



PvuI (7)
SgfI (6) 1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82) **EcoNI (96)**
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACCGCGCCCGCCCTACCTGAGGCC
PvuII (239) **Bsu36I (291)**
EcoNI (287)
301 GCCATCCACGCGGTTGAGTCGCGTTCTGCCGCTCCCGCCTGTGGTGCCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGTCTTTGTTTCGTTT

NcoI (560)
BstEII (555)
AgeI (552) 501 TCTGTTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCCCTACCTGAGATCACCGGTACCATTGGAGAGCAAGGTGCTGCTGGCCGTCGCCCTGTGGCT
1 M E S K V L L A V A L W L

XmaI (611)
601 CTGCGTGGAGACCCGGGCGCCTCTGTGGGTTTGCCTAGTGTCTTCTTGTGATCTGCCAGGCTCAGCATAAAAAAGACATACTTACAATTAAGGCTAAT
13 C V E T R A A S V G L P S V S L D L P R L S I Q K D I L T I K A N
701 ACAACTCTTCAAATTACTTGCAGGGGACAGAGGACTTGGACTGGCTTTGGCCCAATAATCAGAGTGGCAGTGAAGCAAAGGTTGGAGGTGACTGAGTGC
47 T T L G I T C R G Q R D L D W L W P N N Q S G S E Q R V E V T T E C G
801 GCGATGGCCTCTTGTAAAGACTCACAAATCCAAAAGTGCAGAAATGACACTGGAGCCTACAAGTCTTACCGGAAACTTGTGCTTGCCTCGGT
80 S D G L F C K T L T I P K V I G N D T G A Y K C F Y R E T D L A S V
BglII (925)
901 CATTTATGTCTATGTTCAAGATTACAGATCTCCATTTATTGCTTCTGTTAGTGACCAACATGGAGTCGTGACATTACTGAGAACAAAAACAAAAGTGTG
113 I Y V Y V Q D Y R S P F I A S V S D Q H G V V Y I T E N K N K T V

DraIII (1040)
1001 GTGATCCATGTCTCGGGTCCATTTCAAATCTCAACGTGTACCTTTGTGCAAGATACCCAGAAAAGAGATTTGTTCTGATGGTAACAGAATTTCTGGG
147 V I P C L G S I S N L N V S L C A R Y P E K R F V P D G N R I S W

AseI (1171)
1101 ACAGCAAGAAGGGCTTTACTATTCCAGCTACATGATCAGCTATGCTGGCATGGTCTTCTGTGAAGCAAAAATTAATGATGAAAGTTACCAGTCTATTAT
180 D S K K G F T I P S Y M I S Y A G M V F C E A K I N D E S Y Q S I M
1201 GTACATAGTTGCTGTTGAGGGTATAGGATTTATGATGTGGTCTGAGTCCGCTCATGGAATTGAACATCTGTTGGAGAAAAGCTTGTCTTAAATTGT
213 Y I V V V V G Y R I Y D V V L S P S H G I E L S V G E K L V L N C

BstBI (1353)
1301 ACAGCAAGAAGTGAACAAATGTGGGATTGACTTCAACTGGGAATACCTTCTTCAAGCATCAGCATAAGAAAATTGTAACCAGACCTAAAAACCC
247 T A R T E L N V G I D F N W E Y P S S K H Q H K K L V N R D L K T
1401 AGTCTGGGAGTGAAGTGAAGAAATTTTGGAGCCTTAACTATAGATGGTGAACCCGGAGTGACCAAGGATTGTACCTGTGCAGCATCCAGTGGGCT
280 Q S G S E M K K F L S T L T I D G V T R S D Q G L Y T C A A S S G L
1501 GATGACCAAGAAGAACAGCACATTTGTGAGGTCAGGAAACCTTTGTTGCTTTTGGAAAGTGGCATGGAATCTCTGGTGGAGCCAGGTTGGGGGAG
313 M T K K N S T F V R V H E K P F V A F G S G M E S L V E A T V G E

BstEII (1627) **XcmI (1639)**
1601 CGTGTGCAATCCCTGCGAAGTACCTTGGTTACCCACCCAGAAATAAAATGGTATAAAAATGGAATACCCCTTGTGATCCAATCACACAATTAAGCGG
347 R V R I P A K Y L G Y P P P E I K W Y K N G I P L E S N H T I K A
1701 GGCATGTACTGACGATTTGAAAGTGAAGTGAAGGAGACAGGAAATACACTGTATCCTTACCAATCCCATTTCAAAGGAGAAAGCAGAGCCATGTGGT
380 G H V L T T I M E V S E R D T G N Y T V I L T N P I S K E K Q S H V V

BspLU11I (1891)
1801 CTCTCTGGTGTGTATGTCCACCCAGATTGGTGAAGAAATCTCTAATCTCTCTGTTGATTCTACAGTACGGCACCCTCAAACGCTGACATGTACG
413 S L V V Y V P P Q I G E K S L I S P V D S Y Q Y G T T Q T L T C T
1901 GTCTATGCCATTCCTCCCGCATTCCACTCCACTGGTATTGGCAGTTGGAGGAAGAGTGCGCCAACGAGCCAGCCAAAGCTGTCTCAGTGACAAACCAT
447 V Y A I P P P H H I H W Y W Q L E E E C A N E P S Q A V S V T N P
2001 ACCCTTGTGAAGAATGGAGAAGTGTGGAGACTCCAGGGAGGAAATAAAATGAAGTAAATAAAAATCAATTTGCTCTAATTGAAGGAAAAACAAAAC
480 Y P C E E W R S V E D F Q G G N K I E V N K N Q F A L I E G K N K T

BbrPI (2197)
2101 TGTAAGTACCCTTGTATCCAAGCGCAAAATGTGTCAGCTTTGTACAATGTGAAGCGGTCAACAAAAGTGGGAGAGGAGAGGGGTGATCTCCTTCCAC
513 V S T L V I Q A A N V S A L Y K C E A V N K V G R G E R V I S F H

EcoO109I (2208) **PstI (2276)**
2201 GTGACCAGGGTCTGAAATTAATTTGCAACCTGACATGCAGCCACTGAGCAGGAGAGCGTGTCTTTGTGGTCACTGCAGACAGATCTACGTTTGA
547 V T R G P E I T L Q P D M Q P T E Q E S V S L W C T A D R S T F E
2301 ACCTCACATGGTACAAGCTTGGCCACAGCCTCTGCCAATCCATGTGGGAGAGTTGCCACACCTGTTTGAAGAATTTGGATACTCTTTGAAATTTGAA
580 N L T W Y K L G P Q P L P I H V G E L P T P V C K N L D T L W K L N

NsiI (2453) **Tth111I (2476)**
2401 TGCCACCATGTTCTTAATAGCACAAAATGACATTTTGTGATGAGGCTTAAAGAAATGCATCTTGCAGGACCAAGGAGACTATGTCTGCTTGTCTCAAGAC
613 A T M F S N S T N D I L I M E L K N A S L Q D Q G D Y V C L A Q D
2501 AGGAAGACCAAGAAAAGACATTTGCGTGGTCAAGGAGCTCACAGTCTAGAGCGTGTGGCACCCACGATCACAGGAAACCTGGAGAATCAGACGACAAGTA
647 R K T K K R H C V V R Q L T V L E R V A P T I T G N L E N Q T T S

XcmI (2647)
2601 TTGGGAAAGCATCGAAGTCTCATGCACGGCATCTGGGAATCCCCCTCCACAGATCATGTGGTTAAAGATAATGAGACCCTTGTAGAAGACTCAGGCAT
680 I G E S I E V S C T A S G N P P P Q I M W F K D N E T L V E D S G I

2701 TGTATTGAAGGATGGGAACCGAACCTCACTATCCGCAGAGTGAGGAAGGAGGACGAAGGCCTCTACACCTGCCAGGCATGCAGTGTTCTTGGCTGTGCA
713▶ V L K D G N R N L T I R R V R K E D E G L Y T C Q A C S V L G C A
StuI (2757) SphI (2776)
2801 AAAGTGGAGGCATTTTTATAATAGAAGGTGCCTAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGCAAACCACTAGAATGCAGT
747▶ K V E A F F I I E G A •
NheI (2836) MscI (2842)
2901 GAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAAACAACAATTGCATTCAATT
HpaI (2974) MfeI (2985)
3001 TATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATCTAAATACAGCATAGCAAACCTTT
EcoRI (3070)
3101 AACCTCAAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTC
SapI (3252)
3201 ACCTTCTTTCATGGAGTTAAGATATAGTGATTTTTCCCAAGGTTTGAAGTACTCTTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCT
SspI (3309) SmaI (3323) EcoO109I (3384)
3301 TTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATAT
3401 CCCCCAGTTTAGTGTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAG
141▶ • N R T Y K L
SacI (3584)
3501 GGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATG
134▶ P 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
BstXI (3613)
3601 CCACAGGGGTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAG
100▶ G 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
StuI (3748)
3701 CAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGC
67▶ S 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
XmnI (3890)
3801 CGCCCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTC
34▶ A 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
AseI (3956)
3901 ATGGTGGCCCTCCTATAGTGAGTCGATTATACTATGCGGATATACTATGCGGATGATTAATTGTCAAACACAGCGTGGATGGCGTCTCCAGCTTATCTGA
0▶ M
SacI (4013)
4001 CGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACAGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGT
SpeI (4111)
4101 CCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTA
SnaBI (4239)
4201 CTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTATGACTGGGCATAATGCCAG
NdeI (4344)
4301 GCGGGCCATTTACCGTCATTGACGTCAATAGGGGCGTACTTGGCATATGATACACTTGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCC
4401 ATTGACGTCAATGAAAAGTCCCTATTGGCGTACTATGGGAACATACGTATTATTGACGTCAATGGGCGGGGTGTTGGGCGGTGAGCCAGCGGGCC
PacI (4530)
4501 ATTTACCGTAAGTTATGTAACGCCTGCAGGTTAAATAAGAACATGTGAGCAAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGT
PstI (4523) SdaI (4522) BspLU11I (4540)
4601 TTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGCGTTTCC
4701 CCCTGGAAGTCCCTCGTGCCTCTCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTTTCTCATAGC
ApaLI (4854)
4801 TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGCCGACCGCTGCGCTTATCCGGTA
4901 ACTATCGTCTTGTGTCACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCTGTTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTAC
5001 AGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGT
5101 AGCTCTTGATCCGGCAAACAAACCAGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTT
EagI (5290)
5201 TGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGGGATTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGAAT
PacI (5270) SmaI (5279) NotI (5289)
5301 AAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAG
5401 CAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA