



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA MfeI (82)

101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC

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

301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTAAAGTCAAGTGCAGTGCAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGTCTTTGTTTCGTTT

NcoI (560)
BstEII (555)
AgeI (552)
501 TCTGTTCTGCGCGGTTACAGATCCAAGCTGTGACCGCGGCTACCTGAGATCACCGGTACCATTGGCACAGTTTGACACTGAATACCAGCGCCTAGAGGC
1 M A Q F D T E Y Q R L E A

601 CTCCTATAGCGATTGCCCCCGGGGAAGAGGACCTGTTGGTGCATGTGGCCGAGGGGAGCAAATCACCTTGGCACCACATTGAAAACCTTGACCTCTTC
13 S Y S D S P P G E E D L L V H V A E G S K S P W H H I E N L D L F

XhoI (704) **BspLU11I (739)**
701 TTCTCTCGAGTTTATAATCTCCACCAGAAGAATGGCTTACATGTATGCTCATTGGAGAGATGTTGAACTCATGCAGTTCCTCTTTGTGGTTGCCTTCA
47 F S R V Y N L H Q K N G F T C M L I G E M F E L M Q F L F V V A F

801 CCACCTTCTGTTAGCTGTGGACTATGACATCCTCTTTGCCAACAAAGATGGTGAACCACAGTCTCCATCCTACTGAGCCTGTCAAGGTCACTCTGCC
80 T T F L V S C V D Y D I L F A N K M V N H S L H P T E P V K V T L P

BamHI (989)
901 AGATGCCTTTTTGCTGCCCCAAGTCTGTAGTGCAGGATTGAGAAAATGGCTCTTATCACCATCCTGGTCATTGCTGGTGTCTTGGATCCACCGG
113 D A F L P A Q V C S A R I Q E N G S L I T I L V I A G V F W I H R

BamHI (1067)
1001 CTTATCAAGTTTATCTATAACATTTGCTGCTATTGGGAGATCCACTCCTTCTACCTACATGCTCTTCGGATCCCAATGTCTGCCCTTCCATACTGCACAT
147 L I K F I Y N I C C Y W E I H S F Y L H A L R I P M S A L P Y C T

BsaBI (1146) **BglII (1145)**
1101 GGCAGGAAGTTCAGGCCGATTGTGCAGACGCAGAAAGATCAGATCGATCGCATCCACAAGCGTGAGCTGACAGAGTTGGACATCTACCATCGCATCCT
180 W Q E V Q A R I V Q T Q K E H Q I C I H K R E L T E L D I Y H R I L

1201 ACGTTCCAGAATACATGTTGGCCTGGTGAACAAATCCTCCTGCCTGCGCTTCGGTCTGCCCGCCTCGGAGAGTTGTCTTCTTCCACCGTGGC
213 R F Q N Y M V A L V N K S L L P L R F R L P G L G E V V F F T R G

SmaI (1332)
1301 CTCAAGTACAACCTTTGAGCTCATCCTCTTCTGGGACCCGCTCTGTTTCTCAATGAGTGGAGCCTCAAGGCCGAGTACAAACGTGGAGGGCAACGGC
247 L K Y N F E L I L F W G P G S L F L N E W S L K A E Y K R G G Q R

BsaBI (1437)
1401 TAGAGCTGGCCAGCGTCTCAGCAACCGCATCTTGGATTGGCAGTCCCAACTTCTGCTGTGTCCCCTCATCCTCATCTGGCAGATCCTCTATGCCTT
280 L E L A Q R L S N R I L W I G I A N F L L C P L I L I W Q I L Y A F

1501 CTTAGCTATGCCGAGGTGCTGAAGAGAGAGCCGGGGCCTGGGAGCGGTTGCTGCTCACTCTATGGCCGTTGTTACCTCCGCCACTTCAATGAGCTG
313 F S Y A E V L K R E P G A L G A R C W S L Y G R C Y L R H F N E L

BstAPI (1602)
1601 GAGCATGAGCTGCAGTCCCGCTCAACCGAGGCTACAAGCCCGCTCCAAGTACATGAATTGCTTCTTGTACCAGTGTGACTCTGCTGGCCAAGAATG
347 E H E L Q S R L N R G Y K P A S K Y M N C F L S P L L T L L A K N

BstEII (1791)
1701 GTGCCCTTCTCGTGGCTCTATCCTGGCTGTGCTTATTGCCCTCACCATCTATGATGAAGATGTGTTAGCTGTGGAACACGCTCCTCACCACGGTACCCT
380 G A F F A G S I L A V L I A L T I Y D E D V L A V E H V L T T V T L

DraIII (1846)
1801 CCTGGGAGTACGGTGTGTCAGGTCCTTTCATCCAGACAGCATGTTGTTCTGCCCGAGCAGCTGCTCCGAGTGATTCTTGCACACATCCAC
413 L G V T V T V C R S F I P D Q H M V F C P E Q L L R V I L A H I H

1901 TACATGCCGACTGAGGTAATGCCACCGCTCGAGACCCGGGACGAGTTTGGCCAGCTCTCCAGTACAAGGCAAGTGTTCATCTTGGAGGAGT
447 Y M P D H W Q G N A H R S Q T R D E F A Q L F Q Y K A V F I L E E

2001 TGCTGAGTCCCATCGTCACACCCCTATTCTCATCTTTCGCTCCGCCCTGGGAGTCAATGACTTCTCCGCAACTTACGGTCCGAGGTCGT
480 L L S P I V T P L I L I F C L R P R A L E I I D F F R N F T V E V V

2101 GGGTGTGGGAGACCTGCTCCTTTGCTCAGATGGACGTTCCGACATGGCCATCCTCAGTGGCTGTCTGGAGGGCAGACAGAGGCTCAGTGTACCAG
513 G V G D T C S F A Q M D V R Q H G H P Q W L S G G Q T E A S V Y Q

2201 CAAGCCGAGGAGGGAAGCTGAGTGTGCTCATGCACTTGGCAATCCCGGCTGGCAGCCCTCGTGAGAGCACAGCTTCTGGGCTTCC
547 Q A E D G K T E L S L M H F A I T N P G W Q P P R E S T A F L G F

2301 TCAAGGAGCAGGTGCAGCGAGATGGAGCAGTGTGCTGGCCAGGGTGGCTGCTCCCGAGAATGCCCTTTCACATCCATCCAGTCTTACAGTC
580 L K E Q V Q R D G A A A G L A Q G G L L P E N A L F T S I Q S L Q S

SdaI (2475)
2401 TGAGTCCGAGCCACTGAGCCTTATTGCAATGTGGTAGCAGGCTCATCCTGCCGAGGACCTCACTGTCCAGAGACCTGACGGCTCCAGGCACAGGGCT
613 E S E P L S L I A N V V A G S S C R G P S L S R D L Q G G S R H R A

2501 GATGTTGCTTCTGCCCTTCGATCCTTCTCCCTCTGCAGCCTGGAGGGCCCTCAAGGCCGGTTCCAGTACCATGACAGGCTCTGGAGTGGATGCCA
647 D V A S A L R S F S P L Q P G A A P Q G R V P S T M T G S G V D A

MluI (2663) **BsrGI (2693)**
2601 GGACAGCCAGCTCTGGGAGTAGCGTGTGGGAAGGACAGCTGCAGAGCCTGGTGTGTCGAATACCGTCCACCGAGATGAGCCTGCACGCCCTGTACAT
680 R T A S S G S S V W E G Q L Q S L V L S E Y A S T E M S L H A L Y M

2701 GCACCAGCTCCACAAGCAGCAGACCCAGGCTGAGCCCGAGCGGATGTGTGGCACCAGGGGAGAGTGTGAGAGTGGAGAGTGGCCCTGAAGAGGGG
713 H Q L H K Q Q T Q A E P E R H V W H R R E S D E S G E S A P E E G

SrfI (2809)
2801 GGAGAGGGTGGCCGGGCCCCCAACCCATCCCCCGCTCGGCCAGCTATCCCTGTGTACACCCCGCCTGGAGCACCTGAGACCACCGCCCTGCATGGG
747▶ G E G A R A P Q P I P R S A S Y P C A T P R P G A P E T T A L H G
2901 GCTTCAGAGGCGCTACGGGGGCATCACAGATCCTGGCAGAGTCCCCCGTGGCCCTCTCACTTCTCCAGGCTGCCCTGGGAGGATGGGCAGAAGATGG
780▶ G F Q R R Y G G I T D P G T V P R G P S H F S R L P L G G W A E D G

ApaLI (3067) **NheI (3086)**
3001 CCAGCCAGCATCAAGGCACCCAGAGCCGGTGCCAGAGGAGGGCTCAGAGGATGAACTCCCCCTCAAGTGACAAAGGTATAGGCAAGCTAGCTGGCCAGA
813▶ Q P A S R H P E P V P E E G S E D E L P P Q V H K V •
3101 CATGATAAGATACATTGATGAGTTGGACAAACCAACTAGAAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTA

HpaI (3224) MfeI (3235)
3201 ACCATTATAAGCTGCAATAACAAGTTAAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAA

EcoRI (3320)
3301 ACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAG
3401 GCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTGCAGCCTCACCTCTTTCATGGAGTTAAGATATAGTGTATTTCCCAAGGTTTGAAC

SspI (3559) SmaI (3573)
3501 TAGCTCTTCATTTCTTTATGTTTAAATGCACTGACCTCCACATTCCCTTTTATGATAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAT
3601 AAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAA
3701 TTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGGGGGATGAGTTCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAA
1414 • N R T Y K L P I L E E I T T K V L K G N M E I
3801 TGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGCCACCTCATCAGAGTAGGG
117▶ 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 D V E D S Y P
3901 GTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCAGACAGATGACCCTGCCAATGTAG
84▶ 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 V T V R G I Y
4001 GCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTGGTCCGATGGCCGCCCGACATGGTGTCTTGTCTCATAGAGCATGGTATCTTCTCAG
50▶ 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 Y L M T I K E T
4101 TGGCAGCTCCACAGCTCAGATCTGCTGAGAGATGTTGAAGTCTTATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATG
17▶ A V E V L E L D Q Q S I N F T K M

AseI (4206)
4201 CCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCT

SpeI (4361)
4301 ACCGCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGG

SnaBI (4489)
4401 AGACTTGGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGT

NdeI (4594)
4501 ACTGCAAGTAGGAAAGTCCATAAGGTCATGTACTGGCATAATGCCAGGCGGGCCATTTACCGTCAATGAGGGGCGTACTTGGCATATG
4601 ATACACTTGTACTGCCAAGTGGGCGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTC

SdaI (4772) PacI (4780) BspLU11I (4790)
4701 ATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTACGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCCTGCAGGTTAAATAAGAACATGTGAGC
4801 AAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCGCGTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCA
4901 AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCTGGAAGTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCG
5001 GATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGG

ApaLI (5104)
5101 CTGTGTGCACGAACCCCGTTACGCCGACCGCTGCGCCTTATCCGTAACATATCGTCTTGTGTCACCCCGTAAGACACGACTTATCGCCACTGGCA
5201 GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTAT
5301 TTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAACAAACCCGCTGGTAGCGGTGTTTTTTTGT
5401 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAA

EagI (5540)
PacI (5520) SmaI (5529) NotI (5539)
5501 GGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGAATAAAATATCTTTATTTTACATTACATCTGTGTGTTGTTTTTGTGTAATC
5601 GTAACATAACGCTCTCCATCAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGACAGGTGCCAGAACATTTCTCTAT
5701 CGAA