



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAAGTGGCGGGGTAAGTGGAAAGTGTGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203)
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGGCC
Bsu36I (291)
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTTGTTCGTTT

KasI (535)
AgeI (552)
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCCCTACCTGAGATCACCGGTGCATCATGAGGACAGGCTGGCGACCCCTCGCCGCCGGGG
1 M R T G W A T P R R P A G

SapI (612)
601 GCTCCTCATGCTGCTTCTGTTTCTCGATCTCGCGGAGCCCTCTGGCCGCGCAGCTAATGACCCCTTACCACATCGTCCATGGAAATACGGGCAAGTGC
13 L L M L L F W F F D L A E P S G R A A N D P F T I V H G N T G K C
701 ATCAAGCCAGTGTATGGTGGATAGTAGCAGACACTGTGATGAACTGAGGACAAAGTTATGGAAGTGGGTGCCAGCATCGGCTTTTCAATTCGACT
47 I K P V Y G W I V A D D C D E T E D K L W K W V S Q H R L F H L H
801 CCCAAAAGTGCCTTGGCCTCGATATTACCAAATCGGTAATGAGCTGAGAATGTTGAGTGTGACTCCAGTGCATGCTGTGGTGGAAATGTGAGCACCA
80 S Q K C L G L D I T K S V N E L R M F S C D S S A M L W W K C E H H

Acc65I (920)
NsiI (965)
901 CTCTGTACGGAGCTGCCCGTACCGGCTGGCTCTGAAGGATGGACATGGCAGCAATCTCAAATGCATCTGATGTCTGGAAGAAAGGAGGCTCAGAG
113 S L Y G A A R Y R L A L K D G H G T A I S N A S D V W K K G G S E

BglII (1028)
BspHI (1023)
1001 GAAAGCCTTTGTGACCAGCCTTATCATGAGATCTATACAGAGATGGGAACTCTATGGGAGACCTTGTGAATTTCCATTCTTAATTGATGGGACCTGGC
147 E S L C D Q P Y H E I Y T R D G N S Y G R P C E F P F L I D G T W

BspHI (1101)
BsaBI (1117)
1101 ATCATGATTGCATTCTTGATGAAGATCATAGTGGCCATGGTGTGCCACCACCTTAAATTAATGAATATGACCGAAAGTGGGCATCTGCTTAAAGCCTGA
180 H H D C I L D E D H S G P W C A T T L N Y E Y D R K W G I C L K P E
1201 AAACGGTTGTGAAGATAATTGGGAAAAGAACGAGCAGTTTGAAGTGTCTACCAATTAATACTCAGACGGCTTTTCTTGGAAAAGCCTTATGTTTCA
213 N G C E D N W E K N E Q F G S C Y Q F N T Q T A L S W K E A Y V S
1301 TGTCAGAATCAAGGAGCTGATTTACTGAGCATCAACAGTGTCTGATTAACCTTACCTTAAAGAAAAAGAGGCATTGCTAAGATTTCTGGATTGGTT
247 C Q N Q G A D L L S I N S A A E L T Y L K E K E G I A K I F W I G

ApaLI (1484)
1401 TAAATCAGCTATACTCTGCTAGAGGCTGGGAATGGTCAGACCACAAACCATTAACTTTCTCAACTGGGATCCAGACAGGCCCACTGACCTACTATAGG
280 L N Q L Y S A R G W E W S D H K P L N F L N W D P D R P S A P T I G
1501 TGGCTCCAGCTGTGCAAGAAATGGATGCTGAGTCTGGTCTGTGGCAGAGCTTTTCTGTGAAGTCAACTGCCCTATGTCTGCAGGAAACCATTAAATAAT
313 G S S C A R M D A E S G L W Q S F S C E A Q L P Y V C R K P L N N

HpaI (1608)
1601 ACAGTGGAGTTAACAGATGTCTGGACATACTCAGATACCCGCTGTGATGCAGGCTGGCTGCCAAATAATGGATTTTGTCTATCTGCTGGTAAATGAAAGTA
347 T V E L T D V W T Y S D T R C D A G W L P N N G F C Y L L V N E S
1701 ATTCCTGGGATAAGGCACATGCGAAATGCAAAGCCTTCAGTAGTGACCTAATCAGCATTCTCTAGCAGATGTGGAGGTGGTTGTCAAAAACTCCA
380 N S W D K A H A K C K A F S S D L I S I H S L A D V E V V V T K L H

EcoRV (1807)
1801 TAATGAGGATATCAAAGAAGAAGTGTGGATAGGCCTTAAAGAACATAAACATACCAACTTTATTTTCAGTGGTCCAGATGGTACTGAAGTTACTCTAACATAT
413 N E D I K E E V W I G L K N I N I P T L F Q W S D G T E V T L T Y
1901 TGGGATGAGAATGAGCCAAATGTTCCCTACAATAAGACGCCCACTGTGTTTCTACTTAGGAGAGCTAGGTGAGGAAAGTCCAATCATGTGAGGAGA
447 W D E N E P Y N K T P N C V S Y L G E L G W K V S C E E
2001 AACAAAAATGTATGCAAGAGAAAAGGAGAAAAACTGAATGACGCAAGTCTGATAAGATGTCTCCAGATGAGGGCTGGAAGAGACATGGAGAAAC
480 K L K Y V C K R K G E K L N D A S S D K M C P P D E G W K R H G E T
2101 CTGTTACAAGATTTATGAGGATGAGTCCCTTTTGAACAACATGCAATCTGACTATCACTAGCAGATTTGAGCAAGAATACCTAAATGATTTGATGAAA
513 C Y K I Y E D E V P F G T N C N L T I T S R F E Q E Y L N D L M K
2201 AAGTATGATAAATCTCAAAGAAAATCTTGGACTGGCCTGAGAGATGTAGATTCTTGTGGAGAGTAACTGGGCAACTGTTGGTGGAAAGGCGGG
547 K Y D K S L R K Y F W T G G L R D V D S C G E Y N W A T V G G R R R

XmaI (2340)
2301 CTGTAACCTTTTCCAACCTGGAATTTTCTTGGCCAGCTTCCCGGGCGGCTGCGTGGCTATGTCTACTGGAAAGTCTGTTGGAAAGTGGGAGGTGAAGGA
580 A V T F S N W N F L E P A S P G G C V A M S T G K S V G K W E V K D

PshAI (2482)
2401 CTGCAGAAGCTTCAAAGCACTTTCAATTTGCAAGAAAATGAGTGGACCCCTTGGGCTGAAGAAGCATCCCCTAAGCCTGATGACCCCTGTCTGAAAGC
613 C R S F K A L S I C K K M S G P L G P E E A S P K P D D P C P E G
2501 TGGCAGAGTTTCCCGCAAGTCTTTCTGTTATAAGGTATTCCATGCAGAAAAGATTAAGAAAAGAGGAACTGGGAAGAAGCTGAACGATTCTGCCAAG
647 W Q S F P A S L S C Y K V F H A E R I V R K R N W E E A E R F C Q

XmnI (2647)
2601 CCCTTGGAGCACACCTTTCTAGCTTCAGCCATGTGGATGAAATAAAGGAATTTCTTCACTTTTTAACGGACAGTTCAGTGGCCAGCATTGGCTGTGGAT
680 A L G A H L S S F S H V D E I K E F L H F L T D Q F S G Q H W L W I
2701 TGGTTTGAATAAAGGAGCCAGATTTACAAGGATCCTGGCAATGGAGTGTACACCACTGTCTACTATTATCATGCCAAATGAGTTTTCAGCAGGAT
713 G L N K R S P D L Q G S W S D R T P V S T I I M P N E F Q Q D
2801 TATGACATCAGAGACTGTCTGTCTCAAGGATTTTATAGGCCATGGCGAAGAGGCTGGCATTCTATGATGATAGAGAATTTATTTTATTTGAGGCCTT
747 Y D I R D C A A V K V F H R P W R R G W H F Y D D R E F I Y L R P

EcoRI (2990)

2901 TTGCTTGTGATACAAAACCTTGAATGGGTGTGCCAAATCCAAAAGGCCGTACTCCAAAAACACCAGACTGGTACAATCCAGACCGTGGTGAATTCATGG
780>F A C D T K L E W V C Q I P K G R T P K T P D W Y N P D R A G I H G
3001 ACCTCCACTTATAATTTGAAGGAAGTGAATATTGGTTTGTGCTGATCTTCACCTAAACATGAAGAAGCCGCTGTACTGTGCCAGCAATCACAGCTTT
813>P P L I I E G S E Y W F V A D L H L N Y E E A V L Y C A S N H S F
3101 CTTGCAACTATAACATCTTTTGTGGGACTAAAAGCCATCAAAAACAAAATAGCAAATATATCTGGTGATGGACAGAAGTGGTGGATAAGAATTAGCGAGT
847>L A T I T S F V G L K A I K N K I A N I S G D G Q K W W I R I S E

EcoRV (3230)

BsrGI (3275)

3201 GGCCAATAGATGATCATTTTACATACTCACGATATCCATGGCACCGCTTTCTGTGACATTTGGAGAGGAATGCTTGTACATGTCTGCCAAGACTTGGCT
880>W P I D D H F T Y S R Y P W H R F P V T F G E E C L Y M S A K T W L
3301 TATCGACTTAGGTAACCAACAGACTGTAGTACCAAGTTGCCCTTCATCTGTGAAAAATAATAGTTTCTTCTGTAGAGAAAACAGCCAGATTCTGCA
913>I D L G K P T D C S T K L P F I C E K Y N V S S L E K Y S P D S A
3401 GCTAAAGTGAATGTTCTGAGCAATGGATTCTTTTCAAGAATAAGTGTCTTAAAGATCAAACCCGTGTCTCACATTTTCTCAAGCAAGCGATACCT
947>A K V Q C S E Q W I P F Q N K C F L K I K P V S L T F S Q A S D T

BspEI (3570)

3501 GTCACCTCTATGGTGGCACCCCTTCTTCACTGTTGAGCCAGATTGAACAAGACTTTATTACATCTTGTCTCCGGATATGGAAGCTACTTTATGGATTGG
980>C H S Y G G T L P S V L S Q I E Q D F I T S L L P D M E A T L W I G
3601 TTTGCGCTGGACTGCTATGAAAAGATAAAACAAATGGACAGATAACAGAGAGCTGACGTACAGTAACTTTACCCATTATTGGTTAGTGGGAGGCTGAGA
1013>L R W T A Y E K I N K W T D N R E L T Y S N F H P L L V S G R L R
3701 ATACCAGAAAATTTTTTGGAGGAAGAGTCTCGTACCCTGTGCCCTAATACTCAACCTCCAAAATCACCGTTTACTGGGACGTGGAATTTTACATCCT
1047>I P E N F F E E E S R Y H C A L I L N L Q K S P F T G T W N F T S

DraIII (3811)

3801 GCAGTGAACGCGCACTTTGTCTCTCTGTGAGAAATATTCAGAAGTAAAAAGCAGACAGAGCTTGCAGAATGCTTCAAGAACTGTAAAGTATCTAAATAA
1080>C S E R H F V S L C Q K Y S E V K S R Q T L Q N A S E T V K Y L N N
BsrGI (3902)

3901 TCTGTACAAAATAATCCCAAAGACTCTGACTTGGCACAGTGTCTAAAAGGGAGTGTCTGAAAAGTAAACATGCAGCTGGTGAGCATCACGGACCTTACCAG
1113>L Y K I I P K T L T W H S A K R E C L K S N M Q L V S I T D P Y Q
4001 CAGGCATTCTCAGTGTGCGAGGCTCCTTCACTACTTCTTATGGATCGGACTCTTCACTCAAGATGATGAACTCAACTTTGGTTGGTGCAGATGGGA
1147>Q A F L S V Q A L L H N S S L W I G L F S Q D D E L N F G W S D G
4101 AACGTCTTCAATTTAGTGCCTGGCTGAACTAATGGGCAACTCGAAGACTGTGTAGTATTAGACACTGATGGATTCTGGAAAACAGTTGATTGCAATGA
1180>K R L H F S R W A E T N G Q L E D C V V L D T D G F W K T V D C N D
4201 CAATCAACAGGTGCTATTTGCTACTATTGAGAAATGAGACTGAAAAGAGGTCACCAACAGTGCAGTGTAAATGTCATCTCTGTTCTAAACT
1213>N Q P G A I C Y Y S G N E T E K E V K P V D S V K C P S V L N T
4301 CCGTGATACCATTTCAAGAACTGTTGCTACAATTTCAATAACAAAGAATAGGCATATGGCAACAACACAGGATGAAGTTCATACTAAATGCCAGAAAC
1247>P W I P F Q N C C Y N F I I T K N R H M A T T Q D E V H T K C Q K
4401 TGAATCCAAAATCACATATTCTGAGTATTGAGATGAAAAGGAGAATAACTTTGTTCTTGGCAACTGCTGTACTTCAATATATGGCTTCTATGGGTCAT
1280>L N P K S H I L S I R D E K E N N F V L E Q L L Y A M A S W V M
4501 GTTAGAATAAAGTATAGAAAATAGTCTTATGTGGTTTGATAAGACCCCACTGTCATATAACATTGGAGAGCAGGAAGCAACTATAAAAAATGAG
1313>L G I T Y R N K S L M W F D K T P L S Y T H W R A G R P T I K N E

ScaI (4618)

4601 AAGTTTTTGGCTGGTTTAAAGTACTGACGGCTTCTGGGATATTCAAACCTTTAAAGTATTGAAGAAGCAGTTTATTTTACCAGCACAGCATTCTTGCTT
1347>K F L A G L S T D G G F W D I Q T F K V I E E A V Y F H Q H S I L A
4701 GTAAAATGAAATGGTTGACTACAAAGAAGAATAATACTACTACTGCCACAGTTTATGCCATATGAAGATGGTATTTACAGTGTATTCAAAAAAGGT
1380>C K I E M V D Y K E E Y N T T L P Q F M P Y E D G I Y S V I Q K K V
4801 AACATGGTATGAAGCATTAAACATGTGTCTCAAAGTGGAGGTCACCTGGCAAGCGTTCAACAACAAAATGGCCAGCTCTTCTGGAAGATATTGAAAA
1413>T W Y E A L N M C S Q S G G H L A S V H N Q N G Q L F L E D I V K

BspHI (4935)

4901 CGTGATGGATTTCCACTATGGGTTGGGCTCTCAAGTCATGGAAGTGAATCAAGTTTTGAATGGTCTGATGGTAGTACATTTGACTATATCCCATGGA
1447>R D G F P L W V G L S S H D G S E S S F E W S D G S T F D Y I P W
5001 AAGGCCAAACATCTCCTGAAAATTTGTGTCTCTTGGATCCAAAAGGAAGTGGAAACATGAAAATGCAACTCTGTTAAGGATGGTGTATTTGTTATAA
1480>K G Q T S P G N C V L L D P K G T W K H E K C N S V K D G A I C Y K
5101 ACCTCAAAAATCTAAAAAGCTGTCCCGTCTTACATATTTCATCAAGATGTCAGCAGCAAAAAGAGAATGGGTACAGGTCCAGTACAAGGTCAGTGT
1513>P T K S K K L S R L T Y S S R C P A A K E N G S R W I Q Y K G H C
5201 TACAAGTCTGATCAGGCATTGCACAGTTTTTCAAGGCCAAAAAATTTGTTTCAAACATGATCACTCTGCAACTATCGTTTCCATAAAAGATGAAGATG
1547>Y K S D Q A L H S F S E A K K L C S K H D H S A T I V S I K D E D
5301 AGAATAAATTTGTGAGCAGACTGATGAGGGAAAATAATAACATTACCATGAGAGTTTGGCTTGGATTATCTCAACATTCTGTTGACCAGTCTTGGAGTTG
1580>E N K F V S R L M R E N N N I T M R V W L G L S Q H S V D Q S W S W

PshAI (5418)

5401 GTTAGATGGATCAGAAGTACATTTGTCAAATGGGAAAATAAAAGTAAGAGTGGTGTGGAAAGATGTAGCATGTTGATAGCTTCAAATGAAACTTGGAAA
1613>L D G S E V T F V K W E N K S K S G V G R C S M L I A S N E T W K

Bsp120I (5550)

5501 AAAGTTGAATGTGAACATGGTTTTGGAAGAGTTGTCTGCAAGTGCCTTGGCCCTGATTACACAGCAATAGCTATCATAGTTGCCACACTAAGTATCT
1647>K V E C E H G F G R V V C K V P L G P D Y T A I A I I V A T L S I

BstAPI (5640)

5601 TAGTTCTCATGGGCGGACTGATTTGGTTCTCTTCCAAAGGCACCGTTTGACCTGGCGGGTTTCTCATCAGTTCGATATGCACAAGGAGTGAATGAAGA
1680>L V L M G G L I W F L F Q R H R L H L A G F S S V R Y A Q G V N E D

NheI (5738)

5701 TGAGATTATGCTTCTTCTTCCATGACTAAATCTTCCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACACTAGAATGCA
1713>E I M L P S F H D •

HpaI (5876)

MfeI (5887)

5801 GTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTAAACAACAACAAATTGCATTCAT

EcoRI (5972)

5901 TTTATGTTTCAGTTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAAATTTAAAATACAGCATAGCAAACCT

6001 TTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCC

SapI (6154)

6101 TCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTTCCAAGGTTTGAAGTCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCC

SwaI (6225)

6201 CTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAAT

6301 ATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGG
141 • N R T Y K

SacI (6486)

6401 AGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACA
134 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 Y D S I L E R C M

6501 TGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCAC
101 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 F D K Q G N S V

6601 AGCAGACCAATGGCAATGGCTTCCAGCAGACAGTACCCTGCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTGATG
68 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 I E G T K T R I

XmnI (6792)

6701 GCCGCCCCGACATGGTGTCTTGTCTCATAGAGCATGGTGATCTTCTCAGTGGCGACTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGGTCT
34 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 D Q Q S I N F T K

AseI (6858)

6801 TCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACAGCGTGATGGCGTCTCCAGCTTATCT
1 M

SacI (6915)

6901 GACGGTTCATAAACGAGCTCTGTTATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTACGACATTTTGAAAA

SpeI (7013)

7001 GTCCCGTTGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATG

SnaBI (7141)

7101 TACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCC

7201 AGGCGGGCATTACCCTGATTGACGTCAATAGGGGGCTACTGGCATATGATACACTTGATGTACTGCAAGTGGCAGTTTACCCTAAATACTCCAC

7301 CCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGGGGCTGTTGGCGGTGAGCCAGGCGGG

PacI (7432)

SdaI (7424)

7401 CCATTTACGTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGGTGGC
7501 GTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTT
7601 CCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATA

ApaLI (7756)

7701 GCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCCTGCGCTTATCCGG

7801 TAACTATCGTCTTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCT

7901 ACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTG

8001 GTAGCTCTTGATCCGGCAACAAACCACCGCTGGTAGCGGTGTTTTTTTTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCC

EagI (8192)

PacI (8172) SwaI (8181) NotI (8191)

8101 TTTGATCTTTTCTACGGGCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTGGTATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCA

8201 ATAAAATATCTTTATTTTATTACATCTGTGTGTTGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAAACT

8301 AGCAAAATAGGCTGTCCCGAGTCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA