



PvuI (7)
SgfI (6) **MfeI (82)**
1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGTGCCTA
101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGTGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) **HindIII (245)**
PvuII (239) **Bsu36I (291)**
201 GTGAACGTTCTTTTTTCGCAACGGGTTTGGCCGACAGTGAAGCTTCAGAGGGCTCGCATCTCTCTTACAGCGCCCGCCGCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTGCAGTCTGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTTGTTCGTTT

KasI (535) **NcoI (560)**
AgeI (552) **BstEII (555)**
MluI (587)
501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTACCATTGGCCTCGTACCCCGCCATCAACACGCGTCTGCGTT
1 M A S Y P G H Q H A S A F

BsiWI (637) **BstAPI (690)**
601 CGACCAGGCTGCGGTTCTCGCGCCATAGCAACCGACGTACGGCTGCGCCCTCGCCGCGCAGCAAGGCCACGGAAGTCCGCCCCGAGCAGAAAAATG
13 D Q A A R S R G H S N R R T A L R P R R Q Q E A T E V R P E Q K M

XcmI (755) **EcoRV (790)** **SnaBI (799)**
701 CCCACGCTACTGCGGTTTATATAGACGGTCCCCACGGGATGGGAAAACCACCACCAGCAACTGCTGGTGGCCCTGGGTTCCGCGGACGATATCGTCT
47 P T L L R V Y I D G P H G M G K T T T T Q L L V A L G S R D D I V

NruI (851) **EcoRV (894)**
801 ACGTACCCGAGCCGATGACTTACTGGCGGTGCTGGGGGCTTCCGAGACAATCGGAACATCTACACCACACAACACCCGCTCGACCAGGGTGAGATATC
80 Y V P E P M T Y W R V L G A S E T I A N I Y T T Q H R L D Q G E I S

SphI (947)
901 GGCCGGGACGCGCGGTGTAATGACAAGCGCCAGATAACAATGGGCATGCTTATGCCGTGACCGACGCCGTTTGGCTCCTCATATCGGGGGGAG
113 A G D A A V V M T S A Q I T M G M P Y A V T D A V L A P H I G G E

Acc65I (1088)
1001 GCTGGGAGCTCACATGCCCGCCCGCCCTCACCTCATCTTGCACCGCCATCCCATCGCCGCCCTCCTGTGCTACCCGGCCGCGGGTACCTTATGG
147 A G S S H A P P P A L T L I F D R H P I A A L L C Y P A A R Y L M

BspEI (1185)
1101 GCAGCATGACCCCCAGGCCGTGCTGGCGTTCTGGCCCTCATCCCGCCGACCTTGGCCGGCACCAACATCGTGTGGGGCCCTTCCGGAGGACAGACA
180 G S M T P Q A V L A F V A L I P P T L P G T N I V L G A L P E D R H

BsrBI (1234)
1201 CATCGACCGCTGGCCAAACGCCAGCGCCCGGCGGCTGGACTGGCTATGCTGGCTGCGATTGCGCCGTTTTACGGGCTACTTGCCAATACGGTG
213 I D R L A K R Q R P G E R L D L A M L A A I R R V Y G L L A N T V

PstI (1306)
1301 CGGTATCTGCAGTGGCGGGTCTGGCGGGAGGACTGGGGACAGCTTTCGGGGACGGCCGTGCCGCCCCAGGGTGGCAGCCCCAGAGCAACGCGGGCC
247 R Y L Q C G G S W R E D W G Q L S G T A V P P Q G A E P Q S N A G

1401 CACGACCCATATCGGGGACGTTATTTACCCTGTTTCGGGCCCCGAGTTGCTGGCCCCAACGGCGACCTGTATAACGTGTTTGCCTGGGCTTGGGA
280 P R P H I G D T L F T L F R A P E L L A P N G D L Y N V F A W A L D

1501 CGTCTGGCCAAACGCTCCGTTCCATGCAGTCTTTATCCTGGATTACGACCAATCGCCCGCGGCTGCGGGACGCCCTGCTGCAACTTACCTCCGGG
313 V L A K R L R S M H V F I L D Y D Q S P A G C R D A L L Q L T S G

PshAI (1608) **BssHII (1654)**
1601 ATGGTCCAGACCCACGTCAACCCCGGCTCCATACCGACGATATGGACCTGGCGGCACGTTTGGCCGTGAGATGATCAGCGGAGCTAATGGCGTCA
347 M V Q T H V T T P G S I P T I C D L A R T F A R E M I S G A N G V

BssHII (1735) **BsrBI (1751)** **XmaI (1787)**
1701 TGGCCAAGTTGACCAAGTCCCGTTCGGTGCACCGCGCGACGCTCGCCGGAGCGGTGAGTTCTGGACCGACCGGCTCGGGTTCTCCCGGACTTCGT
380 M A K L T S A V P V L T A R D V A G A V E F W T D R L G F S R D F V

SgrAI (1812)
1801 GGAGGACGACTTCGCGGTGTGGTCCGGGACGACGTGACCTGTTTCATCAGCGCGGTCCAGGACCAGGTGGTCCGGACAACACCTGGCCTGGGTGTGG
413 E D D F A G V V R D D V T L F I S A V Q D Q V V P D N T L A W V W

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

DraIII (2047) **RsrII (2094)**
ApaLI (2044)
2001 CGTGGGGCGGGAGTTCCGCTGCGCGACCCGGCCGCAACTGCTGCACTTCGTGGCCGAGGAGCAGGACTGACCGACGCCACCAACCCCGGCTCC
480 P W G R E F A L R D P A G N C V H F V A E E Q D •

ClaI (2135) **NheI (2147)**
StuI (2120) **BglII (2129)** **BsrGI (2141)**
2101 GACGC GGCCCGACGGTCCGAGGCCTCGGAGATCTATCGATTGTACAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTGGACAAACCACAA

HpaI (2285) **MfeI (2296)**
2200 CTAGAATGCAGTAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTAAACAACAACA

EcoRI (2381)
2300 TTGCATTATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCATAAATACAGCA

2400 TAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCT

2500 **SapI (2563)**
 GTTTGCAGCCTCACCTTCTTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAAGTCTCTTCATTTCTTTATGTTTTAAATGCACTGACCT

2600 **SspI (2620)** **SwaI (2634)**
 CCCACATTCCCTTTTATGAAAATATTCAGAAAATATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGC

2700 CCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCT
 141 • N R

2800 GGTGACTTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGC
 138 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 L

2900 **BstXI (2924)**
 TCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCCTTCTGCC
 104 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 Q G

3000 **StuI (3059)**
 CGTTGCTCACAGCAGACCAATGGCAATGGCTTACGACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTT
 71 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 T K

3100 GGCCTGATGGCCGCCGACATGGTGTGTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACTCCACCAGCTCCAGATCTGTGAGAGATG
 38 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 I

3200 **BbsI (3205)** **XmnI (3201)** **AseI (3267)**
 TTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACAGCGTGGATGGCGTCTC
 4 N F T K M

3300 CAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGAC

3400 **SpeI (3422)**
 ATTTTGGAAAGTCCCCTTGGATTACTAGTCAAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCAGC

3500 **SnaBI (3550)**
 CCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCCATAAAGTCAATGTAAGT

3600 **NdeI (3655)**
 GCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTAAGTCCCAAGTGGGCGATTTACCGTA

3700 AATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTCA

3800 **PacI (3841)** **PstI (3834)** **SdaI (3833)** **BspLU11I (3851)**
 GCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAAGTTAATAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCG
 4

3900 CGTTGCTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA

4000 CCAGGCGTTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCG

4100 **ApaI (4165)**
 CTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTGCTCCAAGTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCG

4200 CTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGT

4300 AGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAAGTATTTGGTATCTGCGCTCTGTGAAGCCAGTTACCTTCGGA

4400 AAAAGAGTTGGTAGCTCTTGTATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTC

4500 **PacI (4581)** **SwaI (4590)**
 AAGAAGATCCTTTGATCTTTTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTATGGCTAGTTAATTAACATTTAAATC

4600 **NotI (4600)**
 AGCGGCGCAATAAAATATCTTTATTTTATTACATCTGTGTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAAC

4700 AAAACAACTAGCAAAATAGGCTGTCCCGAGTCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA