



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) **HindIII (245)** **Bsu36I (291)**
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCTGTGGTGCCTCCTGAAGCTCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTTGTTCGTTT

NcoI (560) **BstEII (555)**
501 TCTGTTCTGCGCGCTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTACCATTGGACCAAAGAGAAATTCGCAGAAGTTCCTGGATGA
1 M D Q R E I L Q K F L D E
601 GGCCCAAAGCAAGAAAATTAATAAGAGGAGTTGCAATGAATTTCTGAAGCTGAAAAGGCAATCTACCAAGTACAAGGCAGACAAAACCTATCCTACA
13 A Q S K K I T K E E F A N E F L K L K R Q S T K Y K A D K T Y P T
701 ACTGTGGCTGAGAAGCCCAAGAATATCAAGAAAAACAGATATAAGGATATTTGCCCCTATGATTATAGCCGGGTAGAACCTACCTGATAACCTCTGATG
47 T V A E K P K N I K K N R Y K D I L P Y D Y S R V E L S L I T S D
EcoO109I (870)
801 AGGATTCAGCTACATCAATGCCAATTCATTAAGGGAGTTTATGGACCAAGGCTTATATTGCCACCCAGGGTCTTTATCTACAACCCCTCCTGGACTT
80 E D S S Y I N A N F I K G V Y G P K A Y I A T Q G P L S T T L L D F

SphI (943) **Eco47III (980)**
901 CTGGAGGATGATTTGGGAATATAGTGTCTTATCATTGTTATGGCATGCATGGAGTATGAAATGGGAAAGAAAAGTGTGAGCGCTACTGGGCTGAGCCA
113 W R M I W E Y S V L I I V M A C M E Y E M G K K K C E R Y W A E P
1001 GGAGAGATGCAGCTGGAATTTGGCCCTTCTCTGTATCCTGTGAAGCTGAAAAAGGAAATCTGATTATATAATCAGGACTCTAAAAGTTAAGTTCAATA
147 G E M Q L E F G P F S V S C E A E K R K S D Y I I R T L K V K F N

MscI (1138)
1101 GTGAAACTCGAATCTACAGTTTCATTACAAGAATTGGCCAGACCATGATGTACCTTCATCTATAGACCCTATTCTTGAGCTCATCTGGGATGTACG
180 S E T R T I Y Q F H Y K N W P D H D V P S S I D P I L E L I W D V R

BstAPI (1232)
1201 TTGTTACCAAGAGGATGACAGTGTCCCATATGCATTCAGTGTGCTGGTGTGGAAGGACTGGTGTATTGTTGCTATTGATTATACATGGATGTTG
213 C Y Q E D D S V P I C I H C S A G C G R T G V I C A I D Y T W M L

StuI (1366)
1301 CTAAGAAGATGGGATAATTCCTGAGAACTCAGTGTTCAGTTCAGTTCGCGGAAATGCGGACACAGAGGCCTCATTAGTTCAAACGCAGGAACAATATG
247 L K D G I I P E N F S V F S L I R E M R T Q R P S L V Q T Q E Q Y
1401 AACTGGTCTACAATGCTGTATTAGAATTTAAGAGACAGATGGATGTTATCAGAGATAAACATTCTGGAACAGAGAGTCAAGCAAAGCATTGTATTCC
280 E L V Y N A V L E L F K R Q M D V I R D K H S G T E S Q A K H C I P
1501 TGAGAAAAATCACACTCTCCAAGCAGACTCTTATTCTCCTAATTTACCAAAAAGTACCACAAAAGCAGCAAAAATGATGAACCAACAAAGGACAAAAATG
313 E K N H T L Q A D S Y S P N L P K S T T K A A K M M N Q Q R T K M

SapI (1654)
1601 GAAATCAAAGAATCTTCTCCTTTGACTTTAGGACTTCTGAAATAAGTGCAAAAGAAGAGCTAGTTTTGCACCCTGCTAAATCAAGCACTTCTTTTGACT
347 E I K E S S S F D F R T S E I S A K E E L V L H P A K S S T S F D
1701 TTCTGGAGCTAAATTACAGTTTTGACAAAAATGCTGACACAACCATGAAATGGCAGACAAAGGCATTTCCAATAGTTGGGGAGCCTCTCAGAAGCATCA
380 F L E L N Y S F D K N A D T T M K W Q T K A F P I V G E P L Q A K H Q
1801 AAGTTTGGATTTGGGCTCTCTTTGTTGAGGATGTTCTAATTCTAAACCTGTAATGCAGCAGGAAGATATTTAATTCAAAGTGCCAATKACACGG
413 S L D L G S L L F E G C S N S K P V N A A G R Y F N S K V P I T R
1901 ACCAAATCAACTCTTTGAAATGATACAGCAGAGAGAAACCAAGGAGGTGGACAGCAAGGAAAATTTCTTATTGGAATCTCAACCATGATTCTT
447 T K S T P F E L I Q Q R E T K E V D S K E N F S Y L E S Q P H D S

SnaBI (2087)
2001 GTTTTGTAGAGATGCAGGCTCAAAAAGTAATGCATGTTTCTTCAGCAGAAGTGAATTTCACTGCCATATGACTCTAAACACCAAAATACGTAATGCCTC
480 C F V E M Q A Q K V M H V S S A E L N Y S L P Y D S K H Q I R N A S

Acc65I (2191)
2101 TAATGTAAGACCATGACTCTAGTCTTGGTGTATATTCTTACATACCTTTAGTGGAAAATCCTTATTTTTATCATGGCCTCCAAGTGGTACCAGT
513 N V K H H D S S A L G V Y S Y I P L V E N P Y F S S W P P S G T S
2201 TCTAAGATGTCTCTTGATTTACCTGAGAAGCAAGATGGAAGTGTTCCTTCTTCTGTTGCCAACATCCTCTACATCCCTCTTCTTATTACAATT
547 S K M S L D L P E K Q D G T V F P S S L L P T S S T S L F S Y Y N
2301 CACATGATCTTTATCACTGAATTCCTCAACCAATATTTCTCTACTATTGAACAGGAGTGCAGCTGTACTAGCAACTGCCTCAAGGATAGATGATGAAAT
580 S H D S L S L N S P T N I S S L L N Q E S A V L A T A P R I D D E I
2401 CCCCCCTCACTTCTGTACGGACACCTGAATCTTTATTTGTTGAGGAGCTGGAGAATTTCAACAAATGTTCCAAATCTTATCCTCAGCTGTG
613 P P P L P V R T P E S F I V V E E A G E F S P N V P K S L S S A V
2501 AAGTAAAAATGGAACATCACTGGAATGGGGTGAACATCTGAACCAAGAAAATTTGATGACTCTGTGATCTTAGACCAAGCAAGAGTGTAAAACCTC
647 K V K I G T S L E W G G T S E P K K F D D S V I L R P S K S V K L

XbaI (2664)
2601 GAAGTCTAAATCAGAACTACATCAAGATGTTTCTTCTCCCCACCTCTCTCCAGAAAAGAACTCTAGAGTCTTCTTTCTTCTGCGATGAAGATTGTAT
680 R S P K S E L H Q D R S S P P P L P E R T L E S F F L A D E D C M

NcoI (2744) **BbsI (2778)**
2701 GCAGGCCAATCTATAGAAACATATTCTACTAGCTATCCTGACACCATGGAAAATTCACATCTTCAAAAACAGACACTGAAGACTCCTGGAAAAAGTTTC
713 Q A Q S I E T Y S T S Y P D T M E N S T S S K Q T L K T P G K S F

2801 ACAAGGAGTAAGAGTTTAAAAATTTGCGAAACATGAAAAAGAGTATCTGTAATTCTTGCCACCAAACAAGCCTGCAGAATCTGTTCCAGTCAAATACT
747▶ T R S K S L K I L R N M K K S I C N S C P P N K P A E S V Q S N N
MscI (2992)

2901 CCAGCTCATTCTGAATTTTGGTTTTGCAAACCGTTTTTCAAACCCAAAGGACCAAGGAATCCACCACCAACTTGAATATTTAAGCTAGCTGGCCAGA
780▶ S S S F L N F G F A N R F S K P K G P R N P P P T W N I •
3001 CATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTA

HpaI (3124) MfeI (3135)

3101 ACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAA
3201 ACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAG
3301 GCATAGGCATCAGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTTCATGGAGTTAAGATATAGTGTATTTCCCAAGGTTTGAAC

SapI (3402) SwaI (3473)

3401 TAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAT

EcoO109I (3534)

3501 AAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAA
3601 TTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAA
1+1▶ • N R T Y K L P I L E E I T T K V L K G N M E I
3701 TGAGCACAAGCAGCTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGG
117▶ 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 S Y P
StuI (3898)

3801 GTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAG
84▶ 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 G I Y
3901 GCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCCTGATGGCCGCCCCGACATGGTGCTTGTGCTCATAGAGCATGGTATCTTCTCAG
50▶ 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 I K E T

BbsI (4044) XmnI (4040)

4001 TGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGGTGGCCCTCCTATAGTGAGTCGTATTATACTATGCCGATATACTATG
17▶ A V E V L E L D Q Q S I N F T K M

AseI (4106)

4101 CCGATGATTAATTGTCAAACACAGCGTGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCT

SpeI (4261)

4201 ACCGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCCTTGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGG

SnaBI (4389)

4301 AGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGT
4401 ACTGCCAAGTAGGAAAGTCCATAAGGTGATGACTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATG
4501 ATACACTTGATGACTGCCAAGTGGCAGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTC

PacI (4680) SdaI (4672) BspLU11I (4690)

4601 ATTATTGACGTCAATGGGCGGGGTCGTTGGCGGTCAGCCAGCGGGCCATTTACCCTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGC
4701 AAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCA
4801 AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCG
4901 GATACCTGTCGCCCTTTCTCCCTCGGGGAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGG

ApaLI (5004)

5001 CTGTGTGCACGAACCCCGTTCCAGCCGACCGCTGCGCTTATCCGTAACATATCGTCTTGTAGTCCAAACCGGTAAGACACGACTTATCGCCACTGGCA
5101 GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGCTACACTAGAAGAACAGTAT
5201 TTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTTGT
5301 TTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAA

EagI (5440) PacI (5420) SwaI (5429) NotI (5439)

5401 GGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGAATAAAAATATCTTTATTTTTCATTACATCTGTGTGTTGTTTTTGTGTGAATC
5501 GTAACATAACATACGCTCTCCATCAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTAT
5601 CGAA