

Sequence: p2200 MUC5B-TdTomato-pDTA-TK DV.dna (Circular / 12,086 bp)

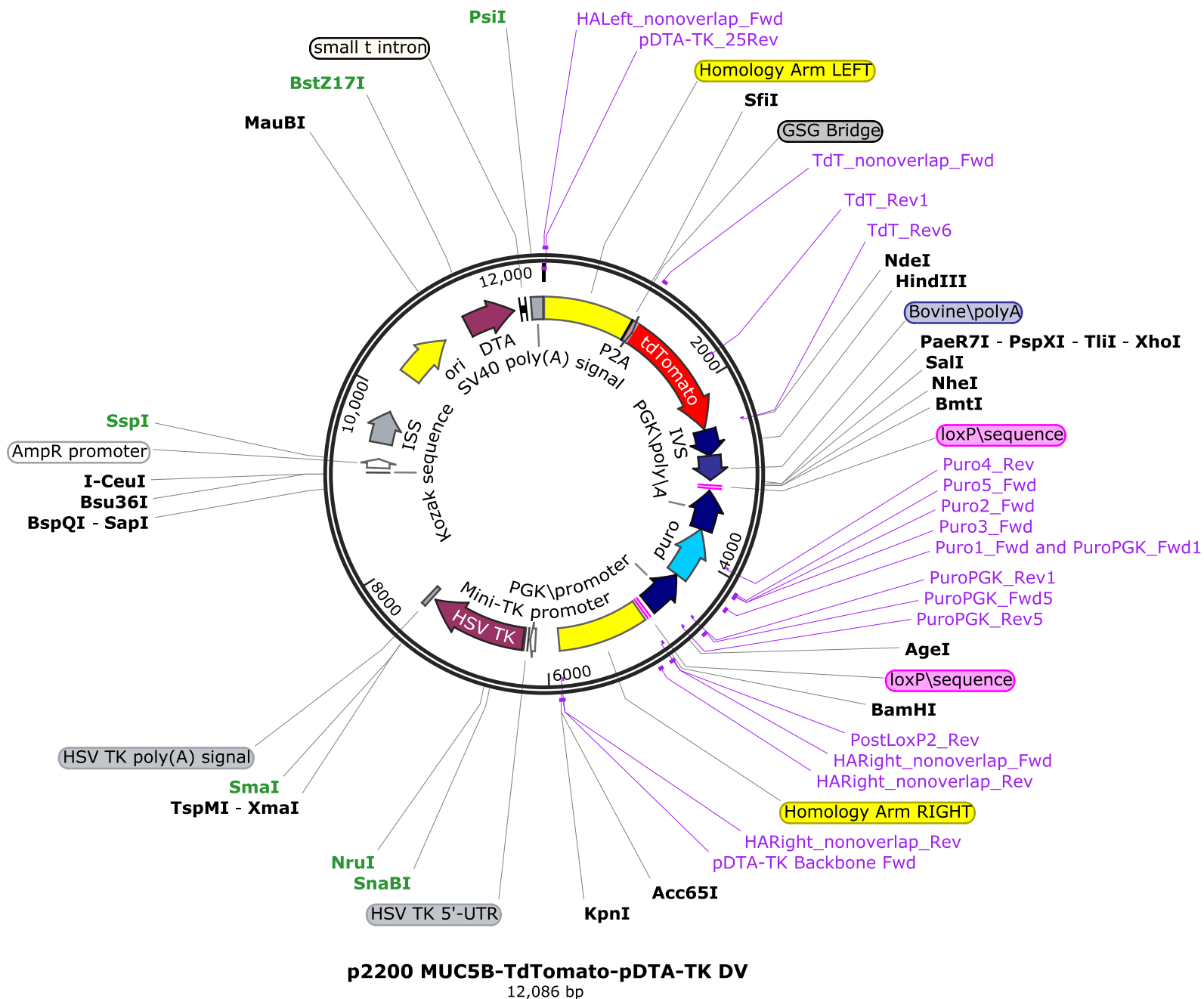
Enzymes: Unique Cutters (27 of 653 total)

Features: 23 visible, 41 total

Primers: 17 visible, 55 total

Blunt Cutters **Green**, Unique Cutters **Bold**

This MUC5B reporter/excisable selection cassette was generated using HiFi Assembly. Intended to be a promoter fusion, it replaces the stop codon of MUC5B Exon 49, in frame, with a GSG bridge fused to a P2A sequence. Left and right homology arms were generated using PCR of genomic DNA from the FHS14030-3 iPSC (rs35705950 TT homozygote). The GSG-P2A sequence were ordered as a duplex from IDT. The TdTomato-Puro sequence extending to the 2nd loxP sequence was PCR amplified from the SPC-DV vector. The backbone vector was PCR amplified from the pDTA-TK plasmid. HiFi assembly was performed using 25 nucleotide overlap sequences. Final sequencing of the donor vector is embedded in this file under the alignment tab for Maxiprep Clone#2 and #7. Embedded chromatogram ab files are from MUC5B T/TdtTdtT Clone 62, Cre-excised Clone 26, which was the iPSC line referred to as MUC5B T/TdtTdtT.



HLeft_nonoverlap_Fwd

TCTGCCTTTCCTCCTCCC

5' TCTGCCTTTCCTCCTCCCAGACTCCTGTCAAGTCCGCATCAACACGACCATCCTG
 ...
 3' AGACGGAAAGGAGGAGGGTCTGAGGACAGTTTCAGGCGTAGTTGTGCTGGTAGGAC

55

Homology Arm LEFT

AGACGGAAAGGAGGAGGGTCTGAGG

pDTA-TK_25Rev

TGGCACCAGGGCTGCGAGACCGAGGTCAACATCACCTTCTGCGAGGGCTCCTGCC
 ++++++
 ACCGTGGTCCCGACGCTCTGGCTCCAGTTGTAGTGGAAGACGCTCCCGAGGACGG

110

Homology Arm LEFT

CCGGAGCGTCCAAGTGAGTGGGCTCCTGGCCCTGTGCCAAGAGCACCTGCGTGTG
 ++++++
 GGCCTCGCAGGTTCACTCACCCGAGGACCGGGACACGGTTCTCGTGGACGCACAC

165

Homology Arm LEFT

GTGGGTCCGCCCTGGCCCGTCAGCTGGCTGGCAGCCTGCTCTGGGTCAGGAGGGC
 ++++++
 CACCCAGGCGGGACCGGGCAGTCGACCGACCGTTCGGACGAGACCCAGTCCTCCCG

220

Homology Arm LEFT

CTACGCCAGCTCCAGGAAGGAGGGAAGGGGCTCCCCAGCAGGGCTCCCCCTCTCC
 ++++++
 GATGCGGTCTGAGGTCCTTCTCCTTCCCCGAGGGGTCGTCCCGAGGGGGAGAGG

275

Homology Arm LEFT

ACGTCCATTGGCAGAAAGCCTGGGGCCGCTCTCCTAGCCAGGCCAGGGCCCGGCT
 ++++++
 TGCAGGTAACCGTCTTTCGGACCCCGGCGAGAGGATCGGTCCGGTCCCGGGCCGA

330

Homology Arm LEFT

GAGCCGAGCCACATGGCACAGAGCTCCCCGCTGCGAACCAGGTCCAGGTGTCCC
 ++++++
 CTCGGCTCGGTGTACCGTGTCTCGAGGGGCGACGCTTGGGTCCAGGTCCACAGGG

385

Homology Arm LEFT

CCTGGGGAGGGTGGGTGGAACCAGAGCCTTCTGTGAAGCCCCCATGGCTTGCCA
 ++++++
 GGACCCCTCCACCCACCTTGGTCTCGGAAGACACTTCGGGGGGTACCGAACGGT

440

Homology Arm LEFT

CTGTGGTTGGCTGTCAGCAGGAGGGGCCAGAAGGGTTGACCCTTAGCTAGAAGAG 495
GACACCAACCGACAGTCGTCCTCCCCGGTCTTCCCAACTGGGAATCGATCTTCTC

Homology Arm LEFT

GCAGGAGGAAGGGTCAAGGGCTGCTTAGGGTCAGCAGGATTGGAGGAGCTGGTTTC 550
CGTCCTCCTTCCCAGTTCCCGACGAATCCCAGTCGTCCTAACCTCCTCGACCAAG

Homology Arm LEFT

AGACAGGGGAAGTGCAGGCCACAGAAGACAGAGCCGGGAGCCAAGGGACCCTGGG 605
TCTGTCCCCTTCACGTCCGGTGTCTTCTGTCTCGGCCCTCGGTTCCCTGGGACCC

Homology Arm LEFT

GGCCCAGAGCTGCCCCATCCTGAGAAGGTGAAGCCTCACGTGCCACAGAGCAGG 660
CCGGGTCTCGACGGGGTAGGACTCTTCCACTTCGGAGTGCACGGGTGTCTCGTCC

Homology Arm LEFT

CCACTGTGGACAGAAGGGGGCGGCCGTCTGGCCCAGGACCCCAGAGGCCCATGTG 715
GGTGACACCTGTCTTCCCCCGCCGGCAGACCGGGTCCTGGGGTCTCCGGGTACAC

Homology Arm LEFT

TCACCATGGCTGGGCAGAGAAACGGCGGGGCCAAGGTGGCTGATGTGAGGGCCAC 770
AGTGGTACCGACCCGTCTCTTTGCCGCCCCGGTTCCACCGACTACACTCCCGGTG

Homology Arm LEFT

CCTGCGTCCACAGGTA CT CAGCAGAGGCCAGGCCATGCAGCACCAGTGACCTG 825
GGACGCAGGTGTCCATGAGTCGTCTCCGGGTCCGGTACGTGCTGGTACGTGGAC

Homology Arm LEFT

CTGCCAGGAGAGGCGGGTCCACGAGGAGACGGTGCCCTTGCACTGTCCTAACGGC 880
GACGGTCCTCTCCGCCAGGTGCTCCTCTGCCACGGGAACGTGACAGGATTGCCG

Homology Arm LEFT

TCAGCCATCCTGCACACCTACACCCACGTGGATGAGTGTGGCTGCACGCCCTTCT 935
AGTCGGTAGGACGTGTGGATGTGGGTGCACCTACTCACACCGACGTGCGGGAAGA

Homology Arm LEFT

SfiI

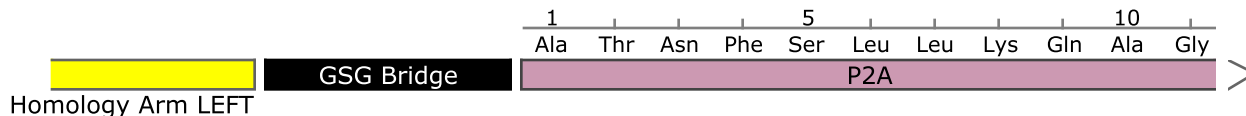
GTGTCCCTGCGCCCATGGCTCCCCACACACCCGTGGCTTCCCGGCCAGGAGGC
 CACAGGGACGCGGGTACCGAGGGGGTGTGTGGGCACCGAAGGGCCGGGTCTCTCG

990

Homology Arm LEFT

CACTGCTGTCTcttggcagcggagCCACGAACCTTCTCTGTAAAGCAAGCAGGA
 GTGACGACAGagaccgtcgcctCGGTGCTTGAAGAGAGACAATTTCTGTTCTCTCT

1045

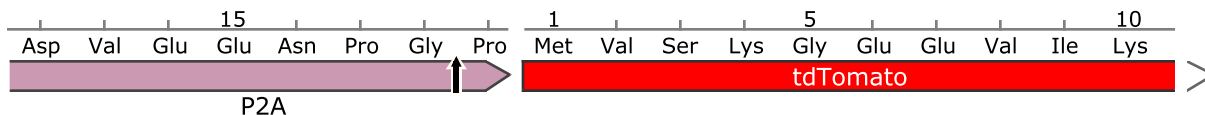


TdT_nonoverlap_Fwd

ATGGTGAGCAAGGGCGAG

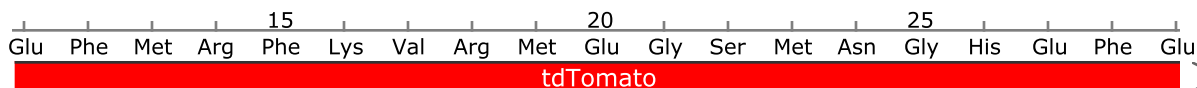
GACGTGGAAGAAAACCCGGTCCTATGGTGAGCAAGGGCGAGGAGgtcatcaaag
 CTGCACCTTCTTTTGGGGCCAGGATACCACTCGTTCCCGCTCCTCcagtagtttc

1100



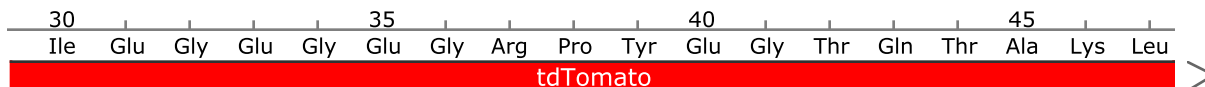
agttcatgcgcttcaagggtgcgcattggagggctccatgaacggccacgagttcga
 tcaagtacgcgaagttccacgcgtacctcccgaggtacttgccggtgctcaagct

1155



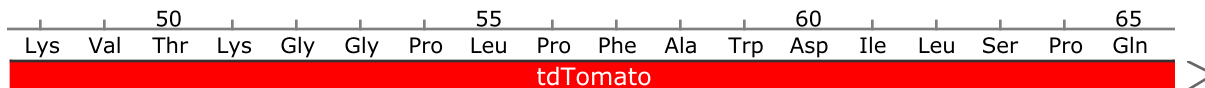
gatcgagggcgagggcgagggcgccctacgagggcaccagaccgccaagctg
 ctagctcccgcctcccgcctcccggcggggatgctcccgtgggtctggcggttcgac

1210



aaggtagaccaaggcgggcccccctgcccttcgcctgggacatcctgtccccccagt
 ttccactggttcccgcgggggacgggaagcggaccctgtaggacaggggggtca

1265



tcatgtacggctccaaggcgtagctgaagcaccgacatcccgattacaa 1320
 agtacatgccgaggttccgcacatgcacttcgtggggcggctgtaggggctaattgtt
 70 75 80
 Phe Met Tyr Gly Ser Lys Ala Tyr Val Lys His Pro Ala Asp Ile Pro Asp Tyr Lys
 tdTomato >

gaagctgtccttccccgagggcttcaagtgggagcgcgtgatgaacttcgaggac 1375
 cttcgacaggaaggggctcccgaagttcacccctcgcgcaactacttgaagctcctg
 85 90 95 100
 Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val Met Asn Phe Glu Asp
 tdTomato >

ggcggctctggtgaccgtgaccaggactcctccctgcaggacggcacgctgatct 1430
 ccgccagaccactggcactgggtcctgaggaggagcgtcctgccgtgcgactaga
 105 110 115 120
 Gly Gly Leu Val Thr Val Thr Gln Asp Ser Ser Leu Gln Asp Gly Thr Leu Ile
 tdTomato >

acaaggtgaagatgcgcggcaccaacttccccccgacggccccgtaatgcagaa 1485
 tgttccacttctacgcgccgtgggttgaagggggggctgccggggcattacgtcctt
 125 130 135
 Tyr Lys Val Lys Met Arg Gly Thr Asn Phe Pro Pro Asp Gly Pro Val Met Gln Lys
 tdTomato >

gaagaccatgggctgggaggcctccaccgagcgcctgtacccccgcgacggcgctg 1540
 cttctggtacccgaccctccggaggtggctcgcggacatgggggcgctgccgcac
 140 145 150 155
 Lys Thr Met Gly Trp Glu Ala Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val
 tdTomato >

ctgaagggcgagatccaccaggccctgaagctgaaggacggcgggccactacctgg 1595
 gacttcccgctctaggtgggtccgggacttcgacttctgcccgggtgatggacc
 160 165 170 175
 Leu Lys Gly Glu Ile His Gln Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu
 tdTomato >

tggagttcaagaccatctacatggccaagaagcccgctgcaactgcccggtacta 1650
 acctcaagttctggtagatgtaccggttcttcgggcacgttgacgggcccgatgat
 180 185 190
 Val Glu Phe Lys Thr Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr Tyr
 tdTomato >

ctacgtggacaccaagctggacatcacctcccacaacgaggactacaccatcgtg 1705
 gatgcacctgtggttcgacctgtagtggagggtgttgctcctgatgtggttagcac
 195 200 205 210
 Tyr Val Asp Thr Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr Thr Ile Val
 tdTomato >

gaacagtacgagcgctccgagggccgccaccacctgttcctggggcatggcaccg 1760
 cttgtcatgctcgcgaggctcccgggcgggtggtggacaaggaccccgtagcgtggc
 215 220 225 230
 Glu Gln Tyr Glu Arg Ser Glu Gly Arg His His Leu Phe Leu Gly His Gly Thr
 tdTomato >

gcagcaccggcagcggcagctccggcaccgcctcctccgaggacaacaacatggc 1815
 cgtcgtggccgctcgcgctcgaggccgtggcggaggaggctcctgttgttgtagccg
 235 240 245
 Gly Ser Thr Gly Ser Gly Ser Ser Gly Thr Ala Ser Ser Glu Asp Asn Asn Met Ala
 tdTomato >

gttgttgtagccg
 TdT_Rev1

cgtcatcaaagagttcatgcgcttcaagggtgcgcacatggagggtcctatgaacggc 1870
 gcagtagtttctcaagtacgcgaagttccacgcgtacctcccgagggtacttgccg
 250 255 260 265
 Val Ile Lys Glu Phe Met Arg Phe Lys Val Arg Met Glu Gly Ser Met Asn Gly
 tdTomato >

gcagtag
 TdT_Rev1

cacgagttcagagatcgagggcgagggcgagggccgccctacgagggcaccacaga 1925
 gtgctcaagctctagctcccgctcccgctcccgggcggggatgctcccggtgggtct
 270 275 280 285
 His Glu Phe Glu Ile Glu Gly Glu Glu Gly Arg Pro Tyr Glu Gly Thr Gln
 tdTomato >

ccgccaagctgaagggtgaccaaggggcggccccctgcccttcgcctgggacatcct 1980
 ggcggttcgacttccactggttcccgggggacgggaagcggaacctgttagga
 290 295 300
 Thr Ala Lys Leu Lys Val Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu
 tdTomato >

gtccccccagttcatgtacgggtccaaggcgtagctgaagcaccgccgacatc 2035
 caggggggtcaagtacatgccgaggttccgcatgcacttcgtggggcggctgtag
 305 310 315 320
 Ser Pro Gln Phe Met Tyr Gly Ser Lys Ala Tyr Val Lys His Pro Ala Asp Ile
 tdTomato >

cccgattacaagaagctgtccttccccgagggttcaagtgggagcgcgtgatga 2090
 gggctaattgttcttcgacaggaaggggtcccgaaagttcacccctcgcgcactact
 325 330 335 340
 Pro Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val Met
 tdTomato >

acttcgaggacggcgggtctggtgaccgtgaccaggactcctccctgcaggacgg 2145
 tgaagctcctgccgccagaccactggcactgggtcctgaggaggacgtcctgcc
 345 350 355
 Asn Phe Glu Asp Gly Gly Leu Val Thr Val Thr Gln Asp Ser Ser Leu Gln Asp Gly
 tdTomato >

cacgctgatctacaaggtgaagatgcgcggcaccaacttccccccgacggcccc 2200
 gtgcgactagatgttccacttctacgcgccgtgggtgaagggggggctgccgggg
 360 365 370 375
 Thr Leu Ile Tyr Lys Val Lys Met Arg Gly Thr Asn Phe Pro Pro Asp Gly Pro
 tdTomato >

gtaatgcagaagaagaccatgggctgggaggcctccaccgagcgcctgtaccccc 2255
 cattacgtcttcttcttgggtaccgaccctccggaggtggctcgcggacatggggg
 380 385 390 395
 Val Met Gln Lys Lys Thr Met Gly Trp Glu Ala Ser Thr Glu Arg Leu Tyr Pro
 tdTomato >

gcgacggcgtgctgaagggcgagatccaccaggccctgaagctgaaggacggcgg 2310
 cgctgccgcacgacttcccgctctaggtgggtccgggacttcgacttcctgccgcc
 400 405 410
 Arg Asp Gly Val Leu Lys Gly Glu Ile His Gln Ala Leu Lys Leu Lys Asp Gly Gly
 tdTomato >

ccactacctgggtggagtccaagaccatctacatggccaagaagcccgtagcactg 2365
 ggtgatggaccacctcaagttctggtagatgtaccgggttcttcgggcacgttgac
 415 420 425 430
 His Tyr Leu Val Glu Phe Lys Thr Ile Tyr Met Ala Lys Lys Pro Val Gln Leu
 tdTomato >

cccggtactactacgtggacaccaagctggacatcacctcccacaacgaggact 2420
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 gggccgatgatgatgcacctgtggttcgacctgtagtggagggtgttgctcctga
 435 440 445 450
 Pro Gly Tyr Tyr Tyr Val Asp Thr Lys Leu Asp Ile Thr Ser His Asn Glu Asp
tdTomato >

acaccatcgtggaacagtacgagcgctccgagggccgccaccacctgttcctgta 2475
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 tgtggtagcaccttgtcatgctcgcgaggctcccggcggtggtggacaaggacat
 455 460 465
 Tyr Thr Ile Val Glu Gln Tyr Glu Arg Ser Glu Gly Arg His His Leu Phe Leu Tyr
tdTomato >

NdeI

cggcattggacgagctgtacaagtaacatatgtctagatggaattaattcgctgtc 2530
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 gccgtacctgctcgacatgttcattgtatacagatctaccttaattaagcgacag
 470 475
 Gly Met Asp Glu Leu Tyr Lys *
tdTomato **IVS** >
cgtacctgctcgacatgt
 TdT_Rev6

tgcgagggccagctgttggggtgagtactccctctcaaaagcgggcatgacttct 2585
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 acgctcccggctcgacaaccccactcatgaggagaggttttcgcccgtactgaaga
IVS >

gcgctaagattgtcagtttccaaaaacgaggaggatttgatattcacctggcccg 2640
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 cgcgattctaacagtcaaagggtttttgctcctcctaaactataagtggaccgggc
IVS >

cggtgatgcctttgaggggtggccgcgtccatctggtcagaaaagacaatcttttt 2695
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 gccactacggaaactcccaccggcgcaggtagaccagtccttttctgtagaaaaa
IVS >

HindIII

gttgtcaagcttgagggtgtggcaggcttgagatctggccatacacttgagtgaca 2750
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 caacagttcgaactccacaccgtccgaactctagaccggtatgtgaactcactgt
IVS >

atgacatccacttttgcccttctctccacaggtgtccactcccaggtccaactgca 2805
 tactgtagggtgaaacggaaagagaggtgtccacaggtgaggggtccaggttgacgt
 IVS >

ggtcggagctcgcgtgatcagcctcgactgtgccttctagttgccagccatctgtt 2860
 ccagcctcgagcgactagtcggagctgacacggaagatcaacgggtcggtagacaa
 IVS > Bovine\polyA

gtttgccccctcccccgctgccttccttgaccctggaaggtgccactcccactgtcc 2915
 caaacggggaggggggcacggaaggaactgggaccttcacgggtgaggggtgacagg
 Bovine\polyA >

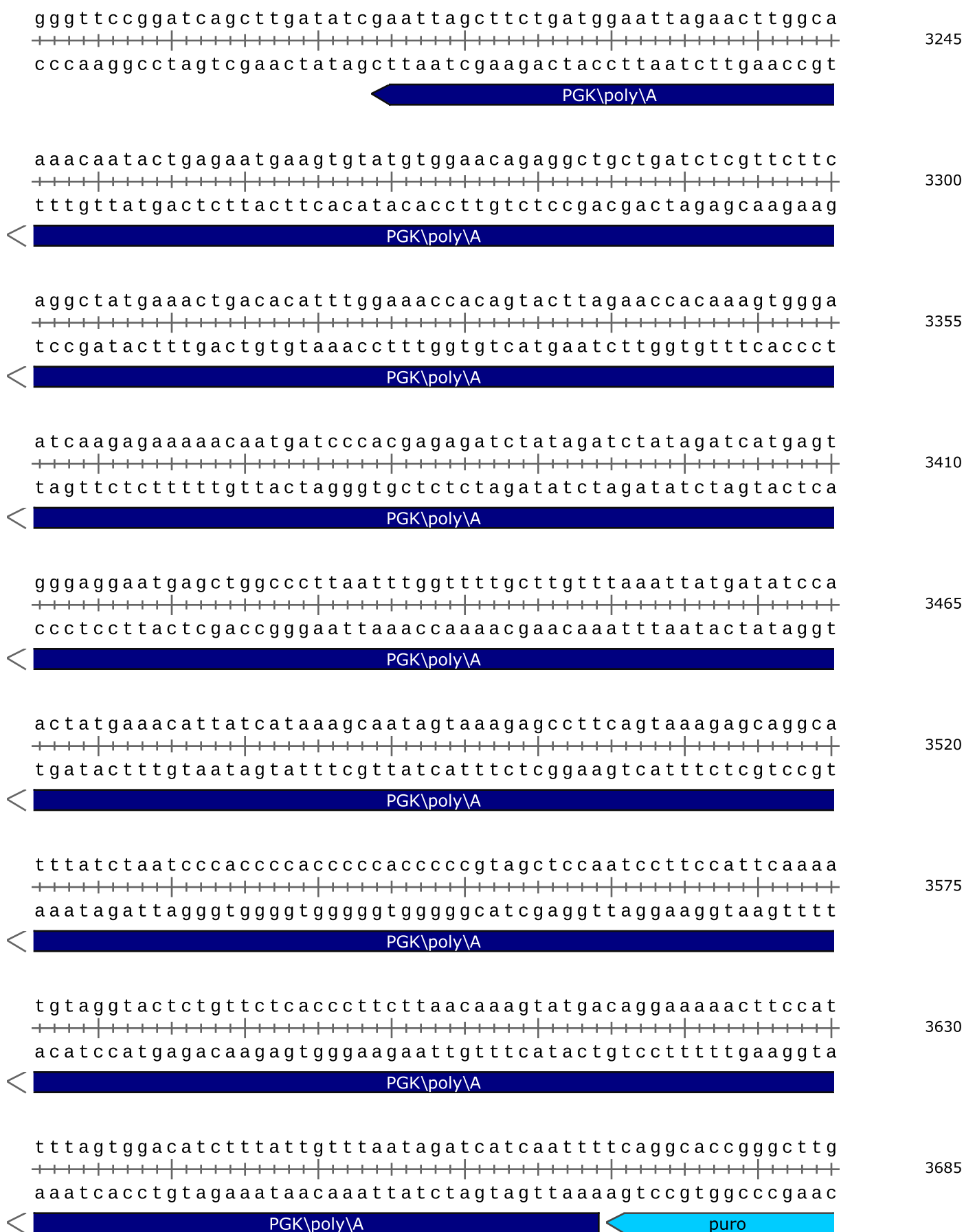
tttcctaataaaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctat 2970
 aaaggattattttactcctttaacgtagcgtaacagactcatccacagtaagata
 Bovine\polyA >

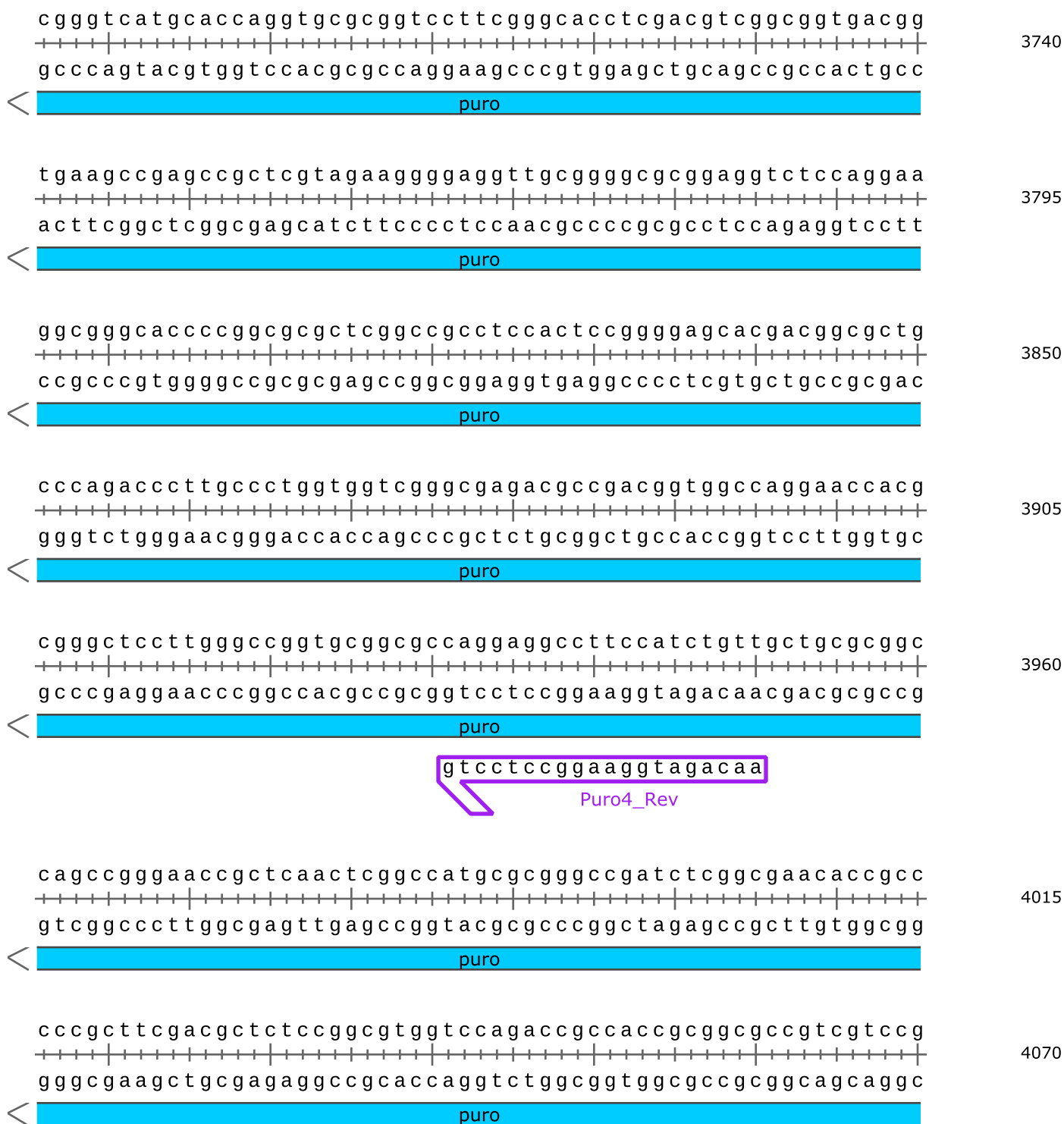
tctgggggggtgggggtggggcaggacagcaagggggaggattgggaagacaatagc 3025
 agacccccccacccccaccccgctcctgtcgttccccctcctaacccttctgttatcg
 Bovine\polyA >

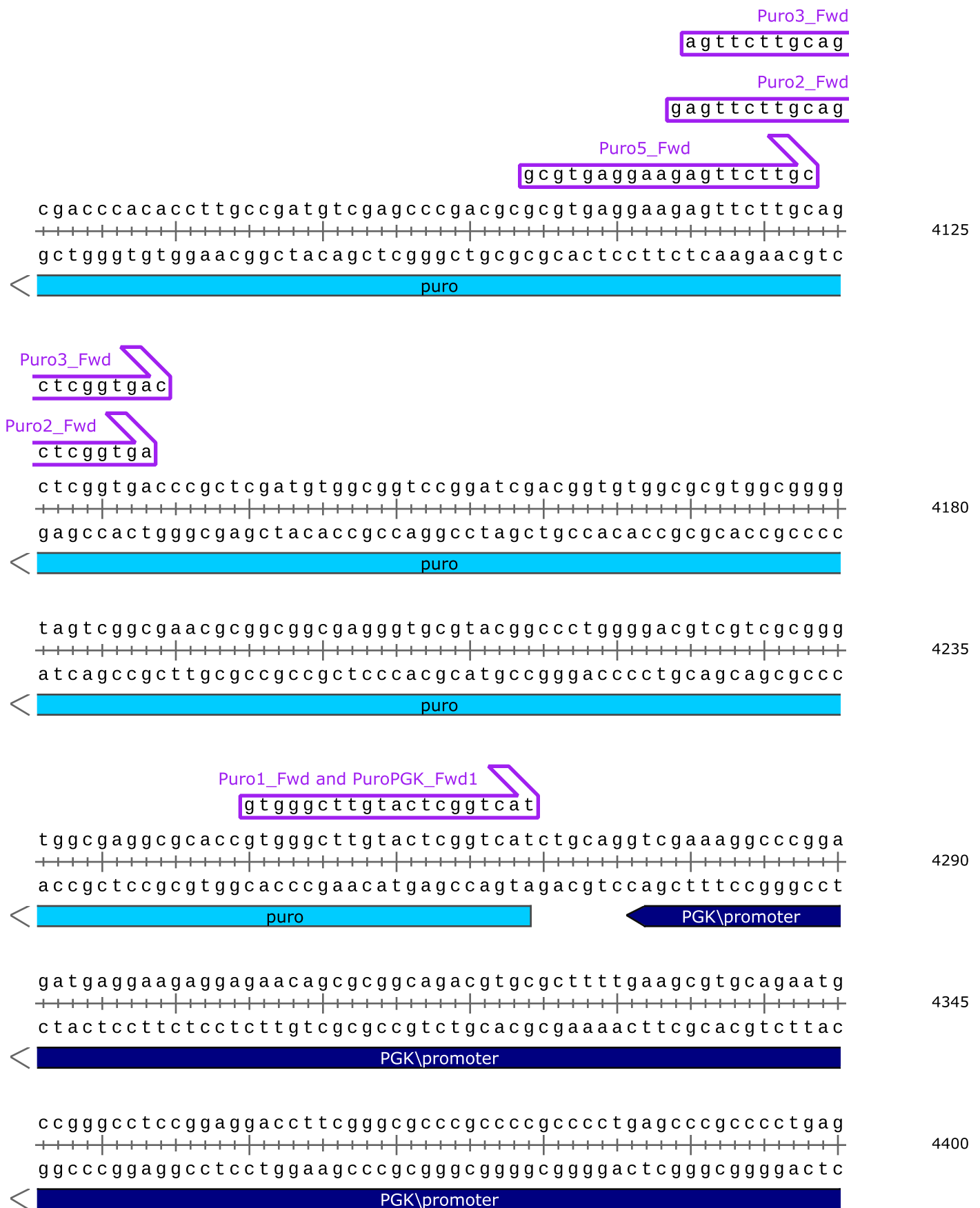
aggcatgctggggatgcggtgggctctatggcttctgaggcggaaagaaccagct 3080
 tccgtacgacccctacgccacccgagataccgaagactccgcctttcttggtcga
 Bovine\polyA >

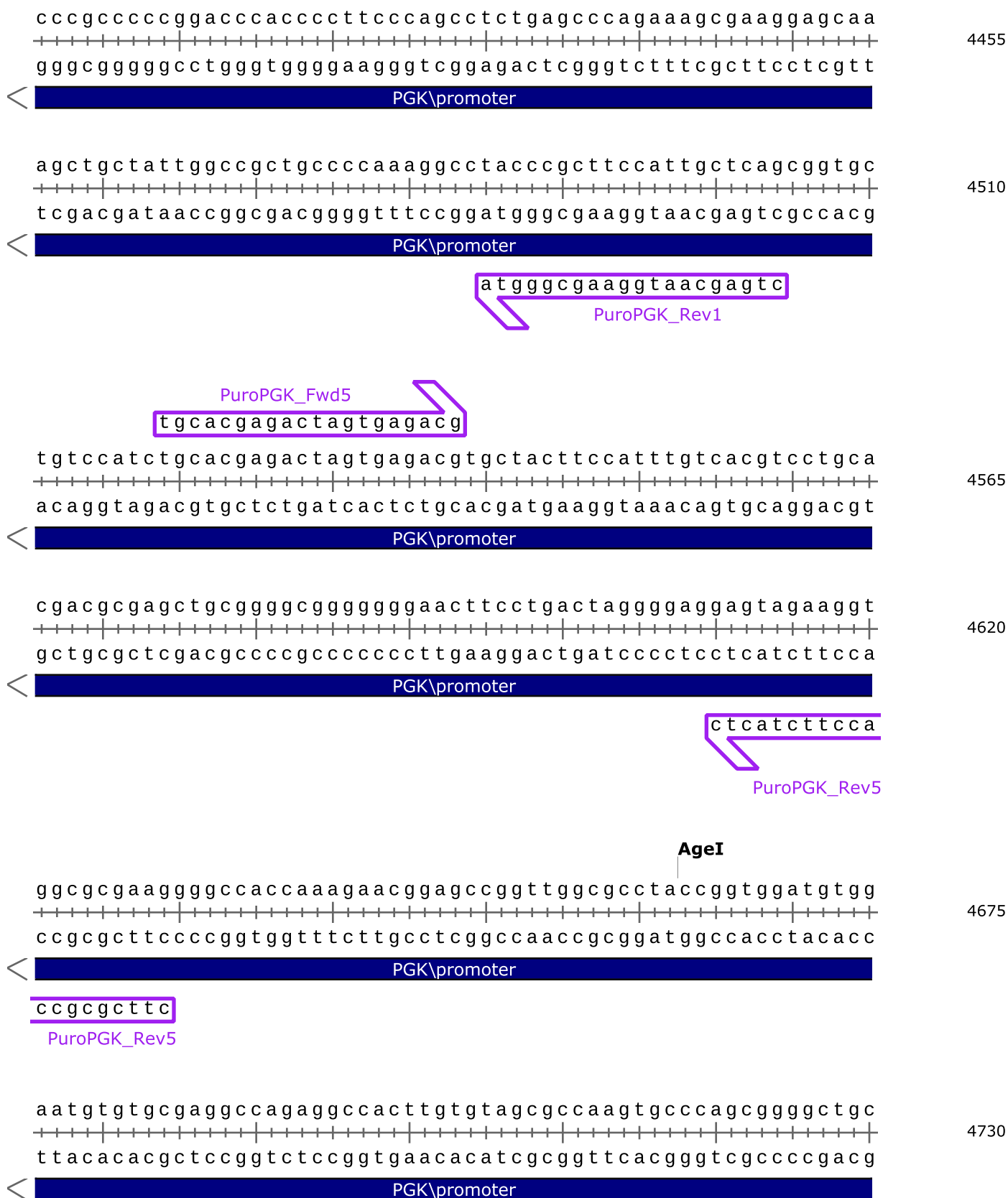
TliI
 XhoI
 PspXI
 PaeR7I SalI NheI BmtI
 ggggctcgatcctctagactcgaggtcgacgggtatcgataagctagcttgggctg 3135
 ccccgagctaggagatctgagctccagctgccatagctattcgatcgaacccgac
 Bovine\polyA >

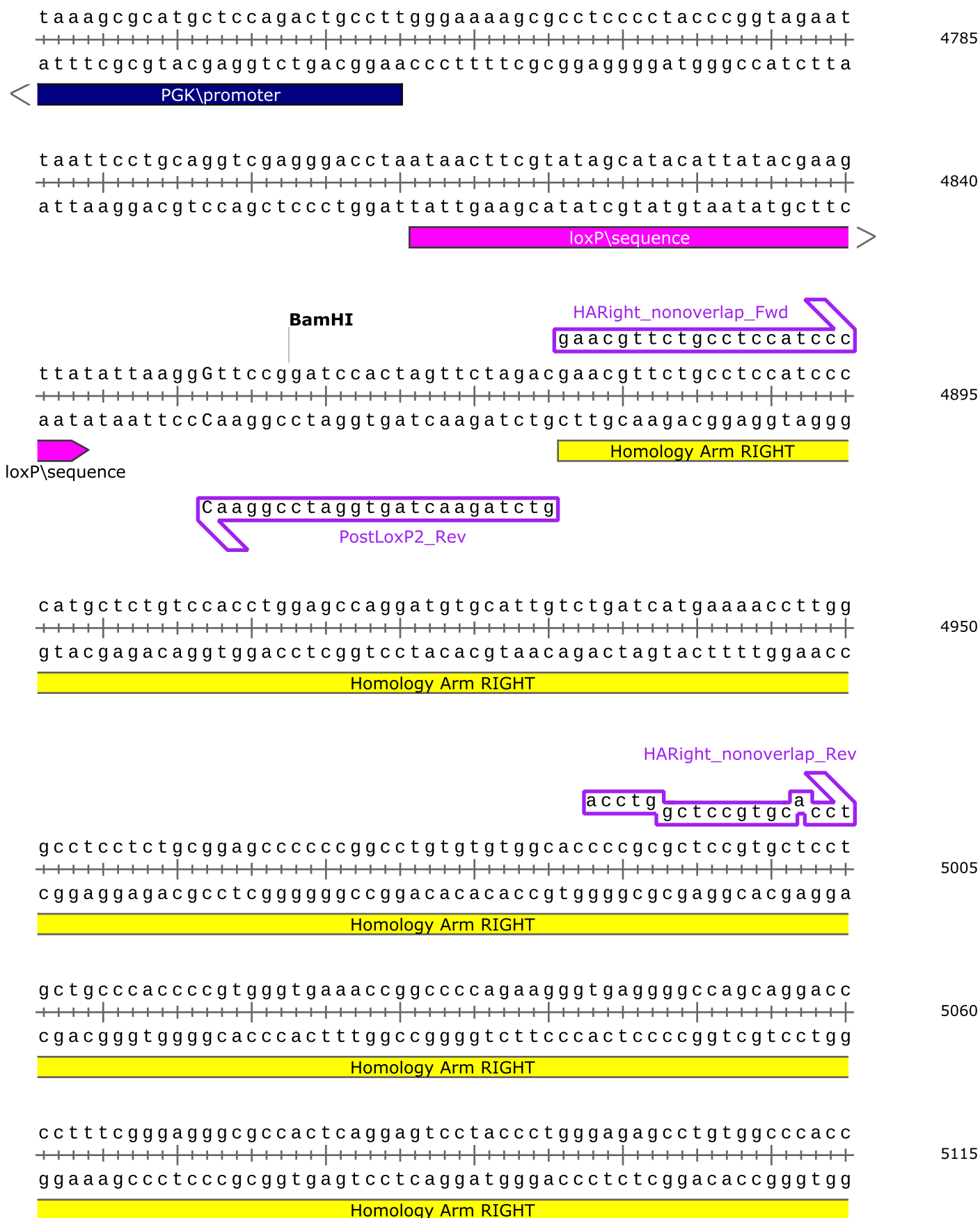
caggctcaggggacctaataacttcgtatagcatatactatacgaagttatatata 3190
 gtccagctccctggattattgaagcatatcgatatgtaatatgcttcaatataatt
 loxP\sequence >











Homology Arm RIGHT

ggggacaaggagccgcttgggtcatccacatgcccctgtgggcaccccgagagtca 5665
 cccctgttcctcggcgaaccagtaggtgtacggggacacccgtggggctctcagt

Homology Arm RIGHT

gtcttgggctgtgggagcactgcccaggggacagggggcagggggccgccaggac 5720
 cagaacccgacaccctcgtgacgggtccctgtcccccgtcccccggcggtcctg

Homology Arm RIGHT

tcagaggtcaccccggtgttatttccagtgtaggggattgtgtgaggaaaccag 5775
 agtctccagtggggccgacaataaaggtcacatcccctaacacactcctttggtc

Homology Arm RIGHT

ctcatgtaggaagatggaggcagtcagtggggcaagcaggcgaccaggcccggtc 5830
 gagtacatccttctacctccgtcagtcaccccggttcgtccgctgggtccgggccag

Homology Arm RIGHT

pDTA-TK Backbone Fwd

tctaggaatt

agcggggaggcagctgccccccagggcaggtgcacggagccaggttctaggaatt 5885
 tcgcccctccgtcgacggggggtcccggtccacgtgcctcgggtccaagatccttaa

Homology Arm RIGHT

tccacgtgcctcgggtcca

HARight_nonoverlap_Rev

Acc65I

KpnI

pDTA-TK Backbone Fwd

ctctagggtacctc

ctctagggtacctctaggagatcgatcctctagagtcgagcagtggtgttttcaa 5940
 gagatcccatggagatcctctagctaggagatctcagctcgtcacaccaaagtt

gaggaagcaaaaagcctctccaccaggcctggaatgtttccaccaaatgtcgag 5995
 ctcttctgtttttcggagaggtgggtccggaccttaciaaagggtgggttacagctc

cagtggtgtttttgcaagaggaagcaaaaagcctctccaccaggcctggaatgtt 6050
 gtcacaccaaagcgttctccttcgtttttcggagaggtgggtccggaccttaciaa

tccacccaatgtcgagcaaaccgcccagcgtcttgtcattggcgaattcgaac 6105
 aggtgggttacagctcgtttggggcgggtcgcagaacagtaaccgcttaagcttg

acgcagatgcagtcggggcggcgcggtcccagggtccacttcgcataattaagggtga 6160
 tgcgctctacgtcagccccgcccgcgcagggtccagggtgaagcgtataattccact

Mini-TK promoter

cgcgtgtggcctcgaacaccgagcgcaccctgcagcgaaccgcttaacagcgtcaa 6215
 gcgcacaccggagcttgtggctcgcctgggacgtcgcctgggcgaattgtcgcagtt

Mini-TK promoter

cagcgtgccgcagatcttgggtggcgtgaaactcccgcacctcttcggccagcgcc 6270
 gtcgcacggcgtctagaaccaccgcactttgagggcgtggagaagccgggtcgcgg

HSV TK 5'-UTR

ttgtagaagcgcgtatggcttcgtaccccgggccatcagcacgcgtctgcgttcga 6325
 aacatcttcgcgcataccgaagcatggggccggtagtcgtgcgcagacgcaagct

1 5 10
 Met Ala Ser Tyr Pro Gly His Gln His Ala Ser Ala Phe Asp
 HSV TK

ccaggctgcgcgttctcgcggccatagcaaccgacgtacggcgttgcgccctcgc 6380
 ggtccgacgcgcaagagcgccgggtatcgttggctgcatgccgcaacgcgggagcg

15 20 25 30
 Gln Ala Ala Arg Ser Arg Gly His Ser Asn Arg Arg Thr Ala Leu Arg Pro Arg
 HSV TK

cggcagcaagaagccacggaagtccgcccggagcagaaaatgcccacgctactgc 6435
 gccgtcgttcttcggtgccttcaggcgggcctcgtcttttacgggtgcgatgacg

35 40 45 50
 Arg Gln Gln Glu Ala Thr Glu Val Arg Pro Glu Gln Lys Met Pro Thr Leu Leu
 HSV TK

gggtttatatagacggtcccccacgggatggggaaaaccaccaccacgcaactgct 6490
 cccaaatataatctgccagggtgcccctacccttttggtgggtggtgcgttgacga

55 60 65
 Arg Val Tyr Ile Asp Gly Pro His Gly Met Gly Lys Thr Thr Thr Thr Gln Leu Leu
 HSV TK

SnaBI

ggtggccctgggttcgcgcgacgatatcgtctacgtacccgagccgatgacttac
 ccaccgggacccaagcgcgctgctatagcagatgcatgggctcggctactgaatg

6545

70 75 80 85
 Val Ala Leu Gly Ser Arg Asp Asp Ile Val Tyr Val Pro Glu Pro Met Thr Tyr
 HSV TK

NruI

tggcgggtgctgggggcttccgagacaatcgcgaaacatctacaccacacaacacc
 accgcccacgacccccgaaggctctgttagcgcttgtagatgtgggtgtgttgtgg

6600

90 95 100 105
 Trp Arg Val Leu Gly Ala Ser Glu Thr Ile Ala Asn Ile Tyr Thr Thr Gln His
 HSV TK

gccttgaccaggggtgagatatcggccggggacgcggcggtggtaatgacaagcgc
 cggaactggtcccactctatagccggcccttgcgccgccaccattactgttcgcg

6655

110 115 120
 Arg Leu Asp Gln Gly Glu Ile Ser Ala Gly Asp Ala Ala Val Val Met Thr Ser Ala
 HSV TK

ccagataacaatgggcatgccttatgccgtgaccgacgccgttctggctcctcat
 ggtctattgttaccgtaggaatacggcactggctgcgggcaagaccgaggagta

6710

125 130 135 140
 Gln Ile Thr Met Gly Met Pro Tyr Ala Val Thr Asp Ala Val Leu Ala Pro His
 HSV TK

atcgggggggaggctgggagctcacatgccccgcccccgccctcaccctcatct
 tagccccccctccgaccctcgagtgtacggggcgggggcgaggagtgggagtaga

6765

145 150 155 160
 Ile Gly Gly Glu Ala Gly Ser Ser His Ala Pro Pro Pro Ala Leu Thr Leu Ile
 HSV TK

tcgaccgccatcccatcgccgccctcctgtgctacccggccgcgcgataccttat
 agctggcggtagggtagcggcgggaggacacgatgggcccggcgcgctatggaata

6820

165 170 175
 Phe Asp Arg His Pro Ile Ala Ala Leu Leu Cys Tyr Pro Ala Ala Arg Tyr Leu Met
 HSV TK

gggcagcatgacccccagggcgtgctggcggttcgtggccctcatcccgccgacc 6875
 cccgtcgtactggggggtccggcacgaccgcaagcaccgggagtagggcggctgg
 180 185 190 195
 Gly Ser Met Thr Pro Gln Ala Val Leu Ala Phe Val Ala Leu Ile Pro Pro Thr
 HSV TK >

ttgcccggcacaaacatcgtgttgggggccccttcggaggacagacacatcgacc 6930
 aacgggccgtgtttgtagcacaacccccgggaaggcctcctgtctgtgttagctgg
 200 205 210 215
 Leu Pro Gly Thr Asn Ile Val Leu Gly Ala Leu Pro Glu Asp Arg His Ile Asp
 HSV TK >

gcctggccaaacgccagcgccccggcgagcggccttgacctggctatgctggccgc 6985
 cggaccgggtttgcggtcgcggggccgctcgccgaactggaccgatacgaccggcg
 220 225 230
 Arg Leu Ala Lys Arg Gln Arg Pro Gly Glu Arg Leu Asp Leu Ala Met Leu Ala Ala
 HSV TK >

gattcgccgcgtttacgggctgcttgccaatacgggtgcggtatctgcagggcggc 7040
 ctaagcggcgcaaatgcccgaacgaacgggttatgccacgccatagacgtcccgcgc
 235 240 245 250
 Ile Arg Arg Val Tyr Gly Leu Leu Ala Asn Thr Val Arg Tyr Leu Gln Gly Gly
 HSV TK >

gggtcgtggcgggaggattggggacagctttcggggacggccgtgccgccccagg 7095
 cccagcaccgccctcctaaccctgtcgaaagccctgccggcacggcggggtcc
 255 260 265 270
 Gly Ser Trp Arg Glu Asp Trp Gly Gln Leu Ser Gly Thr Ala Val Pro Pro Gln
 HSV TK >

gtgccgagccccagagcaacgcggggcccacgaccccatatcggggacacgttatt 7150
 cacgggtcggggtctcgttgcgcccggtgctgggggtatagccctgtgcaataa
 275 280 285
 Gly Ala Glu Pro Gln Ser Asn Ala Gly Pro Arg Pro His Ile Gly Asp Thr Leu Phe
 HSV TK >

taccctgtttcgggcccccgagttgctggcccccaacggcgacctgtacaacgtg 7205
 atgggacaaagcccggggggtcaacgaccgggggttgccgctggacatgttgcac
 290 295 300 305
 Thr Leu Phe Arg Ala Pro Glu Leu Leu Ala Pro Asn Gly Asp Leu Tyr Asn Val
 HSV TK >

ttgacctgggaccttgacgtcttggccaaacgcctccgtcccatgcacgtcttta
 aaacggaccggaacctgcagaaccggtttgcgaggcagggtacgtgcagaaat
 310 315 320 325
 Phe Ala Trp Ala Leu Asp Val Leu Ala Lys Arg Leu Arg Pro Met His Val Phe
 HSV TK

tcctggattacgaccaatcgcccgccggctgccgggacgccctgctgcaacttac
 aggacctaatgctggtttagcgggcgccgacggccctgcgggacgacgttgaatg
 330 335 340
 Ile Leu Asp Tyr Asp Gln Ser Pro Ala Gly Cys Arg Asp Ala Leu Leu Gln Leu Thr
 HSV TK

ctccgggatgatccagacccacgtcaccaccccaggctccataccgacgatctgc
 gaggccctactaggtctgggtgcagtgggtgggtccgaggtatggctgctagacg
 345 350 355 360
 Ser Gly Met Ile Gln Thr His Val Thr Thr Pro Gly Ser Ile Pro Thr Ile Cys
 HSV TK

gacctggcgcgacgtttgcccgggagatgggggaggctaactgaaacacggaag
 ctggaccgcgctgcaaacgggccctctacccccctccgattgactttgtgccttc
 365 370 375
 Asp Leu Ala Arg Thr Phe Ala Arg Glu Met Gly Glu Ala Asn *
 HSV TK

gagacaataccggaaggaacccgcgctatgacggcaataaaaagacagaataaaa
 ctctgttatggccttccttgggcgcgatactgccgttatTTTTTctgtcttatttt
 HSV TK poly(A) signal

cgcacgggtgttgggtcgtttgttcataaacgcgggggttcggtcccagggtggc
 gcgtgccacaaaccagcaaacagtatTTTgcgcccagccagggtcccgaccg
 HSV TK poly(A) signal

actctgtcgataccccaccgagaccccatTTGGGGccaatacgcccgctttcttc
 tgagacagctatgggggtggctctggggtaaccccggttatgcgggcgcaaagaag

cttttccccaccccccccccaagttcgggtgaaggcccagggtcgcagccaac
 gaaaaggggtgggggtggggggttcaagcccacttccgggtcccagcgctcggttg

gtcggggcggcaggccctgccatagccacggggcccggtgggttagggacgggggtc +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ cagccccgccgtccgggacgggtatcggtgcccggggcacccaatccctgccccag	7700
cccatgggggaatggtttatggttcgtgggggttattattttgggcgttgcggtg +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ gggtaccccttaccaaataccaagcaccaccaataataaaacccgcaacgcacc	7755
ggtcagtccacgactggactgagcagacagaccatgggtttttggatggcctggg +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ ccagtcagggtgctgacctgactcgtctgtctgggtacaaaaaacctaccggaccc	7810
catggaccgcatgtactggcgcgacacgaacaccgggcgtctgtggctgccaaac +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ gtacctggcgtagacacgctgtgcttgtggcccgacagacaccgacgggttg	7865
acccccgacccccaaaaaccaccgcgcggaatttctggcgccgcccggacgaactaa +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ tgggggctgggggtttttggtggcgcgccctaaagaccgcggcggcctgcttgatt	7920
acctgactacggcatctctgcccccttcttcgctggtagaggagcgctttttgttt +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ tggactgatgccgtagagacggggaagaagcgaccatgctcctcgcgaaaacaaa	7975
tgtattggtcaccacggccgagtttccgcgggaccccgccagatcaagctagag +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ acataaccagtggtgcccggctcaaaggcgccctgggggcccgtctagttcgatctc	8030
tctgaaagaaaaaacacaatctctgtttgccaacgcatttggctaccctgccact +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ agactttcttttttgtgttagagacaaacggttgcgtaaaccgatgggacgggtga	8085
cacaccattcaggcgcctggccgcgtgaatttgattgggtgaacacaccgactaca +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ gtgtggtaagtccgcggacccggcgcaactaaactaaccacttgtgtggctgatgt	8140
acgacggtttctgttctgccctgcgcgattgattatcaaaccgtgatcagttgtgc +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ tgctgccaaagcaagacgggacgcgctaactaatagtttggcactagtcacacg	8195
accacgcatgaccgtaaagtctcgctgatggcagccgattatgaaaatcagctc +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ tgggtgcgctactggcatttcaagcgcactaccgtcggctaatacttttagtcgag	8250
gacgagttttccctcgatgcgcccatgtcgcacatgaaaactatcaatgggcta +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++ ctgctcaaaaggagctacgcgggtaacagcgtgtacttttgatagttaccgat	8305

actacgttcgtggcgtggtgaaacatctgcaactgcgtaacaacagcttcggcgg	8360
tgatgcaagcaccgcaccactttgtagacgttgacgcattgttgtcgaagccgcc	
cgtggacatggtgatcagcggcaatgtgccgcaggggtgccgggttaagtctctcc	8415
gcacctgtaccactagtcgccgttacacggcgtcccacggcccaattcaagaagg	
gcttcactggaagtcgcggtcggaaccgtattgcagcagctttatcatctgccgc	8470
cgaagtgaccttcagcggcagccttggcataacgtcgtcgaatatagtagacggcg	
tggacggcgcacaaatcgcgcttaacggtcaggaagcagaaaaccagttttagg	8525
acctgccgcgtgttttagcgcgaattgccagtccttcgtcttttgggtcaaacatcc	
ctgtaactgcgggatcatggatcagctaatttccgcgctcggcaagaaagatcat	8580
gacattgacgccctagtagctcgattaaaggcgcgagccgttctttctagta	
gccttgctggcgcgtttcttccagcagccagccctgcgtgatgtcaccattgaag	8635
cggaacgaccgcgcaaagaaggtcgtcggtcgggacgcactacagtggtaacttc	
agttcaacgctgtttgcgcatgaactggaccgatcgtggcaaaacgcgtgcgtca	8690
tcaagttgcgacaacgcgtacttgacctgggctagcacccgttttgcgcacgcagt	
tatactgactgaaaacgcccgcaccgttgaagctgccagcgcgctggagcaaggc	8745
atatgactgacttttgcgggctggcaacttcgacggtcgcgcgacctcgttccg	
gacctgaaacgtatgggcgagttgatggcggagtctcatgcctctatgcgcgatg	8800
ctggactttgcatacccgctcaactaccgcctcagagtacggagatacgcgctac	
atttcgaaatcacccgtgccgcaaattgacactctggtagaaatcgtcaaagctgt	8855
taaagcttttagtggcacggcggtttaactgtgagaccatcttttagcagtttcgaca	
gattggcgacaaagggtggcgtacgcatgaccggcggcggtttggcggctgtatc	8910
ctaaccgctgtttccaccgcatgcgtactggccgcgcctaaaccgccgacatag	
SapI BspQI	
gtcgcgctgatcccggaagagctgggtgcctgccgtacagcaagctgtcgctgaac	8965
cagcgcgactagggccttctcgaccacggacggcatgtcgttcgacagcgacttg	

aatatgaagcaaaaacaggtattaaagagactttttacgtttgtaaaccatcaca 9020
 ttatacttcgtttttgtccataatttctctgaaaaatgcaaacatttggtagtgt

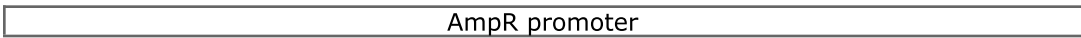
Bsu36I**I-CeuI**

aggagcaggacagtgcctagcttatcggatcttcgctaccttaggaccgttatag 9075
 tcctcgtcctgtcacggatcgaatagcctagaagcgatggaatcctggcaatatc

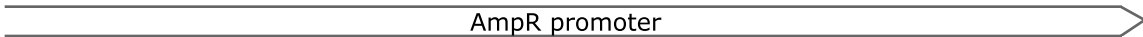
ttagcggccgccatgggtcatagctgtttgacgtcagggtggcacttttcggggaaa 9130
 aatcgccggcggtaccagtatcgacaaactgcagttccaccgtgaaaagccccctt


 Kozak sequence

tgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccg 9185
 acacgcgccttggggataaacaataaaaagatttatgtaagtttatacataggc


 AmpR promoter
SspI

ctcacgagacaataaccctgataaatgcttcaataatatgaaaaaggaagagta 9240
 gagtgctctgttattgggactatttacgaagtattataactttttccttctcat


 AmpR promoter

tgagtattcaacatttccgtgtcgccttattcccttttttgcggcattttgcct 9295
 actcataagttgtaaaggcacagcggaataagggaacaaacgccgtaaaacgga

tcctgtttttgctcaccagaaacgctgggtgaaagtaaaagatgctgaagatcag 9350
 aggacaaaaacgagtggttctttgcgaccactttcattttctacgacttctagtc

ttgggtgcacgagtggttacatcgaactggatctcaacagcggttaagatccttg 9405
 aaccacgtgctcacccaatgtagcttgacctagagttgtcgccatttctaggaa

agagttttcgccccgaagaacgttttccaatgatgagcacttttaagttctgct 9460
 tctcaaaagcggggcttcttgcaaaaggttactactcgtgaaaatttcaagacga


 ISS

Genomic map of the *hprt* gene region on chromosome 6p21.3, showing the location of the ISS (Interspersed Sequence) and the positions of the 10,010 bp and 9515 bp markers.

The map displays the DNA sequence (A, T, C, G) and the positions of the markers (10,010 bp and 9515 bp) relative to the ISS. The ISS is indicated by a grey bar with the label "ISS" below it. The markers are indicated by vertical lines with labels "10,010" and "9515" to their right.

The DNA sequence is shown in a 100 bp window, with the ISS located between the 10,010 bp and 9515 bp markers. The sequence is as follows:

10,010 bp: `gtctcgcggtatcattgcagcactggggggccagatggtaagcccctccngtatcg`

9515 bp: `atgtggcgcggtatttatcccggtattgacgcccgggcaagagcaactcggtcgccgc`

ISS: `atacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatc`

9570 bp: `tatgtgataagagtcttactgaaccaactcatgagtggtcagtgctcttttcgtag`

ISS: `ttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtga`

9625 bp: `aatgcctaccgtactgtcattctcttaatacgtcacgacgggtattgggtactcact`

ISS: `taacactgcggccaaacttacttctgacaacgatcggaggaccgaaggagctaacc`

9680 bp: `attgtgacgccggttgaatgaagactggttgctagcctcctggcttcctcgattgg`

ISS: `gcttttttgcacaacatgggggatcatgtaactcgccctgatcgttgggaaccgg`

9735 bp: `cgaaaaaacgtgttgtagcccttagtacattgagcgggaactagcaacccttggcc`

ISS: `agctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaat`

9790 bp: `tcgacttacttcggtatgggtttgctgctcgcactgtggtgctacggacatcgta`

ISS: `ggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccgg`

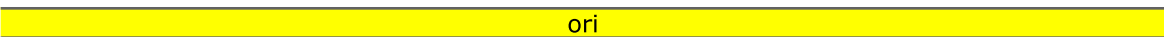
9845 bp: `ccgttggtgcaacgcgtttgataattgaccgcttgatgaatgagatcgaagggcc`

ISS: `caacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgct`

9900 bp: `gttggttaattatctgacctacctccgcctatttcaacgtcctggtgaagacgcga`

9955 bp: `cggcccttccggctggctgggtttattgctgataaatctggagccggtgagcgtgg`

10,010 bp: `gccgggaaggccgaccgaccaataacgactatttagacctcggccactcgcacc`

tagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagat	10,065
atcaatagatgtgctgccccctcagtcggttgatacctacttgctttatctgtcta	
cgctgagataggtgcctcactgattaaagcattggtaactgtcagaccaagtttac	10,120
gcgactctatccacggagtgactaattcgttaaccattgacagtctgggttcaaatg	
tcatatatacttttagattgatttaaacttcatttttaatttaaaggatctagg	10,175
agtatatatgaaatctaactaaattttgaagtaaaaattaaattttcctagatcc	
tgaagatcctttttgataatctcatgacccaaatcccttaacgtgagttttcgtt	10,230
acttctaggaaaaactattagagtactggtttttagggaattgcactcaaagcaa	
ccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttt	10,285
ggtgactcgcagtcctggggcatcttttctagtttcctagaagaactctaggaaaa	
	 >
tttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtag	10,340
aaagacgcgcattagacgacgaacgtttggttttttggtggcgatggctcgccacc	
	 >
tttggttgccggatcaagagctaccaactctttttccgaaggtaactggcttcag	10,395
aaacaaacggcctagttctcgatggttgagaaaaaggcttccattgaccgaagtc	
	 >
cagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccac	10,450
gtctcgcgtctatggtttatgacaggaagatcacatcggcatcaatccggtggtg	
	 >
ttcaagaactctgtagcaccgcctacatacctcgctctgctaatacctgttaccag	10,505
aagttcttgagacatcgtggcggtatgtatggagcgagacgattaggacaatggtc	
	 >
tggctgctgccagtgggcgataagtcgtgtcttaccgggttggaactcaagacgata	10,560
accgacgacggtcaccgctattcagcacagaatggcccaacctgagttctgctat	
	 >

agcttggagcgaacgacctacaccgaactgagatacctacagcggtgagctatgag
+ + + + +
tcgaacctcgcttgctggatgtggcttgactctattggatgtcgcactcgatactc

ori >

10,670

aaagcgccacgcttcccgaaggaggagaagggcggacaggtatccggtaagcggcag
+-----+-----+-----+-----+-----+-----+-----+-----+
tttcgcggtgcgaagggtctccctctttccgcctgtccataggccattcgcgcgtc

ori >

10,725

ggtcggaaacaggagagcgcacgaggaggagcttccagggggaaacgcctggatatctt
 ctagccttgtcctctcgcgtgctccctcgaagggtcccccctttgcggaaccatagaa
 ori

tatagtcctgtcggggttccgccacctctgacttgagcgctcgatttttgtgatgct
+-----+-----+-----+-----+-----+-----+-----+-----+
atatcaggacagcccaaagcgggtggagactgaactcgcagctaaaaaacactacga

ori >

cgtcaggggggcggaacctatggaaaaacgccagcaacgcggcccttttttacggtt
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
gcagtccccccgcctcggataaccttttttgcggtcgtttgcgccggaaaaaatgccaa

ori

10,890

`ctggccttttgcacggaccgaacaactccgccgcgcgcgctccgtgtaggccggt`

+-----+-----+-----+-----+-----+-----+-----+-----+

`gaccggaaaacgtgcctggcttggttagggcggcgcgcgagggcacatccggcca`

10,945

g c g g g c g g c c c c g t a g c g c a a g g g a g g g c g g g a a a g g a a g g g g c g g g a c a c a a g g
+ + + + +
c g c c c g c c g g g g c a t c g c g t t c c c t c c c g c c c t t t c c t t c c c g c c c t g t g t t c c 11,000

g c g a a t c t a t a a a g g g c g t c a c t c a g c c a g t t c t c t c c t c a g a a g c g c c g a g a g c
+
c g c t t a g a t a t t t c c c g c a g t g a g t c g g t c a a g a g a g g a g t c t t c g c g g c t c t c g

11,055

g c g a c c g g g a c g g t t g g a g a a g a a g g t g g c t c c c g g a a g g g g g a g a g a c a a a c t g
+
c g c t g g c c c t g c c a a c c t c t t c t t c c a c c g a g g g c c t t c c c c c t c t c t g t t t g a c

11,110

ccgtaacctctgccggttcaggatctgcgacctgcaggctcctcgccatggaccctg
 11,165
 ggcatctggagacggcaagtcctagacgctggacgtccaggagcgggtacctgggac
 DTA

atgatgttgttgattcttctaaatcttttgtgatggaaaacttttcttcgtacca
 11,220
 tactacaacaactaagaagatttagaaaacactaccttttgaaaagaagcatggt
 5 10 15 20
 Asp Asp Val Val Asp Ser Ser Lys Ser Phe Val Met Glu Asn Phe Ser Ser Tyr His
 DTA

cgggactaaacctgggttatgtagattccattcaaaaagggtatacaaaaagccaaa
 11,275
 gccctgatttggaccaatacatctaaggtaagttttccatatgttttcggtttt
 25 30 35
 Gly Thr Lys Pro Gly Tyr Val Asp Ser Ile Gln Lys Gly Ile Gln Lys Pro Lys
 DTA

tctggtacacaaggaaattatgacgatgattggaaagggttttatagtaccgaca
 11,330
 agaccatgtgttcctttaatactgctactaacctttcccaaaatatcatggctgt
 40 45 50 55
 Ser Gly Thr Gln Gly Asn Tyr Asp Asp Asp Trp Lys Gly Phe Tyr Ser Thr Asp
 DTA

ataaatacgacgctgcgggatactctgtagataatgaaaaccgctctcttgaaa
 11,385
 tatttatgctgcgacgccctatgagacatctattacttttgggcgagagacctt
 60 65 70 75
 Asn Lys Tyr Asp Ala Ala Gly Tyr Ser Val Asp Asn Glu Asn Pro Leu Ser Gly Lys
 DTA

agctggaggcgtgggtcaaagtgacgtatccaggactgacgaagggttctcgacta
 11,440
 tcgacctccgcaccagtttcaactgcataggctcctgactgcttccaagagcgtgat
 80 85 90
 Ala Gly Gly Val Val Lys Val Thr Tyr Pro Gly Leu Thr Lys Val Leu Ala Leu
 DTA

aaagtggataatgccgaaactattaagaaagagttagggtttaagtctcactgaac
 11,495
 tttcacctattacggctttgataattctttctcaatccaaattcagagtgacttg
 95 100 105 110
 Lys Val Asp Asn Ala Glu Thr Ile Lys Lys Glu Leu Gly Leu Ser Leu Thr Glu
 DTA

cgttgatggagcaagtcggaacggaagagtttatcaaaagggttcggtgatgggtgc 11,550
gcaactacctcgttcagccttgccttctcaaatagttttccaagccactaccacg

115 120 125 130
Pro Leu Met Glu Gln Val Gly Thr Glu Glu Phe Ile Lys Arg Phe Gly Asp Gly Ala
DTA

ttcgcgtgtagtgctcagccttcccttcgctgaggggagttctagcgttgaatat 11,605
aagcgcacatcacgagtcggaagggaagcgactccctcaagatcgcaacttata

135 140 145
Ser Arg Val Val Leu Ser Leu Pro Phe Ala Glu Gly Ser Ser Ser Val Glu Tyr
DTA

attaataactgggaacaggcgaaagcggttaagcgtagaacttgagattaattttg 11,660
taattattgacccttgctcgccttctcgcaattcgcatcttgaactctaattaaaac

150 155 160 165
Ile Asn Asn Trp Glu Gln Ala Lys Ala Leu Ser Val Glu Leu Glu Ile Asn Phe
DTA

aaaccctgggaaaacgtggccaagatgcgatgtatgagtatatggctcaagcctg 11,715
tttgggcaccttttgcaccggttctacgctacatactcatataccgagttcggac

170 175 180 185
Glu Thr Arg Gly Lys Arg Gly Gln Asp Ala Met Tyr Glu Tyr Met Ala Gln Ala Cys
DTA

tgcaggaaatcgtgtcaggcgatctctttgtgaagggaaccttacttctgtggtgt 11,770
acgtccttttagcacagtcgctagagaaacacttccttggaatgaagacaccaca

190
Ala Gly Asn Arg Val Arg Arg Ser Leu Cys Glu Gly Thr Leu Leu Leu Trp Cys
DTA (in frame with DTA)

gacataattggacaaactacctacagagatttaaagctctaagggtaaatataaaa 11,825
ctgtatttaacctgtttgatggatgtctctaaatttcgagattccatttataattt

Asp Ile Ile Gly Gln Thr Thr Tyr Arg Asp Leu Lys Leu *
(in frame with DTA) small t intron

tttttaagtgtataatgtgttaaactactgattctaattgtttgtgtatttttaga 11,880
aaaaattcacatattacacaatttgatgactaagattaacaaacacataaaaatct

small t intron

ttccaacctatggaactgatgaatgggagcagtggtggaatgcagatccactagg 11,935
aagggttggaataccttgactacttaccctcgtcaccaccttacgtctaggtgatcc

PsiI

atctaacttggtttattgcagcttataatggttacaaataaagcaatagcatcaca
tagattgaacaaataacgtcgaatattaccaatgtttatttcgttatcgtagtgt

11,990

SV40 poly(A) signal

aatttcacaaataaagcatttttttctactgcatttctagttgtggtttgtccaaac
ttaagtgtttatttcgtaaaaaaagtgcgtaagatcaacaccaaacaggtttg

12,045

SV40 poly(A) signal

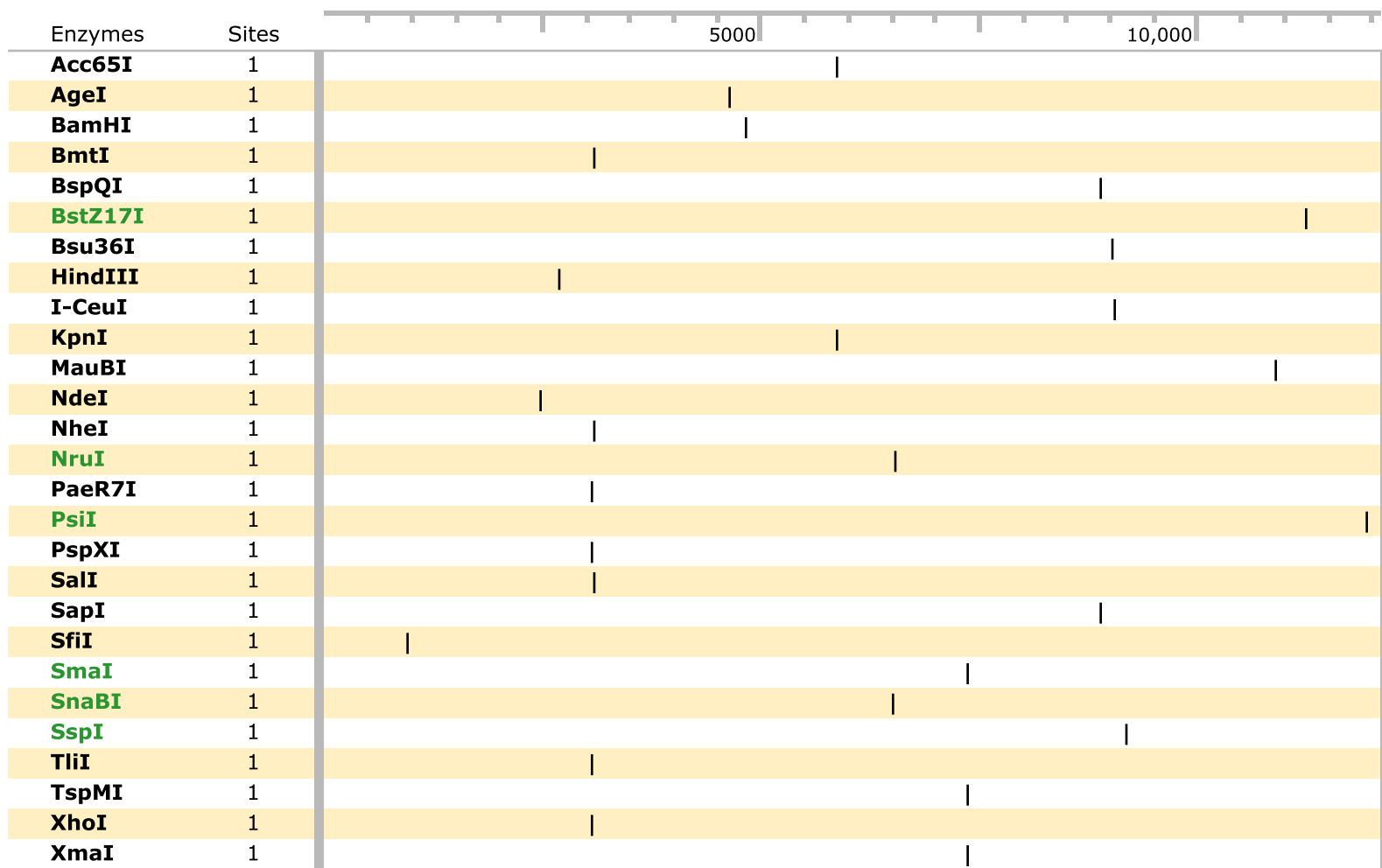
tcatcaatgtatcttatcatgtctggatcgtagttctagag 3'
agtagttacatagaatagtagacacctagcatcaagatctc 5' 12,086

SV40 poly(A) signal

tagtacagacctagcatcaagatctc

pDTA-TK_25Rev

| Enzymes | Sites | |
|----------------|-------|--------|
| Acc65I | 1 | 5892 |
| AgeI | 1 | 4662 |
| BamHI | 1 | 4857 |
| BmtI | 1 | 3126 |
| BspQI | 1 | 8921 |
| BstZ17I | 1 | 11,261 |
| Bsu36I | 1 | 9060 |
| HindIII | 1 | 2702 |
| I-CeuI | 1 | 9064 |
| KpnI | 1 | 5896 |
| MauBI | 1 | 10,923 |
| NdeI | 1 | 2502 |
| NheI | 1 | 3122 |
| NruI | 1 | 6576 |
| PaeR7I | 1 | 3099 |
| PsiI | 1 | 11,959 |
| PspXI | 1 | 3099 |
| SalI | 1 | 3105 |
| SapI | 1 | 8921 |
| SfiI | 1 | 986 |
| SmaI | 1 | 7392 |
| SnaBI | 1 | 6524 |
| SspI | 1 | 9222 |
| TliI | 1 | 3099 |
| TspMI | 1 | 7390 |
| XhoI | 1 | 3099 |
| XmaI | 1 | 7390 |



| Feature | Location | Size (bp) |  |  | Type |
|----------------------------|-----------------|-----------|--|---|--------------|
| ✓ Homology Arm LEFT | 1 .. 1000 | 1000 |  |  | misc_feature |
| ✓ GSG Bridge | 1001 .. 1012 | 12 |  |  | misc_feature |
| Pair 1 | 1001 .. 1012 | 12 |  |  | misc_feature |
| Pair 2 | 1001 .. 1012 | 12 |  |  | misc_feature |
| Pair 02 | 1001 .. 1012 | 12 |  |  | misc_feature |
| Pair 03 | 1001 .. 1012 | 12 |  |  | misc_feature |
| Pair 05 | 1001 .. 1012 | 12 |  |  | misc_feature |
| Reading Frame | 1001 .. 1012 | 12 |  |  | misc_feature |
| /note | = Reading Frame | | | | |
| Reading Frame | 1001 .. 1012 | 12 |  |  | misc_feature |
| /note | = Reading Frame | | | | |
| Reading Frame | 1001 .. 1012 | 12 |  |  | misc_feature |
| /note | = Reading Frame | | | | |
| Reading Frame | 1001 .. 1012 | 12 |  |  | misc_feature |
| /note | = Reading Frame | | | | |
| Reading Frame | 1001 .. 1012 | 12 |  |  | misc_feature |
| /note | = Reading Frame | | | | |
| Reading Frame | 1001 .. 1012 | 12 |  |  | misc_feature |
| /note | = Reading Frame | | | | |
| Reading Frame | 1001 .. 1012 | 12 |  |  | misc_feature |
| /note | = Reading Frame | | | | |
| Reading Frame | 1001 .. 1012 | 12 |  |  | misc_feature |
| /note | = Reading Frame | | | | |
| Reading Frame | 1001 .. 1012 | 12 |  |  | misc_feature |
| /note | = Reading Frame | | | | |
| Seq #64 | 1001 .. 1012 | 12 |  |  | misc_feature |
| /note | = Seq #64 | | | | |

| Feature | Location | Size (bp) |  |  | Type |
|------------------------|---|-----------|--|---|--------------|
| Seq eGFP_F | 1001 .. 1012 | 12 |  |  | misc_feature |
| /note | = Seq eGFP_F | | | | |
| ✓ P2A | 1013 .. 1069 | 57 |  |  | CDS |
| /product | = 2A peptide from porcine teschovirus-1 polyprotein | | | | |
| /note | = Eukaryotic ribosomes fail to insert a peptide bond between the Gly and Pro residues, yielding se polypeptides. | | | | |
| /translation | = ATNFSLLKQAGDVEENPGP
19 amino acids = 2.0 kDa | | | | |
| ✓ tdTomato | 1070 .. 2500 | 1431 |  |  | CDS |
| /product | = tandem dimeric (pseudo-monomeric) derivative of DsRed (Shaner et al., 2004) | | | | |
| /note | = mammalian codon-optimized | | | | |
| /translation | = MVSKGEEVIKEFMRFKVRMEGSMNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSK.
ADIPDYKKLSFPEGFKWERVMNFEDGGLVTVTQDSSLQDGTLIYKVKMRGTNFPDPGPVMQKKTMGWEAS
RDGVLKGEIHQALKLKDGGHYLVEFKTIYMAKKPVQLPGYYYVDTKLDITSHNEDYTIVEQYERSEGRHHLFL
STGSGSSGTASSEDNNMAVIKEFMRFKVRMEGSMNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWI
FMYGSKAYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGLVTVTQDSSLQDGTLIYKVKMRGTNFPDPGPV
MGWEASTERLYPRDGV LKGEIHQALKLKDGGHYLVEFKTIYMAKKPVQLPGYYYVDTKLDITSHNEDYTIVEC
GRHHLFLYGMDELYK* | | | | |
| | 476 amino acids = 54.2 kDa | | | | |
| ✓ IVS | 2513 .. 2810 | 298 |  |  | misc_feature |
| /note | = /ApEinfo_label=IVS - /ApEinfo_fwdcolor=#ff2ddc - /ApEinfo_revcolor=#b3387e - /vntifkey="21
/label=IVS - 2024..2321 | | | | |
| ✓ Bovine\polyA | 2811 .. 3092 | 282 |  |  | polyA_site |
| /note | = /ApEinfo_label=Bovine pA - /ApEinfo_fwdcolor=#3ff9ff - /ApEinfo_revcolor=#348180 - /vntifke
- /label=Bovine\polyA - 2322..2603 | | | | |
| BpA\min | 2823 .. 3029 | 207 |  |  | polyA_signal |
| /note | = /ApEinfo_label=BpA min - /ApEinfo_fwdcolor=cyan - /ApEinfo_revcolor=#0cb0ab - /vntifkey="2
/label=BpA\min - 2334..2540 | | | | |
| ✓ loxP\sequence | 3152 .. 3185 | 34 |  |  | misc_recomb |
| /note | = /ApEinfo_label=LoxP - /ApEinfo_fwdcolor=#f94fcb - /ApEinfo_revcolor=#c700bd - /vntifkey="8
/label=loxP\sequence - 2663..2696 | | | | |
| ✓ PGK\polyA | 3214 .. 3669 | 456 |  |  | polyA_signal |
| /note | = /ApEinfo_label=PGK pA - /ApEinfo_fwdcolor=#12f0ff - /ApEinfo_revcolor=#026f7e - /vntifkey="2
/label=PGK\polyA - complement(2725..3180) | | | | |

| Feature | Location | Size (bp) |  |  | Type |
|-----------------------------------|--|-----------|--|---|--------------|
| Puromycin resistance ge... | 3670 .. 4269 | 600 |  |  | CDS |
| /note | = /vntifkey="4" - /label=Puro - /note="Puromycin resistance gene" - complement(3181..3780) | | | | |
| ✓ puro | 3670 .. 4269 | 600 |  |  | gene |
| /note | = /ApEinfo_label=Puro - /ApEinfo_fwdcolor=#b8b5b3 - /ApEinfo_revcolor=#686363 - /vntifkey='complement(3181..3780) | | | | |
| ✓ PGK\promoter | 4276 .. 4755 | 480 |  |  | promoter |
| /note | = /ApEinfo_label=PGK Prom - /ApEinfo_fwdcolor=#ffce10 - /ApEinfo_revcolor=#b79000 - /vntifkey='complement(3787..4293) - /label=PGK\promoter - complement(3787..4293) | | | | |
| ✓ loxP\sequence | 4811 .. 4844 | 34 |  |  | misc_recomb |
| /note | = /ApEinfo_label=LoxP - /ApEinfo_fwdcolor=#f94fcb - /ApEinfo_revcolor=#c700bd - /vntifkey="8" - /label=loxP\sequence - 4386..4419 | | | | |
| ✓ Homology Arm RIGHT | 4876 .. 5875 | 1000 |  |  | misc_feature |
| ✓ Mini-TK promoter | 6144 .. 6206 | 63 |  |  | promoter |
| /note | = minimal herpes simplex virus (HSV) thymidine kinase promoter | | | | |
| ✓ HSV TK 5'-UTR | 6228 .. 6267 | 40 |  |  | 5'UTR |
| /note | = 5' untranslated region from the herpes simplex virus thymidine kinase gene | | | | |
| ✓ HSV TK | 6285 .. 7415 | 1131 |  |  | CDS |
| /product | = herpes simplex virus thymidine kinase | | | | |
| /note | = confers sensitivity to nucleoside analogs such as acyclovir or ganciclovir | | | | |
| /translation | = MASYPGHQHASAFDQAARSRGHSNRRTALRPRRQQEATEVRPEQKMPTLLRVYIDGPHGMGKTTTTQLLVA
DIVYVPEPMTYWRVLGASETIANIYTTQHRLDQGEISAGDAAVVMTSAQITMGMPYAVTDAVLAPHIGGEAG
PPALTLIFDRHPAAALLCYPAARYLMGSMTPQAVLAFVALIPPTLPGTNIVLGALPEDRHRIDRLAKRQRPGERLD
AIRRVYGLLANTVRYLQGGGSWREDWGQLSGTAVPPQGAEPQSNAGPRPHIGDTLFTLFRAPELLAPNGDL
ALDVLAKRLRPMHVFILDYDQSPAGCRDALLQLTSGMIQTHVTPGSIPTICDLARTFAREMGEAN* | | | | |
| | 376 amino acids = 40.9 kDa | | | | |
| ✓ HSV TK poly(A) signal | 7457 .. 7505 | 49 |  |  | polyA_signal |
| /note | = herpes simplex virus thymidine kinase polyadenylation signal (Cole and Stacy, 1985) | | | | |
| ✓ Kozak sequence | 9082 .. 9091 | 10 |  |  | regulatory |
| /note | = vertebrate consensus sequence for strong initiation of translation (Kozak, 1987) | | | | |
| ✓ AmpR promoter | 9135 .. 9239 | 105 |  |  | promoter |
| /gene | = <i>bla</i> | | | | |

| Feature | Location | Size (bp) |  |  | Type |
|------------------------------|--|-----------|--|---|--------------|
| ✓ ISS | 9424 .. 9802 | 379 |  |  | CDS |
| /note | = immunostimulatory sequence from the AmpR gene; contains unmethylated CpG dinucleotides in context of 5'-AACGTT-3' (Sato et al., 1996) | | | | |
| ✓ ori | 10,273 .. 10,861 | 589 |  |  | rep_origin |
| /direction | = RIGHT | | | | |
| /note | = high-copy-number ColE1/pMB1/pBR322/pUC origin of replication | | | | |
| ✓ DTA | 11,165 .. 11,737 | 573 |  |  | CDS |
| /gene | = <i>Corynebacterium diphtheriae tox</i> | | | | |
| /product | = diphtheria toxin fragment A | | | | |
| /translation | = DDVVDSSKSFVMENFSSYHGTPGYVDSIQKGIQPKPSGTQGNYDDDWKGFYSTDNKYDAAGYSVDNENI
GGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYIN
AKALSVELEINFETRGRGQDAMYEYMAQACAGNRVRR | | | | |
| | 191 amino acids = 21.0 kDa | | | | |
| ✓ small t intron | 11,814 .. 11,879 | 66 |  |  | intron |
| /note | = SV40 (simian virus 40) small t antigen intron | | | | |
| ✓ SV40 poly(A) signal | 11,940 .. 12,074 | 135 |  |  | polyA_signal |
| /note | = SV40 polyadenylation signal | | | | |

| Primer | Length |  | Binding Sites |  | Tm |
|-------------------------------|---|---|------------------------------|--|--------------|
| ✓ HLeft_nonoverlap_Fwd | 18-mer | | 1 .. 18 |  | 56°C |
| /sequence | = TCTGCCTTTCCTCCTCCC
61% GC / 5288.5 Da | | | | |
| HLeft_overlap_Rev | 43-mer | | 983 .. 1037 |  | 66°C |
| /sequence | = GCTTTAACAGAGAGAAGTTCGTGGCGACAGCAGTGGCCTCCTG
56% GC / 13,308.7 Da | | | | |
| HLeft_nonoverlap_Rev | 18-mer | | 983 .. 1000 |  | 58°C |
| /sequence | = GACAGCAGTGGCCTCCTG
67% GC / 5500.6 Da | | | | |
| TdT_overlap_Fwd | 43-mer | | 1045 .. 1087 |  | 74°C |
| /sequence | = AGACGTGGAAGAAAACCCCGGTCCTATGGTGAGCAAGGGCGAG
58% GC / 13,400.8 Da | | | | |
| ✓ TdT_nonoverlap_Fwd | 18-mer | | 1070 .. 1087 |  | 58°C |
| /sequence | = ATGGTGAGCAAGGGCGAG
61% GC / 5653.7 Da | | | | |
| TdT_Fwd1 | 20-mer | | 1083 .. 1102 |  | 56°C |
| /sequence | = GCGAGGAGgtcatcaaagag
55% GC / 6240.1 Da | | | | |
| TdT_Rev7 | 19-mer | | 1738 .. 1754
2464 .. 2482 | 
 | 43°C
56°C |
| /sequence | = catgccgtacaggaacagg
58% GC / 5846.9 Da | | | | |
| TdT_Fwd6 | 19-mer | | 1794 .. 1812 |  | 55°C |
| /sequence | = cctccgaggacaacaacat
53% GC / 5750.8 Da | | | | |
| ✓ TdT_Rev1 | 19-mer | | 1804 .. 1822 |  | 55°C |
| /sequence | = gatgacggccatgttgttg
53% GC / 5874.9 Da | | | | |

| Primer | Length |  | Binding Sites |  | Tm |
|----------------------------------|--|---|------------------------------|--|--------------|
| TdT_Fwd7 | 20-mer | | 1809 .. 1828 |  | 56°C |
| /sequence | = acatggccgtcatcaaagag
50% GC / 6135.1 Da | | | | |
| ✓ TdT_Rev6 | 18-mer | | 2478 .. 2495 |  | 55°C |
| /sequence | = tgtacagctcgtccatgc
56% GC / 5450.6 Da | | | | |
| LoxP1_overlap_Rev | 60-mer | | 3149 .. 3208
4808 .. 4861 | 
 | 65°C
63°C |
| /sequence | = caagctgatccggaacccttaataataacttcgtataatgtatgctatacgaagtattag
35% GC / 18,455.1 Da | | | | |
| TdT_loxP-PuroR_PGK_Lox... | 37-mer | | 3149 .. 3185
4808 .. 4844 | 
 | 56°C
56°C |
| /sequence | = ataacttcgtataatgtatgctatacgaagtattag
27% GC / 11,400.6 Da | | | | |
| TdT_loxP_PuroR_PGK_Lox... | 60-mer | | 3149 .. 3185
4808 .. 4844 | 
 | 57°C
57°C |
| /sequence | = atggggatggaggcagaacgttcataacttcgtataatgtatgctatacgaagtattag
38% GC / 18,656.2 Da | | | | |
| PostLoxP4_Rev | 21-mer | | 3186 .. 3202
4845 .. 4865 | 
 | 51°C
56°C |
| /sequence | = agtggatccggaaCccttaat
48% GC / 6430.2 Da | | | | |
| PostLoxP1_Fwd1 | 23-mer | | 3186 .. 3208 |  | 58°C |
| /sequence | = attaagggtccggatcagcttg
48% GC / 7094.7 Da | | | | |
| Puro4_Fwd | 20-mer | | 3931 .. 3950 |  | 57°C |
| /sequence | = caggaggccttccatctgtt
55% GC / 6084.0 Da | | | | |

| Primer | Length |  | Binding Sites |  | Tm |
|------------------------------------|-------------------------|---|---------------|---|------|
| ✓ Puro4_Rev | 20-mer | | 3931 .. 3950 |  | 57°C |
| /sequence | = aacagatggaaggcctcctg | | | | |
| | 55% GC / 6151.1 Da | | | | |
| ✓ Puro5_Fwd | 20-mer | | 4104 .. 4123 |  | 57°C |
| /sequence | = gcgtgaggaagagttcttgc | | | | |
| | 55% GC / 6213.1 Da | | | | |
| Puro5_Rev | 20-mer | | 4104 .. 4123 |  | 57°C |
| /sequence | = gcaagaactcttcctcacgc | | | | |
| | 55% GC / 6022.0 Da | | | | |
| ✓ Puro2_Fwd | 20-mer | | 4114 .. 4133 |  | 58°C |
| /sequence | = gagtcttcgcagctcggtga | | | | |
| | 55% GC / 6164.1 Da | | | | |
| Puro2_Rev | 20-mer | | 4114 .. 4133 |  | 58°C |
| /sequence | = tcaccgagctgcaagaactc | | | | |
| | 55% GC / 6071.0 Da | | | | |
| ✓ Puro3_Fwd | 20-mer | | 4115 .. 4134 |  | 58°C |
| /sequence | = agttcttcgcagctcggtgac | | | | |
| | 55% GC / 6124.0 Da | | | | |
| Puro3_Rev | 20-mer | | 4115 .. 4134 |  | 58°C |
| /sequence | = gtcaccgagctgcaagaact | | | | |
| | 55% GC / 6111.0 Da | | | | |
| ✓ Puro1_Fwd and PuroPGK_... | 20-mer | | 4250 .. 4269 |  | 57°C |
| /sequence | = gtgggcttgactcggtcat | | | | |
| | 55% GC / 6155.1 Da | | | | |
| Puro1_Rev | 20-mer | | 4250 .. 4269 |  | 57°C |
| /sequence | = atgaccgagtacaagccac | | | | |
| | 55% GC / 6080.0 Da | | | | |

| Primer | Length |  | Binding Sites |  | Tm |
|--------------------------------|---|---|---------------|---|------|
| PuroPGK_Fwd3 | 20-mer | | 4266 .. 4285 |  | 58°C |
| /sequence | = tcatctgcaggtcgaaaggc
55% GC / 6142.1 Da | | | | |
| ✓ PuroPGK_Rev1 | 20-mer | | 4485 .. 4504 |  | 58°C |
| /sequence | = ctgagcaatggaagcgggta
55% GC / 6231.1 Da | | | | |
| ✓ PuroPGK_Fwd5 | 20-mer | | 4519 .. 4538 |  | 57°C |
| /sequence | = tgcacgagactagtgcagacg
55% GC / 6191.1 Da | | | | |
| PuroPGK_Rev3 | 20-mer | | 4549 .. 4568 |  | 58°C |
| /sequence | = tcgtgcaggacgtgacaaat
50% GC / 6166.1 Da | | | | |
| ✓ PuroPGK_Rev5 | 20-mer | | 4610 .. 4629 |  | 59°C |
| /sequence | = cttcgcgccaccttctactc
60% GC / 5939.9 Da | | | | |
| PreLoxP25_overlap20_Rev | 45-mer | | 4780 .. 4824 |  | 65°C |
| /sequence | = ctatacgaagttattaggtccctcgacctgcaggaattaattcta
40% GC / 13,794.0 Da | | | | |
| PreLoxP25_overlap15_Rev | 40-mer | | 4780 .. 4819 |  | 65°C |
| /sequence | = cgaagttattaggtccctcgacctgcaggaattaattcta
43% GC / 12,270.0 Da | | | | |
| PreLoxP20_overlap20_Rev | 40-mer | | 4785 .. 4824 |  | 64°C |
| /sequence | = ctatacgaagttattaggtccctcgacctgcaggaattaa
43% GC / 12,279.1 Da | | | | |
| PreLoxP20_overlap15_Rev | 35-mer | | 4785 .. 4819 |  | 64°C |
| /sequence | = cgaagttattaggtccctcgacctgcaggaattaa
46% GC / 10,755.1 Da | | | | |

| Primer | Length |  | Binding Sites |  | Tm |
|---------------------------------|--|---|---------------|---|------|
| PreLoxP17_overlap20_Rev | 37-mer | | 4788 .. 4824 |  | 65°C |
| /sequence | = ctatacgaagttattaggtccctcgacctgcaggaat
46% GC / 11,348.4 Da | | | | |
| PreLoxP17_overlap15_Rev | 32-mer | | 4788 .. 4819 |  | 65°C |
| /sequence | = cgaagttattaggtccctcgacctgcaggaat
50% GC / 9824.4 Da | | | | |
| PostLoxP2_25ntOLRev2 | 49-mer | | 4852 .. 4900 |  | 73°C |
| /sequence | = gcatggggatggaggcagaacgttcgtctagaactagtgatccggaaC
55% GC / 15,276.0 Da | | | | |
| ✓ PostLoxP2_Rev | 24-mer | | 4852 .. 4875 |  | 57°C |
| /sequence | = gtctagaactagtgatccggaaC
50% GC / 7401.9 Da | | | | |
| PostLoxP3_Rev | 24-mer | | 4853 .. 4875 |  | 54°C |
| /sequence | = agtctagaactagtgatccggaa
46% GC / 7425.9 Da | | | | |
| PostLoxP5_Rev | 24-mer | | 4856 .. 4875 |  | 52°C |
| /sequence | = tcgagtctagaactagtgatccg
50% GC / 7392.9 Da | | | | |
| PostLoxP6_Rev | 24-mer | | 4857 .. 4875 |  | 49°C |
| /sequence | = ctcgagtctagaactagtgatcc
50% GC / 7352.8 Da | | | | |
| PostLoxP1_Rev | 25-mer | | 4861 .. 4875 |  | 42°C |
| /sequence | = ccATActcgagtctagaactagtg
48% GC / 7666.1 Da | | | | |
| ✓ HARight_nonoverlap_Fwd | 20-mer | | 4876 .. 4895 |  | 58°C |
| /sequence | = gaacgttctgcctccatccc
60% GC / 5988.9 Da | | | | |

| Primer | Length |  | Binding Sites |  | Tm |
|---------------------------------|---|---|------------------------------|--|--------------|
| HARight_overlap_Fwd | 60-mer | | 4876 .. 4895 |  | 59°C |
| /sequence | = gacctataacttcgtatagcatatatacgaagttagaacgttctgcctccatccc
40% GC / 18,312.0 Da | | | | |
| ✓ HARight_nonoverlap_Rev | 18-mer | | 4993 .. 5005
5858 .. 5875 | 
 | 42°C
62°C |
| /sequence | = acctggctccgtgcacct
67% GC / 5411.5 Da | | | | |
| pDTA-TK_25Fwd | 49-mer | | 5851 .. 5899 |  | 74°C |
| /sequence | = ccagggcaggtgcacggagccaggttctaggaattctctagggtacctc
59% GC / 15,138.8 Da | | | | |
| ✓ pDTA-TK Backbone Fwd | 24-mer | | 5876 .. 5899 |  | 55°C |
| /sequence | = tctaggaattctctagggtacctc
46% GC / 7318.8 Da | | | | |
| TK-pA-R | 20-mer | | 7413 .. 7432 |  | 55°C |
| /sequence | = TTGTCTCCTTCCGTGTTTCA
45% GC / 6016.0 Da | | | | |
| /note | = Thymidine kinase polyA, reverse primer | | | | |
| Amp-R | 20-mer | | 9458 .. 9477 |  | 56°C |
| /sequence | = ATAATACCGCGCCACATAGC
50% GC / 6055.0 Da | | | | |
| /note | = Ampicillin resistance gene, reverse primer | | | | |
| pBR322ori-F | 20-mer | | 10,762 .. 10,781 |  | 56°C |
| /sequence | = GGGAAACGCCTGGTATCTTT
50% GC / 6148.1 Da | | | | |
| /note | = pBR322 origin, forward primer | | | | |
| SV40pA-R | 20-mer | | 11,977 .. 11,996 |  | 49°C |
| /sequence | = GAAATTTGTGATGCTATTGC
35% GC / 6162.1 Da | | | | |
| /note | = SV40 polyA, reverse primer | | | | |

| Primer | | Length |  | Binding Sites |  | Tm |
|-----------------------------|---|---|---|------------------|---|------|
| EBV-rev | | 20-mer | | 12,031 .. 12,050 |  | 52°C |
| /sequence | = | GTGGTTTGTCCAAACTCATC
45% GC / 6083.0 Da | | | | |
| /note | = | SV40 polyA terminator, reverse primer | | | | |
| ✓ pDTA-TK_25Rev | | 51-mer | | 12,061 .. 25 |  | 71°C |
| /sequence | = | GGAGTCTGGGAGGAGGAAAGGCAGActctagaactacgatccagacatgat
51% GC / 15,895.4 Da | | | | |
| pDTA-TK Backbone Rev | | 26-mer | | 12,061 .. 12,086 |  | 56°C |
| /sequence | = | ctctagaactacgatccagacatgat
42% GC / 7923.2 Da | | | | |

Description: This MUC5B reporter/excisable selection cassette was generated using HiFi Assembly. Intended to be a promoter fusion, it replaces the stop codon of MUC5B Exon 49, in frame, with a GSG bridge fused to a P2A sequence. Left and right homology arms were generated using PCR of genomic DNA from the FHS14030-3 iPSC (rs35705950 TT homozygote). The GSG-P2A sequence were ordered as a duplex from IDT. The TdTomato-Puro sequence extending to the 2nd loxP sequence was PCR amplified from the SPC-DV vector. The backbone vector was PCR amplified from the pDTA-TK plasmid. HiFi assembly was performed using 25 nucleotide overlap sequences. Final sequencing of the donor vector is embedded in this file under the alignment tab for Maxiprep Clone#2 and #7. Embedded chromatogram ab files are from MUC5B T/TtdTtdT Clone 62, Cre-excised Clone 26, which was the iPSC line referred to as MUC5B T/TtdTtdT.

Created: Feb 7, 2020

Last Modified: Jun 8, 2022

Accession Number:

Code Number:

Sequence Author: Sheng Si

DNA Type: Synthetic DNA

Laboratory Host Organism:

Bacterial Transformation Strain: DH5a™
Dam⁺ Dcm⁺ EcoKI⁺

Comments:

References: