



PvuI (7)
SgfI (6)
1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) Bsu36I (291)
201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCCTTACGCGCCCGCCCTACCTGAGGGC
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
NgoMI (441)
NaeI (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGTCTTTGTTTCGTTT

KasI (535) AgeI (552) SphI (560)
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTA CCTGAGATCACCGGTGAGCATGCGATATCTAAATATAAAAGAGGACTGCAATGCCAT
601 GGCTTTCTGTGCTAAAATGAGGAGCTCCAAGAAGACTGAGTGAACCTGGAGGCCCTGAGCCAGGGGTGGAAGTGATCTTCTATCTGTCGGACAGGGAG
13▶ A F C A K M R S S K K T E V N L E A P E P G V E V I F Y L S D R E
1▶ M Q Y L N I K E D C N A M

SphI (760)
701 CCCCTCCGGTGGGCGAGTACACAGCAGAGGAAGTGTGCATCAGGGCTGCACAGGCATGCCGTATCTCTCTCTTTGTGACAACTCTTTGCC
47▶ P L R L G S G E Y T A E E L C I R A A Q A C R I S P L C H N L F A
801 TGTATGACGAGAACCAAGCTCTGGTATGCTCCAATCGCACCATCACCGTTGATGACAAGATGTCCCTCCGCTCCACTACCGGATGAGGTTCTATTT
80▶ L Y D E N T K L W Y A P N R T I T V D D K M S L R L H Y R M R F Y F
901 CACCAATGGCATGGAACCAACGACAATGAGCAGTCAAGTGTGGCCTCATTCTCCAAGAAGCAGAAAAATGGCTACGAGAAAAAAGATTCCAGATGCA
113▶ T N W H G T N D N E Q S V W R H S P K K Q K N G Y E K K K I P D A
1001 ACCCTCTCTTGTGATGCCAGTCACTGGAGTATCTGTTTGTGCTCAGGACAGTATGATTTGGTAAATGCCTGGCTCCTATTGAGACCCCAAGACCGAGC
147▶ T P L L D A S S L E Y L F A Q G Q Y D L V K C L A P I R D P K T E
1101 AGGATGGACATGATATTGAGAACGAGTGTCTAGGGATGGCTGTCTGGCCATCTCACACTATGCCATGATGAAGAAGTGCAGTTGCCAGAAGTCCCAA
180▶ Q D G H D I E N E C L G M A V L A I S H Y A M M K K M Q L P E L P K
1201 GGACATCAGCTACAAGCGATATATTCAGAAACATTGAATAAGTCCATCAGACAGAGGAACCTTCTCACCAGGATGCGGATAAATAATGTTTTCAAGGAT
213▶ D I S Y K R Y I P E T L N K S I R Q R N L L T R M R I N N V F K D
1301 TTCCTAAAGGAATTTAAACAACAAGACATTTGTGACAGCAGCTGTCCACGATGACCTGAAGGTGAAATACTTGGCTACCTTGGAACCTTTGACAAAAAC
247▶ F L K E F N N K T I C D S S V S T H D L K V K Y L A T L E T L T K

BstBI (1470)
Bsp119I (1470)
AsuII (1470) SspI (1413) Psp1406I (1486)
1401 ATTACGGTCTGAAATATTTGAGACTTCCATGTTACTGATTTTCATCAGAAAATGAGATGAATTGGTTTCATTGCAATGACGGTGGAAACGTTCTCTACTA
280▶ H Y G A E I F E T S M L L I S S E N E M N W F H S N D G G N V L Y Y
1501 CGAAGTGTGGTACTGGGAATCTTGAATCCAGTGGAGGCATAAACCAATGTTGTTTCTGTTGAAAAGGAAAAAATAAAGTGAAGCGGAAAAAAGT
313▶ E V M V T G N L G I Q W R H K P N V V S V E K E K N K L K R K K L
1601 GAAAAAACAAGAAGGATGAGGAGAAAAACAAGTCCGGGAAGAGTGGAAACATTTTCTTACTTCCCTGAAATCAGTACATTTGTAATAAAGGAGT
347▶ E N K H K K D E E K N K I R E E W N N F S Y F P E I T H I V I K E
1701 CTGTGGTCAAGCAACAAGCAGGACAACAAGAAAATGAACTGAAGCTCTTCCACGAGGAGGCTTGTCTTTGTGTCCTGGTAGATGGTACTT
380▶ S V V S I N K Q D N K K M E L K L S S H E E A L S F V S L V D G Y F

BsiBI (1813)
BsaBI (1813) BsrGI (1894)
1801 CCGGCTCACAGCAGATGCCATCATTACCTCTGCACCGACGTGGCCCCCGTTGATCGTCCACAACATACAGAATGGCTGTCATGGTCCAATCTGTACA
413▶ R L T A D A H H Y L C T D V A P P L I V H N I Q N G C H G P I C T
1901 GAATACGCCATCAATAAATTGCGGCAAGAAGGAAGCGAGGAGGGGATGTACGTGCTGAGGTGGAGCTGCACCGACTTTGACAACATCCTCATGCCGTC
447▶ E Y A I N K L R Q E G S E E G M Y V L R W S C T D F D N I L M T V

RsrII (2096)
2001 CCTGCTTTGAGAAGTCTGAGCAGGTGCAGGGTCCAGAAAGCAGTTCAAGAACTTTCAGATCGAGGTGCAGAAGGGCCGTACAGTCTGCACGGTTCGGA
480▶ T C F E K S E Q V Q G A Q K Q F K N F Q I E V Q K G R Y S L H G S D

FspI (2153)
BstAPI (2145)
2101 CCGCAGCTTCCCCAGCTTGGGAGACCTCATGAGCCACCTCAAGAAGCAGATCCTGCGCACGGATAACATCAGCTTCATGCTAAAACGCTGCTGCCAGCCC
513▶ R S F P S L G D L M S H L K K Q I L R T D N I S F M L K R C C Q P

XcmI (2216) PvuI (2286) BamHI (2291)
2201 AAGCCCCGAGAAATCTCCAACCTGCTGGTGGCTACTAAGAAAAGCCAGGAGTGGCAGCCCGTCTACCCATGAGCCAGCTGAGTTTCGATCGGATCCTCA
547▶ K P R E I S N L L V A T K K A Q E W Q P V Y P M S Q L S F D R I L

SandI (2359)
2301 AGAAGATCTGGTGCAGGGCAGCACCTTTGGGAGAGGCACGAGAACACACATCTATTCTGGGACCTGATGGATTACAAGGATGACGAAGGAACCTTCTGA
580▶ K K D L V Q G E H L G R G T R T H I Y S G T L M D Y K D D E G T S E
2401 AGAGAAGAAGATAAAAGTATCCTCAAAGTCTTAGACCCAGCCACAGGGATATTTCCCTGGCCTTCTCGAGGCAGCCAGCATGATGAGACAGGTCTCC
613▶ E K K I K V I L K V L D P S H R D I S L A F F E A A S M M R Q V S

DraIII (2506)
2501 CACAAACACATCGTACCTCTATGGCTGTGTGTCGCGACGTGGAGAATATCATGGTGAAGAGTTTGTGGAAGGGGTCCTCTGGATCTTCTCATGC
647▶ H K H I V Y L Y G V C V R D V E N I M V E E F V E G G P L D L F M

SexAI (2687)
Tth111I (2686)

2601 ACCGAAAAGCGATGTCCTTACCACACCATGGAATTCAAAGTTGCCAAACAGCTGGCCAGTGCCTGAGCTACTTGGAGGATAAAGACCTGGTCCATGG
680▶ H R K S D V L T T P W K F K V A K Q L A S A L S Y L E D K D L V H G
2701 AAATGTGTGTAATAAAACCTCCTCCTGGCCCGTGGAGGCATCGACAGTGTGAGTGTGGCCATTATCAAGCTCAGTGACCCCGCATCCCCATTACGGTG
713▶ N V C T K N L L L A R E G I D S E C G P F I K L S D P G I P I T V

Ppu10I (2814)
NsiI (2814)

2801 CTGTCTAGGCAAGAATGCATTGAACGAATCCCATTGATTGCTCCTGAGTGTGTTGAGGACTCCAAGAACCCTGAGTGTGGCTGCTGACAAGTGGAGCTTTG
747▶ L S R Q E C I E R I P W I A P E C V E D S K N L S V A A D K W S F
2901 GAACCACGCTCTGGGAAATCTGCTACAATGGCGAGATCCCCTTGAAGACAAGACGCTGATTGAGAAAGAGAGATTCTATGAAAGCCGGTGCAGGCCAGT
780▶ G T T L W E I C Y N G E I P L K D K T L I E K E R F Y E S R C R P V

VspI (3094)
AseI (3094)

3001 GACACCATCATGTAAGGAGCTGGCTGACCTCATGACCCGCTGCATGAACTATGACCCCAATCAGAGGCCCTTTCTCCGAGCCATCATGAGAGACATTAAT
813▶ T P S C K E L A D L M T R C M N Y D P N Q R P F F R A I M R D I N

Eco47III (3176)

AfeI (3176) BamHI (3191)

3101 AAGCTTGAAGAGCAGAATCCAGATATTGTTTCAGAAAAAAACAGCAACTGAAGTGGACCCACACATTTTGAAGCGCTTCTTAAAGAGGATCCCGT
847▶ K L E E Q N P D I V S E K K P A T E V D P T H F E K R F L K R I R

PstI (3234)

3201 ACTTGGGAGAGGGCCACTTTGGGAAGTTGAGCTCTGCAGGTATGACCCCGAAGGGGACAATACAGGGGAGCAGGTGGCTGTTAAATCTCTGAAGCCTGA
880▶ D L G E G H F G K V E L C R Y D P E G D N T G E Q V A V K S L K P E

BstEII (3307)

BglII (3341)

3301 GAGTGGAGGTAACACATAGCTGATCTGAAAAAGGAAATCGAGATCTTAAGGAACCTCTATCATGAGAACATTGTGAAGTACAAAGGAATCTGCACAGAA
913▶ S G G N H I A D L K K E I E I L R N L Y H E N I V K Y K G I C T E

3401 GACGGAGGAAATGGTATTAAGCTCATCATGGAATTTCTGCCTTGGGAAGCCTTAAGGAATATCTTCAAAGAATAAGAACAAAATAAACCTCAAACAGC
947▶ D G G N G I K L I M E F L P S G S L K E Y L P K N K N K I N L K Q

3501 AGCTAAAATATGCCGTTACAGATTTGTAAGGGGATGGACTATTTGGGTTCTCGGCAATACGTTACCCGGGACTTGGCAGCAAGAAATGCTCTTGTGAGAG
980▶ Q L K Y A V Q I C K G M D Y L G S R Q Y V H R D L A A R N V L V E S

XcmI (3606)

DraIII (3604)

3601 TGAACACCAAGTGAATAATGGAGACTTCGGTTTAAACCAAAGCAATTGAAACCGATAAGGAGTATTACACCGTCAAGGATGACCGGGACAGCCCTGTGTT
1013▶ E H Q V K I G D F G L T K A I E T D K E Y Y T V K D D R D S P V F

BstAPI (3774)

3701 TGGTATGCTCCAGAATGTTAATGCAATCTAAATTTTATATTGCCTCTGACGCTGGTCTTTTGGAGTCACTCTGCATGAGCTGCTGACTTACTGTGATT
1047▶ W Y A P E C L M Q S K F Y I A S D V W S F G V T L H E L L T Y C D

3801 CAGATTCTAGTCCATGGCTTTGTTCTGAAAATGATAGGCCAACCCATGGCCAGATGACAGTCACAAGACTTGTGAATACGTTAAAAGAAGGAAAACG
1080▶ S D S S P M A L F L K M I G P T H G Q M T V T R L V N T L K E G K R

EcoRI (3958)

3901 CCTGCCGTGCCACCTAACTGTCCAGATGAGGTTTATCAACTTATGAGGAAATGCTGGGAATCCAACCATCCAATCGGACAAGCTTTTCAAGACCTTATT
1113▶ L P C P P N C P D E V Y Q L M R K C W E F Q P S N R T S F Q N L I

NheI (4036)

4001 GAAGGATTTGAAGCACTTTTAAAATAAGAAGCATGAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGT
1147▶ E G F E A L L K •

HpaI (4174)

4101 GAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAATTGCATTCAATT

EcoRI (4270)

4201 TATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTT

4301 AACCTCCAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTC

4401 ACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTTGAAGTAGCTCTTCAATTTCTTTATGTTTAAATGCACTGACCTCCACATTCCCT

SspI (4509)

Swal (4523)

4501 TTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATAT

4601 CCCCCAGTTTAGTGTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTCGGTGTAAGTCTGAG
1414 • N R T Y K L

4701 GGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATG
134▶ P I L L E E I T T 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

4801 CCACAGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAG
100▶ G C P S V V R I S R D V E D S Y P H R V A V I T D F D K C T G G N S V A

4901 CAGACCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGC
67▶ S G I A I A E A 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

5001 CGCCCCGACATGGTGTCTTGTCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTC
34▶ A G V H H K N D 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

VspI (5156)

AseI (5156)

5101 ATGATGGCCCTCTATAGTGTGAGTCGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTG
04M ▶

5200 ACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTACGACATTTTGAAG

5300 TCCCGTTGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATG
SpeI (5311) ←

5399 TACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAGGTCATGTACTGGGCATAATGCC
SnaBI (5439)
Eco105I (5439)

5499 AGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCAC
NdeI (5544)

5599 CCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGCGGTGAGCCAGGCGGG

5699 CCATTTACCGTAAGTTATGTAACGCC TGCAGGTT AA TTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTG
PstI (5723) SdaI (5722) PaeI (5730) PciI (5740) BspLU11I (5740) ←

5797 GCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGT

5897 TTCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCGACCCTGCCGTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCA

5997 TAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCGACCGCTGCGCCTTATCC
ApaLI (6054)

6097 GGTAATATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGGT

6197 CTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGT

6297 TGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGAT

6397 CCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAACTCACGTAAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCC
PaeI (6470) SmaI (6479) EagI (6490)
NotI (6489)

6496 GCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAA
6596 ACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA