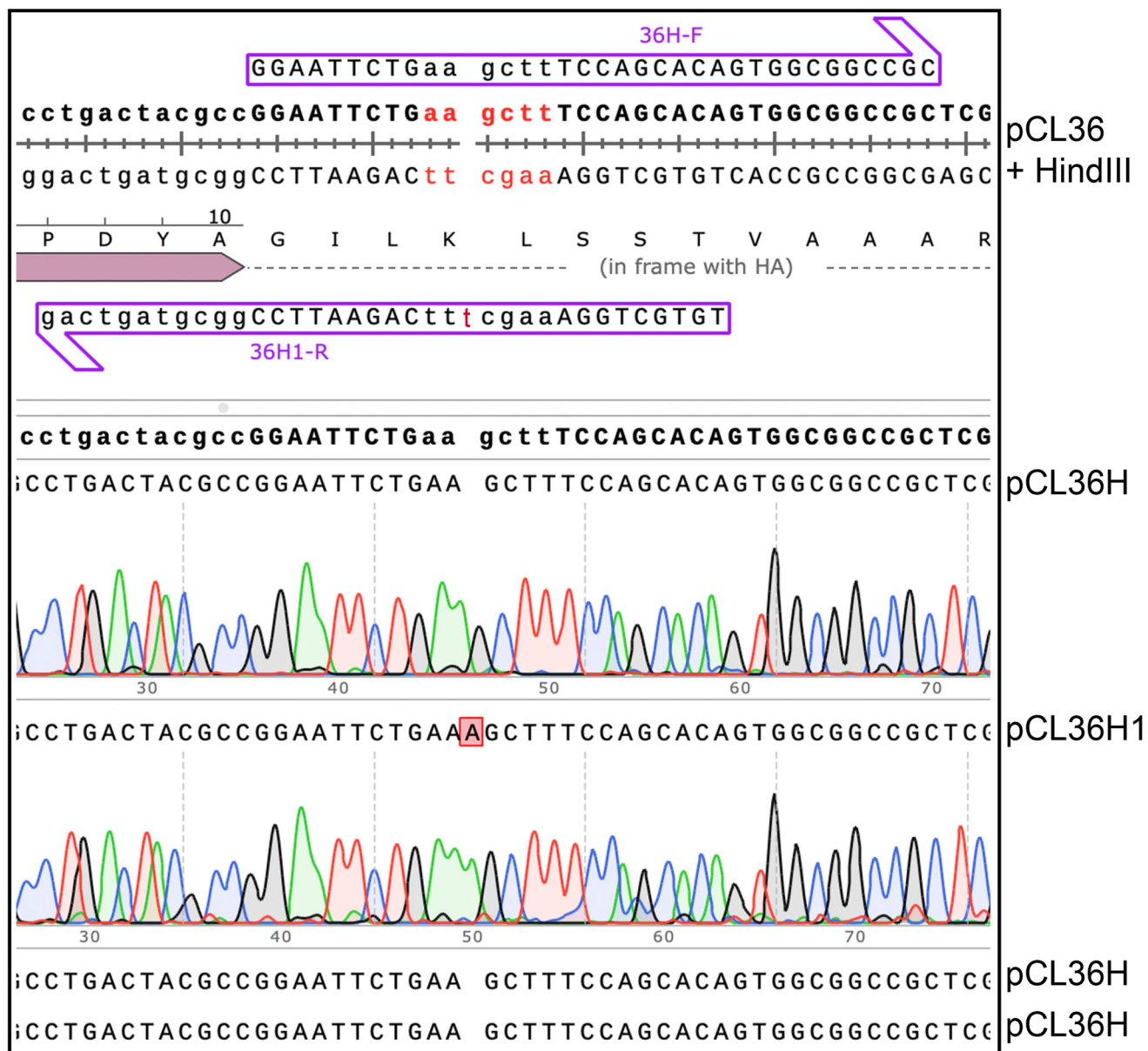
A

New restriction site	Mutants/colonies sequenced (%)
NheI elimination (pAW51)	4/4 (100%)
XhoI/NheI (pAW51a)	4/4 (100%)
HindIII (pCL36)	4/4 (100%)
HindIII (USP10)	3/4 (75%)
HindIII (SIN3B)	4/9 (44.4%)
XhoI/BamHI (Histone H4)	2/4 (50%)
Total	21/29 (72.4%)

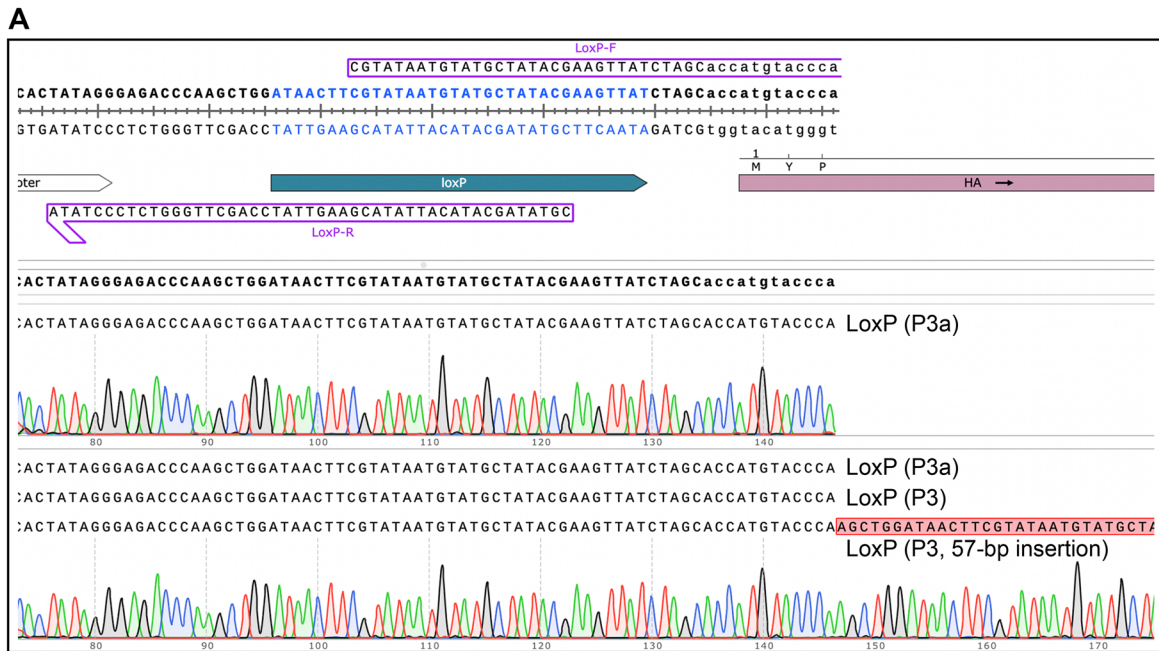
B



Supplementary Figure 7: P3a mutagenesis for engineering restriction sites on plasmids for subcloning. (A) Single or double restriction sites were engineered into (or eliminated from) the indicated vectors for subsequent subcloning. For pAW51, an NheI site upstream from the coding sequence for the HA tag was first eliminated to produce pAW51a, which was subsequently used for the introduction of an XhoI site and an NheI site downstream from the HA coding sequence (panel B), for subcloning of open reading frames flanked by these two sites. For pCL36, an HindIII site was inserted to produce two vectors possessing different reading frames downstream from the coding sequence for the HA tag (See Supplementary Figure 8). Notably, for the histone H4 plasmid, two restriction sites were engineered upstream and downstream from its open reading frame, indicating that the method is applicable for introducing two distant mutations simultaneously. (B) Sequence chromatograms of four representative plasmids sequenced for engineering the XhoI and NheI sites into pAW51a. All were correct mutants, yielding an ideal efficiency of 100%.

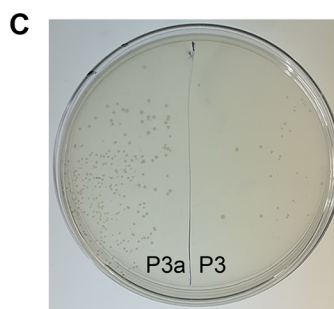


Supplementary Figure 8: P3a mutagenesis for engineering a HindIII site on a plasmid for subcloning. Sequence chromatograms of four representative plasmids sequenced for engineering a HindIII site into the mammalian expression vector pCL36. The two primers (36H-F and 36H1-R) were designed to engineer the HindIII site at different reading frames upstream from the coding sequence for the HA tag, thereby resulting in two plasmids (pCL36H and pCL36H1, respectively).

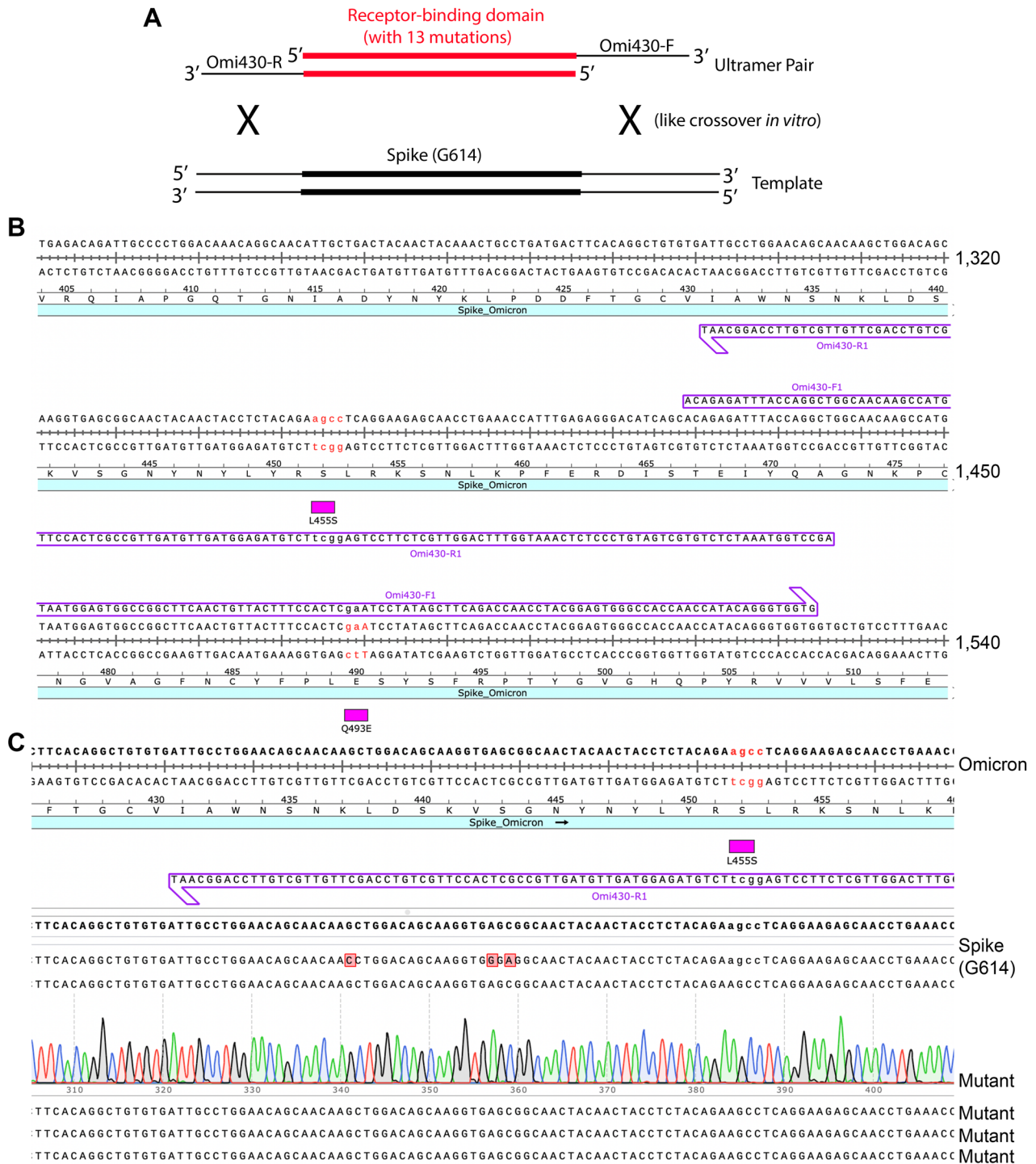


B Efficiency of P3 cassette mutagenesis in engineering deletion and insertion

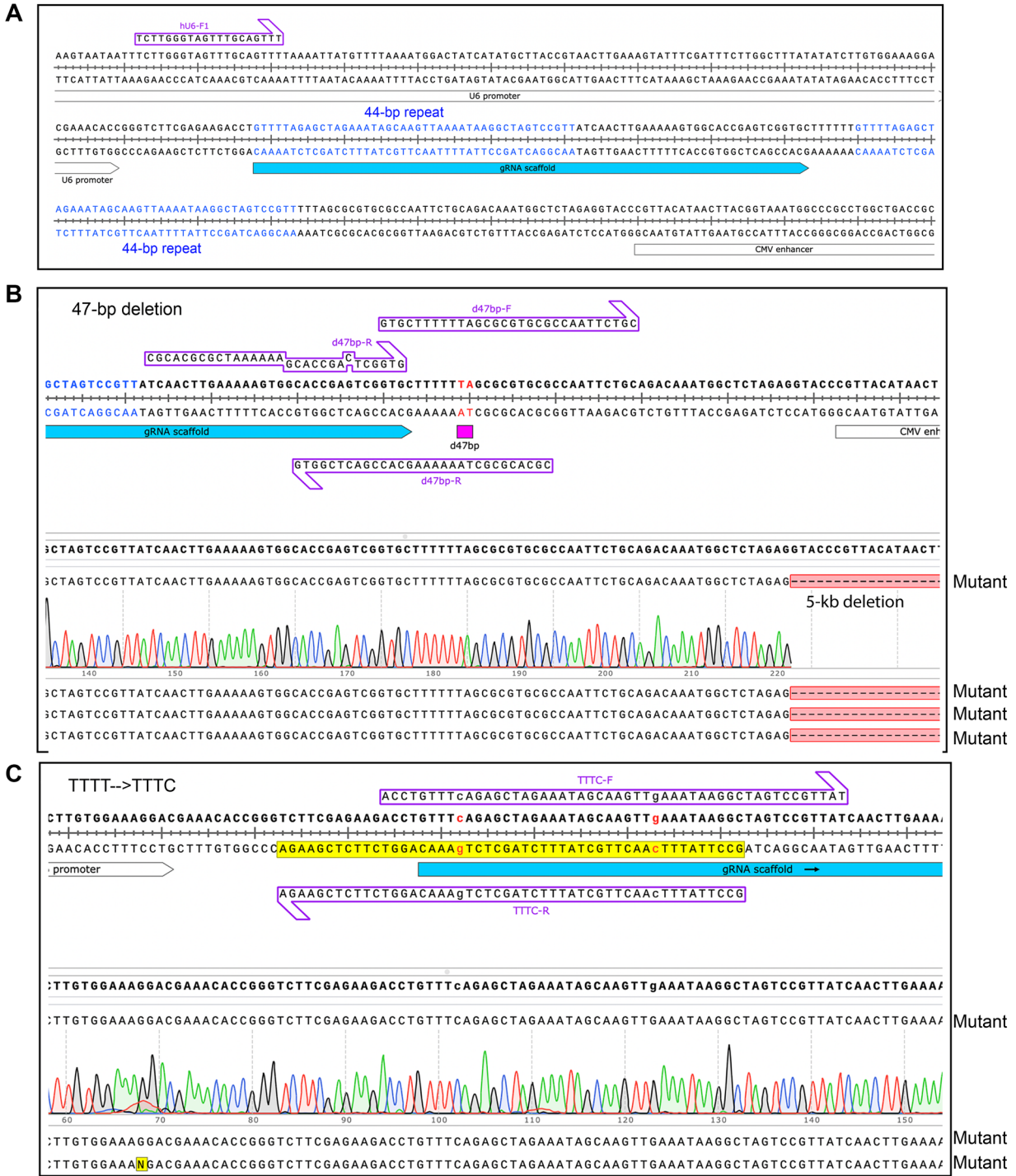
Deletion Mutants	Plasmids Sequenced (%)	Deletion (bp)
BRPF1-dN131	4/6 (66.7%)	393
BRPF1-dN204	3/6 (50%)	612
BRPF2-d31-80	2/6 (33.3%)	150
p300-dN1301	2/5 (40%)	3,903
CBP-dN1068	6/7 (85.7%)	3,204
p300 vector_d2K	3/6 (50%)	1,928
LoxP (P3)	4/6 (66.7%)	34 bp (insertion)
LoxP (P3a)	6/6 (100%)	34 bp (insertion)
Total	24/42 (57.1%; 56.1±16.7%)*	n/a



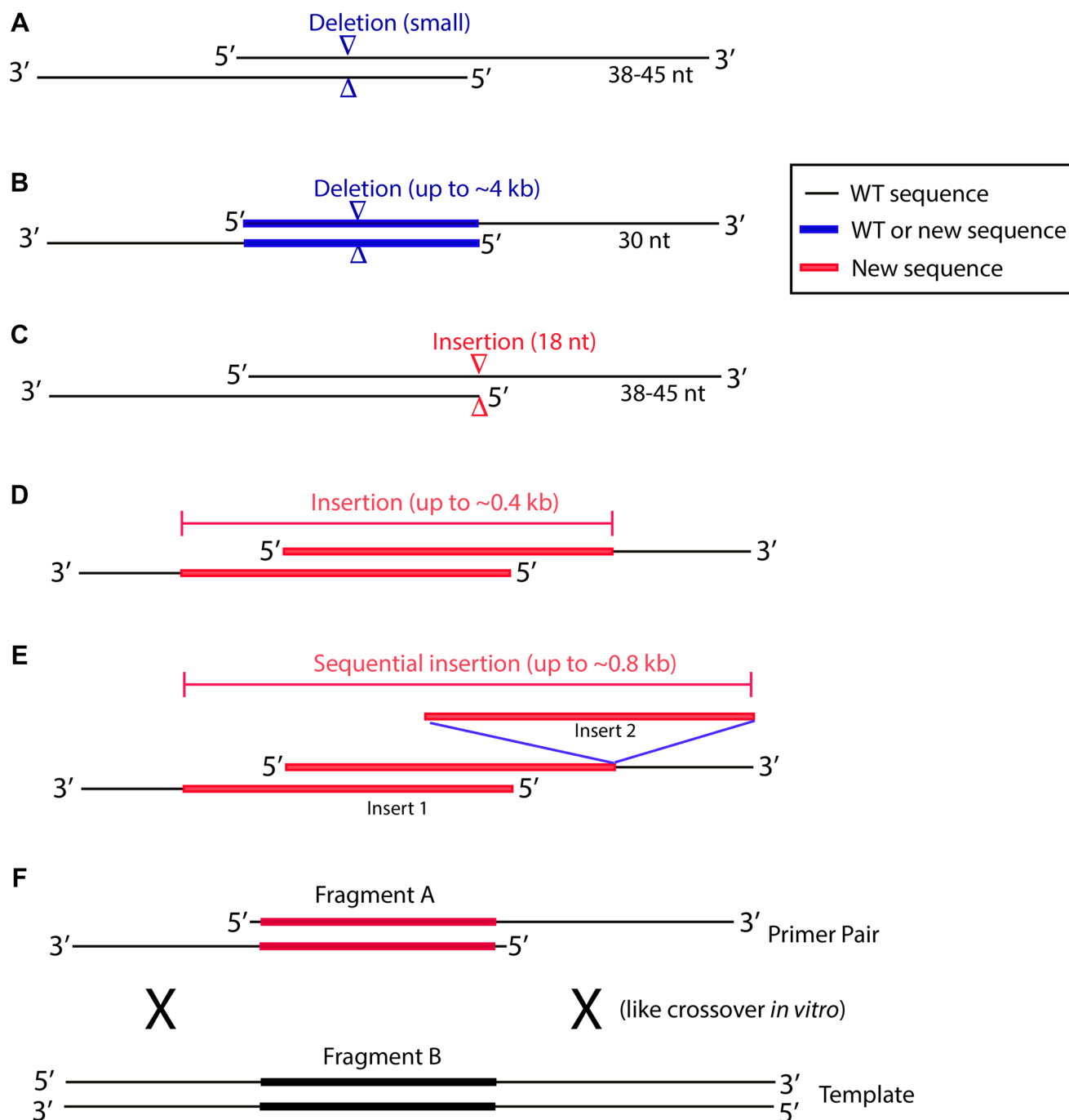
Supplementary Figure 9: P3a and P3 mutagenesis methods for engineering LoxP sites and introducing deletions. (A) Sequence chromatograms of four representative plasmids sequenced for engineering the LoxP site into the mammalian expression vector for HA-tagged BRPF1. The two primers (LoxP-F and LoxP-R) were designed to engineer the LoxP site upstream from the coding sequence for the HA tag. Clones C1 and C2 were from P3a mutagenesis, whereas clones C3 and C4 were from P3 mutagenesis. For each mutagenesis reaction, plasmids from 6 colonies were analyzed by sequencing. All those from P3a mutagenesis contained the LoxP site, resulting in the ideal efficiency of 100%. For those from P3 mutagenesis, 4 harbored the LoxP site, leading to an efficiency of 4/6 (66.7%). (B) Efficiency of P3 cassette mutagenesis to engineer deletion of 6 fragments and insertion of a LoxP site. The asterisk denotes the average efficiency from these 7 mutations. For comparison, the efficiency for insertion of a LoxP site via P3a cassette mutagenesis is also shown here. (C) Photo of bacterial colonies from P3 and P3a mutagenesis reactions to introduce the LoxP site. For each mutagenesis reaction, 2 μ l of the mixture from DpnI digestion was used to transform 20 μ l DH5a competent cells. After transformation, the bacterial cells from the two mutagenesis reactions were plated out on the two sides of one LB-agar plate containing ampicillin.



Supplementary Figure 11: P3a mutagenesis to replace the receptor-binding domain of SARS-CoV-2 spike protein. (A) Cartoon showing an Ultra primer pair (Omi430-F and Omi430-R) to replace the coding sequence for the receptor-binding domain of the spike protein. (B) Two Ultra primers (Omi430-F and Omi430-R) were designed to replace the coding sequence for the receptor-binding domain of the spike protein. (C) Sequence chromatograms of four representative plasmids from P3a mutagenesis to replace the coding sequence replace the coding sequence for the receptor-binding domain of the spike protein. All of them were correct, leading to the ideal efficiency of 100%.



Supplementary Figure 12: P3a mutagenesis of Cas9 and guide RNA vectors for genome editing. (A) The Cas9 expression vectors pX330, pX459 and some of their derivatives possess two 44-bp repeats in the sgRNA scaffold and its downstream region. The repeats are problematic to PCR primers designed to mutate the scaffold or insert sgRNA-coding sequences just from the scaffold. (B) Two primers (d47bp-F and d47bp-R) were designed to remove the second 44-bp repeat along with its downstream TTT sequence. Four colonies were analyzed and all of them contained the correct mutations, leading to the ideal efficiency of 100%. (C) Sequence chromatograms of three representative plasmids from P3a mutagenesis to replace the TTTT stretch in the sgRNA scaffold with TTTC. The corresponding downstream AAAA was changed to GAAA. Two primers (TTTC-F and TTTC-R) were designed to engineer these two point mutations. All three clones contained the correct mutations, resulting in the ideal efficiency of 100%.



Supplementary Figure 13: Different primer design strategies used to introduce deletion and insertion. (A, B) Two distinct strategies to design a primer pair with 3'-overhangs for engineering deletion. The strategy depicted in panel A, as employed in the previous study [1], assumes that the overlapping region (except for the deletion) is identical to the wild-type template. This is not required for the strategy depicted in panel B, employed in the current study; the sequence could be a redesigned junction flanking the deletion. For example, for a protein-encoding gene, the sequence could be codon-optimized or even an epitope tag. The strategy depicted in panel B is actually a special case of replacement mutagenesis (see panel F). (C, D) Two different strategies to design a primer pair with 3'-overhangs for engineering insertion. For the strategy depicted in panel C, as employed in a previous study [1], the entire insertion is located within the overlapping regions of the two primers, with one primer carrying the insertion in the middle and another harboring the insertion at the 5'-end. By contrast, for the strategy shown in panel D, the inserted sequence is encoded within the overlapping region and the two 3'-overhangs. For large insertion, this reduces the costs by almost a half. The strategy depicted in panel D is a special case of replacement mutagenesis (see panel F). (E) Sequential insertion mutagenesis of two pairs of Ultramer primers allows circumventing the size limit of mutagenesis with one pair of such primers. Megamer™ single-stranded DNA fragments from IDT are another viable alternative to circumvent the size limit, although they are more expensive than regular oligos or Ultramers. (F) Scheme showing replacement (i.e., 'cassette') mutagenesis, where a primer pair containing the sequence of fragment A is used to replace fragment B. Conceptually, deletion and insertion are special cases where the respective sequences to be inserted and deleted are zero bp. This is because replacement mutagenesis converts fragment A to fragment B. For deletion, fragment B is zero bp in size, whereas for insertion, fragment A is zero bp.

SUPPLEMENTARY REFERENCES

1. Liu H, Naismith JH. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. BMC Biotechnol. 2008; 8:91. <https://doi.org/10.1186/1472-6750-8-91>. PMID:19055817