Phage-assisted evolution of an adenine base editor with enhanced Cas domain compatibility and activity

Michelle F. Richter[1,2,3,*], Kevin T. Zhao[1,2,3,*], Elliot Eton[1,2,3], Audrone Lapinaite[4], Gregory A. Newby[1,2,3], Benjamin W. Thuronyi[1,2,3], Christopher Wilson[1,2,3], Luke W. Koblan[1,2,3], Jing Zeng[5,6,7], Daniel E. Bauer[5,6,7], Jennifer A. Doudna[4,8,9,10,11], David R. Liu[1,2,3]

SUPPLEMENTARY INFORMATION

Supplementary Figure 1. Base editing efficiencies and indel frequencies for SpABE7.10 and SpABE8a-e-dimer variants at four genomic sites in HEK293T cells.

Supplementary Figure 2. Base editing efficiencies and indel frequencies for SaABE7.10 and SaABE8a-e-dimer variants at four genomic sites in HEK293T cells.

Supplementary Figure 3. Base editing efficiencies and indel frequencies for LbABE7.10 and LbABE8a-e-dimer variants at three genomic sites in HEK293T cells.

Supplementary Figure 4. Base editing efficiencies and indel frequencies for SpABE7.10, SpABE8e, CP1028-ABE7.10, CP1028-ABE8e, CP1028-ABE8e-dimer, CP1041-ABE7.10, CP1041-ABE8e, and CP1041-ABE8e-dimer at seven genomic sites in HEK293T cells**.**

Supplementary Figure 5. Base editing efficiencies and indel frequencies for SaABE7.10, SaABE8e, SaABE8e-dimer, SaKKH-ABE7.10, SaKKH-ABE8e, and SaKKH-ABE8e-dimer in HEK293T cells at five genomic sites with NNGRRT PAMs in HEK293T cells.

Supplementary Figure 6. Base editing efficiencies and indel frequencies for LbABE7.10, LbABE8e, Lb-ABE8e-dimer, enAsABE7.10, enAsABE8e, and enAsABE8e-dimer in HEK293T cells at three genomic sites in HEK293T cells.

Supplementary Figure 7. Base editing efficiencies and indel frequencies for NG-ABE7.10, NG-ABE8e, and NG-ABE8e-dimer in HEK293T cells at five genomic sites in HEK293T cells.

Supplementary Figure 8. Base editing efficiencies and indel frequencies for SaABE7.10, SaABE8e, SaABE8e-dimer, SaKKH-ABE7.10, SaKKH-ABE8e, and SaKKH-ABE8e-dimer in HEK293T cells at nine genomic sites with NNHRRT PAMs in HEK293T cells.

Supplementary Figure 9. Processivity analysis for ABE7.10 and ABE8e.

Supplementary Figure 10. Comparison of ABE8e and Cas9 nuclease at the top three off-target sites for *EMX1* as identified by GUIDE-Seq.

Supplementary Figure 11. On-target DNA base editing efficiencies comparing ABE8e mutants with minimized transcriptome-wide RNA off-target editing.

Supplementary Figure 12. On-target base editing efficiencies and indel frequencies for ABE7.10, ABE8e, and ABE8e (TadA-8e V106W) in HEK293T cells at seven genomic sites in HEK293T cells.

Supplementary Figure 13. On-target base editing efficiencies and indel frequencies for SaABE7.10, SaABE8e, and SaABE8e (TadA-8e V106W) in HEK293T cells at four genomic sites in HEK293T cells.

Supplementary Figure 14. On-target base editing efficiencies and indel frequencies for LbABE7.10, LbABE8e, and LbABE8e (TadA-8e V106W) in HEK293T cells at six genomic sites in HEK293T cells.

Supplementary Figure 15. On-target base editing efficiencies and indel frequencies for ABE7.10, ABE8e, and ABE8e (TadA-8e V106W) in HEK293T cells at site 3 in HEK293T for the orthogonal R-loop assay.

Supplementary Figure 16. Allele compositions following treatment with ABE7.10 or ABE8e at a GATA1 binding site of a *BCL11A* enhancer.

Supplementary Figure 17. Allele compositions following treatment with ABE7.10 or ABE8e at the *HBG1* and *HBG2* promoter.

Supplementary Table 1. Gene circuit components.

Supplementary Table 2. Selection schedule for PANCE.

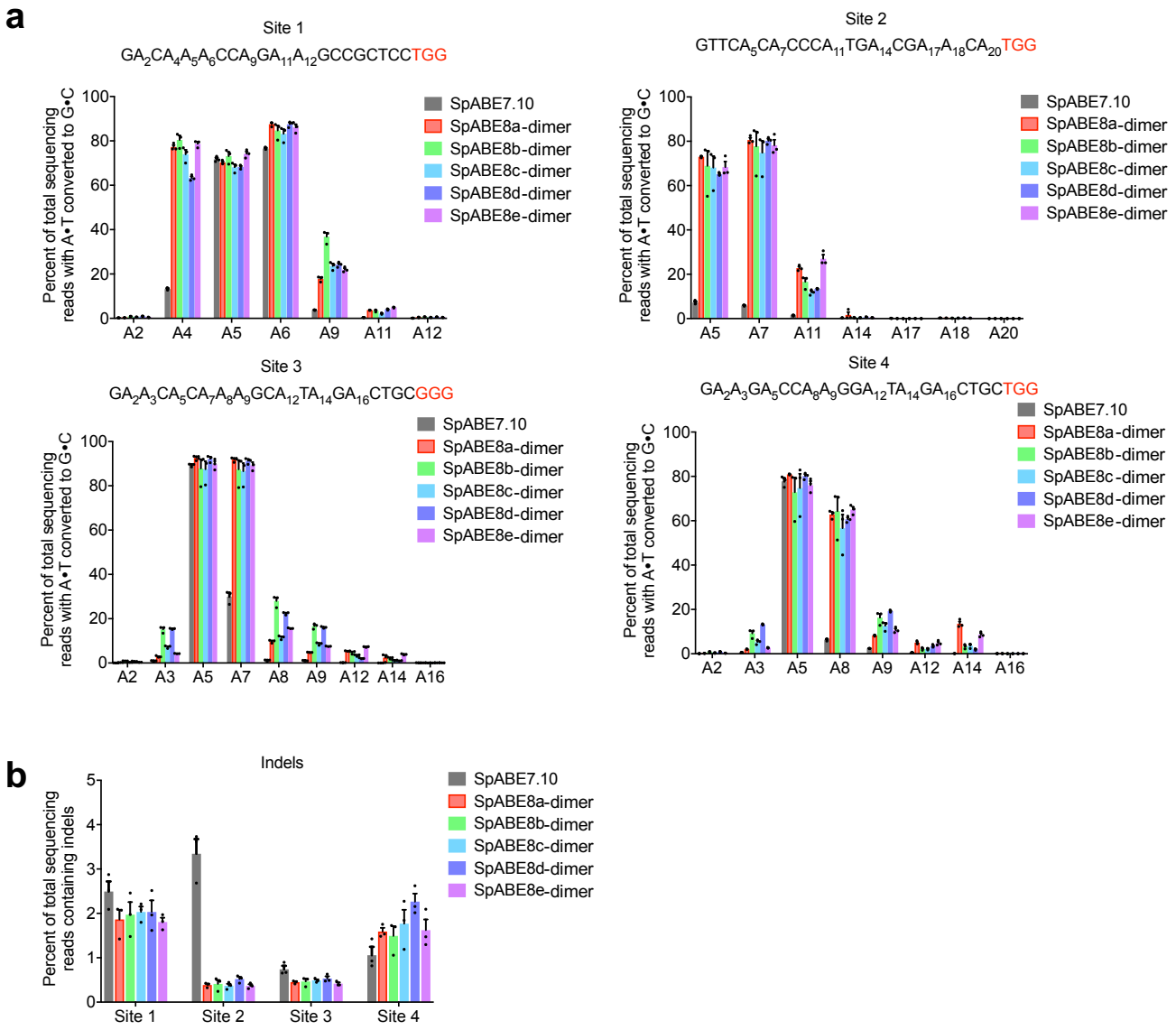
Supplementary Table 3. Protospacer sequences for mammalian genomic sites.

Supplementary Table 4. Primers used for mammalian cell genomic DNA amplification.

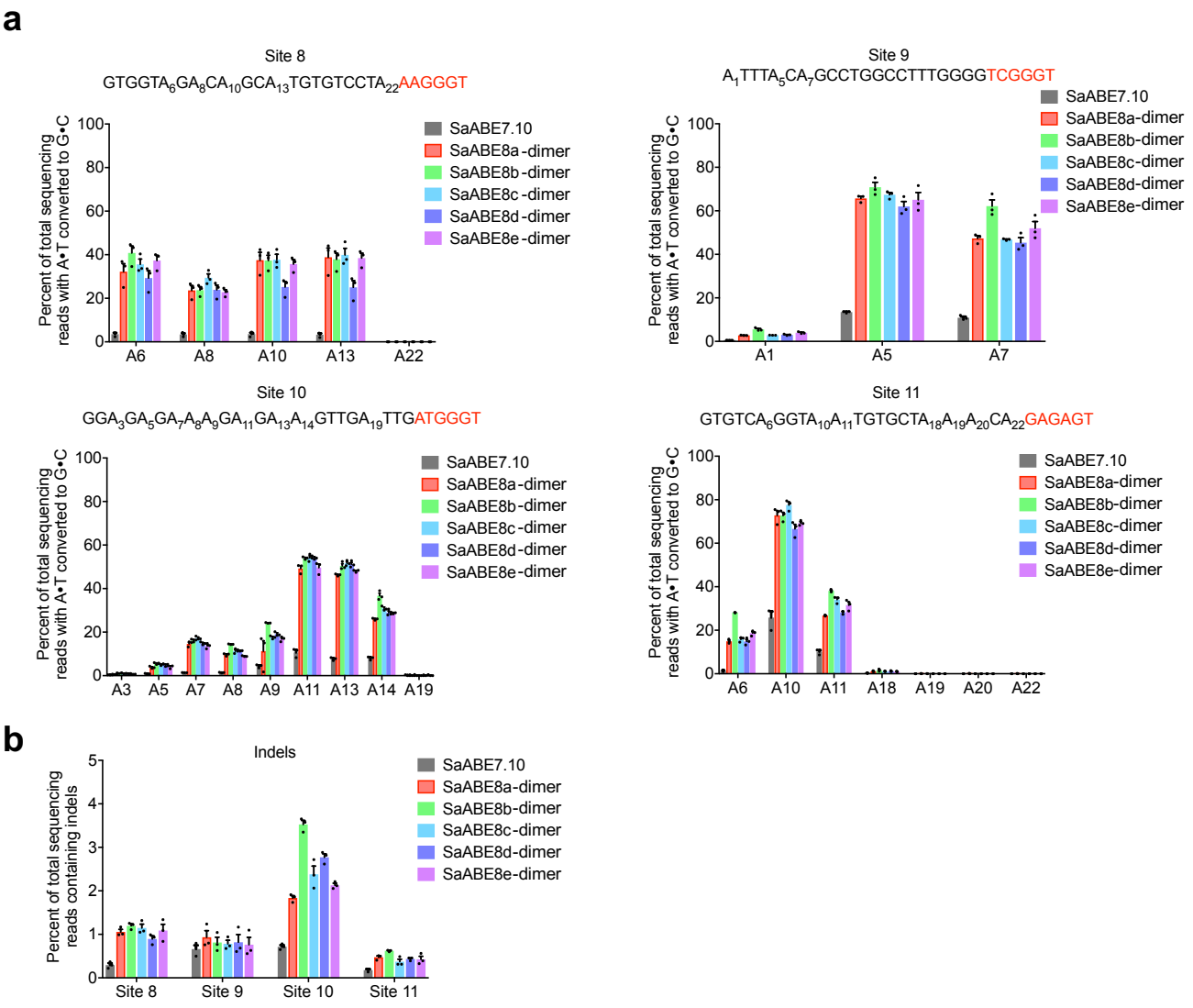
Supplementary Table 5. Amplicons for high-throughput sequencing analysis.

Supplementary Note 1. Python script to extract allele frequencies with two edits.

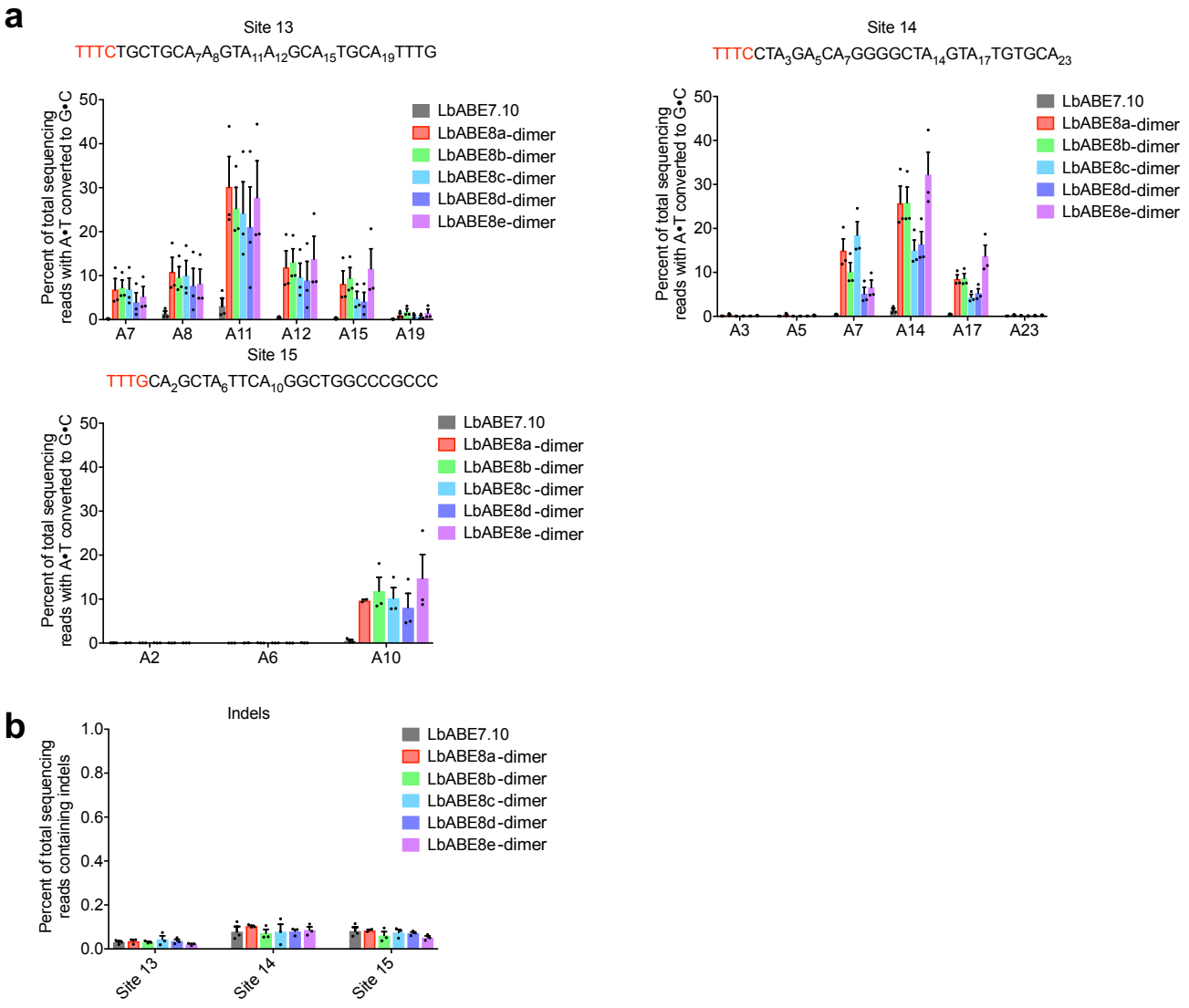
Supplementary Sequences 1. Amino acid sequences for each base editor in this study.



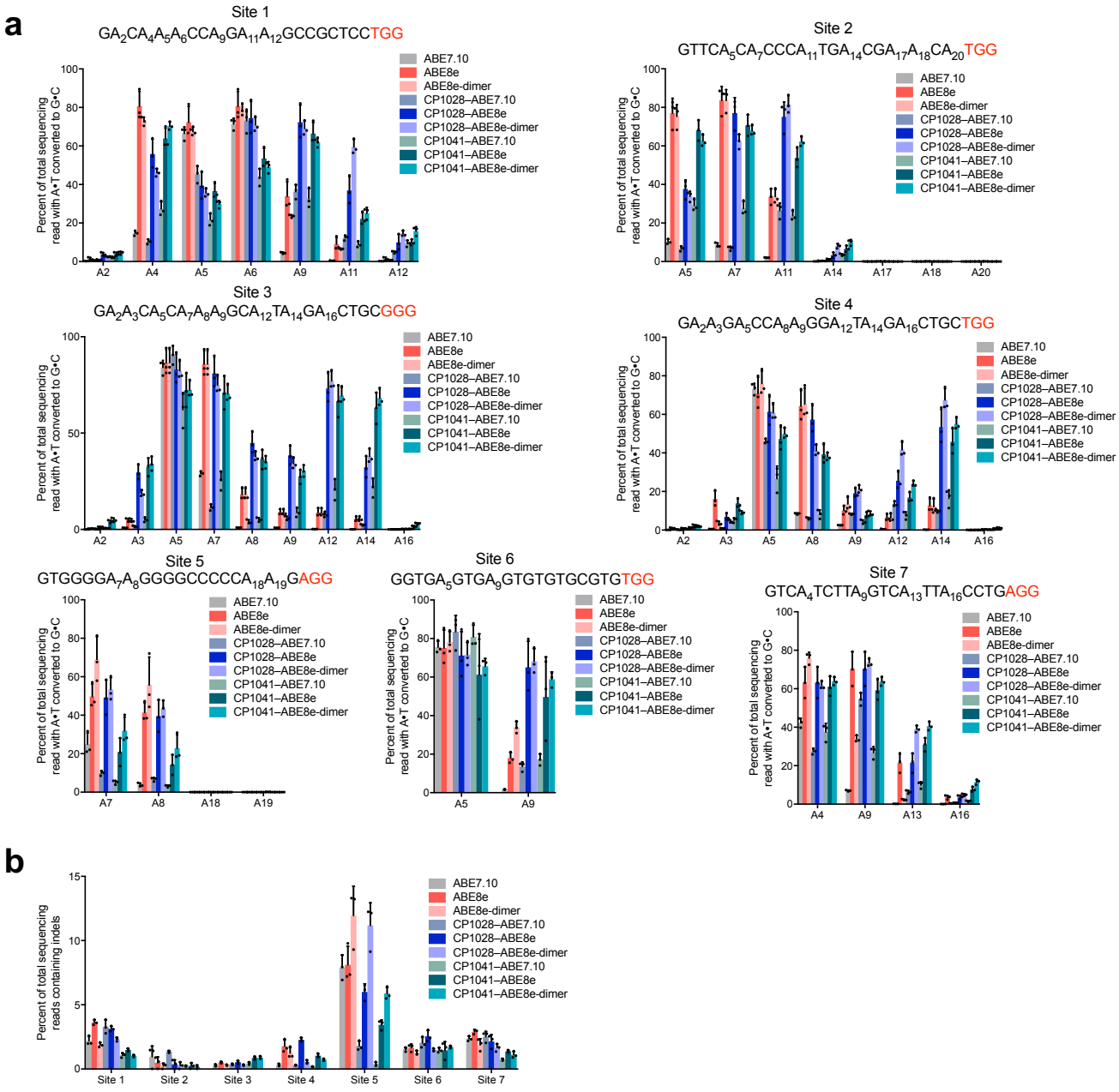
Supplementary Figure 1. Base editing efficiencies and indel frequencies for SpABE7.10 and SpABE8a-e-dimer variants at four genomic sites in HEK293T cells. a, Base editing with SpABE7.10, SpABE8a-dimer, SpABE8b-dimer, SpABE8c-dimer, SpABE8d-dimer, and SpABE8e-dimer at four genomic sites in HEK293T cells. **b**, Percent of all sequencing reads containing an indel following modification by SpABE7.10-dimer, SpABE8a-dimer, SpABE8b-dimer, SpABE8c-dimer, SpABE8d-dimer, or SpABE8e-dimer at four genomic sites in HEK293T cells. All base editors are optimized with the architecture, codon usage, and nuclear localization signals of ABEmax[16]. All untreated controls show no editing or indel formation at all sites. For all plots, dots represent individual biological replicates and bars represent mean±s.d. of three independent biological replicates.



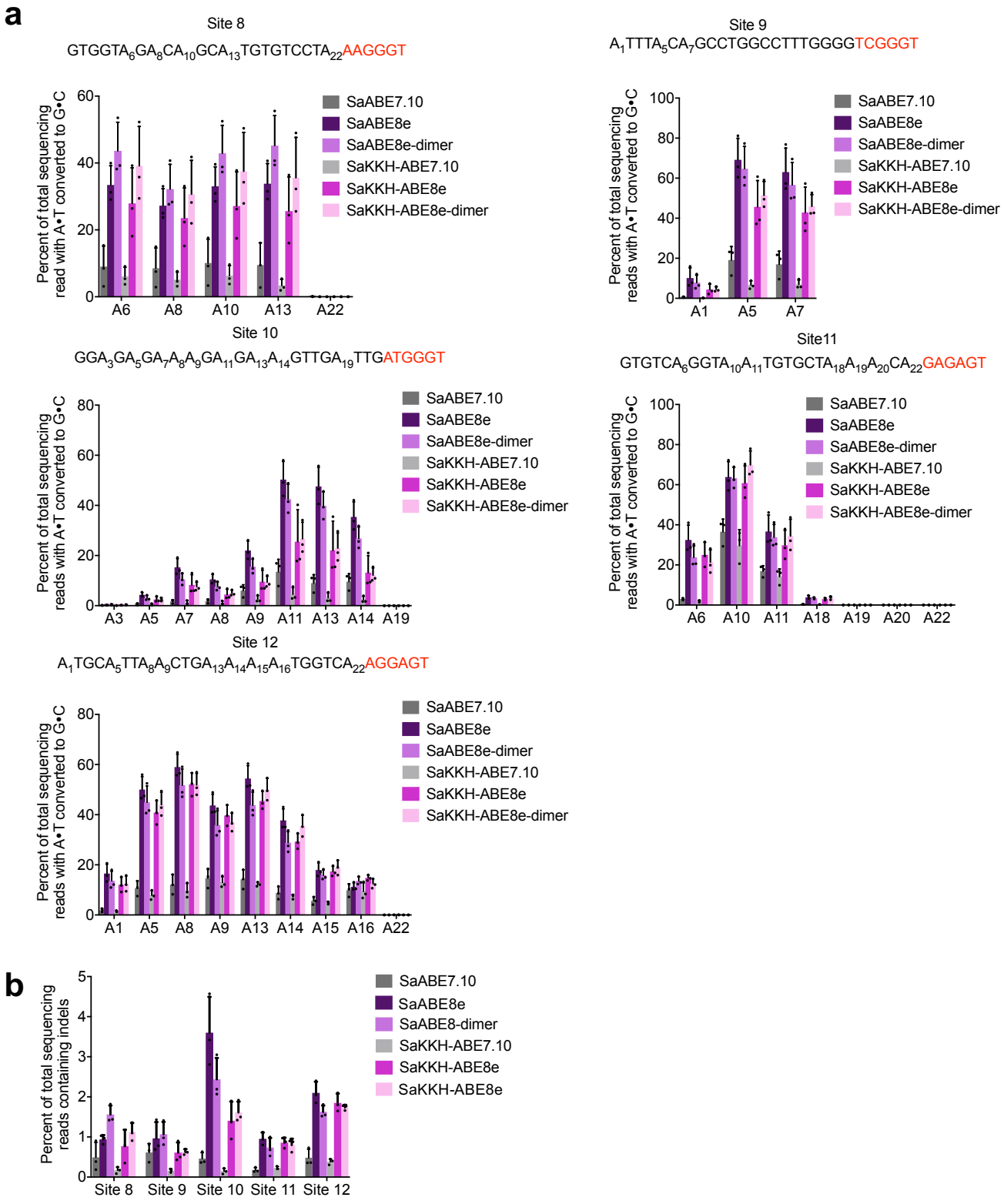
Supplementary Figure 2. Base editing efficiencies and indel frequencies for SaABE7.10 and SaABE8a-e-dimer variants at four genomic sites in HEK293T cells. a, Base editing with SaABE7.10, SaABE8a-dimer, SaABE8b-dimer, SaABE8c-dimer, SaABE8d-dimer, and SaABE8e-dimer at four genomic sites in HEK293T cells. **b**, Percent of all sequencing reads containing an indel following modification by SaABE7.10, SaABE8a-dimer, SaABE8b-dimer, SaABE8c-dimer, SaABE8d-dimer, or SaABE8e-dimer at four genomic sites in HEK293T cells. All base editors are optimized with the architecture, codon usage, and nuclear localization signals of ABEmax[16]. All untreated controls show no editing or indel formation at all sites. For all plots, dots represent individual biological replicates and bars represent mean±s.d. of three independent biological replicates.



Supplementary Figure 3. Base editing efficiencies and indel frequencies for LbABE7.10 and LbABE8a-e-dimer variants at three genomic sites in HEK293T cells. a, Base editing with LbABE7.10, LbABE8a-dimer, LbABE8b-dimer, LbABE8c-dimer, LbABE8d-dimer, and LbABE8e-dimer at three genomic sites in HEK293T cells. **b**, Percent of all sequencing reads containing an indel following modification by LbABE7.10, LbABE8a-dimer, LbABE8b-dimer, LbABE8c-dimer, LbABE8d-dimer, or LbABE8e-dimer at three genomic sites in HEK293T cells. LbCas12a was used in the nuclease inactive form since the desired nickase enzyme for base editing has not been reported. Otherwise, all base editors are optimized with the architecture, codon usage, and nuclear localization signals of ABEmax[16]. All untreated controls show no editing or indel formation at all sites. For all plots, dots represent individual biological replicates and bars represent mean±s.d. of three independent biological replicates.

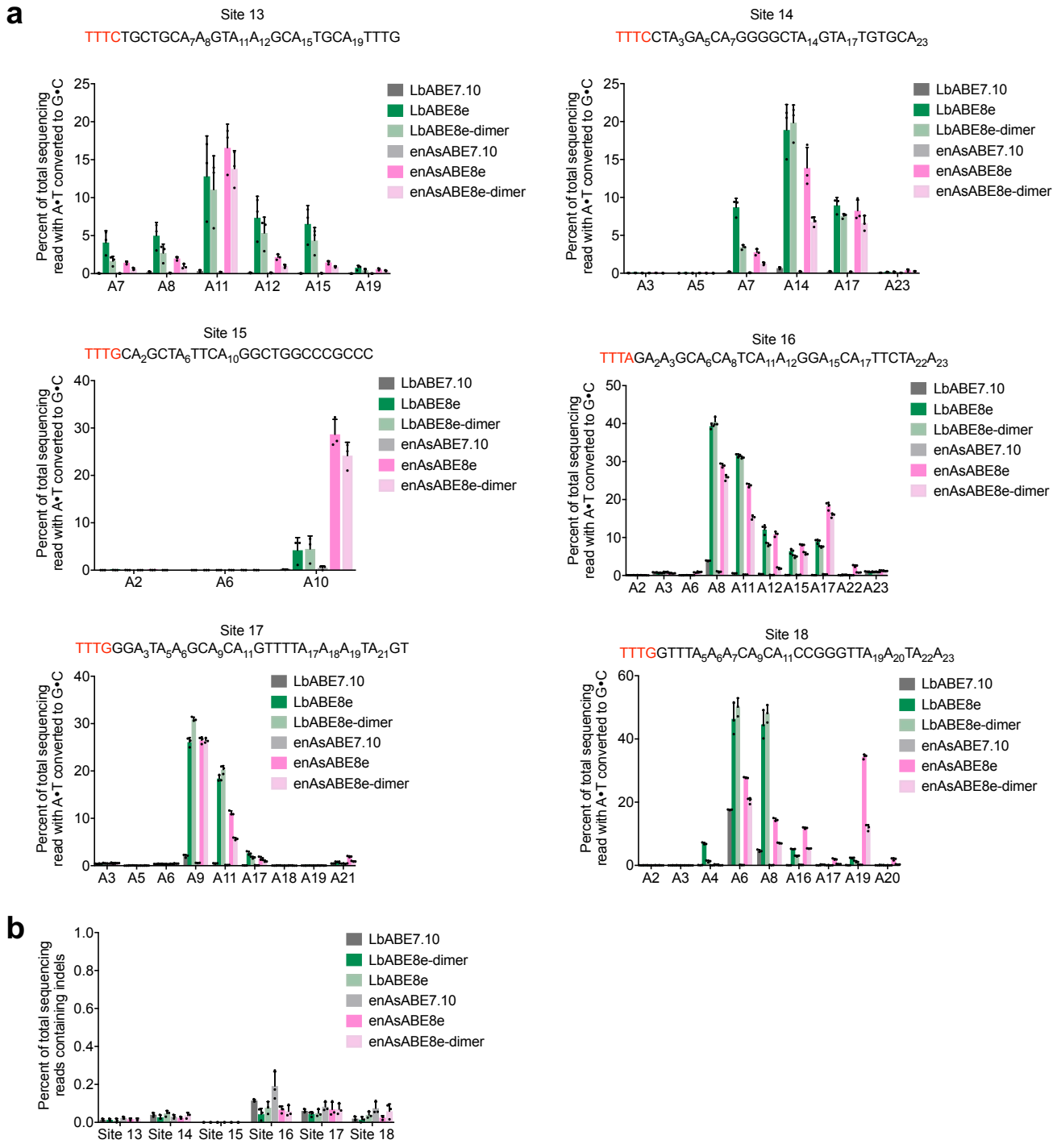


Supplementary Figure 4. Base editing efficiencies and indel frequencies for SpABE7.10, SpABE8e, CP1028-ABE7.10, CP1028-ABE8e, CP1028-ABE8e-dimer, CP1041-ABE7.10, CP1041-ABE8e, and CP1041-ABE8e-dimer at seven genomic sites in HEK293T cells. a, Base editing with SpABE7.10, SpABE8e, CP1028-ABE7.10, CP1028-ABE8e, CP1028-ABE8e-dimer, CP1041-ABE7.10, CP1041-ABE8e, and CP1041-ABE8e-dimer at seven genomic sites in HEK293T cells. **b**, Percent of all sequencing reads containing an indel following modification by SpABE7.10, SpABE8e, CP1028-ABE7.10, CP1028-ABE8e, CP1028-ABE8e-dimer, CP1041-ABE7.10, CP1041-ABE8e, or CP1041-ABE8e-dimer at seven genomic sites in HEK293T cells. All untreated controls show no editing or indel formation at all sites. For all plots, dots represent individual biological replicates and bars represent mean±s.d. of three independent biological replicates.



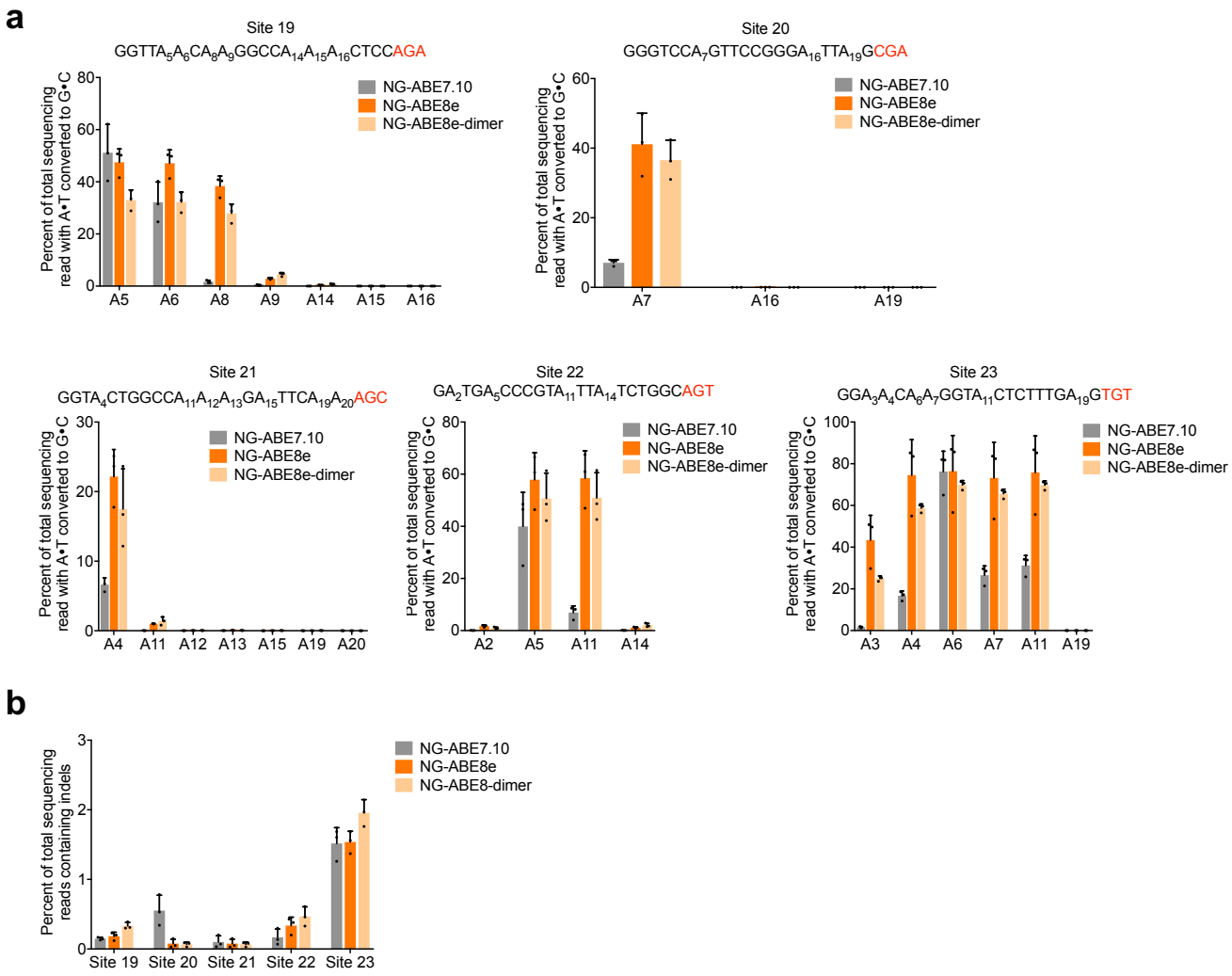
Supplementary Figure 5. Base editing efficiencies and indel frequencies for SaABE7.10, SaABE8e, SaABE8e-dimer, SaKKH-ABE7.10, SaKKH-ABE8e, and SaKKH-ABE8e-dimer in HEK293T cells at five genomic sites with NNGRRT PAMs in HEK293T cells. a, Base editing with SaABE7.10, SaABE8e, SaABE8e-dimer, SaKKH-ABE7.10,

SaKKH-ABE8e, and SaKKH-ABE8e-dimer in HEK293T cells at five genomic sites with NNGRRT PAMs in HEK293T cells. **b**, Percent of all sequencing reads containing an indel following modification by SaABE7.10, SaABE8e, SaABE8e-dimer, SaKKH-ABE7.10, SaKKH-ABE8e, and SaKKH-ABE8e-dimer at five genomic sites in HEK293T cells. All untreated controls show no editing or indel formation at all sites. For all plots, dots represent individual biological replicates and bars represent mean±s.d. of three independent biological replicates.

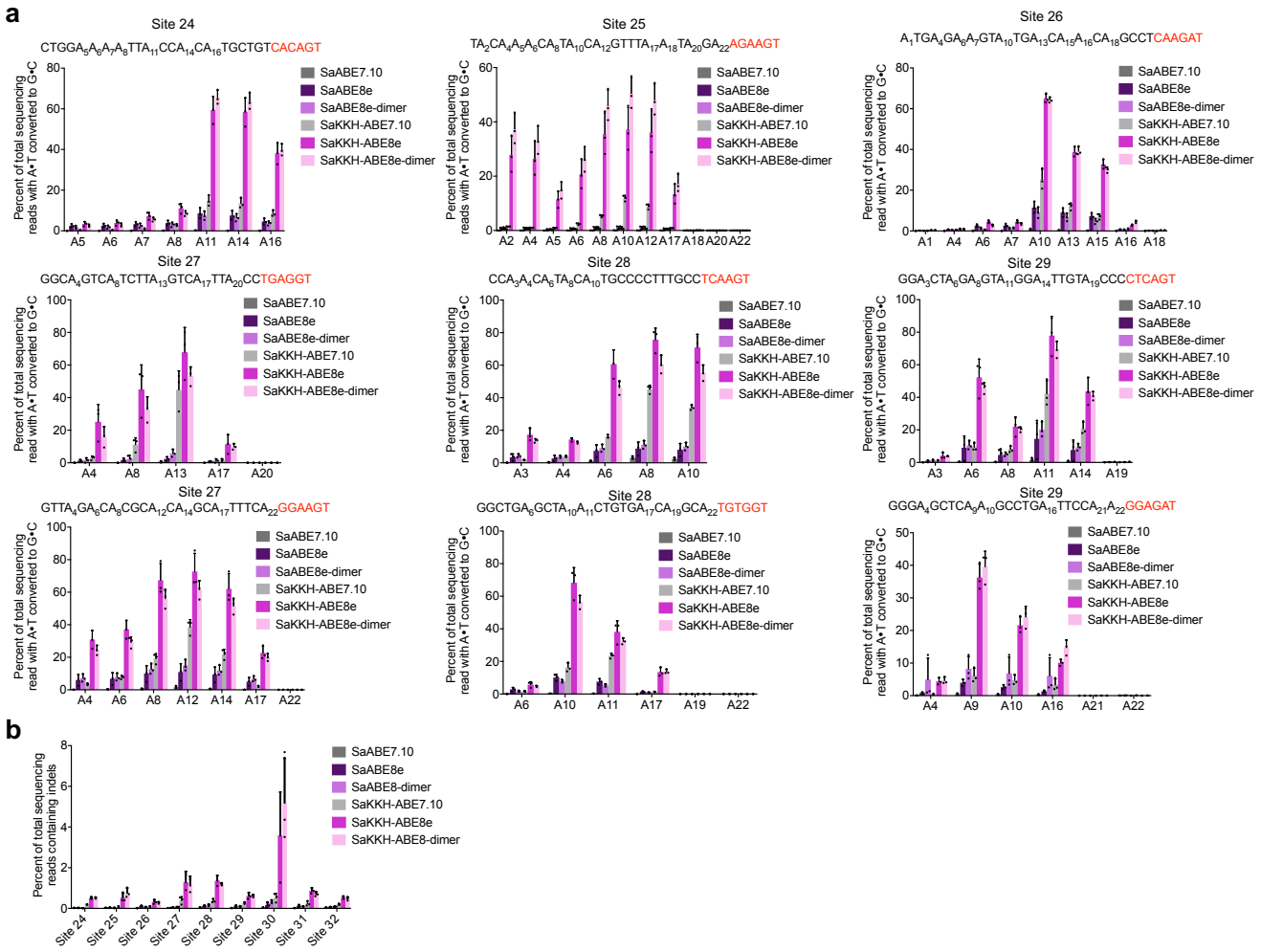


Supplementary Figure 6. Base editing efficiencies and indel frequencies for LbABE7.10, LbABE8e, Lb-ABE8e-dimer, enAsABE7.10, enAsABE8e, and enAsABE8e-dimer in HEK293T cells at three genomic sites in HEK293T cells. a, Base editing with LbABE7.10, LbABE8e, LbABE8e-dimer, enAsABE7.10, enAsABE8e, and enAsABE8e-dimer in HEK293T cells at three genomic sites in HEK293T cells. **b**, Percent of all sequencing reads containing an indel following modification by LbABE7.10, LbABE8e, LbABE8e-dimer, enAsABE7.10, enAsABE8e, and enAsABE8e-dimer at three genomic sites in HEK293T cells.

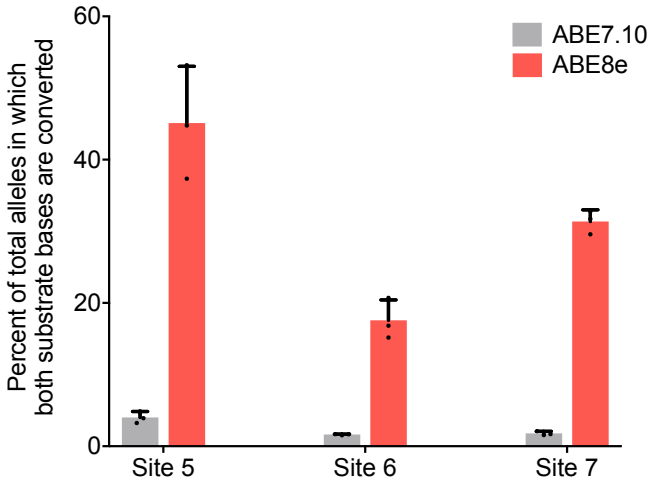
All untreated controls show no editing or indel formation at all sites. For all plots, dots represent individual biological replicates and bars represent mean±s.d. of three independent biological replicates.



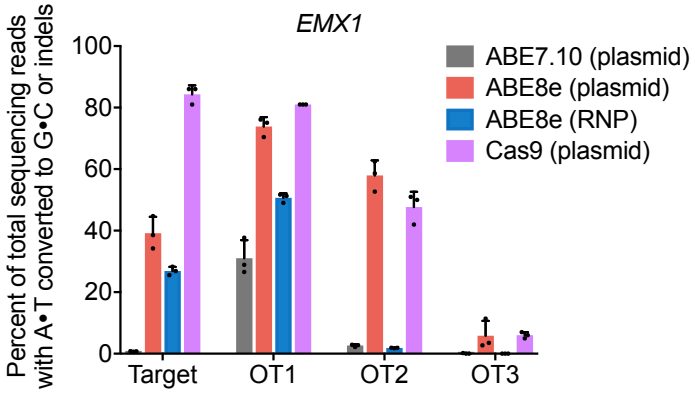
Supplementary Figure 7. Base editing efficiencies and indel frequencies for NG-ABE7.10, NG-ABE8e, and NG-ABE8e-dimer in HEK293T cells at five genomic sites in HEK293T cells. a, Base editing with NG-ABE7.10, NG-ABE8e, and NG-ABE8e-dimer in HEK293T cells at five genomic sites in HEK293T cells. **b**, Percent of all sequencing reads containing an indel following modification by NG-ABE7.10, NG-ABE8e, and NG-ABE8e-dimer at five genomic sites in HEK293T cells. All untreated controls show no editing or indel formation at all sites. For all plots, dots represent individual biological replicates and bars represent mean±s.d. of three independent biological replicates.



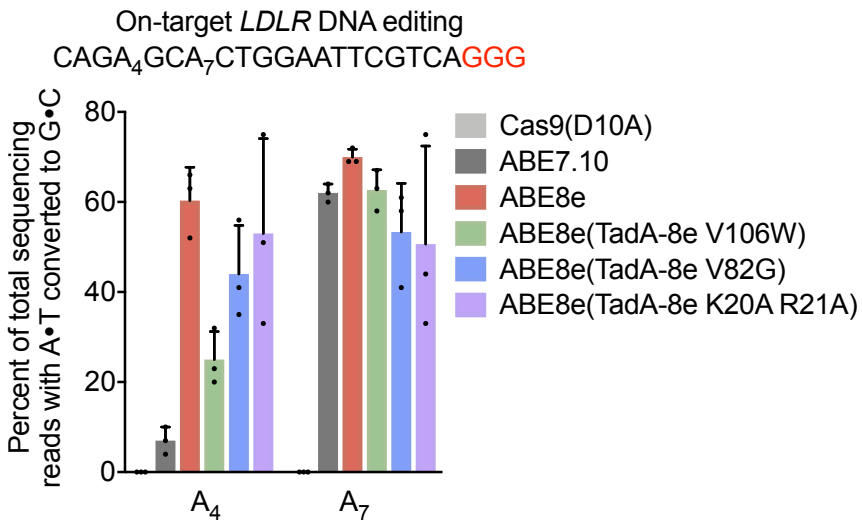
Supplementary Figure 8. Base editing efficiencies and indel frequencies for SaABE7.10, SaABE8e, SaABE8e-dimer, SaKKH-ABE7.10, SaKKH-ABE8e, and SaKKH-ABE8e-dimer in HEK293T cells at nine genomic sites with NNHRRT PAMs in HEK293T cells. a, Base editing with SaABE7.10, SaABE8e, SaABE8e-dimer, SaKKH-ABE7.10, SaKKH-ABE8e, and SaKKH-ABE8e-dimer in HEK293T cells at nine genomic sites with NNHRRT PAMs in HEK293T cells. **b**, Percent of all sequencing reads containing an indel following modification by SaABE7.10, SaABE8e, SaABE8e-dimer, SaKKH-ABE7.10, SaKKH-ABE8e, and SaKKH-ABE8e-dimer at nine genomic sites in HEK293T cells. All untreated controls show no editing or indel formation at all sites. For all plots, dots represent individual biological replicates and bars represent mean±s.d. of three independent biological replicates.



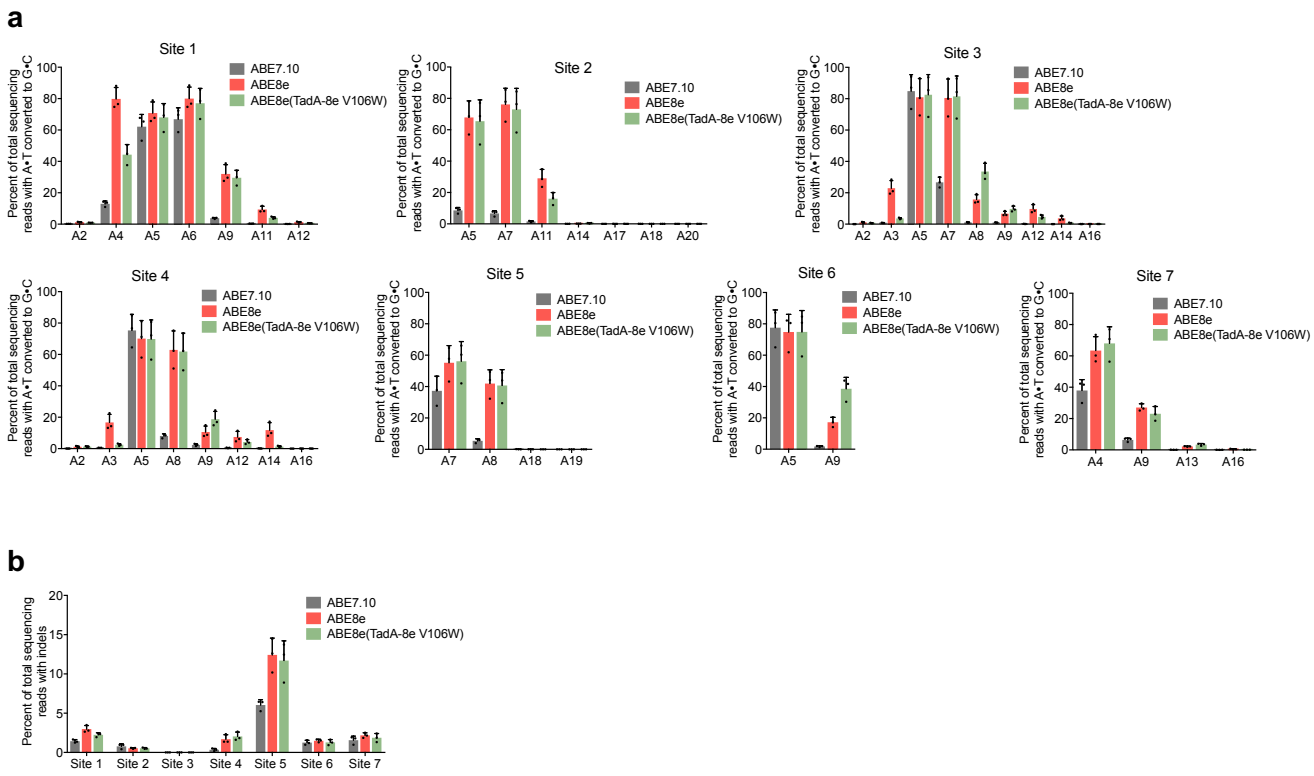
Supplementary Figure 9. Processivity analysis for ABE7.10 and ABE8e. Percent of all sequenced alleles in which the two target As in the editing window are both converted to Gs, following transfection with plasmids expressing ABE7.10 or ABE8e at three genomic sites in HEK293T cells. Analyses were done using the Python script provided in **Supplementary Note 1**. Dots represent individual values and bars represent mean±s.d. of three independent biological replicates.



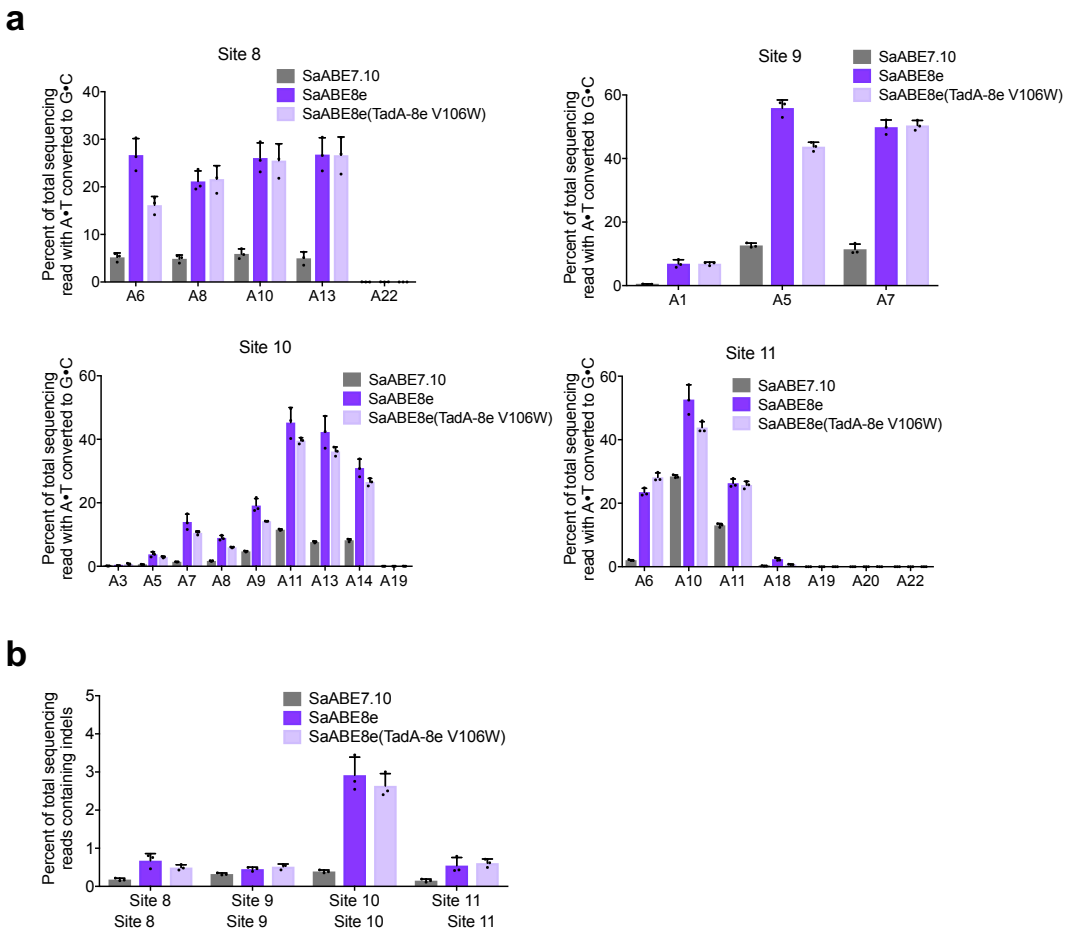
Supplementary Figure 10. Comparison of ABE8e and Cas9 nuclease at the top three off-target sites for *EMX1* as identified by GUIDE-Seq[44]. For all plots, dots represent individual biological replicates and bars represent mean±s.d. of three independent biological replicates.



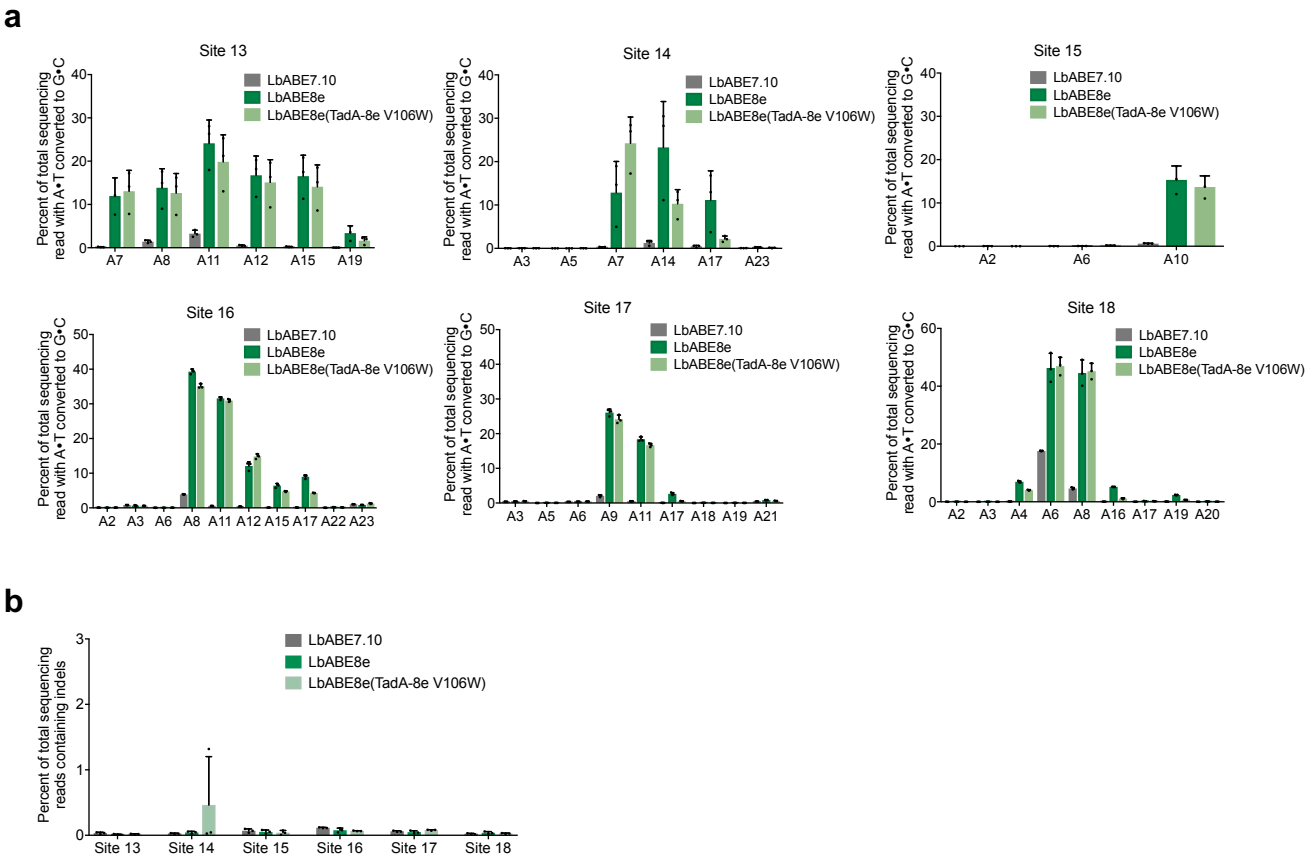
Supplementary Figure 11. On-target DNA base editing efficiencies comparing ABE8e mutants with minimized transcriptome-wide RNA off-target editing. Base editing with Cas9(D10A), ABE7.10, ABE8e, ABE8e(TadA-8e V106W), ABE8e(TadA-8e V82G), and ABE8e(TadA-8e K20A R21A) at the *LDLR* locus in HEK293T cells[47]. Editing was analyzed from RNAseq analysis. All base editors are optimized with the architecture, codon usage, and nuclear localization signals of ABEmax[16]. For all plots, dots represent individual biological replicates and bars represent mean±s.d. of three independent biological replicates.



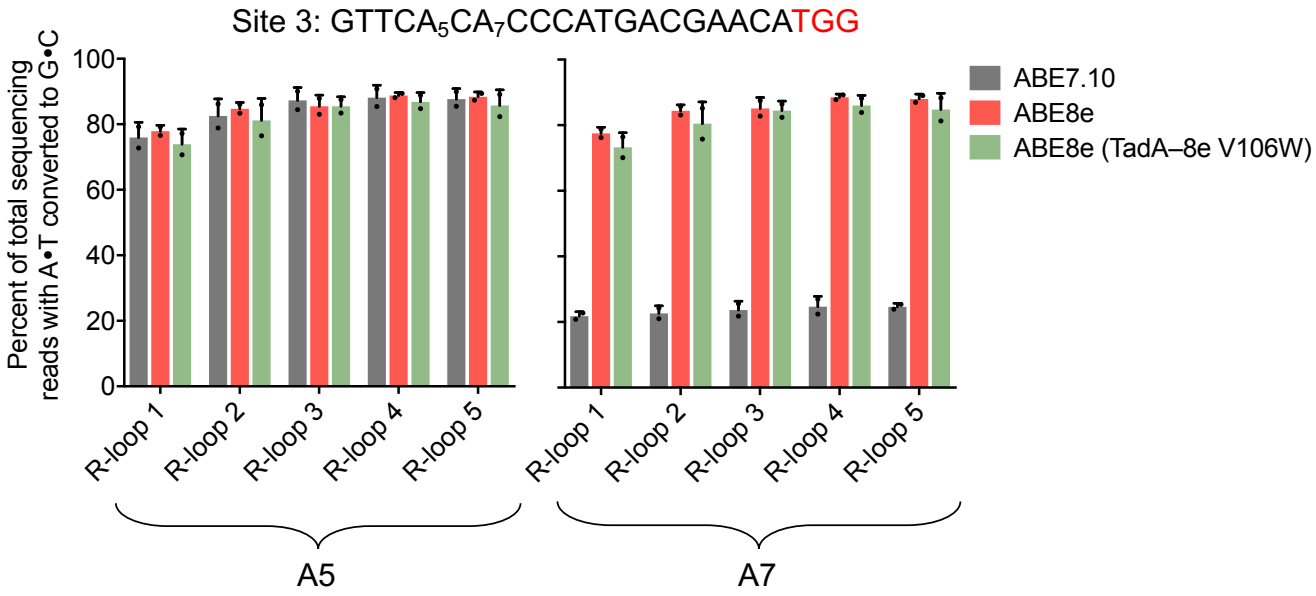
Supplementary Figure 12. On-target base editing efficiencies and indel frequencies for ABE7.10, ABE8e, and ABE8e(TadA-8e V106W) in HEK293T cells at seven genomic sites in HEK293T cells. a, Base editing with ABE7.10, ABE8e, and ABE8e(TadA-8e V106W) in HEK293T cells at seven genomic sites in HEK293T cells. **b**, Percent of all sequencing reads containing an indel following modification by ABE7.10, ABE8e, and ABE8e(TadA-8e V106W) at seven genomic sites in HEK293T cells. All untreated controls show no editing or indel formation at all sites. For all plots, dots represent individual biological replicates and bars represent mean±s.d. of three independent biological replicates.



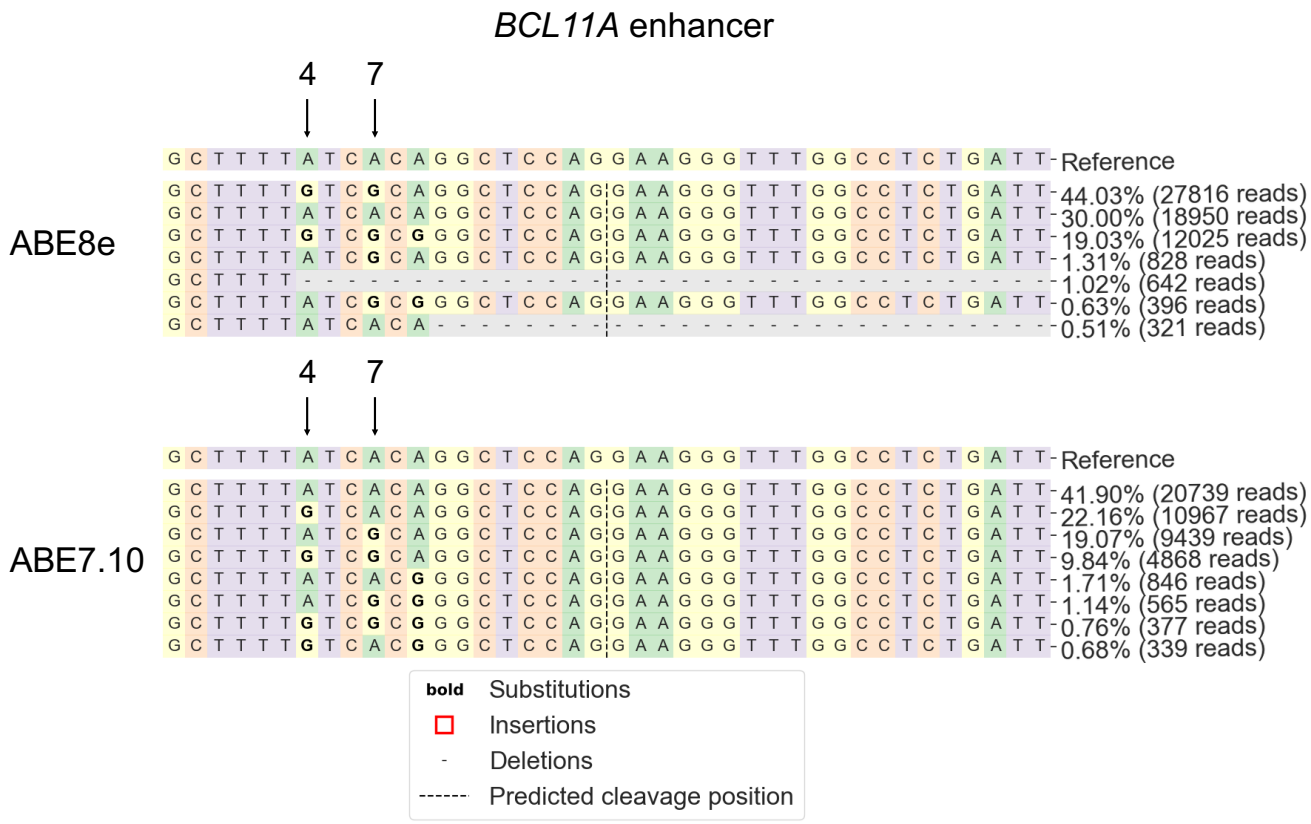
Supplementary Figure 13. On-target base editing efficiencies and indel frequencies for SaABE7.10, SaABE8e, and SaABE8e(TadA-8e V106W) in HEK293T cells at seven genomic sites in HEK293T cells. a, Base editing with SaABE7.10, SaABE8e, and SaABE8e(TadA-8e V106W) in HEK293T cells at seven genomic sites in HEK293T cells. **b**, Percent of all sequencing reads containing an indel following modification by SaABE7.10, SaABE8e, and SaABE8e(TadA-8e V106W) at seven genomic sites in HEK293T cells. All untreated controls show no editing or indel formation at all sites. For all plots, dots represent individual biological replicates and bars represent mean±s.d. of three independent biological replicates.



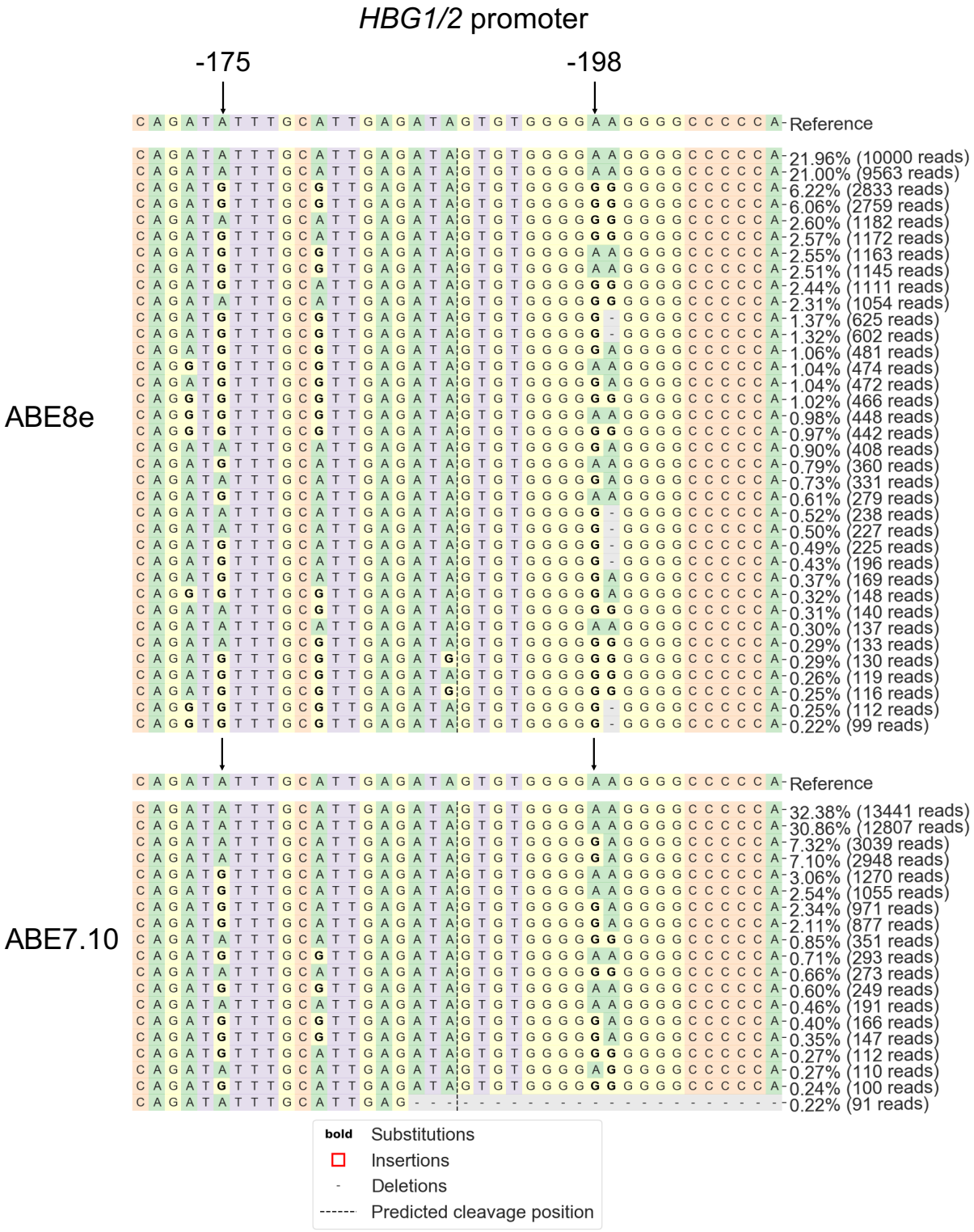
Supplementary Figure 14. On-target base editing efficiencies and indel frequencies for LbABE7.10, LbABE8e, and LbABE8e(TadA-8e V106W) in HEK293T cells at seven genomic sites in HEK293T cells. a, Base editing with LbABE7.10, LbABE8e, and LbABE8e(TadA-8e V106W) in HEK293T cells at seven genomic sites in HEK293T cells. **b**, Percent of all sequencing reads containing an indel following modification by LbABE7.10, LbABE8e, and LbABE8e(TadA-8e V106W) at seven genomic sites in HEK293T cells. All untreated controls show no editing or indel formation at all sites. For all plots, dots represent individual biological replicates and bars represent mean±s.d. of three independent biological replicates.



Supplementary Figure 15. On-target base editing efficiencies and indel frequencies for ABE7.10, ABE8e, and ABE8e(TadA-8e V106W) in HEK293T cells at site 3 in HEK293T cells for the orthogonal R-loop assay. a, Base editing with ABE7.10, ABE8e, and ABE8e (TadA-8e V106W) in HEK293T cells at site 3. For all plots, dots represent individual biological replicates and bars represent mean±s.d. of three independent biological replicates.



Supplementary Figure 16. Allele compositions following treatment with ABE7.10 or ABE8e at a GATA1 binding site of the *BCL11A* enhancer. Due to increased processivity and overall efficiency, we observed an increased percentage of alleles simultaneously containing A•T-to-G•C editing at positions four and seven after treatment with ABE8e compared to ABE7.10. Multiple edits are predicted to more effectively disrupt the GATA1 binding site compared to single edits[49]. The data shown is representative of three biological replicates.



Supplementary Figure 17. Allele compositions following treatment with ABE7.10 or ABE8e at the *HBG1* and *HBG2* promoters. We observed an increased percentage of alleles simultaneously containing A•T-to-G•C editing at positions -195 and -175 after treatment with ABE8e compared to ABE7.10. Multiple edits are predicted to more effectively lead to the expression of fetal hemoglobin[49]. The data shown is representative of three biological replicates.

Circuits	Plasmid 1 (AP)	Plasmid 2 (P2)	Plasmid 3 (P3)
P1+P2+P3a	**pBT120a17** proT7.sd8.gIII.luxAB, proLac.sgRNA, SC101-E93K, Cb^R	**pBT128c** proC.SD8.intC.dCas9, p15a, Km^R	**pMR-A-3** proD.SD8.T7-RNAP(R57* Q58*), ColE1, Sp^R
P1+P2+P3b	**pBT120a17** proT7.sd8.gIII.luxAB, proLac.sgRNA, SC101-E93K, Cb^R	**pBT128c** proC.SD8.intC.dCas9, p15a, Km^R	**pMR-A-4** proC.SD8.T7-RNAP(R57* Q58*), ColE1, Sp^R
P1+P2+P3c	**pBT120a17** proT7.sd8.gIII.luxAB, proLac.sgRNA, SC101-E93K, Cb^R	**pBT128c** proC.SD8.intC.dCas9, p15a, Km^R	**pMR-A-5** proB.SD8.T7-RNAP(R57* Q58*), ColE1, Sp^R
P1+P2+P3d	**pBT120a17** proT7.sd8.gIII.luxAB, proLac.sgRNA, SC101-E93K, Cb^R	**pBT128c** proC.SD8.intC.dCas9, p15a, Km^R	**pMR-A-6** proA.SD8.T7-RNAP(R57* Q58*), ColE1, Sp^R
P1+P2+P3e	**pBT120a17** proT7.sd8.gIII.luxAB, proLac.sgRNA, SC101-E93K, Cb^R	**pBT44c** proC.SD8.intC.dCas9.ugi, p15a, Km^R	**pBT138b-proD** proD.R4.T7-RNAP(R57* Q58*), ColE1, Sp^R
P1+P2+P3f	**pBT120a17** proT7.sd8.gIII.luxAB, proLac.sgRNA, SC101-E93K, Cb^R	**pBT44c** proC.SD8.intC.dCas9.ugi, p15a, Km^R	**pBT138b-proB** proC.R4.T7-RNAP(R57* Q58*), ColE1,Sp^R
P1+P2+P3g	**pBT120a17** proT7.sd8.gIII.luxAB, proLac.sgRNA, SC101-E93K, Cb^R	**pBT44c** proC.SD8.intC.dCas9.ugi, p15a, Km^R	**pBT138b-proA** proB.R4.T7-RNAP(R57* Q58*), ColE1, Sp^R
P1+P2+P3h	**pBT120a17** proT7.sd8.gIII.luxAB, proLac.sgRNA, SC101-E93K, Cb^R	**pBT44c** proC.SD8.intC.dCas9.ugi, p15a, Km^R	**pBT138b-pro1** proA.R4.T7-RNAP(R57* Q58*), ColE1, Sp^R

Supplementary Table 1. Gene circuit components. An asterisks (*) indicates a stop codon.

Passage	PANCE 1, Pool 1 (P1+P2+P3a ; input phage: TadA(7.10))	PANCE 1, Pool 2 (P1+P2+P3b; input phage: TadA(7.10))	PANCE 2, Pool 1 (P1+P2+P3e; input phage: PANCE 1, Pool 1)	PANCE 2, Pool 2 (P1+P2+P33; input phage: PANCE 1, Pool 2)	PANCE 2, Pool 3 (P1+P2+P3g; input phage: PANCE 1, Pool 1)	PANCE 2, Pool 4 (P1+P2+P3g; input phage: PANCE 1, Pool 2)
1	1:100 + drift	1:100 + drift	1:50	1:50	1:50	1:50
2	1:50	1:50	1:50	1:50	1:50	1:50
3	1:50	1:50	1:50	1:50	1:50	1:50
4	1:100	1:100	1:50	1:50	1:50	1:50
5	1:100 + drift	1:100 + drift	1:50	1:50	1:50	1:50
6	1:100	1:100	1:50	1:50	1:50	1:50
7	1:100 + drift	1:100 + drift	1:50	1:50	1:50	1:50
8	1:100	1:100	1:50	1:50	1:50	1:50
9	1:100	1:100	1:50	1:50	1:50	1:50
10	1:100	1:100	1:50	1:50	1:50	1:50
11	1:200	1:200				
12	1:200	1:200				
13	1:500	1:500				
14	1:500	1:500				
15	1:1000	1:1000				

Supplementary Table 2. Selection schedule for PANCE. For the first phase of PANCE (PANCE 1), two selections were performed in parallel using host cells harboring P1, P2, and either P3a or P3b. For the second phase of PANCE (PANCE 2), four selections were performed in parallel, two in host cells harboring P1, P2, and P3e, and two in host cells harboring P1, P2, and P3g. For details on plasmids, see **Supplementary Table 1**. Following the final passage of PANCE 2 (25 total passages), all phage were pooled and used as the starting point for PACE.

Supplementary Table 3. Protospacer sequences for mammalian genomic sites.

Site #	Protospacer	PAM
1	GACAAACCAGAAGCCGCTCC	TGG
2	GTTCACACCCATGACGAACA	TGG
3	GAACACAAAGCATAGACTGC	GGG
4	GAAGACCAAGGATAGACTGC	TGG
5	GTGGGGAAGGGGCCCCCAAG	AGG
6	GGTGAGTGAGTGTGTGCGTG	TGG
7	GTCATCTTAGTCATTACCTG	AGG
8	GTGGTAGACAGCATGTGTCCTA	AAGGGT
9	ATTTACAGCCTGGCCTTTGGGG	TCGGGT
10	GGAGAGAAAGAGAAGTTGATTG	ATGGGT
11	GTGTCAGGTAATGTGCTAAACA	GAGAGT
12	ATGCATTAACTGAAAATGGTCA	AGGAGT
13	TGCTGCAAGTAAGCATGCATTTG	TTTC
14	CTAGACAGGGGCTAGTATGTGCA	TTTC
15	CAGCTATTCAGGCTGGCCCGCCC	TTTG
16	GAAGCACATCAAGGACATTCTAA	TTTA
17	GGATAAGCACAGTTTTAAATAGT	TTTG
18	GTTTAAACACACCGGGTTAATAA	TTTG
19	GGTTAACAAGGCCAAACTCC	AGAT
20	GGGTCCAGTTCCGGGATTAG	CGAA
21	GGTACTGGCCAAAGATTCAA	AGCC
22	GATGACCCGTATTATCTGGC	AGTT
23	GGAACAAGGTACTCTTTGAG	TGTT
24	CTGGAAAATTACCACATGCTGT	CACAGT
25	TACAAACATACAGTTTAATAGA	AGAAGT
26	ATGAGAAGTATGACAACAGCCT	CAAGAT
27	GGCAGTCATCTTAGTCATTACC	TGAGGT
28	CCAACATACATGCCCCTTTGCC	TCAAGT
29	GGACTAGAGTAGGATTGTACCC	CTCAGT
30	GTTAGACACGCACAGCATTTCA	GGAAGT
31	GGCTGAGCTAACTGTGACAGCA	TGTGGT
32	GGGAGCTCAAGCCTGATTCCAA	GGAGAT
Site 5, OT1	GGTGGGATGGGGTCCCCAAG	TGG
Site 5, OT2	GGTAGGGAGAGGCCCCCAGA	GGG
Site 5, OT3	GGTGGGGAGCGGCCCCCCAG	TGG
EMX1	GAGTCCGAGCAGAAGAAGAA	GGG
EMX1, OT1	GAGTTAGAGCAGAAGAAGAA	AGG
EMX1, OT2	GAGTCTAAGCAGAAGAAGAA	GAG
EMX1, OT3	GAGGCCGAGCAGAAGAAAGA	CGG
VEGFA3, OT4	GAGTGAGTGAGTGTGTGTGTG	GGG
VEGFA3, OT6	GTGTGAGTAAGTGTGTGTGTG	TGG
VEGFA3, 12	GGTGAGTGTGTGTGTGCATG	TGG
R-loop 1	GTGGTAGACAGCATGTGTCCTA	AAGGGT
R-loop 2	ATTTACAGCCTGGCCTTTGGGG	TCGGGT
R-loop 3	GTGTCAGGTAATGTGCTAAACA	GAGAGT
R-loop 4	GGTGGAGGAGGGTGCATGGGGT	CAGAAT
R-loop 5	TCTGCTTCTCCAGCCCTGGC	CTGGGT
-198 target	GTGGGGAAGGGGCCCCCAAG	AGG
-175 target	ATATTTGCATTGAGATAGTG	TGG
***BCL11A* enhancer**	TTTATCACAGGCTCCAGGAA	GGG

Supplementary Table 4. Primers used for mammalian cell genomic DNA amplification.

Site #	HTS Forward Primer	HTS Reverse Primer
1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNAGCCCTCTTTTTATTGGAACTGTG	TGGAGTTCAGACGTGTGCTCTTCCGATCTC CGACTGGTCCACTTACCTA
2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNCTGACTCAGCCCTGCAAAGG	TGGAGTTCAGACGTGTGCTCTTCCGATCTC AAGTCAGGGGAGCGTGTC
3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNCCAGCCCCATCTGTCAAACT	TGGAGTTCAGACGTGTGCTCTTCCGATCTT GAATGGATTCCTTGGAAACAATGA
4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNTTGCTTATTGCTGAGGGGCA	TGGAGTTCAGACGTGTGCTCTTCCGATCTA CCTCTCTCCTCCAGCTGAG
5	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGTAGAAATGGGGTCTTGCTTTG	TGGAGTTCAGACGTGTGCTCTTCCGATCTT TGAGTCTATCGAGTGTGTGCAT
6 (VEGFA3)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGTGCAGACGGCAGTCACTAGG	TGGAGTTCAGACGTGTGCTCTTCCGATCTC TATTGGAATCCTGGAGTGACCC
7	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNAACGGAACTCAACCATTAAGCA	TGGAGTTCAGACGTGTGCTCTTCCGATCTC CAACATACAGAAGTCAGGAATGC
8	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNTGCAGTCTCCTGCTTCTCTG	TGGAGTTCAGACGTGTGCTCTTCCGATTGG TGGAGTGCTCTGTGTTTG
9	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGGACATTTCCACCGCAAAATG	TGGAGTTCAGACGTGTGCTCTTCCGATGCT ACAGAAAGGTCAGCAGC
10	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNTCACTTCAGCCCAGGAGTAT	TGGAGTTCAGACGTGTGCTCTTCCGATCTT GTGTATGGTGAGAGGTAGGGA
11	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNCTGCACCTAGCCTCCATGTC	TGGAGTTCAGACGTGTGCTCTTCCGATCTG CTGTGGCATCCAGAGACAT
12	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGATTCAGTGCTCAAAATAAGTTGTC	TGGAGTTCAGACGTGTGCTCTTCCGATCTA TTGACATTCCCTAACTTCAACGTA
13	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNATGTGGGCTGCCTAGAAAGG	TGGAGTTCAGACGTGTGCTCTTCCGATCTC CCAGCCAAACTTGTCAACC
14	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNATGTGGGCTGCCTAGAAAGG	TGGAGTTCAGACGTGTGCTCTTCCGATCTC CCAGCCAAACTTGTCAACC
15	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNCCAGCCCCATCTGTCAAACT	TGGAGTTCAGACGTGTGCTCTTCCGATCTT GAATGGATTCCTTGGAAACAATGA
16	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNAAAAGGGGCAAGCTTCAGAT	TGGAGTTCAGACGTGTGCTCTTCCGATCTA GTGAGGAGAAGGCAGGAGG
17	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNTGTTCTGCCCTCACAGAGGT	TGGAGTTCAGACGTGTGCTCTTCCGATCCC AAAGGACATACGGGGAG
18	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGTGCGTGCTTCTTACATGCC	TGGAGTTCAGACGTGTGCTCTTCCGATCCA AGTATGCCTTAAGCAGAACAA
19	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGCTTTAAACATTTGTCTGTGCG	TGGAGTTCAGACGTGTGCTCTTCCGATCTG TTTTCTGTCCCTCCCTCAGTA
20	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGCTTTAAACATTTGTCTGTGCG	TGGAGTTCAGACGTGTGCTCTTCCGATCTG TTTTCTGTCCCTCCCTCAGTA
21	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNCAGAGAGAGCAGGACGTCACA	TGGAGTTCAGACGTGTGCTCTTCCGATCTA GCACTACCTACGTCAGCACCT

22	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNCACTGCTGAACCAGTCAAACTC	TGGAGTTCAGACGTGTGCTCTTCCGATCTG GCATGGGGAAATATAAACTTG
23	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGACAGAGGGAGAGAAACAGAGC	TGGAGTTCAGACGTGTGCTCTTCCGATCTT TCTAGATGCCGACAAAAGGAT
24	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGTAGAAATGGGGTCTTGCTTTG	TGGAGTTCAGACGTGTGCTCTTCCGATCTT TGAGTCTATCGAGTGTGTGCAT
25	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNTCACTTCAGCCCAGGAGTAT	TGGAGTTCAGACGTGTGCTCTTCCGATCTT GTGTATGGTGAGAGGTAGGGA
26	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNCTGACTCAGCCCTGCAAAGG	TGGAGTTCAGACGTGTGCTCTTCCGATCTC AAGTCAGGGGAGCGTGTC
27	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNACGTCTCATATGCCCCTTGG	TGGAGTTCAGACGTGTGCTCTTCCGATCTA CGTAGGAATTTTGGTGGGACA
28	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNTGCAGTCTCCTGCTTCTCTG	TGGAGTTCAGACGTGTGCTCTTCCGATTGG TGGAGTGCTCTGTGTTTG
29	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNCCCTGTTCCTAAAGCCCACC	TGGAGTTCAGACGTGTGCTCTTCCGATCTA CTGGTTCTGTTTGTGGCCA
30	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNAGAGACTGATTGCGTGGAGT	TGGAGTTCAGACGTGTGCTCTTCCGATCTC ACTCCAGCCTAGGCAACAA
31	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNCCAGCCCCATCTGTCAAACT	TGGAGTTCAGACGTGTGCTCTTCCGATCTT GAATGGATTCCTTGGAAACAATGA
32	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNCGCGGGCTGAAGTAGATCAA	TGGAGTTCAGACGTGTGCTCTTCCGATCTC CTGTCTCTGCTCCTTTGTCCCC
Site 5, OT1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNATGGCTGCAAATCCAAGGGT	TGGAGTTCAGACGTGTGCTCTTCCGATCTA AATGCTTCTCGGGCTCTCC
Site 5, OT2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGAGGTTGAAACTCCTCGCCA	TGGAGTTCAGACGTGTGCTCTTCCGATCTG GAATTAAGATGCAACTGAGAGTA
Site 5, OT3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGGCTGTCCCTGGTTGTCTGG	TGGAGTTCAGACGTGTGCTCTTCCGATCTC GAGCACTGAGGCCTGGTTA
EMX1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNCAGCTCAGCCTGAGTGTTGA	TGGAGTTCAGACGTGTGCTCTTCCGATCTC TCGTGGGTTTGTGGTTGC
EMX1, OT1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGTGGGGAGATTTGCATCTGTGGAGG	TGGAGTTCAGACGTGTGCTCTTCCGATCTG CTTTTATACCATCTTGGGGTTACAG
EMX1, OT2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNCAATGTGCTTCAACCCATCACGGC	TGGAGTTCAGACGTGTGCTCTTCCGATCTC CATGAATTTGTGATGGATGCAGTCTG
EMX1, OT3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGAGAAGGAGGTGCAGGAGCTAGAC	TGGAGTTCAGACGTGTGCTCTTCCGATCTC ATCCCGACCTTCATCCCTCCTGG
VEGFA3, OT4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGCTCATTTCCTACGGCCCAG	TGGAGTTCAGACGTGTGCTCTTCCGATCTC TGCAGTGAGGAGGTGGTTC
VEGFA3, OT6	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGCCACAGGCACTAACTTCTTCA	TGGAGTTCAGACGTGTGCTCTTCCGATCTG ATGAAGCTGCCTTTCCTAAGC
VEGFA3, OT12	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGCGCTTTCCCTTTGCTAGAATC	TGGAGTTCAGACGTGTGCTCTTCCGATCTC TCAGCAATGCTTATATTACTGGC
R-loop 1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNTGCAGTCTCCTGCTTCTCTG	TGGAGTTCAGACGTGTGCTCTTCCGATTGG TGGAGTGCTCTGTGTTTG

R-loop 2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGGACATTTCCACCGCAAAATG	TGGAGTTCAGACGTGTGCTCTTCCGATGCT ACAGAAAGGTCAGCAGC
R-loop 3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNCTGCACCTAGCCTCCATGTC	TGGAGTTCAGACGTGTGCTCTTCCGATCTG CTGTGGCATCCAGAGACAT
R-loop 4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNGGAGGTGGAGAGAGGATGT	TGGAGTTCAGACGTGTGCTCTTCCGATCTT CCTGAGGTCTAGGAACCCG
R-loop 5	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNATGTGGGCTGCCTAGAAAGG	TGGAGTTCAGACGTGTGCTCTTCCGATCTC CCAGCCAAACTTGTCAACC
-198 target	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNTCTTAGACATAACACACCAGG	TGGAGTTCAGACGTGTGCTCTTCCGATCTG TCTGCCAGTCCTCTTC
-175 target	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNTCTTAGACATAACACACCAGG	TGGAGTTCAGACGTGTGCTCTTCCGATCTG TCTGCCAGTCCTCTTC
BCL11A enhancer	ACACTCTTTCCCTACACGACGCTCTTCCGATCTN NNNCCTGGCCTCACTGGATACTC	TGGAGTTCAGACGTGTGCTCTTCCGATCTC TGACAAAAGAAGTCCTGGTATC

Supplementary Table 5. Amplicons for high-throughput sequencing analysis.

Site #	Amplicon
1	CCGACTGGTCCACTTACCTATCACAATCACAACTGCAATATTACATAAATAGTCTGTTAGAGAAGG ACAACTGAATCTGCAAATCAGTATCTTCTCAATGAGAGTACAAGGAGGCAATGGCTACATACGAT GGACAAACCAGAAGCCGCTCCTGGGCTATGTTTACTATTTACTTTTATGGTATAAAAATGTTCTCA GTCAAGCAGTTTCAACAATAGATACCACAGTTCCAATAAAAAGAGGGCT
2	GCAAGTCAGGGGAGCGTGTCCATAGGGTGCCAGGCTGGCCGCTTCTCCACGGGCCCTGCCTTC CTCACCTGATGATCTTGAGGCTGTTGTCATACTTCTCATGGTTCACACCCATGACGAACATGGGG GCATCAGCAGAGGGGGCAGAGATGATGACCCTTTTGGCTCCCCCCTGCAAATGAGCCTACAGCA GAGAAGCAGACAGTTATGAACCCGGGTCCTGCCTTTGCAGGGCTGAGTCAG
3	TAACAAGACCTGGCTGAGCTAACTGTGACAGCATGTGGTAATTTTCCAGCCCGCTGGCCCTGTA AAGGAAACTGGAACACAAAGCATAGACTGCGGGGCGGGCCAGCCTGAATAGCTGCAAACAAGT GCAGAATATCTGATGATGTCATACGCAC
4	AAGTGAAGACCAAGGATAGACTGCTGGGCTTGACAGCATGGAGGGTGCTGGGGGCCTAGGCAA GTGCAATATTCATGTGGTGTCATTGGGCCAAAGGTGTCATTGGAGAGTACTCAAAAAACACTAGG AGAGGAGCTGAGCTTTGTCAGGCGTGTAGGAAGATGGAGGAGAAGAAAAGGGGGAGTGGTGGT GGGCT
5	GTGGAGTTTAGCCAGGGACCGTTTCAGACAGATATTTGCATTGAGATAGTGTGGGGAAGGGGCC CCCAAGAGGATACTGCTGCTTAATTTTTTTTATAGCCTTTGCCTTGTTCCGATTCAGTCATTCCAA TTTTTCTCTAATTTATTCTTCCCTTTAGCTAGTTTTCTTCTCCCACCA
6	GTGCAGACGGCAGTCACTAGGGGGCGCTCGGCCACCACAGGGAAGCTGGGTGAATGGAGCGA GCAGCGTCTTCGAGAGTGAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGTGTGCGTGTGGGG TTGAGGGCGTTGGAGCGGGGAGAAGGCCAGGGGTCACTCCAGGATTCCAATAG
7	CCAACATACAGAAGTCAGGAATGCTTGAATATAAATTTATTATTACTCTATGTTCTATTTAAGTTTT CATGTTCTAAAAATGTATCCCAGTTTACACGTCTCATATGCCCCTTGGCAGTCATCTTAGTCATTA CCTGAGGTGTTCGTTGTAACTCATATAAACTGAGTTCCCATGTTTTGCTTAATGGTTGAGTTCCGT T
8	TGGTGGAGTGCTCTGTGTTTGTCTTTATAAACCCAGATGAGAGGATGAAGGCAACAAGCTTCTGT ACCAACATACATGCCCCTTTGCCTCAAGTCTGGTTATTTTAGGGGGATGCTAGGTTGCTTTGGGT CTACCTTACTGAGAAAATGGCCCCAGGTCATTGTCATGTCCAGTTGTGGTAGACAGCATGTGTCC TAAAGGGTATATTCACATGCATGTGCAAAAATACAGGGGTCCTTCTAACCCTATCACAGAGAAGC AGGAGACTGCA

9	GCTACAGAAAGGTCAGCAGCTATATTTAACCTCAGACCAGGGTGCGGTGGGAGATCTGGTTTCC GGAAGACGGAATGGGGAGAAGGGCAGGTTCCCCGAGGCGCCCAGACACCCAATCCTCCCGGT GACATTTACAGCCTGGCCTTTGGGGTCGGGTCAACGCTAGGCTGGCAGGGGAAGGGCGGGGC CGTGAGGTGAGCCGGCGCTGCAGGAAGGGGCCACCACCAGAGGGGCCATTTTGCGGTGGAAA TGTCC
10	AACAAATGTGAAAAAAAACAAGTATCAAATGTTCTCACTCACATGTGGAAACTAAAAAAGTGGATC TCATGAAGATAGAGAATAGACTGCTGGTTACTAGAGGTCAGGAAGAGTAGAGGGGTGGGGGAG AGAAAGAGAAGTTGATTGATGGGTACAAACATACAGTTTAATAGAAGAAGTATGACCCAGTGTTT GATAGATCAGTAGGGTGACTACAGTTTACAACCATGTA
11	AAGTGTTCAGCTGCTTTTCTTTCATTTATTCCACATATAATTACTATAATTGCTAAACATTTATTTAG TGTCAGGTAATGTGCTAAACAGAGAGTTACTGCTCAGACATGTAATAATAATAAATAACACATCAA ATAACCATACCATTTTAAGCTGTAGTATTATGAAGGGAAATCTGGAGCAAAGAGAATAGACTGTA GGGAAACCAGTTAAGAAATAGGACATGGAGGCTAGGTGCAG
12	ATTGACATTCCCTAACTTCAACGTACCGAGTAAGGAAATTCTTTGTCCTGAGATCACAGACAAATG GGATTGTATGCATTAACTGAAAATGGTCAAGGAGTCAGTACTGAAGATCAGCAATGCCAAGGGTA AACGGAATATCCCAGTATAATAGGATAAGCAATGGACAACTTATTTTGAGCACTGAATC
13	ATGTGGGCTGCCTAGAAAGGCATGGATGAGAGAAGCCTGGAGACAGGGATCCCAGGGAAACGC CCATGCAATTAGTCTATTTCTGCTGCAAGTAAGCATGCATTTGTAGGCTTGATGCTTTTTTTCTGC TTCTCCAGCCCTGGCCTGGGTCAATCCTTGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAA GGAAGCCCTGCTTCCTCCAGAGGGCGTCGCAGGACAGCTTTTCCTAGACAGGGGCTAGTATGT GCAGCTCCTGCACCGGGATACTGGTTGACAAGTTTGGCTGGG
14	ATGTGGGCTGCCTAGAAAGGCATGGATGAGAGAAGCCTGGAGACAGGGATCCCAGGGAAACGC CCATGCAATTAGTCTATTTCTGCTGCAAGTAAGCATGCATTTGTAGGCTTGATGCTTTTTTTCTGC TTCTCCAGCCCTGGCCTGGGTCAATCCTTGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAA GGAAGCCCTGCTTCCTCCAGAGGGCGTCGCAGGACAGCTTTTCCTAGACAGGGGCTAGTATGT GCAGCTCCTGCACCGGGATACTGGTTGACAAGTTTGGCTGGG
15	GTGCGTATGACATCATCAGATATTCTGCACTTGTTTGCAGCTATTCAGGCTGGCCCGCCCCGCA GTCTATGCTTTGTGTTCCAGTTTCCTTTACAGGGCCAGCGGGCTGGAAAATTACCACATGCTGTC ACAGTTAGCTCAGCCAGGTCTTGTTA
16	AAAAGGGGCAAGCTTCAGATTGTTTAGATTTTGTTTGTGAAATACTTCTCATTCCTCATTATTAAAA TATAAAATGAGTTTCTATTAATTTTTCTTAGTTTCTTTATAGATACTGATATAATTCATAATAATACC ATTCTTATCTTAAAACCTTGTCACACACAATGAAACTTTGCTGTTCACTGTCAGTTATAACTTACAT GAGGTGACCCATTTCCATTCAAGGGTTTTAGAAGCACATCAAGGACATTCTAAGGATGATTGACT TACACAATGATCTCTGAACATGCCTCCTGCCTTCTCCTCACT
17	TGTTCTGCCCTCACAGAGGTTATGATCCAATGAGACAGATGAGGTGGTTGTGGATTAATTAAATA TTCATAAAAAAGCAAATTAAGTGCTATAATGTTAGAATTACAGAAAAATAAAGGATGTAATTTTGGG ACTTTAATTTTTTAAATACTTATATTCACTTTTATAACGAAGAACTCTTTGTGGAAAATGGTAATTTT CTGTTACCATTTGGGATAAGCACAGTTTTAAATAGTTCTGGAATTATAGAGGCACCTCCCCGTATG TCCTTTGGG
18	GTGCGTGCTTCTTACATGCCTTATTAACCCGGTGTGTTTAAACCAAACACTGTTCATATTTTTCCA GGAGGAAAACAAAACAATAAAAAACATTATTCAGATAAAATATTATAGGTTTATTTAAAACTTAATT CTCACCTTGAGTATGCAAAATACAAACTCCACAAAATGTTCATTTTACTTTGTAGTTTACAAATATA CAAAATAGACGTTTGCTTAAATTTATATTACATATTTATTAAGGCAAGGAACTATATAGAAAAACAC ATTTGTTCTGCTTAAGGCATACTTGG
19	GCTTTAAACATTTGTCTGTGCGTCAATGCTTTAAAGGGACATACTATGTGGTGATTCATTTTCAAA GTAAGAATTTGCAGAAAATGAGAAACAAACTTTCTGAACAATACAATTTTAATGCATACTAAGGCT GGTTAACTGGTTAGAATACTAGGTTAACAAGGCCAAACTCCAGATAGGCCAACAGCTTTTCAAAC CAATATTCCTGACACTGCCAGGAATAAAGAAAACTCTTCACTACTGAGGGAGGGACAGAAAAC
20	AGCACTACCTACGTCAGCACCTGGGACCCCGCCACCGTGCGCCGGGCCTTGCAGTGGGCGCG CTACCTGCGCCACATCCATCGGCGCTTTGGTCGGCATGGCCCCATTCGCACGGCTCTGGAGCG GCGGCTGCACAACCAGTGGAGGCAAGAGGGCGGCTTTGGGCGGGGTCCAGTTCCGGGATTAG CGAACTTCCAGGCCCTCGGTCACTGTGACGTCCTGCTCTCTCTG
21	CACTGCTGAACCAGTCAAACTCCAACTCTAAGCATGGAGTTTCCTAAGAGATGGAAGAAAAGCTA TTATATATACATAAGAGAGAAGGTTTGACTGCCATAAAAAATATCTAATTTATGACAATAAAAACCT TACTTTATTTGGATTTGATCCAGTAACACCAATAGGGTTCAGCAAATCTTCTAATCCATGAGGTAC TGGCCAAAGATTCAAAGCCATTTTTCCAGATACTAGAGTGTCTGTGTAATCAAACAAGTTTATATT TCCCCATGCC

22	TTCTAGATGCCGACAAAAGGATCAAGGTGGCGAAGCCCGTGGTGGAGATGGATGGTGATGAGAT GACCCGTATTATCTGGCAGTTCATCAAGGAGAAGGTAGTGCCCCCTCCTGAAGTGGGTGGCTCT CCAGGTGGGCTGGCCAGGGATTGTTCTGTCCCACAGGGTCTTCTGGACTGCAGGTCCCTAGGA CCCCCCCGTTGTCCTGGTAGGCAGCAGCAGCTCTGTTTCTCTCCCTCTGTC
23	GTAGAAATGGGGTCTTGCTTTGTTGCCCAGGCTGGTCTAAAAAAATATACTACTTTTATGGATCAT ACTGCTAAACACTAATATAACCTTTGGAAATATAAATCTATATACTTCCTTACCTGGGATTGGAACA AGGTACTCTTTGAGTGTTCACATTGTCACATAAGGGTTCTCCTCCATGGTAGATACCTGTTCGAA CATAGATCTAAAAGAAAAAAGTAGGTATATACTAATGTATACACTCAACATACACATATGCACACA CTCGATAGACTCAA
24	GTGCGTATGACATCATCAGATATTCTGCACTTGTTTGCAGCTATTCAGGCTGGCCCGCCCCGCA GTCTATGCTTTGTGTTCCAGTTTCCTTTACAGGGCCAGCGGGCTGGAAAATTACCACATGCTGTC ACAGTTAGCTCAGCCAGGTCTTGTTA
25	AACAAATGTGAAAAAAAACAAGTATCAAATGTTCTCACTCACATGTGGAAACTAAAAAAGTGGATC TCATGAAGATAGAGAATAGACTGCTGGTTACTAGAGGTCAGGAAGAGTAGAGGGGTGGGGGAG AGAAAGAGAAGTTGATTGATGGGTACAAACATACAGTTTAATAGAAGAAGTATGACCCAGTGTTT GATAGATCAGTAGGGTGACTACAGTTTACAACCATGTA
26	CTGACTCAGCCCTGCAAAGGCAGGACCCGGGTTCATAACTGTCTGCTTCTCTGCTGTAGGCTCA TTTGCAGGGGGGAGCCAAAAGGGTCATCATCTCTGCCCCCTCTGCTGATGCCCCCATGTTCGTC ATGGGTGTGAACCATGAGAAGTATGACAACAGCCTCAAGATCATCAGGTGAGGAAGGCAGGGCC CGTGGAGAAGCGGCCAGCCTGGCACCCTATGGACACGCTCCCCTGACTTGC
27	ACGTCTCATATGCCCCTTGGCAGTCATCTTAGTCATTACCTGAGGTGTTCGTTGTAACTCATATAA ACTGAGTTCCCATGTTTTGCTTAATGGTTGAGTTCCGTTTGTCTGCACAGCCTGAGACATTGCTG GAAATAAAGAAGAGAGAAAAACAATTTTAGTATTTGGAAGGGAAGTGCTATGGTCTGAATGTATG TGTCCCACCA
28	TGGTGGAGTGCTCTGTGTTTGTCTTTATAAACCCAGATGAGAGGATGAAGGCAACAAGCTTCTGT ACCAACATACATGCCCCTTTGCCTCAAGTCTGGTTATTTTAGGGGGATGCTAGGTTGCTTTGGGT CTACCTTACTGAGAAAATGGCCCCAGGTCATTGTCATGTCCAGTTGTGGTAGACAGCATGTGTCC TAAAGGGTATATTCACATGCATGTGCAAAAATACAGGGGTCCTTCTAACCCTATCACAGAGAAGC AGGAGACTGCA
29	AGAATTGCTGTCAGTACTAACCAGCTATTTCATAGCTGTTGCATGAGGAAAGGGACTAGAGTAGG ATTGTACCCCTCAGTCTATGCTTTGTTTACTCTGAGTGTACAAAAGAT
30	TCATGGAGTATGAGGCATAGACTGCAGGAGACATCAAACCATGACTTGCAGATGAAGAAGCATTT TAAAAGTTAGACACGCACAGCATTTCAGGAAGTTATATATAAGGAGTGTTATAGAAGATAACACCT CTAAAATGTGTAGTATATATATCAGGGGTTTTTTTTGAGACAGAGTTTTGCTC
31	TGAATGGATTCCTTGGAAACAATGATAACAAGACCTGGCTGAGCTAACTGTGACAGCATGTGGTA ATTTTCCAGCCCGCTGGCCCTGTAAAGGAAACTGGAACACAAAGCATAGACTGCGGGGCGGGC CAGCCTGAATAGCTGCAAACAAGTGCAGAATATCTGATGATGTCATACGCACAGTTTGACAGATG GGGCTGG
32	AGTCCCCAGGGAGCTCAAGCCTGATTCCAAGGAGATTGCCAATATTTTAGGAGGGAGTAAACATT TTCTGAAGTTTTTTGTTGTCTAAAGCCAAAAAATGCACTTGCGCCTCCTGAGAGGACAAAGAGGA AGAGAGACGGGGAGA
Site 5, OT1	ATGGCTGCAAATCCAAGGGTGCACACTGGGGTGTCCTTGGCCTGAGTGCCAGCCTGCAGAGAG CAGGAGTGGGCCACTTGGGGACCCCATCCCACCAGGAAGCCTGAGAAAGCCCAGGGTGAGCC CGGCCCTGTGGAATTCTGGAGAGCCCGAGAAGCATTT
Site 5, OT2	GAGGTTGAAACTCCTCGCCACAGTCCACTCTCCATCTCCATCCCTACCCCATCAATAAGCCCTTC CCCCACCATCTCAAGACACTGGAATGCAGGTAGGGAGAGGCCCCCAGAGGGAGTATTTGAGAAT TTGCTCCCCTGTCCCTTGCCTCAACTACTCTCAGTTGCATCTTAATTCC
Site 5, OT3	GGCTGTCCCTGGTTGTCTGGTACCTGGCTTTGGAGCAATGAGGGCTGGTGGGGAGCGGCCCCC CAGTGGGGCCGTCTGATTAGCGCAGGGCTGCCGTGTGGACCAATCAGCTGCTCAGGAAGGGGC CGGCCCCTAACCAGGCCTCAGTGCTCG
EMX1	CAGCTCAGCCTGAGTGTTGAGGCCCCAGTGGCTGCTCTGGGGGCCTCCTGAGTTTCTCATCTGT GCCCCTCCCTCCCTGGCCCAGGTGAAGGTGTGGTTCCAGAACCGGAGGACAAAGTACAAACGG CAGAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCATCACATCAAC CGGTGGCGCATTGCCACGAAGCAGGCCAATGGGGAGGACATCGATGTCACCTCCAATGACTAG GGTGGGCAACCACAAACCCACGAG

EMX1, OT1	GTGGGGAGATTTGCATCTGTGGAGGCAATAAAGTGAAATAAATAACAGATGCAAGCAAGCTTTTC CTGACGCCCCGCTTGTCCATGTCTAGGAAAGATTAACAGAGAGTCTGACACCTTTTAAGATCTGA CAGAGAAACATTTACCATAGACTATCACCTATTTTTTCTGAGGGCTGCTACCTGTACATCTGCACA AGATTGCCTTTACTCCATGCCTTTCTTCTTCTGCTCTAACTCTGACAATCTGTCTTGCCATGCCAT AAGCCCCTATTCTTTCTGTAACCCCAAGATGGTATAAAAGC
EMX1, OT2	CAATGTGCTTCAACCCATCACGGCCTTTGCAAATAGAGCCCTTTATTCATAGTAGACAAGAGTCT AAGCAGAAGAAGAAGAGAGCCACTACCCAACCATCTACTCTTCTAATGGTGTTTTCCTACAAAGG CCAAGTCATGAGACTGCATCCTTGTGAAAGCCAACACTGATGATAATGAGGCTTACCTTGAGTAC AATGAAGTAGAGGAAGGTAGGCAGTGAAACAGTAGAAAAAAGTCCCCCCCCCAAAAAGGCAGACT GCATCCATCACAAATTCATGG
EMX1, OT3	GAGAAGGAGGTGCAGGAGCTAGACTCCGAGGGGAGGCTGCGAGCCGCAAGCGCAGGAGCCGG GTGGGAGAGAGACCCCTTCTTCTGCAAATGAGGAGGCCGAGCAGAAGAAAGACGGCGACAGAT GTTGGGGGGAGGGGACGGTTTGTGAGGGATAGGGAGAGAAAGTCTAAGTGAGAGCAGGACGA GGAGGGGAGGGTGGTTGGGGAAGGGGATGAAATAAAGTGGCTGAGGGAAGGCTGAGGTGGGG TGAAGCTCCGCCAGGAGGGATGAAGGTCGGGATG
VEGFA3, OT4	GCTCATTTCCTACGGCCCAGGGAGCGGCCACTGCAGCGGCGGCGGGGAGGGGAAGGGGTGAA GGGGAGGGGGAAGTCACCGACAACAACAAGCCGAGTCCCCCCCACACACACACTCACTCACTC ACTCACTCGCTCTCTCACTCACACACACACACACAACACGGGCAAGAACCACCTCCTCACTGCAG
VEGFA3, OT6	GCCACAGGCACTAACTTCTTCAGCCTATCTCCTATCTGCTCTTCCCTCCCACCCCCTCCACACAC ACACACTTACTCACACGCTCACAAACACATGCCCAACCTAATTCAGGGACTCACATACAGCTTAG GAAAGGCAGCTTCATC
VEGFA3, OT12	GCGCTTTCCCTTTGCTAGAATCTAATTTTATGAAACACACTCATGTGTGTGCCCACATGCACACAC ACACTCACCAAGAGACACACATATACCAAAGCAACCACTGAACCCTCTCAAGAAGAAAAATTGAA GGCTGGAATTTTCTCTATATAATGAAGTTTTTTTAAATGTTGCCAGTAATATAAGCATTGCTGAG
R-loop 1	TGGTGGAGTGCTCTGTGTTTGTCTTTATAAACCCAGATGAGAGGATGAAGGCAACAAGCTTCTGT ACCAACATACATGCCCCTTTGCCTCAAGTCTGGTTATTTTAGGGGGATGCTAGGTTGCTTTGGGT CTACCTTACTGAGAAAATGGCCCCAGGTCATTGTCATGTCCAGTTGTGGTAGACAGCATGTGTCC TAAAGGGTATATTCACATGCATGTGCAAAAATACAGGGGTCCTTCTAACCCTATCACAGAGAAGC AGGAGACTGC
R-loop 2	GCTACAGAAAGGTCAGCAGCTATATTTAACCTCAGACCAGGGTGCGGTGGGAGATCTGGTTTCC GGAAGACGGAATGGGGAGAAGGGCAGGTTCCCCGAGGCGCCCAGACACCCAATCCTCCCGGT GACATTTACAGCCTGGCCTTTGGGGTCGGGTCAACGCTAGGCTGGCAGGGGAAGGGCGGGGC CGTGAGGTGAGCCGGCGCTGCAGGAAGGGGCCACCACCAGAGGGGCCATTTTGCGGTGGAAA TGTCC
R-loop-3	GCTGTGGCATCCAGAGACATGGTTTCTTATCTCCTTAAGTGTTCAGCTGCTTTTCTTTCATTTATT CCACATATAATTACTATAATTGCTAAACATTTATTTAGTGTCAGGTAATGTGCTAAACAGAGAGTTA CTGCTCAGACATGTAATAATAATAAATAACACATCAAATAACCATACCATTTTAAGCTGTAGTATTA TGAAGGGAAATCTGGAGCAAAGAGAATAGACTGTAGGGAAACCAGTTAAGAAATAGGACATGGA GGCTAGGTGCAG
R-loop-4	GGAGGTGGAGAGAGGATGTTTTGCTTATCCAGAAAAGGGAGTGATTGCTTCCAGGGGCCTCAGG GGAATAAATCATAGAATCCTGGACAAGGTTTGAAGGACAGGTAGGATTTGGGTGGGTGGAGGAG GGTGCATGGGGTCAGAATTGTAACCGAAAACTCATTCCAGGTGGATAGAGAAAATTTCTAGTGTT GTTGTTTTTAAACTATTTGGGGGACTGGCACAGACCCTTTTTGAATACCTGATGGGCTCACATTTC TGTCGAATCCCAGCGGGTTCCTAGACCTCAGGA
R-loop-5	ATGTGGGCTGCCTAGAAAGGCATGGATGAGAGAAGCCTGGAGACAGGGATCCCAGGGAAACGC CCATGCAATTAGTCTATTTCTGCTGCAAGTAAGCATGCATTTGTAGGCTTGATGCTTTTTTTCTGC TTCTCCAGCCCTGGCCTGGGTCAATCCTTGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAA GGAAGCCCTGCTTCCTCCAGAGGGCGTCGCAGGACAGCTTTTCCTAGACAGGGGCTAGTATGT GCAGCTCCTGCACCGGGATACTGGTTGACAAGTTTGGCTGGG
-198 target	CTGACAAAAGAAGTCCTGGTATCTTCTATGGTGGGAGAAGAAAACTAGCTAAAGGGAAGAATAAA TTAGAGAAAAATTGGAATGACTGAATCGGAACAAGGCAAAGGCTATAAAAAAAATTAAGCAGCAG TATCCTCTTGGGGGCCCCTTCCCCACACTATCTCAATGCAAATATCTGTCTGAAACGGTCCCTGG CTAAACTCCACCCATGGGTTGGCCAGCCTTGCCTTGACCAATAGCCTTGACAAGGCAAACTTGA CCAATAGTCTTAGAGTATCCAGTGAGGCCAGG

-175 target	CTGACAAAAGAAGTCCTGGTATCTTCTATGGTGGGAGAAGAAAACTAGCTAAAGGGAAGAATAAA TTAGAGAAAAATTGGAATGACTGAATCGGAACAAGGCAAAGGCTATAAAAAAAATTAAGCAGCAG TATCCTCTTGGGGGCCCCTTCCCCACACTATCTCAATGCAAATATCTGTCTGAAACGGTCCCTGG CTAAACTCCACCCATGGGTTGGCCAGCCTTGCCTTGACCAATAGCCTTGACAAGGCAAACTTGA CCAATAGTCTTAGAGTATCCAGTGAGGCCAGG
BCL11A enhancer	TCTTAGACATAACACACCAGGGTCAATACAACTTTGAAGCTAGTCTAGTGCAAGCTAACAGTTGC TTTTATCACAGGCTCCAGGAAGGGTTTGGCCTCTGATTAGGGTGGGGGCGTGGGTGGGGTAGA AGAGGACTGGCAGAC

Supplementary Note 1. Python script to extract allele frequencies with two edits.

```
#!/usr/bin/env python
# coding: utf-8

# In[1]:


#import python packages
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np


# In[2]:


#import data

#names is a tab-delimited file containing the CRISPResso2.0 allele frequency table filenames.
#here, names.txt contains 18 filenames, sorted by site, editor, and biological replicate.
names = "names.txt"
df = pd.read_table(names,delim_whitespace=True)
df.set_index('file',inplace=True)

#group the files according to their corresponding genomic locus.
files_site5 = list(df.index.values)[0:6] #six files: three for three ABE7.10 replicates; three for three ABE8e replicates.
files_site6 = list(df.index.values)[6:12]
files_site7 = list(df.index.values)[12:18]


# In[3]:


def get_data(site,D):
    '''Function inputs: site number (here: 5,6, or 7), and allele frequencies table (D).

       Function output: the percent of total alleles in which both of the adenine (A) substrates
       are converted to the desired guanine (G) products.

       Gu and Gd correspond to a G product at the upstream and downstream target positions, respectively.
    '''

    if site == 5:
        total_reads = D[D["Aligned_Sequence"].str.contains('.....................', regex=True)]["#Reads"].sum()
```

```
        #Gu = G at upstream position; Gd = G at downstream position
        Gu_Gd =     D[D["Aligned_Sequence"].str.contains('.....GG.....',
regex=True)]["#Reads"].sum()
        percent_Gu_Gd = 100*Gu_Gd/total_reads
        data = {'% of total reads with both substrates edited': percent_Gu_Gd}

    if site == 6:
        total_reads = D[D["Aligned_Sequence"].str.contains('.....',
regex=True)]["#Reads"].sum()
        Gu_Gd =     D[D["Aligned_Sequence"].str.contains('.....G...G.....',
regex=True)]["#Reads"].sum()
        percent_Gu_Gd = 100*Gu_Gd/total_reads
        data = {'% of total reads with both substrates edited': percent_Gu_Gd}

    if site == 7:
        total_reads = D[D["Aligned_Sequence"].str.contains('.....',
regex=True)]["#Reads"].sum()
        Gu_Gd =     D[D["Aligned_Sequence"].str.contains('.....G....G.....',
regex=True)]["#Reads"].sum()
        percent_Gu_Gd = 100*Gu_Gd/total_reads
        data = {'% of total reads with both substrates edited': percent_Gu_Gd}

    return data


# In[4]:


def makedfs(site,names):
    '''Function inputs: Site number (here: 5,6, or 7) and the names.txt file.

        Please note: ABE7.10 is henceforth referred to as ABE7 in the code.

        Function output: Two tables, one for ABE7 and one for ABE8e, that contain the get_data
function output for
        each of the three biological replicates at the designated site.
    '''

    #i, below, is initialized and subsequently used to track iterations of the for loop
    #and to appropriately assign the allele frequency table to the corresponding editor
(ABE7.10 or ABE8e)
    i = 0
    global df_ABE7
    global df_ABE8e

    for file in names:
        D = pd.read_table(file,delim_whitespace=True)
        #the output after the first three iterations (0<=i<=2) is the ABE7.10 dataframe
        if i == 0:
```

```
            df_ABE7 = pd.DataFrame(get_data(site,D),index=[str(file)])
        if 1 <= i <= 2:
            D_output_new = pd.DataFrame(get_data(site,D),index=[str(file)])
            df_ABE7 = pd.concat([df_ABE7, D_output_new], axis =0)
        #the output after the final three iterations (3<=i<=5) is the ABE8e dataframe
        if i == 3:
            df_ABE8e= pd.DataFrame(get_data(site,D),index=[str(file)])
        if 4<=i<=5:
            D_output_new = pd.DataFrame(get_data(site,D),index=[str(file)])
            df_ABE8e= pd.concat([df_ABE8e, D_output_new], axis =0)
        i+=1 #tracking iteration
    return df_ABE7,df_ABE8e


# In[5]:


#create dataframes
ABE7Site5,ABE8eSite5 = makedfs(5,files_site5)
ABE7Site6,ABE8eSite6 = makedfs(6,files_site6)
ABE7Site7,ABE8eSite7 = makedfs(7,files_site7)


# In[6]:


#for the given editor and site, obtain the average, across three biologial replicates,
#percent of total alleles in which both substrate bases are converted.
ABE7_Site5_GuGd_average = float(ABE7Site5.mean())
ABE8e_Site5_GuGd_average = float(ABE8eSite5.mean())

ABE7_Site6_GuGd_average = float(ABE7Site6.mean())
ABE8e_Site6_GuGd_average = float(ABE8eSite6.mean())

ABE7_Site7_GuGd_average = float(ABE7Site7.mean())
ABE8e_Site7_GuGd_average = float(ABE8eSite7.mean())

#store averages by editor for subsequent plotting
means_ABE7 =
(ABE7_Site5_GuGd_average,ABE7_Site6_GuGd_average,ABE7_Site7_GuGd_average)
means_ABE8e =
(ABE8e_Site5_GuGd_average,ABE8e_Site6_GuGd_average,ABE8e_Site7_GuGd_average)

#for the given editor and site, obtain the standard deviation, across three
#biological replicates, in the percent of total alleles in which both substrate bases are
converted.

ABE7_Site5_GuGd_std = float(ABE7Site5.std())
ABE8e_Site5_GuGd_std = float(ABE8eSite5.std())
```

```
ABE7_Site6_GuGd_std = float(ABE7Site6.std())
ABE8e_Site6_GuGd_std = float(ABE8eSite6.std())

ABE7_Site7_GuGd_std = float(ABE7Site7.std())
ABE8e_Site7_GuGd_std = float(ABE8eSite7.std())

#store standard deviations by editor for subsequent plotting
std_ABE7 = (ABE7_Site5_GuGd_std,ABE7_Site6_GuGd_std,ABE7_Site7_GuGd_std)
std_ABE8e = (ABE8e_Site5_GuGd_std,ABE8e_Site6_GuGd_std,ABE8e_Site7_GuGd_std)


# In[7]:


#number of sites
n_sites = 3

#create plot
fig, ax = plt.subplots()
index = np.arange(n_sites)
bar_width = 0.35
opacity = 1

#plot data
ABE7bars = plt.bar(index, means_ABE7,
bar_width,alpha=opacity,color='#A0A0A3',label='ABE7.10',yerr=std_ABE7,capsize=5)
ABE8ebars = plt.bar(index + bar_width, means_ABE8e, bar_width,
alpha=opacity,color='#F94040',label='ABE8e',yerr=std_ABE8e,capsize=5)

#label plot
plt.ylabel('Percent of total alleles in which both \n substrate bases are converted')
plt.xticks(index + bar_width/2, ('Site 5', 'Site 6', 'Site 7'))

#display plot with legend
plt.legend()
plt.tight_layout()
plt.show()


# In[ ]:
```

Supplementary Sequences 1. Amino acid sequences for each base editor in this study.

ABE8e

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGE
GWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG
VRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSI
NSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS
KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK
VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL
ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR
RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG
DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI
PHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT
PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL
KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL
SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA
NLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL
DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFF
YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQT
GGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELL
GITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELAL
PSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLS
AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE
TRIDLSQLGGDSGGSKRTADGSEFEPKKKRKV

ABE8e-dimer

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIG
EGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVF
GARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSS
TDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDER
EVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCV
MCAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFY
RMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIG
TNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRR
KNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLR
KKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINA
SGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS
KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD
LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNR
EDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGN
SRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVE

DRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ
VSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQ
KNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDY
DVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKL
VSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVE
KGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLA
SAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSK
RVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL
DATLIHQSITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKV

SaABE8e
MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGE
GWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG
VRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSI
NSGGSSGGSSGSETPGTSESATPESSGGSSGGSGKRNYILGLAIGITSVGYGIIDYETRDVID
AGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEAR
VKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQ
LERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGE
GSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYY
EKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIEN
AELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHT
NDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIEL
AREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIP
LEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFK
KHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNN
LDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMEN
QMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKD
DKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYK
YYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLD
NGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIG
VNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQI
IKKGSGGSKRTADGSEFEPKKKRKV

SaABE8e-dimer
MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIG
EGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVF
GARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSS
TDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDER
EVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCV
MCAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFY
RMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPESSGGSSGGSGKRNYILGLAI
GITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDY
NLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKE
QISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFI
DTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNAL

NDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFT
NLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTG
THNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQS
IKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKI
KLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRT
PFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRY
ATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFI
FKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDK
KPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKL
KLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNK
VVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASF
YNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYST
DILGNLYEVKSKKHPQIIKKGSGGSKRTADGSEFEPKKKRKV

LbABE8e
MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGE
GWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG
VRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSI
NSGGSSGGSSGSETPGTSESATPESSGGSSGGSSKLEKFTNCYSLSKTLRFKAIPVGKTQE
NIDNKRLLVEDEKRAEDYKGVKKLLDRYYLSFINDVLHSIKLKNLNNYISLFRKKTRTEKENKE
LENLEINLRKEIAKAFKGNEGYKSLFKKDIIETILPEFLDDKDEIALVNSFNGFTTAFTGFFDNRE
NMFSEEAKSTSIAFRCINENLTRYISNMDIFEKVDAIFDKHEVQEIKEKILNSDYDVEDFFEGEF
FNFVLTQEGIDVYNAIIGGFVTESGEKIKGLNEYINLYNQKTKQKLPKFKPLYKQVLSDRESLS
FYGEGYTSDEEVLEVFRNTLNKNSEIFSSIKKLEKLFKNFDEYSSAGIFVKNGPAISTISKDIFG
EWNVIRDKWNAEYDDIHLKKKAVVTEKYEDDRRKSFKKIGSFSLEQLQEYADADLSVVEKLK
EIIIQKVDEIYKVYGSSEKLFDADFVLEKSLKKNDAVVAIMKDLLDSVKSFENYIKAFFGEGKET
NRDESFYGDFVLAYDILLKVDHIYDAIRNYVTQKPYSKDKFKLYFQNPQFMGGWDKDKETDY
RATILRYGSKYYLAIMDKKYAKCLQKIDKDDVNGNYEKINYKLLPGPNKMLPKVFFSKKWMA
YYNPSEDIQKIYKNGTFKKGDMFNLNDCHKLIDFFKDSISRYPKWSNAYDFNFSETEKYKDIA
GFYREVEEQGYKVSFESASKKEVDKLVEEGKLYMFQIYNKDFSDKSHGTPNLHTMYFKLLF
DENNHGQIRLSGGAELFMRRASLKKEELVVHPANSPIANKNPDNPKKTTTLSYDVYKDKRFS
EDQYELHIPIAINKCPKNIFKINTEVRVLLKHDDNPYVIGIARGERNLLYIVVVDGKGNIVEQYSL
NEIINNFNGIRIKTDYHSLLDKKEKERFEARQNWTSIENIKELKAGYISQVVHKICELVEKYDAVI
ALEDLNSGFKNSRVKVEKQVYQKFEKMLIDKLNYMVDKKSNPCATGGALKGYQITNKFESFK
SMSTQNGFIFYIPAWLTSKIDPSTGFVNLLKTKYTSIADSKKFISSFDRIMYVPEEDLFEFALDY
KNFSRTDADYIKKWKLYSYGNRIRIFRNPKKNNVFDWEEVCLTSAYKELFNKYGINYQQGDI
RALLCEQSDKAFYSSFMALMSLMLQMRNSITGRTDVDFLISPVKNSDGIFYDSRNYEAQENA
ILPKNADANGAYNIARKVLWAIGQFKKAEDEKLDKVKIAISNKEWLEYAQTSVKSGGSKRTAD
GSEFEPKKKRKV

LbABE8e-dimer
MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIG
EGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVF
GARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSS
TDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDER
EVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCV
MCAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFY
RMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPESSGGSSGGSSKLEKFTNCY
SLSKTLRFKAIPVGKTQENIDNKRLLVEDEKRAEDYKGVKKLLDRYYLSFINDVLHSIKLKNLN

NYISLFRKKTRTEKENKELENLEINLRKEIAKAFKGNEGYKSLFKKDIIETILPEFLDDKDEIALV
NSFNGFTTAFTGFFDNRENMFSEEAKSTSIAFRCINENLTRYISNMDIFEKVDAIFDKHEVQEI
KEKILNSDYDVEDFFEGEFFNFVLTQEGIDVYNAIIGGFVTESGEKIKGLNEYINLYNQKTKQK
LPKFKPLYKQVLSDRESLSFYGEGYTSDEEVLEVFRNTLNKNSEIFSSIKKLEKLFKNFDEYS
SAGIFVKNGPAISTISKDIFGEWNVIRDKWNAEYDDIHLKKKAVVTEKYEDDRRKSFKKIGSFS
LEQLQEYADADLSVVEKLKEIIIQKVDEIYKVYGSSEKLFDADFVLEKSLKKNDAVVAIMKDLL
DSVKSFENYIKAFFGEGKETNRDESFYGDFVLAYDILLKVDHIYDAIRNYVTQKPYSKDKFKLY
FQNPQFMGGWDKDKETDYRATILRYGSKYYLAIMDKKYAKCLQKIDKDDVNGNYEKINYKLL
PGPNKMLPKVFFSKKWMAYYNPSEDIQKIYKNGTFKKGDMFNLNDCHKLIDFFKDSISRYPK
WSNAYDFNFSETEKYKDIAGFYREVEEQGYKVSFESASKKEVDKLVEEGKLYMFQIYNKDFS
DKSHGTPNLHTMYFKLLFDENNHGQIRLSGGAELFMRRASLKKEELVVHPANSPIANKNPDN
PKKTTTLSYDVYKDKRFSEDQYELHIPIAINKCPKNIFKINTEVRVLLKHDDNPYVIGIARGERN
LLYIVVVDGKGNIVEQYSLNEIINNFNGIRIKTDYHSLLDKKEKERFEARQNWTSIENIKELKAG
YISQVVHKICELVEKYDAVIALEDLNSGFKNSRVKVEKQVYQKFEKMLIDKLNYMVDKKSNPC
ATGGALKGYQITNKFESFKSMSTQNGFIFYIPAWLTSKIDPSTGFVNLLKTKYTSIADSKKFISS
FDRIMYVPEEDLFEFALDYKNFSRTDADYIKKWKLYSYGNRIRIFRNPKKNNVFDWEEVCLTS
AYKELFNKYGINYQQGDIRALLCEQSDKAFYSSFMALMSLMLQMRNSITGRTDVDFLISPVK
NSDGIFYDSRNYEAQENAILPKNADANGAYNIARKVLWAIGQFKKAEDEKLDKVKIAISNKEW
LEYAQTSVKSGGSKRTADGSEFEPKKKRKV

LbABE7.10

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIG
EGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVF
GARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSS
TDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDER
EVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCV
MCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFR
MPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSKLEKFTNCYS
LSKTLRFKAIPVGKTQENIDNKRLLVEDEKRAEDYKGVKKLLDRYYLSFINDVLHSIKLKNLNN
YISLFRKKTRTEKENKELENLEINLRKEIAKAFKGNEGYKSLFKKDIIETILPEFLDDKDEIALVN
SFNGFTTAFTGFFDNRENMFSEEAKSTSIAFRCINENLTRYISNMDIFEKVDAIFDKHEVQEIK
EKILNSDYDVEDFFEGEFFNFVLTQEGIDVYNAIIGGFVTESGEKIKGLNEYINLYNQKTKQKL
PKFKPLYKQVLSDRESLSFYGEGYTSDEEVLEVFRNTLNKNSEIFSSIKKLEKLFKNFDEYSS
AGIFVKNGPAISTISKDIFGEWNVIRDKWNAEYDDIHLKKKAVVTEKYEDDRRKSFKKIGSFSL
EQLQEYADADLSVVEKLKEIIIQKVDEIYKVYGSSEKLFDADFVLEKSLKKNDAVVAIMKDLLD
SVKSFENYIKAFFGEGKETNRDESFYGDFVLAYDILLKVDHIYDAIRNYVTQKPYSKDKFKLYF
QNPQFMGGWDKDKETDYRATILRYGSKYYLAIMDKKYAKCLQKIDKDDVNGNYEKINYKLLP
GPNKMLPKVFFSKKWMAYYNPSEDIQKIYKNGTFKKGDMFNLNDCHKLIDFFKDSISRYPKW
SNAYDFNFSETEKYKDIAGFYREVEEQGYKVSFESASKKEVDKLVEEGKLYMFQIYNKDFSD
KSHGTPNLHTMYFKLLFDENNHGQIRLSGGAELFMRRASLKKEELVVHPANSPIANKNPDNP
KKTTTLSYDVYKDKRFSEDQYELHIPIAINKCPKNIFKINTEVRVLLKHDDNPYVIGIARGERNL
LYIVVVDGKGNIVEQYSLNEIINNFNGIRIKTDYHSLLDKKEKERFEARQNWTSIENIKELKAGY
ISQVVHKICELVEKYDAVIALEDLNSGFKNSRVKVEKQVYQKFEKMLIDKLNYMVDKKSNPCA
TGGALKGYQITNKFESFKSMSTQNGFIFYIPAWLTSKIDPSTGFVNLLKTKYTSIADSKKFISSF
DRIMYVPEEDLFEFALDYKNFSRTDADYIKKWKLYSYGNRIRIFRNPKKNNVFDWEEVCLTSA
YKELFNKYGINYQQGDIRALLCEQSDKAFYSSFMALMSLMLQMRNSITGRTDVDFLISPVKN
SDGIFYDSRNYEAQENAILPKNADANGAYNIARKVLWAIGQFKKAEDEKLDKVKIAISNKEWL
EYAQTSVKSGGSKRTADGSEFEPKKKRKV

enAsABE8e

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGE
GWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG
VRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSI
NSGGSSGGSSGSETPGTSESATPESSGGSSGGSMTQFEGFTNLYQVSKTLRFELIPQGKTL
KHIQEQGFIEEDKARNDHYKELKPIIDRIYKTYADQCLQLVQLDWENLSAAIDSYRKEKTEETR
NALIEEQATYRNAIHDYFIGRTDNLTDAINKRHAEIYKGLFKAELFNGKVLKQLGTVTTTEHEN
ALLRSFDKFTTYFSGFYRNRKNVFSAEDISTAIPHRIVQDNFPKFKENCHIFTRLITAVPSLREH
FENVKKAIGIFVSTSIEEVFSFPFYNQLLTQTQIDLYNQLLGGISREAGTEKIKGLNEVLNLAIQK
NDETAHIIASLPHRFIPLFKQILSDRNTLSFILEEFKSDEEVIQSFCKYKTLLRNENVLETAEALF
NELNSIDLTHIFISHKKLETISSALCDHWDTLRNALYERRISELTGKITKSAKEKVQRSLKHEDI
NLQEIISAAGKELSEAFKQKTSEILSHAHAALDQPLPTTLKKQEEKEILKSQLDSLLGLYHLLD
WFAVDESNEVDPEFSARLTGIKLEMEPSLSFYNKARNYATKKPYSVEKFKLNFQMPTLARG
WDVNREKNNGAILFVKNGLYYLGIMPKQKGRYKALSFEPTEKTSEGFDKMYYDYFPDAAKMI
PKCSTQLKAVTAHFQTHTTPILLSNNFIEPLEITKEIYDLNNPEKEPKKFQTAYAKKTGDQKGY
REALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYHISFQRIAEKEIMD
AVETGKLYLFQIYNKDFAKGHHGKPNLHTLYWTGLFSPENLAKTSIKLNGQAELFYRPKSRM
KRMAHRLGEKMLNKKLKDQKTPIPDTLYQELYDYVNHRLSHDLSDEARALLPNVITKEVSHEI
IKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEHPETPIIGIARGERNLIYITVIDSTG
KILEQRSLNTIQQFDYQKKLDNREKERVAARQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQ
AVVVLENLNFGFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGVLNPYQLTDQF
TSFAKMGTQSGFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHESRKHFLEGFDFLHYDVKTG
DFILHFKMNRNLSFQRGLPGFMPAWDIVFEKNETQFDAKGTPFIAGKRIVPVIENHRFTGRYR
DLYPANELIALLEEKGIVFRDGSNILPKLLENDDSHAIDTMVALIRSVLQMRNSNAATGEDYIN
SPVRDLNGVCFDSRFQNPEWPMDADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQDWL
AYIQELRNSGGSKRTADGSEFEPKKKRKV

enAsABE8e-dimer

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIG
EGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVF
GARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSS
TDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDER
EVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCV
MCAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFY
RMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPESSGGSSGGSMTQFEGFTNL
YQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARNDHYKELKPIIDRIYKTYADQCLQLVQLDWE
NLSAAIDSYRKEKTEETRNALIEEQATYRNAIHDYFIGRTDNLTDAINKRHAEIYKGLFKAELFN
GKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYRNRKNVFSAEDISTAIPHRIVQDNFPKFK
ENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFYNQLLTQTQIDLYNQLLGGISRE
AGTEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQILSDRNTLSFILEEFKSDEEVIQSFC
KYKTLLRNENVLETAEALFNELNSIDLTHIFISHKKLETISSALCDHWDTLRNALYERRISELTG
KITKSAKEKVQRSLKHEDINLQEIISAAGKELSEAFKQKTSEILSHAHAALDQPLPTTLKKQEE
KEILKSQLDSLLGLYHLLDWFAVDESNEVDPEFSARLTGIKLEMEPSLSFYNKARNYATKKPY
SVEKFKLNFQMPTLARGWDVNREKNNGAILFVKNGLYYLGIMPKQKGRYKALSFEPTEKTSE
GFDKMYYDYFPDAAKMIPKCSTQLKAVTAHFQTHTTPILLSNNFIEPLEITKEIYDLNNPEKEP
KKFQTAYAKKTGDQKGYREALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGEYYAEL
NPLLYHISFQRIAEKEIMDAVETGKLYLFQIYNKDFAKGHHGKPNLHTLYWTGLFSPENLAKT
SIKLNGQAELFYRPKSRMKRMAHRLGEKMLNKKLKDQKTPIPDTLYQELYDYVNHRLSHDLS
DEARALLPNVITKEVSHEIIKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEHPETP

IIGIARGERNLIYITVIDSTGKILEQRSLNTIQQFDYQKKLDNREKERVAARQAWSVVGTIKDLK
QGYLSQVIHEIVDLMIHYQAVVVLENLNFGFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDY
PAEKVGGVLNPYQLTDQFTSFAKMGTQSGFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHES
RKHFLEGFDFLHYDVKTGDFILHFKMNRNLSFQRGLPGFMPAWDIVFEKNETQFDAKGTPFI
AGKRIVPVIENHRFTGRYRDLYPANELIALLEEKGIVFRDGSNILPKLLENDDSHAIDTMVALIR
SVLQMRNSNAATGEDYINSPVRDLNGVCFDSRFQNPEWPMDADANGAYHIALKGQLLLNHL
KESKDLKLQNGISNQDWLAYIQELRNSGGSKRTADGSEFEPKKKRKV

enAsABE7.10
MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIG
EGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVF
GARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSS
TDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDER
EVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCV
MCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFR
MPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSMTQFEGFTNL
YQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARNDHYKELKPIIDRIYKTYADQCLQLVQLDWE
NLSAAIDSYRKEKTEETRNALIEEQATYRNAIHDYFIGRTDNLTDAINKRHAEIYKGLFKAELFN
GKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYRNRKNVFSAEDISTAIPHRIVQDNFPKFK
ENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFYNQLLTQTQIDLYNQLLGGISRE
AGTEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQILSDRNTLSFILEEFKSDEEVIQSFC
KYKTLLRNENVLETAEALFNELNSIDLTHIFISHKKLETISSALCDHWDTLRNALYERRISELTG
KITKSAKEKVQRSLKHEDINLQEIISAAGKELSEAFKQKTSEILSHAHAALDQPLPTTLKKQEE
KEILKSQLDSLLGLYHLLDWFAVDESNEVDPEFSARLTGIKLEMEPSLSFYNKARNYATKKPY
SVEKFKLNFQMPTLARGWDVNREKNNGAILFVKNGLYYLGIMPKQKGRYKALSFEPTEKTSE
GFDKMYYDYFPDAAKMIPKCSTQLKAVTAHFQTHTTPILLSNNFIEPLEITKEIYDLNNPEKEP
KKFQTAYAKKTGDQKGYREALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGEYYAEL
NPLLYHISFQRIAEKEIMDAVETGKLYLFQIYNKDFAKGHHGKPNLHTLYWTGLFSPENLAKT
SIKLNGQAELFYRPKSRMKRMAHRLGEKMLNKKLKDQKTPIPDTLYQELYDYVNHRLSHDLS
DEARALLPNVITKEVSHEIIKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEHPETP
IIGIARGERNLIYITVIDSTGKILEQRSLNTIQQFDYQKKLDNREKERVAARQAWSVVGTIKDLK
QGYLSQVIHEIVDLMIHYQAVVVLENLNFGFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDY
PAEKVGGVLNPYQLTDQFTSFAKMGTQSGFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHES
RKHFLEGFDFLHYDVKTGDFILHFKMNRNLSFQRGLPGFMPAWDIVFEKNETQFDAKGTPFI
AGKRIVPVIENHRFTGRYRDLYPANELIALLEEKGIVFRDGSNILPKLLENDDSHAIDTMVALIR
SVLQMRNSNAATGEDYINSPVRDLNGVCFDSRFQNPEWPMDADANGAYHIALKGQLLLNHL
KESKDLKLQNGISNQDWLAYIQELRNSGGSKRTADGSEFEPKKKRKV

NG-ABE8e
MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGE
GWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG
VRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSI
NSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS
KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK
VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL
ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR
RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG
DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI

PHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT
PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL
KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL
SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA
NLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL
DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFF
YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQT
GGFSKESIRPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELL
GITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASARFLQKGNELAL
PSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLS
AYNKHRDKPIREQAENIIHLFTLTNLGAPRAFKYFDTTIDRKVYRSTKEVLDATLIHQSITGLYE
TRIDLSQLGGDSGGSKRTADGSEFEPKKKRKV

NG-ABE8e-dimer

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIG
EGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVF
GARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSS
TDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDER
EVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCV
MCAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFY
RMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIG
TNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRR
KNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLR
KKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINA
SGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS
KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD
LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNR
EDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGN
SRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVE
DRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ
VSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQ
KNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDY
DVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFD
NLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKL
VSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
MPQVNIVKKTEVQTGGFSKESIRPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVE
KGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLA
SARFLQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSK
RVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPRAFKYFDTTIDRKVYRSTKEVL
DATLIHQSITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKV

SaKKH-ABE8e

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGE
GWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG
VRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSI
NSGGSSGGSSGSETPGTSESATPESSGGSSGGSGKRNYILGLAIGITSVGYGIIDYETRDVID
AGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEAR
VKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQ
LERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGE
GSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYY
EKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIEN
AELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHT
NDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIEL
AREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIP
LEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFK
KHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNN
LDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMEN
QMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKD
DKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYK
YYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLD
NGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIG
VNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQI
IKKGSGGSKRTADGSEFEPKKKRKV

SaKKH-ABE8e-dimer:

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIG
EGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVF
GARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSS
TDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDER
EVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCV
MCAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFY
RMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPESSGGSSGGSGKRNYILGLAI
GITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDY
NLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKE
QISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFI
DTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNAL
NDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFT
NLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTG
THNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQS
IKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKI
KLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRT
PFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRY
ATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFI
FKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDK
KPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKL
KLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNK
VVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASF
YNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYST
DILGNLYEVKSKKHPQIIKKGSGGSKRTADGSEFEPKKKRKV

CP1028-ABE8e

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGE
GWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG
VRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSI
NSGGSSGGSSGSETPGTSESATPESSGGSSGGSEIGKATAKYFFYSNIMNFFKTEITLANGE
IRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLI
ARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK
GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENII
HLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGSGGS
GGSGGSGGSGGSGGMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI
GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLN
PDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF
GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILL
SDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY
PFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIE
RMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKT
NRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV
LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL
KSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDE
LVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQN
EKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV
PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVA
QILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQSGGSKRTADGSEFEPKKKRKV

CP1028-ABE8e-dimer

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIG
EGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVF
GARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSS
TDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDER
EVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCV
MCAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFY
RMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPESSGGSSGGSEIGKATAKYF
FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQ
TGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKEL
LGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELA
LPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLY
ETRIDLSQLGGDGGSGGSGGSGGSGGSGGSGGSGGMDKKYSIGLAIGTNSVGWAVITDEYKVP
SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMA
KVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIY
LALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKS
RRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQI
GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK
YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGS
IPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT

PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL
KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL
SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA
NLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL
DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQSGGSKRTADG
SEFEPKKKRKV

CP1041-ABE8e
MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGE
GWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG
VRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSI
NSGGSSGGSSGSETPGTSESATPESSGGSSGGSNIMNFFKTEITLANGEIRKRPLIETNGET
GEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLI
IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQL
FVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPA
AFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGSGGSGGSGGSGGSG
GSGGDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE
ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV
DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQ
LVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPN
FKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA
PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKP
ILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI
LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNE
KVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDY
FKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQL
IHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENI
VIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNY
WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE
NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFV
YGDYKVYDVRKMIAKSEQEIGKATAKYFFYSSGGSKRTADGSEFEPKKKRKV

ABE8e(TadA-8e V106W)
MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGE
GWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG
WRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSI
NSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS
KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK
VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL
ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR
RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG
DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY

KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI
PHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT
PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL
KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL
SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA
NLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL
DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFF
YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQT
GGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELL
GITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELAL
PSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLS
AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE
TRIDLSQLGGDSGGSKRTADGSEFEPKKKRKV

SaABE8e(TadA-8e V106W)
MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGE
GWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG
WRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSI
NSGGSSGGSSGSETPGTSESATPESSGGSSGGSGKRNYILGLAIGITSVGYGIIDYETRDVID
AGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEAR
VKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQ
LERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGE
GSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYY
EKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIEN
AELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHT
NDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIEL
AREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIP
LEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFK
KHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNN
LDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMEN
QMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKD
DKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYK
YYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLD
NGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIG
VNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQI
IKKGSGGSKRTADGSEFEPKKKRKV

LbABE8e(TadA-8e V106W)
MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGE
GWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG
WRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSI
NSGGSSGGSSGSETPGTSESATPESSGGSSGGSSKLEKFTNCYSLSKTLRFKAIPVGKTQE
NIDNKRLLVEDEKRAEDYKGVKKLLDRYYLSFINDVLHSIKLKNLNNYISLFRKKTRTEKENKE
LENLEINLRKEIAKAFKGNEGYKSLFKKDIIETILPEFLDDKDEIALVNSFNGFTTAFTGFFDNRE
NMFSEEAKSTSIAFRCINENLTRYISNMDIFEKVDAIFDKHEVQEIKEKILNSDYDVEDFFEGEF

FNFVLTQEGIDVYNAIIGGFVTESGEKIKGLNEYINLYNQKTKQKLPKFKPLYKQVLSDRESLS
FYGEGYTSDEEVLEVFRNTLNKNSEIFSSIKKLEKLFKNFDEYSSAGIFVKNGPAISTISKDIFG
EWNVIRDKWNAEYDDIHLKKKAVVTEKYEDDRRKSFKKIGSFSLEQLQEYADADLSVVEKLK
EIIIQKVDEIYKVYGSSEKLFDADFVLEKSLKKNDAVVAIMKDLLDSVKSFENYIKAFFGEGKET
NRDESFYGDFVLAYDILLKVDHIYDAIRNYVTQKPYSKDKFKLYFQNPQFMGGWDKDKETDY
RATILRYGSKYYLAIMDKKYAKCLQKIDKDDVNGNYEKINYKLLPGPNKMLPKVFFSKKWMA
YYNPSEDIQKIYKNGTFKKGDMFNLNDCHKLIDFFKDSISRYPKWSNAYDFNFSETEKYKDIA
GFYREVEEQGYKVSFESASKKEVDKLVEEGKLYMFQIYNKDFSDKSHGTPNLHTMYFKLLF
DENNHGQIRLSGGAELFMRRASLKKEELVVHPANSPIANKNPDNPKKTTTLSYDVYKDKRFS
EDQYELHIPIAINKCPKNIFKINTEVRVLLKHDDNPYVIGIARGERNLLYIVVVDGKGNIVEQYSL
NEIINNFNGIRIKTDYHSLLDKKEKERFEARQNWTSIENIKELKAGYISQVVHKICELVEKYDAVI
ALEDLNSGFKNSRVKVEKQVYQKFEKMLIDKLNYMVDKKSNPCATGGALKGYQITNKFESFK
SMSTQNGFIFYIPAWLTSKIDPSTGFVNLLKTKYTSIADSKKFISSFDRIMYVPEEDLFEFALDY
KNFSRTDADYIKKWKLYSYGNRIRIFRNPKKNNVFDWEEVCLTSAYKELFNKYGINYQQGDI
RALLCEQSDKAFYSSFMALMSLMLQMRNSITGRTDVDFLISPVKNSDGIFYDSRNYEAQENA
ILPKNADANGAYNIARKVLWAIGQFKKAEDEKLDKVKIAISNKEWLEYAQTSVKSGGSKRTAD
GSEFEPKKKRKV

ABE8e(TadA-8e V82G)
MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGE
GWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG
VRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSI
NSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS
KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK
VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL
ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR
RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG
DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI
PHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT
PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL
KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL
SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA
NLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL
DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFF
YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQT
GGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELL
GITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELAL
PSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLS
AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE
TRIDLSQLGGDSGGSKRTADGSEFEPKKKRKV

ABE8e(TadA-8e K20A R21A)
MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGE
GWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG

VRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSI
NSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS
KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK
VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL
ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR
RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG
DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI
PHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT
PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL
KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL
SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA
NLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGI
KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD
SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL
DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFF
YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQT
GGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELL
GITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELAL
PSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLS
AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE
TRIDLSQLGGDSGGSKRTADGSEFEPKKKRKV