



PvuI (7)
SgfI (6)
1 GGATCTGCATCGCTCCGGTCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGTGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC

Psp1406I (203) **HindIII (245)** **Bsu36I (291)**
201 GTGAACGTTCTTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACCGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCGGTCTTCCCGCTCCCGCTGTGGTGCCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

AgeI (552) **BspHI (568)**
501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTAGGAGGGCCATCATGAGACAGACTTTCCTTGTATCTACTTT
1 M R Q T L P C I Y F

ApaLI (650)
DraIII (644)
601 TGGGGGGCCTTTTGGCCTTTGGGATGCTGTGTCATCTCCACCACCAAGTGCAGTGTAGCCATGAAGTGTGCTGACTGCAGCCACCTGAAGTTGACTC
11 W G G L L P F G M L C A S S T T K C T V S H E V A D C S H L K L T
701 AGGTACCCGATGATCTACCCACAACATAACAGTGTGAACCTTACCATAATCAACTCAGAAGATTACCAGCCGCAACTTCAAGGTATAGCCAGCT
44 Q V P D D L P T N I T V L N L T H N Q L R R L P A A N F T R Y S Q L

XcmI (826)
801 AACTAGCTTGGATGTAGGATTTAACACCATCTCAAACTGGAGCCAGAATTGTCCAGAACTTCCCATGTTAAAAGTTTTGAACTCCAGCACAATGAG
77 T S L D V G F N T I S K L E P E L C Q K L P M L K V L N L Q H N E
901 CTATCTCAACTTTCTGATAAAACCTTTGCCTTCTGCACGAATTTGACTGAACTCCATCTCATGTTCAACTCAATCCAGAAAATTAATAATCCCTTTG
111 L S Q L S D K T F A F C T N L T E L H L M S N S I Q K I K N N P F

BglII (1025)
1001 TCAAGCAGAAGAATTTAATCACATTAGATCTGTCTCATAATGGCTTGTATCTACAAAATTAGGAACTCAGGTTGAGTGGAAAATCTCCAAGAGCTTCT
144 V K Q K N L I T L D L S H N G L S S T K L G T Q V Q L E N L Q E L L

Eco47III (1121) **EcoRV (1143)**
1101 ATTATCAAACAATAAAATTCAGCGCTAAAAAGTGAAGAACTGGATATCTTTGCCAATTCATCTTTAAAAAATTAGAGTTGTCATCGAATCAAATTTAAA
177 L S N N K I Q A L K S E E L D I F A N S S L K K L E L S S N Q I K

SmaI (1268)
1201 GAGTTTTCTCCAGGTGTTTTACGCAATTGGAAGATTATTTGGCCTCTTTCTGAACAATGTCCAGCTGGTCCCAGCCTTACAGAGAAGCTATGTTGG
211 E F S P G C F H A I G R L F G L F L N N V Q L G P S L T E K L C L
1301 AATTAGCAAACACAAGCATTCCGAATCTGTCTCTGAGTAACAGCCAGCTGTCCACCACCAGCAATACAACCTTTCTGGGACTAAAAGTGACAAATCTCAC
244 E L A N T S I R N L S L S N S Q L S T T S N T T F L G L K W T N L T
1401 TATGCTCGATCTTCTACAACAACCTAAATGTGGTGGTAACGATTCCTTTGCTTGGCTTCCACAACCTAGAATATTTCTCTAGAGTATAATAATATA
277 M L D L S Y N N L N V V G N D S F A W L P Q L E Y F F L E Y N N I
1501 CAGCATTTGTTTTCTACTCTTTGCACGGGCTTTTCAATGTGAGGTACCTGAATTTGAAACGGTCTTTTACTAAACAAAGTATTTCCCTTGCCTCACTCC
311 Q H L F S H S L H G L F N V R Y L N L K R S F T K Q S I S L A S L
1601 CCAAGATTGATGATTTTTCTTTTCAGTGGCTAAAATGTTGGAGCACCCTTAACATGGAAGATAATGATATCCAGGCATAAAAAGCAATATGTTCCAGG
344 P K I D D F S F Q W L K C L E H L N M E D N D I P G I K S N M F T G
1701 ATTGATAAACCTGAAATACTTAAGTCTATCCAACCTCTTTACAAGTTTGCAACTTTGACAAATGAAACATTTGTATCACTTGTCTATTCTCCCTTACAC
377 L I N L K Y L S L S N S F T S L R T L T N E T F V S L A H S P L H

ScaI (1871)
1801 ATACTCAACCTAACCAAGAATAAAATCTCAAAAATAGAGAGTGATGCTTTCTCTTGGTTGGGCCACCTAGAAGTACTTGACCTGGGCCTTAATGAAATTG
411 I L N L T K N K I S K I E S D A F S W L G H L E V L D L G L N E I

XbaI (1931)
1901 GGCAAGAACTCACAGGCCAGGAATGGAGAGGTCTAGAAAAATTTTTGAAATCTATCTTCTCACAACAAGTACTGCAGCTGACTAGGAACCTCTTTGC
444 G Q E L T G Q E W R G L E N I F E I Y L S Y N K Y L Q L T R N S F A
2001 CTTGGTCCCAAGCCTTCAACGACTGATGCTCCGAAGGGTGGCCCTTAAAAATGGGATAGCTCTCTTACCATTCCAGCCTCTTCTGAACTTGACCATT
477 L V P S L Q R L M L R R V A L K N V D S S P S P F Q P L R N L T I

BspLU11I (2140)
2101 CTGGATCTAAGCAACAACAACATAGCCAACATAAATGATGACATGTTGGAGGGTCTTGGAGAACTAGAATTTCTCGATTTGCAGCATAACAACCTTAGCAC
511 L D L S N N N I A N I N D D M L E G L E K L E I L D L Q H N N L A
2201 GGCTCTGAAACACGCAAAACCTGGTGGTCCCATTTATTTCCATAAAGGCTGTCTCACTCCACATCCTTAACTTGGAGTCCAACGGCTTTGACGAGAT
544 R L W K H A N P G G P I Y F L K G L S H L H I L N L E S N G F D E I

BbsI (2310) **ClaI (2340)**
2301 CCCAGTTGAGGCTTCAAGGATTTATTTGAACTAAAGATCATCGATTTAGGATTGAATAATTTAAACACACTTCCAGCATCTGTCTTTAATAATCAGGTG
577 P V E V F K D L F E L K I I D L G L N N L N T L P A S V F N N Q V

XmnI (2450)
2401 TCTCTAAAGTCATTGAACCTTCAAGAAGTCTCATAACATCCGTTGAGAAGAAGGTTTTCCGGCCAGCTTTCAGGAACCTGACTGAGTTAGATATGCGCT
611 S L K S L N L Q K N L I T S V E K K V F G P A F R N L T E L D M R

BbrPI (2516)
2501 TTAATCCCTTTGATTGCACGTGTGAAAGTATTGCTGGTTTGTAAATTGGATTAACGAGACCCATACCAACATCCCTGAGCTGTCAAGCCACTACCTTTG
644 F N P F D C T C E S I A W F V N W I N E T H T N I P E L S S H Y L C

BspHI [m] (2686)

2601 CAACACTCCACCTCACTATCATGGGTTCCAGTGAGACTTTTTGATACATCATCTTGCAAAGACAGTGCCCCCTTTGAACTCTTTTTCATGATCAATACC
677▶ N T P P H Y H G F P V R L F D T S S C K D S A P F E L F F M I N T

BsaBI (2711)

EcoRV (2756)

2701 AGTATCCTGTTGATTTTTATCTTTATTGTACTTCTCATCCACTTTGAGGGCTGGAGGATATCTTTTATTGGAATGTTTCAGTACATCGAGTTCTTGGTT
711▶ S I L L I F I F I V L L I H F E G W R I S F Y W N V S V H R V L G

MscI (2841)

NheI (2835)

2801 TCAAAGAAATAGACAGACAGACAGAACAGTTTTAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTG
744▶ F K E I D R Q T E Q F •

HpaI (2973)

2901 AAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAAACAAGTTAAACAACAACAATTGCATTCATTTT

EcoRI (3069)

3001 ATGTTTCAGGTTCAAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAAACTTTA

3101 ACCTCAAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTGCAGCCTCA

SapI (3251)

3201 CCTTCTTTCATGGAGTTAAGATATAGTGATTTTTCCAAGGTTTGAAGTACTGCTTCTTCTTTATGTTTTAAATGCAGTACCTCCACATTCCCTT

SwaI (3322)

3301 TTTAGTAAAATATTCAGAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATC

3401 CCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGGAGG

141◀ • N R T Y K L

SacI (3583)

3501 GGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGC

133◀ P I 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 G

BstXI (3612)

3601 CACAGGGGCTGACCACCTGATGGATCTGTCCACTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGC

100◀ 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 Q G N S V A

StuI (3747)

3701 AGACCCAAATGGCAATGGCTTCCAGCACAGACAGTGACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCC

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

BbsI (3893)

XmnI (3889)

3801 GCCCCGACATGGTGCTTGTGCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCA

33◀ 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 M

AseI (3955)

3901 TGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGAC

0◀

SacI (4012)

4001 GGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTC

SpeI (4110)

4101 CCGTTGATTTACTAGTCAAAAACAACTCCCATTTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAC

SnaBI (4238)

4201 TGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGG

NdeI (4343)

4301 CGGGCCATTTACCGTCATTGACGTCAATAGGGGCGTACTTGGCATATGATACTTGTACTGCTGCAAGTGGGCGAGTTTACCGTAAATACTCCACCCA

4401 TTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGGCGTCCAGGCCAGGGGCCA

PacI (4529)

SdaI (4521)

BspLU11I (4539)

4501 TTTACCGTAAGTTATGTAACGCTGCGAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTT

4601 TTTCCATAGGCTCCGCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCC

4701 CCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCT

ApaLI (4853)

4801 CACGCTGTAGGTATCTCAGTTCCGGTGTAGGTCGTTCCGCTCAAAGCTGGGCTGTGTGCACGAACCCCGTTACGCCGACCGCTGCGCCTTATCCGGTAA

4901 CTATCGTCTTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACA

5001 GAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTA

5101 GCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTT

EagI (5289)

PacI (5269) SwaI (5278) **NotI (5288)**

5201 GATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATA

5301 AAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGC

5401 AAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA