



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

101 GAGAAAGTGGCGGGGTAAGTGGAAAGTGTGCTGTACTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGC

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

301 GCCATCCACGCCGGTTGAGTGCCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

KasI (535) **AgeI (552)** **SphI (560)**
501 TCTGTTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCCCTA CTTGAGATCACCGGTACAGTCCCAAGCACCAGCTTTCCAGTCCCTTCCAAGTTTCC 1 M P S T S F P V P S K F P

SgrAI (657) **BssHIII (654)**
601 ACTTGGCCCTGCGGTGCGGTCTTCCGGGAGAGGAGAACTTTGGGGCCCGCGCCGCGCCGCGCACCATGAAGTACAGCGGAGGAAGAACAATATGGC 13 L G P A A A V F G R G E T L G P A P R A G G T M K S A E E E H Y G

Ppu10I (701) **NsiI (701)** **BsrBI (730)**
701 TATGCATCCTCAACGTCAGCCCCGCCCTGCCGCTCCACAGGCGCACTCCACCTGCCGGCCCGTGCACAACTTCAGACCTCCACACCGGGCATCA 47 Y A S S N V S P A L P L P T A H S T L P A P C H N L Q T S T P G I

SgrAI (803)
801 TCCGCGCGCGGATCACCCCTCGGGTACGGAGCAGCTTTGGACGGTGGGCGCGGGCTACTTCTCTCTCCGGCCACACCAGGCCTGATGGGGCCCC 80 I P P A D H P S G Y G A A L D G G P A G Y F L S S G H T R P D G A P

DraIII (970)
901 TGCCTGGAGAGTCTCGCATCGAGATAACCTCGTGTGGGCTGTACCACAACAATAACCAGTTTTTCCACGATGTGGAGGTGGAAGAGCTCCTCCCT 113 A L E S P R I E I T S C L G L Y H N N N Q F F H D V E V E D V L P

1001 AGTCCAAACGGTCCCCCTCCACGGCCAGCTGAGTCTGCCAGCCTGGAGGCTACAGAGACCCCTCGTGCCTGAGCCCGGCCAGCAGCCTGTCTCCCTCC 147 S S K R S P S T A T L S L P S L E A Y R D P S C L S P A S S L S S

MluI (1151) **BsiWI (1148)** **NcoI (1171)**
1101 GGAGCTGCAACTCAGAGGCTCCTCCTACGAGTCCAACACTACTCGTACCCGTACGCGTCCCGCCAGACGTCGCCATGGCAGTCTCCCTGCGTGTCTCCCAA 180 R S C N S E A S S Y E S N Y S Y P Y A S P Q T S P W Q S P C V S P K

1201 GACGACGGACCCCGAGGAGGCTTTCCCGCGGGCTGGGGGCTGCACACTGCTGGTTCCCGCGGCACTCCCTCCACCTCGCCCCGCGCCAGCGTC 213 T T D P E E G F P R G L G A C T L L G S P R H S P S T S P R A S V

BsrBI (1323) **PshAI (1334)**
1301 ACTGAGGAGAGCTGGCTGGGTGCCCGCTCCTCCAGACCCGCGTCCCGGTGCAACAAGAGGAAGTACAGCCTCAACGGCCGGCAGCCGCTACTCACCCC 247 T E E S W L G A R S S R P A S P C N K R K Y S L N G R Q P P Y S P

SfiI (1491)
1401 ACCACTCGCCACGCGTCCCGCAGCGTCCCGCGGGTTCAGCGTACCCGACGACTCGTGGTGGGCAACACCACCCAGTACACCAGCTCGGCCATCGT 280 H H S P T P S P H G S C P R V S V T D D S W L G N T T Q Y T S S C A I V

1501 GGCCGCACTCAACGCGCTGACCACTGACAGCAGCTGGACCTGGGAGATGGCGTCCCTGTCAAGTCCCGCAAGACCACCTGAGGACCGCCCTCAGTG 313 A A I N A L T T D S S L D L G D G V P V K S R K T T L E Q P P S V

1601 GCGCTCAAGGTGGAGCCCGTCCGGGAGGACCTGGGAGCCCGCCCGCGGCGACTTCGCGCCGAAGACTACTCCTCTTTCCAGCACATCAGGAAGG 347 A L K V E P V G E D L G S P P P P A D F A P E D Y S S F Q H I R K

1701 GCGCTTTCGCGACAGTACCTGCGGTGCCGAGCACCCCTACCAGTGGGCGAAGCCCAAGCCCTGTCCCTACGTCCTACATGAGCCGACCCCTGCC 380 G G F C D Q Y L A V P Q H P Y Q W A K P K P L S P T S Y M S P T L P

1801 CGCCCTGACTGGCAGCTGCCCTCCACTCAGGCGCGTATGAGCTTCGGATTGAGGTGAGCCCAAGTCCACCAGCCACTACGAGACCGGAGGGG 413 A L D W Q L P S H S G P Y E L R I E V Q P K S H R A H Y E T E G

1901 AGCCGGGGGGCGGTGAAGGCGTCCGGCGGAGGACACCCATCGTGCAGCTGCATGGCTACTTGGAGAATGAGCCGCTGATGCTGCAGCTTTTTCATTGGGA 447 S R G A V K A S A G G H P I V Q L H G Y L E N E P L M L Q L F I G

ApaLI (2041) **Tth111I (2061)**
2001 CGGCGGACGACCGCTGCTGCGCCGACGCGCTTCTACAGGTGCACCGCATCACAGGGAAGACCGTGTCCACCACGACGACGAGGCCATCCTCTCCAA 480 T A D D R L L R P H A F Y Q V H R I T G K T V S T T S H E A I L S N

BstXI (2102) **SphI (2138)**
2101 CACCAAAGTCTGGAGATCCCACTCCTGCCGAGAAACAGCATGCGAGCCGTCATTGACTGTGCCGGAATCCTGAAACTCAGAACTCCGACATTGAACTT 513 T K V L E I P L L P E N S M R A V I D C A G I L K L R N S D I E L

NotI (2276)
2201 CGGAAAGGAGAGACGGACATCGGGAGGAAGAACACACGGGTACGGTGGTGTCCGCGTTACAGTCCCGCAACCCAGCGGCGGACGCTGTCCCTGCAGG 547 R K G E T D I G R K N T R V R L V F R V H V P Q P S G R T L S L Q

Eco47III (2327) **AfeI (2327)**
2301 TGGCTCCAACCCATCGAATGCTCCAGCGCTCAGCTCAGGAGTGCCTCTGGTGGAGAAGCAGAGCACGGACAGCTATCCGGTCTGGGCGGGGAGAA 580 V A S N P I E C S Q R S A Q E L P L V E K Q S T D S Y P V V G G K K

2401 GATGGTCTGTCTGGCCAACTTCTGCAAGGACTCCTGAGGACTCCTGAGGCTGATTTTCGTTGGAGAAAGCCCAAGTGGCCACCATGTCTGGGAGATGGAAGCGAAAAC 613 M V L S G H N F L Q D S K V I F V E K A P D G H H V W E M E A K T

EcoRI (2520)
2501 GACCGGGACCTGTGCAAGCCGAATTCTGTTGGTGGTGGAGATCCCGCGTTTCGGAATCAGAGGATAACCAGCCCGTTTACGTCAGTTTCTACGTCGTGCA 647 D R D L C K P N S L V V E I P P F R N Q R I T S P V H V S F Y V C

2601 ACGGGAAGAGAAAAGCGAAGCCAGTACCAGCGTTTCACCTACCTTCCC GCCAACGGTAACGCCATCTTTCTAACCGTAAGCCGTGAACATGAGCGCGTGGG
680▶ N G K R K R S Q Y Q R F T Y L P A N G N A I F L T V S R E H E R V G
NheI (2731)

2701 GTGCTTTTTCTAAAGACGCAGAAACGACGTCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAC TAGAATGCAGTGAAAA
713▶ C F F •

2801 AAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAATTGCATTCATTTTATGT
HpaI (2869) MfeI (2880)

2901 TTCAGGTT CAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTTAACCT
EcoRI (2965)

3001 CCAAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTT
3101 CTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGTTTGAAGTAGCTCTTCATTTCTTTATGTTTTAAATGACTGACCTCCACATTCCCTTTTAA

3201 GTAATAATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCC
SspI (3204) SwaI (3218)

3301 AGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGA
141 • N R T Y K L P I

3401 TGAGTTCCTCAATGGTGGTTTTGACCAGCTTCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACA
132▶ 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 C M G C
SacI (3479)

3501 GGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGAC
BstXI (3508)

3601 CCAATGGCAATGGCTTCAGCACAGACAGTGCCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTGGTCCCTGATGGCCGCC
99▶ 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 S V A S
65▶ 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 I A A G
BspHI (3793)

3701 CGACATGGTGCTTGTTCCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGAT
32▶ 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 T K M
XmnI (3785)

3801 GGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGT
AseI (3851)

3900 TCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCG
SacI (3908)

4000 TTGATTTACTAGTCAAAAACAACTCCATTGACGTCAATGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTG
SpeI (4006)

4099 CCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAAGTCATGTACTGGCATAATGCCAGGGC
SnaBI (4134)

4199 GGCCATTTACCGTATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCTCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATT
NdeI (4239)

4299 GACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATT

4399 TACCGTAAGTTATGTAACGCTG CAG GTT AA TTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGTTGCTGGCGTT
PacI (4425) BspLU11I (4435)

4497 TTTCCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCC
4597 CCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCT

4697 CACGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTT CAGCCGACCGCTGCGCTTATCCGGTAA
ApaLI (4749)

4797 CTATCGTCTT GAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACA
4897 GAGTCTTTGAAGTGGTGGCCTAACTACGCTACACTAGAAGAAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTA
4997 GCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTT

5097 GATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATA
PacI (5165) SwaI (5174) NotI (5184)

5197 AAATATCTTTATTTTATTACATCTGTGTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGC
5297 AAAATAGGCTGTCCCGAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA