



PvuI (7)
SgfI (6) 1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGTGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTTCGAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTCACGGCCCGCCCTACCTGAGGCC
HindIII (245)
301 GCCATCCACGCCGGTTGAGTCGGTCTGCGCCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

AgeI (552) **BspHI (560)** **NcoI (581)**
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCATCATGAGCGACCAAGATCACTCCATGGATGAAGTGACAGC
1 M S D Q D H S M D E V T A
601 TGTGGTGAAGATTGAAAAAGATGTTGGTGGCAATAATGGGGGTAGCGGCAATGGTGGCGGTGCCGCTTTTCTCAGACTCGAAGCAGCAGCAGGCAGT
13 V V K I E K D V G G N N G G S G N G G G A A F S Q T R S S S T G S
701 AGCAGCAGCAGTGGTGGCGGAGGAGGCAGGAATCCCAGCCATCTCCTTTGGCTCTGCTGGCAGCAACTGCAGCAGAATTGAGTCACCCAATGAGAACA
47 S S S S G G G G Q E S Q P S P L A L L A A T C S R I E S P N E N
DraIII (869)
801 GCAACAACCTCCAGGGTCCGAGTCAGTCAGGGGACAGGTGAACCTGACCTCACAGCCGACAACTTTCACAGGGTGCCAATGGCTGGCAGATCATCTC
80 S N N S Q G P S Q S G G T G E L D L T A A Q L S Q G A N G W Q I I S
901 TTCCTCCTGGGGCTACCCCTACCTCAAAGGAACAGAGTGGCAACAGTACCAATGGCAGCGAGTCTTCAAGAACCAGCAGTCTCTGGTGGCAGTAT
113 S S S G A T P T S K E Q S G N S T N G S E S S K N R T V S G G Q Y
SspI (1066)
1001 GTTGTGGCTGCTACCCCAACTTACAGAACCAGCAAGTCTGACAGGTCTCCCTGGAGTAATGCCTAATATTAGTATCAAGTAATCCACAGTTCAGAA
147 V V A A T P N L Q N Q Q V L T G L P G V M P N I Q Y Q V I P Q F Q

BstAPI (1110) 1101 CTGTTGATGGGCAGCAGCTGCATTTGCTGCCATGGGCGCCAAAGTCAGCAGGATGGTTCTGGTCAAATACAGATCATAACAGGTGCAAACCAACAGAT
Bsp120I (1136)
180 T V D G Q Q L Q F A A T G A Q V Q Q D G S G Q I Q I I P G A N Q Q I
1201 CATCCAAATAGAGGAAGTGGGGCAACATTATTGCTGCTATGCCAAATCTACTCCAGCAGGCTGTCCCCCTTCAAGGCTTGCTAATAATGTGCTCTCA
213 I P N R G S G G N I I A A M P N L L Q Q A V P L Q G L A N N V L S
1301 GGACAGACTCAGTATGTACCAATGTACCAGTGGCCCTGAATGGGAACATCACCTGTACCTGTCAACAGGTTTCTGCAGTACCCTGACTCCCAGCT
247 G Q T Q Y V T N V P V A L N G N I T L L P V N S V S A A T L T P S

BamHI (1426) 1401 CTCAGGCAGGCACTATCAGCAGCTCTGGATCCCAGGAGAGCAGCTCACAGCCTGTACCTCAGGGACTGCCATCAGTCTGCCAGCTTGGTGTATCACA
280 S Q A G T I S S S G S Q E S S S Q P V T S G T A I S S A S L V S S Q
1501 AGCTAGTTCAGCTCCTTTTTTACCAATGCCAATAGTATTCAACAATACTACCACCAGCAACATGGGAATTATGAACCTTACCAGCAGTGGCTCATCA
313 A S S S S F F T N A N S Y S T T T T S N M G I M N F T S S G S S
SpeI (1603)
1601 GGGACTAGTTCTCAAGGCCAGACGCCCCAGAGGGTTGGTGGGCTACAAGGTCTGATTCTCTGAACATCCAGCAGAACCAGACATCAGGAGGCTCGCTGC
347 G T S S Q G Q T P Q R V G G L Q G S D S L N I Q Q N Q T S G G S L
1701 AAGGAAGTCAGCAGAAAGAGGAGCAAAGTCAGCAGACACAGCAACAACAACTTATTCAGCCTCAGTAGTTCAAGGAGGACAAGCTCTTCAGGC
380 Q G S Q Q K E G E Q S Q Q T Q Q Q I L I Q P Q L V Q G G Q A L Q A

BspEI (1819) 1801 CCTTCAAGCAGCACCATTGTCCGACAGACCTTCAACAATACTCAAGCTATTTCCAGGAAACCTTCAGAACCTCAGCTTCCAGGCTGTTCAAACTCTGGT
413 L Q A A P L S G Q T F T T Q A I S Q E T L Q N L Q L Q A V Q N S G
XmnI (1855)
1901 CCCATCATCATTGACACCAACCGTGGGCGCCAAATGGACAGGTGAGTGGCAGACCTTCCAGCTGCAGAATCTTCAAGTTCAAGTCCAGAACCCACAAGCCCAGA
447 P I I I R T P T V G P N G Q V S W Q T L Q L Q N L Q V Q N P Q A Q
2001 CAATCACCTTGGCCCTATGCAGGGTGTTCCTTGGGCGAGCTAGCAGCAGTAATACCACCTAACCCATTGCTCAGCTGCCTCCATTCTGCTGG
480 T I T L A P M Q G V S L G Q T S S S N T T L T P I A S A A S I P A G
ApaLI (2188)
2101 CACAGTCACTGTGAATGCTGCTCAACTCTCCTCCATGCCAGGCTCCAGACCATTAACCTCAGTGCATTGGGTACTTCAGGATCCAGGTGCACCAGCTT
513 T V T V N A A Q L S S M P G L Q T I N L S A L G T S G I Q V H Q L
2201 CCAGGCCCTGCTTTGGCTATAGCAAACACCCAGGTGATCATGGAACCTCAACTTGGTCTTCATGGATCTGGTGGTGGTGGATACATGAGACAGCAG
547 P G L P L A I A N T P G D H G T Q L G L H G S G G D G I H D E T A
BamHI (2180)
2301 GTGGAGAAGGAGAGAATAGTCTGATCTCCAACCCCAAGCTGGACGCGAGGACTCGTCGGGAAGCATGTACATGCCCTATTGCAAAGACAGTGAGGGAAG
580 G G E G E N S S D L Q P Q A G R R T R R E A C T C P Y C K D S E G R
2401 AGCCTCAGGAGATCCTGGCAAAAAGAACAGCACATTTGTACATCCAAGGATGCGGCAAAGTATATGGCAAGACCTCACATCTCCGAGCACACTTGGCG
613 A S G D P G K K K Q H I C H I Q G C G K V Y G K T S H L R A H L R
BsrGI (2365)
2501 TGGCATAAGGGGAGAGGCCATTGTAATTGGTCAATTTGTTGGGAAAGCGCTTTACAGTTCGGACGAGCTTCAGAGACATAAACGTACACATACAG
647 W H T G E R P F M C N W S Y C G K R F T R S D E L Q R H K R T H T
Eco47III (2549)

2601 GAGAGAAGAAATTTGCCCTGCCCTGAGTGCCCTAAGCGTTTCATGAGGAGTGATCACCTGTCAAAGCATATCAAGACTCACCAGAACAAGAAGGGAGGCC
680▶ G E K K F A C P E C P K R F M R S D H L S K H I K T H Q N K K G G P

2701 AGTGTTAGCCCTGAGTGTGGCACATTGCCCTGGACAGTGGGGCAGGTTTCAAGGCACTGCCACTCCCTCAGCCCTTATTACCACCAATATGGTAGCC
713▶ G V A L S V G T L P L D S G A G S E G T A T P S A L I T T N M V A

2801 ATGGAGGCCATCTGTCCAGAGGGTATTGCCCGTCTTGCCAACAGTGGCATCAACGTCATGCAGGTGACAGAGCTGCAGTCCATTAATATCAGTGGCAATG
747▶ M E A I C P E G I A R L A N S G I N V M Q V T E L Q S I N I S G N

2901 GTTCTAAGATTAGACACCCAGTGCCAGAGACAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGA
780▶ G F •

3001 AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACAAATTGCATTCATTTAT

3101 GTTTCAGGTTTCAGGGGAGGTGTGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTTAAAATACAGCATAGCAAACTTTAAC

3201 CTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTCAGCCTCAC

3301 TTCTTTCATGGAGTTTAAGATATAGTGATTTTCCCAAGTTTGAAGTACTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTT

3401 TAGTAAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCC

3501 CCAGTTTAGTAGTTGGACTTAGGGAACAAGGAACCTTAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGG
141▶ • N R T Y K L P

3601 GATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGACATGCCA
133▶ I 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

3701 CAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAG
99▶ C P 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

3801 ACCCAATGGCAATGGCTTCCAGCAGACAGTACCCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTGGTCTGATGGCCGC
66▶ G 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

3901 CCCGACATGGTGCTTGTGCTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATG
33▶ G 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

4001 GTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGG

4101 TTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCC

4201 GTTGATTTACTAGTCAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGTATCCACGCCATTGATGTACTG

4301 CCAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGACTGGGCATAATGCCAGGGC

4401 GGCCATTTACCGTATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTACCGTAAATACTCCACCCATT

4501 GACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGCTGTTGGCGGTGAGCCAGGCGGGCCATT

4601 TACCGTAAGTTATGTAACGCTGCAGGTTAATTAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCGCTTGTGGCGTTTTT

4701 TCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCC

4801 TGGAACTCCCTCGTGCCTCTCCTGTTCCGACCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTGGGAAGCGTGGCGCTTTCTCATAGCTCA

4901 CGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGCCGACCGCTGCGCTTATCCGGTAACT

5001 ATCGTCTTGAGTCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGA

5101 GTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGC

5201 TCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCCGAGAAAAAAGGATCTCAAGAAGATCCTTTGA

EagI (5387)

PacI (5367) SwaI (5376) **NotI (5386)**

5301 TCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAA

5401 ATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAA

5501 AATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA