



PvuI (7)
SgfI (6)
1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA

101 GAGAAAGTGGCGGGGTAACGGAAAGTGTGTCTGTACTGGCTCCGCCTTTTTCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC

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

301 GCCATCCACGCCGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCGCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGTCTTTGTTTCGTTT

KasI (535) **AgeI (552)** **SphI (560)**
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTA CTTGAGATCACCGGTGACAGCGAGCCGATTCCGAGCAGCCCTCCAAGCGTCC
1 M Q R A D S E Q P S K R P

XmaI (660) **BspEI**
601 CCGTTGCGATGACAGCCGAGAACCCCTCAAACACCCCTTCCGAGAGGCAGACTGGTCCCGGGCCTGGAACCTCCATCCCGACTACAAGACATGGGGT
13 R C D D S P R T P S N T P S A E A D W S P G L E L H P D Y K T W G

Bsu36I (720)
701 CCGGAGCAGGTGTGCTCCTTCTCAGCGCGGTGGCTTTGAAGAGCCGGTGTGCTGAAGAACATCCGAGAAAATGAAATCACAGGCGCATTACTGCCTT
47 P E Q V C S F L R R G G F E E P V L L K N I R E N E I T G A L L P

801 GTCTTGATGAGTCTGTTTTGAAAATCTTGAGTAAGTTCTTGGGGGAGAGGAAGCTGCTTAGTTATATCCAGCGATTGGTTCAAATCCACGTTGA
80 C L D E S R F E N L G V S S L G E R K K L L S Y I Q R L V Q I H V D

NcoI (927) **BbsI (989)**
901 TACAATGAAGTAATTAATGATCCTATCCATGGCCACATTGAGCTCCACCTCTCCTCGTCCGAATCATTGATACACCTCAATTTCAACGTCTTCGATAC
113 T M K V I N D P I H G H I E L H P L L V R I I D T P Q F Q R L R Y

ClaI (1049)
1001 ATCAAACAGCTGGGAGGTGGTTACTATGTTTTTCCAGGAGCTTCGCACAATCGATTTGAGCATAGTCTAGGGGTGGGGTATCTAGCAGGATGTCTAGTTC
147 I K Q L G G G Y Y V F P G A S H N R F E H S L G V G Y L A G C L V

DraIII (1103) **Tth111I (1169)**
1101 ACGCACTGGGTGAAAAACAACAGAGCTGCAGATAAGTGAACGAGATGTTCTCTGTGTTGAGATTGCTGGACTTTGTGATGATCTCGGTGATGGCCATT
180 H A L G E K Q P E L Q I S E R D V L C V Q I A G L C H D L G H G P F

1201 TTCTCACATGTTTGTAGGACGATTTATCCACTTGCTCGCCCGAGGTGAAATGGACGCATGAACAAGGCTCAGTTATGATGTTTGGACACCTTATTAAT
213 S H M F D G R F I P L A R P E V K W T H E Q G S V M M F E H L I N

1301 TCTAATGGAATTAAGCCTGTGATGGAACAATATGGTCTCATCCCTGAAGAAGATTTGCTTTATAAAGGAACAAATTTAGGACCACCTTGAATCACCTG
247 S N G I K P V M E Q Y G L I P E E D I C F I K E Q I V G P L E S P

HindIII (1443)
1401 TAGAAGATTCATTGTGGCCATATAAAGGGCGTCTGAAAAACAAAAGCTTCTTTATGAGATAGTATCTAATAAAGAAATGGCATTGATGTGGACAAATG
280 V E D S L W P Y K G R P E N K S F L Y E I V S N K R N G I D V D K W

BstXI (1521) **Eco47III (1556)** **AfeI (1556)**
1501 GGATTATTTTGGCAGGGACTGCCATCATCTTGGAAATCCAAAATAATTTTGATTACAAGCGCTTTATTAAGTTTGGCCGTGCTGTGAAGTAGACAATGAG
313 D Y F A A R D C H H L G I Q N N F D Y K R F I K F A R V C E V D N E

1601 TTGCGTATTTGTGCTAGAGATAAGGAAGTTGAAATCTGTATGACATGTTCCACACTCGCAACTCTTTACACCGTAGAGCTTATCAACACAAAGTTGGCA
347 L R I C A R D K E V G N L Y D M F H T R N S L H R R A Y Q H K V G

1701 ACATTATTGATACAATGATTACAGATGCTTCTCAAAGCAGATGACTACATAGAGATTACAGGTGCTGGAGA AAAAAAGTATCGCATTCTACAGCAAT
380 N I I D T M I T D A F L K A D D Y I E I T G A G E K K Y R I S T A I

1801 TGACGACATGGAAGCCTATACTAAGCTGACAGATAACATTTTTCTGGAGATTTTATACTACTGATCCCAAATTTGAAAGACGCACGAGAGATTTAAAA
413 D D M E A Y T K L T D N I F L E I L Y S T D P K L K D A R E I L K

XmnI (1981)
1901 CAAATGAATACCATACTATTCAAGTATGTGGGTGAGACGCAGCCAACAGGACAAATAAAGATTAAGGGAGGACTATGAATCTCTTCCAAAAGAGG
447 Q I E Y R N L F K Y V G E T Q P T G Q I K I K R E D Y E S L P K E

2001 TTGCCAGTGCTAAACCCAAAGTATTGCTAGACGTGAAACTGAAGGCTGAAGATTTTATAGTGGATGTTATCAACATGGATTATGGAATGCAAGAAAAGAA
480 V A S A K P K V L L D V K L K A E D F I V D V I N M D Y G M Q E K N

2101 TCCAATTTGATCATGTTAGCTTCTATTGTAAGACTGCCCAACAGAGCAATCAGGATTAATAAACCAGGTTTCACTTCTGCCAGAGAAATTTGTA
513 P I D H V S F Y C K T A P N R A I R I T K N Q V S Q L L P E K F V

SspI (2261)
2201 GAGCAGCTGATTCGAGTATATTGTAAGAAGGTGGACAGAAAGAGTTTGTATGCCGCAAGACAATTTTGTTCAGTGGTGTGCAGACAGAAATTTACCA
547 E Q L I R V Y C K K V D R K S L Y A A R Q Y F V Q W C A D R N F T

ScaI (2362)
2301 AGCCGAGGATGGCGATGTTATAGCCCACTCATAACACCTCAAAAAAGGAATGGAACGACAGTACTTCACTCCAAAATCCAACCTCGCTCCGAGAAGC
580 K P Q D G D V I A P L I T P Q K K E W N D S T S V Q N P T R L R E A

NheI (2482)
2401 ATCCAAAAGCAGAGTCCAGCTTTTTAAAGATGACCCAATGTGAATGTCTGTAGTCAGTTGTTTACAAACTCCCTCTCCTGCAGCTAGCTGGCCAGACATG
613 S K S R V Q L F K D D P M •

2501 ATAAGATACATTGATGAGTTTGACAAAACCAACTAGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCA

HpaI (2620)
2601 TTATAAGTGAATAAACAAGTTAACAACAACAAATGCATTCTTTTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAACCT

2701 **EcoRI (2716)**
 CTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCAT
 2801 AGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACG

2901 **SspI (2955)** **SwaI (2969)**
 TCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAATAAT

3001 **EcoO109I (3030)**
 GTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGG

3101 ACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAG
 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

3201 **BstXI (3259)**
 CACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGC
 116 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 H

3301 **StuI (3394)**
 CTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTGACCTGCCAATGTAGGCCT
 82 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 A E

3401 CAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTGTTGTCCTCATAGAGCATGGTGATCTTCTCAGTGGC
 49 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 A

3501 **BspHI (3544)** **XmnI (3536)**
 GACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGCTTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGA
 16 V E V L E L D Q Q S I N F T K M

3601 TGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCG

3701 **SpeI (3757)**
 CCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAAACAAACTCCCATTGACGTCAATGGGGTGGAGAC

3801 **SnaBI (3885)**
 TTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGGATGACTAATACGTAGATGACTG

3901 **NdeI (3990)**
 CCAAGTAGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATAC

4001 ACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTA

4101 **SdaI (4168)** **PacI (4176)**
 TTGACGTCAATGGGCGGGGTGCTTGGGCGGTACGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAA

4201 GGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTC

4301 AGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGCGTTTTCCCTGGAAGTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGCTTACCGGATA

4401 CCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGT

4501 **ApaLI (4500)**
 GTGCACGAACCCCCGTTACGCCGACCGCTGCGCTTATCCGGTAACATCGTCTTGTAGTCCAAACCGGTAAGACACGACTTATCGCCACTGGCAGCAG

4601 CCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGG

4701 TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGATTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGGGTGGTTTTTTTGTTC

4801 AAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGGGA

4901 **EagI (4936)** **PacI (4916)** **SwaI (4925)** **NotI (4935)**
 TTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGGCCGAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTA

5000 ACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAAGTGCCAGAGTCCAGAACATTTCTCTATCGA
 5100 A