

Supplementary Information

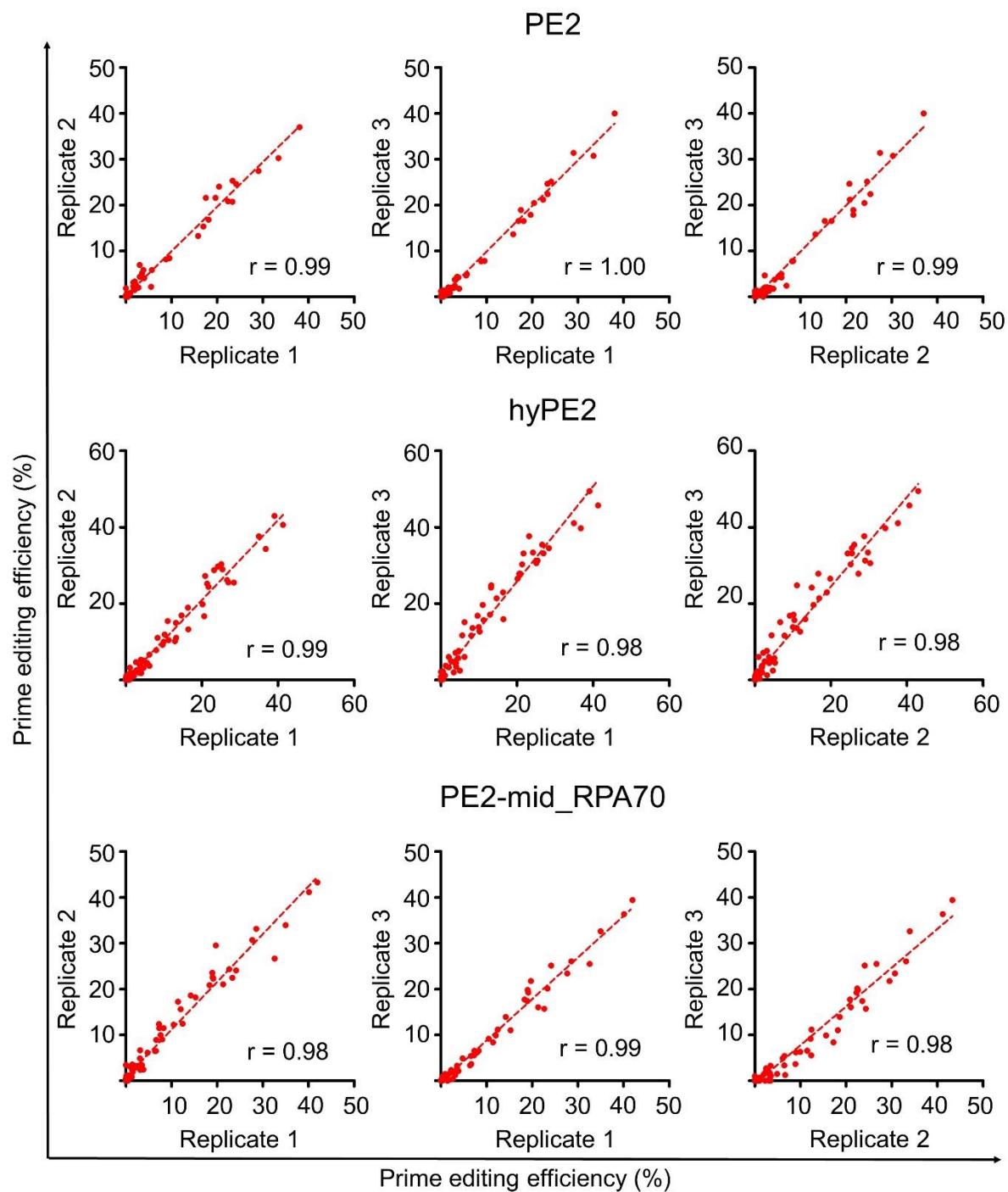
Generation of a more efficient prime editor 2 by addition of the Rad51 DNA-binding domain

Contents:

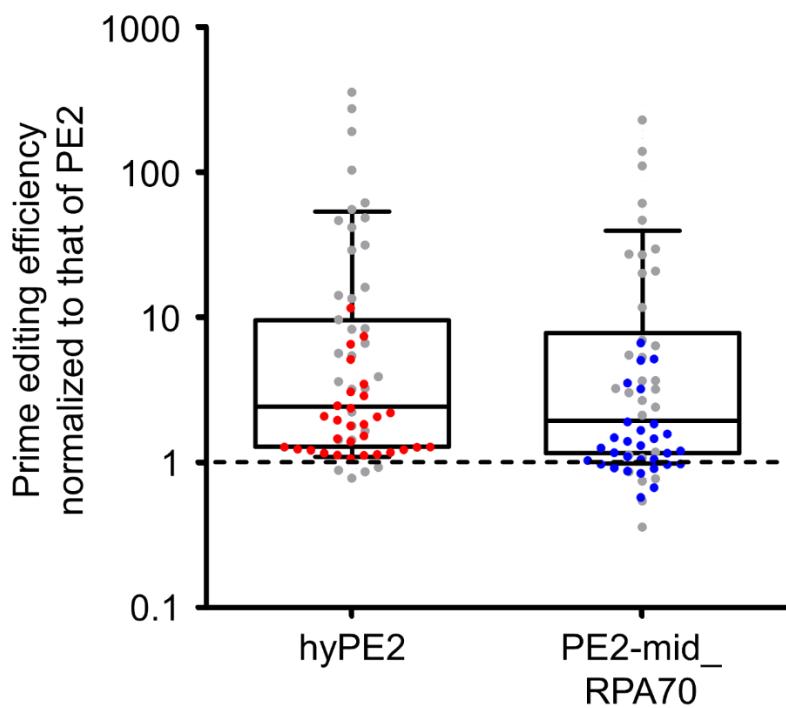
Supplementary Figures 1 – 5

Supplementary Tables 1 – 4

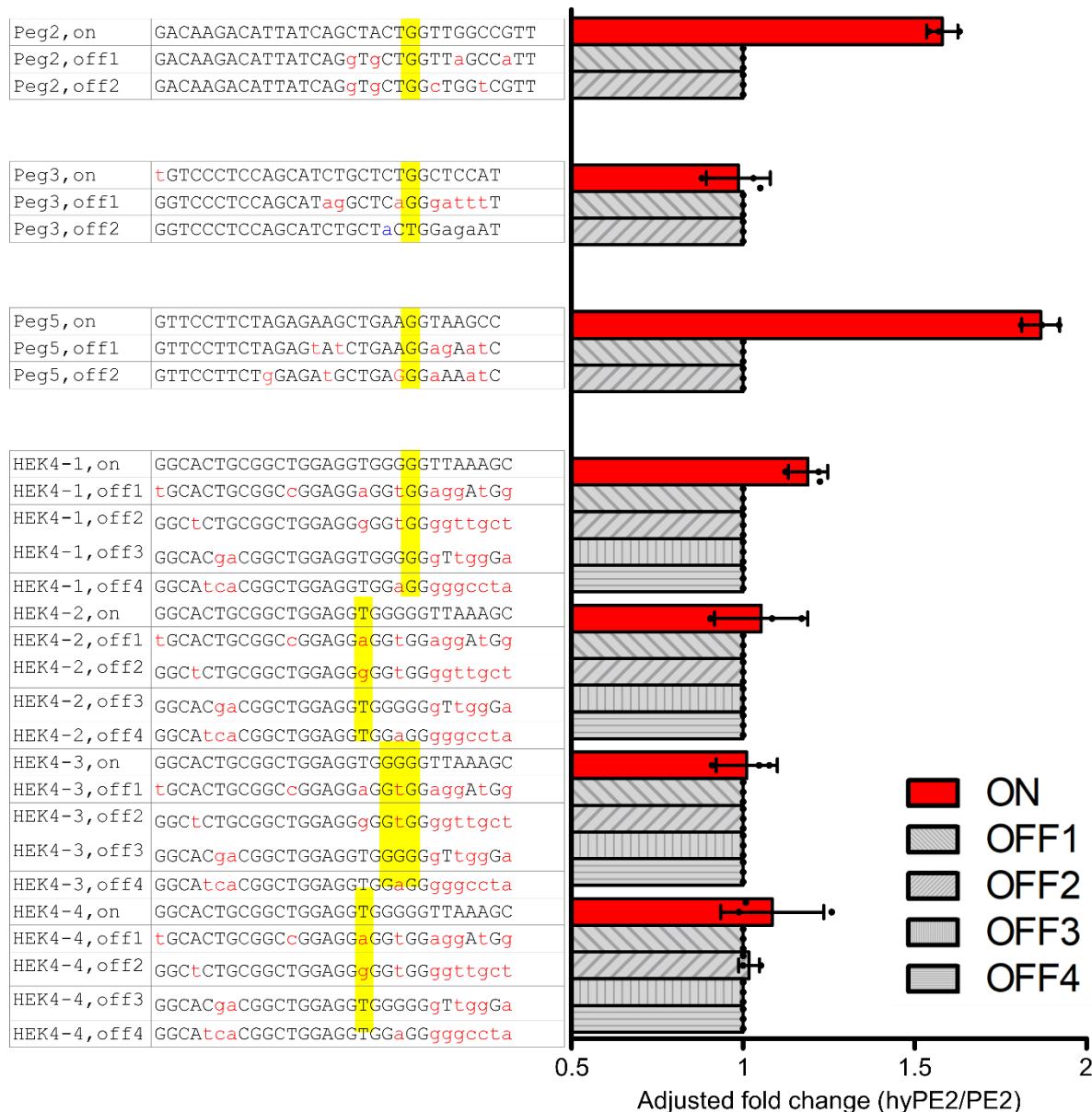
Supplementary Note



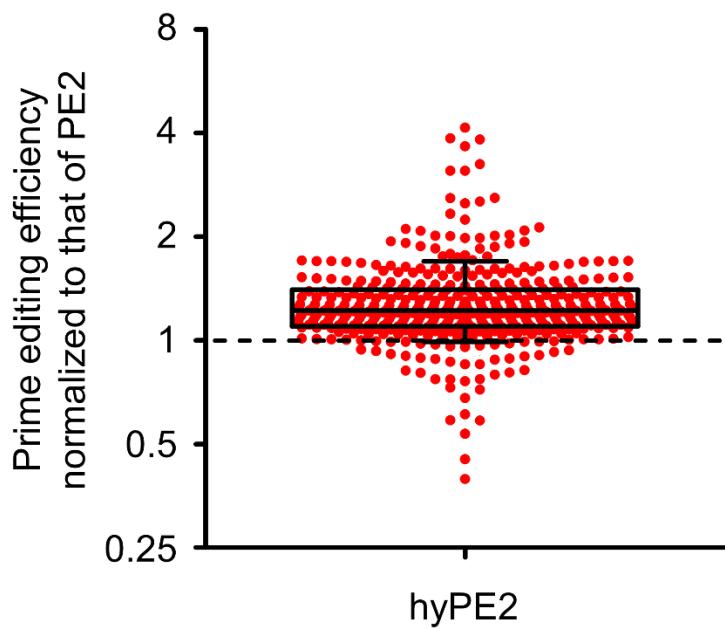
Supplementary Figure 1. Correlations between prime editing efficiencies from biological replicates for high-throughput evaluations of prime editing activities. The biological replicates were independently cultured, transfected with plasmids encoding prime editor 2 (variants), and analyzed. The Pearson correlation coefficients (r) and trend lines are shown. The number of analyzed target sequences $n = 83$. Source data are provided as a Source Data file.



Supplementary Figure 2. Prime editing efficiencies of PE2 variants normalized to the efficiency of PE2 at the same target sequences, which had been lentivirally integrated in HEK293T cells. Target sequences with PE2-induced prime editing efficiencies less than 1% are shown as gray dots. The number of target sequences $n = 64$. Data of minimum to maximum values are presented. For the boxes, the top, middle, and bottom lines represent the 25th, 50th, and 75th percentiles, respectively. The whiskers indicate the 10th and 90th percentile values. Source data are provided as a Source Data file.

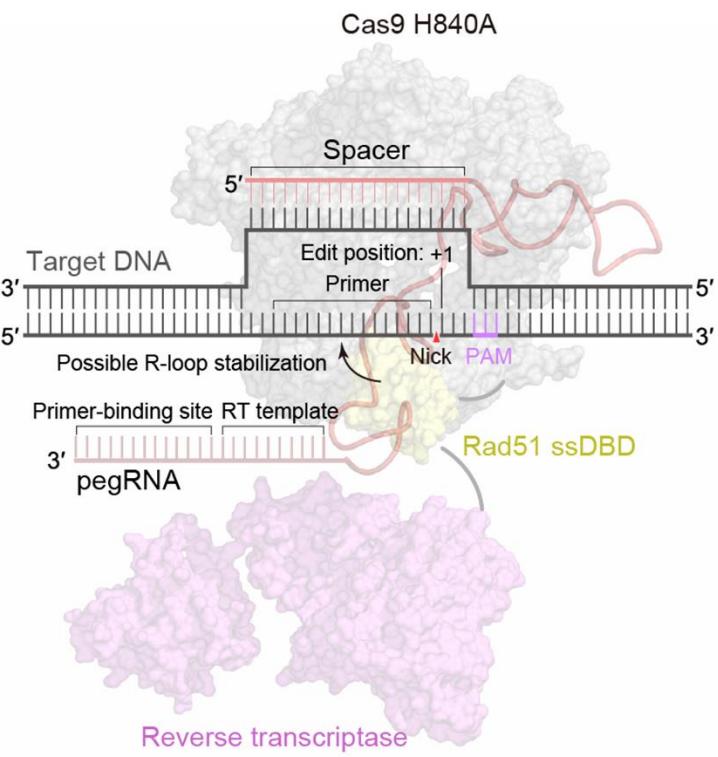


Supplementary Figure 3. Off-target effects of hyPE2 and PE2 at the potential off-target sites for pegRNAs 2, 3, and 5 and for four different pegRNAs targeting HEK4 in HEK293T cells. The intended editing positions are highlighted in yellow; mismatched and bulged nucleotides are indicated in red and blue lower-case fonts, respectively. Data points are overlaid. Data are means \pm S.D. for three independent biological replicates. on, on-target site; off, off-target site. Source data are provided as a Source Data file.

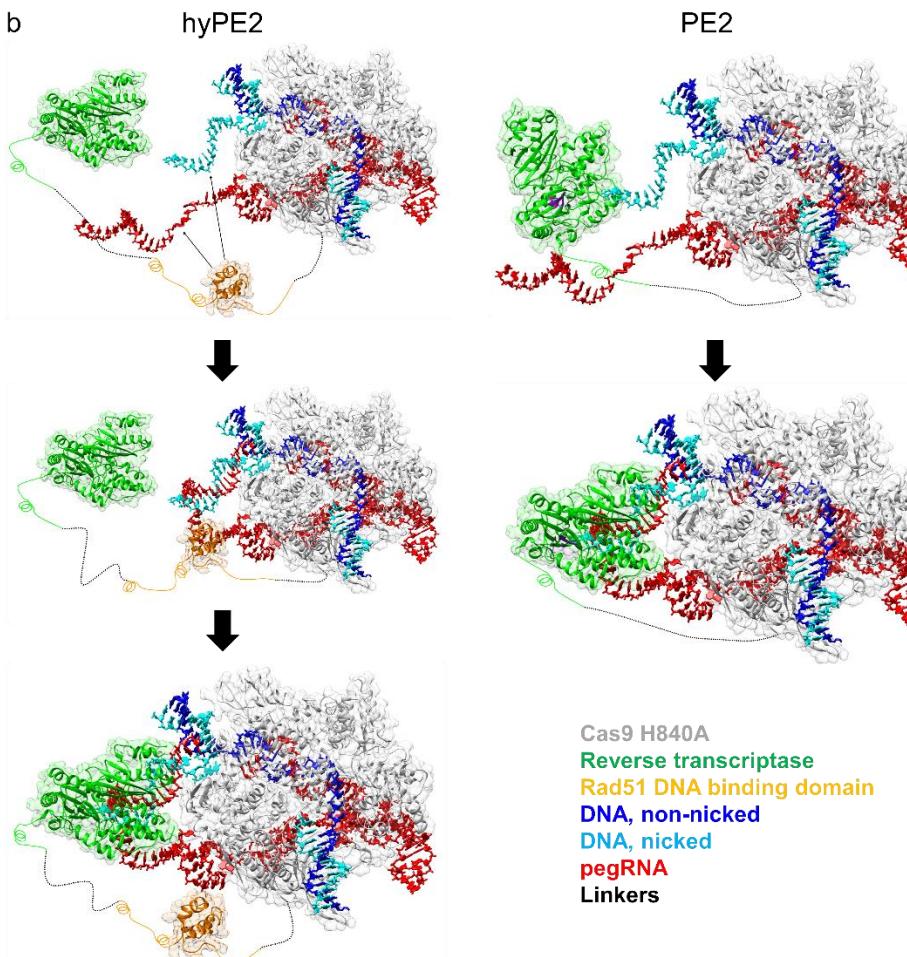


Supplementary Figure 4. Prime editing efficiencies of PE2 variants normalized to the efficiency of PE2 at the same target sequences, which had been lentivirally integrated in HEK293T cells. The number of target sequences $n = 423$. PegRNAs that resulted in PE2-directed prime editing efficiencies higher than 1% are shown. Data of minimum to maximum values are presented. For the boxes, the top, middle, and bottom lines represent the 25th, 50th, and 75th percentiles, respectively. The whiskers indicate the 10th and 90th percentile values. Source data are provided as a Source Data file.

a



b



Supplementary Figure 5. Structure of hyPE2. **a**, Schematic representation of hyPE2. **b**, Three-dimensional structures predicted for hyPE2 (left) and PE2 (right) before and after binding of the reverse transcriptase domain to the nicked target ssDNA/pegRNA hybrid. Modeling of a hypothetical interaction between Rad51, the nicked target ssDNA, and the pegRNA primer binding site is shown for hyPE2. Gray, Cas9 H840A; Green, Reverse transcriptase; Yellow, Rad51 DNA binding domain; Blue, DNA, non-nicked; Light blue, DNA, nicked; Red, pegRNA; Black, Linkers.

Supplementary Table 1. Characteristics (sequences of target, pegRNA spacer, and extension) of pegRNAs used in this study and relevant editing efficiencies (provided as a separate file).

Supplementary Table 2. Information about endogenous target sites and associated potential off-target sites (provided as a separate file). Mismatched and bulged nucleotides in the off-target sites are indicated in red and blue lower-case fonts, respectively. Mismatched nucleotides at the 5' end of target sequences that are 20-nt away from the PAM were not counted as mismatches.

Supplementary Table 3. Training and test datasets used for the development of PEselector (provided as a separate file).

Supplementary Table 4. Oligonucleotides used in this study (provided as a separate file).

Supplementary Note

PE2 (addgene # 132775)

NLS

Cas9 H840A

Linker A

Reverse transcriptase

atgaaaacggacagccgacggaagcgagttcgagtca**ccaaagaagaagcggaaagtgcacaagaagtacacg**
atccggcctggacatggcaccaactctgtgggctggccgtatcaccgacgagtacaagggtcccagcaagaaa
ttcaagggtctggcaacaccgaccggcacagcatcaagaagaacctgatcggagccctgcttgcacagcggc
gaaacagccgaggccacccggctgaagagaaccggcagaagaagatacaccagacggaagaaccggatctg
ctatctgcaagagatctcagcaacgagatggccaagggtggacagacttctccacagacttgcagacttcct
ggtggaaagaggataagaagcagcggcacccatctcggcaacatctggacgagggtggctaccacgaga
agtacccaccatctaccacctgagaaagaactggtgacagcaccgacaaggccgacctgcggctatctatct
ggccctggcccacatgtcaagttccccccacttcctgatcgagggcgacctgaaccccgacaacagcgacgt
ggacaagctttcatccagctggcagacactacaaccagctgtcgaggaaaacccatcaacgcgcggcgtg
gacgccaaggccatctgtccagactgagcaagagcagacggctggaaaatctgatcggccagctgcggcc
gagaagaagaatggcttgcggaaacctgattgccctgagcctggccctgaccccaactcaagagcaacttcg
acctggccgaggatgccaactgcagctgagcaaggacacctacgacgacgacctggacaacctgtggcc
atcgccgaccagtacgcccaccttgcggccaaagaaacctgtccgacgccatctgtctgagcgacatcctgag
agtgaacaccgagatccaaggccccctgagcgcttatgtcaagagatacgacgagcaccaccaggacc
tgaccctgctgaaagctctgtgcggcagcagctgcctgagaagtttgcaccagagcaagaac
ggctacgcggctacattgacggcgagccagccaggaagagttctacaagttcatcaagccatcctggaaaag
atggacggcaccgaggaactgctgtgaagctgaacagagaggacctgtcggaagcagcggacattcgacaa
cggcagcatccccaccagatccacctggagagactgcacgcattctgcggcggcaggaagatttacccattcc
tgaaggacaaccggaaaagatcgagaagatcctgacccatcccctactacgtggccctggcaggg
gaaacagcagattgcctggatgaccagaaaagagcggagaaaccatccccctggaaacttcgaggaagtgg
gacaaggcgctcccccagagcttcatcgagcggatgaccaacttcgataagaacctgcacgagaagg
ctgcccacagcacgcctgtacgacttcaccgtgtataacgagctgaccaactgtgaaatacgacgagg
gaatgagaaagcccccttcgtacggcgagcagaaaaaggccatctggacacttcgatgtcaagaccaaccgg
aaagtgaccgtgaagcagctgaaagaggactactcaagaaaatcgagtgttcgactccgtggaaatctccggc
tggaaagatcggtcaacgcctccctggcacataccacgtctgtaaaattatcaaggacaaggacttcctggac
aatgagaaaacgaggacattcttggaaagatatcgctgtgaccctgacactgtttgaggacagagatgtcagg
aacggctgaaaacctatgcccacctgtcgacgacaaagtgtatcgaaagcagctgaaacgcggcgagatacaccgg
ggggcaggctgagccggaaagctgtatcaacggcatccggacaaggcgtccggcaagacaatctggattcctga
agtccgacggctcgccaacagaaactcatgcagctgatccacgacgacgacgacccatcaaagggacatcca
gaaagcccaagggtccggccaggcgatagcctgcacgagcacattgccaatctggccggcagccccccattaa
gaagggcatctgcagacagtgaaggtggacgagctgtgaaaagtgtatggccggcacaagcccgagaac
atcgatcgaaatggccagagagaaccagaccaccagaagggacagaagaacagccgcagagaatgaa
gcggatcgaaagagggcatcaaagagctggcagccagatctgtgaaagaacacccctggaaaacaccggctg
cagaacgagaagctgtacctgtactacctgcagaatggcggatatgtacgtggaccaggaactggacatcaac
cggctgtccactacgatgtggacgctatcgtccctcagagcttctgtgaaaaggacgactccatcgacaaacaagg
accagaagcgacaagaaccggggcaagagcgacaaacgtgcctccgaagaggtcgtaagaagatgaaagaa
ctactggccggcagctgtgaacccaagctgattaccagagaaagttcgacaaatctgaccaaggccgagagagg
cgccctgagcgactggataaggccggctcatcaagagacagctggaaaaccggcagatcacaaggc

tggcacagatcctggactcccgatgaacactaagaactacgacgagaatgacaagctgatccggaaagtgaaagtga
tcaccctgaagtccaagctgggtccgattccggaaaggattccagtttacaaagtgcgcgagatcaacaactacca
ccacgcccacgcacgcctacctaacaacgcgcgtggaaaccgcctgatcaaaaagtaccctaagctggaaagcg
agttcgtgtacggcactacaagggtacgacgtgcggaaagatgatgcacaagagcgcagcaggaaatcgcaag
gctaccgccaagtacttctacagcaacatcatgaactttcaagaccgagattaccctggcaacggcggagatcc
ggaagcggcctctgatcgagacaaacggcaaaccgggagatcgtgtggataaggggcgggatttgccacc
gtgcggaaagtgtcagcatgccccaaagtgaatatcgtaaaaagaccgaggtgcagacaggcggctcagcaa
agagtctatcctgccaagaggaacagcgataagctgatgcgcagaaagaaggactgggaccctaagaagtagc
gcggctcgcacagccccaccgtggctattctgtgtggggccaaagtggaaaaggcaagttcaagaaact
gaagagtgtgaaagagactgtgtgggatcaccatcatgaaagaagcagcttcgagaagaatccatcgacttctg
gaagccaaggctacaaagaagtggaaaaggacactgatcatcaagctgcctaagtactccctgtcgagctggaa
aacggccgaaagagaatgctggctctgcggcgaactgcagaaggaaacgaactggccctccctccaaata
tgtgaacctcctgtacctggccagccactatgagaagctgaaagggtcccccggaggataatgagcagaaacagctg
ttgtggAACAGCACAAGCACTACCTGGACGAGATCATCGAGCTTCAGAGAGTGTCCCTGG
cgacgctaattggacaaagtgtgtccgcctacaacaaggccggataagcccatcagagagcaggccgaga
atatcatccacctgtttaccctgaccaatctggagccctgcccctcaagtacttgacaccaccatcgaccggaa
gaggtacaccagcaccaaaagagggtgtggacgccaccctgatccaccagagcatcaccggctgtacgagacac
ggatcgacctgtctcagctggaggtgactctggaggatctcgaggatctcgaggatcttcggcagcagacaccaggaaac
aagcgagtcaacaccagagagcagtggcggcagcagcggcggcagcagcaccctaaatataagaagatga
gtatcggtcatcatgagacactcaaaagagccagatgttctctgggtccatgggtctgtatccctcaggcctgg
cgggaaaccggggcatggactggcagttcgcctactggcactggggatcaagccccacatacagagactgtggaccag
tccataaaaacaatacccatgtcacaagaagccagactggggatcaagccccacatacagagactgtggaccag
ggaatactggtacccctggcagtccttggggatcaacagccgggtggaaagacatccacccaccgtgcccacccctacaac
cttgagcgggctcccaccgtcccaccgtggatctgtgtgtttttctgcctgagactccaccc
caccagtcaacttcgccttgccttggatctgtggggatccagagatggaaatctcaggacaattgacctggaccagact
cccacagggttcaaaaacagtcccaccctgttaatgaggcactgcacagagactccggatccac
acccagactgtacttcgcctactgtggatgacttactgtgtggccacttctgagactgccaacaaggta
tcggccctgtacaaacccttagggaaacctcggtatcgccctcgccaaagaaagccaaatttgcctgg
ggtaagtatctgggtatcttctaaaagagggtcagagatggctactgaggccagaaaagagactgtatgggg
cagcctactccgaagaccctcgacaactaaggagttccctagggaaaggcaggctctgcgccttcattccctgg
tttgcagaaatggcagccccctgtaccctctcaccaaaccggggactctgttaattggggccagaccaacaaaa
ggcctatcaagaaatcaagcaagcttctactgccttgcctggggatctgtttacttgccttgcact
ttgtcgacgagaagcaggctacgccaagggtgtccctacgcacaaaactggggacttgcgtcgccgggtggccta
cctgtccaaaaagcttagaccctgtggatcttgccttgcctacggatggtagcagccattccgtactga
caaaggatgcaggcaagctaaccatgggacagccactagtcatctggcccccattgcacttagaggcacttagtca
aacaaccccccggaccgtggcttccaacgcccggatgactcaactatcaggcctgtttggacacggaccgggtcc
agttcgacccgggtggtagccctgaaccggctacgctgtcccactgcctgagggctgcaacacaactgcctt
gatattccctggccgaagccacggaaaccggaccctacggatggggactccggcgtccctggccggatgg
ggtaacacggatggaaagcagtcttacaagaggacagcgttaaggcggggactgcgggtgaccacccggagaccga
ggtaatctggctaaagccctgcccggacatccgtcagcggctgactgtatagcactcaccctggccct
aagatggcagaaggtaagaagactaaatgttatactgtatagccgtatgccttgcactgcccataatccatggagaaat
atacagaaggcgtgggtggctcacatcagaaggcaaaagagatcaaaaataaagacgagatttggccctactaaa
agccctttctgcctccaaaagacttagcataatccattgtccaggacatcaaaaggacacagcggccggaggctaga
ggcaaccggatggctgaccaagcggccggaaaggcagccatcacagagactccagacaccccttacccctccat
agaaaattcatcaccctctggcggctcaaaaagaaccggccacggcagcgaattcgagcccaagaagaagagg

aaagtctaa

PE2-mid_RPA70

NLS

Cas9 H840A

Linker A

Linker B

RPA70 ssDBD

Reverse transcriptase

atgaaaacggacagccacggaagcgagttcgagtcaccaaagaagaagcggaaagtcgacaagaagtacagc
atccggcctgcacatcgccaccaactctgtggctggccgtatcaccgacgagtacaaggtgcccagcaagaaaa
ttcaaggtgtggcaacaccgaccggcacagcatcaagaagaacctgatcggagccctgctgtcgacagcggc
gaaacagccgaggccacccggctaagagaacccgccaagaagaagatacaccagacggaagaaccggatctg
ctatctgcaagagatctcagcaacgagatggcaaggttgacgacagcttccacagactggaaaggatcttcc
ggtggaagaggataagaagcacgagcggcacccatctcggcaacatcgggacgaggtgcctaccacgaga
agtaccccaccattaccactgagaaagaactggtgacgacgaccgacaggccgactcggcttatct
ggccctggcccacatgatcaagtccggggccactcctgatcggagggcgactgaaccccacagcggactg
ggacaagcttctatcccagctggtcgaggactacaaccgactgttcgaggaaaacccatcacggccgaggggctg
gacgccaaggccatcctgtgccgactgaggacagaggacgggctggaaaatctgatcggccgactgcccggc
gagaagaagaatggcttcggaaacctgattgcctgaggctggcctgaccccaacttcaaggaggcaacttcg
actggccgaggatgccaaacttcggactgaggcaagggacgacctacgacggactggcccag
atcggcgaccgactgccgactgtttgggccgccaagaacctgtccgacgccatctgctgaggcggactctgag
agtgaacaccgagatgccaaggccccctgaggcggcttatgatgcaaggagatacgacgggaccgaccgagggacc
tgaccctgtaagacttctgtggcggcgaggcgacttcggctgggaccgaggcaagagac
ggctacgccggctacatgacgggggcaggccgagagaggatttctgacagccgatctgaaagag
atggacggcaccggcgaactgctgtaagctgacgaggagggacgctgcgggagacgttcggacgaagaa
cggcgactcccccaccgatccacctgggggaggactgcggccgatctgcggccgaggattttacccattc
tgaaggacaaccggaaaagatgagagatcctgacgcctgatcctactgttggccctgccgaggg
gaaacgaggattgcctggatgaccgaaaaggaggcggagggaaaccatgcctttgggactgtggagggatgtggtgg
gacaagggcggttccgccgccgaggactgttctgacgcgggatgacgactgtcgatagaagacgccgaagatgtggtggtggtgg
ctgccgcagacgcctgctgactgtgcactgttcctaccgtgttaacggcactgtggaccgaggg
gaatgaggaaggcccgccgttccgtgaggcgggcaggaaaaaagggccatgtgggcactgtggttccagacgacgccggc
aaagtgaccgtggaagggactgactgcaaaaatgggatgtggactgccgtggaaatgtccggg
tggaaggatgggttcacgcgccgccgacgatgaccgatgctgaaaattatgcaagggacaagggactgcgtggac
aatgaggaaaacgagggacgattgtggaagatgatgtggctggcaccgactgtttggagggacgaggaggatgtggaggg
aacgggctgaaaacgctgatgcccgaccgtgtcgacgggaaaatgggatgacgccgggggatgacgccggg
ggggcgagggtggccgggacgactggtgatgcaacgcggcatgccggggacagacagcgtgggatgttcctgga
agtgccgacggggctgccgcagacgaaaactgcatgtggactgggatgccgacgtggacgatgccgatgca
gaaaggccgagggtggccgggacgatgccgtggacgggacgatgtggccggggacagggccgccgattgaa
gaagggggatgccgtggacgggactgtggagggtggggacgggactgctggtgaaagatggggagggacgatgccggg
atgctgtggaaatggggccgggaggagagaccgaccgccgggacagggacgccgggggacagggatgtgggggaagatgtggtgg
gcggggatgcggaagggactgtggttcactgtggactgccgtggacgggacagacgccgggggatgctggg
cagaacgggacagggactgtggttcactgtggactgccgtggacgggacagacgccgggggacagacgccggg
cggggctgtggccgactgacgtggtggacgggctgtggactgccgatgcggacagacagggatgtggctggg
accgggacagggacagaccggggggcagaggacgggacagacgccgggggacagacgccggg

ctactggcggcagctgtaaccccaagctgattaccaggagaaaagttcgacaatctgaccaaggcccagagagg
cggcctgagcgaactggataaggccggctcatcaagagacagctggtgaaaccgcgcagatcacaaagcac
tggcacagatcctggactcccgatgaacactaagtacgacgagaatgacaagctgatccggaaagtgaaagtga
tcacctgaagtccaagctgggtccgattccggaggattccagttacaatgacgacgatcaacaactacca
ccacgcccacgcacgcctacctgaacgcgcgtggaccgcctgatcaaaaagtaccctaagctggaaagcg
agttcggtacggcgactacaagggtacgcacgtcgaaagatgatgcacaagagcgcagcaggaaatcgcaag
gctaccgccaagtacttctacagcaacatcatgaactttcaagaccgagattaccctggcaacggcgagatcc
ggaaggccgcctctgatcgagacaaaacggcggaaaccggggagatcggtggataagggccggatttgcacc
gtcgaaaagtgtgagcatgccccaaagtgaatatcgtaaaaaagaccgagggtgcagacaggcggttcagcaa
agagtatctgcccagaggaacagcgataagctgatgcacaaagaaggactggaccctaagaagtacg
gcggcttcgacagccccaccgtggcttctgtgtggccaaagtggaaaaggcaagtccaaact
gaagagtgtgaaagagactgctggggatcaccatcatgaaagaagcagacttcgagaagaatccatcgacttctg
gaagccaagggtacaaagaagtgaaaaaggacactcatcaagctgcctaagtactccctgtcgagctggaa
aacggccggaagagaatgctggctctgcccgaactgcagaaggaaacgaactggccctgcccctccaaata
tgtgaacttccgtacctggccagccactatgagaagctgaaaggctcccccggggataatgagcagaaacagctg
ttgtggacacgcacaaggactacccgtggacgagatcatcgagcagatcagcggatctccaagagagatgatctggc
cgacgctaactggacaaaagtgtccgcctacaacaaggccaccggataagcccatcagagagcaggccgaga
atcatccacccgtttaccctgaccaatctggggccctggccctcaagtacttgcacaccatcgaccggaa
gaggtacaccagcaccaaagagggtctggacgcccaccctgatccaccagagcatcaccggctgtacgagacac
ggatcgaccgtctcagctggagggtgac **TCTGGTGGTTCTCCCAAGAAGAAGAGGAAAGTC**
GGTAGTTCCGGATCT ggagggagtaacaccaactggaaaaccttgtatgaggtaatccgagaacctg
ggccaaggcgacaagccggactacttagttctgtggccacagtgggttatctcgcaaagagaactcgatgtaccaa
gcctgcccactcaggactgcaataagaaagtgtatgtatcaacagaatggattgtaccgtgtgaaagtgcaca
ccgaattcccaattcaagtaccgcatgtctgtcgtcgtaaatattgcagatccaaatgacgatgtgggtgacttgc
ccaggagtctgtgaaactatccgtggacaaaatgctgttatctggggaaattaaagacaagaatgacgatgttgc
gaagaagtttccagaatgccaactccgatttcatattcagagtcaagggtcaaagtggagacactacaacgac
tctcgaattaaggccactgtgatggacgtgaagccgtggactacagagactatggccgaaggctggcatgagcat
caggagaagtgcattgtatgtctggaggatctaggcggaggatctctggcagcggagacaccaggaaacgcgatc
agcaacaccagagcagtggccgcagcagccggcggcagcaccctaaatataagaagatgatgtatccgcta
catgagacccaaaagagccagatgttctctgggtccacatggctgtgatccctcaggcctggccggaaaccc
ggggcatggactggcagttcccaagctccctgtatcataccctctgaaagcaacccctacccctgtccataaaac
aatccccatgtcacaagaagccagactggggatcaagccccacatacagagactgtggaccaggaaatactgg
taccctgccagtccccctggAACACGCCCTGTACCCGTTAAGAAACCAGGGACTAATGATTAGGCCTGTCCAG
GATCTGAGAGAAGTCAACAAGCGGGTGGAAAGACATCCACCCACCGTGCCTAACCCCTTACAACCTCTGAGCGGG
CTCCCTCCACCGTGGTACACTGTGCTGATTAAAGGATGCCCTTCTGCTGAGACTCCACCCACCGTCA
CCTCTCTGCCCTTGAGTGGAGAGATCCAGAGATGGGAATCTCAGGACAATTGACCTGGACCAGACTCCCACAGGGT
TTCAAAAACAGTCCCACCCCTTTAATGAGGCACTGCACAGAGACCTAGCAGACTTCCGGATCCAGCACCCAGACTT
GATCCGTACAGTACGTGGATGACTTACTGCTGGCCGCACCTCTGAGACTGCCAACAAGGTACTCGGGCCCTG
TTACAAACCTAGGAAACCTCGGGTATCGGGCTCGCCAAGAAAGCCAAATTGCCAGAAACAGGTCAAGTATC
TGGGGTATCTCTAAAAGAGGGTCAAGAGATGGCTGACTGAGGCCAGAAAAGAGAGACTGTGATGGGGCAGCCTACTCC
GAAGACCCCTCGACAACTAAGGGAGITCCCTAGGGAAGGCAGGCTCTGCGCTCTCATCCCTGGGTTGAGAAAT
GGCAGCCCCCTGTACCCCTCTCACCAACCGGGGACTCTTAAATGGGGCCAGACCAACAAAAGGCCTATCAA
GAAATCAAGCAAGCTCTCAACTGCCAGCCCTGGGTTGCCAGATTGACTAAGCCCTTGAACCTTGTGACGA
GAAGCAGGGCTACGCCAAAGGTGTCTAACGCAAAAACGGGACCTGGCGTGGCCGGTGGCCTACCTGTCCAA
AAAGCTAGACCCAGTAGCAGCTGGGTGGCCCCCTGCCACCGTGGTAGCAGCCATTGCCGTACTGACAAAGGAT
GCAGGCAAGCTAACCATGGACAGCCACTAGTCATTCTGGCCCCCTGCCACCGTGGTAGCAGCTAAACAACCC

cccgaccgctggcttccaacgcccggatgactcactatcaggccttgctttggacacggaccgggtccagttcgg
ccggtgttagccctgaacccggotacgctgctcccactgcctgaggaagggtgcaacacaactgcctgataatcct
ggccgaagcccacggaacccgacccgacctaacggaccagccgctccagacgcccaccacac
ggatggaagcagtcttacaagagggacagcgttaaggcggagctgcggaccaccgagaccgaggtaatct
ggctaaagccctgccagccggacatccgctcagcggctgaactgatagcactcaccaggccctaaagatgg
cagaaggtaagaagctaaatgtttatactgatagccgtatgctttgctactgcccataatccatggagaaaatatacaga
aggcgtgggtggctcacatcagaaggcaaagagataaaaataaagacgagatctggccctactaaaagccctc
ttctgcccataaaagacttagcataatccattgtccaggacatcaaaaggacacagcggccgaggctagaggcaacc
ggatggctgaccaaggccccgaaaggcagccatcacagagactccagacacaccttacccctcatagaaaaatt
catcaccctctggcggctcaaaaagaaccggcagcgaattcgagcccaagaagaagaggaaagtcta
a

hyPE2 = PE2-mid_Rad51

NLS

Cas9 H840A

Linker A

Linker B

Rad51 ssDBD

Reverse transcriptase

ataaaacggacagccgacggaagcgagttcgagtcacccaaaagaagaagcggaaagtcgacaagaagtacagc
atcggcctggacatcgccaccaactctgtgggctggccgtgtcaccgacgagtcacagggtgccagcaagaaaa
ttcaagggtgtggcaacaccgaccggcacagcatcaagaagaacctgtatcgaggccctgtgtcgacagcggc
gaaacagccgaggccacccggctgaagagaaccggcagaagaagatacaccagacggaagaaccggatctg
ctatctcaagagatctcagcaacgagatggcaagggtggacgacagcttccacagactggaagagtcctcc
ggtggaaagaggataagaaggcacgagcggcacccatctcggcaacatcggtggacgagggtggctaccacgaga
agtacccaccatctaccacctgagaaagaaaactggtagacagcaccgacaaggccgacctgcggctgttatct
ggccctggcccacatgtatcaagttccggggcacttcctgtatcgagggcgacctgaaccccgacaacagcgcacgt
ggacaagcttcatccagctggcagacactacaaccagctgtcgaggaaaacccatcaacgcccagcggcgt
gacgccaaggccatctgtccagactgagcaagagcagacggctggaaaatctgtatcgcccagctgcccgg
gagaagaagaatggctttcgaaacctgttgcggcttgcggccatccatctgtatcgcccagctgcccgg
acccggccgaggatgccaactgcagctgagcaaggacacctacgacgacgacctgtggacaacctgtggcc
atcgccgaccagtacgcccacctgttctggccccaagaacctgtccgacgccatctgtatcgacatcctgag
agtgaacaccgagatccaaggccccctgagcgcctatgtatcaagagatacgacgagcaccaccaggacc
tgaccctgtgaaagctctgtcgccagcagctgcctgagaagtacaagagatttctcgaccagagcaagaac
ggctacgcccgtacattgacggcgagccagccaggaaagatctacaagttcatcaagccatctggaaaag
atggacggcaccgaggaactgctgtgtgttt
ccgc
tgaaggacaaccggaaaagatcgagaagatcctgacccctgcgttgcgttgcgttgcgttgcgttgcgtt
gaaacagcagattcgccctggatgaccagaaaagagcggagggaaaccatcaccctggaaacttcgaggg
gacaaggcgctcccccagagcttcatcgagcggatgaccaacttcgataagaacctgcccacgagaagg
ctgcccacacgcgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt
gaatgagaaagccgccttcctgagcggcgagcagaaaaaggccatcggtggacctgtgttcaagaccaacc
aaagtgaccgtgaagcagacttcaagaaaatcgagtgctcgactccgtggaaatctccgc

tggaaagatcggttcaacgcctccctggcacataccacgatctgctgaaaattatcaaggacaaggactcctggac
aatgaggaaaaacgaggacattctggaagatatcgctgaccctgacactgtttgaggacagagatgatcgagg
aaccgctgaaaacctatgcccacctgtcgacgacaaagtgtatcgaaagcagctgaacgcggcggagatacaccggct
ggggcaggctgagccggaagctgatcaacggcatccggacaagcagtcggcaagacaatcctggattcctga
agtccgacggctcgccaacagaaaacttcatcgacgtatccacgacgacagcctgacccttaaagaggacatcca
gaaagcccagggtgtccggccaggcgatagcctgcacgagcacattgccaatctggccggcagccccgccattaa
gaagggcatcctgcagacagtgaaggtggacgagctcgtaaaagtgtatggccggcacaagccccgagaac
atcgatcgaaatggccagagagaaccagaccacccagaagggacagaagaacagccgcagagaatgaa
gccccatcgaaagagggcatcaaagagactggcagccagatctgaaagaacacccccgtggaaaacacccagctg
cagaacgagaagctgtacctgtactacctgcagaatggccggatgtacgtggaccaggaactggacatcaac
cggtctccactacgatgtggacgctatcgctcagagcttctgaaggacgactccatcgacaacaaggctg
accagaagcgacaagaaccggggcaagagcgacaacacgtgccctccgaagaggctgtaagaagatgaa
ctactggccgcagctgtaacgccaagctgattaccagagaaaagttcgacaatctgaccaaggccgagagagg
cggtctgagcgaactggataaggccggctcatcaagagacagctggtgaaaacccggcagatcacaagcac
tggcacagatcctggactcccgatgaacactaagtacgacgagaatgacaagctgatccggaaagtgaaagtga
tcacctgaagtccaagctgggtccgattccggaggattccagtttacaatgtcgagatcaacaactacca
ccacgcccacgacgcctacctaagccgtcggtggaaaccgcctgatcaaaaagtaccctaagctggaaagcg
agttcggtacggcgactacaagggtacgacgtcgaaagatgatcgccaaagagcgagcaggaaatcgcaag
gctaccgccaagtacttctacagcaacatcatgaaactttcaagaccgagattaccctggcaacggcgagatcc
ggaagcggcctctgatcgagacaaaacggcggaaaccggggagatcggtggataagggccggatttgccacc
gtcgaaagtgtgagcatgccccaaagtgaaatatcgtaaaaagaccgaggtgcagacaggcggctcagcaa
agagtctatctgcccagaggaacagcgataagctgatcgccagaaagaaggactggaccctaagaagtacg
cggtctcgacagccccaccgtggctattctgtcggtggccaaagtggaaaaggcaagtccaaact
gaagagtgtgaaagagactgctgggatcaccatcatgaaagaagcagcgttcgagaagaatccatcgacttctg
gaagccaaggctacaagaagtgaaaaggacgactcatcaagctgcctaaagtactccctgtcgagctggaa
aacggccggaagagaatgctggctctgcccggcaactgcagaaggaaacgaactggccctgcccctccaaata
tgtgaacttccgtacctggccagccactatgagaagctgtaaggctcccccggataatgagcagaaacagctg
tttgtgaacacgacaagcactacctggacgagatcatcgagcagatcagcgagttctccaagagagatgatctggc
cgacgctaactggacaaaagtgtcgccctacaacaaggccggataagcccatcagagagcaggccgaga
atatcatccacctgtttaccctgaccaatctggagccctgcccctcaagtacttgcacaccatcgaccggaa
gaggtacaccagacccaaagagggtctggacgcccaccctgatccaccagagcatcaccggctgtacgagacac
ggatcgaccctgtctcagctggaggtgac TCTGGTGGTTCTCCCAAGAAGAAGAGGAAAGTC
GGTAGTTCCGGATCT gcaatgcagatgcagctgaagcaaattcgatctcgatggaaagaagaaagct
ttggcccaacccatttcacggtagagcagtgatcgatcgatggccataatgcacqatgtgaaagaaattggaaagaagctgga
ttccatactgtggaggctgtgcctatgcgc当地ggactaaaggagactaaataatattaaggaaattgtgaaagccaaagct
gataaaattctggctgaggcagctaaatttagtccaatggttcaccactgcaactgaaatccaccaaaaggcggctca
gagatcatacagattactactggctccaaagagactgatcgatctgacaaactactcaatctggaggatctcgaggatct
ggcagcgcgagacaccaggaacaagcgagtcagcaacaccagagcagtgccggcagcggcggcagca
gcaccctaaatataagaagatgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgatcg
ctgtctgatccctcaggcctggccggaaaccggggcatggactggcagttcgccaaatcgatcgatcgatcgatcgatcg
gaaagcaaccttaccccggtccataaaacaataccctgtcacaagaagccagactggggatcaagccca
catacagagactgtggaccagggaaatctggatccctgcccgtccatggacacgcggccctgctaccctgtaaga
aaccaggactaatgattataggcctgtccaggatctgagagaagtcacaacaagcgggtggaaagacatccacc
ccgtgc当地ggccaaacccttacaacccttgcagcggctccaccgtccaccgtggatctgtgatccctgatcg
ctttctgcctgagactccaccctaccgtcagcccttgccttgccttgccttgccttgccttgccttgccttgccttgccttgc
gacaattgacctggaccagactcccacagggtttcaaaaacagttccaccctgtttaatgaggactgcacagaga

cctagcagacttccggatccaggcaccccagacttgtatccgtacagtacgtggatgacttaactgtggccgcactct
gagctagactgccaacaaggtaactcgccctgttacaaaaccctagggAACCTCGGGTATCGGGCCTCGGCCAAG
aaAGCCCAATTGCCAGAAACAGGTCAAGTATCTGGGTATCTCTAAAAGAGGGTCAGAGATGGCTGACTGAGGC
CAGAAAAGAGACTGTGATGGGCAGCCTACTCCGAAGACCCCTGACAACTAAGGGAGTTCCTAGGAAAGGCAG
GCTTCTGTCGCCCTCTCATCCCTGGGTTGAGAAAATGGCAGCCCCCTGTACCCCTCACCAAACCGGGACTCTGTTA
ATGGGGCCAGACCAACAAAAGGCCTATCAAGAAAATCAAGCAAGCTCTAACTGCCCCAGCCCTGGGGTTGCC
AGATTGACTAAGCCCTTGAACTCTTGTGACGAGAAGCAGGGCTACGCCAAAGGTGTCTAACGCAAAAACCTGG
GACCTGGCGTCGGCCGGTGGCCTACCTGTCCAAAAGCTAGACCCAGTAGCAGCTGGGTGGCCCTTGCTACGG
ATGGTAGCAGCCATTGCCGTACTGACAAAGGGATGCAGGCAAGCTAACCATGGACAGCCACTAGTCATTCTGGCC
CCCATGCACTAGAGGCAGTAGTCAAACAAACCCCCCGACCCTGGCTTCCAACGCCGGATGACTCCTATCAGGC
CTGCTTTGGACACGGACCGGGTCCAGTTCGACCGGTGGTAGCCCTGAACCCGGTACGCTGCTCCCCTGCCTGA
GGAAGGGCTGCAACACAACCTGCTGATACTCTGGCCGAAGCCACCGAACCCGACCTAACGGACCCAG
CGCTCCAGACGCCGACCACACCTGGTACACGGATGGAAGCAGTCTCTTACAAGAGGGACAGCGTAAGGCCGA
GOTGCGGTGACCACCGAGACCGAGGTAATCTGGCTAAAGCCCTGCCAGCCGGACATCCGCTAGCGGGCTGAA
CTGATAGCACTCACCCAGGCCCTAAAGATGGCAGAAGGTAAGAAGCTAAATGTTATACTGATAGCCGTTGCTTTG
CTACTGCCCATATCCATGGAGAAATATACTAGAAGCGTGGGTGGCTCACATCAGAACGGCAAAGAGATCAAATA
AAAGACGAGATCTGGCCCTACTAAAAGCCCTTCTGCCAAAAGACTAGCATAATCCATTGTCAGGACATCAA
AGGGACACAGCGCCGAGGCTAGAGGCAACCGGATGGCTGACCAAGCGGCCGAAAGGCAGCCATCACAGAG
ACTCCAGACACCTCTACCCCTCATAGAAAATTCTACCCCTGGCGCTAAAAAGAACCGCCGACGGCAGCG
AATTCTGAGCCAAGAAGAAGAGGAAAGTCAA

PE2-N_Rad51

NLS

Cas9 H840A

Linker A

Linker B

Rad51 ssDBD

Reverse transcriptase

atgaaaacggacagccgacggaagcgagttcgagtca**ACCCAAAAGAAGAAGCGGAAAGTCGCAATCGAGATCGAG**
CTTGAAGCAAATCGAGATACTTCAGTGGAGAAAGAAAGCTTGGCCACAAACCCATTCACGGTTAGAGCAGTGTGG
CATAAAATGCCAACGATGTGAAGAAAATTGGAAGAAGCTGGATTCCATACTGTGGAGGCTGTGCCTATGCCAAAGA
AGGAGCTAATAAATATTAGGGAAATTAGTGAAGCCAAGCTGATAAAATTCTGGCTGAGGCAGCTAAATTAGTCCAA
TGGGTTTACCAACTGCAACTGAATTCCACCAAAAGGCAGGCTAGAGATCATACTAGATTACTGCTCCAAGAGCTG
ACAAACTACTCAATCTGGTGGTTCTCCCAAGAAGAAGAGGAAAGTCGGTAGTTCCGG****
ATCTgacaagaagtacagcatggcctggacatcgccaccaactctgtggctggccgtatcaccgacagtg
acaagggtgcccgaccaagaaattcaagggtgtggcaacaccgaccggcacagcatcaagaagaacctgatcg
agccctgttcgacagcggcggaaacagccgaggccacccggctgaagagaaccggcagaagaagatacacc
agacggaagaaccggatctgttatctgcaagagatcttcagcaacgagatggcaagggtggacagacagttctcc
acagactggaagagtccctctgggaagaggataagaagcagcagcggcacccatctcggcaacatctgtgg
acgaggtggcctaccacgagaagtacccaccatctaccacctgagaagaaaactggtgacagcaccgacaa
ggccgacctcgccgtatctgtggccctggccacatgtcaagttccggggccacttctgtatcgagggcggact
gaaccccgacaacacagcgacgtggacaagctgtttcatccagctggcagacactacaaccagctgtcgaggaaaa
ccccatcaacgcccagccggcgtggacgccaaggccatctgtccagactgagcaagagcagacggctggaaaa

actgtatcgcccagctgccccggcagaagaagaatggcctgtcgaaacctgattgcctgagccctggcctgacc
cccaacttcaagagcaactcgacacctggccgaggatgccaaactgcagctgagcaaggacacacgacgacga
cctggacaacctgctggcccagatcgccgaccagtacgcccacctgtggcccaagaacacctgtccgacgcc
atccctgctgagcgacatcctgagagtgaacaccgagatccaaggccccctgagcgccttatgtatcaagagat
acgacgagcaccaccaggacacctgctgaaagctctcgccgagcagctgcctgagaagttacaagat
atttcctcgaccagagaagaacccgtacgcccgtacattgacggccgagccaggcaggaagatctacaagtt
catcaagccatcctgaaaagatggacggcaccgaggaactgctgtgaagctgaacagagaggacctgctgc
ggaaggcagccgacccgtgacaacccgagcatccccaccagatccacctggagagctgcacccattctgccc
cgccggcaggaagatttaccatcctgaaaggacaaccggaaaagatcgagaagatccgtacccatcccata
ctacgtggccctctggccagggaaaacagcagattccctggatgaccagaaaagagcgaggaaaccatcaccc
cctggaaacttcgaggaagtggtgacaaggcgctcccccagagcttcatcgagcggatgaccaacttcgataa
gaacctgccaacgagaagggtgtcccaagcacagccgtgtacgagttacccgtataacgagctgacc
aaagtaaaatcgtgaccgagggaaatgagaaaagccgccttcctgagcggcagcagaaaaaggccatctgg
acctgctgttcaagaccaaccggaaaagtgaccgtgaagcagctgaaagaggactacttcaagaaaatcgagtgc
cgactccgtggaaatctccggcgtgaaagatcggtcaacgcctccctggcacataccacgatctgtgaaaattat
caaggacaaggacttcctggacaatgagggaaaacgaggacattctgaaagatatcgctgaccctgacacttgtt
aggacagagagatgtcgaggaacccgtgaaaacctatgcccacctgttcgacgacaaagtgtatgaaagcagctg
aagccgcgagatacaccgcgtggggcaggctgagccgaaagctgatcaacggcatccggacaaggcgtcc
gcaagacaatcctggattcctgaaagtccgacggcgtcccaacagaaaacttcatcgagctgatccacgacgacgc
ctgacctttaaagaggacatccagaaagcccgatgtccggccaggcgatagcctgacgacatggccaaat
ctggccggcagcccgccattaagaagggcatcctgcagacagtgaaggtggacgagctgtaaagtgtatg
ggccggcacaagcccgagaacatcgtgatcgaaatggccagagagaaccagaccaccagaagggacagaa
gaacagccgcgagagaatgaagcggatcgaaagagggcatcaaagagactggcagccagatccgtaaagaaca
cccccgtggaaaacaccagctgcagaacgcgagaagctgtacctgtactacccgtcagaatggcgggatgtacgt
gaccaggaactggacatcaaccggctgtccgactacgatgtggacgctatcgccctcagagcttctgaaaggacga
ctccatcgacaacaagggtgtgaccagaagcgacaagaaccggggcaagagcgacaacgtgcctccgaaaga
ggcgtgaaagatgaaactactggccgcagctgtgaaacgcctgattaccctgagaaaagttcgacaa
tctgaccaaggccgagagaggccctgagcgaaactggataaggccgctcatcaagagacagctgggaaa
cccgccagatcacaaggcgtggcacagatccctggactcccgatgaaactaagtacgacgagaatgacaag
ctgatccggaaagtgtatcaccctgaaagtccaaagctgggtccgatcccgaaaggattcccgatccatca
tgccgagatcaacaactaccaccacgcccacgacgcctacctgaaacgcgtgtggaaaccgcctgatcaaaa
agtaccctaagctggaaagcgagttcggtacggcactacaagggtgtacgacgtgcggaaagatgtaccc
gcgagcagggaaatcgcaaggctaccgcaagtacttcttacagcaacatcatgaaactttcaagaccc
ccctggccaaacggcgagatccggaaagcgccctgtatcgagacaaaacggcgaaaccggggagatgtgtgg
aaggccggatttgcaccgtcgaaaaagtgtgacgatgccccaaagtgtaaatatcgaaaaagaccgagg
cagacaggccgtcagaaagagtctatctgtcccaagagaaacagcgataagctgatgcctgagaaaagg
actgggaccctaagaagtacggccgttcgacagccccaccgtggctattctgtgtgggtggccaaagtgg
aaggcaagtccaaagaaactgaagagtgaaagagactgtgtggatcaccatcatgaaagaagcagctcg
gaagaatccatcgacttctgaaagccaaaggctacaaagaagtgtaaaaaggacctgatcatcaagctgc
gtactccctgtcgagctggaaaacggccggaaagagaatgtggctctgccccgcaactgcagaaggaaacga
actggccctgcctccaaatatgtgaaactccctgtacccgtccagccactatgagaagctgaaaggctccccc
ataatgagcagaaacagctttggaaacagcacaagcactacccgtggacgagatcatcgagcagatcg
ctccaagagagtgatccctggccgacgctaattggacaaaagtgtccgcctacaacaaggccaccgg
atcagagagcaggccgagaatcatccacccgttaccctgaccaatctggagcccccgtccctcaagtactt
acaccaccatcgaccggaaagaggtacaccagcacccaaagaggtgtggacgcccaccctgatcc
caccggcctgtacgagacacggatcgacccgtctcagctggaggtgactctggaggatctg
cggaggatctg

gcagcgagacaccaggaacaagcgagtcaagcaacacccagagagcagtggcggcagcagcggcggcagcag
caccctaaatatagaagatgagtatcggtacatgagacactcaaagagccagatgttctctagggccacatggct
gtctgatttcctcaggcctggcgaaaccggggcatggactggcagttccaagactcctctgatcataccctg
aaagcaacctctacccctgtccataaaacaatacccatgtcacaagaagccagactgggatcaagccccac
atacagagactgtggaccaggaaactgttccactgtccaggatctgagagaagtcacaacaagcgggtggaagacatccacccac
cgtgcccaaccctacaacctcttgagcgggctccaccgtcccaccgtggtacactgtgctgattaaaggatgcc
tttctgcctgagactcoaccccaccagtcaaccccttcgccttgagtggagagatccagagatggaatctcagg
acaattgacctggaccagactcccacagggttcaaaaacagtcccaccctgttaatgaggcactgcacagagacc
tagcagacttccggatccaggccactgttccactgttccactgtggatgacttactgtgctggccacttctga
gctagactgccaacaaggtaactcggcccttttacaaacccttaggaaacctcgggtatcggcctcggccaagaa
agcccaaatttgcagaaacaggtaagtatctgggtatcttctaaaaagagggtaagatggctgactgaggcca
gaaaagagactgtgatggggcagcctactcgaagacccctcgacaaactaaggagttccttagggaaaggcaggc
tttgcgccttcatccctgggttgcagaaatggcagccccctgtaccctctcaccaaaccgggactctgttaatt
ggggcccaagccaacaaaaggcttatcaagaaatcaagcaagcttctaactgccccagccctgggttgccag
atttgcactaagcccttgaactcttgcacgagaagcaggctacgccaagggtgtcttaacgcaaaaactggac
cttgcgtccggcgtggcctaccgtccaaaaagcttagaccctgtcagctgggtggccctgcctacggatg
gtacgcatttgcgtactgacaaaggatgcaggcaagcttaccatggacagccactgttgcatttgccttgc
atgcagttagggacttagtcaaacaacccccccaccgtggctttcaacgcgggatgactcactatcaggcctg
ctttggacacggaccgggtccagttcggaccgggttagccctgaacccggctacgctgtccactgcctgagga
agggctgcaacacacaactgcctgatattcctggcgaagcccacggaacccgacccgacctaacggaccagccgct
cccgacgcggaccacaccgttacacggatgaaacggcttccatcgccatcgcttgcggacttgcggact
cggtgaccaccggagaccggaggttaatctggctaaaggccctggccacgggacatccgctcggactgactga
tagcactcaccctggccctaaagatggcagaaggtaaagaagactaaatgttatactgtatggcgtatgctttgctact
gcccattatccatgggagaatacacaaggctgggtggctcacatcagaaggcaaaagagatcaaaaataaaga
cgagatctggccctactaaaaggccctttctggccaaaagacttagcataatccattgtccaggacatcaaaagg
acacagcggcggaggttagggcaaccggatggctgaccaagcggccgaaaggcagccatcacagagactcc
agacaccttaccctcctcatagaaaattcatcaccctctggcggctaaaaagaaccggcagggcagcgaattc
gagcccaagaagaagaggaaagtctaa

PE2-C_Rad51

NLS
Cas9 H840A
Linker A
Linker B
Rad51 ssDBD
Reverse transcriptase

atgaaaacggacagccacggaagcgagttcgagtccaaagaagaagcgaaagtgcacaagaagtacagc
atcgccctggacatcggtaccaactctgtggctggccgtgtacccgtacggactacaagggtcccagcaagaaa
ttcaagggtggcaacaccgaccggcacagcatcaagaagaacactgtatggagccctgttgcacagcggc
gaaacagccaggccaccggctgaagagaaccggcagaagaagatacaccagacggaagaaccggatctg
ctatctcaagagatctcagcaacgagatggcaagggtggacacagcttccacagactggaaagagtccct
ggtggaaagaggataagaagcaccgagcggcacccatctcggcaacatcggtggacaggtggctaccacgaga
agtacccaccatctaccacctgagaaagaactggtagggacagcaccgacaaggccacctgcggctgtatctatct

ggccctggcccacatgatcaagttccggggcacttcgtatcgaggcgacctgaacccgacaacagcgacgt
ggacaagctgtcatccagctggcagacactacaaccagctgtcgaggaaaacccatcaacgcgcggcgtg
gacgccaaggccatcctgtcgccagactgagcaagagcagacggctggaaaatctgatcgcccagctgcccgc
gagaagaagaatggccttgcggaaacctgattgccctgagcctgggcctgaccccaactcaagagcaactcg
acctggccgaggatgccaactgcagctgagcaaggacacctacgacgacgacctggacaacctgctggccag
atcgcgaccagtacgccgacctgttctggcccaagaacctgtccgacgccatcctgctgagcgacatcctgag
agtgaacaccgagatcaccaaggccccctgagcgccttatgatcaagagatacgacgagcaccaccaggacc
tgaccctgctgaaagctctgtcgccagcagctgcctgagaagtacaaagagatttctgaccagagcaagaac
ggctacgcccggctacattgacggcggagccagccaggaagagttctacaagttcatcaagccatcctggaaaag
atggacggcaccgaggaactgctcgtgaagctgaacacagagaggacctgctcggaagcagcggacccgacaa
cggcagcatccccaccagatccacctggagagactgcacgccattctgcggcggcaggaagatttacccattcc
tgaaggacaaccggaaaagatcgagaagatcctgaccctccgcatccctactacgtggccctggccaggg
gaaacagcagattcgcctggatgaccagaaagagcggagaaaccatcaccctggaaactcgaggaagtggtg
gacaagggcgctccgcccagagctcatcgagcggatgaccaacttcgataagaacctgccaacgagaaggtg
ctgcccacagcagctgtacgacttccgtataacgagctgaccaaagtgaaatacgtgaccgagg
aatgagaaagccgcctcctgagcggcggcagcagaaaaaggccatcgacccgtgtcaagaccaccgg
aaagtgaccgtgaagcagctgaaagaggacttcaagaaaatcgagtgcctgactccgtggaaatctccggcg
tggaaagatcggtcaacgcctccctggcacataccacgatctgctgaaaattatcaaggacaaggactcctggac
aatgagggaaaacgaggacattctggaagatctgtgctgaccctgacactgtttgaggacagagatgtcgg
aacggctgaaaacctatgcccacctgtcgacgacaaagtgtatgaaagcagctgaaacggcggagataccggct
ggggcaggctgagccggaaagctgatcaacggcatccgggacaagcagtcggcaagacaatcctggattcctga
agtccgacggctcgccaacagaaaactcatgcagctgatccacgacgacgacgacccgtgaccttaagaggacatcca
gaaagcccagggtccggccagggcgatagcctgcacgagcacattgccaatctggccggcagcccccgcattaa
gaagggcatcctgcagacagtgaaggtggacgagctcgtaaaagtgtatggccggcacaagcccgagaac
atcgatcgaaatggccagagagaaccagaccacccagaagggacagaagaacagccgcgagagaatgaa
gcggatcgaaagagggcatcaaagagctggcagccagatcctgaaagaacacccctggaaaacacccagctg
cagaacgagaagctgtacctgtactacctgcagaatggcgggatgtacgtggaccaggaactggacatcaac
cggtccgactacgatgtggacgctatcgccctcagagcttctgtaaggacgactccatcgacaacaaggctg
accagaacgacaagaaccggggcaagagcgacaaacgtgcctccgaagaggtcgtaagaagatgaagaa
ctactggcggcagctgtgaacgccaagctgtattaccagagaagttcgacaatctgaccaaggccgagagagg
cgccctgagcgaactggataaggccggctcatcaagagacagctggtaaaacccggcagatcacaagcacg
tggcacagatcctggactccggatgaaactaagtacgacgagaatgacgatccggaaatgtgaaatgaa
tcaccctgaagtccaaagctgggtccgattccggaaaggattccagtttacaatgtgcgcgagatcaacaactacca
ccacgcccacgacgcctacctgaacgcccgtggaaaccgcctgatcaaaaagtaccctaagctggaaagcg
agttcgtgtacggcgactacaagggtacgacgtcgaaagatgtccaaagcgacgaggaaatcgcaag
gctaccgccaagtacttctacagcaacatcatgaactttcaagaccgagattaccctggcaacggcgagatcc
ggaagcggcctctgatcgagacaacggcggaaaccggggagatcggtggataagggccgggatttgcccacc
gtgcggaaagtgtgagcatccccaaagtgaatatcgtaaaaagaccgaggtgcagacaggcggctcagcaa
agagtctatcctgcccagaggaacagcgataagctgatcgccagaaagaaggactggaccctaagaagtacg

gcggcttcgacagccccaccgtggcctattctgtgtgggtggccaaagtggaaaaggcaagtccaaagaact
gaagagtgtgaaagagactgctgggatcaccatcatggaaagaagcagcgttcgagaagaatccatcgacttctg
gaagccaagggtacaaaagaagtaaaaaggacactgatcatcaagctgcctaagtactccctgtcgagctggaa
aacggccggaagagaatgctggcctctgccggcgaactgcagaaggaaacgaactggccctgccctccaaata
tgtgaacttcctgtacctggccagccactatgagaagctgaaggctccccgaggataatgagcagaaacagctg
ttgtggAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCAGTTCTCCAAGAGAGTGTACCTGGC
CGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGA
ATATCATCCACCTGTTACCTGACCAATCTGGAGCCCTGCCGCCTCAAGTACTTGACACCACATCGACCGGAA
GAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCTGATCCACCAGAGCATCACCGGCCTGTACGAGACAC
GGATCGACCTGTCTCAGCTGGAGGTGACTCTGGAGGATCTGGAGGATCTGGAGGATCTGGAGACACCAAGGAAC
AAGCGAGTCAGCAACACCAAGAGCAGTGGCGGCAGCAGCGGGCAGCAGCACCTAAATAGAAGATGA
GTATCGGCTACATGAGACCTCAAAAGAGCCAGATGTTCTCTAGGGTCCACATGGCTGTGATTTCTCAGGCCTGGG
CGGAAACCGGGGGCATGGGACTGGCAGTCGCCAAGCTCCTCTGATCATACCTCTGAAAGCAACCTCACCCCGTG
TCCATAAAACAATACCCATGTCACAAGAAGCCAGACTGGGATCAAGCCCCACATAAGAGACTGTTGGACCA
GGAATACTGGTACCCCTGCCAGTCCCCCTGGAACACGCCCTGCTACCCGTTAAGAAAACCAGGGACTAATGATTATAG
GCCGTCCAGGATCTGAGAGAAGTCAACAAGCGGGTGGAAAGACATCCACCCACCCTGCCAACCTTACAACCT
CTTGAGCGGGCTCCCACCGTCCCACCACTGGTACACTGTGCTGATTAAAGGATGCCCTTCTGCCAGACTCCACCC
CACCAGTCAGCCTCTCGCCTTGGAGGAGATCCAGAGATGGGAATCTCAGGACAATTGACCTGGACCAGACT
CCACAGGGTTCAAAAACAGTCCCACCTGTTAATGAGGCACTGCACAGAGACCTAGCAGACTCCGGATCCAGC
ACCCAGACTGATCCTGCTACAGTAGCTGGATGACTTACTGCTGGCCGCCACTCTGAGCTAGACTGCCAACAGGTAC
TCGGGCCCTGTTACAAACCCCTAGGAAACCTCGGGTATCGGGCCTCGGCCAAGAAAGCCCAAATTGCCAGAAACA
GGTCAAGTATCTGGGTATCTCTAAAAGAGGGTCAGAGATGGCTGACTGAGGCCAGAAAAGAGACTGTGATGGG
CAGCCTACTCCGAAGACCCCTGACAACACTAAGGGAGTTCTAGGGAAGGCAGGCTCTGCCCTTCATCCCTGG
TTGAGGAAATGGCAGCCCCCTGTACCCCTCACCACCGGGACTCTGTTAATTGGGCCAGACCAACAAAA
GGCCTATCAAGAAATCAAGCAAGCTCTAACTGCCAGGCCCTGGGTGCCAGATTGACTAAGCCCTTGAACTCT
TTGTCGACGAGAAGCAGGGCTACGCCAAAGGTGCTTAACGCAAAAACGGACCTGGCCTGGCGTGGCCGGTGGCCTA
CCTGCTCCAAAAGCTAGACCCAGTAGCAGCTGGGTGGCCCCCTGCCACGGATGGTAGCAGCCATTGCCGTACTGA
CAAAGGATGCAGGCAAGCTAACCATGGGACAGCCACTAGTCATTGCCCTGGCTTGGACACGGACGGGTCC
AGTTCGGACCGGTGGTAGCCCTGAACCCGGCTACGCTGCTCCACTGCCCTGGGTGCCAGATTGCAACACAACTGCC
GATATCCTGGCCGAAGCCCACGGAACCCGACCCGACCTAACGGACCGCCCTCCAGACGCCGACCAACCT
GGTACACGGATGGAAGCAGTCTCTTACAAGAGGGACAGCGTAAGGCGGGAGCTCGGGTACCGACCGAGACCGA
GGTAATCTGGCTAAAGCCCTGCCAGCGGGACATCCGCTAGCGGGCTGAACTGATAGCAGTACCTACCCAGGCC
AAAGATGGCAGAAGGTAAGAAGCTAAATGTTACTGATAGCCGTTAGCTTTGCTACTGCCCATATCCATGGAGAAAT
ATACAGAAGGCGTGGGTGGCTCACATCAGAAGGCAAAAGAGATCAAAATAAGACGAGATCTGCCCTACTAA
AGCCCTTCTGCCAAAAGACTTAGCATAATCCATTGTCAGGACATCAAAAGGGACACAGCGCCGAGGCTAGA
GGCAACCGGATGGCTGACCAAGCGGCCGAAAGGCAGCCATCACAGAGACTCCAGACACCTCACCCCTC
AGAAAATTCATCACCCCTGGCGGCTCAAAAAGAACCGCCGACGGCAGCGAATTGAGCTCTGGTTCT
CCCCAAGAAGAAGAGGAAAGTCGGTAGTTCCGGATCTGCAATGCGAGATGCAAGCTGAAGCA

aatgcagataactcagtggagaagaagaaagcttggcccacaaccatTCACGGTAGAGCAGTGTGCAATAATGC
caacgatgtgaagaaatttggagaagaagctggattccataCTGTGGAGGCTGTGCCATGCGCCAAGAAGGAGCTA
ataaaatattaaggaaatttgtaagccaaagctgataaaaattctggctgaggcagctaaatttagttccaatgggttca
ccactgcaactgaattccaccaaaaggcggtcagagatcatacagattactactggctccaaagagcttgcacaaacta
cttcAACCCAAAGAAGAGGAAAGTCTAA