



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGTGTCGTGACTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)
Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACAGCGCCCGCCCTACCTGAGGCC
PvuII (239)
Bsu36I (291)
301 GCCATCCACGCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTGGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

BstEII (555)
AgeI (552) 501 TCTGTTCTGCGCGCTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTACCATTGGTGTGAATAACGCTTTAGAAACAATTATTAACGC
NcoI (560) 1 M V S N N A L E T I I N A

BstEII (602)
FspI (666)
601 CCGTTACCAGGCAAGAGGGGCTGTGGCAGATTCATCTGCAGACGGAAAAATCAGCGCCATTGATCGCAATCCGGGTGATGCCATAACTGAAAC
13 R L P G E E G L W Q I H L Q D G K I S A I D A Q S G V M P I T E N
701 AGCCTGGATGCCAACAAGTTTAGTTATACCGCGTTTGTGGAGCCACATATTCACCTGGACACCACGAAACCGCCGACAACCGAACTGGAATCAGT
47 S L D A E Q G L V I P P F V E P H I H L D T T Q T A G Q P N W N Q
801 CCGCAGCTGTTTGAAGCATTGAACGCTGGGCGAGCGAAAGCTTATTAACCCATGACGATGTGAAACAACCGCATGGCAAACGCTGAAATGGCA
80 S G T L F E G I E R W A E R K A L L T H D D V K Q R A W Q T L K W Q
901 GATTGCCAACGGCATTGACGATGTGCGTACCCATGTGATGTTTCCGATGCAACGCTAACTGCCTGAAAGCAATGCTGGAAGTGAAGCAGGAAGTCGCG
113 I A N G I Q H V R T H V D V S D A T L T A L K A M L E V K Q E V A

Bsu36I (1030)
1001 CCGTGGATTGATCTGCAAACTGTCGCCTCCCTCAGGAAGGGATTTTGTGATCCCAACGGTGAAGCGTTGCTGGAAGAGGCGTTACGCTTAGGGCAG
147 P W I D L Q I V A F P Q E G I L S Y P N G E A L L E E A L R L G A
1101 ATGTAGTGGGGGCGATTCCGCATTTTGAATTTACCCGTGAATACGGCGTGGAGTGCATGATAAAACCTTCGCCCTGGCGCAAAAATACGACCGTCTCAT
180 D V V G A I P H F E F T R E Y G V E S L H K T F A L A Q K Y D R L I
1201 CGACGTTCACTGTGATGAGATCGATGACGAGCAGTCGCGCTTTGTCGAAACCGTTGCTGCCCTGGCGCACCATGAAGGCATGGCGCGCGAGTACCGCC
213 D V H C D E I D D E Q S R F V E T V A A L A H H E G M G A R V T A

Bst1107I (1335)
BstXI (1388)
1301 AGCCACACACGGCAATGCACTCCTATAACGGGGCGTATACCTCACGCCTGTTCCGCTTGTGAAATGTCCGGTATTAACCTTTGTCGCCAACCCGCTGG
247 S H T T A M H S Y N G A Y T S R L F R L L K M S G I N F V A N P L

SnaBI (1428)
BsaBI (1426)
SspI (1402)
1401 TCAATATTCATCTGCAAGGACGTTTCGATACGTATCCAAAACGTCGCGGCATCACGGCGTTAAAGAGATGCTGGAGTCCGGCATTACGCTGCTTTGG
280 V N I H L Q G R F D T Y P K R R G I T R V K E M L E S G I N V C F G

BbsI (1510)
BstAPI (1551)
NdeI (1561)
1501 TCACGATGATGCTTCGATCCGTTGGTATCCGCTGGGAACGGCAATATGCTGCAAGTGTGCATATGGGGCTGCATGTTTCCAGTTGATGGGCTACGGG
313 H D D V F D P W Y P L G T A N M L Q V L H M G L H V C Q L M G Y G
1601 CAGATTAACGATGGCCTGAATTTAATCACCCACACAGCGCAAGGACGTTGAATTTGACGATTACGGCATTGCCCGGAAACAGCGCAACCTGATTA
347 Q I N D G L N L I T H H S A R T L N L Q D Y G I A A G N S A N L I

NgoMIV (1704)
1701 TCCTGCCGGTGAATAAGGGTTTGTGCGCTGCGCGTCAGGTTCCGGTACGTTATTCGGTACGTGGCGGCAAGGTGATTGCCAGCACACAACCGGCACA
380 I L P A E N G F D A L R R Q V P V R Y S V R G G K V I A S T Q P A Q

XcmI (1802)
1801 AACCCGTTATCTGGAGCAGCCAGAAGCCATCGATTACAAACGATACAGCGGAGCTAATGGCGTCATGGCTAAGATCGTGAAGTCAAACACCCACTC
413 T T V Y L E Q P E A I D Y K R I S G A N G V M A K I V E V K H P L

NruI (1953)
EcoRV (1933) 1901 GTCAAACACAAGCTGGGACTGATGCGTGAGCAAGATATCAGCACCAAGCGTTTCGCGAACTCGCTTCCGAAGTGGGTAGCCTGCTGACTTACGAAGCGA
12 V K H K L G L M R E Q D I S T K R F R E L A S E V G S L L T Y E A
Eco47III (1946)
2001 CCGCCGACCTCGAAACGGAAAAAGTAACTATCGAAGGCTGGAACGGCCCGGTAGAAATCGACCAGATCAAAGTAAAGAAAATTACCGTTGTGCCAATTCT
45 T A D L E T E K V T I E G W N G P V E I D Q I K G K K I T V V P I L

Psp1406I (2137)
XmnI (2134)
Acc65I
2101 GCGTGGGGTCTTGGTATGATGGACGGTGTGCTGGAAAACGTTCCGAGCGCGCATCAGCGTTGTCGGTATGTACCCTAATGAAGAAACGCTGGAGCCG
78 R A G L G M M D G V L E N V P S A R I S V V G M Y R N E E T L E P

AgeI (2272)
NruI (2288)
2201 GTACCGTACTCCAGAACTGGTTTCTAACATCGATGAGCGTATGGCGCTGATCGTTGACCAATGCTGGCAACCGGTGGTTCGGTTATCCGACCATCG
112 V P Y F Q K L V S N I D E R M A L I V D P M L A T G G S V I A T I
2301 ACCTGCTGAAAAAGCGGCTGCAGCAGCATCAAAGTTCTGGTCTGGTAGCTGCGCCAGAAGGTATCGCTGCGCTGAAAAAGCGCACCCGGACGTCGA
145 D L L K K A G C S S I K V L V L V A A P E G I A A L E K A H P D V E

Bst1107I (2403)
HindIII (2495)
2401 ACTGTATACCGCATCGATTGATCAGGACTGAACGAGCAGGATACATTATTCGGGCCCTCGCGATGCCGGTGACAAAATCTTTGGTACGAAATAAGCT
178 L Y T A S I D Q G L N E H G Y I I P G L G D A G D K I F G T K •

MscI (2517)
NheI (2511)
 2501 TGTCGAAATTCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAAT

HpaI (2649) **MfeI (2660)**
 2601 TTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTG

EcoRI (2745)
 2701 TGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTG

2801 AATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTTCATGGAGTTTAAGATA

SapI (2927) **SspI (2984)** **Swal (2998)**
 2901 TAGTGATTTTCCCAAGTTTGAAGTAGCTCTTCATTTCTTTATGTTTAAATGCACTGACCTCCACATTCCCTTTTGTAGTAAATATTCAGAAATAAT

EcoO109I (3059)
 3001 TTAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAG

3101 GGAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCCTCAATGGTGGTT

141 • N R T Y K L P I L E E I T T
SacI (3259) **BstXI (3288)**
 3201 TTGACCAGCTTGCCATTCACTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGATGG

125 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 C P S V V R I S
 3301 ATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGC

92 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 G I A I A E A
StuI (3423)
 3401 ACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGCTTGTGTCC

59 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 G V H H K N D
BbsI (3569) **XmnI (3565)**
 3501 TCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGGTGGCCCTCCTATAGTGAGTCG

25 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
AseI (3631) **SacI (3688)**
 3601 TATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTA

SpeI (3786)
 3701 TATAGACCTCCCACGCTACACGCCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAACAA

3801 ACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTA

SnaBI (3914)
 3901 ATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGGGGCCATTTACCGTCATTGACGT

NdeI (4019)
 4001 CAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCCTAAATACTCCACCATTGACGTCAATGAAAAGTCCCTAT

SdaI (4197)
 4101 TGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGGGGGGTCTGTTGGGCGGTGAGCCAGGGGGCCATTTACCCTAAGTTATGTAACGCCCT

PacI (4205) **BspLU11I (4215)**
 4201 GCAGGTTAAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGA

4301 CGAGCATCAGAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCT

4401 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG

ApaLI (4529)
 4501 TGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTGAGCCGACCGCTGCGCCTTATCCGTAACATATCGTCTTGAGTCAACCCGGT

4601 AAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC

4701 TACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA

4801 CCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGGTCTGACGC

EagI (4965) **NotI (4964)**
PacI (4945) **Swal (4954)**
 4901 TCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCAATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTTCATTACATC

5001 TGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAG

5101 TGCAGGTGCCAGAACATTTCTCTATCGAA