



PvuI (11)
SgfI (11) 1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA

MfeI (83)
101 GAGAAGGTGGCGCGGGTAAACTGGAAAAGTATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (205)
201 GTGAACGTTCTTTTTTCGAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACCGCGCCGCCCTACCTGAGGCC

301 GCCATCCACGCCGGTTGAGTCGGTTCGCGCCTCCGCCTGTGGTGCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (442)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

AgeI (553)
SgrAI (553) 501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTGCAACATGTTGTCTGGGAAGAAGCGGGCGGGCGGACAGCGG

BspLU11I (562)
1▶ M L S G K K A A A A A A

SfiI (689)
601 CGGCGGCGGGCGGCTGCTGGACCGAGGCCGGTCCGGGGCGGGCGGTTGCCGAGAAGGCTCTGAGGTGGCCGCGCCCGCCGGGCTGACGGG

13▶ A A A A A A A G T E A G S G A A G G A E N G S E V A A P P A G L T G

BstXI (710)
701 CCCACCGACATGGCTACGGGGCGGGCGGAGCGCACTCCCGAAAAGAAGAGCCTCCGCGGGCTCGCCGCGGGGGCTAGCCGAGCCGCGGGG

46▶ P T D M A T G A A G E R T P R K K E P P R A S P P G G L A E P P G

XmaI (774)
801 TCTGCTGGGCCCCAGGCGGGGCCACAGCCGGGCCGCTCCGCGACGCCATGGAGACCGGAATAGCCGAGACCCCGAGGGCCGACGGACCAGCCGGC

80▶ S A G P Q A G P T A G P G S A T P M E T G I A E T P E G R R T S R

NgoMIV (895)
901 GCAAGCGGCCAAGGTAGAATACAGAGAATGGATGAAAGCTTGCCAACCTCTCAGAAGTGAATATTATTCGGAAGAAGAAGAAATGCTAAAGCAGA

113▶ R K R A K V E Y R E M D E S L A N L S E D E Y Y S E E E R N A K A E

1001 GAAGGAAAAGAAGCTTCCCCACCACCTCTCAAGCCCCACCTGAGGAAGAAAATGAAAGTGAGCCAGAAGAGCCGCTCTGGTGTGGAGGGTGCAGCTTTT

146▶ K E K K L P P P P P Q A P P E E E N E S E P E E P S G V E G A A F

1101 CAAAGCCGACTTCCCCATGACCGAATGACCTCTCAGGAAGCAGCCTGTTCCAGACATCATCAGTGGGCTCAGCAGACACAGAAGTTTTTCTGTTCA

180▶ Q S R L P H D R M T S Q E A A C F P D I I S G P Q Q T Q K V F L F

1201 TCAGGAATCGCATTGCAAGTTATGGCTGGACAACCCAAAGATCCAGCTGACGTTTGAAGCCACTCTCCAGCAGCTGGAAGCCCTTACAACAGCGATAC

213▶ I R N R T L Q L W L D N P K I Q L T F E A T L Q Q L E A P Y N S D T

1301 TGTGCTTGCACCGAGTTCACAGTTACTTAGAGCCGCTGTTTCAACTTCGGCATCTATAAGAGGATAAAAACCTTACCAATTA AAAAGACAGGA

246▶ V L V H R V H S Y L E R H G L I N F G I Y K R I K P L I A K T G

1401 AAGGTGATTATTATAGTTTCAGTGTGTTCTGGCTTGGCAGCAGCTCGACAGCTACAGAGTTTTGGGATGGATGTCACACTTCTGGAAGCCAGGGATCGTG

280▶ K V I I I G S G V S G L A A A R Q L Q S F G M D V T L L E A R D R

BstBI (1523)
1501 TAGGTGGACGAGTTGCTACATTTGAAAAGGAAACTATGTAGCTGATCTTGGCGCCATGGTTGTAACAGGTCTTGAGGGGAATCCCATGGCTGTCGTCAG

313▶ V G G R V A T F R K G N Y V A D L G A M V V T G L G G N P M A V V S

1601 CAAACAAGTAAATGGAACCTGGCAAGATCAAGCAAAAATGCCACTTTATGAAGCAATGGACAAGCTGTTCCAAAAGAAAAGATGAAATGGTAGAA

346▶ K Q V N M E L A K I K Q K C P L Y E A N G Q A V P K E K D E M V E

AgeI (1711)
1701 CAAGAATTTAACGGTGTCTAGAAGCCACTTCTACCTTAGTACCAGTTAGACTTCAACGTCTCAATAATAAACCTGTATCCCTTGGCCAGGCATTGG

380▶ Q E F N R L L E A T S Y L S H Q L D F N V L N N K P V S L G Q A L

1801 AGGTTGTCACTAGCTGCAAGAAAAGCATGTCAAGATGAGCAGATTGAACATTGGAAGAAGATAGTAAAACCTCAGGAGGAGTTGAAAGAGCTTCTTAA

413▶ E V V I Q L Q E K H V K D E Q I E H W K K I V K T Q E E L K E L L N

EcoRV (1982)
1901 TAAGATGGTAAATTTGAAGGAGAAAATTAAGAGCTCCATCAGCAATACAAAAGAAGCTTCAGAAGTGAAGCCGCCAGAGATATCACAGCCGAGTTCTCTG

446▶ K M V N L K E K I K E L H Q Q Y K E A S E V K P P R D I T A E F L

XmnI (2078)
2001 GTGAAGAGCAAGCACAGGACCTGACTGCCCTCTGCAAGGAATATGATGAATTAGCTGAAACACAAGGAAAGCTAGAAGAAAACCTCAAGAATTGGAAG

480▶ V K S K H R D L T A L C K E Y D E L A E T Q G K L E E K L Q E L E

Bst1107I (2120)
2101 CCAATCCCCAAGTGTATACCTCTCATCAAGACAGACAATACTTACTGAGCATTGCAAACTCTGAAATTTGCAAACTCTGAAATTTGCAACGCCACACCTCTCTAC

513▶ A N P P S D V Y L S S R D R Q I L D W H F A N L E F A N A T P L S T

2201 CCTCTCTTAAACATTGGGATCAGGATGATGACTTTGAGTTTACTGGAAGCCACCTGACAGTAAGGAATGGCTACTCATGTGTGCCTGTGGCTTTAGCT

546▶ L S L K H W D Q D D D F E F T G S H L T V R N G Y S C V P V A L A

2301 GAAGGCTTGGACATTAACCTGAACACAGCAGTCCGGCAGGTTCCGCTACACAGCCTCAGGATGTGAAGTATTGCTGTGAACACACGTTCCACAAGTCAAA

580▶ E G L D I K L N T A V R Q V R Y T A S G C E V I A V N T R S T S Q

2401 CCTTTATTATAAGTGTGATGCACTTCTGTACACTTCTTTGGGAGTGTGAAGCAGCAGCCACAGCTGTTTCAAGTTGTGCCACCTTCTCTGAGTG

613▶ T F I Y K C D A V L C T L P L G V L K Q Q P P A V Q F V P P L P E W

SandI (2574)
2501 GAAAACATCTGCAGTCCAAGGATGGGATTTGGCAACCTTAACAAGGTGGTGTATGCTTTGACCGTGTGTTCTGGGACCCAAGTGTCAATTTGTTGGG

646▶ K T S A V Q R M G F G N L N K V V L C F D R V F W D P S V N L F G

NheI (2619)
2601 CACGTTGGCAGTACAACCTGCTAGCAGGGGTGAGCTCTTCTCTTCTGGAACCTATATAAAGCTCCAATACTATTGGCCCTGGTAGCAGGAGAAGCTGCTG
680▶ H V G S T T A S R G E L F L F W N L Y K A P I L L A L V A G E A A

EagI (2735) XcmI (2798)
2701 GCATTATGGAGAACATTAGTGACGATGTGATTGTCGGCCGGTGCCTGGCCATTCTCAAAGGGATTTTTGGCAGCAGTGCAGTCCCACAGCCCAAGGAAAC
713▶ G I M E N I S D D V I V G R C L A I L K G I F G S S A V P Q P K E T

XmaI (2833) BamHI (2862)
2801 TGTGGTATCTCGTTGGCGTGCTGATCCGTGGGCCCCGGGCTCCTATTCTTATGTGGCTGCAGGATCCTCTGGAATGACTATGATTTAATGGCTCAGCCG
746▶ V V S R W R A D P W A R G S Y S Y V A A G S S G N D Y D L M A Q P

BspEI (2974)
2901 ATCACTCCTGGCCCTCAATTCCAGGTGCCCCACAGCCAATCCAAGACTCTTCTTTGCTGGAGAACACACAATCCGGAACCTACCCAGCTACAGTCCATG
780▶ I T P G P S I P G A P Q P I P R L F F A G E H T I R N Y P A T V H

DraIII (3090)
3001 GTGCTCTGTTGAGTGGGCTTCGAGAAGCAGGAAGGATTGCCGACCAGTTTTGGGAGCCATGTACACTTTGCCTCGTACAGCCACACAGGTGTCCTGC
813▶ G A L L S G L R E A G R I A D Q F L G A M Y T L P R Q A T P G V P A

NheI (3131) EcoRI (3125)
3101 ACAGCAGTCCCCAAGTATGTGACAGAATTCGCTAGCTCGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATG
846▶ Q Q S P S M •

HpaI (3296)
3201 CTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTAACA

MfeI (3305) EcoRI (3390)
3301 ACAACAATTGCATTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATCTAATA
3401 TACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGC
3501 ATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGTTTGAAGTACTGCTTTCATTTCTTTATGTTTTAAATGCA

SspI (3631) SmaI (3646)
3601 CTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGC
3701 TCAAGGCCCTTCATAATATCCCCAGTTTGTAGTGGACTTAGGGAACAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTT
141◀ •

BstXI (3940)
3801 AGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGA
140◀ 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 C D P A Y D S
3901 GATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCC
107◀ 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 V I T D F D

StuI (4070)
4001 TTCTGCCGTTGCTCACAGCAGACCCAAATGGCAATGGCTTCCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCC
73◀ 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 V A S I I E G
4101 CAGTCTTGGTCTGATGGCCGCCGACATGGTGTGTTGCTCCTCATAGAGCATGGTGTCTTCTCAGTGGCAGCTCCACCAGTCCAGATCCTGCTG
40◀ 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 V L E L D Q Q

BbsI (4207) XmnI (4209) AseI (4277)
4201 AGAGATGTTGAAGGTCTTCATGGTGGCCCTCTATAGTGAGTCGATTATACTATGCCGATATACTATGCCGATGATTAATTTGTCAAACAGCGTGGATG
7◀ S I N F T K M ◀

4301 GCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTG

SpeI (4431)
4401 TTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCATTTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCT

SnaBI (4561)
4501 ATCCACGCCCATTTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAGGTCAT

NdeI (4665)
4601 GTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTT
4701 TACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGGGGGGTCTGTTGG

PacI (4854) SdaI (4847) BspLU11I (4860)
4801 GCGGTGAGCCAGCGGGCCATTTACCGTAAGTTATGTAACGCTGCGAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAA
4901 AAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCAGAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTAT
5001 AAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAG

ApaI (5174)

5101 CGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCGAC

5201 CGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGA

5301 GGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC

5401 CTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAA

PacI (5594) SwaI (5602)

5501 GGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACAT

EagI (5610)

NotI (5610)

5601 TTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAA

5701 ACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA