



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA **MfeI (82)**
101 GAGAAAGTGGCGGGGTAACGGAAAGTGTGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCGCTACCTGAGGCC **HindIII (245)** **Bsu36I (291)**
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGCTAGGTAAGTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTCTCAACTCTACGCTTTTGTTCGTTT **NgoMIV (441)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTA **KasI (535)** **AgeI (552)** **BamHI (570)** **BspHI (560)**
601 CCTCTGCGGGCTGCTGCTGGCCGCTCGGCTGCTGCGCCAGCGGCCCGCACCTCGTGTGCTGGACAACCTGAACGTGCGGGACAGCACTCG **1 M K M D P R L A V R A W P**
701 CTGTTCTCCGAGCCTGAAGGACCGGGCTTTGAGCTCACCTTAAGACCGCAGATGACCCAGCTTGTCCCTCATTAAAGTACGGGGAGTTCCTCTATG **13 L C G L L L A A L G C V C A S G P R T L V L L D N L N V R D T H S**
801 ACAACCTTATCATTTTTCCCGTGGTGAAGATTTGGAGGCAACATCAACGTGGAGACCATCAGTGCCTTATTGATGGTGGCGGAGCGTTTTGGT **47 L F F R S L K D R G F E L T F K T A D D P S L S L I K Y G E F L Y**
80 D N L I I F S P S V E D F G G N I N V E T I S A F I D G G G S V L V
901 GGCTGCCAGCTCTGATATTGGTACCCTTTCGGGAACTGGGCGAGTGTGGATTGAGTTTGTGAAGAGAAAACAGCTGTCATCGACCACCACAAC **BstEII (919)**
113 A A S S D I G D P L R E L G S E C G I E F D E E K T A V I D H H N
MscI (1017)
1001 TATGATGCTCTGACCTTGCCAGCACACTCATTGTGGCTGACACTGAGAACCTGCTGAAGGCCCGACCATTGTTGGCAAGTCACTCTGAACCCCA **14 Y D V S D L G Q H T L I V A D T E N L L K A P T I V G K S S L N P**
1101 TTCTTTCCGAGGAGTTGGAATGGTGGCAGACCCGACAATCCCTTGGTTTGGACATCCTAACAGGCTCTTCAACCTTACTCCTTCCAGATAA **18 I L F R G V G M V A D P D N P L V L D I L T G S S T S Y S F F P D K**
XmaI (1274)
1201 ACCAATCACCCAGTACCCCATGCGGTGGGAGGAACACTCTGCTGATTGCCGGGCTCCAGGCCAGGAACAACGCCCGGGTCACTTTCAGTGGCTCTCTG **21 P I T Q Y P H A V G R N T L L I A G L Q A R N N A R V I F S G S L**
FspI (1353)
DraIII (1346)
1301 GATTTCTTCAGCGATGCCTTCTCAACTCGGCAGTGCAGAAGGCCACCCCGTGCAGAGGTTATTCTCAGACAGGCAACTGAAGTACTGCTGGCCC **24 D F F S D A F F N S A V Q K A T P G A Q R Y S Q T G N Y E L A V A**
PshAI (1442) **DraIII (1456)**
1401 TCTCAGCTGGGTGTTCAAGGAGGAGGGTGTCTTCGAGTAGGACCTGTGTCATCACCGGGTGGGCGAGATGGCTCCACCAATGCCTACACTGTCAC **28 L S R W V F K E E G V L R V G P V S H H R V G E M A P P N A Y T V T**
SandI (1550) **FspI (1589)**
1501 CGACTTGGTGGAGTATAGCATCGTCATAGAACAGCTCTCCAATGGCAAGTGGTCCCTTTGATGGTGTGACATTGAGTGGAGTTCGTGGCCTCGAC **31 D L V E Y S I V I E Q L S N G K W V P F D G D D I Q L E F V R I D**
1601 CCCTTCGTGAGGACCTTCTGAAGAGGAAAGGTGGCAAGTACAGTGTCCAGTTCAGGCTGACGTGTATGGTGTATTCCAGTTAAAGTGGATTACA **34 P F V R T F L K R K G G K Y S V Q F K L P D V Y G V F Q F K V D Y**
Eco47III (1772)
1701 ACCGGCTGGCTACACCCACTGTACTCTCCACCCAGGTGTGAGGCCACTCCAGCACACAGATGAGCGCTTACCCCTCGGCTATCCCTA **38 N R L G Y T H L Y S S T Q V S V R P L Q H T Q Y E R F I P S A Y P Y**
BbsI (1849)
1801 CTATGCCAGTGCCTTCTCATGATGGCCGGGCTTTCATCTTCCAGCATCGTCTTCTTGCACATGAAGGAGAAGGAGAAGTCTGACTGAGCTAGAACCTGG **41 Y A S A F S M M A G L F I F S I V F L H M K E K E K S D •**
MscI (1916)
NheI (1910)
1901 GGTCCACACAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATT

2001 TGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACATTGCATTCAATTTATGTTTCAGGTTTCAGGGGAGGTGT **HpaI (2048)** **MfeI (2059)**

2101 GGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATCTAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGA **EcoRI (2144)**
2201 ATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATAT

2301 AGTGATTTTCCCAAGGTTTGAAGTACTGCTCTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAATATTCAGAATAATT **SspI (2383)** **SwaI (2397)**
2401 TAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGG
2501 GAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTT **141 • N R T Y K L P I L E E I T T K**
BstXI (2687)
2601 TGACCAGCTTGCATTCACTCAATGAGCAAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGTCACATGCCACAGGGGCTGACCACCTGATGGA **125 V L K G N 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**

2701 TCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCA
92 R D V E D 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
StuI (2822)

2801 CAGACAGTGACCCGTGCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTGGTCTGATGGCCGCCCGACATGGTCTTGTGTCTC
58 C V T V R 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
BspHI (2972)
BbsI (2968)
XmnI (2964)

2901 CATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCGT
25 Y L M T I K E T A V E V L E L D Q Q S I N F T K M
AseI (3030)

3001 ATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTAT
SpeI (3185)

3101 ATAGACCTCCCACGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAACAAA
3201 CTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAA
SnaBI (3313)

3301 TAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTC
NdeI (3418)

3401 AATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATT
PstI (3597)
SdaI (3596)

3501 GGCCTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCTGTTGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCCCTG
PacI (3604) BspLU11I (3614)

3601 CAGGTTAATTAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGAC
3701 GAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTC
3801 CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGT
ApaLI (3928)

3901 GTAGGTCGTTGCTCCAAGTGGGTGTGTGCACGAACCCCGTTGAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTA
4001 AGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT
4101 ACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCAC
4201 CGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCT
EagI (4364)
PacI (4344) SmaI (4353) NotI (4363)

4301 CAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAATAAAATATCTTTATTTTATTACATCT
4401 GTGTGTTGGTTTTTTTGTGTAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGT
4501 GCAGGTGCCAGAACATTTCTCTATCGAA