



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

101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) **HindIII (245)** **EcoNI (287)**
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCCTTACGCGCCCGCGCCCTACCTGAGGCC

301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
NgoMI (441)
NaeI (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGTCTTTGTTTCGTTT

NcoI (560)
BstEII (555)
KasI (535) **AgeI (552)**
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCCTACCTGAGATCACCGGTACCATTGGGCAGGGTCCCGCTGGCCTGGTGGTTGGCCTGTG

601 CTGCTGGGGGTGTGCAGCCATAAGGACACACAGACCGAGGCTGGCAGCCGTTTGTGGGAAACCCAGGGAATATCACAGGTGCCAGAGGACTCACGGGG
13> C W G C A A H K D T Q T E A G S P F V G N P G N I T G A R G L T G
701 ACACCTCGGTGTGAGCTCCAGGTTACAGGGGAAACCCCTGAGGTGGTGTGGCTTCGAGATGGACAGATCCTAGAACTGGCTGATAACCCAGACCCAGG
47> T L R C E L Q V Q G E P P E V V W L R D G Q I L E L A D N T Q T Q
801 TGCCTCTGGGCGAAGACTGGCAAGATGAATGAAAAGTTGTGTCAGTCAGATCAGAATCTCAGCCCTGCAACTTTTCAGATGCAGGGGAGTACCAGTGTATGGT
80> V P L G E D W Q D E W K V V S Q L R I S A L Q L S D A G E Y Q C M V

XbaI (903)
901 GCATCTAGAAGGACGGACCTTTGTGTCTCAGCCGGGCTTTGTAGGGCTGGAAGGTCTCCCGTACTTCTGGAGGAGCCTGAGGACAAAGCTGTGCCTGCC
113> H L E G R T F V S Q P G F V G L E G L P Y F L E E P E D K A V P A
1001 AACACCCCTTTCAACCTAAGCTGCCAGGCCAGGGACCCCGGAACCCGTGACCTACTCTGGCTTCAAGATGCTGTCCCCCTGGCCCAAGTACAGGAC
147> N T P F N L S C Q A Q G P P E P V T L L W L Q D A V P L A P V T G
1101 ACAGCTCCAGCACAGTCTGCAACTCCAGGCCTGAACAAGACATCTTCTTCTCATGTGAAGCCACAATGCCAAGGGAGTACCACCTCCCGCACAGC
180> H S S Q H S L Q T P G L N K T S S F S C E A H N A K G V T T S R T A

Eco72I (1237)
BbrPI (1237)
1201 CACCATCACAGTGTCTCCCCAGAGCCTCACCATCTCCAGTGGTTTCCAGACAACTACGGAGCTAGAGGTAGCTTGGACCCCTGGCCTGAGTGGCATC
213> T I T V L P Q R P H H L H V V S R Q P T E L E V A W T P G L S G I

BsrBI (1303) **SdaI (1320)**
1301 TACCCGCTCACCCACTGCAACCTGCAGGCCGTGCTGTCAGACGATGGGGTGGGTATCTGGCTGGGAAAGTCAGATCCTCCTGAAGACCCCTCACCTTGC
247> Y P L T H C N L Q A V L S D D G V G I W L G K S D P P E D P L T L

EcoRV (1467)
Eco32I (1467)
BspEI (1463) **Bsp120I (1488)**
1401 AAGTATCAGTGCCCCCACCAGCTTCGGCTGGAAAAGCTCCTTCTCACACCCCGTATCACATCCGGATATCCTGCAGCAGCAGCCAGGGCCCTCACC
280> Q V S V P P H Q L R L E K L L P H T P Y H I R I S C S S S Q G P S P

Psp1406I (1558)
1501 TTGGACCCACTGGCTTCTGTGGAGACCACAGAGGAGTGCCTTGGGTCCCCCTGAGAACGTTAGCGCCATGCGGAATGGGAGCCAGGTCTCTGTCGCT
313> W T H W L P V E T T E G V P L G P P E N V S A M R N G S Q V L V R

Acc65I (1640)
1601 TGGCAGGAGCCAAGGGTGGCCCTGCAAGGCACCCTGTTAGGGTACCGGCTGGCATATCGAGGCCAGGACACCCCGAGGTACTTATGGATATAGGGCTAA
347> W Q E P R V P L Q G T L L G Y R L A Y R G Q D T P E V L M D I G L

BstEII (1707)
XhoI (1700) **Tth111I (1749)** **Bsp120I (1786)**
1701 CTCGAGAGGTGACCTTGGAACTGCGGGGGACAGGCCTGTGGTAACCTGACTGTCTGTGACAGCCTATACCTCGGCTGGGGATGGGCCCTGGAGCCT
380> T R E V T L E L E R G D R P V A N L T V S V T A Y T S A G D G P W S L

KasI (1820)
1801 TCCTGTGCCCTAGAGCCCTGGCGCCAGGGCAAGGACAGCCACTCCACCATCTGGTGTGAGTGAACCCCACTCGCGCCTTCTCGTGGCCTTGGTGGTAT
413> P V P L E P W R P G Q G Q P L H H L V S E P P P R A F S W P W W Y
1901 GTACTGCTGGGAGCACTTGTGGCTGCCGCTGCTCATCTTGGCCCTGTTCTTGTCCATCGGAGGAAGAAGGAGACTCGATATGGGGAGGTGTTTG
447> V L L G A L V A A A C V L I L A L F L V H R R K K E T R Y G E V F

Acc65I (2033) **NgoMIV (2057)**
NgoMI (2057)
NaeI (2057)
2001 AGCCAACCGTGGAAAGAGGTGAAGTGGTAGTCAGGTACCGTGTCCGAAAGTCTACAGCCGGCGGACCCTGAAGCCACCTTGAACAGTCTGGGCATCAG
480> E P T V E R G E L V V R Y R V R K S Y S R R T T E A T L N S L G I S
2101 TGAAGAGCTGAAGGAGAACTACGAGACGTCATGGTAGATCGGCATAAGTGGCCTTGGGGAAGACCTGGGAGAAGGAGAATTTGGCGCTGTGATGGAA
513> E E L K E K L R D V M V D R H K V A L G K T L G E G E F G A V M E
2201 GGTCAAGTCAATCAGGATGACTCCATCCTCAAGGTGCTGTGAAGACCATGAAAATGGCCATGTCACAAAGATCAGAGCTGGAGGATTTCTGAGTGAAG
547> G Q L N Q D D S I L K V A V K T M K I A I C T R S E L E D F L S E
2301 CTGCTGCATGAAGGAATTTGACCACCCCAAGTCATGAGGCTCATTGGCGTCTGTTTTAGGGCTCTGACAGAGAGGGTTTCCAGAACCTGTGGTCAT
580> A V C M K E F D H P N V M R L I G V C F Q G S D R E G F P E P V V I
2401 CTTGCCTTTCATGAAACACGGAGACCTACACAGTTTCTCCTGTACTCCCGGCTCGGGGACCAGCCAGTGTCTTCTGCCACTCAGATGCTAGTGAAGTTC
613> L P F M K H G D L H S F L L Y S R L G D Q P V F L P T Q M L V K F

2501 ATGGCCGACATTGCCAGTGGTATGGAGTACCTGAGTACCAAGAGATTACATACATCGGGACCTGGCTGCCAGGAAGTGCATGCTGAATGAGAACATGTCCG
 647▶ M A D I A S G M E Y L S T K R F I H R D L A A R N C M L N E N M S

BglIII (2630)

2601 TGTGTGTGGCAGACTTCGGGCTCTCCAAGAAGATCTACAACGGGGATTACTACCGCCAAGGGCGCATTGCCAAGATGCCAGTCAAGTGGATTGCTATTGA
 680▶ V C V A D F G L S K K I Y N G D Y Y R Q G R I A K M P V K W I A I E
 2701 GAGTCTGGCAGATCGGGTCTACACCAGCAAGAGCGATGTGTGGTCTCTCGGTGTGACAATGTGGGAGATCGCCACCCGAGGCCAACTCCCTATCCAGGG
 713▶ S L A D R V Y T S K S D V W S F G V T M W E I A T R G Q T P Y P G
 2801 GTGGAGAACAGTGAATTTACGACTACCTGCGTCAAGGAAATCGGCTGAAACAGCCTGTGGACTGTCTGGACGGCCTGTATGCCCTGATGTCTCGGTGCT
 747▶ V E N S E I Y D Y L R Q G N R L K Q P V D C L D G L Y A L M S R C

XhoI (2912)

2901 GGGAACTGAACCTCGAGACCGCCAAAGTTTTGCGGAGCTCCGGGAAGACTTGGAGAACACACTGAAGGCTCTGCCCCCTGCTCAGGAGCCAGATGAAAT
 780▶ W E L N P R D R P S F A E L R E D L E N T L K A L P P A Q E P D E I
 3001 CCTCTATGCAACATGGATGAGGGCGGAAGCCACCTTGAACCCCGTGGGGCTGCTGGAGGAGCTGACCCCAACCCAACTGATCCTAAGGATTCTGT
 813▶ L Y V N M D E G G S H L E P R G A A G G A D P P T Q P D P K D S C
 3101 AGTGTCTCACTGCAGCTGACGTCCACTCAGCTGGACGCTATGCTTTTGTCTTCTACAGCCCCAGGACCCACTCTGTCTGTGACAGAGGCTGCCAG
 847▶ S C L T A A D V H S A G R Y V L C P S T A P G P T L S A D R G C P

MscI (3263)
BalI (3263)
NheI (3257)

3201 CACCTCCAGGGCAGGAGGACGGAGCCTGAGACAATCTTCCACCTGGGACATCCTCTCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGAC
 880▶ A P P G Q E D G A •

HpaI (3395)

3301 AAACCACAAC TAGAATGCAGTGA AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAGTTAA

MfeI (3406)

3401 CAACAACAATTGCATTCTTTTATGTTTCAGGTT CAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAATGTGGTATGGAATTCTAA

EcoRI (3491)

3501 AATACAGCATAGCAAACCTTAACTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTTTGCCAATGT
 3601 GCATTAGCTGTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGTTTGAAGTACTGCTCTTCATTCTTTATGTTTTAAATG

SspI (3730) **SwaI (3744)**

3701 CACTGACCTCCCACATTCCTTTTATGTA AAAATATT CAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGAT
 3801 GCTCAAGGCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTCTAGCT

3901 TTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCA
 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 V F C D P A Y D
 4001 GAGATGAGCTCTGCACATGCCACAGGGGTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGT
 107▶ 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 R V A V I T D F D
 4101 CCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTT CAGCACAGACAGT GACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTC
 74▶ 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 I H V A S I I E
 4201 CCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGCTTGTTCCTCATAGACATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGC
 41▶ 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 V E V L E L D Q

XmnI (4311)

4301 TGAGAGATGTTGAAGGTCTTCATGATGGCCCTCTATAGTGAGTCTGATTATACTATGCCGATATACTATGCCGATGATTAATGTCAAACAGCGTGA
 7▶ Q S I N F T K M

VspI (4377)
AseI (4377)

4401 TGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGT

SpeI (4532)

4501 TGTTACGACATTTTGGAAAGTCCC GTTGATTTACTAGTCAA ACAAACCTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACC

SnaBI (4660)
Eco105I (4660)

4600 GCTATCCACGCCATTGATGTA CTGCAAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTA CTGCAAGTAGGAAAGTCCCATAAGGT

NdeI (4765)

4700 CATGTA CTGGCATAATGCCAGGCGGGCATTACC GTGATTGACGTCAATAGGGGGCTACTTGGCATATGATACTTGTACTGCAAGTGGGCA
 4800 GTTACC GTAAACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGT

SdaI (4943) **PacI (4951)** **BspLU11I (4961)**

4900 TGGGCGGT CAGCCAGGCGGGCATTACC GTAAGTTATGTAACGCTT C G A G T T A A T T A A G A A C A T G T G A G C A A A A G G C C A G C A A A G G C C A G G A A C C
 4998 GTAAAAAGGCCGCTTGTGCGCTTTTCCATAGGCTCCGCCCCCTGACGAGCATC AAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG
 5098 ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCG

ApaLI (5275)

5198 GGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGC
 5298 CCGACCGTGCGCCTTATCCGGTAACTATCGTCTT GAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAG

5398 AGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA

5498 GTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTTGGTTTGGCAAGCAGCAGATTACGCGCAGAA

5598 AAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATT PacI (5691)

EagI (5711)

Swal (5700) **NotI (5710)**

5698 AACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAA

5798 ACAAACGAAACAAACAAACTAGCAAATAGGCTGTCCCAAGTGCAAGTGCAAGTGCCAGAACATTTCTCTATCGAA