



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA **EcoNI (96)**
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)
Psp1406I (203) 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCTTACCGCGCCCGCCCTACCTGAGGCC **EcoNI (287)**
PvuII (239) 301 GCCATCCACGCGGTTGAGTGCAGTTCGCCCTCCCGCCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
NaeI (441) 401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTTGTTCGTTT

KasI (535) 501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCGCCTACCTGAGATCACCGGTCATCATGATTGACAGACAGAGAATGGGACTTTGGGCTTTGGC **AgeI (552)**
601 AATTCTGACACTTCCCATGATTTGACAGTTACGAGGGCAGTAAATCGTCTGGGGTCTGGAAAATGAGGCTTTAATTGTGAGATGCCCCAAAGAGGA **BspHI (560)**
13> I L T L P M Y L T V T E G S K S S W G L E N E A L I V R C P Q R G
701 CGCTCGACTTATCCTGTGGAATGGTATTACTCAGATACAAATGAAAGTATTCCTACTCAAAAAAGAAATCGGATCTTTGTCTCAAGAGATCGTCTGAAGT
47> R S T Y P V E W Y Y S D T N E S I P T Q K R N R I F V S R D R L K A
801 TTCTACCAGCCAGAGTGGAGACTCTGGGATTTATGCTGTGTTATCAGAAGCCCAACTTGAATAAGACTGGATACTTGAATGTCACCATACATAAAA
80> F L P A R V E D S G I Y A C V I R S P N L N K T G Y L N V T I H K K
901 GCCGCAAGCTGCAATATCCCTGATTATTTGATGACTCGACAGTACGTGGATCAGATAAAAAATTTCAAGATAACGTGTCCAACAATTGACCTGTATAAT
113> P P S C N I P D Y L M Y S T V R G S D K N F K I T C P T I D L Y N

XmnI (742) 1001 TGGACAGCACCTGTTTCAGTGGTTAAGAAGCTCCTCAAGAGCCAAGGTTTCAGGGCACACAGGTCCTACTTGTTCATTGACAACGTGACTCATG **BspHI (1095)**
147> W T A P V Q W F K N C K A L Q E P R F R A H R S Y L F I D N V T H
1101 ATGATGAAGGTGACTACACTTGTCAATTCACACAGTGGAGAATGGAACCAACTACATCGTGACC GCCACCAGATCATTACAGTTGAAGAAAAAGGCTT
180> D D E G D Y T C Q F T H V E N G T N Y I V T A T R S F T V E E K G F
1201 TTCTATGTTTCCAGTAATTACAAATCCTCCATACAACACACAATGGAAGTGGAAATAGGAAAACAGCAAGTATTGCCTGTTTCAGCTTGTCTTGGCAA
213> S M F P V I T N P P Y N H T M E V E I G K P A S I A C S A C F G K

EcoRI (1370) 1301 GGCTCTCACTTCTGGCTGATGTCCTGTGGCAGATTAACAAAACAGTAGTTGAAATTTTGGTGAAGCAAGAATTCAAGAAGAGGAAGTTCGAAATGAAA
247> G S H F L A D V L W Q I N K T V V G N F G E A R I Q E E E G R N E

XcmI (1403) 1401 GTTCCAGCAATGACATGGATTGTTAACCTCAGTGTTAAGGATAACTGGTGTGACAGAAAAGGACCTGTCCCTGGAATATGACTGTCTGGCCCTGAACCT
280> S S S N D M D C L T S V L R I T G V T E K D L S L E Y D C L A L N L
1501 TCATGGCATGATAAGGCACACATAAGACTGAGAAGGAAACAACCAATTGATCACCGAAGCATCTACTACATAGTTGCTGGATGATGTTTATTGCTAATG
313> H G M I R H T I R L R R K Q P I D H R S I Y Y I V A G C S L L M
1601 TTTATCAATGCTTGGTATAGTCTTAAAAGTGTCTGGATTGAGGTTGCTCTGTTCTGGAGAGATATAGTGACACCTTACAAAACCCGGAACGATGGCA
347> F I N V L V I V L K V F W I E V A L F W R D I V T P Y K T R N D G

ScaI (1769) 1701 AGCTCTACGATGCGTACATCATTACCCTCGGGTCTTCCGGGGCAGCGCGGGGAACCCACTCTGTGGAGTACTTTGTTCCACACACTCTGCCCGACGT
380> K L Y D A Y I I Y P R V F R G S A A G T H S V E Y F V H H T L P D V

DraIII (1759) 1801 TCTTAAAAATAAATGTGGCTACAAATTGTGCATTTATGGGAGAGACCTGTTACCTGGGCAAGATGCAGCCACCGTGGTGGAAAGCAGTATCCAGAATAGC
413> L E N K C G Y K L C I Y G R D L L P G Q D A A T V V E S S I Q N S
1901 AGAAGACAGGTGTTTGTCTGGCCCTCACATGATGCACAGCAAGGAATTTGCTACGAGCAGGAGATTGCTCTGCACAGCGCCCTCATCCAGAACAAC
447> R R Q V F V L A P H M M H S K E F A Y E Q E I A L H S A L I Q N N
2001 CCAAGGTGATTCTTATTGAAATGGAGCCTCTGGGTGAGGCAAGCCGACTACAGGTTGGGGACCTGCAAGATTCTCCAGCATCTTGTGAAAATTCAGGG
480> S K V I L I E M E P L G E A S R L Q V G D L Q D S L Q H L V K I Q G

Asp718I (2171)
Acc65I (2171) 2101 GACCATCAAGTGGAGGGAAGATCATGTGGCCGACAAGCAGTCTCTAAGTTCAAATCTGGAAGCATGTGAGGTACCAATGCCAGTGCAGAAAGAGCC
513> T I K W R E D H V A D K Q S L S S K F W K H V R Y Q M P V P E R A

MscI (2273) 2201 TCCAAGACGGCATCTGTTGCGGCTCCGTTGAGTGGCAAGGCATGCTTAGACCTGAAACACTTTTGAAGCTAGCTGGCCAGACATGATAAGATACATTGAT
547> S K T A S V A A P L S G K A C L D L K H F •

SphI (2239) 2301 GAGTTTGGACAAACCACAACCTAGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATA

HpaI (2405) 2401 AACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTTGGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAT

EcoRI (2501) 2501 GGAATTCTAAAATACAGCATAGCAAACCTTAACTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTG

2601 TTGCCAATGTCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGTTTGAAGTCTTTCATTCTTTAT

2701 GTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTTTATTAGGCA

2801 GAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGA

2901 GCTTCTAGCTTTAGTTCTGGTGTACTTGGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGG
 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 V F C D P

3001 AGCATAGTCAGAGATGAGCTCTCGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATG
 111 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 V A V I

3101 GTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCCAGCACAGACAGTACCCTGCCAATGTAGGCCCTCAATGTGGACAGCAG
 77 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 I H V A S

3201 AGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTC
 44 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 V E V L E

3301 CAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATGTCAA
 11 L D Q Q S I N F T K M

3401 ACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGTTATATAGACCTCCACCGTACACGCCTACCGCCATTGCGTCAAT

3501 GGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCATTGACGTCAATGGGTGGAGACTTGAAATCCCCGT

3600 GAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGT

3700 CCCATAAGGTCATGTACTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGC

3800 CAAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATTATTGACGTCAATGGG

3900 CGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTTGCAGGTTAA TTAAGAACATGTGAGCAAAAGCCAGCAAAAG
 SdaI (3953) PaeI (3961) BspLU11I (3971)

3998 GCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAA

4098 ACCCGACAGGACTATAAAGATACAGGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCCTT

4198 TCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCC

4298 CCGTTCAGCCCGACCGTGCCTTATCCGGTAACTATCGTCTTGTAGTCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA

4398 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCT

4498 GCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATT

4598 ACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCAATGG

4698 CTAGTTAATTAACATTTAAATC AGCGGCCCAATAAAATATCTTTATTTTATTACATCTGTGTGGTTTTTTTGTGTGAATCGTAACTAACATACGCT

4798 CTCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCGAGTCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA