



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

Psp1406I (203)
201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTACTCAACTCTACGCTTTTGTTCGTTT

MluI (564)
BspEI (558) 501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGCTCCGGAACGCGTACCATGGGGGAGAAGCTGGAGCTGAGACTG
PshAI (595) 1 M G E K L E L R L

BsaBI (654)
601 AAGTCGCCCGTGGGGGCTGAGCCCCGCGTCTACCGTGGCCGCTGCCGCTACGATAAACATCAGATGCTGCTCATGAAATCATCGAGACCATCCGAT
10 K S P V G A E P A V Y P W P L P V Y D K H H D A A H E I I E T I R
BspHI (674)

SphI (781)
701 GGGTCTGTGAAGAAATCCCGATCTCAAGCTCGTATGGAGAATTACGTTTTAATTGACTATGACACAAAAGCTTCGAGAGCATGCAGAGGCTCTGCGA
43 W V C E E I P D L K L A M E N Y V L I D Y D T K S F E S M Q R L C D
801 CAAGTACAACCGTCCCATCGACAGCATCCACCAGCTGTGGAAGGGCACCACGACGCCATGAAGCTGAACACGCGGCCGTCCACTGGACTCCTGCGCCAT
76 K Y N R A I D S I H Q L W K G T T Q P M K L N T R P S T G L L R H

BstEII (926)
DraIII (921)
901 ATCCTGCAGCAGGTCTACAACCACTCGGTGACCGACCCGAGAAGCTCAACAACACGAGCCCTTCTCCCGGAGGTGTACGGGGAGACCTCCTTCGACC
110 I L Q Q V Y N H S V T D P E K L N N Y E P F S P E V Y G E T S F D
1001 TGGTGGCCAGATGATTGATGAGATCAAGATGACCGACGACGACCTGTTTGTGGACTTGGGGAGCGGTGTGGCCAGGTGCTGCTCCAGTTGCTGCTGC
143 L V A Q M I D E I K M T D D D L F V D L G S G V G Q V V L Q V A A A
1101 CACCAACTGCAAACATCACTATGGCGTCGAGAAAGCAGACATCCCGCCAAAGTATGCGGAGACCATGGACCGGAGTTTCAAGGAGTGAATGGTAT
176 T N C K H H Y G V E K A D I P A K Y A E T M D R E F R K W M K W Y

SphI (1208)
1201 GAAAAAAGCATGCAGAATACATTGGAGAGAGCGATTTCCTCTCAGAAGAGTGGAGGGAGCGAATCGCCAACACGAGTGTATATTTGTGAATAATT
210 G K K H A E Y T L E R G D F L S E E W R E R I A N T S V I F V N N
1301 TTGCCTTTGGTCTGAGGTGGATCACCAGCTGAAGGAGCGTTTGCAAACATGAAGGAAGGTGGCAGAATCGTGTCTCGAAACCTTTGACCTCTGAA
243 F A F G P E V D H Q L K E R F A N M K E G G R I V S S K P F A P L N
1401 CTTCAGAATAAACAGTAGAACTTGGTACATCGGCACCATCATGCGCGTGGTGGAGCTCTCGCCCTGAAGGGCTCGGTGCTGTTGGAGGCAAGCCA
276 F R I A N S R N L S D I G T I M R V E L S P L K G S V L K G S W T G K K P
1501 GTCTCCTACTNCTGCACACTATCGACCGACCATATTGAAAACATTTTTCTAGTCTGAAAAACCCAAAACCTCAGGGAGSAACAGGAGGCAGCCGGC
310 V S Y Y L H T I D R T I L E N Y F S S L K N P K L R E E Q E A A R
1601 GCCGCCAGCAGCGGAGAGCAAGAGCAACGCGCCACGCCCCTAAGGGCCAGAGGGCAAGTGGCCGGCCCGCCGACGCCCCATGGACTCTGGTGC
343 R R Q Q R E S K S N A A T P T K G P E G K V A G P A D A P M D S G A
1701 TGAGGAAGAGAAGGCGGGAGCAGCCACCGTGAAGAAAGCGTCTCCCTCAAAGCCGCAAGAAGAAGCTAAACAAGAAGGGGAGGAAGATGGCTGGCCG
376 E E E K A G A A T V K K P S P S K A R K K K L N K K G R K M A G R

Tth111I (1889)
1801 AAGCGCGGCCCCCAAGAAGATGAACACTGCGAACCCGAGCGGAAGCCCAAGAAGAACCAACTGCACTGGATGCCCTGCACGCTCAGACCTGTCTC
410 K R G R P K K M N T A N P E R K P K K N Q T A L D A L H A Q T V S
1901 AGACGGCGCCTCTCAGCCAGATGCCTACAGTCCCTCACAGCCGTTTACAGNCTACTCCGAGCTGCAGGGCACTCCCCAGCTCCGCTGCT
443 Q T A A S S P Q D A Y R S P H S P F Y Q L P P S V Q R H S P N P L L

XbaI (2033)
2001 GGTGGCGCCACCCGCGCGTGCAGAAGCTTCTAGAGTCTTCAAGATCCAGTACCTGCAGTTCTTGGCATAACAAAGACCCCCAGTACAAGGCC
476 V A P T P P A L Q K L L E S F K I Q Y L Q F L A Y T K T P Q Y K A
2101 AGCCTGCAGAGCTCTGGGCGAGGAGAAGGAGAAGCAACGCCAGCTCCTGGGTGCGGCTCAGCAGTCTCAGCCACTGCCAGGCCAGAAGGAGGAGA
510 S L Q E L L G Q E K E K N A Q L L G A A Q Q L L S H C Q A Q K E E

BglII (2281)
2201 TCAGGAGGCTGTTTCAGCAAAAATTGGATGAGCTGGGTGTGAAGGCGCTGACCTACAACGACCTGATTCAAGCGCAGAAGGAGATCTCCGCCATAACCA
543 I R R L F Q Q K L D E L G V K A L T Y N D L I Q A Q K E I S A H N Q

BssHII (2341)
2301 GCAGCTGCGGGAGCAGTCGGAGCAGCTGGAGCAGGACAACCGCGCTCCGCGCCAGAGCTTGCAGTCTCAAGGCTCGTCCGAGGAGCTGCAGCTG
576 Q L R E Q S E Q L E Q D N R A L R G Q S L Q L L K A R C E E L Q L

BglII (2458)
2401 GACTGGGCCACGCTGTGCTGGAGAAGCTGTTGAAGGAGAAGCAGGCCCTGAAGAGCCAGATCTCGGAGAAGCAGAGGCACTGCTGGAGCTGCAGATCA
610 D W A T L S L E K L L K E K Q A L K S Q I S E K Q R H C L E L Q I
2501 GCATTGTGGAGCTAGAGAAGAGCCAGCGCAGCAGGAGCTCCTGCAGTCAAGTCTGTGTGCCGCTGACGACGCCCTGCTCCCTGCACCTGCGTGGAA
643 S I V E L E K S Q R Q Q E L L Q L K S C V P P D D A L S L H L R G K
2601 GGGCGCCTGGGCGGAGCTGGAGCCTGACGCCAGCCGGCTGCACCTGGAGCTGGACTGCACCAAGTCTCGCTGCTTACCTGAGCAGCAGTGGAGCCG
676 G A L G R E L E P D A S R L H L E L D C T K F S L P H L S S M S P
2701 GAGCTCTCCATGAACGGCCAGGCTGTGGCTATGAGCTCTGCGGTGTGCTGAGCCGGCTTCTGCGAAGCAGAACACGCCCCAGTACCTGGCCTACCCC
710 E L S M N G Q A A G Y E L C G V L S R P S S K Q N T P Q Y L A S P
2801 TGGACCAGGAGGTGGTGCCTGTACCCCTAGCCACGTCCGCGCGCCGCGCTGGAGAAGCTGTCTGGCCTAGCCGACCCGACTACACTAGGCTGTCCCC
743 L D Q E V V P C T P S H V G R P R L E K L S G L A A P D Y T R L S P

DraII (2937)

2901 GGCCAAGATTGTGCTGAGGCGGCACCTGAGCCAGGACACACGGTGCCCGGCAGGCGGCTGCCAGTGAGCTGCATTGAGAGCTGAGCACACCAAGGAG
776▶ A K I V L R R H L S Q D H T V P G R P A A S E L H S R A E H T K E
3001 AACGGCCTTCCCTACCAGAGCCCGAGCGTGCCTGGCAGCATGAAAGCTGAGCCCTCAGGACCCGCGCCCTGTCCCTGGGGCTTGCAGCTTGTGGAG
810▶ N G L P Y Q S P S V P G S M K L S P Q D P R P L S P G A L Q L A G

BssHII (3128)

3101 AGAAGAGCAGTGAGAAGGGCCTGAGAGAGCGCGCTACGGCAGCAGCGGGGAGCTCATCACCAGCCTGCCATCAGCATCCCCCTCAGCACCGTGCAGCC
843▶ E K S S E K G L R E R A Y G S S G E L I T S L P I S I P L S T V Q P
3201 CAACAAGCTCCCGGTGAGCATTCCCTGGCCAGCGTGGTGTGCCAGCCGCGCCGAGAGGGGAGGAGCACCCCCAGTCCCGTGTGCAGCCCCGTGAC
876▶ N K L P V S I P L A S V V L P S R A E R A R S T P S P V L Q P R D
3301 CCCTCGTCCACACTTGAAAAGCAGATTGGTGTAAATGCCACGGTGTGGGAGCAGAAGCCTTGCCCTGGCCCCGAGGCTTCTCTACGCTGGCTCGG
910▶ P S S T L E K Q I G A N A H G A G S R S L A L A P A G F S Y A G S
3401 TGGCATCAGCGGGGCTTGGCGGGCAGCCGGCTCTCTCACACCTGGAGCCGAGCCGGCCACCTTGGATGAGTCTCCAGCTCTGGGAGCCTTTTTC
943▶ V A I S G A L A G S P A S L T P G A E P A T L D E S S S S G S L F A

SmaI (3507)

3501 CACCGTGGGGTCCCGCAGCTCCACGCCACAGCACCCCTGCTGTGGCACAGCCCCGAACTGCTTCTGCCTCCTCCGCCCCACAGCTCTCTCCAGT
976▶ T V G S R S S T P Q H P L L L A Q P R N S L P A S P A H Q L S S S
3601 CCCCGCTTGGTGGGGCCCGCCAGGGCCGTGGCCGAGCCAGCAAGGGAGACTGCCCTCCGATTCCGGCTTCTCAGATCCTGAGAGTGAAGCCAAGA
1010▶ P R L G G A A Q G P L P E A S K G D L P S D S G F S D P E S E A K
3701 GGAGGATTGTGTTACCATCACCCTGGTGGGGCAGTGCCAAGCAGTGCCTCCAGCAAGCACAGCCCCCTGACCGCCAGCGCCCGTGGGGACTGTGT
1043▶ R R I V F T I T T G A G S A K Q S P S S K H S P L T A S A R G D C V
3801 GCCGAGCCACGGGAGGACAGTGCAGGCGCGGCGGGAAGCGAGCATCTGCGGGGACGCCAGCTTGGAGCGAGCGTGTCCCCAAGCGCCGAGCC
1076▶ P S H G Q D S R R R R K R A S A G T P S L S A G V S P K R R A
3901 CTGCGTCCGTCGCTGGCCTTTTACACAGCCTTGGGGTCTCCCTCAACTCAACTCCATGGTCAGTAACATCAACCAGCCCTGGAGATTACAGCCA
1110▶ L P S V A G L F T Q P S G S P L N L N S M V S N I N Q P L E I T A
4001 TCTCGTCCCCGGAGACCTCCCTGAAGAGTCCCTGTGCCCTACCAGGACCACGACCAGCCCCCGTGTCAAGAAGGAGCGGCTCTGAGCCAGACAA
1143▶ I S S P E T S L K S S P V P Y Q D H D Q P P V L K K E R P L S Q T N
4101 TGGGGCACACTCTCCCACTCAGACGAGGAGCCAGGCTGTGAGGACGAGCCAGCAGTGTGGAATTGAGAGAAAAATTGCAACAATCTCTTA
1176▶ G A H Y S P L T S D E E P G S E D E P S S A R I E R K I A T I S L

SgrAI (4265)

4201 GAAAGCAAATCTCCCCGAAAACCTTGGAAAATGGTGGTGGCTTGGCGGGAAGGAAGCCCGCGCCGCGGCGAGCCAGTCAATAGCAGCAAGTGAAGT
1210▶ E S K S P P K T L E N G G G L A G R K P A P A G E P V N S S K W K
4301 CCACCTTCTCGCCATCTCCGACATCGGCTGGCCAAGTCCGGCGAGCCCGCTGCAGGCGAGCTCCGCCCTCAGCCAGAACTCCCTGTTACAGTTCCG
1243▶ S T F S P I S D I G L A K S A D S P L Q A S S A L S Q N S L F T F R

XmaI (4493)

SrfI (4492)

4401 GCCCGCCTTGAGAGGCCCTCTGCCGATGCCAAGTGGCCGCTCACCCAGGAAAAGGCTTCCCGGCTCCCTGTGGGGGCTGACGGACTCAGCCCGGC
1276▶ P A L E E P S A D A K L A A H P R K G F P G S L S G A D G L S P G
4501 ACCAACCTGCCAACGGTGCACCTTCCGGCGGGGCTGGCCGCGACCTGAGTTTACACAGCTTCAAGTATGATGGTCTTCTTCCCCACAAGGGCCCCG
1310▶ T N P A N G C T F G G G L A A D L S L H S F S D G A S L P H K G P
4601 AGGCGCCCGGCTGAGCTCCCCGCTGAGCTTCCCTCGCAGCGCGCAAGGAGGGCTCGGACGCCAACCTTCTGAGCAAGAGGACGCTGGACGGCT
1343▶ E A A G L S S P L S F P S Q R G K E G S D A N P F L S K R Q L D G L

SfiI (4799)

4701 GGCTGGGCTGAAGGGCGAGGGCAGCCGCGCAAGGAGGAGGGGAGGGCGGCTACCCTGTGCGGGCCACGGACAAGACCCACTGCTGAGCGGCAAG
1376▶ A G L K G E G S R G K E A G E G G L P L C G P T D K T P L L S G K

XmaI (4810)

SalI (4824)

4801 GCCGCAAGGCCCCGGGACCGAGGTCGACCTCAAGAATGGCCACAACCTTTCATCTCTGCGCGGCGCTGCCCTCCCGAAGCCTCCTCAGCGGCCCGG
1410▶ A A K A R D R E V D L K N G H N L F I S A A A V P P G S L L S G P

AgeI (4954)

4901 GCCTGGCCCCGGCGGCTCTCCGACGGCGCGCGGCTCTCCGCCAGACGACCGGTCTTCTGGGCCCTTCCCGCGGGACCGCAGTTCGGCT
1443▶ G L A P A A S S A G G A A S S A Q T H R S F L G P F P P G P Q F A L

ApaLI (5094)

StuI (5087)

BbrPI (5097)

5001 CGGCCCATGTCCCTGCAGGCCAACCTCGGCTCCGTGGCCGGCTCCTCCGTGTGCAGTGCCTGTTAGCTCTGTGCCGGCCCGCGCAGGCTGGTGCAC
1476▶ G P M S L Q A N L G S V A G S S V L Q S L F S S V P A A A G L V H

NheI (5187)

5101 GTGTCGTCGCTGCCACCAGACTGACCAACTCGCAGCCATGGGAGCTTTTCCGGGGTGGCAGGCGGCACAGTTGGAGGTAAGTGGCTAGCTGGCCAG
1510▶ V S S A A T R L T N S H A M G S F S G V A G G T V G G N •
5201 ACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGT

HpaI (5325) MfeI (5336)

5301 AACCATTAAGCTGCAATAAACAGTTAAACAACAACATTGCATTATTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAA

EcoRI (5421)

5401 AACCTCTACAAATGTGGTATGGAATTCTAAAAATACAGCATAGCAAACTTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAA

5501 GGCATAGGCATCAGGGGCTTGGCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAA

SspI (5660)

SwaI (5674)

5601 CTAGCTCTTCATTTCTTTATGTTTTAAATGCAGTACCTCCACATTCCCTTTTATGATAAATTCAGAAATAATTTAAATACATCATTGCAATGAAAA

5701 TAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAA

5801 ATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTTCATCTCA

5901 ATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGG
 117 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 S Y P

6001 GGTGCCTGACAGCCACAATGGTGTCAAAGTCCTTCTGCCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTA
 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

6101 GGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGCTCTGATGGCCGCCGACATGGTGCTTGTTCCTCATAGAGCATGGTGATCTTCTCA
 51 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

BbsI (6245)
XmnI (6241)

6201 GTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTAT
 17 T A V E V L E L D Q Q S I N F T K M

AseI (6307)

6301 GCCGATGATTAATTGTCAAACACGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCC

SpeI (6462)

6401 TACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAACAAACTCCCATTGACGTCAATGGGGTG

SnaBI (6590)

6501 GAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATG

NdeI (6695)

6601 TACTGCCAAGTAGAAAGTCCATAAAGTCATGTACTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGCCATAT

6701 GATACACTTGATGTACTGCCAAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGT

PacI (6881) **BspLU11I (6891)**

6801 CATTATTGACGTCAATGGGCGGGGTCGTTGGCGGTGACCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAG

6901 CAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTC

7001 AAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGTTACC

7101 GGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGG

ApaLI (7205)

7201 GCTGTGTGCACGAACCCCGTTGACGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGC

7301 AGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTA

7401 TTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTTG

7501 TTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACTCACGTTA

PacI (7621) **SwaI (7630)** **NotI (7640)**

7601 AGGGATTTGGTTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAAT

7701 CGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTA

7801 TCGAA