



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

101 GAGAAAGTGGCGGGGTAACCTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) **PvuII (239)** **EcoNI (287)**
201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTGTTTCGTTT

KasI (535) **AgeI (552)** **NcoI (560)**
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTACCATTGGAGCAAGGTGCTGCTGGCCGTCCGCCTGTGGCT
1 M E S K V L L A V A L W L

XmaI (611)
601 CTGCGTGGAGACCCGGGCGCCTCTGTGGTGGCTAGTGTCTTCTGATCTGCCAGGCTCAGCATAAAAAAGACATACTTACAATTAAGGCTAAT
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 ACAACTTTCAAATTAAGTGCAGGGGACAGAGGGACTGGACTGGCTTTGGCCCAATAATCAGAGTGGCAGTGGCAAAAGGGTGGAGGTGACTGAGTGCA
47 T T L Q I T C R G V Q R D L D W L W P N N Q S G S E Q R V E V T E C
801 GCGATGGCCTCTTCTGTAAGACACTCACAATTCAAAAGTGATCGAAATGACACTGGAGCCTACAAGTGTCTTACCAGGAACTGACTTGGCCTCGGT
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
901 CATTATGTCATGTTCAAGATTACAGATCTCCATTTATTGCTTCTGTAGTGACCAACATGGAGTCTGTACATTACTGAGAACAAAAACAAACTGTG
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 GTGATCCATGTCTCGGGTCCATTTCAAATCTCAACGTGTCACTTTGTGCAAGATACCCAGAAAAGAGATTTGTTCTGATGGTAACAGAATTTCTGGG
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 ACAGCAAGAAGGGCTTACTATTCCAGCTACATGATCAGCTATGCTGGCATGGTCTTCTGTGAAGCAAAAATTAATGATGAAAGTTACCAGTCTATTAT
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 GTACATAGTTGCTGTTGATAGGATTTATGATGTGGTCTGAGTCCGTCTCATGAATGAACATCTGTTGGAGAAAAGCTTGTCTTAAATTTG
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 ACAGCAAGAAGTGAACAAATGTGGGATTGACTTCAACTGGGAATACCCCTTCTCGAAGCATCAGCATAAGAAAATTGTAAACCGAGACCTAAAAACCC
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 AGTCTGGGAGTGAAGTGAAGAAATTTTGGACACCTTAACATAGATGGTGTAAACCCGGAGTGACCAAGGATTGTACACCTGTGCAGCATCCAGTGGGCT
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 GATGACCAAGAAGAACAGCACATTTGTGAGGTCATGAAAACCTTTGTTGCTTTTGGAAAGTGGCATGGAATCTCTGGTGGAAAGCCACGGTGGGGAG
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
1601 CGTGTGCAATCCCTGCGAAGTACCTGGTTACCCACCCAGAAATAAAATGGTATAAAAATGGAATACCCCTTGTGTTCAATCACACAATTAAGCGG
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 GGCATGACTGACGATTATGGAAGTGAAGTGAAGAGACACAGGAAATTAACACTGTCATCCTTACCAATCCCATTTCAAAGGAGAAGCAGAGCCATGTGGT
380 G H V L 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

BspLUII (1891)
1801 CTCTCTGGTTGTGTATGTCCACCCAGATTGGTGAAGAACTCTAATCTCTCTGTTGATTCTACCAGTACGGCACCCTCAAACGCTGACATGTACG
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 GTCTATGCCATTCCTCCCGCATCACATCCACTGGTATTGGCAGTTGGAGGAAGAGTGCGCCAACGAGCCAGCCAGTGTCTCAGTGACAAAACCCAT
447 V Y A I C P P H H I H W Y Q L E E E C A N E P S Q A V S V T N P
2001 ACCCTTGTGAAGAATGGAGAAGTGGAGGACTTCCAGGGAGAAATAAATTAAGTGAATTAATAAATCAATTTGCTCTAATTGAAGGAAAAACAAAC
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 TGTAAGTACCCTTGTATCCAAGCGCAAAATGTGTCAGCTTTGTACAAATGTGAAGCGGTCAACAAAGTGGGAGAGGAGAGGGGTGATCTCCTTCCAC
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 V I S F H

ApaI (2271)
2201 GTGACCAGGGTCTGAAATTACTTTGCAACCTGACATGCAGCCACTGAGCAGGAGAGCGTGTCTTTGGTGCAGTGCAGACAGATCTACGTTTGA
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 ACCTCACATGGTACAAGCTTGGCCACAGCCTCTGCCAATCCATGTGGGAGAGTTGCCACACCTGTTTGAAGAACTGGATACTCTTTGAAATTTGAA
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 TGCCACCATGTTCTTAATAGCACAAATGACATTTTATCATGGAGCTTAAGAATGCATCCTTGCAGGACCAAGGAGACTATGTCTGCCTTGCTCAAGAC
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 AGGAAGACCAAGAAAAGACATTTGCGTGGTCAGGAGCTCAGCTCCTAGAGCGTGTGGCACCACCGATCACAGGAAACCTGGAGAATCAGACGACAAGTA
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
2601 TTGGGAAAGCATCGAAGTCTCATGCACGGCATCTGGAAATCCCTCCACAGATCATGTGGTTAAAGATAATGAGACCCTGTAGAAGACTCAGGCAT
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

StuI (2757) **SphI (2776)**
2701 TGTATTGAAGGATGGAAACCGAACCTCACTATCCGAGAGTGAGGAAGGAGGACGAAGGCGCTTACACCTGCCAGGCATGCAGTGTCTTGGCTGTGCA
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
2801 AAAGTGGAGGCAATTTTTCATAATAGAAGGTGCCAGGAAAAGACGAACCTGGAATCATTATTCTAGTAGGCACGGCGGTGATTGCCATGTTCTTCTGGC
747 K V E A F F I I E G A Q E K T N L E I I I L V G T A V I A M F F W

RsrII (2920) **BstXI (2969)** **BamHI (2979)**
2901 TACTTCTGTCATCATCTACGGACCGTTAAGCGGGCCAAATGGAGGGAACTGAAGACAGGCTACTTGTCCATCGTCATGGATGCAGATCAAGTCCCAT
780 L L L V I I L R T V K R A N G G E L K T G Y L S I V M D P D E L P L

3001 GGATGAACATTGTGAACGACTGCCTTATGATGCCAGCAATGGGAATCCCCAGAGACCGGCTGAAGCTAGGTAAGCCTCTTGGCCGTGGTGCCTTTGGC
813▶ D E H C E R L P Y D A S K W E F P R D R L K L G K P L G R G A F G

DraIII (3185)

3101 CAAGTGATTGAAGCAGATGCCTTTGGAATTGACAAGACAGCAACTTGCAGGACAGTAGCAGTCAAATGTTGAAAGAAGGAGCAACACACAGTGAGCATC
847▶ Q V I E A D A F G I D K T A T C R T V A V K M L K E G A T H S E H

3201 GAGCTCTCATGTCTGAACTCAAGATCCTCATTCAATTTGGTCAACATCTCAATGGTCAACCTTCTAGGTGCCTGTACCAAGCCAGGAGGGCCACTCAT
880▶ R A L M S E L K I L I H I G H L N V V N L L G A C T K P G G P L M

3301 GGTGATTGTGAATTCTGCAAAATTTGGAACCTGTCCACTTACCTGAGGAGCAAGAGAAATGAAATTTGCCCTACAAGCAAAAGGGGCAGATTCCGT
913▶ V I V E F C K F G N L S T Y L R S K R N E F V P Y K T K G A R F R

3401 CAAGGGAAAGACTACGTTGGAGCAATCCCTGTGGATCTGAAACGGCGCTTGGACAGCATCACCAGTAGCCAGAGCTCAGCCAGCTCTGGATTTGTGGAG
947▶ Q G K D Y V G A I P V D L K R R L D S I T S S Q S S A S S G F V E

3501 AGAAGTCCCTCAGTGATGAGAAGAAGGAAGCTCCTGAAGATCTGTATAAGGACTTCTGACCTTGGAGCATCTCATCTGTTACAGCTTCCAAGTGGC
980▶ E K S L S D V E E E E A P E D L Y K D F L T L E H L I C Y S F Q V A

NruI (3622)

3601 TAAGGGCATGGAGTTCTTGGCATCGCGAAAGTGTATCCACAGGACCTGGCGGCACGAAATATCCTCTTATCGGAGAAGAAGCTGGTTAAAATCTGTGAC
1013▶ K G M E F L A S R K C I H R D L A A R N I L L S E K N V V K I C D

XmaI (3710)

3701 TTTGGCTTGGCCGGGATATTTATAAAGATCCAGATTATGTCAGAAAAGGAGATGCTCGCCTCCCTTTGAAATGGATGGCCCCAGAAACAATTTTGGACA
1047▶ F G L A R D I Y K D P D Y V R K G D A R L P L K W M A P E T I F D

SspI (3852)

3801 GAGTGTACACAATCCAGAGTGACGTCTGGTCTTTTGGTGTGTTGCTGTGGAAATATTTTCTTAGGTGCTTCTCCATATCCTGGGGTAAAGATTGATGA
1080▶ R V Y T I Q S D V W S F G V L L W E I F S L G A S P Y P G V K I D E

Bsp120I (3938)

3901 AGAATTTTGTAGGCGATTGAAAGAAGGAAGTAAAGTAGAATGAGGGCCCTGATTATACTACACCAGAAATGTACCAGACCATGCTGGACTGCTGGCACGGGGAG
1113▶ E F C R R L K E G T R M R A P D Y T T P E M Y Q T M L D C W H G E

4001 CCCAGTCAGAGACCAGTCTTTCAGAGTTGGTGAACATTTGGGAAATCTTGGCAAGTCTTGGCAAGTCTCAGCAGGATGGCAAAGACTACATTTGTTCTCCGA
1147▶ P S Q R P T F S E L V E H L G N L L Q A N A Q Q D G K D Y I V L P

4101 TATCAGAGACTTTGAGCATGGAAGAGGATTCTGGACTCTCTGCTACCTCCTGTTTCTGTATGGAGGAGGAGGAAGTATGTGACCCCAATTTCCA
1180▶ I S E T L S M E E D S G L S L P T S P V S C M E E E E V C D P K F H

NgoMIV (4253)

4201 TTATGACAACACAGCAGGAATCAGTCAGTATCTGCAGAACAGTAAAGCGAAAGAGCCGGCCTGTGAGTGTAAAAACATTTGAAGATATCCCGTTAGAAGAA
1213▶ Y D N T A G I S Q Y L Q N S K R K S R P V S V K T F E D I P L E E

SapI (4360)

4301 CCAGAAGTAAAAGTAATCCAGATGACAACAGACGGACAGTGGTATGGTTCTTGCCTCAGAAGAGCTGAAAACCTTTGGAAGACAGAACCAATTTATCTC
1247▶ P E V K V I P D D N Q T D S G M V L A S E E L K T L E D R T K L S

BspEI (4480)

4401 CATCTTTTGGTGAATGGTGGCCAGCAAAAGCAGGGAGTCTGTGGCATCTGAAGGCTCAAACAGACAAGCGGCTACCAGTCCGGATATCACTCCGATGA
1280▶ P S F G G M V P S K S R E S V A S E G S N Q T S G Y Q S G Y H S D D

AgeI (4567)

4501 CACAGACACCACCGTGTACTCCAGTGAGGAAGCAGAACCTTTAAAGCTGATAGAGATTGGAGTGCAAAACCGGTAGCACAGCCAGATTCTCCAGCCTGAC
1313▶ T D T T V Y S S E E A E L L K L I E I G V Q T G S T A Q I L Q P D

MscI (4681)

4601 TCGGGACCACACTGAGCTCTCCTCCTGTTTAAAAGGAAGCATCCACACCCCCAACTCCCGGACATCACATGAGAGCTAGCTGGCCAGACATGATAAGAT
1347▶ S G T T L S S P P V •

NheI (4675)

4701 ACATTGATGAGTTTGGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAG

HpaI (4813) MfeI (4824)

4801 CTGCAATAAACAAGTTAAACAACAACATTGCATTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAAACCTCTACAAA

4901 TGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTGAAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATC

SapI (5091)

5001 AGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAAGATATAGTGTATTTTCCCAAGTTTGAACATGCTCTTCA

SspI (5148)

SwaI (5162)

5101 TTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTATAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTTTT

5201 ATTAGGCAGAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTAAATAGAAATTTGGACAGCAA

5301 GAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAG
141▶ N R T Y K L P I L E E I T T K V L K G N M E I L V F

BstXI (5452)

5401 CAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAG
113▶ 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 H R V A

StuI (5587)

5501 CCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCCAGCAGACAGTACCCTGCCAATGTAGGCCTCAATGTG
80▶ 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 A E I H

5601 GACAGCAGAGATGATCTCCCGAGTCTGGTCTGATGGCCGCCCGACATGGTCTTGTCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCC
47▶ 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 A V E

BspHI (5737)

XmnI (5729)

AseI (5795)

5701 ACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTTCATGATGGCCCTCTATAGTGAGTCGATTATACTATGCCGATATACTATGCCGATGATTA
13▶ V L E L D Q Q S I N F T K M

5801 **TTGTCAA**AAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTT
←

5901 GCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTT**ACTAGT**CAAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGGAA
SpeI (5950) ←

6000 ATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGT
SnaBI (6078)

6100 AGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGA
NdeI (6183)

6200 TGTACTGCCAAGTGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACG

6300 TCAATGGGCGGGGTCGTTGGGCGGTGACCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCC**TGCAGGTTAA**TTAAGAACATGTGAGCAAAGGCC
SdaI (6361) Pacl (6369) BspLU11I (6379) ←

6398 AGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAG

6498 GTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTG

6598 TCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGC
ApaLI (6693)

6698 ACGAACCCCCGTTAGCCCGACCCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCAC

6798 TGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATC

6898 TGGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGCAAGC

6998 AGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACACTCACGTTAAGGGATTTT

EagI (7129)
Pacl (7109) SwaI (7118) NotI (7128)

7098 **GGTCATGGCTAGTTAATTAACATTTAAATC** AGCGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAA

7198 CATACGCTCTCCATCAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAAGTGCCAGAGTCCAGAACATTTCTCTATCGAA