



PvuI (7)
SgfI (6) 1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA **EcoNI (96)**

101 GAGAAGGTGGCGGGGTAAGTGGGAAAGTGTGTCGTGTAAGTGGTCCGCCTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCGCTACCTGAGGCC
HindIII (245)
PvuII (239)
Bsu36I (291)
EcoNI (287)

301 GCCATCCACGCGGTTGAGTCGCGTTCTGCCGCTCCGCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGGCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTTGTTCGTTT

NgoMIV (441)

501 TCTGTTCTGCGCGGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCATCATGAGCAGGAGTTTGCCTTGTATATCTACTCCTTTTG
KasI (535) **AgeI (552)**
1 M S R S L P C H I Y S F W

601 GGTACTGTTGCCCTTTGGATACTGTACACAACCTTCTACCAACAAATGTACTGTTAGACATGAAATAGCTGACTGCAGCCATCTGAAGTTGACTCAAATA
BsrGI (623) **PstI (672)**
13 V L L P F W I L Y T T S T N K C T V R H E I A D C S H L K L T Q I
701 CCCGATGACCTCCCGGCAAAATATAACAGTGTGAACTTACCCACAATCAACTCAGAAGATTACCACCTGCCAATTTACAATATACAGCCAACCTACTA
47 P D D L P A N I T V L N L T H N Q L R R L P P A N F T I Y S Q L T

801 CCTTGATGGAGGATTTAACACCATCTCAAAGCTGGAGCCAGAATTGTCCAAAGTCTCCCTTTGTTGGAAATTTGAACCTCCAACACAATGAGCTCTC
XcmI (821)
80 T L D G G F N T I S K L E P E L C Q S L P L L E I L N L Q H N E L S
901 TCAGCTTTCTGATAAAACCTTTATCTTTCGCATGAATTTGATTGAAGTCCATCTAATGTCCAACCTCAATCCAGAAGATTCAAAAATAATCCCTTTAAAAAC
113 Q L S D K T F I F C M N L I E L H L M S N S I Q K I Q N N P F K N

BglII (1020)
1001 CTGAAGAATTTAATCAAATTAGATCTATCTCATAATGGTTTATCATCTACTAAATTAGGAACGCAGCTCCAACCTGAAAATCTCCAAGAGCTTCTATTAG
147 L K N L I K L D L S H N G L S S T K L G T Q L Q L E N L Q E L L L
1101 CAAATAATAAAATTTCTGCACTAAAACGTGAAGAACTTGATTTCCCTTGGCAATTTCTTTTAAAAAGATTAGAGTTGTCATCAAATCAAATACAAGGTT
180 A N N K I S A L K R E E L D F L G N S S L K R L E L S S N Q I Q E F
1201 CTCCTCAGGGTGTTCATGCCATTGGAAAATTTGGCCTCTCTGAAACAATGTCAAACCTGAGCCCACTGACAGAGAAAATTTGCTTAGAATTA
213 S P G C F H A I G K L F G L S L N N V K L S P S L T E K L C L E L

BsrGI (1343)
PvuII (1339) 1301 TCCAACACAAGCATTGAGAATCTATCCCTGAGCAACATCCAGCTGTACAAAACAGCAACACGACTTTTCTTGGACTAAAGCAGACAAAACCTCAGCATGC
SphI (1394)
247 S N T S I E N L S L S N I Q L Y K T S N T T F F G L K Q T N L S M
1401 TTGATCTTTCCATAACAGCTTAAAGTGTGATTGGTAATGATTCCTTTGCTTGGCTTCCACATCTGAAATATTTCTTCTGGAGTATAACAATATAGAGCG
280 L D L S H N S L S V I G N D S F A W L P H L K Y F F L E Y N N I E R
1501 TTTGTCTTCTCGCTCTTTATGGGCTTTCCAATGTGAAATACCTGAATCTGAGACGATCTTTTACTAAACAAAGCATTTTCTTGGCTTCACTTCCAAA
313 L S S R S L Y G L L S N V K Y L N L R R S F T K Q S I S L A S L P K
1601 ATAGAGGACTTTTCTTCAATGGCTAAAATCTTTGGAGTATCTCAACATGGAAGATAACAACCTTCCAGGCATAAAAAGGAATACTTTCACAGGATTGA
347 I E D F S F Q W L K S L E Y L N M E D N N F P G I K R N T F T G L
1701 TAAAGCTGAAAAGTTAAGTCTATCCAACCTCTTCTCAAGTTTGGCACTTTAACAATGAAACATTTATACTCTCGCTGATTCTCCTCTTCTCATACT
380 I K L K S L S L S N S F S S L R T L T N E T F I S L A D S P L L I L
1801 CAACCTAACCAAAAATAAAATCTCAAAAATAGAGAGTGGTCTTTTCTTGGTGGCCACCTAAAGGTAAGTACTTGCCTCGCCTTAATGAAATTTGGGCAA
413 N L T K N K I S K I E S G A F S W L G H L K V L D L G L N E I G Q
1901 GAACCTACAGGACAGGAATGGAGAGGTCAAAAATATTGTTGAAGTCTACCTTCTTACAACAGATACTAGAGTTGACTACCGACTCTTTCGCTTGG
447 E L T G Q E W R G L K N I V E V Y L S Y N R Y L E L T T D S F A L
2001 TTCCAAGCTTCAACAACCTGATGCTCCGAAGGGTGGCCTTAGAAACATGGATTGCTCCCTTCCACCTTTTACCCTTTTAACTTTGACCATTTTGG
480 V P S L Q Q L M L R V A L R N M D C S P S P F H P L F N L T I L D
2101 TCTAAGTAACAACAATATAGCCAACATAAATGATGAAGTGTGAAAGGCTTGGAGAACTACAATTTGATTTGCAGCATAACAATTTAGCACGGCTC
513 L S N N N I A N I N D E L L K G L E K L Q I L D L Q H N N L A R L

EcoO109I (2240)
2201 TGGAAACACGAAAACCTGGCGGTCTGTTCAGTTTCTGAAGGCCTTCTCACCTCCACATCTTAACCTTAGAGTCTAATGGCTTTGATGAAATCCCAG
547 W K H A N P G G P V Q F L K G L S H L H I L N L E S N G F D E I P
2301 CAGATGCCTTCAGGGACTTATCTGAATTAAGAGCATTGATTTAGGATTGAATAATTTAAACATCTTCCACCATCTGTCTTTGATAATCAAGTGTCTCT
580 A D A F R D L S E L K S I D L G L N N L N I L P P S V F D N Q V S L
2401 GAAGTCATTAAGCCTTCAAGAAGTCTCATAACGTCAGTCAAGAAGACTGTTTTTGGACCTGCATTCCAGAAGTTGAGTAATTTAGATATGCGCTTCAAT
613 K S L S L Q K N L I T S V K K T V F G P A F Q K L S N L D M R F N

ScaI (2551)
AseI (2545)
2501 CCGTTTACTGTACTGTGAAAGCATTGCCTGGTTTGTAGTTGGATTAATAGTACTCACACCAACATCTCTGAATTATCCAGCCATTACCTCTGCAATA
647 P F D C T C E S I A W F V S W I N S T H T N I S E L S S H Y L C N

2601 CTCCACCTCAATATCATGGTCTCCAGTGATTCTTTTTGATACATCACCTGCAAAGACAGTGGCCCATTTGAACTCTTTTTCATGATAACTGCCAGTAT
680 T P P Q Y H G L P V I L F D T S P C K D S A P F E L F F M I T A S M

BsaBI (2706) 2701 GCTACTGATTTTTATTACTGCTCATCCATTTCAAGGCTGGCGGATATCTTTTTATTGGAATGTTTCAGTGCATCGAGTTCTTGGTTTCAA
BstBI (2737) **EcoRV (2751)**
713 L L I F I F I I L L I H F E G W R I S F Y W N V S V H R V L G F K
2801 GAAATAGACAAGCAGCCAGAGCAGTTTGAATATGCAGCATATAAATTCATGCCTATAAAGATAGGGATTGGGCTGGGAACACTTTGCCCAATGGAAG
747 E I D K Q P E Q F E Y A A Y I I H A Y K D R D W V W E H F A P M E

2901 AAAAAGATGAAACTCTCAGATTTTGTCTGGAAGAAAGGACTTTGAGGCAGGTGCCCTTGAACCTGAAGCAATTGTTAATAGCATCAAAAAGGAGCAGAAA
780▶ E K D E T L R F C L E E R D F E A G A L E L E A I V N S I K R S R K
ApaLI (3058)

3001 AATTATTTTTGTCTATAACACAACATCTACTGAAAGATCCATTGTGCAAAAAGATTCAAGGTGCACCATGCAGTTCAGCAAGCTATTGAACAAAATCTGGAT
813▶ I I F V I T Q H L L K D P L C K R F K V H H A V Q Q A I E Q N L D
3101 TCCATTATACTGATCTTTCTTGAGGAGATTCCAGATTATAAAGTGAACCATGCCTCTGTTTGCAGAGGAAATGTTTAAATCTACTGCATCTTGAAC
847▶ S I I L I F L E E I P D Y K L N H A L C L R R G M F K S H C I L N
MscI (3287)

3201 GGCCCGTGCAGAAAGAACGGATAAATGCCTTTCATCACAACTGCAAGTAGCATTGTTTCCAGAAATTCAGTACATTAAGGCTAGTGGCCAGACATGA
880▶ W P V Q K E R I N A F H H K L Q V A L G S R N S V H •
NheI (3281)

3301 TAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCAT

HpaI (3419)

3401 TATAAGCTGCAATAAACAAGTTAAACAACAATTGCATTCTTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTC

EcoRI (3515)

3501 TACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATA
3601 GGCATCAGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTTCATGGAGTTTAAAGATATAGTGATTTTCCCAAGGTTTGAACCTAGCT

Swal (3768)

3701 CTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATG

EcoO109I (3829)

3801 TTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGA
3901 CAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGGGGGATGAGTTTCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGC
1414 • N R T Y K L P I L E E I T T K V L K G N M E I L
BstXI (4058)

4001 ACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCC
1154 V F C D P 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
StuI (4193)

4101 TGACAGCCACAATGGTGTCAAAGCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTC
824 V A V I 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
4201 AATGTGGACAGCAGAGATGATCTCCCGAGTCTGGTCTGATGGCCGCCCGACATGGTGCTTGTTCCTCATAGAGCATGGTGATCTTCTCAGTGGCG
494 I H V A S 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

4301 ACCTCCACAGCTCCAGATCCTGTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGAT
154 V E V L E L D Q Q S I N F T K M

AseI (4401)

4401 GATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGC

SpeI (4556)

4501 CCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGGATTTACTAGTCAAAAACAACTCCATTGACGTCAATGGGGTGGAGAC

SnaBI (4684)

4600 TTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGTAGCTAATAACGTAGATGACTG

NdeI (4789)

4700 CCAAGTAGGAAAGTCCATAAGGTCATGTACTGGCATAATGCCAGGCGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATAC
4800 ACTTGATGACTGCCAAGTGGGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTA

PstI (4968)

4900 TTGACGTCAATGGGCGGGGCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCCTG C A G G T T A A T T A A G A A C A T G T G A C A A
SdaI (4967) PacI (4975) BspLU11I (4985)

4998 AAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAG
5098 TCAGAGGTGGCGAAACCCGACAGGACTATAAGATAACAGGCGTTTCCCCTGGAAAGTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGCTTACCGGA
5198 TACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCT

ApaLI (5299)

5298 GTGTGCACGAACCCCGTTTCCAGCCGACCCTGCGCCTTATCCGGTAACTATCGTCTTGGTCCAAACCCGTAAGACACGACTTATCGCCACTGGCAGC
5398 AGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTT
5498 GGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAACAAACCCGCTGGTAGCGGTGGTTTTTTTGT
5598 GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGG

EagI (5735)

5698 GATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCGCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGT
PacI (5715) Swal (5724) NotI (5734)

5798 AACTAACATACGCTCTCCATCAAAAACAAACGAAACAAACAACTAGCAAAATAGGCTGTCCCGAGTCAAGTGCAGGTGCCAGAACATTTCTCTATCG
5898 AA