

5' CTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGA 65
 3' GACCGACCGACCGTATTTATAGAGTGAGCGTTAGTTTAAAGTCGGCTATCGCCTTGCCCTTCCGCT

L A G W H K Y L T R N Q I Q P I A E R E G D

lacI

CTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTG 130
 GACCTCACGGTACAGGCCAAAAGTTGTTTGGTACGTTTACGACTTACTCCCCTAGCAAGGGTGAC

W S A M S G F Q Q T M Q M L N E G I V P T

lacI

CGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTG 195
 GCTACGACCAACGGTTGCTAGTCTACCGCGACCCGCGTTACGCGCGGTAATGGCTCAGGCCCGAC

A M L V A N D Q M A L G A M R A I T E S G L

lacI

CGCGTTGGTGCGGATATCTCGGTAGTGGGATACGACGATACCGAAGATAGCTCATGTTATATCCC 260
 GCGCAACCACGCCTATAGAGCCATCACCTATGCTGCTATGGCTTCTATCGAGTACAATATAGGG

R V G A D I S V V G Y D D T E D S S C Y I P

lacI

HpaI

GCCGTTAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGC 325
 CGGCAATTGGTGGTAGTTTGTCTAAAAGCGGACGACCCCGTTTGGTTCGCACCTGGCGAACGACG

P L T T I K Q D F R L L G Q T S V D R L L

lacI

AACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCAGTCTCACTGGTGAAAAGAAAA 390
 TTGAGAGAGTCCCGGTCCGCCACTTCCCGTTAGTCGACAACGGTTCAGAGTGACCACTTTTCTTTT

Q L S Q G Q A V K G N Q L L P V S L V K R K

lacI

ACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCGCGCGTGGCCGATTCATTAATGCAGCT 455
 TGGTGGGACCGCGGGTTATGCGTTTGGCGGAGAGGGGCGCGCAACC GGCTAAGTAATTACGTCGA

T T L A P N T Q T A S P R A L A D S L M Q L

lacI

GGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGACTCATGACCAAATCCCTTAACGTGAGT 520
 CCGTGCTGTCCAAAGGGCTGACCTTTCGCCCGTCACTGAGTACTGGTTTTAGGGAATTGCACTCA

A R Q V S R L E S G Q

lacI

MauBI

TACGCGCGCGTTCGTTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATC
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 ATGCGCGCGCAGCAAGGTGACTCGCAGTCTGGGGCATCTTTTCTAGTTTCTAGAAAGAACTCTAG

585

ori

CTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCACCGCTACCAGCGGTGGTTTGT
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GAAAAAAGACGCGCATTAGACGACGAACGTTTGT TTTT TTTGGTGCGATGGTTCGCCACCAAACA

650

ori

TTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACC
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 AACGGCCTAGTTCTCGATGGTTGAGAAAAAGGCTTCCATTGACCGAAGTCGTCTCGCGTCTATGG

715

ori

AAATACTGTTCTTCTAGTGTAGCCGTAGTTAGCCCACCACTTCAAGAACTCTGTAGCACCGCCTA
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TTTATGACAAGAAGATCACATCGGCATCAATCGGGTGGTGAAGTTCTTGAGACATCGTGCGGGAT

780

ori

CATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACC
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GTATGGAGCGAGACGATTAGGACAATGGTCACCGACGACGGTCACCGCTATTCAGCACAGAATGG

845

ori

GGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTTCGGGCTGAACGGGGGGTTTCGTG
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CCCAACCTGAGTTCTGCTATCAATGGCCTATTCCGCGTCGCCAGCCCGACTTGCCCCCAAGCAC

910

ori

CACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAG
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GTGTGTCGGGTCGAACCTCGCTTGTGGATGTGGCTTGA CTCTATGGATGTCGCACTCGATACTC

975

ori

AAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACA
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TTTGCGGGTGC GAAGGGCTTCCCTCTTTCCGCCTGTCCATAGGCCATTGCGCGTCCAGCCTTGT

1040

ori

GGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCCTGGTATCTTTATAGTCCTGTGGGTTTTCG
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CCTCTCGCGTGCTCCCTCGAAGGTCCCCCTTTGCGGACCATAGAAATATCAGGACAGCCCAAAGC

1105

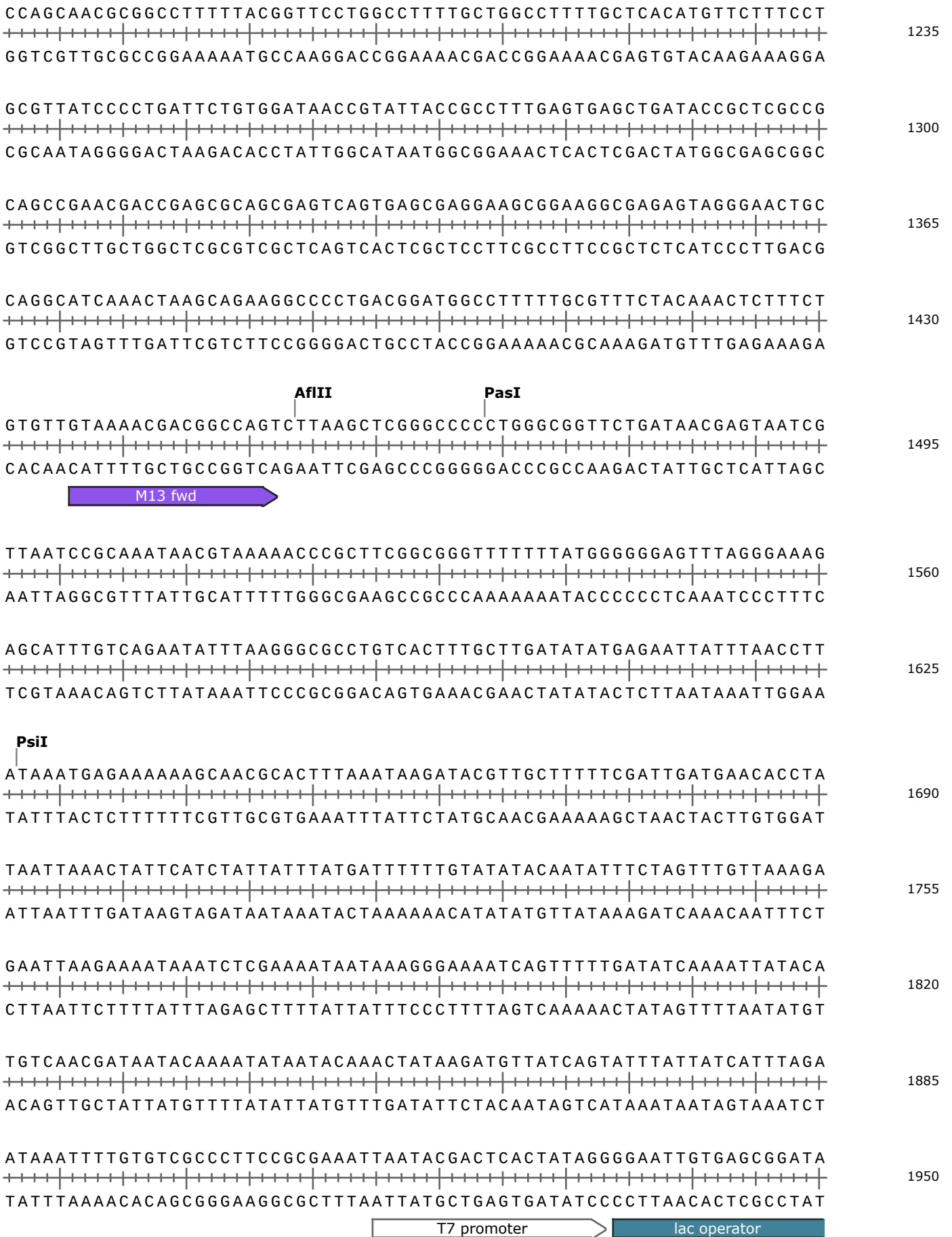
ori

DrdI

CCACCTCTGACTTGAGCGTTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGA AAAACG
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GGTGGAGACTGAACTCGCAGCTAAAACACTACGAGCAGTCCCCCGCCTCGGATACCTTTTGTG

1170

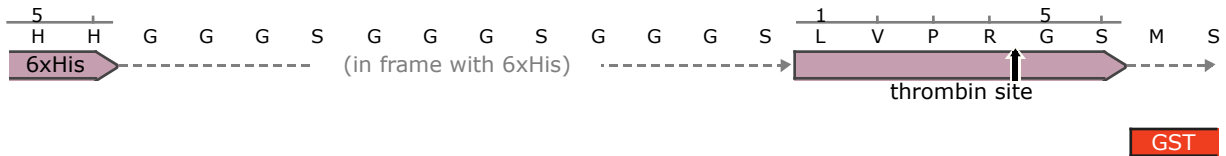
ori



XbaI **NdeI** **BfrBI** **NsiI**
 ACAATTCCCCTCTAGAAATAATTTTGTTTAACTTTTAGGAGATAAAACATATGCATCATCACCAT
 TGTTAAGGGGAGATCTTTATTAACAAATTGAAAATCCTCTATTTTGTATACGTAGTAGTGTA



CATCACGGCGGGCGGTAGCGGGCGGGTTCAGGTGGCGGCTCATTAGTTCGCGTGGTTCCATGTC
 GTAGTGCCGCGCCATCGCCGCGCCAAGTCCACCGCCGAGTAATCAAGGCGCACCAAGGTACAG



StuI
 CCCAATCCTCGGCTACTGGAAGATCAAAGGCCTCGTGCAGCCCACTCGACTTCTTTTGAATATC
 GGGTTAGGAGCCGATGACCTTCTAGTTTCCGGAGCACGTCGGGTGAGCTGAAGAAAACCTTATAG
 P I L G Y W K I K G L V Q P T R L L L E Y
 (in frame with thrombin site)



SapI **BspQI**
 TTGAAGAAAAATATGAAGAGCATTGTATGAGCGCGATGAAGGTGATAAATGGCGAAACAAAAG
 AACTTCTTTTATACTTCTCGTAAACATACTCGCGCTACTTCCACTATTTACCGCTTTGTTTTTC
 L E E K Y E E H L Y E R D E G D K W R N K K
 (in frame with thrombin site)



TTTGAATTGGGTTTGGAGTTTCCAATCTTCCTTATTATATTGATGGTGATGTTAAATTAACACA
 AAACCTAACCCAAACCTCAAAGGGTTAGAAGGAATAATAACTACCACTACAATTTAATTGTGT
 F E L G L E F P N L P Y Y I D G D V K L T Q
 (in frame with thrombin site)



MscI
 GTCTATGGCCATCATACGTTATATAGCTGACAAGCACAACATGTTGGGTGGTTGTCCAAAAGAGC
 CAGATACCGGTAGTATGCAATATATCGACTGTTCTGTGTTGTACAACCCACCAACAGGTTTTCTCG
 S M A I I R Y I A D K H N M L G G C P K E
 (in frame with thrombin site)



GTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAGATACGGTGTTCGAGAATTGCA 2405
 CACGTCTCTAAAGTTACGAACTTCTCGCCAAAACCTATAATCTATGCCACAAAGCTCTTAACGT
 R A E I S M L E G A V L D I R Y G V S R I A
 ----- (in frame with thrombin site) ----->

GST

TATAGTAAAGACTTTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTACCTGAAATGCTGAAAAT 2470
 ATATCATTCTGAAACTTTGAGAGTTTCAACTAAAAGAATCGTTCGATGGACTTTACGACTTTTA
 Y S K D F E T L K V D F L S K L P E M L K M
 ----- (in frame with thrombin site) ----->

GST

BstBI **SwaI**
 GTTCGAAGATCGTTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCA 2535
 CAAGCTTCTAGCAAATACAGTATTTTGTATAAATTTACCACTAGTACATTGGGTAGGACTGAAGT
 F E D R L C H K T Y L N G D H V T H P D F
 ----- (in frame with thrombin site) ----->

GST

TGTTGTATGACGCTCTTGATGTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAAAA 2600
 ACAACATACTGCGAGAACTACAACAAAATATGTACCTGGGTTACACGGACCTACGCAAGGGTTTT
 M L Y D A L D V V L Y M D P M C L D A F P K
 ----- (in frame with thrombin site) ----->

GST

TTAGTTTGTTTTAAAAACGTATTGAAGCTATCCACAAATTGATAAGTACTTGAAATCCAGCAA 2665
 AATCAAACAAAATTTTTGCATAACTTCGATAGGGTGTTTAACTATTCATGAACTTTAGGTCGTT
 L V C F K K R I E A I P Q I D K Y L K S S K
 ----- (in frame with thrombin site) ----->

GST

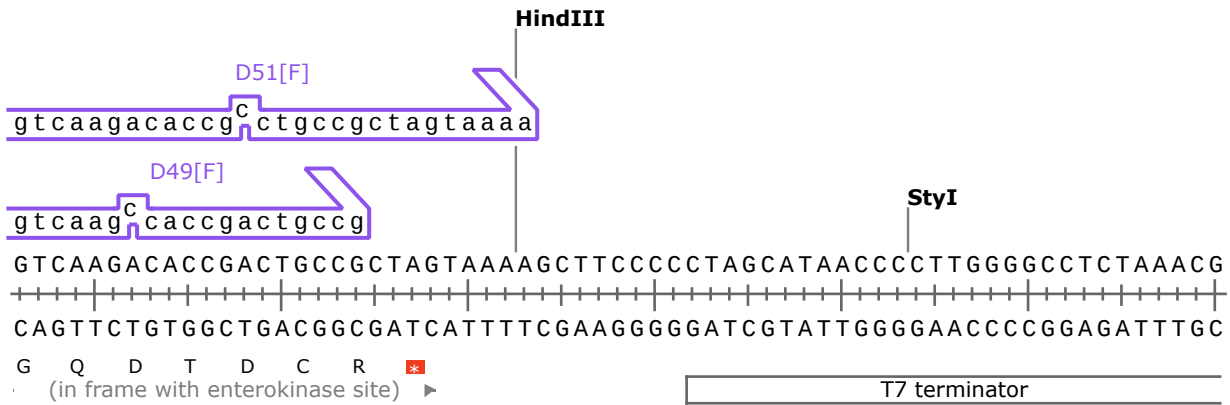
EcoNI

SEQ[F]

TGGCGACCATCCTCCAAAAG

GTATATAGCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAAAAG 2730
 CATATATCGTACCGGAAACGTCCCGACCGTTTCGGTGCAAACCACCACCGCTGGTAGGAGGTTTTC
 Y I A W P L Q G W Q A T F G G G D H P P K
 ----- (in frame with thrombin site) ----->

GST



PLADWT

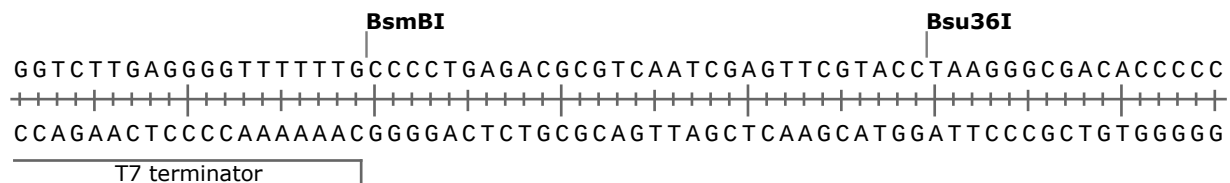
D49 D51

cagttcgtggctgacggc

D49[R]

cagttctgtggcgacggcgatcatTTT

D51[R]



XmaI SrfI
 TspMI SmaI

TAATTAGCCC GGGCGAAAGGCCAGTCTTTCTGACTGAGCCTTTTCGTTTTATTTGATGCCTGGCAG
 ATTAATCGGGCCCGCTTTCCGGGTCAGAAAGCTGACTCGGAAAGCAAATAAACTACGGACCGTC

TTCCGGGTCAGAAAGCTGAC

SEQ[R]

TTCCCTACTCTCGCATGGGGAGTCCCCACACTACCATCGGGCGCTACGGCGTTTCACTTCTGAGTT
 AAGGGATGAGAGCGTACCCCTCAGGGGTGTGATGGTAGCCGCGATGCCGCAAAGTGAAGACTCAA

3185

CGGCATGGGGTCAGGTGGGACCACCGCGCTACTGCCGCCAGGCAAACAAGGGGTGTTATGAGCCA
 GCCGTACCCAGTCCACCCTGGTGGCGCGATGACGGCGGTCCGTTTGTTCCCACAATACTCGGT

3250

NruI

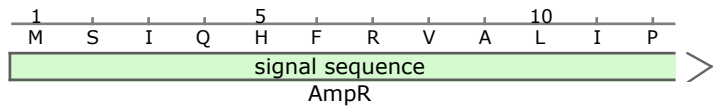
TATTCAGGTATAAATGGGCTCGCGATAATGTTTCAGAATTGGTTAATTGGTTGTAACACTGACCCC
 ATAAGTCCATATTTACCCGAGCGCTATTACAAGTCTTAACCAATTAACCAACATTGTGACTGGGG

3315

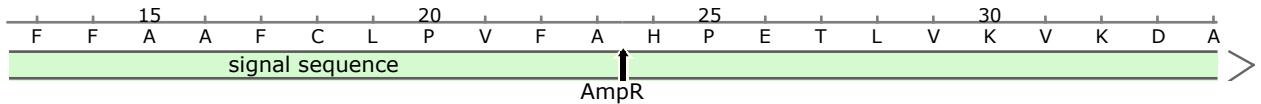
TATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAA
 ATAAACAATAAAAAGATTTATGTAAGTTTATACATAGGCGAGTACTCTGTTATTGGGACTATTT

3380

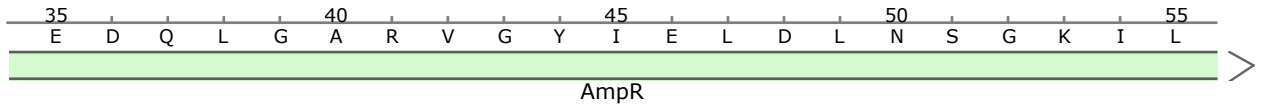
TGCTTCAATAATATTGAAAAAGGAAGAATATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCC
 ACGAAGTTATTATAACTTTTTCTTCTTATACTCATAAGTTGTAAAGGCACAGCGGGAATAAGGG



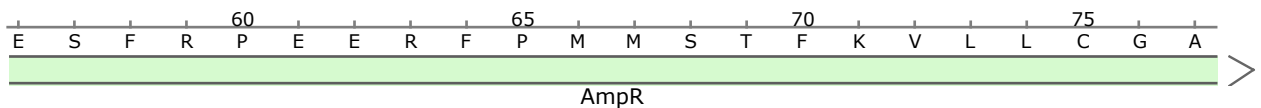
TTTTTTGCGGCATTTTGCCTTCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAAGATGC
 AAAAAACGCCGTAAAACGGAAGGACAAAAACGAGTGGGTCTTTGCGACCACTTTTCATTTTCTACG



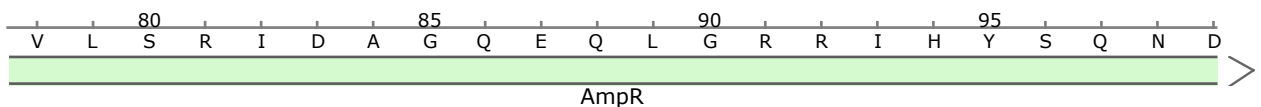
TGAAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTG
 ACTTCTAGTCAACCCACGTGCTCACCCAATGTAGCTTGACCTAGAGTTGTGCGCCATTCTAGGAAC



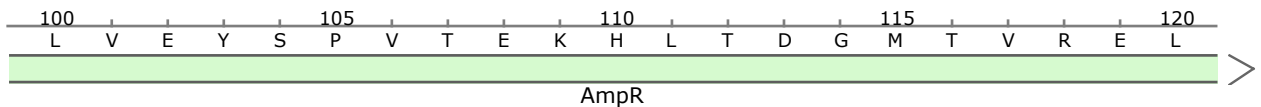
AGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCG
 TCTCAAAGCGGGGCTTCTTGCAAAGGTTACTACTCGTGAAAATTTCAAGACGATACACCGCGC



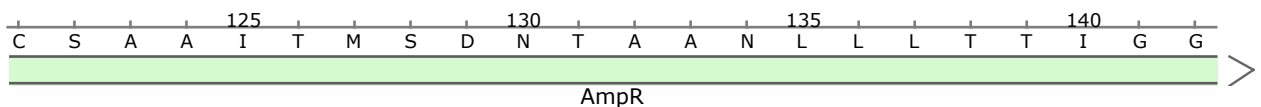
GTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCCGCATACACTATTCTCAGAATGA
 CATAATAGGGCATAACTGCGGCCCGTTCTCGTTGAGCCAGCGGCGTATGTGATAAGAGTCTTACT



CTTGGTTGAGTACTCACAGTACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTAT
 GAACCAACTCATGAGTGGTCAAGTGTCTTTTCGTAGAATGCCTACCGTACTGTCAATTCTTTAATA



GCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGA
 CGTCACGACGGTATTGGTACTCACTATTGTGACGCCGGTTGAATGAAGACTGTTGCTAGCCTCCT



CCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTA ACTCGCCTTGATCGTTGGGA
 3900
 GGCTTCCTCGATTGGCGAAAAAACGTGTTGTACCCCTAGTACATTGAGCGGAACTAGCAACCCT
 145 150 155 160
 P K E L T A F L H N M G D H V T R L D R W E
 AmpR

ACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCGATGGCAA
 3965
 TGGCCTCGACTTACTTTCGGTATGGTTTGCTGCTCGCACTGTGGTGCTACGGACATCGCTACCGTT
 165 170 175 180 185
 P E L N E A I P N D E R D T T M P V A M A
 AmpR

FspI
 CAACGTTGCGCAAAC TATTA ACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGAC
 4030
 GTTGCAACGCGTTTGATAATTGACCGCTTGATGAATGAGATCGAAGGGCCGTTGTTAATTATCTG
 190 195 200 205
 T T L R K L L T G E L L T L A S R Q Q L I D
 AmpR

TGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTAT
 4095
 ACCTACCTCCGCTATTTC AACGTCCTGGTGAAGACGCGAGCCGGAAGGCCGACCGACCAATA
 210 215 220 225
 W M E A D K V A G P L L R S A L P A G W F I
 AmpR

BspEI **AfeI**
 TGCTGATAAATCCGGAGCCGGTGAGCGTGGTTCTCGCGGTATCATCGCAGCGCTGGGGCCAGATG
 4160
 ACGACTATTTAGGCCTCGGCCACTCGCACCAAGAGCGCCATAGTAGCGTCGCGACCCCGGTCTAC
 230 235 240 245 250
 A D K S G A G E R G S R G I I A A L G P D
 AmpR

GTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAAT
 4225
 CATTCCGGGAGGGCATAGCATCAATAGATGTGCTGCCCTCAGTCCGTTGATACCTACTTGCTTTA
 255 260 265 270
 G K P S R I V V I Y T T G S Q A T M D E R N
 AmpR

AscI
 AGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAAGCGGCGGCCATCGAATGGC
 4290
 TCTGTCTAGCGACTCTATCCACGGAGTGACTAATTCGTAACCATTCGCCGCGCGGTAGCTTACCG
 275 280 285
 R Q I A E I G A S L I K H W *
 AmpR

GCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATATG 4355
 CGTTTTGGAAAGCGCCATACCGTACTATCGCGGGCCTTCTCTCAGTTAAGTCCCACCACTTATAC

1
M
lacI

AAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTTATCAGACCGTTTCCCAGCGT 4420
 TTTGGTCATTGCAATATGCTACAGCGTCTCATACGGCCACAGAGAATAGTCTGGCAAAGGGCGCA

5 10 15 20
K P V T L Y D V A E Y A G V S Y Q T V S R V
lacI

GGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGAAGCGGCGATGGCGGAGC 4485
 CCACTTGGTCCGGTCCGGTCAAAGACGCTTTTTCACCTTCGCGGCTACCGCCTCG

25 30 35 40
V N Q A S H V S A K T R E K V E A A M A E
lacI

TGAATTACATTCCCAACCGCGTGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTT 4550
 ACTTAATGTAAGGGTTGGCGCACCGTGTGTTGACCGCCCGTTTGTGTCAGCAACGACTAACCGCAA

45 50 55 60 65
L N Y I P N R V A Q Q L A G K Q S L L I G V
lacI

GCCACCTCCAGTCTGGCCCTGCACGCGCCGTGCGAAATTGTCGCGGCGATTAAATCTCGCGCCGA 4615
 CGGTGGAGGTCAGACCGGGACGTGCGCGGCAGCGTTTAACAGCGCCGCTAATTTAGAGCGCGGCT

70 75 80 85
A T S S L A L H A P S Q I V A A I K S R A D
lacI

TCAACTGGGTGCCAGCGTGGTGGTGTGCGATGGTAGAACGAAGCGGCGTTCGAAGCCTGTAAAGCGG 4680
 AGTTGACCCACGGTTCGCACCACCACAGCTACCATCTTGCTTCGCCGACGCTTCGGACATTTGCC

90 95 100 105
Q L G A S V V V S M V E R S G V E A C K A
lacI

CGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAG 4745
 GCCACGTGTTAGAAGAGCGCGTTGCGCAGTCAACCGACTAGTAATTGATAGGCGACCTACTGGTC

110 115 120 125 130
A V H N L L A Q R V S G L I I N Y P L D D Q
lacI

GATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCA 4810
 CTACGGTAACGACACCTTCGACGGACGTGATTACAAGGCCGAATAAAGAACTACAGAGACTGGT

135 140 145 150
D A I A V E A A C T N V P A L F L D V S D Q
lacI

GACACCCATCAACAGTATTATTTTCTCCCATGAGGACGGTACGCGACTGGGCGTGGAGCATCTGG
 ++++++
 CTGTGGGTAGTTGTCATAATAAAAGAGGGTACTCCTGCCATGCGCTGACCCGCACCTCGTAGACC
 155 160 165 170
 T P I N S I I F S H E D G T R L G V E H L
 lacI >

4875

TCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCATTAAGTTCTGTCTCGGCGCGTCTG
 ++++++
 AGCGTAACCCAGTGGTCGTTTAGCGCGACAATCGCCCGGGTAATTCAAGACAGAGCCGCGCAGAC
 175 180 185 190 195
 V A L G H Q Q I A L L A G P L S S V S A R L
 lacI >

4940

CGT 3'
 +++ ••• 4943
 GCA 5'
 R
 lacI >