



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82) **EcoNI (96)**
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) **HindIII (245)** **Bsu36I (291)**
EcoNI (287) 201 GTGAACGTTCTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACCGCCGCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCGCGTTCTGCCGCTCCCGCCTGTGGTGCCTCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441) 401 GGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

NcoI (560) **BstEII (555)**
KasI (535) **AgeI (552)** 501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCCCTACCTGAGATCACCGGTCCACCATGGTGCGGTCCGGGAATAAGGCAGCTGTTGTGCTGTG
1 M V R S G N K A A V V L C
BsrGI (684) 601 TATGGACGTGGGCTTTACCATGAGTAATCCATTCTGGTATAGAATCCCATTTGAACAAGCAAAGAAGGTGATAACCATGTTTGTACAGCGACAGGTG
13 M D V G F T M S N S I P G I E S P F E Q A K K V I T M F V Q R Q V

BsaBI (775) **ApaLI (796)** 701 TTTGCTGAGAACAAGGATGAGATTGCTTTAGTCTGTTTGGTACAGATGGCACTGACAATCCCCCTTCTGGTGGGGATCAGTATCAGAATCAGACATCAGTGC
47 F A E N K D E I A L V L F G T D G T D N P L S G G D Q Y Q N I T V
801 ACAGACATCTGATGCTACCAGATTTTGAATTTGCTGGAGACATTGAAAGCAAATCCAACAGGTTCTCAACAGGCTGACTTCTGGATGCACATAATCGT
80 H R H L M L P D F D L L E D I E S K I Q P G S Q Q A D F L D A L I V

SplI (960) 901 GAGCATGGATGTGATTCAACATGAAACAATAGGAAAGAAGTTTGAAGAAGGCATATTGAAATATTCACTGACCTCAGCAGCCGATTGAGCAAAAGTCAG
113 S M D V I Q H E T I G K K F E K R H I E I F T D L S S R F S K S Q
1001 CTGGATATTATAATTCATAGCTTGAAGAAATGTGACATCTCCCTGCAATTTCTTCTGCTTTCTCACTGGCAAGGAAGTGAAGTGGGACAGAGGAG
147 L D I I I H S L K K C D I S L Q F F L P F S L G K E D G S G D R G

Neol (1122) **EcoRV (1197)** 1101 ATGGCCCCCTTCGCTTAGGTGGCCATGGGCCTTCCCTTCCACTAAAAGGAATTACCGAACAGCAAAAAGAAGTCTTGAGATAGTGAAAATGGTATGAT
180 D G P F R L G G H G P S F P L K G I T E Q Q K E G L E I V K M V M I
1201 ATCTTTAGAGGTGAAGATGGGTTGGATGAAATTTATTCATTCAGTGAGAGTCTGAGAAAAGTGTGCGTCTTCAAGAAAATTGAGAGGCATTCCATTCAC
213 S L E G E D G L D E I Y S F S E S L R K L C V F K K I E R H S I H

ClaI (1356) 1301 TGGCCTGCCGACTGACCATTGGCTCCAATTTGTCTATAAGGATTGCAGCCTATAAATCGATTCTACAGGAGAGAGTTAAAAAGACTTGGACAGTTGTGG
247 W P C R L T I G S N L S I R I A A Y K S I L Q E R V K K T W T V V
1401 ATGCAAAAACCTAAAAAAGAAGATATACAAAAGAAGCAGTTTATTGCTTAAATGATGATGATGAAACTGAAGTTTTAAAAGAGGATATTATTCAAGG
280 D A K T L K K E D I Q K E T V Y C L N D D D E T E V L K E D I I Q G

XmnI (1572) 1501 GTTCCGCTATGGAAGTGATATAGTTCCTTTCTCAAAGTGGATGAGGAACAAATGAAATATAAATCGGAGGGGAAGTCTTCTGTTTTGGGATTTTGT
313 F R Y G S D I V P F S K V D E E Q M K Y K S E G K C F S V L G F C

PstI (1679) 1601 AAATCTTCTCAGGTTTCAAGAGATTCTTCATGGGAAATCAAGTTCTAAAGTCTTTGCAGCAAGAGATGATGAGGCAGCTGCAGTTGCACTTTCTCC
347 K S S Q V Q R R F F M G N Q V L K V F A A R D D E A A A V A L S S
1701 TGATTCATGCTTTGGATGACTTAGACATGGTGGCCATAGTTCGATATGCTTATGACAAAAGAGCTAATCCTCAAGTCGGCGTGGCTTTTCTCATATCAA
380 L I H A L D D L D M V A I V R Y A Y D K R A N P Q V G V A F P H I K

BspLU11I (1859) 1801 GCATAACTATGAGTGTGTTAGTGTATGTGCAGCTGCCTTTTCATGGAAGACTTGGGCAATACATGTTTTCATCCTTAAAAACAGTAAGAATATGCTCCC
413 H N Y E C L V Y V Q L P F M E D L R Q Y M F S S L K N S K K Y A P
1901 ACCGAGGCACAGTTGAATGCTGTTGATGCTTTGATTGACTCCATGAGCTTGGCAAAGAAGATGAGAAGACAGACACCCTTGAAGACTGTTTCCAACCA
447 T E A Q L N A V D A L I D S M S L A K K D E K T D T L E D L F P T

XmaI (2063) 2001 CAAAATCCAAATCCTCGATTTTCAAGATTATTTTCAAGTGTCTGCTGCACAGAGCTTTACATCCCGGAGCCTACACCCCAATTCAGCAGCATATTTG
480 T K I P N P R F Q R L F Q C L L H R A L H P R E P L P P I Q Q H I W
2101 GAATATGCTGAATCCTCCGCTGAGGTGACAACGAAAAGTCAAGTTCCTCTCTCTAAAATAAAGACCTTTTCTCTGATTGAAGCCAAGAAAAGGAT
513 N M L N P P A E V T T K S Q I P L S K I K T L F P L I E A K K K D
2201 CAAGTACTGCTCAGGAAATTTTCAAGACAACCATGAAGATGGACCTACAGCTAAAATTAAGACTGAGCAAGGGGAGCCACTTCAGCTCTCCA
547 Q V T A Q E I F Q D N H E D G P T A K K L K T E Q G G A H F S V S
2301 GTCTGGCTGAAGGCAAGTGCACCTCTGTTGGAAGTGTGAATCTGCTGAAAACCTCCGTGTTCTAGTGAACAGAAGAAGGCCAGCTTTGAGGAAGCGAG
580 S L A E G S V T S V G S V N P A E N F R V L V K Q K K A S F E E A S
2401 TAACCAGCTCATAAATCAGTCAACAGTTTTGGATACTAATGAAACACCGTATTTTATGAAGAGCATAGACTGCATCCGAGCCTCCGGGAAGAAGCC
613 N Q L I N H I E Q F L D T N E T P Y F M K S I D C I R A F R E E A

Eco47III (2519) 2501 ATTAAGTTTTTCAAGAGCAGCGCTTAAACAATCTCTGAAAGCCCTTCAAGAGAAAGTGAATTAACAATTAATCATTCTGGGAAATTGTTGTCC
647 I K F S E E Q R F N N F L K A L Q E K V E I K Q L N H F W E I V V

XcmI (2663) 2601 AGGATGGAATTAATCTGATCACCAAGAGGAAGCCTCTGGAAGTCTGTGACAGCTGAGGAAGCAAAAAGTTTCTGGCCCCAAGACAACCAAGTGG
680 Q D G I T L I T K E E A S G S S V T A E E A K K F L A P K D K P S G

2701 AGACACAGCAGCTGTATTTGAAGAAGGTGGTGTATGTGGACGATTTATTGGACATGATATAGGTCGTGGATGTATGGGGAATCGCTAGCTGGCCAGACATG
713▶ D T A A V F E E G G D V D D L L D M I •

2801 ATAAGATACATTGATGAGTTTGGACAAACCACAACACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCA

2901 **HpaI (2920)** MfeI (2931)
TTATAAGCTGCAATAAACCAAGTTAACAAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCT

3001 **EcoRI (3016)**
CTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCAT

3101 AGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAAGTACG

3201 **SspI (3255)** **SwaI (3269)**
TCTTCATTTCTTTATGTTTTAAATGCACCTGCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAATAAT

3301 **EcoO109I (3330)**
GTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGG

3401 ACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTTCATCTCAATGAG
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

3501 **SacI (3530)**
CACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGC
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

3601 **StuI (3694)**
CTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCT
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

3701 CAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTTCCTCATAGAGCATGGTGTCTTCTCAGTGGC
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

3801 **XmnI (3836)**
GACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGCTTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGA
16◀ V E V L E L D Q Q S I N F T K M

3901 **AseI (3902)** **SacI (3959)**
TGATTAATTGTCAAACACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACC

4000 **SpeI (4057)**
GCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCATTGACGTCAATGGGGTGGAG

4099 **SnaBI (4185)**
ACTTGAAATCCCGTGAGTCAAACCGTATCCACGCCATTGATGTAAGTCCGAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAC

4199 **NdeI (4290)**
TGCCAAGTAGGAAAGTCCATAAGTTCATGTAAGTCCGAGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGCTACTTGGCATATGAT

4299 ACACCTGATGTAAGTCCGAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAT

4399 **PstI (4469)** **SdaI (4468)** **PacI (4476)** **BspLU11I (4486)**
TATTGACGTCAATGGGCGGGGCTGTTGGGCGGTGAGCCAGCGGGCCATTTACCGTAAAGTTATGTAACGCCTGCAGGTTAA TTAAGAACATGTGAGC
4497 AAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCA
4597 AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGCGTTTTCCCTGGAAGTCCCTCGTGCCTCTCTGTTCCGACCTGCCGTTACCG
4697 GATACCTGTCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGATAGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGG

4797 **ApaLI (4800)**
CTGTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCCTTATCCGTAAGTATCGTCTTGTGCAACCCGTAAGACACGACTTATCGCCACTGGCA

4897 GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTAT

4997 TTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTGT

5097 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACTCACGTTAA

5197 **EagI (5236)** **PacI (5216)** **SwaI (5225)** **NotI (5235)**
GGGATTTTGGTCAATGGCTAGTTAATTAACATTTAAATC AGCGGCCGAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGTTTTTTGTGTGAATC

5297 GTAACAAACATACGCTCTCCATCAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTAT

5397 CGAA