



PstI (6)
SdaI (6) SspI (26)
1 CCTGCAGGAGCTTGAATAAAATGAATATTAGAAGCTGTTAGAATAAGAGAAAATGACAGAGGAAAACCTGAAAGGGAGAAGCTGAAAGTGGGAAATTCCTCT

PstI (163)
101 GAGGCAGAAAGGACCATCCCTTATAAATAGCACAGGCCATGAAGGAAGATCATTCTCACTGCAGCCTTTGACAGCCTTTGCCTCATCTTGCAGGTAGCAG

BspHI (222) NheI (260)
201 CCGACACCAGCTGGCTTCCATCATGAGCGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGACAGCAAATGGGTGCGGATCTGTA
1 M S G S H H H H H H G M A S M T G G Q Q M G R D L Y

Bsu36I (321)
Acc65I (316)
301 CGACGATGACGATAAGGTACCTAAGGATCAGCTGGAGTGTATCCCGTCTTTACAACGCTGACTGGGAAAACCTGGCGTTACCCAACCTAATCGC
26 D D D D K V P K D Q L G V D P V V L Q R R D W E N P G V T Q L N R

FspI (476)
401 CTTGCAGCACATCCCTTTCCAGCTGGCGTAATAGCGAAGAGGCCGACCGATCGCCCTTCCCAACAGTTGCCAGCCTGAATGGCGAATGGCGCT
60 L A A H P P F A S W R N S E E A R T D R P S Q Q L R S L N G E W R

Bsu36I (558)
501 TTGCTGGTTTCCGGACCAGAAGCGGTGCCGAAAAGCTGGCTGGAGTGCATCTTCTGAGGCCGATACTGTCGTGCTCCCTCAAACCTGGCAGATGCA
93 F A W F P A P E A V P E S W L E C D L P E A D T V V V P S N W Q M H
601 CGGTTACGATGCGCCATCTACCAACGTAACCTATCCATTACGTCATCCGCCGTTGTTCCACGGAGAATCCGACGGGTTGTTACTCGCTCACA
126 G Y D A P I Y T N V T Y P I T V N P P F V P T E N P T G C Y S L T
701 TTTAATGTTGATGAAAGCTGGCTACAGGAAGCCAGACGCGAATATTTTTGATGGCGTAACTCGCGTTTCATCTGTTGCAACGGGCGTGGGTG
160 F N V D E S W L Q E G Q T R I I F D G V N S A F H L W C N G R W V
801 GTTACGGCCAGGACAGTCTGTTGCGCTGTAATTTGACCTGAGCGCATTTTTACGCGCCGAGAAAACCGCCTCGCGGTGATGGTGGCTGCGTTGGAGTGA
193 G Y G Q D S R L P S E F D L S A F L R A G E N R L A V M V L R W S D

AatII (957)
901 CGGCAGTTATCTGGAAGATCAGGATATGTGGCGGATGAGCGGATTTTCCGTGACGTCCTGTTGCTGATAAACCGACTACACAATCAGCGATTTCCAT
226 G S Y L E D Q D M W R M S G I F R D V S L L H K P T T Q I S D F H
1001 GTTGCACCTCGCTTAATGATGATTTTCAGCCCGCTGTACTGGAGGCTGAAGTTCAGATGTGCGCGAGTTGCGTACTACCTACGGGTAACAGTTTCTT
260 V A T R F N D D F S R A V L E A E V Q M C G E L R D Y L R V T V S

ClaI (1158)
1101 TATGGCAGGTTGAAACGAGGTCGCCAGCGCACCGCCCTTTCGGCGGTGAAATTATCGATGAGCGTGGTGGTTATGCCGATCGCGTACACTACGCTC
293 L W Q G E T Q V A S G T A P F G G E I I D E R G G Y A D R V T L R L
1201 GAACCTCGAAAACCCGAACTGTGGAGCGCCGAAATCCCAATCTCTATCGTGGCGTGGTTGAACTGCACACCGCCGACGGCAGCTGATTGAAGCAGAA
326 N V E N P K L W S A E I P N L Y R A V V E L H T A D G T L I E A E
1301 GCCTGCGATGTCGGTTTCCGCGAGGTGCGGATTGAAATGGTCTGCTGCTGCTGAACGGCAAGCCGTTGCTGATTGAGGGCTTAACCGTACGAGCATC
360 A C D V G F R E V R I E N G L L L L N G K P L L I R G V N R H E H

EcoRV (1447)
1401 ATCCTCTGCATGGTCAGGTATCGATGAGCAGACGATGGTGCAGGATATCTGCTGATGAAGCAGAACAACCTTAACGCCGTGCGTGTTCGATTATCC
393 H P L H G Q V M D E Q T M V Q D I L L M K Q N N F N A V R C S H Y P

DraIII (1524) SspI (1564)
1501 GAACCATCCGCTGTGGTACACGCTGTGCGACCGCTACCGCCTGTATGTGGTGGATGAAGCCAATATTGAAACCCACGGCATGGTCCCAATGAATCGTCTG
426 N H P L W Y T L C D R Y G L Y V V D E A N I E T H G M V P M N R L

BsaBI (1660)
1601 ACCGATGATCCGCGTGGCTACCGGCGATGAGCGAAGCGTAACGCGAATGGTGCAGCGGATCGTAATCACCCGAGTGTGATCATCTGGTCCGCTGGGGA
460 T D D P R W L P A M S E R V T R M V Q R D R N H P S V I I W S L G
1701 ATGAATCAGGCCACGGCGCTAATCAGCAGCGCTGTATCGTGGATCAAACTCTGTCATCCTCCCGCCGGTGCAGTATGAAGCGCGGAGCGCGACAC
493 N E S G H G A N H D A L Y R W I K S V D P S R P V Q Y E G G G A D T

BssHII (1832) BbsI (1851)
1801 CACGGCCACCGATATTATTTGCCGATGTACGCGCGTGGATGAAGACCAGCCCTCCCGGCTGTGCCGAAATGGTCCATCAAAAATGGCTTTCGTA
526 T A T D I I C P M Y A R V D E D Q P F P A V P K W S I K K W L S L
1901 CCTGAGAGACGCGCCGCTGATCCTTTGCCAATACGCCACCGATGGTAAACAGTCTTGGCGTTCGCTAAATACTGGCAGCGCTTTCGTCAGTATC
560 P G E T P L I L C E Y A H A M G N S L G G F A K Y W Q A F R Q Y
2001 CCCGTTTACAGGGCGCTTCTGCTGGGACTGGTGGATCAGTCTGATTAATATGATGAAAACGGCAACCCGTTGCTGCGCTTACGGCGGTGATTTTGG
593 P R L Q G G F V W D W V D Q S L I K Y D E N G N P W S A Y G G D F G

Eco47III (2169)
2101 CGATACGCCGAACGATCGCCAGTCTGTATGAACGGTCTGGTCTTTGCCACCGCACGCCGATCCAGCGCTGACGGAAGCAAAACACCAGCAGCAGTTT
626 D T P N D R Q F C M N G L V F A D R T P H P A L T E A K H Q Q Q F

SacI (2274)
2201 TTCCAGTCCGTTTATCCGGGCAACCATCGAAGTACCAGCGAATACCTGTTCCGTATAGCGATAACGAGCTCCTGCACTGGATGGTGGCGTGGATG
660 F Q F R L S G Q T I E V T S E Y L F R H S D N E L L H W M V A L D
2301 GTAAGCCGCTGGCAAGCGTGAAGTGCCTCGATGTCGCTCCAAGGTAACAGTTGATTGAATGCCTGAATACCGCAGCCGAGCGCCGGGCA
693 G K P L A S G E V P L D V A P Q G K Q L I E L P E L P Q P E S A G Q
2401 ACTCTGGCTCACAGTACGCTAGTGAACCGAAGCGACCGCATGGTGAAGCGGGCACATCAGCGCTGGCAGCAGTGGCTGCTGGCGAAAACCTC
726 L W L T V R V V Q P N A T A W S E A G H I S A W Q Q W R L A E N L
2501 AGTGTGACGCTCCCGCCGCTCCACGCCATCCCGCATCTGACCACAGCGAAATGGATTTTTGCATCGAGCTGGGTAATAAGCGTTGGCAATTTAAC
760 S V T L P A A S H A I P H L T T S E M D F C I E L G N K R W Q F N
2601 GCCAGTCAGGCTTCTTCCACAGATGTGGATTGGCGATAAAAAACACTGCTGACGCGCTGCGCGATCAGTTACCCGTCACCGCTGGATAACGACAT
793 R Q S G F L S Q M W I G D K K Q L L T P L R D Q F T R A P L D N D I
2701 TGGCGTAAGTGAAGCGACCCGATGACCTAACGCTGGTGAAGCGTGAAGCGCGGGCCATTACCAGGCCGAAGCAGCGTTGTTGAGTGCAGC
826 G V S E A T R I D P N A W V E R W K A A G H Y Q A E A A L L Q C T
2801 CGACATACACTGCTGATGCTGATTACGACCGCTCAGCGTGGCAGCATCAGGGGAAAACCTTATTTATCAGCCGAAAACCTACCGGATTGATG
860 A D T L A D A V L I T T A H A W Q H Q G K T L F I S R K T Y R I D
2901 GTAGTGGTCAAAATGGCGATTACCGTTGATGTTGAAGTGGCGAGCGATACCCGATCCGGCGCGGATTGGCTGAACTGCCAGCTGGCGCAGGTAGCAGA
893 G S G Q M A I T V D V E V A S D T P H P A R I G L N C Q L A Q V A E

Bst1107I (3096)
BspLU11I (3093)
3001 GCGGGTAAACTGGCTCGGATTAGGGCCGAAGAAAACCTACCCAGCCGCTTACTGCCGCTGTTTGGCCGCTGGGATCTGCCATTGTGACAGCATGTAT
926 R V N W L G L G P Q E N Y P D R L T A A C F D R W D L P L S D M Y

BbsI (3103)
BsiWI (3104)
3101 ACCCCGTACGCTTCCCGAGCGAAAAACGGTCTGCGCTGCGGGACGCGCAATTGAATTATGGCCACACAGTGGCGGGGACTTCCAGTTCAACATCA
960▶ T P Y V F P S E N G L R C G T R E L N Y G P H Q W R G D F Q F N I

NdeI (3291)
3201 GCCGCTACAGTCAACAGCAACTGATGAAACAGCCATCGCCATCTGCTGCACGGGAAGAAGGCACATGGCTGAATATCGACGGTTTCCATATGGGGAT
993▶ S R Y S Q Q Q L M E T S H R H L L H A E E G T W L N I D G F H M G I
3301 TGGTGGCGACGACTCCTGGAGCCCGTCAGTATCGGCGGAATTACAGCTGAGCGCCGGTCGCTACCATACCAAGTTGGTCTGGTGTCAAAAATAAATCT
1026▶ G G D D S W S P S V S A E L Q L S A G R Y H Y Q L V W C Q K •

NheI (3413)
EcoRI (3407)
3401 AGTCGAGAATTCGCTAGCTGACATGATAAGATACATTGATGAGTTTGACAAACCACAAGTGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTG

MfeI (3587)
3501 TGATGCTATTGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAAACAAGTTAAACAACAACAAATTCATTCAAT

DraI (3675)
Swal (3678)
3601 TTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTAGATCCATTTAAATGTTAATTAAGTACAGCCATG
3701 ACCAAAAATCCCTTAACGTGAGTTTTCTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCCGTAATCT
3801 GCTGCTTCAAAACAAAAAACCCGCTACCAGCGGTGGTTTGTTCGCCGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGC
3901 GCAGATACAAATACTGTTCTTCTAGTGAGCGTAGTTAGGCCACCCTCAAGAACTCTGTAGCACCAGCTACATACCTCGCTCTGTAATCCTGTTA
4001 CCAGTGGCTGCTGCCAGTGGCGATAAGTGTCTTACCAGGTTGGACTCAAGACGATAGTTACCAGGATAAGGCGCAGCGGTCCGGCTGAACGGGGGGTT
4101 CGTGACACAGCCAGCTTGAGCGAAGCAGCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCAGCTTCCCGAAGGGAGAAAGGC
4201 GGACAGGTATCCGGTAAGCGGCAGGGTCCGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCTGGTATCTTTATAGTCCTGTCCGGTTTTCCG
4301 CACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGAAAAACGCCAGCAACGCGGCCTTTTTACGGTCTCGCCTTTTT

BspLU11I (4416)
AseI (4454)
4401 GCTGGCCTTTTGTCCACATGTTCTTAATTAATTTTTCAAAGTAGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGACTCACTAT

SfiI (4505)
MseI (4516)
4501 AGGAGGGCCATCATGGCCAAGTTGACCAGTGTGCCAGTGTCCAGTGTCCAGCCAGGGATGTGGCTGGAGCTGTTGAGTTCTGGACTGACAGGTTGGGGTTCT
▶ 1▶ M A K L T S A V P V L T A R D V A G A V E F W T D R L G F
4601 CCAGAGATTTTGTGGAGGATGACTTTGCAGGTGTGGTCAGAGATGATGTACCCCTGTTTCTCAGCAGTCCAGGACCAGGTGGTGCCTGACAACACCCT
30▶ S R D F V E D D F A G V V R D D V T L F I S A V Q D Q V V P D N T L
4701 GGCTTGGGTGTGGGTGAGAGGACTGGATGAGCTGTATGCTGAGTGGAGTGGAGTGGTCTCCACCACTTCAGGGATGCCAGTGGCCCTGCCATGACAGAG
63▶ A W V W V R G L D E L Y A E W S E V V S T N F R D A S G P A M T E

DraIII (4866)
4801 ATTGGAGAGCAGCCCTGGGGAGAGAGTTTGCCTGAGAGACCCAGCAGGCAACTGTGTGCACTTTGTGGCAGAGGAGCAGGACTGAGGATAAGAATTGA
97▶ I G E Q P W G R E F A L R D P A G N C V H F V A E E Q D • —

SfiI (4914)
EcoO109I (4914)
4901 GTTTCAGAAAAGGGGCCTGAGTGGCCCTTTTTTCAACTTAATTAA