



PvuI (7)
SgfI (6)
1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGTGCCTA
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203)
201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTTGTTCGTTT

KasI (535) **AgeI (552)**
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCCTACCTGAGATCACGGTCCAGCCATGGACCCCGTTGGCCTCCAGCTCGGCAACAAGAACCT
601 GTGGAGCTGCTTGTGAGGCTGCTACCAAAAGACCCAGAATGGCTGAACGCCAAGATGAAGTTCTTCTCCCAACACGGACCTGGATTCCAGGAACGAG
13> W S C L V R L L T K D P E W L N A K M K F F L P N T D L D S R N E
701 ACCTTGGACCTGAACAGAGAGTATCCTGCAACTCAACAAGCTGCATGTCCAGGGTTCGGACACCTGGCAGTCTTTTATTTCATTGTGTGCATGCAGC
47> T L D P E Q R V I L Q L N K L H V Q G S D T W Q S F I H C V C M Q

SphI (791)
801 TGGAGGTGCCTTGACCTGGAGGTGCTGCTGCTGAGTACTTTTGGCTATGATGATGGTTCCACAGCCAGCTGGGAGCTGAGGGGAAAAGCCAACTGA
80> L E V P L D L E V L L L S T F G Y D D G F T S Q L G A E G K S Q P E
901 ATCTCAGTCCACCATGGCCTGAAGCGCCACACTCAGAGCTGTGGTCTCACCCGCGGAAGCAGTGAAGAAGCAGCAGCTAGAGTTGGCCAAGAAG
113> S Q L H H G L K R P H Q S C G S S P R R K Q C K K Q Q L E L A K K

ScaI (835)
1001 TACCTGCAGCTCCTGCGGACCTCTGCCAGCAGCGCTACAGGAGCCAAATCCCTGGGTGAGGGACGCCCCACGCTTCCACAGGTCTATGTCCTCCAA
147> Y L Q L L R T S A Q Q R Y R S Q I P G S G Q P H A F H Q V Y V P P
1101 TCCTGCGCCGGCCACAGCATCCTTAGACTCCGAGGGGGCCATTATGGGGACGTCAAGGTGGAAGATGGTGTGACGTGAGCATCTCGGACCTCT
180> I L R R A T A S L D T P E G A I M G D V K V E D G A D V S I S D L F
1201 CAACACCAGGGTTAAACAAGGGCCGAGGGTACCGTGTCTTTGGGAAGGCTGGCATGGGCAAGACCACGCTGGCCACCGGCTCTGCCAGAAGTGGGA
213> N T R V N K G P R V T V L L G K A G M G K T T L A H R L C Q K W A

Eco4VII (1031)
1301 GAGGGCCATCTGAACTGTTTCCAGGCCCTGTTCTTTTTGAATTCGCCAGCTCAACTTGATCACGAGTTTCTGACACCGTCCGAGCTCCTTTTTGATC
247> E G H L N C F Q A L F L F E F R Q L N L I T R F L T P S E L L F D
1401 TGTACCTGAGCCCTGAATCGGACCACGACTGTCTTCCAGTACCTGGAGAAGACGCTGACCAAGTCTGCTGATCTTTGATGGCTAGATGAGGCCCT
280> L Y L S P E S D H D T V F Q Y L E K N A D Q V L L I F D G L D E A L

EcoRI (1339)
1501 CCAGCCTATGGGTCTGATGGCCAGGCCAGTCTCACCTTTTTCTCCATCTCTGCAATGGGACCCCTCGCTGGCTGCCGGGTGATGGCTACCTCC
313> Q P M G P D G P G P V L T L F S H L C N G T L L P G C R V M A T S
1601 CGTCCAGGGAAGTGCCTGCCTGCCTGCCTGCAGAGGCAGCCATGGTCCACATGTTGGCTTTGATGGCCACGGGTGGAAGAATATGTAATCACTTCT
347> R P G K L P A C L P A E A A M V H M L G F D G P R V E E Y V N H F
1701 TCAGCGCCAGCCATCGCGGGAGGGGCCCTGGTGGAGTTACAGACAAATGGAGCTCTCCGAGCCTGTGTGGGTGCCCGACTGTGCCAAGTCGCCTG
380> F S A Q P S R E G A L V E L Q T N G R L R S L C A V P A L C Q V A C
1801 TCTCTGCCTCCACCATCTGCTTCTGACCACGCCCCAGGCCAGTCTGTGGCCCTCTGCCAACATGACTCAGCTCTATATGCAGATGGTGTCTGCCCTC
413> L C L H H L L P D H A P G Q S V A L L P N M T Q L Y M Q M V L A L
1901 AGCCCCCTGGGCACTTGCCACCTCGTCCCTACTGGACCTGGGGGAGGTGGCCCTGAGGGGCCCTGGAGACAGGAAGTTATCTTCTATGAAAAGATA
447> S P P G H L P T S S L L D L G E V A L R G L E T G K V I F Y A K D
2001 TTGCTCCACCCTTGATAGCTTTTGGGGCACTCACAGCTGCTGACTTCTTCTGCTGTCAGAGCCCTGGGCACCAAGCAGACAGGCTATGCTTTTAC
480> I A P P L I A F G A T H S L L T S F C V C T G P G H Q Q T G Y A F T
2101 CCACCTCAGCCTGAGGAGTTTCTGCTGCCCTGCACCTGATGGCCAGCCCCAAGGTGAACAAAGACACACTTACCCAGTATGTTACCTCCATTCCCGC
513> H L S L Q E F L A A L H L M A S P K V N K D T L T Q Y V T L H S R
2201 TGGGTACAGCGGACAAAGCTAGACTGGCCTCTCAGACCCTCCCCACCTTCTGGCGGGCTGGCATCTGCACCTGCCGCCCTTCTTAGCCAC
547> W V Q R T K A R L G L S D D H L P T F L A G L A S C T C R P F L S H
2301 TGGCGCAGGGCAATGAGGACTGTGTGGGTGCCAAGCAGGCTGTGTAGTGCAGGTGTTGAAGAAGTTGGCCACCCGCAAGCTCACAGGGCCAAAGTTGT
580> L A Q G N E D C V G A K Q A A V V Q V L K K L A T R K L T G P K V V
2401 AGAGCTGTGTCAGTGTGGATGAGACACAGGAGCTGAGCTGGCCAGTCTCACCGACAAAGCTCCCTATCAACTGCCCTTCCACAATTTCCCACTG
613> E L C H C V D E T Q E P E L A S L T A Q S L P Y Q L P F H N F P L
2501 ACCTGCACCGACTGGCCACCCTGACCAACATCTAGAGCACAGGGAGGCCCATCCACTGGATTTTGTGGCTGTCCCTGGAGCCCACTGCGCCCTG
647> T C T D L A T L T N I L E H R E A P I H L D F D G C P L E P H C P
2601 AGGCTCTGGTAGGCTGTGGGAGATAGAGAATCTCAGCTTTAAGAGCAGGAAGTGTGGGGATGCCTTTGCAGAAGCCCTCCAGGAGCTTCCGACAAT
680> E A L V G C G Q I E N L S F K S R K C G D A F A E A L S R S L P T M
2701 GGGGAGGCTGCAGTGTGGGTTAGCAGGAAGTAAATCACTGCCGAGGCACTCAGCCACCTGGTGAAGCTTTGCCTCTGTCCACAGCTGAAAGAA
713> G R L Q M L G L A G S K I T A R G I S H L V K A L P L C P Q L K E
2801 GTCAGTTTTCCGGACAACAGCTCAGTGACCAGGTGGTGTGAACATTGTGGAGTTTCTCCCTCACCTACCACGGCTCCGGAAGCTTGACCTGAGCAGCA
747> V S F R D N Q L S D Q V V L N I V E V L P H L P R L R K L D L S S
2901 ACAGCATGTGCTGCAACCTACTCTGCTTGGCAAGGTTGGCAGTCACTGTCTCACCTCAGGATGCTTACAGGCAAGGAGGGGACCTCATCTTCTC
780> N S I C V S T L G L A R V A V T C P T V R M L Q A R E A D L I F L
3001 TCTTTCCCGCCACAGAGACAACCTGCAGAGCTACAAAGAGCTCCAGCCTGCAGGAAAGTGCAGGCCAGAGGAAAGGGGCTCAGAGCAGAAGCTTGACG
813> L S P P T E T T A E L Q R A P D L Q E S D G Q R K G A Q S R S L T
3101 CTCAGGCTCAGAAGTGTGAGCTCAGGTCACAGATGCGGAGGCCCTCATAGCCCTGCTCCAGGAAGGCCCTCAGTGGAGGAGTGGACCTCTCAGGGA
847> L R L Q K C Q L Q V H D A E A L I A L L Q E G P H L E E V D L S G

3201 ACCAGCTGGAAGATGAAGGCTGTCGGCTGATGGCAGAGGCTGCATCCCAGCTGCACATCGCCAGGAAGCTGGACCTCAGTGACAACGGGCTTTCTGTGGC
880▶ N Q L E D E G C R L M A E A A S Q L H I A R K L D L S D N G L S V A
3301 CGGGGTGCATTGTGTGCTGAGGGCCGTGAGTGCCTGGACCTGGCAGAGCTGCACATCAGCCTGCAGCACAAAACCTGTGATCTTCACTGTTTGCCCA
913▶ G V H C V L R A V S A C W T L A E L H I S L Q H K T V I F M F A Q
3401 GAGCCAGAGGAGCAGAAGGGGCCAGGAGAGGGCTGCATTTCTGACAGCCTCATGTCCAGATGCCCTCTGAGCTGCCTCTGAGCTCCCGAAGGATGA
947▶ E P E E Q K G P Q E R A A F L D S L M L Q M P S E L P L S S R R M
3501 GGCTGACACATTGTGGCCTCAAGAAAAGCACCTAGAGCAGCTCTGCAAGGCTCTGGGAGGAAGCTGCCACCTCGGTACCTCCACCTCGACTTCTCAGG
980▶ R L T H C G L Q E K H L E Q L C K A L G G S C H L G H L H L D F S G
3601 CAATGCTCTGGGGATGAAGGTGCAGCCGGCTGCTCAGCTCCTCCAGGGCTGGGAGCTCTGCAGTCCCTTGAACCTCAGTGAGAACGGTTTGTCCCTG
1013▶ N A L G D E G A A R L A Q L L P G L G A L Q S L N L S E N G L S L
3701 GATGCCGTGTTGGGTTTGGTTCGGTCTTCTCCACTCTGCAGTGGCTCTTCCGCTTGACATCAGCTTTGAAAGCCAACACATCTCCTGAGAGGGGACA
1047▶ D A V L G L V R C F S T L Q W L F R L D I S F E S Q H I L L R G D

Eco47III (3875)

3801 AGACAAGCAGGGATATGTGGCCACTGGATCTTTGCCAGACTCCAGCTGCAGCAAGTTCTTAGGGTTCCTGCAGCGCTGCATCCCCAGGAGCCTCTG
1080▶ K T S R D M W A T G S L P D F P A A A K F L G F R Q R C I P R S L C

SmaI (3966)

XmaI (3963)

3901 CCTCAGTGAGTGTCTCTGGAGCCCCAACCTCACCCGCTCTGTGCCACTCTGAAGACTGCCGGGACCCCTGGAATGCAATTGTCTGTGAGTTC
1113▶ L S E C P L E P P S L T R L C A T L K D C P G P L E L Q L S C E F
4001 CTGAGTGACCAGAGCCTGGAGACTCTACTGGACTGCTTACCTCAACTCCCTCAGCTGAGCCTGTGCAGCTGAGCCAGACGGGACTGTCCCCGAAAAGCC
1147▶ L S D Q S L E T L L D C L P Q L P Q L S L L Q L S Q T G L S P K S
4101 CCTTCTGTCTGGCAACCTTAAAGCCTGTGTCCACGGGTTAAAAAGTGGATCTCAGTCCCTGCACCATGCAACTTTGACTTTCAGATCCAACGAGGA
1180▶ P F L L A N T L S L C P R V K K V D L R S L H H A T L H F R S N E E
4201 GGAGGAAGCGTGTGCTGTGGCAGGTTACAGGCTGCAGCCTCAGCCAGGAGCACGTAGAGTCACTCTGCTGGTGTGAGCAAGTGTAAAGACCTCAGC
1213▶ E E G V C C G R F T G C S L S Q E H V E S L C W L L S K C K D L S
4301 CAGGTGGATCTCTCAGCAAACTGCTGGGCAGACGGACTCAGATGCCTTCTGGAATGCTCGCCGAGGTGCCCATCTCCGGTTTGTCTGATCTGAGTC
1247▶ Q V D L S A N L L G D S G L R C L L E C L P Q V P I S G L L D L S
4401 ACAACAGCATTTCTCAGGAAAGTGCCTGTACTGTGGAGACTGCCCTCCTGCCACGTGTCGGGAGGCTCAGTGAACCTGGGCTCTGAGCAGAG
1280▶ H N S I S Q E S A L Y L L E T L P S C P R V R E A S V N L G S E Q S
4501 CTTCCGATTCACTTCTCAGAGAGGACCAGGCTGGGAAGACTCAGGCTAAGTGTGAGTGCAGCTTCCGGCCAGAGCAGTGTCCAGGCTGCCACCGGC
1313▶ F R I H F S R E D Q A G K T L R L S E C S F R P E H V S R L A T G
4601 TTGAGCAAGTCCCTGCAGCTGACGGAGCTCAGCTGACCCAGTGTGCTGGCCAGAGCAGCTGGCCATCTCCTGAGCTTGGTGGGGCAGCCCGCA
1347▶ L S K S L Q L T E L T L T Q C C L G Q K Q L A I L L S V G R P A
4701 GGCTGTTACGCTCAGGTTGCAGGAGCGTGGGCGGACAGAGCAGGTTCTCTCCCTGTTAGAAGTCTGCGCCAGGCTCAGGCACTGACTGAAAT
1380▶ G L F S L R V Q E P W A D R A R V L S L L E V C A Q A S G S V T E I
4801 CAGCATCTCCGAGACCCAGCAGAGCTCTGTGTCCAGCTGGAATTTCTCCTGCCAGGAAGAAATCCAGAAGCTGTGGCACTCAGTGTGGCTCACTGTGAC
1413▶ S I S E T Q Q Q L C V Q L E F P R Q E E N P E A V A L R L A H C D
4901 CTTGGAGCCACCAAGCCTTCTGTGGGAGCTGATGGAGACATGTCCAGCTGCAGCAGCTCAGCTTGTCTCAGGTTAACCTCTGTGAGGAGCATG
1447▶ L G A H H S L L V G Q L M E T C A R L Q Q L S L S Q V N L C E D D
5001 ATGCCAGTTCCTGCTGTCAGAGCCTCTGCTGTCCCTCTGAGCTGAAGACATTTCCGCTGACCTCCAGCTGTGTGAGCACCGAGGGCCTCGCCA
1480▶ D A S S L L L Q S L L L S L S E L K T F R L T S S C V S T E G L A H
5101 CCTGGCATCTGGTCTGGGCCACTGCCACCACTGGAGGAGCTGGACTTGTCTAACAATCAATTTGATGAGGAGGGCACCAGGCGCTGATGAGGGCCTT
1513▶ L A S G L G H C H L E E L D L S N N Q F D E E G T K A L M R A L

PshAI (5227)

5201 GAGGGAAATGGATGCTAAAGAGGCTGGACCTCAGTCACTTCTGCTGAACAGCTCCACCTTGGCCTTGTACTCACAGACTAAGCCAGATGACCTGCC
1547▶ E G K W M L K A R L D L S H L L L N S S T L A L L T H R L S Q M T C
5301 TGCAGAGCCTCAGACTGAACAGGAAACAGTATCGGTGATGTCCGTTGCTGCCACCTTTCTGAGCTCAGGGCTGCCACCAGCTAGAGGCTGGACTT
1580▶ L Q S L R L N R N S I G D V G C C H L S E A L R A A T S L E E L D L
5401 GAGCCACAACAGATTGGAGAGCTGGTGTCCAGCACTTAGCTACCATCTGCTGGGCTGCCAGAGCTCAGGAAGATAGACCTCAGGGAATAGCATC
1613▶ S H N Q I G D A G V Q H L A T I L P G L P E L R K I D L S G N S I

KasI (5546)

BamHI (5586)

5501 AGCTCAGCCGGGGAGTGCAGTTGGCAGAGTCTCTCGTCTTTGACGGCGCTGGAGAGTTGATGCTTGGCTGCAATGCCCTGGGGGATCCCACAGCC
1647▶ S S A G G V Q L A E S L V L C R L E E L M L G C N A L G D P T A
5601 TGGGGCTGGCTCAGGAGCTGCCAGCAGCTGAGGGTCTACACCTACCATTAGCCATCTGGGCCAGGTGGGGCCTGAGCCTGGCCAGGCTCTGGA
1680▶ L G L A Q E L P Q H L R V L H L P F S H L G P G G A L S L A Q A L D

BamHI (5701)

5701 TGGATCCCCCATTGGAAGAGATCAGCTTGGCGGAAAACAACCTGGCTGGAGGGTCTGCGTTTCTGTATGGAGCTCCGCTGCTCAGACAGATAGAC
1713▶ G S P H L E E I S L A E N N L A G G V L R F C M E L P L L R Q I D
5801 CTGGTTTCTGTAAGATTGACAACAGACTGCCAAGCTCTCACCTCCAGCTTACAGAGCTGCCCTGCCCTGGAAGTAATCTTGTCTGCTGGAATCTCC
1747▶ L V S C K I D N Q T A K L L T S S F T S C P A L E V I L L S W N L

NgoMIV (5948)

5901 TCGGGGATGAGGAGCTGCCAGCTGGCCAGGTGCTGCCAGATGGGCGGCTGAAGAGAGTGGACCTGGAGAAGAATCAGATCACAGCTTTGGGGG
1780▶ L G D E A A A E L A Q V L P Q M G R L K R V D L E K N Q I T A L G A
6001 CTGGCTCCTGGCTGAAGGACTGGCCAGGGTCTAGCATCAAGTCAATCCGCTCTGGAATAACCCATTCCCTGCGACATGGCCAGCAGCTGAAGAGC
1813▶ W L L A E G L A Q G S S I Q V I R L W N N P I P C D M A Q H L K S

NheI (6163)

6101 CAGGAGCCAGGCTGACTTTGCCTTTTGAACAACAGCCCAAGGCCCTTGGGGTACTTGAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGT
1847▶ Q E P R L D F A F F D N Q P Q A P W G T •
6201 TTGGACAACCAACTAGAATGAGTGAAGAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACA

EcoRI (6397)

6301 AGTTAACAACAACAAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAA
6401 TTCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGC

6501 CAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTTCCAAGGTTTGAAGTACTCTTCATTTCTTTATGTTT

6601 TAAATGCACTGACCTCCACATTCCCTTTTATAGTAAAAATATTCAGAAAATATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAAT

6701 CCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAAATTGGACAGCAAGAAAAGCGAGCTT

6801 CTAGCTTTAGTTCCTGGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTTCATCTCAATGAGCACAAAGCAGTCAGGAGCA
141 • N R T Y K L P I L E E I T T K V L K G N M E I L V F C D P A

6901 TAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGT
109 Y D S I L E R C M G C P S V V R I S R D V E D S Y P H R V A V I T D

7001 CAAAGTCTTCTGCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCCAGCAGACAGTGCCTGCCAATGTAGGCTCAATGTGGACAGCAGAGAT
76 F D K Q G N S V A S G I A I A E A C V T V R G I Y A E I H V A S I

7101 GATCTCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGCCTTGTTCCTCATAGAGCATGGTGCCTTCTCAGTGGCGACCTCCACAGCTCCAGA
43 I E G T K T R I A A G V H H K N D E Y L M T I K E T A V E V L E L

7201 TCCTGCTGAGAGATGTTGAAGGTCCTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAG
9 D Q Q S I N F T K M

7301 CGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGTTATATAGACCTCCCACCGTACACGCTACCGCCATTTGCGTCAATGGGG

7401 CGGAGTTGTTACGACATTTTGGAAAGTCCCCTGTTGATTTACTAGTCAAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTC

7501 AAACCGTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCAAAGTAGGAAAGTCCCAT

7601 AAGGTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGT

7701 GGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGAAACATACGTCAATATTGACGTCAATGGCGGGG

7801 GTCGTTGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGAGGTTAAITTAAGAACATGTGAGCAAAAAGGCCAGCAAAAAGGCCAGGA
PacI (7857)

7901 ACCGTA AAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGAC

8001 AGGACTATAAAGATACCAGCGTTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTTCGACCCTGCCGCTTACCGGATACCTGTCGCTTTCTCCCT

8101 TCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTT
ApaLI (8181)

8201 AGCCCGACCGCTGCGCTTATCCGGTAAGTATCGTCTTGTAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAG

8301 CAGAGCGAGGTATGTAGGCGGTCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGTGAAG

8401 CCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAAACAAACCACCGTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATTACGGCA

8501 GAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTA
PacI (8597)

8601 ATTAACATTTAAATC AGCGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATC
EagI (8617) Swal (8606) NotI (8616)

8701 AAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTCCCAGTGCAGGTGCCAGAACATTTCTCTATCGAA