



PvuI (7)
SgfI (6) 1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGTGCTGTACTGGCTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTCACGCGCCCGCCCTACCTGAGGCC
HindIII (245)
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BstEII (555)
AgeI (552) **NcoI (560)**
501 TCTGTTTCTGCCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGTCACCATGGATCTGACAAAGATGGGGATGATCCAGTGCAGAA
1 M D L T K M G M I Q L Q N
601 CCCTAGCCACCCACGGGCTGTGTGCAAGGCCAATCAGATGCGACTGGCTGGGACTTTGTGCGATGGTGCATCATGGTGGACAGCCAGGAGTCCAT
13 P S H P T G L L C K A N Q M R L A G T L C D V V I M V D S Q E F H

NheI (714)
701 GCCCACCAGCGGTGCTAGCCTGCACCAGCAAGATGTTTGTGATCCTCTCCACCGAAACAGCCAGCACTATACTCTGGACTTCCTCTGCCAAAAACCT
47 A H R T V L A C T S K M F E I L F H R N S Q H Y T L D F L S P K T
801 TCCAGCAGATCCTGGAGTACGCCTACACGGCCACACTGCAAGCCAAGCGGAGGACCTGGATGACCTGCTGTATGCAGCTGAGATTTAGAGATCGAATA
80 F Q Q I L E Y A Y T A T L Q A K A E D L D D L L Y A A E I L E I E Y
NcoI (968)
901 CCTGGAGGAGCAGTGCCTGAAGATCCTGGAGACCATCCAGGCATCTGATGACAATGACACAGAGGCCACCATGGCTGACGGTGGGGGCAAGAAGAAGAG
113 L E E Q C L K I L E T I Q A S D D N D T E A T M A D G G G E E E E

NcoI (1097)
Bsp120I (1093)
1001 GACCGTAAGGCTCGATACCTCAAGAACATCTTTATCTCGAAGCATTCCAGCGAGGAGAGTGGCTACGCCAGTGTGGCTGGACAGAGCCTCCCTGGGCCA
147 D R K A R Y L K N I F I S K H S S E E S G Y A S V A G Q S L P G P
1101 TGGTGGACAGAGCCCTCAGTCTCCACCTCTTCGGTCTCTCAGCCATGAGTCTACCAAGGCGAGCTGGGACAGCTTGTGAGTATAGGACAGTCACT
180 M V D Q S P S V S T S F G L S A M S P T K A A V D S L M S I G Q S L

SdaI (1222)
1201 CCTGCAAGGAACCTTTCAGCCACCTGCAGGGCCTGAGGAGCCACACTGGCAGGGGGTGGGCGACACCCTGGGGTGGCTGAGGTGAAGATGGAATGATG
213 L Q G T L Q P P A G P E E P T L A G G G R H P G V A E V K M E M M

HindIII (1372)
1301 CAGGTAGATGAAGCCCTGGCCAGGACAGCCCTGGGGCAGCTGAGTCTAGCATCTCAGGAGGGATGGGGGACAAGCTTGAAGAGAGGAGCAAAGAGGGGC
247 Q V D E A P G Q D S P G A A E S S I S G G M G D K L E E R S K E G

XhoI (1412) **SpeI (1429)**
1401 CTGGGACCCGACTCGAGGCAGCGTCATCACTAGTGCCAGAGAGCTGATTATGGGAGAGAGGAGAGTGGTGAAGCAGCTCTCGCCTCCTGTTGAAGCTGG
280 P G T P T R G S V I T S A R E L H Y G R E E S G E Q L S P P V E A G
1501 CCAGGGACCCCTGGGCGGAGGAGCCCTGGCACCCAGTGGAGAAGCATTGGGTATCTACTCGGTGCTGCCAACCAAGCCGATGCTGTGTTG
313 Q G P P G R Q E P L A P P V E K H L G I Y S V L P N H K A D A V L

SphI (1601) **NcoI (1655)**
1601 AGCATGCCGTCTTCAGTGACGTCIGGTCTCCATGTGCAACCTGCCCTGGCAGTCTCCATGGACTTCAGCACCTACGGGGTCTGCTGCCTCAGGGCTTCA
347 S M P S S V T S G L H V Q P A L A V S M D F S T Y G G L L P Q G F
1701 TCCAGAGGGAGCTGTTAGCAAGCTGGGGGAGCTGGCTGTGGGCATGAAGGCTGAGAGCCGCCTCTGGGGGAGCAGTGCAGIGTGTGGGGTGCAGCT
380 I Q R E L F S K L G E L A V G M K A E S R P L G E Q C S V C G V E L

BspEI (1800)
1801 TCCGGACAACGAGGAGTGGAGCAGCACAGGAACTGCACAGTGGGATGAAAACATACGGGTGTGAACCTGCGGAAAACGGTTCCTGGACAGTTTGCGA
413 P D N E A V E Q H R K L H S G M K T Y G C E L C G K R F L D S L R

NsiI (1906)
1901 CTGAGAATGCATTTACTGGCTCATTGAGCGGGTGCCAAAGCCTTTGTGTGTGATCAGTIGGTGCCAGTTCTCAAAGGAGGATGCCCTGGAGACACACA
447 L R M H L L A H S A G A K A F V C D Q C G A Q F S K E D A L E T H
2001 GACAGACCCACTACTGGCAGGACATGGCTGTCTTCTGTCTGTGTGGGAAACGCTTTCAGGCACAAAGCGCACTCCAGCAGCAGTGGAGGTCACGC
480 R Q T H T G T D M A V F C L L C G K R F Q A Q S A L Q Q H M E V H A

FspI (2105) **BstEII (2185)** **NdeI (2195)**
2101 AGGCGTGCGCAGTATATTTGCAAGTGAACCGCACCTTCCCCAGCCACAGGGCTCTCAAGCGCCACCTTCGCTCACATACAGGTGACCCCATAT
513 G V R S Y I C S E C N R T F P S H T A L K R H L R S H T G D H P Y
2201 GAGTGTGAGTTCTGCGGCAGTCTCCGGGATGAGAGCACACTCAAGAGCCACAAGCGCATCCACACAGGGGAGAAGCCCTATGAGTGTAAACGGCTGTG
547 E C E F C G S C F R D E S T L K S H K R I H T G E K P Y E C N G C

Eco47III (2387)
2301 GCAAGAAGTTCAGCCTCAAGCACCAGTTGGAGACGCACTACAGGGTTCACACAGGTGAGAAGCCCTTTGAGTGCAAACTCTGCCACCAGCGCTCCCGAGA
580 G K K F S L K H Q L E T H Y R V H T G E K P F E C K L C H Q R S R D
2401 CTACTCGCCATGATCAAGCACCTGAGAACACACAACGGTGCCTGCCATACAGTGTACCATCTGCACGGAATACTGCCAGCCTCTCCTCCATGCAG
613 Y S A M I K H L R T H N G A S P Y Q C T I C T E Y C P S L S S M Q

BstXI (2533) AvrII (2590)

2501 AAACACATGAAGGGCCACAAGCCGAGGAGATCCCACCTGACTGGAGGATAGAGAAGACGTACCTCTACCTGTGTTATGTGTGAAGGGGGCCTAGGACTA
647▶ K H M K G H K P E E I P P D W R I E K T Y L Y L C Y V •

2601 GCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGC

HpaI (2734) MfeI (2745)

2701 TTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAACAAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTTAA

EcoRI (2830)

2801 AGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACCTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGG
2901 GATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCCA

SapI (3012) SspI (3069) SwaI (3083)

3001 AGGTTTGAAGTACTCTTCATTTCTTTATGTTTAAATGCACTGACCTCCCACATTCCCTTTTATGAAAATATTCAGAAATAATTTAAATACATCATTG
3101 CAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCT
3201 TTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCA
141◀ • N R T Y K L P I L E E I T T K V L K G

SacI (3344) BstXI (3373)

3301 TTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCAT
120◀ N M E I L 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
3401 CAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTGACCT
87◀ S Y P H R 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

StuI (3508)

3501 GCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGCTTGTGCTCATAGAGCATGGTG
54◀ G I Y A E 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

XmnI (3650)

3601 ATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCG
20◀ I K E T A V E V L E L D Q Q S I N F T K M

AseI (3716) SacI (3773)

3701 ATATACTATGCCGATGATTAATTGCAAACACAGCGTGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCCACC

SpeI (3871)

3801 GTACACGCTACCGCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAACAAACTCCCATTGACGTC

SnaBI (3999)

3901 AATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAAT
4001 ACGTAGATGTAAGTCCAAGTAGGAAAGTCCATAAGGTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTAC

NdeI (4104)

4101 TTGGCATATGATACACTTGTACTGCCAAGTGGGCAGTTTACCCTAAATACTCCACCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGG

SdaI (4282)PacI (4290)

4201 AACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGCGGTGACCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTCGAGGTTAATTAAGA

BspLU1II (4300)

4301 ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAA
4401 TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGACTATAAAGATACCAGGCGTTTCCCCTGGAAGTCCCTCGTGCCTCTCCTGTTCCGACCTG
4501 CCGCTTACCGGATACCTGTCGCTTTTCTCCCTCGGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTGCT

ApaLI (4614)

4601 CCAAGCTGGGCTGTGTGCACGAACCCCCGTTAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAAGTGGTGCCTAACTACGGCTACACTAGA
4701 GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGCCTAACTACGGCTACACTAGA
4801 AGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTATCGGCAACAAACACCCTGGTAGCGGTG
4901 GTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACAAAA

EagI (5050)

PacI (5030) SwaI (5039) NotI (5049)

5001 CTCACGTTAAGGGATTTTGGTACGGCTAGTTAATTAACATTTAAATCAGCGCGCAATAAAAATATCTTTATTTTTCATTACATCTGTGTGTTGTTTTT
5101 TGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAGAAC
5201 ATTTCTCTATCGAA