



PvuI (7)
SgfI (6) 1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82) **EcoNI (96)**
101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)
Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTGGCCGAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACAGCGCCCGCCGCTACCTGAGGCC
PvuII (239) **EcoNI (287)**
301 GCCATCCACGCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTGGCTGACCCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

KasI (535) **AgeI (552)** **BspHI (560)**
501 TCTGTTCTGCGCGGTTACAGATCCAAGCTGTGACCGCGGCTACCTGAGATCACCGGTCATCATGATGGCGGTGCCCGCTGGCTGCGGCTCTGATCCC
1▶ M M A R A R L A A A L I P

SandI (638)
601 AGCCACGGCCATCCTCTCCTGCCTGAGAACCAGAGAGTTGGACCTTGGTACAGGTTGTTCTAACATTAGTTACCAGTGCATGGAGCTGAATCTCTAC
13▶ A T A I L S C L R T E S W D P C V Q V V P N I S Y Q C M E L N L Y
701 AAAATCCCGACAACATCCCATATCAACCAAGATGCTGGACCTGAGCTTTAACTACCTGAGACATTTAGGCAGCCATAACTTCTCCAGCTTCCGAGAAC
47▶ K I P D N I P I S T K M L D L S F N Y L R H L G S H N F S S F P E
801 TGCAAGTGTGGATTTATCCAGATGTGAAATTAAGATTATTGAAGACGACACATTTTCAGGGCCTAAACCACCTCTCCACCTTGATACTGACGGAAACCC
80▶ L Q V L D L S R C E I K I I E D D T F Q G L N H L S T L I L T G N P
901 TATCCAGAGTTTAGCCTGGGAGCCTTTTCTGGCTATCAAGTTTACAGAAGCTGGTGGCGGTGGAGACAAACCTAGTCTCTCTAAATGACTTCCCATT
113▶ I Q S L A W G A F S G L S S L Q K L V A V E T N L V S L N D F P I
1001 GGACATCTCAAAAACCTGAAAGAGCTTAATGTGGCTCACAAATTTATCCATTCCTTCAAGTTACCTGAATATTTTCTAACCTGCCAACCTGGAGCACT
147▶ G H L K N L K E L N V A H N F I H S F K L P E Y F S N L P N L E H
1101 TGGATCTTTCTAACACAAAATCCAAAATATTTATGAAGATGTGAAAGTTCTACATCAAATGCCCTACTCAACCTCTTTAGATTGTCTTGA
180▶ L D L S N N K I Q N I Y Y E D V K V L H Q M P L L N L S L D L S L N

Acc65I (1222) **BspHI (1286)**
1201 CCCTTGAAGCTTTATTGAACCAAGTACCTTTAAAGAAATTAAGCTCAATGGATTGACTCTGCGAAGTAATTTTAAACAGTTTACATGTCAATGAAAACCTGT
213▶ P L D F I E P G T F K E I K L N G L T L R S N F N S S H V M K T C
1301 ATTCAAGGCTGGCTGGTTTAAAAACCAACCGGCTGGTTTGGGAGAATTTAAAAATGAAAGGAGTTGCAAAGATTGACAGATCTTCTGGAGGGAC
247▶ I Q G L A G L K T N R L V L G E F K N E R K L Q R F D R S F L E G

BspEI (1424)
1401 TGTCAACCTGACCATTGAACAATTCGGATAGCTACTTGGACAAATCTCAGGGGACGATACAGACTTATTTAATTGTTTGGCAAATGTTTCTGTGAT
280▶ L C N L T I E Q F R I A Y L D K F S G D D T D L F N C L A N V S V I
1501 TTCTCTGTGAGTATATCTTTAGGAAGTCTACAAGCCCTTCTAAAGATTTTAGATGGCAACACTTAGAAATTTAACTGTGACTTTGACAAGTTTCT
313▶ S L L S I S L G S L Q A L L K D F R W Q H L E I I N C D F D K F P

XbaI (1692)
1601 GCATTGAAGCTCAGTTCTCTCAAAAAGTTGTTTTACAGACAACAAAGATATAAGCACTTTTACTGAGTTTACAGCTACCAAGCCTTACAGTATCTAGATC
347▶ A L K L S S L K K F V F T D N K D I S T F T E F Q L P S L Q Y L D
1701 TCAAAAGAAATCACTTGAGTTTCAAGGGTGTGCTTCTCACACTGATTTTGGGACAACCAACCTGAAGCATTGATCTGAGCTTCAATGATGTCATTAC
380▶ L K R N H L S F K G C C S H T D F G T T N L K H L D L S F N D V I T
1801 CTTAGGTTCAAACCTCATGGGCTTAGAGCAACTAGAACACCTGGATTTTTCAGCTTCCACTCTGAAACAGATCAATGCTTTTTTTCAGCATTCTATCACTC
413▶ L G S N F M G L E Q L E H L D F Q H S T L K Q I N A F S A F L S L

NcoI (1950)
1901 AGAAACCTCCGCTACCTTGATATTTCTTATACCAACATCCGATTGTCTTCCATGGCATCTTTACTGGCTTAGTCACTGCTGCAACCTTGAAAATGGCAG
447▶ R N L R Y L D I S Y T N I R I V F H G I F T G L V S L Q T L K M A
2001 GCAACTCTTTTTCAGAACAACTTGTCCCTGACATCTTACAGAACTGACTAATTAACCGTCTTGGACCTCTAAGTGTCAACTGGAACAGGTAGCCCA
480▶ G N S F Q N N L L P D I F T E L T N L T V L D L S K C Q L E Q V A Q

BspEI (2198)
2101 GACAGATTTCACTCCCTCTAGCCTTCAAGGTCGATGATGAGTCAACAACAACTCTTGTCACTGGATACATTTCTTTATGAACCACTCCACTCGCTC
513▶ T A F H S L S S L Q V L N M S H N K L L S L D T F L Y E P L H S L

BamHI (2201) **NsiI (2297)**
2201 CGGATCCTAGACTGCAGTTTCAACCGTATCATGGCCTCTAAGGAGCAAGAACTACAGAATTTGCCAAGGAGCCTCACTTGGCTAAATCTTACTCAGAATG
547▶ R I L D C S F N R I M A S K E Q E L Q N L P R S L T W L N L T Q N
2301 CATTGCTTGTGTTTGTGAACATCAGAGTTTCTGAGTGGGTCAAGGACAGAGGACGCTCTTGGTGGGAGCTGAGCAAATGATGTGTGAGAGCCTTT
580▶ A F A C V C E H Q S F L Q W V K D Q R Q L L V G A E Q M M C A E P L

PvuII (2446)
2401 AGATATGGAGGACATGCCAGTGTAGTTTTCAGGAACGCCACTTGTGACGTGAGCAAGACGATCATCAGTGTGTCGGTGGTCACTGTGCTCCTGGTGTCT
613▶ D M E D M P V L S F R N A T C Q L S K T I I S V S V V T V L L V S
2501 GTGGTAGGGTCTAGTCTACAAGTTCTATTTCCACCTGATGCTTCTTGTGCTGCAAAAAGTATGGCAGGGGCGAGAGCACCTATGATGCCTTTGTCA
647▶ V V G V L V Y K F Y F H L M L L A G C K K Y G R G E S T Y D A F V

XhoI (2604)
2601 TCTACTCGAGCCAGGATGAAGACTGGGTGCGGAATGAAGTGGTAAAGAACTTGGAGGAGGGCGTCCCCCTTTTCACTCTGCTTCACTACAGGGACTT
680▶ I Y S S Q D E D W V R N E L V K N L E E G V P P F Q L C L H Y R D F

MscI (2711)
2701 TATTCTGGGGTGGCCATCGCCGCAATATCATCCAGGAGGTTTCCACAAAAGCCGTAAGGTTATTGTCGTGGTGTCCAGCACTTCACTCAGAGCCGA
713▶ I P G V A I A A N I I Q E G F H K S R K V I V V V S Q H F I Q S R

2801 TGGTGTATCTTCGAGTATGAGATTGCCAGACCTGGCAGTTTCTGAGCAGCCGTGCTGGCATCATCTTCATCGTCTGCAGAAGCTGGAGAAGTCTCTCC
747▶ W C I F E Y E I A Q T W Q F L S S R A G I I F I V L Q K L E K S L

Tth111I (2959)

2901 TGGCGCAGCAGGTGGAGCTCTATCGCCTTCTGAGCAGGAACACCTACCTGGAGTGGGAGGACAGTGTCTTGGGCGGCATGTCTTCTGGAGAAGACTCAG
780▶ L R Q Q V E L Y R L L S R N T Y L E W E D S V L G R H V F W R R L R

3001 AAAAGCCTTGCTGGCTGGTAAACCCAGAGTCCAGAAGGAACAGCAGATGCAGAAACCAACCCGAAGAAGCGACAACCTCCACCTGAGGAGGAGAATCC
813▶ K A L L A G K P Q S P E G T A D A E T N P Q E A T T S T •

MseI (3114)

3101 CCTGATGTGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAATTTG

HpaI (3246) MfeI (3257)

3201 TGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACCAAGTTAAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGG

EcoRI (3342)

3301 GAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGAATTCATAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAAT

3401 CCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAG

SapI (3524) SwaI (3595)

3501 TGTATTTTCCAAGGTTTGAAGTACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACATTCCTTTTATGATAAATTCAGAAAATATTTA

3601 AATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGA

3701 ACAAAGGAACCTTAATAGAAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTG

3801 ACCAGCTTGCCATTCATCTCAATGAGCACAAGCAGTCAAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGCTGACCACCTGATGGATC
124▶ 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 C P S V V R I S R

3901 TGCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACA
91▶ D V E D S Y 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

StuI (4020)

4001 GACAGTGACCCTGCCAATGTAGCCCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCTGTTGCCTCA
58▶ V T V R G I Y 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

XmnI (4162)

4101 TAGAGCATGGTATCTTCTCAGTGGCGACCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGGTGGCCCTCCTATAGTGAGTCGTAT
24▶ 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 (4228)

4201 TATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATAT

SpeI (4383)

4301 AGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAAGTCCCGTTGATTTACTAGTCAAAACAAACT

4401 CCCATTGACGTCAATGGGGTGGAGACTTGAAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATA

SnaBI (4511)

4501 GCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCCATAAGGTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCAATGACGTCAA

NdeI (4616)

4601 TAGGGGGCGTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAAGTCCCTATTGG

SdaI (4794)

4701 CGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCA

PacI (4802) BspLU11I (4812)

4801 GGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGA

4901 GCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGCGCTTCCCCCTGGAAGCTCCCTCGTGGCTCTCCT

5001 GTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGT

ApaLI (5126)

5101 AGGTGCTTCGCTCCAAGTGGGCTGTGTGCACGAACCCCGTTCAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGGTAAG

5201 ACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTAC

5301 GGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAAACAAACCCG

5401 CTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGCTGTGACGCTCA

EagI (5562) NotI (5561)

5501 GTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCAATGGCTGATTAATTAACATTTAAATCAGCGGCCGCAATAAAAATATCTTTATTTTATTACATCTGT

5601 GTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAAGTGC

5701 AGGTGCCAGAACATTTCTCTATCGAA