



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGGCCGCAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCGCTACCTGAGGCC
HindIII (245)
Bsu36I (291)
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTGTTTCGTTT

KasI (535) 501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACGGTCCATCATGAACGCCCCGAGCGGCAGCCCCAACCCGACGGCGG
AgeI (552) 1▶ M N A P E R Q P Q P D G G
601 GGACGCCCCAGGCCACGAGCCTGGGGCAGCCCCAAGACGAGCTTGACTTCTCCATCCTCTTTCGACTATGAGTATTTGAATCCGAACGAAGAAGAGCCG
13▶ D A P G H E P G G S P Q D E L D F S I L F D Y E Y L N P N E E E P

BspEI (727) 701 AATGCACATAAGGTGCCAGCCACCCTCCGGACCCGCATACCCCGATGATGTCCTGGACTATGGCCTCAAGCCATACAGCCCCCTTGCTAGTCTCTCTG
47▶ N A H K V A S P P S G P A Y P D D V L D Y G L K P Y S P L A S L S

BsaBI (897) 801 GCGAGCCCCCGCGGATTCGGAGAGCCGATAGGGTAGGGCCGAGAAGTTTCTGAGCGCGCAAGCCAGCAGGGGCTCGGGCTGAGCCCTCGGAT
80▶ G E P P G R F G E P D R V G P Q K F L S A A K P A G A S G L S P R I

SfiI (987) 901 CGAGATCACTCCGTCACGAACTGATCCAGGCAGTGGGGCCCTCCGCATGAGAGACGCGGGCTCCTGGTGGAGCAGCCGCCCCCTGGCCGGGGTGGCC
113▶ E I T P S H E L I Q A V G P L R M R D A G L L V E Q P P L A G V A

NheI (1069) 1001 GCCAGCCGAGGTTACCCTGCCGTGCCGGCTTCGAGGGCTACCGGAGCCGCTTTGCTTGGAGCCCGTAGCAGCGGCTCCTCTGCCAGCTTCATT
147▶ A S P R F T L P V P G F E G Y R E P L C L S P A S S G S S A S F I

PshAI (1159) 1101 CTGACACCTTCTCCCCACACCTCGCCCTGCGTCTCGCCCAATAACGGCGGGCCGACGACCTGTGTCCGAGTTTCAAACATCCCTGCTCATTATTC
180▶ S D T F S P Y T S P C V S P N N G G P D D L C P Q F Q N I P A H Y S
1201 CCCCAGAACTCGCAATAATGTACCTCGAACAGCCTCGCCGAGGACAGCTGCCTGGCCGCGCACTGCCCCGTGCCCGCTCCGCGCTCCGCTCCTCA
213▶ P R T S P I M S P R T S L A E D S C L G R H S P V P R P A S R S S

Eco47III (1370)
AfeI (1370) 1301 TCGCCTGGTGCCAAAGCGGAGGCATTGTCGCCGAGGCCCTTGGTTGCCCTGCCGCCGAGCCTCACCCAGCGCTCCCGGAGCCCTCGCCGAGCCCT
247▶ S P G A K R R H S C A E A L V A L P P G A S P Q R S R S P S P Q P

Asp718I (1439)
Acc65I (1439) 1401 CATCTCACGTGGCACCCAGGACCACGGCTCCCCGGCTGGGTACCCCTGTGGTGGCTGTCGGTGATCATGGATGCCCTGAACAGCCTGCCACGGA
280▶ S S H V A P Q D H G S P A G Y P P V A G S A V I M D A L N S L A T D

BamHI (1511) 1501 CTGCCTTGTGGGATCCCCCAAGATGTGGAAGACCAGCCCTGACCCCTCGCCGGTGTCTGCCGCCCATCAAGGCCGGCCTGCTCGCCACATCTAC
313▶ S P C G I P P K M W K T S P D P S P V S A A P S K A G L P R H I Y
1601 CCGCCGTGGAGTTCTGGGGCCCTGCGAGCAGGGCAGAGAGAACTCGCTCCAGAATCCATCCTGCTGGTTCGCCCCACTTGGCCCAAGCCGCTGG
347▶ P A V F L G P C E Q G E R R N S A P E S I L L V P P T W P K P L
1701 TGCCTGCCATTCCATCTGCAGCATCCAGTGAATCCCTCCCTCCACTTGAAGTGGCCGCTGTCCAGTCACTCAGGCTCTTACGAGCTGCGGATCGA
380▶ V P A I P I C S I P V T A S L P P L E W P L S S Q S G S Y E L R I E

DraIII (1879) 1801 GGTGCAGCCAAAGCCACATCACCGGCCACTATGAGACAGAAGCGAGCCGAGGGGTGTCAAAGCTCAAAGTGGAGGCCACCTGTGGTTTCACTCCAT
413▶ V Q P K P H H R A H Y E T E G S R G A V K A P T G G H P V V Q L H

BglII (1931) 1901 GGCTACATGAAAACAAGCCTCTGGGACTTCAGATCTTATTGGGACAGCTGATGAGCGGATCCTTAAGCCGACGCCTTCTACCAGTGCACCGAATCA
447▶ G Y M E N K P L G L Q I F I G T A D E R I L K P H A F Y Q V H R I
2001 CGGGAAAAGTGTACCACCACGCTATGAGAAGATAGTGGGCAACCAAAGTCTTGAGATACCCCTGGAGCCAAAACAACATGAGGGCAACCAT
480▶ T G K T V T T T S Y E K I V G N T K V L E I P L E P K N N M R A T I

HindIII (2119) 2101 CGACTGTGCGGGATCTTGAAGCTTAGAAAACGCCGACATTGAGCTGCGGAAAGCGAGACGGACATTGGAAGAAAGAACACGCGGGTGGAGCTGGTTTTC
513▶ D C A G I L K L R N A D I E L R K G E T D I G R K N T R V R L V F
2201 CGAGTTCACATCCAGATCCAGTGGCAGAAATGTCCTTTACAGACTGCATCTAACCCATCGAGTGTCCAGCGATCTGCTCAGAGCTGCCATGG
547▶ R V H I P E S S G R I V S L Q T A S N P I E C S Q R S A H E L P M
2301 TTAAAGACAAGACACAGACAGTGCCTGGTCTATGGCGCCAGCAAATGATCCTCACGGGCGAGAATTTACATCCGAGTCCAAAGTTGTGTTACTGA
580▶ V E R Q D T D S C L V Y G G Q Q M I L T G Q N F T S E S K V V F T E
2401 GAAGACCACAGATGGACAGCAAATTTGGGAGATGGAAGCCAGGTTGGATAAGGACAAGAGCCAGCCCAACATGCTTTTTGTTGAGATCCCTGAATACCG
613▶ K T T D G Q Q I W E M E A T V D K D K S Q P N M L F V E I P E Y R
2501 AACAAAGCATATCCGCACACCTGTAAAAGTGAACCTTCTACGTCAATCAATGGGAAGAGAAAACGAAGTCAAGCCTCAGCCTTACCTACCACCCAGTCCAG
647▶ N K H I R T P V K V N F Y V I N G K R K R S Q P Q H F T Y H P V P
2601 CCATCAAGACGGAGCCACGGATGAATATGACCCACTCTGATCTGCAGCCCCACCATGGAGGCGTGGGGAGCCAGCCTTACTACCCCGAGCACCCGAT
680▶ A I K T E P T D E Y D P T L I C S P T H G G L G S Q P Y Y P Q H P M

MscI (2726) 2701 GGTGGCCGAGTCCCCCTCCTGCCTCGTGGCCACCATGGCTCCCTGCCAGCAGTTCGCGACGGGGTCTCATCCCTGACGCCGCTACCAGCAACAGAAC
713▶ V A E S P S C L V A T M A P C Q Q F R T G L S S P D A R Y Q Q Q N

SfiI (2865)

2801 CCAGCGGCCGTACTCTACCAGCGGAGCAAGAGCCTGAGCCCCAGCCTGCTGGGCTATCAGCAGCCGGCCCTCATGGCCGCCCGCTGTCCCTTGGCGACG
747▶ P A A V L Y Q R S K S L S P S L L G Y Q Q P A L M A A P L S L A D

2901 CTCACCGCTCTGTGCTGGTGACGCCGGCTCCCAGGGCCAGAGCTCAGCCCTGCTCCACCCTCTCCGACCAACCAGCAGGCCTCGCCTGTGATCCA
780▶ A H R S V L V H A G S Q G Q S S A L L H P S P T N Q Q A S P V I H Y

3001 CTCACCCCAACAGCAGCTGCGCTGCGGAAGCCACAGGAGTTCAGCACATCATGTACTGCGAGAATTTGCGACCAAGCACCACAGACCTGGCCCG
813▶ S P T N Q Q L R C G S H Q E F Q H I M Y C E N F A P G T T R P G P

XmaI (3129)
Srfl (3128)

3101 CCCCCGGTCAGTCAAGGTCAGAGGCTGAGCCCGGCTCCCTACCCACAGTCAATCAGCAGCAGAATGCCACGAGCCTAAAGAGCCGCAAAAACGGACCCC
847▶ P P V S Q G Q R L S P G S Y P T V I Q Q Q N A T S Q R A A K N G P

BstEII (3234)

3201 CGGTCAGTGACAAAAAGGAGTATTACCTGCGGGGTGACCATTAACAGGAGCAGAATTGGACCAGACCTACTTGGATGATGTTAATGAAATTATCAG
880▶ P V S D Q K E V L P A G V T I K Q E Q N L D Q T Y L D D V N E I I R

AvrII (3349) MscI (3361)

3301 GAAGGAGTTTTCAGGACCTCCTGCCAGAAATCAGACGTAAAAGAAGCCACCTAGGACTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGACAA
913▶ K E F S G P P A R N Q T •

HpaI (3493)

3401 ACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAGTTAA

MfeI (3504) EcoRI (3589)

3501 ACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTTCAAAA

▶ ◀

3601 TACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGC

3701 ATTAGCTGTTTGCAGCCTCACCTCTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTGAACTAGCTCTTCATTTCTTTATGTTTTAAATGCA

SspI (3828) SmaI (3842)

3801 CTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGC

3901 TCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTT

141 ▶ •

4001 AGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCACTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGA
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

4101 GATGAGCTCTCGACATGCCACAGGGGCTGACCACCTGTAGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCC
107▶ I L E R C M G C P S V R I S R D V E D S Y P H R V A V I T D F D

4201 TTCTGCCGTTGCTCACAGCAGCCCAATGGCAATGGCTTCAGCACAGCAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCC
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

4301 CAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTGTCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCTGTCTG
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

BspHI (4417)

XmnI (4409) AseI (4475)

4401 AGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATG
7▶ S I N F T K M ◀

4501 GCGTCTCCAGCTTATCTGACGGTTCCTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTG

SpeI (4630)

4601 TTACGACATTTTGAAAAGTCCCCTGTGATTTACTAGTCAAACAAACTCCCATTTGACGTCAATGGGGTGGAGACTTGAAATCCCCTGAGTCAAACCGC

◀

SnaBI (4758)

4700 TATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCA

NdeI (4863)

4800 TGTACTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGCGTACTTGGCATATGATACACTTGTACTGCCAAGTGGGCAGT

4900 TTACCGTAAATACTCCACCATTGACGTCAATGAAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTG

SdaI (5041) PacI (5049) BspLU11I (5059)

5000 GCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCC T G 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 G G C C A G G A A C C G T

◀

5098 AAAAAAGCCCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAGAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC

5198 TATAAAGATACCAGCGTTTTCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGG

5298 AAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTAGGCC

5398 GACCGCTGCGCTTATCCGGTAACATATCTGCTTGTAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG

5498 CGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT

5598 TACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAA

PacI (5789)

5698 AAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTATGGCTAGTTAATTA

SmaI (5798) NotI (5808)

5798 CATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGTTTTTGTGTAATCGTAACTAACATACGCTCTCCATCAAAAC
5898 AAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCCAGTCAAGTGACAGTGCCAGAACATTTCTCTATCGAA