



PvuI (7)
SgfI (6) 1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)
Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTGGCCGAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACCGCGCCCGCCGCTACCTGAGGCC
PvuII (239)
Bsu36I (291)
301 GCCATCCACGCGGTTGAGTGCAGTCTGCCGCTCCCGCCTGTGGTGCCTCTGAAGTGCCTCCGCGCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTGGCTGACCCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

BstEII (555)
KasI (535) 501 TCTGTTCTGCGCGGTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTACCATTGGCGCGTTCTGCGCATCGCCTCAACTCCTATGA
AgeI (552) 1▶ M A P F L R I A F N S Y E
KasI (564)
FspI (575)

SdaI (609)
601 GCTGGGCTCCCTGCAGCGGAGGACGAGGCGAACCAGCCCTTCTGTGCCGTGAAGATGAAGGAGGCGCTCAGCACAGAGCGTGGGAAAACACTGGTGCAG
13▶ L G S L Q A E D E A N Q P F C A V K M K E A L S T E R G K T L V Q

SalI (729)
701 AAGAAGCCGACCATGTATCCTGAGTGAAGTGCAGCTTCGATGCCACATCTATGAGGGGCGCGTCATCCAGATTGTGCTAATGCGGGCAGCAGAGGAGC
47▶ K K P T M Y P E W K S T F D A H I Y E G R V I Q I V L M R A A E E

BstEII (810) 801 CAGTGTCTGAGGTGACCGTGGGTGTGTCGGTGTGGCCGAGCGCTGCAAGAAGAACAATGGCAAGGCTGAGTTCTGGCTGGACCTGCAGCCTCAGGCCAA
Eco47III (839) 80▶ P V S E V T V G V S V L A E R C K K N N G K A E F W L D L Q P Q A K
Bsu36I (889)

FspI (953) 901 GGTGTTGATGCTGTTTCAGTATTTCTGGAGGACGTGGATTGCAAACAGTCTATGCGCAGTGAGGACGAGGCCAAGTTCCCAACGATGAACCGCCGGGA
SacII (993) 113▶ V L M S V Q Y F L E D V D C K Q S M R S E D E A K F P T M N R R G

Tth111I (1090)
1001 GCCATCAAACAGGCCAAAATCCACTACATCAAGAACCATGAGTTTATCGCCACCTTCTTTGGGCAACCCACCTTCTGTTCTGTGTGCAAAGACTTTGTCT
147▶ A I K Q A K I H Y I K N H E F I A T F F G Q P T F C S V C K D F V

NsiI (1158) 1101 GGGGCTCAACAAGCAAGGCTACAAATGCAGGCAATGTAACGCTGCCATCCACAAGAAATGCATCGACAAGATCATCGGCAGATGCACCTGGCACCGCGGC
180▶ W G L N K Q G Y K C R Q C N A A I H K K C I D K I I G R C T G T A A
SacII (1193)
1201 CAACAGCCGGGACTATATCCAGAAAGAAGCTTCAACATCGACATGCCGACCGCTTCAAGGTTCAACACTACATGAGCCCCACCTTCTGTGACCC
213▶ N S R D T I F Q K E R F N I D M P H R F K V H N Y M S P T F C D H
1301 TGGCGCAGCTGCTCTGGGGACTGGTGAAGCAGGGATTAAGTGTGAAGACTGCGGCATGAATGTGCACCATAAATGCCGGGAGAAGTTGGCCAACTCT
247▶ C G S L L W G L V K Q G L K C E D C G M N V H H K C R E K V A N L

HindIII (1414) 1401 GCGGCATCAACGAGAAGCTTTTGGCTGAGGCTTGAACCAAGTCAACAGAGAGCCTCCCGGAGATCAGACTCAGCCTCCTCAGAGCCTGTTGGGATATA
StuI (1427) 280▶ C G I N Q K L L A E A L N Q V T Q R A S R R S D S A S S E P V G I Y

BglIII (1568) 1501 TCAGGTTTTCGAGAAGAAGACCGGAGTTGCTGGGGAGGACATGCAAGACAACAGTGGGACCTACGGCAAGATCTGGGAGGGCAGCAGCAAGTGAACATC
BstAPI (1583) 313▶ Q G F E K K T G V A G E D M Q D N S G T Y G K I W E G S S K C N I

ScaI (1679)
1601 AACAACTTCATCTCCACAAGGCTCTGGGCAAAGGCGAGCTTCGGGAAGGTGCTGCTTGGAGAGCTGAAGGGCAGAGGAGTACTTTGGCCATCAAGGCC
347▶ N N F I F H K V L G K G S F G K V L L G E L K G R G E Y F A I K A
1701 TCAAGAAGGATGTGGTCTGATCGACGACGACGTGGAGTGCACCATGTTGAGAAGCGGTGCTGACACTTGGCCGAGAGAATCCCTTCTCACCACCT
380▶ L K K D V V L I D D D V E C T M V E K R V L T L A A E N P F L T H L
1801 CATCTGCACCTCCAGACCAAGGACCACTGTTCTTTGTGATGGAGTCTCAACGGGGGGACCTGATGTACCACATCCAGGACAAAGGCCGCTTTGAA
413▶ I C T F Q T K D H L F F V M E F L N G G D L M Y H I Q D K G R F E
1901 CTCTACCGTCCACGTTTATGCCGCTGAGATAATGTGTGAGTGCAGTTTCTACACAGCAAGGCATCATTTACAGGACCTCAAACCTGGACAATGTGC
447▶ L Y R A T F Y A A E I M C G L Q F L H S K G I I Y R D L K L D N V

XmnI (2052) 2001 TGCTGACCGGGATGGCCACATCAAGATTGCCACTTTGGGATGTGCAAAGAGAATATTCGGGGAGAGCCGGGCCAGCCTTCTGCGGCCACCCCTGA
BstAPI (2078) 480▶ L L D R D G H I K I A D F G M C K E N I F G E S R A S T F C G T P D
2101 CTATATCGCCCTGAGATCTACAGGGCTGAAGTACACATTCTGTGGACTGGTGGCTTTTGGGGTCTTCTGTACGAGATGCTCATTGGCCAGTCC
513▶ Y I A P E I L Q G L K Y T F S V D W W S F G V L L Y E M L I G Q S

XcmI (2285)
2201 CCCTTCATGTTGATGATGAGGATGAACCTTTCGAGTCCATCCGTGTGGACACGCCACATTATCCCGCTGGATCACCAAGGATCCAAAGACATCCTGG
547▶ P F H G D D E D E L F E S I R V D T P H Y P R W I T K E S K D I L
2301 AGAAGCTCTTTGAAAGGGAACCAACCAAGAGGCTGGGAGTGACCGGAAACATCAAATCCACCCCTTCTTCAAGACCATAAAGTGGACTCTGCTGAAAA
580▶ E K L F E R E P T K R L G V T G N I K I H P F F K T I N W T L L E K

BssHII (2485)
AscI (2484)
2401 GCGGAGTTGGAGCCACCTTTCAGGCCAAAGTGAAGTCAACAGAGACTACAGTAACTTTGACCAGGAGTTCCTGAACGAGAAGGCGCGCTCTCCTAC
613▶ R R L E P P F R P K V K S P R D Y S N F D Q E F L N E K A R L S Y
2501 AGCGACAAGAACCTCAGTCCATGGACCAGTCTGCATTGCTGGCTTCTCCTTTGTGAACCCAAATTCGAGCACCTCCTGGAAGATTGAGGTTCT
647▶ S D K N L I D S M D Q S A F A G F S F V N P K F E H L L E D •

NheI (2611)
2601 GGACAGATCAGGCTAGCCCTGCCTAGTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAACCAACTAGAATGCAGTGAATAAATGCTTTA

2701 **HpaI (2759)** **MfeI (2770)**
TTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTTCATTTTATGTTTCAGGTTCA

2801 **EcoRI (2855)**
GGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAG

2901 CCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCCTCACCTTCTTTCATGGA

3001 **SspI (3094)**
GTTAAGATATAGTGATTTTCCCAAGTTTGAAGTAGCTCTTCATTTCTTTATGTTTAAATGCACTGACCTCCACATTCCCTTTTATAGTAAATATT

3101 **SwaI (3108)**
CAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAG

3201 TTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCCT
141 • N R T Y K L P I L E E

3301 **SacI (3369)**
AATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACC
129 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 G C P S V

3401 ACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAA
95 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 A S G I A I

3501 **StuI (3533)**
TGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTG
62 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 A A G V H H

3601 **XmnI (3675)**
CTTGTGTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTA
29 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 M

3701 **AseI (3741)** **SacI (3798)**
TAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTATCTGACGGTTCACTAAACGA

3801 **SpeI (3896)**
GCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTGCGTCAATGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGATTTACTA

3901 GTCAAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCAT

4001 **SnaBI (4024)**
CATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCG

4101 **NdeI (4129)**
TCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCCTAAATACTCCACCCATTGACGTCAATGGA

4201 AAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTGTTGGGCGGTCAGCCAGGCGGGCCATTTACCCTAAGTTA

4301 **SdaI (4307)** **PacI (4315)** **BspLU11I (4325)**
TGTAACGCCTGCAGGTTAAATAAGAACATGTGAGCAAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCC

4401 GCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCTTTCCCTGGAAGCTCCCT

4501 CGTGCGCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTAT

4601 CTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCTTATCCGTAACATATCGTCTTGAGT

4701 CCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTG

4801 GTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTTTGATCCGGC

4901 AAACAACCACCGCTGGTAGCGGTGTTTTTTGTTTGAAGCAGCAGATTACCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGG

5001 **EagI (5075)** **PacI (5055)** **SwaI (5064)** **NotI (5074)**
GGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGCCGCAATAAAATATCTTTATTT

5101 TCATTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAAACAAAACGAAAACAAAACAACTAGCAAAATAGGCTGTCC

5201 CCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA