



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

101 GAGAAAGTGGCGGGGTAACGGAAAGTGTGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

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

301 GCCATCCACGCGGTTGAGTCGCGTTTCTGCCGCCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGCTAGGTAAGTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
NaeI (441) 401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

NcoI (560)
BstEII (555)
KasI (535) 501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTA CCGT GAGATCACCGGTCCACCATGGGGAGGCGGAGAAGTTCCACTACATCTACAGCTG **AgeI (552)**

601 TGACCTGGACATCAACGTGCAGCTGAAGATAGGGAGCTTGAAGGGAAGAGAGAACAACAAAGAGCTATAAAGCTGTCTAGAAGATCCCATGTTAAAGTTT
13▶ D L D I N V Q L K I G S L E G K R E Q K S Y K A V L E D P M L K F
701 TCTGGCTATAACAGAGACATGCTCCGACCTCTATGTGACTTGTCAAGTGTGCTGAAGGGAAGCCCTGGCCTTACCAGTCAGAATCCTACAAGG
47▶ S G L Y Q E T C S D L Y V T C Q V F A E G K P L A L P V R T S Y K
801 CGTTTAGTACAAGATGGAATTGGAACGAGTGGCTGAAACTTCCGGTGAAGTACCTGACCTGCCAGGAATGCCAAGTGGCCTTGACTATATGGGATGT
80▶ A F S T R W N W N E W L K L P V K Y P D L P R N A Q V A L T I W D V

Eco47III (913) 901 GTATGGACCGGGAAGCGTGTGCCAGTGGGAGGAACAACCGTGTGCCTCTTTGGAAAATATGGCATGTTTCGCCAAGGAATGCATGACTGAAGGTCTGG
AfeI (913) 113▶ Y G P G S A V P V G G T T V S L F G K Y G M F R Q G M H D L K V W **Ppu10I (979)**
NsiI (979) 1001 CCTAACGTGGAGGAGATGTTCTGAACCTACAAGAACTCCTGGCAGAACAAGCAGCACACTGTGAGAAGATCAGATGAGCCGCTCGCAAGCTCACCA
147▶ P N V E A D G S E P T R T P G R T S S T L S E D Q M S R L A K L T

DraIII (1117) 1101 AGGCTCATCGGCAAGGACATGGTGAAGTGGATTGGCTGGACAGATTAACATTTAGAGAGATAGAAATGATAAATGAGAGTGAAGAACGAAGCTCTAA
180▶ K A H R Q G H M V K V D W L D R L T F R E I E M I N E S E K R S S N
1201 TTTTACTGTTGATGTTGAGTTTTCGCTGTGTCAAGTGCAGTACAAGGAGTATGGCATTGTTTACTATGAAAAGGATGGTACGAATCATCTCCAATT
213▶ F M Y L M V E F R C V K C D D K E Y G I V Y Y E K D G D E S S P I

HindIII (1375) 1301 TTAACCAGTTTTGAGTTAGTGAAGTTCCTGATCCTCAAATGTCTATGGAGAATTTAGTGGAGAGCAACACCACAAGCTTGCCTCGGAGCTTAAAGAGCG
247▶ L T S F E L V K V P D P Q M S M E N L V E S K H H K L A R S L R S

SspI (1450) 1401 GGCCATCTGACCACGATCTCAAACCCAACGCTACCACAAGAGATCAGCTAAATATTATTGTGAGTTACCCACCAACCAAGCAACTCACATATGAAGAACA
280▶ G P S D H D L K P N A T T R D Q L N I I V S Y P P T K Q L T Y E E Q **NdeI (1487)**

BglII (1500) 1501 AGATCTTGTGGAAATTTAGATATTATCTTACTAATCAAGAAAAAGCTCTGACGAAGTCTTGAAGTGTGTTAATTGGGACCTGCCCCAGGAGGCCAAG
313▶ D L V W K F R Y Y L T N Q E K A L T K F L K C V N W D L P Q E A K **XcmI (1595)**
SfiI (1593) 1601 CAGGCCCTGGAACCTCTGGGAAAGTGAAGCCAATGGATGTAGAGGACTCCCTGGAGCTGCTCCTCTCATTACCAACCCACCGTGAGGCGCTACG
347▶ Q A L E L L G K W K P M D V E D S L E L L S S H Y T N P T V R R Y

Bsu36I (1710) 1701 CTGTGGCCCGCTGAGGCAAGCCGATGATGAGGATTTGCTGATGACTTACTGAGCTGGTCCAGGCTCTGAAATATGAAAACCTTGTGACATAAAGAA
380▶ A V A R L R Q A D D E D L L M Y L L Q L V Q A L K Y E N F D D I K N **PstI (1750)**
1801 TGGTTTGAACCTACCAAGAAAGATAGTCAAACCTCAGCATCAGAAAGTCTGTCAAATCTGGAGTCAAGTCTGGAGACATAGATAGCTCCCAATTATA
413▶ G L E P T K K D S Q T S A S E S L S N S G V S S G D I D S S Q I I

BglII (1980) 1901 ACCAACCTCTTCTCCCGTGGCCTCCCTCCTCCTGCATCTAAAGCAAAGGAAGTTCCGATGGGGAAAATCTCGAGCAAGATCTATGTACGTTCTTGA
447▶ T N P L P P V A S P P P A S K A K E V S D G E N L E Q D L C T F L **XhoI (1972)** **EcoRV (1998)**

2001 TATCAAGAGCCTGTAAAGAACTCAACACTGGCTAATTATTTACTGTTATGTGATAGTGGAGTGTGAAGATCAGGACCCAGCAGCGGACCCAAAGAC
480▶ I S R A C K N S T L A N Y L Y W Y V I V E C E D Q D T Q Q R D P K T **SandI (2087)**
2101 TCACGAGATGACCTGAACGTGATGAGGCGCTTACGCAAGCCTTGTCAAGGGTGAAGTCTGTGAGAGTCATGCGCTCCCTGCTGGCGGCTCAGCAG
513▶ H E M Y L N V M R R F S Q A L L K G D K S V R V M R S L L A A Q Q

StuI (2282) 2201 ACCTTTGTAGATCGCTGGTCCATCTGATGAAGCGGTCAGAGAGAAAGTGGGAACCGCAAGAAGAAGTGAAGACTTACGCGCTTGTGGAGATA
547▶ T F V D R L V H L M K A V Q R E S G N R K K K N E R L Q A L L G D
2301 ATGAGAAAATGAACCTTACAGATGGAAGTATCCATTGCCGCTGGAGCCACAGGTGAAAATAAGGGGCATCATCCCGAAAACAGCTACCTGTTCAA
580▶ N E K M N L S D V E L I P L P L E P Q V K I R G I I P E T A T L F K **PshAI (2479)**

2401 GAGTGTCTTATGCCTGCCAGCTGTTCTTCAAGACTGAAGTGGAGGCAAAATACCCAGTTATTTTCAAGCATGGAGAGACTTGCCTCAAGATCAGCTT
613▶ S A L M P A Q L F F K T E D G G K Y P V I F K H G D L R Q D Q L
2501 ATTCTTCAAGATCATCTCCCTCATGGACAAGCTGTTACGGAAAAGAAAACCTGGATTTGAAATGACCCATAAAGGTGTTAGLCCACTAGCACAAAACATG
647▶ I L Q I I S L M D K L L R K E N L D L K L T P Y K V L A T S T K H
2601 GCTTCATGAGTTCATCCAGTCCGTTCTGTGCAGAAGTCTTGTGACACAGAGGAAGCATTGAGAACTTTTTCAGGAAATATGCCCCAGTGGAGACTGG
680▶ G F M Q F I Q S V P V A E V L D T E G S I Q N F F R K Y A P S E T G

2701 ACCCTATGGCATCAGTGCAGAGGTCATGGACACTTACGTTAAAAGCTGTGCTGGATACTGTGTGATTACGTACATTCTTGGAGTTGGAGACCCGGCACCTG
713▶ P Y G I S A E V M D T Y V K S C A G Y C V I T Y I L G V G D R H L

HpaI (2811)

2801 GATAACCTTTTGTAAACAAAGACAGGCAAACCTCTCCATATAGATTTCCGGGTATATTTTGGGTCGAGATCCGAAGCCCCTCCCTCCTCCGATGAAGCTGA
747▶ D N L L L T K T G K L F H I D F G Y I L G R D P K P L P P P M K L

2901 ACAAGGAGATGGTGAAGGGATGGGTGGCACCCAGAGTGAGCAGTACCAAGAGTTCCGAAAGCAGTGTACACAGCCTTCTCCACCTGCGAAGGTATTC
780▶ N K E M V E G M G G T Q S E Q Y Q E F R K Q C Y T A F L H L R R Y S

3001 CAATCTGATCTTGAACCTGTTCTCTTGTGTTGATGCAAAACATTCCAGATATTGCTCTTGAGCCAGATAAAACTGTAAAAAGGTTCCAGGATAAGTTC
813▶ N L I L N L F S L M V D A N I P D I A L E P D K T V K K V Q D K F

ApaLI (3127)

3101 CGCCTGGACCTGTCAGATGAGGAGGCTGTGCACTACATGCAGAGTCTGATCGACGAGAGCGTCCATGCACTGTTCCGCCGCGTGGTGGAGCAGATCCATA
847▶ R L D L S D E E A V H Y M Q S L I D E S V H A L F A A V V E Q I H

MscI (3259)

ScaI (3209) NheI (3253)

3201 AGTTTGCCAGTACTGGAGAAAATGAAAGGGGATTTGACCCATCCTGATGCCTGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAAC
880▶ K F A Q Y W R K •

HpaI (3391)

3301 CACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAAACAAC

MfeI (3402) EcoRI (3487)

3401 AACAAATGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAAATTCATAAAATA

3501 CAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCAT

3601 TAGCTGTTGCAGCCTCACCTTCTTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTGAACTAGCTCTTCATTTCTTTATGTTTTAAATGCAT

SspI (3726) SwaI (3740)

3701 GACCTCCACATTCCTTTTTAGTAAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTC

3801 AAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAG
141▶ •

3901 TTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGA
139▶ 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 I

SacI (4001) BstXI (4030)

4001 TGAGTCTCTGCACATGCCACAGGGGCTGACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCCTT
106▶ 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 K

StuI (4165)

4101 CTGCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCA
73▶ 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

4201 GTCTTGGTCTGATGGCCGCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAG
39▶ 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 S

BspHI (4315) BbsI (4311) AseI (4373)

4301 AGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGC
6▶ I N F T K M

SacI (4430)

4401 GTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGCGGAGTTGTT

SpeI (4528)

4501 ACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTA

SnaBI (4656)

4600 TCCACGCCATTGATGTAAGTCCCAAACCGCATCATGGAATAGCGATGACTAATACGTAGATGTAAGTCCCAAAGTAGGAAAGTCCCATAGGTCATG

NdeI (4761)

4700 TACTGGGCATAATGCCAGGCGGGCCATTTACCGTCAATGAGGCGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCAGTTT

4800 ACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGAAACATACGTCAATTATTGACGTCAATGGGCGGGGCTGTTGGG

PstI (4940) SdaI (4939) PacI (4947) BspLU111 (4957)

4900 CGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCCTG C A G G T T A A T T A A G A A C A T G T G A G C A A A A G G C C A G A A A A G G C C A G A A C C G T A A
4998 AAAGCGCGCTTGTGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTA

5098 TAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAA

ApaLI (5271)

5198 GCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGCCGA

5298 CCGCTGCGCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCG

5398 AGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTA

5498 CCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAA

5598 AGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACA

EagI (5707)

NotI (5706)

5698 TTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAA

5798 AACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAAGTGCCAGAACATTTCTCTATCGAA