



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203)
201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGTCTTTGTTTCGTTT

KasI (535) **BstEII (555)**
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCCCTAAGCTGAGATCACCGGTACCATGGCTGCAGCAGCAGCTTCTCACCTGAACCTGGATGC
1 M A A A A A S H L N L D A

SapI (646)
601 CCTCCGGGAAGTGCTAGAATGCCCATCTGCATGGAGTCTTACAGAAAGCAGCTGCGTCCCAAGCTTCTGCACTGTGGCCATACCCTGCGCCAG
13> L R E V L E C P I C M E S F T E E Q L R P K L L H C G H T I C R Q
701 TGCCTGGAGAAGCTATTGGCCAGTAGCATCAATGGTGTCCGCTGTCCCTTTGAGCAAGATTACCCGCATAACCAGCTTGACCCAGCTGACAGACAATC
47> C L E K L L A S S I N G V R C P F C S K I T R I T S L T Q L T D N
801 TGACAGTGCTAAAGATCATTGATACAGCTGGGCTCAGCGAGGCTGTGGGGCTGCTCATGTGTCGGTCTGTGGCGGGCTGTGCCCGGCAATCTGCCG
80> L T V L K I I D T A G L S E A V G L L M C R S C G R R L P R Q F C R

BsrGI (961) **BsrBI (991)**
901 GAGCTGTGGTTTTGGTGTATGTGAGCCCTGCCGGAGGAGCAGCATCAGCCTCTGGCCACTGTACACTCCCTGTCAAAGAAGCAGCTGAGGAGCGGCT
113> S C G L V L C E P C R E A D H Q P P G H C T L P V K E A A E E R R
1001 CGGACTTTGGAGAGAAGTAACTCGTCTGCGGAACTTATGGGGAGCTGCAGCGGCGGAAGGCAGCCTTGGAAAGTGTCTCCAAGGACCTTCAAGCAA
147> R D F G E K L T R L R E L M G E L Q R R K A A L E G V S K D L Q A
1101 GGTATAAAGCAGTTCTCCAGGATATGGGCATGAGGAGCGCAGGTTCCAGATGAGCTGCTCGCTCTCGGAAGTCTTACAGGCTCTTTGGCTGAAGT
180> R Y K A V L Q E Y G H E E R R V Q D E L A R S R K F F T G S L A E V
1201 TGAGAAGTCCAATAGTCAAGTGGTAGAGGAGCAGATTTACCTGCTTAACATTGCAGAGTGCAGGCTGTGCTCGCTGTGACTACTTCTGGCCAAGATC
213> E K S N S Q V V E E Q S Y L L N I A E V Q A V S R C D Y F L A K I
1301 AAGCAGGCAGATGTAGCACTACTGGAGGAGACAGCTGTGAGGAGGAGCCAGAGCTCACTGCCAGCTTGCCTCGGGAGCTCACCTGCAAGATGTGGAGC
247> K Q A D V A L L E E T A D E E E P E L T A S L P R E L T L Q D V E

SfiI (1482)
1401 TCCTTAAGGTAGGTCATGTTGGCCCCCTCAAATTGGACAAGCTGTTAAGAAGCCCCGGACAGTTAAGTGAAGATTCTGGCCATGGAGGCCACAGC
280> L L K V G H V G P L Q I G Q A V K K P R T V N V E D S W A M E A T A

AvrII (1567)
1501 GTCTGCTGCCTCTACCTCTGTTACTTTTAGAGAGATGGACATGAGCCCCGAGGAAGTGGTTGCCAGCCCTAGGGCTCACCTGTAAACAGCGGGTCTCT
313> S A A S T S V T F R E M D M S P E E V V A S P R A S P A K Q R G P
1601 GAGGCAGCTCCAATATCCAGCAGTGCCTCTTCTCAAGAAGATGGGGCCAAAGGCAGCACTCCAGGAATGTTCAATCTTCCAGTCACTCTACGTGA
347> E A A S N I Q Q C L F L K K M G A K G S T P G M F N L P V S L Y V

SpeI (1716) **ScaI (1713)** **Bst1107I (1742)**
1701 CCAGTCAAGGTGAAGTACTAGTCGCTGACCGTGGTAACTATCGTATACAAGTCTTTACCCGCAAAGGCTTTTTGAAGGAAATCCGCCGACGCCCCAGTGG
380> T S Q G E V L V A D R G N Y R I Q V F T R K G F L K E I R R S P S G

BglII (1833)
1801 CATTGATAGCTTTGTGCTAAGCTTCTTGGGGCAGATCTACCCAACTCACTCTCTCTCAGTGGCAATGAACTGCCAGGGGCTGATTGGTGTGACTGAC
413> I D S F V L S F L G A D L P N L T P L S V A M N C Q G L I G V T D

DraIII (1939)
1901 AGCTATGATAACTCCCTCAAGGTATATACCTGGATGGCCACTGCTGGCCTGTACAGGAGCCAGCTGAGCAAACCATGGGGTATCACAGCCTTCCAT
447> S Y D N S L K V Y T L D G H C V A C H R S Q L S K P W G I T A L P
2001 CTGGCCAGTTTGTAGTAACCGATGTGGAAGGTGAAAGCTTTGGTGTTCACAGTTGATCGAGGATCAGGGTGGTCAAATACAGCTGCCTATGTAGTGC
480> S G Q F V V T D V E G G K L W C F T V D R G S G V V K Y S C L C S A

StuI (2160)
2101 TGTGCGGCCCAAATTTGTCACCTGTGATGCTGAGGGCACCGTCTACTTCAACCCAGGGCTTAGGCCTCAATCTGGAGAATCGGCAGAATGAGCACCACTG
513> V R P K F V T C D A E G T V Y F T Q G L G L N L E N R Q N E H H L
2201 GAGGGTGGCTTTTCCATTGGCTCTGTAGGCCCTGATGGCAGCTGGGTGCCAGATTAGCCACTTCTTCTCGGAGAATGAGGATTTCCGCTGCATTGCTG
547> E G G F S I G S V G P D G Q L G R Q I S H F F S E N E D F R C I A

PshAI (2338)
2301 GCATGTGTGGATGCTCGTGGTATCTCATCGTGGCTGACAGTAGTCGAAGGAAATCTCCATTTTCTAAGGGTGGGGGCTATAGTGCCTTATTTCG
580> G M C V D A R G D L I V A D S S R K E I L H F P K G G G Y S V L I R

BglII (2483)
2401 AGAGGGACTTACCTGTCCGGTGGGCATAGCCCTAACTCCTAAGGGGACAGCTGCTGGTCTTGGACTGTTGGGATCATTGCATCAAGATCTACAGCTACCAT
613> E G L T C P V G I A L T P K G Q L L V L D C W D H C I K I Y S Y H

NheI (2529)
2501 CTGAGAAGATATTCCACCCCATAGGGGATCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAA
647> L R R Y S T P •

MfeI (2678)
2601 ATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCTATAAGCTGCAATAAACAAGTTAACAACAACCAATTGCATTTCATTATGTTT

2701 CAGGTTCCAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAAATCTAAAATACAGCATAGCAAACTTTAACCTCC
EcoRI (2763)

2801 AAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCT

2901 TTCATGGAGTTTAAAGATATAGTGTATTTTCCCAAGTTTGAAC TAGCTCTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTATG
SapI (2945)

3001 AAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAG
SspI (3002) SwaI (3016)

3101 TTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGACTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATG
141 • N R T Y K L P I

3201 AGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGG
131 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 G C P

3301 GGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCC
98 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 A S G

3401 AATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCCG
65 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 A A G
StuI (3441)

3501 ACATGGTCTTGTGTCTCATAGAGCATGGTGTCTTCTCAGTGGGACCTCCACCAGTCCAGATCTGCTGAGAGATGTTGAAGTCTTCATGGTGG
31 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 M
BbsI (3587)

3601 CCTCCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCA
AseI (3649)

3701 CTAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTG

3801 ATTTACTAGTCAAAACAACTCCCAATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAA
SpeI (3804)

3901 AACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAGGTCATGTACTGGGCATAATGCCAGGCGGGCC
SnaBI (3932)

4001 ATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGCAGTTTACCGTAAATACTCCACCCATTGACG
NdeI (4037)

4101 TCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGGGGGGTCTTGGGCGGTGAGCCAGGCGGGCCATTACC

4201 GTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGTTGCTGGCGTTTTTCCA
SdaI (4215) PacI (4223) BspLU11I (4233)

4301 TAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGA

4401 AGCTCCCTCGTGCCTCTCTGTTCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCT

4501 GTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTATCG
ApaLI (4547)

4601 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTC

4701 TTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTT

4801 GATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTT

4901 TTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATAT
PacI (4963) SwaI (4972) EagI (4983) NotI (4982)

5001 CTTTATTTTTCATTACATCTGTGTGTTGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATA

5101 GGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTATCGAA