



PvuI (7)
SgfI (6) 1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGTGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTTCGAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACCGCGCCCGCCCTACCTGAGGCC
HindIII (245)
Bsu36I (291)
301 GCCATCCACGCCGGTTGAGTCGGTTCGCGCCTCCCGCTGTGGTGCCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCCTGCTTCTCAACTCTACGCTTTTGTTCGTTT

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCACCATGGTGGAGGAGATACCCAGCAGCTGCGACACCATCT
BstEII (555) **AgeI (552)** **NcoI (560)** **Bst1107I (573)**
1 M V S R Y T Q Q L R H H L
BglII (659)
601 GGGCCGTGACTCCAAGTTCGTGCTGTGCTATGCCAGAAGGAGGAGCTGCTGCTGGAGGAGATCTACATGGACACCATCATGGAGCTGGTTGGCTTCAGC
13 G R D S K F V L C Y A Q K E E L L L E E I Y M D T I M E L V G F S
701 AATGAGAGCCTGGGAGCCTGAACAGCCTGGCTGCCTCCTGGACCACACCACCGCATCCTCAATGAGCAGGGTGGAGCCATCTTATCCTGGGTGATG
47 N E S L G S L N S L A C L L D H T T G I L N E Q G E T I F I L G D
801 CTGGGGTGGGCAAGTCCATGCTGCTACAGCGCTGCAGAGCCTCTGGGCCACGGCGGCTAGACGAGGGGTCAAATTTCTTCCACTTTCCGCTGCCG
80 A G V G K S M L L Q R L Q S L W A T G R L D A G V K F F F H F R C R

901 CATGTTACAGTGTCTCAAGAAAGTGCAGGCTGTGTCTGCAGGACCTGCTCTCAAGCACTACTGCTACCCAGAGCGGGACCCGAGGAGGTGTTTGCC
113 M F S C F K E S D R L C L Q D L L F K H Y C Y P E R D P E E V F A
BsrBI (973)
BsrPI (1018)
1001 TTCCTGCTGCGCTTCCCCACGTGGCCCTTTCACCTTCGATGGCCTGGACGAGCTGCACTCGGACTTGGACCTGAGCCGCGTGCCTGACAGCTCCTGCC
147 F L L R F P H V A L F T F D G L D E L H S D L D L S R V P D S S C

1101 CCTGGGAGCTGCCACCCCTGGTCTTGTGGCAACCTGCTCAGTGGAAAGCTGCTCAAGGGGCTAGCAAGCTGCTCACAGCCCGCACAGGCATCGA
180 P W E P A H P L V L L A N L L S G K L L K G A S K L L T A R T G I E
MscI (1130) **NheI (1165)**
1201 GGTCCCGCGCAGTTCCTGCGGAAGAAGGTGCTTCTCCGGGGCTTCTCCCCAGCCACTGCGCGCTATGCCAGGAGGATGTTCCCCGAGCGGGCCCTG
213 V P R Q F L R K K V L L R G F S P S H L R A Y A R R M F P E R A L
1301 CAGGACCGCCTGCTGAGCCAGCTGGAGGCCAACCCCAACTCTGCAGTCTGTGCTCTGTGCCCTTCTGCTGGATCATCTTCCGGTGCTTCCAGCAGT
247 Q D R L L S Q L E A N P N L C S L C S V P L F C W I I F R C F Q H

1401 TCCGTGCTGCCTTGAAGGCTCACACAGCTGCCCCACTGCACGATGACCCTGACAGATGTCTTCTCCTCTGGTCACTGAGGTCATCTGAACAGGATGCA
280 F R A A F E G S P Q L P D C T M T L T D V F L L V T E V H L N R M Q
PshAI (1452)
DraIII (1563)
1501 GCCAGCAGCCTGGTGCAGCGGAACACACGCGCCAGTGGAGACCCTCCACGCCGGCGGACACTGTGCTCGTGGGCGAGGTGGCCACCGGGGC
313 P S S L V Q R N T R S P V E T L H A G R D T L C S L G Q V A H R G
1601 ATGGAGAAGAGCCTTTTGTCTTCAACCAGGAGGAGGTGCAGGCTCCGGGCTGCAGGAGAGAGACATGCAGCTGGGCTTCTGCGGGCTTTGCCGGAGC
347 M E K S L F V F T Q E E V Q A S G L Q E R D M Q L G F L R A L P E

1701 TGGCCCCGGGGTGACCAGCAGTCTATGAGTTTTTCCACCTCACCTCCAGGCCTTCTTACAGCCTTCTCCTCGTGTGGACGACAGGGTGGGCAC
380 L G P G G D Q Q S Y E F F H L T L Q A F F T A F F L V L D D R V G T
BstEII (1711)
XmaI (1705)
Tth111I (1851)
1801 TCAGGAGTGTCTCAGGTTCTCCAGGAGTGGATGCCCCCTGCGGGGCGAGCACCAGTCTGCTATCCTCCCTTCTCCCGTTCAGTGCCTGCAGGGC
413 Q E L L R F F Q E W M P P A G A A T T S C Y P P F L P F Q C L Q G
1901 AGTGGTCCGGCGCGGAAGACCTCTTCAAGAACAAGGATCACTTCCAGTTCACCAACCTTCTCCTGTGCGGGCTGTTGTCAAAGCCAAACAGAACTCC
447 S G P A R E D L F K N K D H F Q F T N L F L C G L L S K A K Q K L

2001 TGCGGCATCTGGTCCCGCGGAGCCCTGAGGAGAAAGCGCAAGGCCCTGTGGGCACACCTGTTTTCCAGCCTGCGGGGTACCTGAAGAGCCTGCCCCG
480 L R H L V P A A A L R R K R K A L W A H L F S S L R G Y L K S L P R
Bsu36I (2025)
HindIII (2112)
2101 CGTTCAGTTCGAAAGCTTCAACCAGGTGCAGGCCATGCCACGTTTCATCTGGATGCTGCGCTGCATCTACGAGACACAGAGCCAGAAGTGGGGCAGCTG
513 V Q V E S F N Q V Q A M P T F I W M L R C I Y E T Q S Q K V G Q L
2201 GCGCCAGGGGATCTGCGCCAACTACTCAAGCTGACCTACTGCAACGCTCTCGGCCACTGCAGCGCCCTCTCCTGCTGCTGCTCACTTCCCCA
547 A A R G I C A N Y L K L T Y C N A C T S A D C S A L S F V L H H F P
2301 AGCGGCTGGCCCTAGACCTAGACAACAACAATCTAACGACTACGGCGTGGGGAGCTGCAGCCCTGCTTACGCCCTCACTGTTCTCAGACTCAGCGT
580 K R L A L D L D N N N L N D Y G V R E L Q P C F S R L T V L R L S V
2401 AAACCAGATCACTGACGGTGGGTAAAGGTGCTAAGCGAAGAGCTGACCAAATACAAAATTGTGACCTATTTGGTATACAAACAACAGATCACCGAT
613 N Q I T D G G V K V L S E E L T K Y K I V T Y L G L Y N N Q I T D

2501 GTCGGAGCCAGGTACGTACCAAAATCCTGGATGAATGCAAAGCCTCACGCATCTTAACTGGGAAAAACAAAATAACAAGTGAAGGAGGGGAAGTATC
647▶ V G A R Y V T K I L D E C K G L T H L K L G K N K I T S E G G K Y

2601 TCGCCCTGGCTGTGAAGAACAGCAAATCAATCTCTGAGTTGGGATGTGGGGCAATCAAGTTGGGGATGAAGGAGCAAAGCCTTCGAGAGGCTCTCGC
680▶ L A L A V K N S K S I S E V G M W G N Q V G D E G A K A F A E A L R

XbaI (2796)

2701 GAACCACCCAGCTTGACCACCCTGAGTCTTGCCTCAACGGCATCTCCACAGAAGGAGGAAAGAGCCTTGCAGGGCCCTGCAGCAGAACACGTCTCTA
713▶ N H P S L T T L S L A S N G I S T E G G K S L A R A L Q Q N T S L

2801 GAAATACTGTGGCTGACCCAAAATGAACTCAACGATGAAGTGGCAGAGAGTTTGGCAGAAATGTTGAAAGTCAACCAGACGTTAAAGCATTATGGCTTA
747▶ E I L W L T Q N E L N D E V A E S L A E M L K V N Q T L K H L W L

2901 TCCAGAATCAGATCACAGCTAAGGGGACTGCCAGCTGGCAGATGCGTTACAGAGCAACACTGGCATAACAGAGATTTGCCTAAATGGAAACCTGATAAA
780▶ I Q N Q I T A K G T A Q L A D A L Q S N T G I T E I C L N G N L I K

NheI (3058)

EcoRI (3052) MscI (3064)

3001 ACCAGAGGAGGCCAAAGTCTATGAAGATGAGAAGCGGATTATCTGTTTCTGAGAATTC**CGTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGG**
813▶ P E E A K V Y E D E K R I I C F •

HpaI (3196)

3101 **CAAACCACAAC**TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTA

MfeI (3207) EcoRI (3292)

3201 **ACAACAACAATTGCATT**CATTTTATGTTTCAGGTT**CAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATCTA**
▶

3301 **AAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGCTGTTGCCAATG**

3401 **TGCATTAGCTGTTTGACGCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTTGAAGTACTAGCTCTTCATTTCTTTATGTTTTAAAT**

SspI (3531) SwaI (3545)

3501 **GCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGA**

3601 **TGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAGGAACCTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGC**

3701 **TTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTC**
141▶ • 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

SaeI (3806)

3801 **AGAGATGAGCTCTGACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAG**
108▶ S 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

3901 **TCCTTCTGCCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCCAGCAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCT**
74▶ D 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

4001 **CCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGTGTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCTCTG**
41▶ G 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

BspHI (4120)

XmnI (4112) AseI (4178)

4101 **CTGAGAGATGTTGAAGGCTTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGG**
8▶ Q S I N F T K M ◀

SaeI (4235)

4201 **ATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAG**

SpeI (4333)

4301 **TTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACC**
▶

SnaBI (4461)

4401 **GCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAGGT**

NdeI (4566)

4501 **CATGTAAGTGGCATAATGCCAGGCGGGCATTACCCTGATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTAAGTGGGCA**

4601 **GTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGT**

PacI (4752) BspLU11I (4762)

4701 **TGGCGGTGAGCCAGGCGGGCATTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGT**
▶

4801 **AAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC**

4901 **TATAAAGATACCAGCGTTTTCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTCTCCCTTCGGG**

ApaLI (5076)

5001 **AAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCC**

5101 **GACCGCTGCGCTTATCCGGTAACTATCGTCTTGAAGTCCAACCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG**

5201 CGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT

5301 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAA

5401 AAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAA

PacI (5492)

SwaI (5501) NotI (5511)

5501 CATTAAATCAGCGGCCGAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAC

5601 AAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA