



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

101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGCTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

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

301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCTCCTGAACTCGTCCGCGCTTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
NgoMI (441)
NaeI (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGTCTTTGTTTCGTTT

AgeI (552) **NcoI (568)**
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCCTACCTGAGATCACCGGTAGGAGGGCCACCATGGAAGGGAGAGGACCTACGGGATCTAC
1 M E G R G P Y G I Y
601 GACCCAGGGGGCAGCAGCCTCTGGGAGAGGTGCCGAGCTTTTGAACGTCTAGTGGAGGAGAATACTCGGCTGAAGGAAAAATGCAAGGGATAAAGA
11 D P G G S T P L G E V S A A F E R L V E E N T R L K G K M Q G I K
701 TGTTAGGGGAGCTTCTGGAGGAGTCTCAGATGGAAGCGTCCAGACTCCGGCAGAAGGCAGAGGAGCTGGTCAAGGACGCGAGCTGTCCACCACCGACATC
44 M L G E L L E E S Q M E A S R L R Q K A E E L V K D S E L S P P T S
801 TGCCCCCTCTTGGTCTCCTTTGATGACCTGGCTGAGCTCACAGGACAGGATACAAAGTCCAGGTACATCCTGCTACCAGCACTGCCGCCACCACC
77 A P S L V S F D D L A E L T G Q D T K V Q V H P A T S T A A T T T

NcoI (922)
901 GCCACCGCCACCAGGGAAATCCATGGAGAAGCCGAGCCAGCCTCCAAATCTCCGTCGAATGGCGCTCCTCGACTTTGAAGTGGTCCCTACTGAGG
11 A T A T T G N S M E K P E P A S K S P S N G A S S D F E V V P T E

EcoRI (1004) **Bsp120I (1050)**
1001 AGCAGAATTCACCCGAAACTGGCAGCCACCTACGAACATGATGGACCTGGGGCCCCACCCACAGAGGACGCAACCTGAAGCTCCACCTGCAGCGCCT
14 E Q N S P E T G S H P T N M M D L G P P P P E D S N L K L H L Q R L

XhoI (1173)
1101 GGAGACCACCTTAGCGTGTGTGCAGAGAGCCAGACCACAGCCAGCTTTCACCCACCTGGGCCGCATGGCCCTCGAGTTCAACAGGTTGGCCTCCAAA
17 E T T L S V C A E E P D H S Q L F T H L G R M A L E F N R L A S K
1201 GTGCATAAAAAATGAGCAGCGCACCTCCATCCTGCAGACTTATGTGAGCAGCTGCAGGAGGAGGAAAGCCCTGAAGGCCAAGCTGGACAAGGGCCTGG
21 V H K N E Q R T S I L Q T L C E Q L R Q E N E A L K A K L D K G L
1301 AACAGCGGGATCTGGCTGCTGAGAGGCTGCGGGAGGAAAAACGGAGCTCAAGAACTGTTGATGAACAGCAGCTGCAAAGAGGGACTCTGTGGGAGCC
24 E Q R D L A A E R L R E E N T E L K K L L M N S S C K E G L C G Q P

BstXI (1456)
1401 CAGCTCCCCAAAGCCAGAGGGTGTGGCAGAAGGGCGTGGCTGGACAGCAGCAGCCAGTGTGATGGCGAGTAAAGTCCCTGAAGCGGGGCCCTTTGGA
27 S S P K P E G A G K K G V A G Q Q Q A S V M A S K V P E A G A F G
1501 GCAGCTGAGAAGAAGGTGAAGTTGCTAGAACAGCAACGCATGGAGCTGTGGAAGTGAACAAGCAGTGGGACCAGCATTCCGGTCCATGAAGCAGCAGT
31 A A E K K V K L L E Q Q R M E L L E V N K Q W D Q H F R S M K Q Q
1601 ATGAGCAGAAGATCACAGAGCTTCGCCAGAAGCTGGTGGACCTGCAGAAACAGGTAAGTGAAGTGGAGGCCAAGGGAGCAGAAGCAGCGTACTTTGA
34 Y E Q K I T E L R Q K L V D L Q K Q V T E L E A E R E Q K Q R D F D

MscI (1714) **Acc65I (1795)**
1701 CCGAAAACCTCCTCGCCAAATCGAAGATAGAGATGGAAGAGACCGACAAGGAGCAGCTGACAGCAGAGGCCAAGGAACTGCCCGAGAAGGTCAGGTAC
37 R K L L L A A K S K I E M E E T D K E Q L T A E A K E L R Q K V R Y

BsrBI (1818)
1801 CTACAGGATCAGCTGAGCCCGCTCAAGGCAACGAGAATACCAGGAGAAGGAGATCCAGCGGCTCAATAAGGCCCTGGAGGAGGCCCTCAGCATCCAGG
41 L Q D Q L S P L T R Q R E Y Q E K E I Q R L N K A L E E A L S I Q

SpeI (1970)
1901 CCTCTCCATCATCTCCCGCTGCAGCTTTTGGGAGTCCAGAAGCGTGGGGGCCATCTGAGGAAGCAGGAACTAGTGACACAGAATGAGTTGCTGAAACA
44 A S P S S P P A A F G S P E G V G G H L R K Q E L V T Q N E L L K Q

BglIII (2008)
2001 GCAGGTAAGATCTTTGAAGAGGACTTCCAGAGGGAACGGAGTGACCGTGAACGCATGAATGAAGAGAAGGAGGAGCTGAAGAAGCAAGTAGAGAAGCTG
47 Q V K I F E E D F Q R E R S D R E R M N E E K E E L K K Q V E K L

BstEII (2108)
2101 CAGGCCAGGTACCCCTGACTAATGCCAGCTCAAACCTCAAAGAGGAGGAGAAGGCCAAGGAAAGCCCTCAAACAGCAGAAGGAAAGCAAAAGCTT
51 Q A Q V T L T N A Q L K T L K E E E K A K E A L K Q Q K R K A K A

Eco47III (2206) **XcmI (2269)** **Acc65I (2282)**
2201 CGGGAGAGCGCTACACATGGAACCCACCCTGAGCACGCTGCGCGCCTATCCCTATGCTACCCACCCATGCCAGCCATGGTACCTACCATGCCTA
54 S G E R Y H M E P H P E H V C G A Y P Y A Y P P M P A M V P H H A Y

NcoI (2341)
2301 CAAGGACTGGTCCAGATCCGATACCCTCCACCCCTGTGCCATGGAGCACC CGCCCAACCCCAACTCTCGCTCTTCCATCTGCCGGAGTACCC
57 K D W S Q I R Y P P P P V P M E H P P P H P N S R L F H L P E Y T

DraIII (2407)
2401 TGGCGTCCACCCTGTGCAGGATTCGGAATCAGAGCTCTCAAGTATGACCCGCCCCAGACAGGCTGCAGAGCCAGAGTCTGCAGACAATGACTGTG
61 W R P P C A G I R N Q S S Q V M D P P P D R P A E P E S A D N D C

MscI (2549)
Bsp120I (2502) **NheI (2543)**
2501 ATGGGCCCAGTGAGGCTGCAGTGGGTCAATTTGGTTCCACCTTGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGCAAAACCACAACTAGA
64 D G P Q •

2601 **HpaI (2681)** MfeI (2692)
 ATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAAACAAGTTAACACAACAATTGCA

2701 EcoRI (2777)
 TTCATTTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAAATTCTAAAATACAGCATAGCA

2801
 AAACCTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTG

2901
 CAGCCTCACCTTCTTTCATGAGGTTAAGATATAGTGATTTTTCCAAGGTTTGAAGTCTCTTCTTTCTTTATGTTTTAAATGACTGACCTCCCAC

3001 **SspI (3016)** SwaI (3030)
 ATTCCCTTTTATAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTC

3101
 ATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTAAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGT

3201
 ACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTCCATTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCT
 136 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 E R

3301 **BstXI (3320)**
 GCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGGTTG

3401
 103 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 N
 CTCACAGCAGACCAATGGCAATGGCTTCAAGCAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCC
 69 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 T R

3501 **XmnI (3597)**
 TGATGGCCGCCCCGACATGGTGTCTGTGCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAA
 36 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 N F

3601 **BspHI (3605)** **VspI (3663)**
AseI (3663)
 GGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGC T
 3 T K M

3701
 TATCTGACGGTCTACTAAACGAGCTCTGTTATATAGACCTCCACCCTACACGCCTACCGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTT

3801 **SpeI (3818)**
 GGAAAGTCCCCTGATTACTAGTCAAACAACCTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCCA

3900 **SnaBI (3946)**
Eco105I (3946)
 TTGATGTAAGTCCAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGTAAGTGGGCAT

4000 **NdeI (4051)**
 AATGCCAGGGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTAAGTGGCAGTTTACCGTAAATA

4100
 CTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATATTGACGTCAATGGGGGGGGTCTTGGGGGTCAGCCA

4200 **SdaI (4229)** **PacI (4237)** **BspLU11I (4247)**
 GCGGGCCATTTACCGTAAAGTTATGTAACGCTGCAGGTTAA TTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAGGCCGCG

4298
 TTGCTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACC

4398
 AGGCGTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCT

4498 **ApaLI (4561)**
 TTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCC

4598
 TTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAG

4698
 GCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAA

4798
 AAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCTGGTAGCGGTGTTTTTTTTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAA

4898 **PacI (4977)** **SwaI (4986)** **NotI (4996)**
 GAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AG

4998 **EagI (4997)**
 CGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGGTTTTTTTGTGTAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAA

5098
 AACAAACTAGCAAATAGGCTGTCCCGAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA