

Ssp I

GCACATTTTCGGGAAATGTGCGCGGAACCCCTATTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCGTATAAATGCTTCAATAATAITGAAAAAGGAGAGTATGAGTATTC AACATTTCCGTGTCGCCCTTATTTCC 160  
CGTGA AAAAGCCCTTTACACGCGCCTTGGGGATAAAACAAATAAAAAGATTATGTAGITTTATACATAGGCGAGTACTCTGTTATTGGGACTATTTACGAAGITATTATAACTTTTCTCTCTCATACTCATAAGTGTAAAGGCACAGCGGGAATAAGG

H F S G K C A R N P Y L F I F L N T F K Y V S A H E T I T L I N A S I I L K K E E Y E Y S T F P C R P Y S

CTTTTTGCGGCATTTTGCCTTCTGTTTTGCTCACCCAGAAACGCTGGTGAAGTAAAAGATGCTGAAGATCAGTTGGGTGACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTGCGCCCGAAGAACGTTTTCCA 320  
GAAAAACGCCGTA AACCGAAGGACAAAAACGAGTGGGTCTTTGCGACCACTTTTCATTTTCTACGACTTCTAGTCAACCCACGTTGCTCACCCAATGTAGCTTGACCTAGAGTTGTGCGCATTTCTAGGAACTCTCAAAAAGCGGGGCTTTCTTGCAAAAAGGT

L F C G I L P S C F C S P R N A G E S K R C . R S V G C T S G L H R T G S Q Q R . D P . E F S P R R T F S

Sca I

ATGATGAGCACTTTTAAAGTTCCTGCTATGTTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCCGATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAAGTACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAG 480  
TACTACTCGTGA AAAITTTCAAGACGATACACCGCGCCATAATAGGGCATAACTGCGGCCGTTTCTCGTTGAGCCAGCGCGTATGTGATAAGAGTCTTACTGAACCAACTCATGAGTGGTCAGTGTCTTTTCGTAGAATGCCTACCCTACTGTCAITCTC

N D E H F . S S A M W R G I I P Y . R R A R A T R S P H T L F S E . L G . V L T S H R K A S Y G W H D S K R

Pvu I

AATTTATGCACTGCTGCCATAACCATGAGTGATAAACA CTGCGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACC 640  
TTAATACGTCACGACGGTATTGGTACTACTATTTGAGCGCGGTTGAATGAAGACTGTTGCTAGCCTCTCGCTTCTCGATTTGGCGAAAAACGTTGTGATACCCCTAGTACATTGAGCGGA ACTAGCAACCCCTTGGCCTCGACTTACTTTCGGTATGG

I M Q C C H N H E . . H C G Q L T S D N D R R T E G A N R F F A Q H G G S C N S P . S L G T G A E . S H T

AAACGACGAGCGTGACACCAAGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCT 800  
TTTGTGCTCGCACTGTGGTGTACGGACATCGTTACCGTTGTTGCAACCGGTTTGATAAATGACCGCTTGATGAATGAGATCGAAGGGCGGTTGTTAATATCTGACCTACCTCCGCTATTTCAACGTCCTGGTGAAGACGCGAGCCGGGAAGGCCGA

K R R A . H H D A C S N G N N V A Q T I N W R T T Y S S F P A T I N R L D G G G . S C R T T S A L G P S G

GGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGGAGGTTGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAG 960  
CCGACCAATAACGACTAITTTAGACCTCGGCCACTCGACCCAGAGCGCCATAGTAAAGTCTGTTGACCCCGGTTCTACAITTCGGGAGGGCATAGCATCAATAGATGTGCTGCCCCCTCAGTCCGTTGATACTACTTGTCTTATCTGTCTAGCGACTCTATC

W L V Y C . . I W S R . A W V S R Y H C S T G A R W . A L P Y R S Y L H D G E S G N Y G . T K . T D R . D R

GTGCCTCACTGATTAAGCATTGGTAACTGTGACACCAAGTTTACTCATATATACTTTAGATTTGATTTAAAACCTTCATTTTAAATTTAAAAGGATCTAGGTGAAGATCCTTTTGTATAATCTCATGACCAAAATCCCTTAACTGAGTITTTGCTTCCACTG 1120  
CACGGAGTGACTAATTCGTAACCAITGACAGTCTGGTTCAAATGAGTATATATGAAATCTAACTAAATTTTGAAGTAAAAAITAAATTTTCTAGATCCACTTCTAGGAAAAACTAATTAGAGTACTGGTITTTAGGGAATGCACTCAAAAAGCAAGGTGAC

C L T D . A L V T V R P S L L I Y T L D . F K T S F L I . K D L G E D P F . . S H D Q N P L T . V F V P L

AGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTCTCTGCGGTAATCTGCTGCTTGCAAACAACCAACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAAGTGGCTTCAGC  
TCCGAGTCTGGGCATCTTTCTAGTTTCTAGTAAAGAACTCTAGGAAAAAAGACCGCATTAGACGACGAACGTTTGTGTTTGGTGGCGATGGTCGCCACCAACAACCGGCTAGTTCTCGATGGTTGAGAAAAAGGCTTCCATTGACCGAAGTCG

1280

S V R P R R K D Q R I F L R S F F S A R N L L L A N K K T T A T S G G L F A G S R A T N S F S E G N W L Q

AGAGCGCAGATACCAAACTACTGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCCGCTACATACCTCGCTCTGCTAATCTGTGTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCAGGTTGGACTCAAGAC  
TCTCGGCTCTATGGTTTATGACAGGAAGATCACATCGGCATCAATCCGGTGGTGAAGTTCTTGAGACATCGTGGCGGATGTATGGAGCGAGACGATTAGGACAATGGTACCGACGACGGTACCCGCTATTCAGCACAGAATGGCCAACTGAGTTCTG

1440

Q S A D T K Y C P S S V A V V R P P L Q E L C S T A Y I P R S A N P V T S G C C Q W R . V V S Y R V G L K T

GATAGTTACCGATAAGGCGCAGCGTCCGGCTGAACGGGGTTCGTGACACAGCCAGCTTGGAGCGAACGACCTACACGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGT  
CTATCAATGGCTATTCGCGTCCAGCCGACTTGCCTCCCAAGCAGGTGTGTGCGGTGAACTCGCTTGTCTGGATGTGGCTTGTACTCTATGGATGTGCACTCGATACTCTTTCGCGTGGGAGGCGCTTCCCTCTTTCGCGCTGTCCATAGGCCA

1600

I V T G . G A A V G L N G G F V H T A Q L G A N D L H R T E I P T A . A M R K R H A S R R E K G G Q V S G

AAGCGCGAGGTCGGAACAGGAGCGCACAGGGAGCTTCCAGGGGAAACCGCTGGTATCTTTATAGTCTGTGCGGTTTCGCCACCTCTGACTTGGAGCTCGATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAAAAACGCCAGCAACCGC  
TTCGCGTCCAGCCTTGTCTCTCGGTGCTCCCTCGAAGGTCCCTTTGCGGACCATAGAAAATATCAGGACAGCCCAAGCGGTGGAGACTGAACTCGCAGCTAAAAACACTACGAGCAGTCCCCCGCTCGGATACCTTTTTGCGGTGCTTGGC

1760

K R Q G R N R R A H E G A S R G K R L V S L . S C R V S P P L T . A S I F V M L V R G A E P M E K R Q Q R

GCCTTTTTACGGTCTCGGCTTTTTGCTGGCTTTTTGCTCACATGTTCTTTCTGCGTTATCCCTGATTTCTGTGGATAACCGTATTACCCTTTTTGAGTGTAGCTGATACCCTCGCCGAGCCGAAACGACCGAGCGCAGCGAGTCAAGTGTAGCGGAGGAAG  
CGAAAAATGCCAAGACCGGAAAACGACCGAAAACGAGTGTACAAGAAAGGACGCAATAGGGGACTAAGACACCTATTGGCATAATGGCGGAACTCACTCGACTATGGCGAGCGGCTGGCTTGTCTGGCTCGGTCTAGTCACTCGCTCCTTC

1920

G L F T V P G L L L A F C S H V L S C V I P . F C G . P Y Y R L . V S . Y R S P Q P N D R A Q R V S E R G S

Pvu II

CGGAAGAGCGCCAAATACGCAAACCGCTCTCCCGCGCGTTGGCCGATTCAITTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGCGAGTGTAGCGCAACGCAATTAATGTGAGTTAGCTCACTCAITTAGGCACCCAGGCTTTACACTTTAT  
GCCITCTCGCGGTTATGCGTTTGGCGGAGAGGGCGCGCAACCGGCTAAGTAATACGTGACCGTGTCTGTCCAAAGGCTGACCTTTCCCGCTCACTCGCGTTGCGTTAATTACACTCAATCGAGTGTAGTAAATCCGTGGGGTCCGAAATGTGAAATA

2080

G R A P N T Q T A S P R A L A D S L M Q L A R Q V S R L E S G Q . A Q R N . C E L A H S L G T P G F T L Y

Acc65 I  
Kpn I

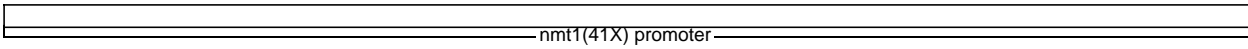
GCTTCCGGCTCGTATGTGTGTGGAATTTGTAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCTCGAAATTAACCTCACTAAAGGGAACAAAAGCTGGTACCCCGCATTTGATGGTTTCGCCATAAAAGATGTTTAAAT  
CGAAGGCCGAGCATAACAACACCTTAACACTCGCCTATTGTTAAAGTGTGTCTTTGTGATACTGGTACTAATGCGGTTGAGCTTTAATTTGGGAGTGTATTTCCCTTGTGTTTTCGACCATGGGCGTAAACTACCAAAGCGGATTTTCTACAAAATTA

2240

A S G S Y V V W N C E R I T I S H R K Q L . P . L R Q A R N . P S L K G T K A G T P H L M V S P K K M F .

EcoR I Pvu I Xba I  
AGCAAATTAATACTACACCTTTAAAAGGAAAAGGAACGTTCCACTATGCAGTGGAAATTCGGTCGATCGACTCTAGAGGATCAGAAAATTATCGCCATAAAAGACAGAATAAGTCATCAGCGGTGTGTTTCATTTCTCTATATTTTTTTTTTATTTTTTATTT  
TCGTTTAATTAATGATGTGAAATTTTCCTTTTCCTTGCAAGGTGATACGTCACCTTAAGCCAGCTAGCTGAGATCTCCTAGTCTTTTAAATAGCGGTATTTTCTGTCTTATTCAGTAGTCGCCAACAAAGTAAAGGATATAAAAAAAAAAATAAAAAATAAA

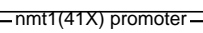
2400



. Q I N T T L . K E K E R S T M Q W N S V D R L . R I R K L S P . K T E . V I S G C F I S Y I F F L F F Y F

TTTAATAAGGAAAATTTAACGTTAAGGATACAGAAGATTGTTAGCACATTAAAGTAATAAAGGCTTAAGTAGTAAGTCCCTTAGCATGTTATTTGTAATTTCAAAGGACATAATCTAAAATAATAACAATATCATTTCTCACAGTTATTCAATTTTCTT  
AAATTAATCCCTTTTAAATTCAGATTCCATGTCCTTAAACAATCGTGAATTTCAATTAATTTCCGAATTCATCATTCACGGAATCGTACAATAACATAAAGTTTCCGTATTAGATTTTATTAATTTGTTATAGTAAAGAGTGTCAATAAGTTAAAGAA

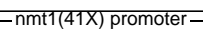
2560



L I R E N L T S K D T E D C . H I K V I K A . V V S A L A C Y C I S K D I I . N N N N I I S H K L F N F L

TTTTTTTCTAATAATATCAAGAATGTAATTTGTTTIGACATAAGTCAACTAATTTAATTTAATATGCTGGATTAATCTTGCAGACATGTAATTAACAAGTTTGTAGTCAATAACGTTGAAGTTTCAATGAACTCAAATAATTTCTCTTTTTTTTATA  
AAAAAAAAGATTATTTATAGTCTTACATAATAACAACACTGTATTTCAGTTGATTAAATAAATTTATACGACCTAATTTAGAACGTCGTACATTTAATTTGTTCAAATCAGTTTATTTGCAACTTCAAAGTTACTTGAGTTTATTTAAAGAGAAAAAATAAT

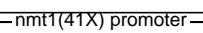
2720



F F F . . Y Q E C I I C L T . V N . F I . Y A G L I L Q T C K L T S F S Q I T L K F Q . T Q I I S L F F Y

TAACCATAAGTCTAATCTGATTTATATTTTCCGCAGGGATCAACTGAAGTTATGACATTTGGATTGGATCACTTATAACCTTGGTCGCCAAATAATACAAAAATCAGCGTTATAAAAACAAAGAAGGTTTTTGTITAAAGAAATTAATCCTCTTCTTTGATAA  
ATTGGTATTCAGATTAGACTAAATATAAAAGGCGTCCCTAGTTGACTTCAATACGTAAACCTAACCTAGTGAATATTGGAACCAGCGGTTTATTTATGTTTITTAGTCGCAATATTTTGTTCCTTCCAAAAACAATCTTTAATTTAGGAGAAAGAACTATT

2880

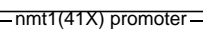


I T I S L I . F I F S A G I N . S Y D I W I G S L I T L V A K . Y K N Q R Y K T K K V F V K K L I L F L D K



GAAAGTTGAACCGAAATTCAGATACTGATATATGAAAATAATACCCACAATTTTGGGAATAGCGCAAGCCTCAATTTAAACAATAGGTGAGGACACATGATAATGACCTCAATGATTGTTAGAAGAAAAGAGCCTCATTTACAAAATCGAAAAATGAATG  
CTTTCAACTTGGCTTTAACGCTCTATGACTATATACCTTTTATTTATGGGTGTTAAAACCTTATCGCGTTCGGAGTTAAATTTGTTATCCACTCCTGTGTACTTACTGAGGTTACTAACAATCTTCTTTTCTCGGAGTAATGTTTGTAGCTTTTACTTTAC

3040



K V E P K L Q I L I Y E N N T H N F G N S A S L N L N N R . G H M I M T S M I V R R K E P H Y K I E K . M

Xba I

GT TGGGTACAAGTTTCCAAAACATGGTAAAGTGGACTTTGCGTATGAGACGTAATAAGAAAAACACTGTATATATGTTTCTAGAAATATTTGTGTCTCTTTATGGTTGGATGATGCAAATAGTAAATTCGGTTAGTTGCTGTAAAACACCACGAGA  
 CAACCCATGTTCAAAGGTTTGTACCATTTACCTGAAACGCATACTCTGCATTTATCTTTTGTGGAACAATATACAAAAGATCTTAATAACAACAGAGAAAATACCAACCTACTACGTTTATCAITAAAGCCAATCAACGACATTTTGTGGTGTCTCT

3200

nmt1(41X) promoter

V G Y K F P K H G K V D F A Y E T . I E K N T C Y M F S R I I V V S L W L D D A K . . F R L V A V K H H E

Hind III

CAATAGATATGGATATTTAATTAATCAGGAAAAACGTAACCTCTCGGCTACTGGATGGTTTCAGTCACCCAAACGATTAAGTGGGAGAGAAAAACAGGGCAAAGCTTAAAGGAATCCGATTTGTCATTTCGGCAATGTGCAGCGAAACTAAAAACCGGA  
 GTTTATCTATACTATAAATAAATTTAGTCCCTTTTGCATTTGAGAGCCGATGACCTACCAAGTCAGTGGTTGCTAATGACCCCTCTCTTTTGTGCCGTTTTCGTTTCGAAATTCCTTAGGCTAACAGTAAAGCCGTTACAGCTGCTTTGATTTTGTGGCTCT

3360

nmt1(41X) promoter

T N R Y G Y L L N Q E K R N S R L L D G S V T Q R L L G R E N R A K A K L K G I R L S F G N V Q R N . K P D

Nde I      Apa I

TAATGGACCTGTAAATCGAAACATTGAAGATAAAGGAAGAGGAATCCIGGCATATCATCAATTTGAATAGTTGAATTAATTAATTTCAATCTCATCTCACCTTCTGACTTATAGTCCGCTTTGTAAATCATATGGGGCCCTGCTTATCTTCTCTCTCC  
 ATTACCTGGACAATTAGCTTTGTAACTTCTATTTCCCTTCTCCCTTAGGACCGTATAGTAGTTAACTTATTTCAACTTAATTAATAAAGTTAGAGTAAGAGTGAAAGACTGAATATCAGCGAAACAATTTAGTATACCCCGGGACGAATAGAAAGAAAGAGG

3520

nmt1(41X) promoter

N G P V N R N I E D K G R G I L A Y H Q L N K L N . L F Q S H S H F L T Y S R F V K S Y G A P A Y L S S S

GAAAGTGTACTTTCAGAAAGAGCAAGAAGTGTACGGATAAGGATAAGTATATCTTCTATCTACTACAATACGACAAGAAGTCTCGAACTCTCCCTCCACCTTTACTTTAACTAATAGATTATGCAGATTTTCGTCAAGACTTTGACCGGTAAAACCATAA  
 CTTTCACGATGAAGTCTTCTCGTCTTGCATATGCTATTCCTATTCATATAGAAGATAGATGATGTTATGCTGTCTTGTAGAGCTTGTAGAGGGAGGGTGAATGAATGATTATCTAATACGCTTAAAAGCAGTTCTGAAACTGGCCATTTTGGTATT

3680

ubiquitin

E S A T S E R A R T V R I R I S I S S I Y Y N T T R T L E L S L P L Y F N . . I M Q I F V K T L T G K T I

BstX I

CATTGGAAGTTGAATCTTCCGATACCATCGACAACGTTAAGTCGAAAATTCAGACAAGGAAGGTATCCCTCCAGATCAACAAGATTGATCTTTGCGGTAAGCAGCTAGAAGACGGTAGAACGCTGTCTGATTACAACATTCAGAAGGAGTCCACCTT  
 GTAACCTTCAACTTAGAAGGCTATGGTAGCTGTGCAATTCAGCTTTAAGTCTGTCTCCATAGGGAGGCTAGTTGTCTTCACTAGAAACGGCCATTCGTCGATCTTCTGCCATCTTGCACAGACTAATGTTGTAAGTCTTCCCTCAGGTGGAA

3840

ubiquitin

T L E V E S S D T I D N V K S K I Q D K E G I P P D Q Q R L I F A G K Q L E D G R T L S D Y N I Q K E S T L

BamH I Sca I

ACATCTTGTGCTAAGGCTAAGAGGTGGTAGGCACGGATCCGGCATCATGGTTCGACCATTGAAC TGCATCGTCGCGTGTCCCAAATATGGGGATTGGCAAGAACGGAGACCTACCC TGGCCTCCGCTCAGGAACGAGTTC AAGTACTTCCAAGAATG 4000

TGTAGAACACGATTCCGATTCTCCACCATCCGTGCCTAGGCCGTAGTACCAAGCTGGTAACTT GACGTAGCAGCGGCACAGGGTTTATACCCCTAACCGTTC TTGCCTCTGGATGGGACCGAGGCGAGTCTTGTCTCAAGTTCATGAAGGTTCTTTAC

ubiquitin N L DHFR ts degrpm

H L V L R L R G G R H G S G I M V R P L N C I V A V S Q N M G I G K N G D L P W P P L R N E F K Y F Q R M

Sac I Sst I

ACCACAACCTCTTTCAGAGGAAGGTAACAGAACTCTGGTGAATTATGGGTAGGAAAACCTGGT TCTCCATTCCTGAGAAGAATCGACTTTTAAAGGACAGAATTAATATAGTTCCTCAGTAGAGAACTCAAAGAACCAACCAGGAGCTCATTTTCTTGCCA 4160

TGGTGTGGAGAAGTCTCCTTCCATTTGTCTTAGACCACTAATACCCATCCTTTTGGACCAAGAGGTAAAGACTCTTCTTAGCTGAAAATTTCTGTCTTAATATATATCAAGAGTCAITCTCTTGAGT TCTTGGTGGTGTCTCTCGAGTAAAAGAACGGT

DHFR ts degrpm

T T T S S E E G K Q N L V I M G R K T W F S I P E K N R L L K D R I N I V L S R E L K E P P R G A H F L A

BstX I

AAAGTTTGGATGATGCCCTTAAGACTTATTGAACAACCGGAATTGGCAAGTAAAGTAGACATGGT TTTGGATAGTCGGAGGCAGTTC TGTTTACCAGGAAGCCATGAATCAACCAGGCCACCTCAGACTCTTTGTGACAAGGATCATGCAGGAATTTGAAAG 4320

TTTCAAACCTACTACGGAATTCGAAATACTTGTGGCCTTAACCGTTCATTTTCATCTGTACCAAACCTATCAGCCTCCGTCAAGACAAATGGTCTTCCGTACTTAGTTGGTCCGGTGGAGTCTGAGAAACTGTTCCTAGTACGTCTTAAACTTTC

DHFR ts degrpm

K S L D D A L R L I E Q P E L A S K V D M V W I V G G S S V Y Q E A M N Q P G H L R L F V T R I M Q E F E S

EcoN I Acc65 I Kpn I

TGACACGTTTTTCCCAGAAATGATTTGGGGAAATATAAACTTCTCCAGAAATACCCAGGCGTCTCTCTGAGGTC CAGGAGGAAAAAGGCATCAAGTATAAGTTTGAAGTCTACGAGAAGAAGACGGTACCTACCCATACGATGTTCCAGATTACGCT 4480

ACTGTGCAAAAAGGGTCTTTAACTAAACCCCTTTATATTGTAAGAGGGTCTTATGGGTCCG CAGGAGACTCCAGGTCCTCTTTTCCGTAGTTTCATATTCAAACCTCAGATGCTCTTCTTTCTGCCATGGATGGGTATGCTACAAGGTTAATGCGA

DHFR ts degrpm HA epitope

D T F F P E I D L G K Y K L L P E Y P G V L S E V Q E E K G I K Y K F E V Y E K K D G T Y P Y D V P D Y A

Ava I Xho I Bgl II Ava I Sma I Hind III

GCAITTTTAGGAGGGCAAGCCTCGAGGGTAGATCTGGTGC CCGGGTGGTGTCTGGTGC CCGGAGCCGGTGTGGTGTCTGAAGCTTTAATGAGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATCTTGTGTAATTAGATGGT GATGTTAATGGGCA 4640

CGTAAAAATCCCTCCCGTTCGGAGCTCCCATCTAGACCA CGGGCCCCACACGACCACGGCC TCGGCCACGACCACTTCGAAATTACTCATTTCTCTTTTGAAAAGT GACCTCAACAGGGTTAAGAACAAC TTAATCTACCACTACAATTAACCCGT

long oligo 2 (reverse)

sites for inserting n terminus

A F L G G Q A S R V D L V P G V V L V P E P V L V L K L . . V K E K N F S L E L S Q F L L N . M V M L M G

Nco I Nde I

CAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATACGAAAACCTACCCCTTAAATTTATTTGCACTACTGGAAAACCTGTTCCATGGCCAACACTTGTCACTACTTTCCTTATGGTGTTCATGCTTTTCAAGATACCCAGATCATATG 4800  
GTTTAAAAGACAGTCACCTCTCCACTTCCACTACGTTGTATGCCTTTTGAATGGGAATTTAAATAAACGTGATGACCTTTTGATGGACAAGGTACCGGTGTGAACAGTGTGAAAGAGAATACCACAAGTTACGAAAAGTTCTATGGGTCTAGTATAC  
T N F L S V E R V K V M Q H T E N L P L N L F A L L E N Y L F H G Q H L S L L S L M V F N A F Q D T Q I I .

Cla I

AAGCGCACGACTTCTTCAAGAGCGCCATGCCTGAGGGATACGTCAGGAGAGGACCATCTTCTTCAAGGACGACGGGAACCTACAAGACAGTGTCTGAAGTCAAGTTTGGGGAGACACCCCTGTCACAGGATCGAGCTTAAAGGAATCGAATTTCAAGG 4960  
TTGCGCGTGTCTGAAGAAGTCTCGCGGTACGGACTCCCTATGCAGTCTCTCTGGTAGAAGAAGTTCCTGCTGCCCTTGATGTTCTGTGACGACTTCAGTTCAAACTCCCTCTGTGGGAGCAGTTGTCTAGCTCGAATTCCTTAGCTAAAGTTCC  
S G T T S S R A P C L R D T C R R G P S S S R T T G T T R H V L K S S L R E T P S S T G S S L R E S I S R

AGGACGGAACATCCTCGGCCACAAGTTGGAATACAACCTACAACCTCCACAACTATACATCATGGCCGACAAGCAAAAAGAACGGCATCAAGCCAACCTTCAAGACCCGCCACAACATCGAAGACGGCGCGTGCACCTCGCTGATCAITATCAACAAAA 5120  
TCCTGCCCTTGTAGGAGCGGTGTCAACCTTATGTTGATGTTGAGGGTGTGTCATATGTAGTACCGGCTGTCTGTTTCTTTCGCCGTAGTTTCGGTGAAGTTCGGCGGTGTGTAGCTTCTGCCGCGCACGTTGAGCGACTAGTAATAGTGTGTTT  
R T E T S S A T S W N T T T T P T T Y T S W P T S K R T A S K P T S R P A T T S K T A A C N S L I I I N K

Pvu II

TACTCCAATTTGGCGATGGCCCTGTCCCTTTTACCAGACAACCAATACCTGTCCACACAATCTGCCCTTTTCGAAAGATCCCAACGAAAAGAGAGACCACATGGTCCCTTCTGTAGTTTGTAAACAGCTGCTGGGATTACACATGGCATGGATGAACATATAAAA 5280  
ATGAGGTTAACCGCTACCGGGACAGGAAAATGTCTGTTGGTAAATGGACAGGTGTGTTAGACGGGAAAGCTTTCTAGGGTTGCTTTTCTCTCTGGTGTACCAGGAAGAACTCAAACAATGTGACGACCCCTAATGTGTACCGTACCTACTTGATATGTTT  
I L Q L A M A L S F Y Q T T I T C P H N L P F R K I P T K R E T T W S F L S L . Q L L G L H M A W M N Y T N

Sal I   BamH I   Sac I  
Sst I Pvu II

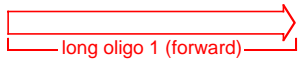
CATGATGAGCTTGTTCGACGGATCCTGAGCTCCAATTCGCCCTATAGTGTGAGTGTATTACAATTCCTGCGCGTGTGTTTACAACGTCGTGACTGGGAAAAACCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTTCGCCAGCTGGCGTAA 5440  
GTACTACTCGAACAGCTGCCTAGGACTCGAGGTTAAGCGGATATCACTCAGCATAATGTTAAGTGTACCGGACGAAAATGTTGCAGCACTGACCCCTTTTGGGACCGCAATGGGTTGAATTAGCGGAACGTCGTGTAGGGGAAAGCGGTGACCGCATT  
M M S L S T D P E L Q F A L . . V V L Q F T G R R F T T S . L G K P W R Y P T . S P C S T S P F R Q L A .

Pvu I

TAGCGAAGAGGCCCGCACCAGTTCGCCCTTCCCAACAGTTGCGCAGCCCTGAATGGCGAATGGGACGCGCCCTGTAGCGCGCATTAAGCGCGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCT 5600  
ATCGCTTCTCCGGCGTGGCTAGCGGGAAGGGTTGTCAACGCGTGGACTTACCCTTACCCTGCGCGGACATCGCCCGTAAATTCGCGCCGCCACACCACCAATCGCGTTCGCACTGGCGATGTGAACGTCGCGGGATCGCGGCGAGGAAAGCGA  
. R R G P H R S P F P T V A Q P E W R M G R A L . R R I K R G G C G G Y A Q R D R Y T C Q R P S A R S F R

NgoM I      Sac I  
Sst I

TTCTTCCCTTCCTTTCTCGCCACGTTTCGCGGGCGAGCTCGTTTAAACTGGATGGCGGCGTTAGTATCGAATCGACAGCAGTATAGCGACCAGCATTACATACGATTGACGCATGATATTACTTTCTGCGCACTTAACTTCGCATCTGGGCAGATGATGT  
AAGAAGGGAAAGAAAGAGCGGTGCAAGCGGCCGCTCGAGCAAAATTTGACCTACCGCCGCAATCATAGCTTAGCTGTGTCATATCGCTGGTCTGAAGTGTATGCTAACTGCGTACTATAATGAAAGACCGGTGAATTTGAAGCGTAGACCCGCTCTACTACA



————— KanR —————

F L P F L S R H V R R R A R L N W M A A L V S N R Q Q Y S D Q H S H T I D A . Y Y F L R T . L R I W A D D V

Sca I

CGAGGCGAAAAAATATAAATCACGCTAACATTTGATTAAAATAGAACAACTACAATATAAAAAAATATACAAATGACAAGTTCCTTGAAAACAAGAATCTTTTATTGTGTCAGTACTGATTAGAAAACTCATCGAGCATCAAATGAAACTGCAATTTA  
GCTCGCTTTTTTTTATAATTTAGTGGGATGTGAAACTAAATTTTATCTTGTGTGATGTTATAATTTTTTTGATATGTTTACTGTTCAGAACITTTTGTCTTAGAAAAATAACAGTCATGACTAATCTTTTTGAGTAGCTCGTAGTTTACTTTGACGTTAAAT

————— KanR —————

E A K K N I N H A N I . L K . N N Y N I K K L Y K . Q V L E N K N L F I V S T D . K N S S S I K . N C N L

TTTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTTCTGTAATGAAGGAGAAAACTCACCGAGGAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCTCGTCAA  
AAGTATAGTCTAATAGTTATGGTATAAAAACTTTTTGCGCAAAGACATTACTTCCCTTTTTGAGTGGCTCCGTCAAGGTATCCTACCGTTCTTAGGACCATAGCCAGACGCTAAGGCTGAGCAGGTTGTAGTTATGTTGGATAATTAAGGGGAGCAGTT

————— KanR —————

F I S G L S I P Y F . K S R F C N E G E N S P R Q F H R M A R S W Y R S A I P T R P T S I Q P I N F P S S

Hind III      Nsi I

AAATPAGGTTATCAAGTGAGAAATCACCATGAGTGCAGACTGAATCCGGTGAGAAATGGCAAAGCTTATGCATTTCTTTCCAGACTTGTTCACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTG  
TTTTATCCAATAGTTTACTCTTTAGTGGTACTCACTGCTGACTTTAGGCCACTTTACCGTTTTTCGAATACGTAAAGAAAGGTCGAACAAGTTGTCCGGTCGGTAATGCGAGCAGTAGTTTTAGTGGCGTAGTTGGTTTGGCAATPAGTAAGCACTAAC

————— KanR —————

K I R L S S E K S P . V T T E S G E N G K S L C I S F Q T C S T G Q P L R S S S K S L A S T K P L F I R D C

Pvu I      Ssp I      EcoN I

CGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAAATCAGGATAITTTCTTAATACCTGGAATGCTGTTTTGCGGGGATCGCA  
CGGACTCGCTCTGCTTTATGCGCTAGCGACAATTTTCCGTAAATGTTTGTCTTAGCTTACGTTGGCCGCTCCTTGTGACGGTCGCGTAGTTGTATAAAAGTGGACTTAGTCTATAAGAAGATTATGGACCTTACGACAAAACGGCCCTAGCGT

————— KanR —————

A . A R R N T R S L L K G Q L Q T G I E C N R R R N T A S A S T I F S P E S G Y S S N T W N A V L P G I A

Nsi I

GTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAAATCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCTGGGCGCATCGGGCT  
CACCACCTCATTGGTACGTAGTAGTCTCATGCCTAATTTACGAACCTACCAGCCCTCTCCGTATTTAAGGCAGTCGGTCAAATCAGACTGGTAGAGTAGACATTGTAGTAACCGTTGCGATGGAAACGGTACAAAGTCTTTGTTGAGACCGGTAGCCCGA

6560

KanR

V V S N H A S S G V R I K C L M V G R G I N S V S Q F S L T I S S V T S L A T L P L P C F R N N S G A S G

Cla I

Nru I

Nco I

TCCCATACAATCGATAGATTTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATAACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCTCGAAACGTGAGTCTTTTCCTTACCCATGGTTGTTATGTTCCGGATGATGAGAGAA  
AGGGTATGTTAGCTATCTAACAGCGTGGACTAACGGGCTGTAATAGCGCTCGGGTAAATATGGGTATAATTTAGTCTGTTAGGTACAACCTTTAAATTAGCGCCGGAGCTTTGCACCTCAGAAAAGGAATGGGTACCAACAAATACAAGCCCTACACTACACTCTT

6720

KanR

F P Y N R . I V A P D C P T L S R A H L Y P Y K S A S M L E F N R G L E T . V F S L P M V V Y V R M . C E N

Pst I

CTGTATCCTAGCAAGATTTTAAAGGAAGTATAATGAAAGAAGAACCCTCAGTGGCAAATCCTAACCTTTTATAATTTCTCTACAGGGGCGCGCGGTGGGACAAATCAACGCGTCTGTGAGGGGAGCGTTTCCCTGCTCGCAGGTCTGCAGCGAGGAGCCGT  
GACATAGGATCGTTCTAAATTTTCCCTTCAATACCTTCTCTTGGAGTACCCGTTTAGGATTGGAAAATATAAAGAGATGTCCTCCCGCGCGCACCCCTGTTAAGTTGCGCAGACACTCCCTCGCAAAGGGACGAGCGTCCAGACGTGCTCTCTCGGCA

6880

KanR

C I L A R F . K E V Y E R R T S V A N P N L L Y F S T G A R R G D N S T R L . G E R F P A R R S A A R S R

BstX I

AATTTTTCCTTCGCGCGGTGCGCCATCAAAATGTATGGATGCAAAATGATATACATGGGGATGATGGGCTAAATGTACGGGCGACAGTACATCATGCCCTGAGCTGCGCACGTCGAAGACTGTCAAGGAGGGTATTCTGGGCCCTCCATGTCGCTGGC  
TTAAAAACGAAGCGCGCACCGGTAGTTTACATACCTACGTTTACTAATATGTAOCCCTACATACCCGATTTACATGCCCGCTGTGAGTGTAGTACGGGACTCGACCGGTGCAGTTCGACAGTTCCTCCCATTAAGACCCGGAGGTACAGCGACCG

7040

KanR

N F C F A P C G H Q N V W M Q M I I H G D V W A K C T G D S H I M P L S C A R Q D C Q G G Y S G P P C R W

BstE II

NgoM I

CGGGTACCCCGGGGACGAGGCAAGCTAAACAGCCGGCTTTCCCGGTCAAGCTCTAAATCGGGGCTCCCTTTAGGGTTCCGATTTAGTGTCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGTATGTTTCACGTAGTGGGCCATCGCCCTGA  
GCCCACTGGGCGCCCTGCTCCGTTGATTTGTGCGCCGAAAGGGGAGTTTCGAGATTTAGCCCGGAGGAAATCCCAAGGCTAAATCACGAAATGCCGTGGAGCTGGGGTTTTCGAACTAATCCCACTACCAAGTGCATCACCCGGTACGCGGACT

7200

KanR

P G D P A G T R Q A K Q P A F P V K L . I G G S L . G S D L V L Y G T S T P K N L I R V M V H V V G H R P D



TAGACGGTITTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAAC TGGAAACAACACTCAACCCATCTCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAAC  
ATCTGCCAAAAAGCGGAAACTGCAACCTCAGGTGCAAGAAATTATCACCTGAGAACAAGGTTTGACCTTGTGTGAGTGGGATAGAGCCAGATAAGAAAACTAAATATTCCCTAAAACGGCTAAAGCCGGATAACCAATTTTTTACTCGACTAAATTG

7360

R R F F A L . R W S P R S L I V D S C S K L E Q H S T L S R S I L L I Y K G F C R F R P I G . K M S . F N

Ssp I  
AAAAATTTAACGCGAATTTTAAACAAAATATTAACGCTTACAATTTAGGTG  
TTTTTAAATGCGCTTAAAATTGTTTTATAAATGCGAATGTTAAATCCAC

7410

K N L T R I L T K Y . R L Q F R W