



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

HindIII (245)
Psp1406I (203) 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCGCCCTACCTGAGGGC
PvuII (239)
Bsu36I (291)
301 GCCATCCACGCCGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

BstEII (555)
AgeI (552) 501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATGGCAGGAGCTACAATGAACTCAGCAGCAGGACAAT
NcoI (560)
1 M A G G Y N E L S S R T I

BsaBI (628) 601 CCCAGAGGACCAGAATGAAGAGGAGTCAGATGACTATCAGCTCATGTTCAAGCACTTCAAAGAAAACAAGGTGGAGATCGCAAGTGCCATCACCAAGCCG
13 P E D Q N E E E S D D Y Q L M F K H F K E N K V E I A S A I T K P
BspHI (710) 701 TTTCTTTCTCATGAGCCTTCGAGACCGAGACTTCATCTCGGAGCAGAAGTTTCAGGAATATCAAGAAAACATGTAAGAACCCTGGTCCAGTGAGAGAG
47 F P F L M S L R D R D F I S E Q K F Q E Y Q E T C K N L V P V E R
BspLU11I (769)
801 TCGTGTACGACATCCTCAGCAATGTACAGAAGAAGTTTAGCCGGGATCTTCTGAAAGTGTATTCAGCAAAACACATCTGAAGGCCTACCCCGACTTAAA
80 V V Y D I L S N V Q K K F S R D L L K V I F S K T H L K A Y P D L K
StuI (881)
901 GGAGACTCTAAAACACTTCTTCTAAATGCATCCAAAACGAATGATGAACAAGCAGAGGAGATGCAATGCAATGGAGGAGAGGGAAGTAATTCTGTGAA
113 E T L K H F F L N A S K T N D E Q A E E M Q C N G G E G S N S C E
NsiI (926)
1001 CAAACAAGACAACAAGACAAGAGCCCGAGGATGACCTGCCCTCATCCCTGAGACAAGAAGCAGGAGCACAGCAACCCACAGTGAAGGAAAGTGTCTCT
147 Q T R Q T R Q E P Q D D L P S S L R Q E A G A Q Q P T R E K K C S
1101 GTGTCATATGTTCCCAACATATGTGCCAGAAGACCTGGAAGCAAGGATGGGAAGCAGCAAGGAGAATGTGTGGGCACCAGGAGGAACATTCCTGCAAG
180 C V I C S P T Y V P E D L E A R M G S S Q G E C V G T R R N I P A R
1201 AAAACCCAAGAAGCAGAGAAGAAAAAGAAAGGCCACAACCTGGTCAAAACCAAGATGGAGAAAACCAATGAGTGTACAGAGAAAACGCTAAGAAGACGC
213 K P K K Q R R K K K G H N W S K P R W R K P M S V Q R K R L R R R
BbrPI (1079)
1301 AGAAAAGAAAATGCCAATTTTAGTGCTGAGTTGCTCCCGTGACGTGTGGTAACTGAAGGGAGTGTACACAAGGAAAAATCAAACAAGGGATATCTG
247 R K E N A N F S A E L L P V T C G N L K G V L H K E K F K Q G I S
1401 TGAAGTCCATACAGTGTGAGAATGGAAGTGGTTACACCCCTCGGAATTTGAAATGATGGGAGGCTATGGAAGTCAAAGAAGTGGAACTGAGCCTGCG
280 V K S I Q C Q N G N W F T P S E F E M M G G Y G K S K N W K L S L R
1501 CTGCCATAATTTGCCCTGAAACTCCTGATCCAGAGAAAATTTCTACCAATCCTCCACGGATACATCGAAAAGAAAAGCAGAGAACACAGAAGCTG
313 C H N W P L K L L I Q R N F L P N P P R I H R K R K E Q R T Q N L
EcoRV (1392)
1601 CATCGTTCAGCTGACCTAGTATTGAGAACTCAGATAAATGTGAGGTGTGCCGGGATGGAGGTTTACTCTTCTGTGTGATACTGTTTCAGAGCCT
347 H R S P A D P S I Q N S D K C E V C R D G G L L F C C D T C F R A
PvuII (1610)
1701 TCCATGAGGACTGTGATATCCCAACAGTGAAGCTGAGATAACACCATGGAGTTGCATCTTCTGTAGGATGCAGTCCCTAGGAAGTCAACCGAGTCTTCC
380 F H E D C H I P T V E A E I T P W S C I F C R M Q S L G S Q P S L P
NcoI (1744) **Bsu36I (1775)**
1801 AGAATCTGAGATACTGCAGAGGGAGATGGCTCCCAGGAACAGTTGAAGTGTGAGTATGTTCTCTTGGAGTCTATTGCTGTTCTGAGAGCTCCTTTTTT
413 E S E I L Q R E M A P Q E Q L K C E Y V L L R V Y C C S E S S F F
1901 TCCAAGATGCCATATTATTATTATTTAGAGAGATGACTGTGAGTGTGCAAGAGCCTATGTGGTTGGACATAATCAAGAAAAAGCTGAGTGACCAGGCTT
447 S K M P Y Y Y Y F R E M T V S V Q E P M W L D I I K K K L S D Q A
BspEI (2042) **EcoO109I (2066)** **PshAI (2096)**
2001 ACTGCCAAGTAGAAGACTTTGTGCAAGACATGAGGCTCATCTCCGGAATCACAAAACACGTTCAAGGACCCTAAGTTTGGCGAAATGGGATTTAGACT
480 Y C Q V E D F V Q D M R L I F R N H K I T F K D P K F G E M G F R L
NheI (2195)
2101 GGAGTCCAAGTTTGAAGAAGTTTCAAGGAAGTGGTTGCTATTCAGGAAACAAATGAAAAGAGTTGACTAAAAATCAGAGACATGGCTGGTGTACGCTAG
513 E S K F E K S F K E V F A I Q E T N E K S •
MscI (2201)
2201 CTGCCAGACATGATAAGATACATTGATGAGTTTGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCT

HpaI (2333) **MfeI (2344)**
2301 TTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGTTTTTTAA

2401 **EcoRI (2429)**
GCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACCTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGG

2501 ATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCCAA

2601 **SapI (2611)** **SspI (2668)** **SwaI (2682)**
GGTTTGAAGTACTGCTCTTCATTTCTTTATGTTTTAAATGCAGTACCTCCACATTCCTTTTATAGTAAATATTCAGAAATAATTTAAATACATCATTGC

2701 **EcoO109I (2743)**
AATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGACCTT

2801 TAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCAT
141 • N R T Y K L P I L E E I T T K V L K G N

2901 **BstXI (2972)**
TCATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATC
120 M E I 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

3001 AGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAAATGGCAATGGCTTCAGCACAGACAGTGACCTG
87 S Y P 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

3101 **StuI (3107)**
CCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTGCTCATAGAGCATGGTGA
53 G I Y 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

3201 TCTTCTCAGTGGCGACCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGATTATACTATGCCGA
20 K E T A V E V L E L D Q Q S I N F T K M

3301 **AseI (3315)**
TATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCG

3401 **SpeI (3470)**
TACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCATTGACGTCA

3501 **SnaBI (3598)**
ATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATA

3601 CGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACT

3701 TGGCATATGATACTTGTACTGCAAGTGGGCAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGA

3801 **PacI (3889)** **PstI (3882)** **SdaI (3881)** **BspLU11I (3899)**
ACATACGTCATTATTGACGTCAATGGCGGGGGTCTGTTGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACCGCTGCAGGTTAATTAAGAA

3901 CATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAAT

4001 CGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCGTGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGC

4101 CGTTACCGGATACTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGTCTGTTCCGCTC

4201 **ApaLI (4213)**
CAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGTAAGACACGACTTATCG

4301 CCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAA

4401 GAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGG

4501 TTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAAAC

4601 **EagI (4649)** **PacI (4629)** **SwaI (4638)** **NotI (4648)**
TCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTT

4701 GTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCGAGTGCAAGTGCAGGTGCCAGAACA

4801 TTTCTCTATCGAA