



SgfI (6) 1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAGGTGGCGGGGTAAGTGGGAAAGTGTGCTGTACTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC

HindIII (245)
Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCGCTACCTGAGGCC
PvuII (239)
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGCTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
Bsu36I (291)

NgoMIV (441)
401 GGGCTTTGTCCGGGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BstEII (555)
KasI (535) 501 TCTGTTTGTCCGCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTACCCATGGCGGACCCGAGGTGTGCTGCTTCATACCAAAAT
AgeI (552) **NeoI (560)**
1 M A D P E V C C F I T K I
Bsp120I (694)

601 CCTGTGCGCCACGGGGCCGCATGGCCCTGGACGCGCTGCTCCAGGAGATCGCGTGTCTGAGCCGAGCTCTGTGAGGTGCTGCAGGTGGCCGGGCC
13 L C A H G R M A L D A L L Q E I A L S E P Q L C E V L Q V A G P

XmaI (770)
XhoI (764) 701 GACCGTTTGTGGTGTGGAGACCGGCGGAGCCGGGATCACCCGATCGGTGGTGGCCACCACTCGAGCCCGGGTCTGCCGTCGAAAGTACTGCCAGA
47 D R F V V L E T G G E A G I T R S V V A T T R A R V C R R K Y C Q
ScaI (788)

801 GACCTGCGATAACCTGCATCTCTGCAAACTCAACTTGTGGGCGGTGCAACTATTCGAGTCCGAGCGGAATTTATGCAAATATTCTCATGAGTTCT
80 R P C D N L H L C K L N L L G R C N Y S Q S E R N L C K Y S H E V L
901 CTCAGAAGAGAATTCAAAGTCTGAAAATACGAACTCTCTGGACTGAACAAAGAGGAATTAGCAGTGTCTCTCTCAAAGTATCCTTTTTTATG
113 S E E N F K V L K N H E L S G L L N K E E L A V L L L Q S D P F F M
1001 CCCGAGATATGAAAAGTTATAAGGGAGAGGGTCGCGCAGCAGATTTGTAACCGCAGCCACCGTGTTCAGACTCCACATCTGTGACCACTTCCCGGAG
147 P E I C K S Y K G E G R Q Q I C N Q Q P P C S R L H I C D H F T R

Tth111I (1189)
1101 GAACTGTCGTTTTCCAACTGCCTCCGGTCCATAACCTGATGGACAGAAAGTGTGGCCATCATGAGGGAGCACGGGCTGAACCCCGACGTGGTCCA
180 G N C R F P N C L R S H N L M D R K V L A I M R E H G L N P D V V Q

Bsp120I (1249)
1201 GAACATCCAGGACATCTGCAACAGCAAGCACATGCAGAAGAATCCCCAGGGCCAGAGCTCCTTCTTACATCGTAGAAACATGGCATATAGGGCTAGA
213 N I Q D I C N S K H M Q K N P P G P R A P S S H R R N M A Y R A R
1301 AGCAAGAGTAGAGATCGGTTCTTTCAGGCGAGCAAGAATTTCTTGGCTGTGCTTCAGCGTGTGCTGAGAGGTCTGCACACCTAGTCCAGATCAGATCA
247 S K S R D R F F Q G S Q E F L A S A S A S A E R S C T P S P D Q I

Bsu36I (1487)
1401 GCCACAGGGCTTCCCTGGAGGACGCGCCCTGGACGATCTCACCCGCAAGTTCACGTATCTGGGGAGTCAGGATCGCGCTCGGCTCCCTCAGGCTCGTC
280 S H R A S L E D A P V D D L T R K F T Y L G S Q D R A R P P S G S S
1501 CAAGGCTACTGATCTTGGAGAACAAAGTCAGGCCGGGACAAGCCAGAGTTTTAGAGAACGGCAGTCAAGAGGACCTTTCATGAAATCCAGGCAGC
313 K A T D L G G T S Q A G T S Q R F L E N G S Q E D L L H G N P G S

KasI (1669)
1601 ACTTACCTTGCTTCCAATCAACATCAGCCCCAACTGGAAGAGCCTCACATCTGGACGAATGACCAAGGCGCCAGGAGAAAGACTGTGTTTTCTCCCA
347 T Y L A S N S T S A P N W K S L T S W T N D Q G A R R K T V F S P

BsrBI (1713)
1701 CGTACTCTGCCGCCGCTCTTCTCTTGGCTCTCTGCAAACACCTGAAGCTGTGACCACCAGAAAGGGCACAGGCTTGTCTTCTCAGACTACAGGATCAT
380 T L P A A R S S L G S L Q T P E A V T T R K G T G L L S S D Y R I I
1801 CAATGGCAAAAGTGAACCTCAGGACATCCAGCCTGGCCCTTTTTAATAATAATGCTGATGGAGTGGCCACAGATATAAATCTTACCAGATCCTTAAT
413 N G K S G T Q D I Q P G P L F N N N A D G V A T D I T S T R S L N

AvrII (1939) 1901 TACAAAAGCACTAGCAGCGGTACAGAGAAATATCATCACCTAGGATTCAGGATGCTGGACCTGCTTCCCGAGATGTCCAGGCCACTGGCAGAATCGCAG
447 Y K S T S S G H R E I S S P R I Q D A G P A S R D V Q A T G R I A
BstXI (1977)

HpaI (2026)
2001 ATGATGCTGACCCAAGAGTAGCACTTGTAAACGATTCTTTATCTGATGTACAAAGTACCACATCTTCTAGGGTGGATGATCATGACTCAGAGAAATTTG
480 D D A D P R V A L V N D S L S D V T S T T S S R V D D H D S E E I C

AgeI (2168)
2101 TCTTGACCATCTGTGAAGGGTTGTCCGCTTAATGGTAGCTGCAGCAAAGTCCACTTCCATCTGCCTTACCGTGGCAGATGCTTATTGGTAAAACCTGG
513 L D H L C K G C P L N G S C S K V H F H L P Y R W Q M L I G K T W
2201 ACGGACTTTGAGCACATGGAGACGATCGAGAAAAGGCTACTGTAACCCCGGAATCCACCTCTGTTCTGTAGGAAGTTATAACAATTTTCGGTAATGA
547 T D F E H M E T I E K G Y C N P G I H L C S V G S Y T I N F R V M
2301 GTTGATTCCTTTCCATCCGACGCTCTCCACTCTTCTTCTGTCACCAAGCCAGCAATCTGTCTTCCACCAAAATGGATTTGGTATTGGAAGAA
580 S C D S F P I R R L S T P S S V T K P A N S V F T T K W I W Y W K N

SalI (2458)
2401 TGAATCTGGCACATGGATTAGTATGGAGAAGAGAAAGACAAACGGAAAAATCAAACGTCGACTTTCATACCTGGAGTCTCTCTATCAATCTGTCCG
613 E S G T W I Q Y G E E K D K R K N S N V D S S Y L E S L Y Q S C P
2501 AGGGGAGTTGTGCCATTTAGGCGGGCTCACGGAATATGAGCTGAGTTTCAAAGGGATGATTCAGACAAACATAGCTTCCAAAACCTAAAAGGATGTCA
647 R G V V P F Q A G S R N Y E L S F Q G M I Q T N I A S K T Q K D V

NheI (2662)
2601 TCAGAAGACCAACATTTGTGCCTCAGTGGTATGTGCAGCAGATGAAGAGAGGGCCAGAGTAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTT
680 I R R P T F V P Q W Y V Q Q M K R G P E •

2701 TGGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAA

HpaI (2800) MfeI (2811) EcoRI (2896)
2801 GTTAAACAACAACAATTGCATTCATTTTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAAT

2901 TCTAAAATACAGCATAGCAAAC TTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCC

3001 AATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGATTTTTCCCAAGTTTGAAGTACTGCTCTTCATTTCTTTATGTTTT

SspI (3135) SmaI (3149)
3101 AAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTTTTATTAGCGAGAATC

3201 CAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTC

3301 TAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCAT
-1-1-1 • 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 Y

BstXI (3439)
3401 AGTCAGAGATGAGCTCTGACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTG
109 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

StuI (3574)
3501 AAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTT CAGCACAGACAGT GACCCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATG
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

3601 ATCTCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTCTTGTCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGAT
42 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 D

XmnI (3716) AseI (3782)
3701 CCTGCTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCTGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGC
9 Q Q S I N F T K M

3801 GTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTACCGCCATTTCGCTCAATGGGGC

SpeI (3937)
3901 GGAGTTGTTACGACATTTTGAAAGTCCCCTTGATTTACTAGTCAAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCTGAGTCA

SnaBI (4065)
4001 AACCGCTATCCACGCCATTGATGTAAGTCCAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCAAAGTAGAAAGTCCCAT

NdeI (4170)
4101 AGGTCATGTAAGTGGCATAATGCCAGGCGGCCATTTACCCTGATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTAAGTCCAAAGT
4201 GGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGGGGG

PacI (4356)
SdaI (4348) BspLU11I (4366)
4301 TCGTTGGGCGGTACGCCAGGCGGGCCATTTACCCTAAGTTATGTAACGCTGACAGGTTAATTAAGAACATGTGAGCAAAGGCCAGAAAAGGCCAGGAA
4401 CCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA
4501 GGACTATAAAGATACCAGGCTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTT

ApaLI (4680)
4601 CGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTCA
4701 GCCCGACCGCTGCGCTTATCCGGTAATATCGTCTTGTAGTCCAAACCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGC
4801 AGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGC
4901 CAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAG

PacI (5096)
5001 AAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAA

EagI (5116)
SwaI (5105) NotI (5115)
5101 TTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGGTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCA
5201 AAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA