



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA **MfeI (82)**

101 GAGAAAGTGGCGGGGTAACGGAAAGTGTGTCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) **HindIII (245)** **Bsu36I (291)**
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCCTACCTGAGGCC

301 GCCATCCACGCGGTTGAGTCGCGTTTCTGCCGCCTCCCGCTGTGGTGCCTCCTGAAGTCGCTCCGCCGTCTAGGTAAGTTAAAGCTCAGGTGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGTCTTTGTTTCGTTT

NcoI (560)
BstEII (555) **AgeI (552)** **Bsp120I (592)**
501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGGCTACCTGAGATCACCGGTACCATTGGAGCGGACGGCGACCAGGCAGGGCACGGGCCCGG
1▶ M E R D G D Q A G H G P R

XhoI (628) **SacII (641)** **NcoI (662)** **AvrII (669)**
601 GCACGGGTGACGGGAAACGGTCCGAGCTCGAGTCACCAGCCGGGCTTCGCTGCTCGCGCCCATGGACCTAGGAGAGAGCCGCTGGAGAAGCGGGAG
13▶ H G S A G N G R E L E S P A A A S L L A P M D L G E E P L E K A E
701 CGTCCCCGCCCTGCCAAAGACCCCAACACCTACAAAGTGTGTGCTGGTTTTGTGCTGCTAACAACAATTCTGGTTGTATATTTGGGTTGA
47▶ R A R P A K D P N T Y K V L S L V L S V C V L T T I L G C I F G L
801 AACCAAGCTGTGCCAAAGAAGTAAAGTTGCAAAGGCCGCTGCTTGAAGGACGTTTCAAGCACTGCTGCTGTGATGCTGCTGTGTCAGCCTTGGAA
80▶ K P S C A K E V K S C K G R C F E R T F S N C R C D A A C V S L G N

XcmI (934)
901 CTGCTGTCTGGATTTCCAGGAGACTGTGTGGAACCAACATATATGGACTTGAACAAATTCAAGTGGTGAAGAAGGGCTGTCCAGGTTCTGTGTG
113▶ C C L D F Q E T C V E P T H I W T C N K F R C G E K R L S R F V C

ClaI (1099)
1001 TCCTGTGACAGACTGCAAAACCCACAATGACTGTTGCATCAACTACAGTTCTGTGTGCCAAGATAAGAAGATTGGGTAGAAGAAACATGTGAAAGCA
147▶ S C A D D C K T H N D C C I N Y S S V C Q D K K S W V E E T C E S
1101 TCGATACCCCGAGTGTCCAGCAGAGTTTGAATCACCCCTACTCTCTGTTTTCTTTGGATGGATTGAGAGTGAATTTGCACACCTGGGGTGGACT
180▶ I D T P E C P A E F E S P P T L L F S L D G F R A E Y L H T W G G L
1201 TCTTCTGTCAATAGCAAGCTGAAAAACTGTGGAACATACACTAAAAACATGAGGCCATGTACCCCTACCAAGACGTTTCCAATCATTACAGCATCGTT
213▶ L P V I S K L K N C G T Y T K N M R P M Y P T K T F P N H Y S I V

NcoI (1320)
BspEI (1311)
1301 ACAGGACTTATCCGGAATCCCATGGCATAATTGATAACAAGATGTATGATCCCAAATGAATGCTTCTTTCTCGCTTAAAAGTAAAGAGAAATTC AAC
247▶ T G L Y P E S H G I I D N K M Y D P K M N A S F S L K S K E K F N

BstEII (1428) **MscI (1471)**
1401 CTTTGTGGTACAAAGGACAGCCGATTTGGGTGACCGCTAATCATCAGGAGGTCAAGTCCGGCACATACTTTTGGCCAGGATCAGATGTGGAGATTGACGG
280▶ P L W Y K G Q P I W V T A N H Q E V K S G T Y F W P G S D V E I D G

EcoRV (1510)
1501 GATTCTACCAGATATCTACAAAGTGTATAATGGGTGAGTACCATTTGAAGAAAGGATTTTAGCTGTTCTTGTGAGTGGCTACAGCTTCTAGCCATGAAAGA
313▶ I L P D I Y K V Y N G S V P F E E R I L A V L E W L Q L P S H E R
1601 CCACACTTTTACACTCTGTATTTAGAAGAACCAGATTCTTCAAGGCATTACATGGACAGTCAAGCAGGAGTCAAGGCCTTGCAGAAGGTTGACC
347▶ P H F Y T L Y L E E P D S S G H S H G P V S S E V I K A L Q K V D

SphI (1709)
1701 GCCTGGTCCGATGCTGTGAGGACCGCTGAAGGACCTGGGCTTGATAAATGCCTGAACCTCATCTCATTTCAGATCACGGCATGGAACAAGGCAGCTG
380▶ R L V G M L M D G L K D L G L D K C L N L I L I S D H G M E Q G S C

ScaI (1823)
1801 TAAGAAGTATGTGTACCTGAATAAGTACTTGGGGATGTGAACAATGTGAAAGTTGTGTATGGACCTGCTGCTCGGTTAAGACCCACTGATGTTCCAGAG
413▶ K K Y V Y L N K Y L G D V N N V K V V Y G P A A R L R P T D V P E

XmnI (1980) **Eco47III (1997)**
1901 ACATACTATTCATTAACATGAAGCCCTTCAAAAAATCTTTCTTGGCCGGAGCCAAACCAGCATTTCCGGCCTTACCTGAAACCCCTTACCCAAGC
447▶ T Y Y S F N Y E A L A K N L S C R E P N Q H F R P Y L K P F L P K
2001 GCTTACACTTGCATAAAGTGACCGGATTGAGCCCTGACCTTCTACCTGACCCCTCAGTGGCAACTTGCCTTGAATCCATCAGAGAGGAAATATTGTTG
480▶ R L H F A K S D R I E P L T F Y L D P Q W Q L A L N P S E R K Y C G
2101 AAGTGGATTTGATGGCTGACAACTTGTTCACAAATGCAAGCTCTTTTATTGGCTATGGACCTGCCTTCAAGCATGGTGTGAAAGTTGACTCCTTT
513▶ S G F H G S D N L F S N M Q A L F I G Y G P A F K H G A E V D S F
2201 GAAATATTGAAGTCTATAACTAATGTGTGACTTACTGGTTTGTATCCAGCTCCCAATAATGGAAGTCATGGCAGCCTCAACCACCTTCTAAAGAAAC
547▶ E N I E V Y N L M C D L L G L I P A P N N G S H G S L N H L L K K
2301 CCATTTACAACCAAGTCAATCCCAAGAGAGGGCTTCTGTCTCAGTGTCCAATCAAAATCAACATCAATGACCTCGGCTGCACATGTGACCTTGGAT
580▶ P I Y N P S H P K E E G F L S Q C P I K S T S N D L G C T C D P W I

NdeI (2461)
2401 TGTGCCAATCAAGGACTTTGAGAAACAGCTTAATCTGACCACAGAAGATGATGACATTTACCATATGACTGTACCTACGGGAGGCCCGGATTCTACTG
613▶ V P I K D F E K Q L N L T T E D D D I Y H M T V P Y G R P R I L L

Bsu36I (2592)
2501 AAGCAGCACCAGCTGTCTACTCCAGCAGCAACAGTTCTTGACCGGATACAGCCTGGACCTCCTCATGCCCCTCTGGCGTCTTACACCTTCTCAGGA
647▶ K Q H H V C L L Q Q Q Q F L T G Y S L D L L M P L W A S Y T F L R

ApaI (2668)
2601 ATGACCAATTCTCCAGAGATGACTTTTCCAAGTGTGTACCAGGATCTTCGATTCCCTTAGTCCCGTGCACAAATGTTCTATTACAAAAGCAATTC
680▶ N D Q F S R D D F S N C L Y Q D L R I P L S P V H K C S Y Y K S N S

Eco47III (2766)
2701 TAAACTAAGTTATGGGTTCTCACGCCACCGAGACTAAATAGAGTTTCAAATCACATATACTCTGAAGCGCTGTACATCTAATATAGTGCCAATGTAT
713▶ K L S Y G F L T P R L N R V S N H I Y S E A L L T S N I V P M Y

BspHI (2829)
2801 CAGAGTTTCAAGTTATATGGCACTACCTTCATGACACCTACTGCAAAGGTATGCCCATGAAAGGAACGGCATCAATGTTGTCAGCGGTCCCCTGTTTG
747▶ Q S F Q V I W H Y L H D T L L Q R Y A H E R N G I N V V S G P V F
2901 ACTTTGATTATGACGGACGCTATGATTCCTTAGAGATCCTGAAACAAAACAGCAGAGTCATCCGAAGCCAGGAAATCTGATTCGACTCACTTCTTCAT
780▶ D F D Y D G R Y D S L E I L K Q N S R V I R S Q E I L I P T H F F I
3001 TGTGCTCACAGCTGCAAGCAGCTGTCTGAGACTCCCTTGGAGTGCTCCGCTTAGAGTCCTCAGCTTACATACTGCCTCATAGGCTGATAACATTGAG
813▶ V L T S C K Q L S E T P L E C S A L E S S A Y I L P H R P D N I E
3101 AGCTGTACGCATGGGAAGCGGAGTCTTCATGGGTCGAGGAGTTGCTGACCTTGACAGAGCTCGGGTCACAGACGTGGAACCTCATCTGGCCTCAGCT
847▶ S C T H G K R E S S W V E E L L T L H R A R V T D V E L I T G L S

MscI (3294)
PshAI (3212) NheI (3288)
3201 TCTACCAGGACCGACAAGAGTCAGTTTCAGAAGTCTGAGGTTGAAAACACATTTGCCAATCTTCAGCCAAGAAGACTGATTGTTTTTGTAGCTGGCCA
880▶ F Y Q D R Q E S V S E L L R L K T H L P I F S Q E D •
3301 GACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTG

HpaI (3426) MfeI (3437)
3401 TAACCATTATAAGCTGCAATAAAACAAGTTAAACAACAACATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTA

EcoRI (3522)
3501 AAACCTCTACAAATGTGGTATGGAATTCAAAAACAGCATAGCAAACTTTAACCTCCTCAAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATA
3601 AGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGA

SapI (3704) SmaI (3775)
3701 ACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACATTCCTTTTATGAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAA
3801 ATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGA
3901 AATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCCTC
141▶ N R T Y K L P I L E E I T T K V L K G N M E

BstXI (4065)
4001 AATGAGCACAAAGCAGTCAGGAGCATGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAG
118▶ 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
4101 GGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCCAGCAGACAGTGCCTGCAATGT
84▶ 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 C V T V R G I Y
4201 AGGCCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTGGTCTGATGGCCGCCCCGACATGGTGCTTGTGCTCCTATAGAGCATGGTATCTTCTC
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

XmnI (4342)
4301 AGTGGCGACCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTA
18▶ T A V E V L E L D Q Q S I N F T K M

AseI (4408)
4401 TGCCGATGATTAATTGTCAAACACAGCGTGATGGCGTCTCCAGCTTATCTGACGGTTCACATAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGC

SpeI (4563)
4501 CTACCGCCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCATTGACGTCAATGGGGT

SnaBI (4691)
4601 GGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGAT

NdeI (4796)
4701 GTACTGCCAAGTAGGAAAGTCCCATAAGTTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCAATTGACGTCAATAGGGGGCGTACTTGGCATA
4801 TGATACACTTGTACTGCAAGTGGGCGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACG

Pacl (4982)
PstI (4975) SdaI (4974)
4901 TCATTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGA
5001 GCAAAAGGCCAGCAAAAGCCAGGAACCGTAAAAAGCCGCTTGGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCT
5101 CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCTGCCGCTTAC
5201 CGGATACCTGTCCGCTTTCTCCCTTCCGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCGCTCCAAGCTG

ApaLI (5306)
5301 GGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCTTATCCGGTAACATATCGTCTTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGG
5401 CAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGT
5501 ATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCCGAAAAAGAGTTGGTAGCTCTTATCCGGCAAAACAAACCCGCTGGTAGCGGTGTTTTTTT

5601 GTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACACGTT

EagI (5742)

PacI (5722) SwaI (5731) NotI (5741)

5701 AAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTGTGTGAA

5801 TCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCT

5901 ATCGAA