



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

HindIII (245)
Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCTTACAGCGCCCGCCGCCCTACCTGAGGCC
PvuII (239)
Bsu36I (291)
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCCTGTGGTGCCTCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
NgoMI (441)
NaeI (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGTCCTTTGTTTCGTTT

KasI (535) 501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCCTACCTGAGATCACCGGTCATCATGAGTCATGTGGCAGTGGAAAATGCGCTCGGGCTGGA
AgeI (552) 1► M S H V A V E N A L G L D
BspHI (560)

XcmI (600) 601 CCAGCAGTTTGTGGCCTAGACCTGAAGTCTTTCAGATAATCAGAGTGGAGGAAGTACAGCAAGCAAAGGGCGTTATATCCACCTCATTAAAGAACAGA
13► Q Q F A G L D L N S S D N Q S G G S T A S K G R Y I P P H L R N R

Ppu10I (761)
NsiI (761) 701 GAAGCTACTAAAGGATTCTATGACAAAGACAGTTCAGGGTGGAGTTCTAGTAAAGATAAGGATGCATACAGCAGTTTTGGATCACGGGTGATTCAAGAG
47► E A T K G F Y D K D S S G W S S S K D K D A Y S S F G S R G D S R
DraIII (782)
801 GGAAGTCTAGCTTCTTTGGAGATCGTGAAGTGGATCAAGGGGAAGTTTGTATGATCGTGGACGGGAGACTATGATGGCATTGGTGGCCGTGGAGATAG
80► G K S S F F G D R G S G S R G R F D D R G R G D Y D G I G G R G D R

DraIII (983)
901 AAGTGGCTTTGGCAAATTTGAAAGAGGTGAAATAGTGCCTGGTGTGACAAATCAGATGAAGATGACTGGTCAAAGCCACTCCACCAAGTGAACGATTG
113► S G F G K F E R G G N S R W C D K S D E D D W S K P L P P S E R L
1001 GAACAGAACTCTTTTCTGGAGGCAATACTGGGATTAACCTTTGAGAAAATATGATGACATTCCAGTCAAGCAACAGGCAACAACCTGCTCCACACATTG
147► E Q E L F S G G N T G I N F E K Y D D I P V E A T G N N C P P H I

SphI (1185)
1101 AAAGTTTTCAGTGTGTCGAGATGGGAGAAATTATTATGGGAAACATTGAGCTTACTCGTTATACTCGCCAACTCCAGTGCAGAAAGCATGCTATTCTCTAT
180► E S F S D V E M G E I I M G N I E L T R Y T R P T P V Q K H A I P I

PstI (1253) 1201 TATCAAAGAGAAAAGAGACTTGTAGGCTGTGCTCAAACAGGCTCTGGAAAACTGCAGCATTCTCTTGGCCATCTTGGATCAGATCTATGCTGATGGT
213► I K E K R D L M A C A Q T G S G K T A A F L L P I L S Q I Y A D G
BglII (1283)
1301 CCAGGAGAAGCTCTGAGGGCTATGAAGGAAAATGGAAGATATGGCCGTCGTAACAGTATCCAATCTCTTTGGTACTGGCACCAACGAGAGAATTGGCAG
247► P G E A L R A M K E N G R Y G R R K Q Y P I S L V L A P T R E L A

BglIII (1403) 1401 TGCAGATCTATGAGGAAGCCAGAAAATTCTCATACCGATCTAGAGTCCGTCCTTGCCTGGTTTATGGTGGTGTGAAATTTGGCCAGCAGATTCGAGACT
280► V Q I Y E E A R K F S Y R S R V R P C V V Y G G A E I G Q Q I R D L
XbaI (1438)
1501 AGAAGCTGGATGCCACTGTTAGTACCCACTCCAGGACGCTCTAGTGGATATGATGGAGAGAGGGAAGATCGGGTTAGACTTCTGCAAATACCTGGTGTTA
313► E R G C H L L V A T P G R L V D M M E R G K I G L D F C K Y L V L
1601 GATGAAGCTGACCGGATGTTAGATATGGGGTTTGAACCTCAGATACGAAGAATAGTTGAACAAGACACTATGCCTCAAAGGTTGTCYCCACACTATGA
347► D E A D R M L D M G F E P Q I R R I V E Q D T M P P K G V R H T M

Bsu36I (1717) 1701 TGTTTGTGCTACTTTTCTAAGGAAATACAGATGCTGGCCCGTGAATTTCTAGATGAGTACATATTTCTGGCTGTAGGAAGAGTTGGGTCTACTTCAGA
380► M F S A T F P K E I Q M L A R D F L D E Y I F L A V G R V G S T S E
1801 GAACATCACACAAAAAGTGGTTTGGGTGGAGGATAGACAAAAGTCAATTTCTGCTTACCTTCTAAATGCAACAGGCAAGGATTCCTGACTCTAGTG
413► N I T Q K V V W V E E I D K R S F L L D L L N A T G K D S L T L V

NcoI (1974) 1901 TTTGTGGAGACAAAAAGGGGGCAGATTGCTGGAGGATTTCTTATACCATGAAGGATATGCTTGTACCAGTATCCATGGAGACCCTTCTCAGAGAGATA
447► F V E T K K G A D S L E D F L Y H E G Y A C T S I H G D R S Q R D

EcoO109I (2006) 2001 GGGAAGAGCCCTTACCAGTTCGCTCAGGAAAAAGCCAAATTTAGTGGCTACAGCAGTAGCAGCAAGAGGACTGGATATTTCAAATGTGAAGCATGT
480► R E E A L H Q F R S G K S P I L V A T A V A A R G L D I S N V K H V
BsrBI (2022)
AseI (2101) 2101 TATTAATTTTGCCTGCTAGTGTATCGAAGAATATGTGCATCGCATAGGCCGTACAGGCCGTGGGAAACCTTGGTCTTGCACCTCATTCTTTAAT
513► I N F D L P S D I E E Y V H R I G R T G R V G N L G L A T S F F N
EcoRV (2122) 2201 GAAAGGAATATAAATATCACAAAGGATTTACTGGATCTTCTTGTGAAGCAAAAACAAGAAGTGCCTTCTTGGTTAGAGAACATGGCTTTTGAACACCACT
547► E R N I N I T K D L L D L L V E A K Q E V P S W L E N M A F E H H
SfiI (2149) 2301 ACAAGGTAGCAGTCGTGGACGTTCTAAGAGCAGTCGATTTAGTGGAGGGTTGGTGCAGAGACTACCGACAGTAGCGGTGCCAGCAGTTCCAGCTT
580► Y K G S S R G R S K S S R F S G G F G A R D Y R Q S S G A S S S S F

NcoI (2442) 2401 CAGCAGCAGCCGTGCAAGCAGCAGTCGAAGTGGTGGAGGTGGCCATGGCGCAGTCGAGGATTTGGTGGAGGTGGCTACGGAGGCTTTTACAACAGTGAT
613► S S S R A S S S R S G G G G H G G S R G F G G G Y G G F Y N S D
BstEII (2566) 2501 GGATATGGAGGAATTATAACTCCAGGGGTTGACTGGTGGGTAAGTGAAGCTGCTCTGTAGTAGTACCTGCCAAAGCTAGCTGGCCAGACATGA
647► G Y G G N Y N S Q G V D W W G N •

2601 TAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCAT

2701 TATAAGCTGCAATAACAAGTTAAACAACAACAATTGCATTCATTTTATGTTTCAGGTTACGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTC

2801 TACAAATGTGGTATGGAATTTCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATA

2901 GGCATCAGGGGCTTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTTTCATGGAGTTTAAAGATATAGTGATTTTCCCAAGGTTTGAAGTAGCT

3001 CTTTCTTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTAGTAAAATATTGAGAAATAATTTAAATACATCATTGCAATGAAAATAAATG

3101 TTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGA

3201 CAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGC
 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

3301 ACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCC
 115 V F C D P A 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

3401 TGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTC
 82 V A V I T D 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

3501 AATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGTAGGCCGCCCGACATGGTGTCTTGTCTCATAGAGCATGGTATCTTCTCAGTGGCG
 49 I H V A S I 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

3601 ACCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGGTTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGAT
 15 V E V L E L D Q Q S I N F T K M

3701 GATTAATTGTCAAACACAGCGTGGATGGCGTCTCCAGCTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGC

3801 CCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACCTCCATTGACGTCAATGGGGTGAGAC

3900 TTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTG

4000 CCAAGTAGGAAAGTCCATAAGGTCACTGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATAC

4100 ACTTGATGTACTGCCAAGTGGGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTA

4200 TTGACGTCAATGGGCGGGGTCGTTGGGCGGTACGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTG C A G G T T A A T T A A G A A C A T G T G A G C A A
 SdaI (4267) PstI (4268) PacI (4275) BspLU11I (4285)

4298 AAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAG

4398 TCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGGCTCTCTGTCCGACCCTGCCGTTACCGGA

4498 TACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCT

4598 GTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGC
 ApaLI (4599)

4698 AGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCCTAACACGGCTACACTAGAAGAACAGTATTT

4798 GGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA AAAAGATTGGTAGCTCTTGATCCGGCAAACAAACCAGCTGGTAGCGGTGTTTTTTGTTT

4898 GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGG

4998 GATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGTTTTTTGTGTAATCGT
 EagI (5035) PacI (5015) SwaI (5024) NotI (5034)

5098 AACTAACATACGCTCTCCATCAAAAACAAAACGAAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCG
 5198 AA