



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCCTACCTGAGGCC
HindIII (245)
301 GCCATCCACGCGGTTGAGTCGCGTTTCTGCCGCCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTGTTTCGTTT

NgoMIV (441)
NaeI (441)
501 TCTGTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTGTTTCGTTT

AgeI (552)
BspHI (560)
501 TCTGTTTGTCCGGCGTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTCATCATGAGCCTCTCCATGAGAGATCCGGTTATCCCTGGGAC
1 M S L S M R D P V I P G T
601 AAGCATGGCCTATCATCCCTTCTACCTCACCGGGCGCGGACTTCGCATGAGCGCGGTACTGGGCCACCAGCCGCTTTCTTCCCCGCGCTAACGCTG
13 S M A Y H P F L P H R A P D F A M S A V L G H Q P P F F P A L T L

XcmI (743)
701 CCTCCAACGGCGCGCGGCTCTCGCTCCCCGAGCTCTGGCCAAGCCATCATGGATCAGTTAGTGGGGGCTGCTGAGACCGGCATCCCTTTCTCAT
47 P P N G A A A L S L P G A L A K P I M D Q L V G A A E T G I P F S

SnaI (804)
801 CCCTGGGACCCCGCAGCAGCATCTGAGGCTCTGAAGACCATGGAACCCGAAGAAGAGCTAGAAGATGACCTAAGTGCACCTGGAGGCGAAGGAACCTTTG
80 S L G P Q A H L R P L K T M E P E E D V E D D P K V H L E A K E L W

BsaBI (927)
901 GGACCAGTTTACAAGCGGGTACAGAGATGGTCATCAGAAAGTCAAGGCGAATGTTCCCTCCGTTAAAGTGAAGTGTCTGGACTGGATAAAAAG
113 D Q F H K R G T E M V I T K S G R R M F P P F K V R C S G L D K K

PshAI (1036)
1001 GCCAAGTATATTTATTGATGGACATTATAGTGTGACGACTGTCGATATAAATTTCAACTCTCGGTGGATGGTGGCCGGTAAGGCAGACCCCGAAA
147 A K Y I L L M D I I A A D D C R Y K F H N S R W M V A G K A D P E
1101 TGCCAAAGAGAATGTATATACACCCGGACAGCCCGCTACGGGGGAGCAGTGGATGTCCAAAGTGTCACTTTCCACAACTGAAACTCACCAACAACAT
180 M P K R M Y I H P D S P A T G E Q W M S K V V T F H K L K L T N N I

Ppu10I (1234)
NsiI (1234)
EcoRV (1276)
ScaI (1297)
1201 ATCGGATAAACACGGATTTACTATACTAACTATGCATAAGTACCAGCCGGTTCACATCGTCAGAGCAACGATATCCTGAAACTGCCTTACAGT
213 S D K H G F T I L N S M H K Y Q P R F H I V R A N D I L K L P Y S

BstBI (1304)
EcoRI (1330)
1301 ACTTTTCGAACCTACCTGTTCCCGGAAACAGAATTCATCGCCGTTACTGCCTATCAGAATGACAAGATAACTCAGTAAAAATAGACAACAATCCCTTTG
247 T F R T Y L F P E T E F I A V T A Y Q N D K I T Q L K I D N N P F
1401 CAAAGGGTTTTCGAGACACTGGCAATGGCAGGAGAGAAAAAGAAAGCAGCTCACACTGCAGTCCATGAGAGTGTGAGCAGAGGCATAAGAAGGAGAC
280 A K G F R D T G N G R R E K R K Q L T L Q S M R V F D E R H K K E T

BglIII (1593)
1501 TTCGGATGAGTCTCCAGCGAGCAGGCAGCCTTCAACTGCTTTGCCAGGCATCTCTCTGCTGTCTCCATCGTGGGGACATCCAACCTCAAAGATCTG
313 S D E S S S E Q A A F N C F A Q A S S P A V S I V G T S N L K D L
1601 TGTCCCAGCGAAGCCGAGAGTACGCGGAGGCTGAGAGCAAGGAGGACAGCCGCTGAGGCTGCGACGCGGCAAGATTTCCACCACCGCCGATG
347 C P S E A E S D A E A E S K E E H G P E A C D A A K I S T T T A D
1701 AGCCAGGCGCGACAAGGGCAGCCAGCAACCAGGGCGCAGCTGTTCCCGCGGAGCCAGCCGGGCCGCGACCCGCGCTTGGACAAGGCATCGCC
380 E P G R D K G S P A T R A Q L F P A E P S R A R D T A R L D K A S P
1801 AGACTCGGCCATAGCCCGCCACCATCTCGTCCAGCAGCGAGTCCGGGAGCTGATGAGCGCAGGAGTCCCGGGCGGAGGGTCCAGTCCGCCACCAAG
413 D S R H S P A T I S S S T R V P G A D E R R S P G R E G P V A T K

BssHII (1911)
BssHII (1909)
1901 GTGGATGAGGCGCGCGATTCCCTGCCAAGGACGCTTCGCTCCACTGTGCGTGCAGACAGACGCCACCAGCGCACCTAGCGCAGGGGGCCCTCCCGGGCT
447 V D E A R A I P A K D A F A P L L S V Q T D A T A H L A Q G P L P G

SfiI (2014)
BsrBI (2047)
2001 TGGGTTTCGCCCCAGGCTTGGGGCCAGCAGTCTTCAACGGGCACCCGCTCTTCTGCACCCGGGCCAGTTGCGCATGGGGCGGCGCTTCTCTAGCAT
480 L G F A P G L A G Q Q F F N G H P L F L H P G Q F A M G G A F S S M

BspEI (2134)
SgrAI (2145)
2101 GGCTGCAGGCATGGGGCCCTGCTGGCCACAGTATCCGGAGATCCACGGGCTCTCAGGCTAGAAATCCACAGCCATGGCCTCGGCTGCCGAGCGCAG
513 A A G M G P L L A T V S G A S T G V S G L E S T A M A S A A A A Q
2201 GGAAGTGTGGGGCGTGGCTGCCACCTGCCCTTCCACCTCCAACAACAGTCTTGGCTCCAGGGCTTGGCTATGTCGCCTTTCCGGAGCCTGTTTC
547 G L S G A S A A T L P F H L Q Q H V L A S Q G L A M S P F G S L F

SphI (2393)
2301 CTTACCCCTACACATACATGGCTGCAGCAGCTGCAGCCTCCACCGCGCGCCTCCAGTCCGTGCACCGCCACCCGTTCCCTCAATTTGAACAGCATCGC
580 P Y P Y T Y M A A A A A A S T A A A S S S V H R H P F L N L N S M R
2401 CCCAGGCTCGGTTACAGCCCTATTCCATCCCGTCCAGTCCGGATAGCAGCAGTTTGTGGCCACAGCTTACCATCAATGGCTCCGCGCGGGG
613 P R L R Y S P Y S I P V P V P D S S S L L A T A L P S M A S A A G

BsrBI (2577)
2501 CCCCTAGACGGCAAAGCGCCGCTGGCAGCCAGCCAGCCTCGGTGGCTGTGGACTCGGGTGGAACTGAACAGCCGCTCGTCCACGCTGTCTCTG
647 P L D G K A A A L A A S P A S V A V D S G S E L N S R S S T L S S
2601 GCTCAGTGTCTTGTACCCAACTCTGCTCCGAGAAGGAGACGGCTACCAGCGAACTGCAGAGTATCCAGCGGCTGGTCAAGTGGCTTGAAGCAAGCC
680 G S V S L S P K L C S E K E T A T S E L Q S I Q R L V S G L E A K P

2701 AGACAGGTCTTGCAGCGGGTCCCTTAAAAACAAGAAAAACAAATCGCCCCCTCGTAGCTGGCCAGACATGATAAGATACATTGATGAGTTGGACAA
 713▶ D R S C S G S P • NheI (2755)

2801 ACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAGTTAACA HpaI (2893)

2901 ACAACAATTGCATTCATTTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCATAAAA MfeI (2904) EcoRI (2989)

3001 TACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGC

3101 ATTAGCTGTTTGCAGCCTCACCTCTTTTCATGGAGTTAAGATATAGTGATTTTTCCCAAGGTTGAACTAGCTCTTCATTTCTTTATGTTTTAAATGCA

3201 CTGACCTCCACATTCCCTTTTTAGTAAATATTCAGAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTTTATTAGGCAGAATCCAGATGC SspI (3228) SwaI (3242)

3301 TCAAGGCCCTTCATAATATCCCCAGTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTT 141▶ •

3401 AGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGA
 140▶ 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 S

3501 GATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCC
 107▶ 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

3601 TTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCC
 73▶ 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 G

3701 CAGTCTTGGTCTGATGGCCGCCGACATGGTGTGTTGTCTCATAGAGCATGGTATCTTCTCAGTGGGACCTCCACCAGCTCCAGATCTGTCTG
 40▶ 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 Q

3801 AGAGATGTTGAAGTCTTCTATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATG XmnI (3809) AseI (3875)

3901 CGCTCCAGCTTATCTGACGGTCTACTAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTG

4001 TTACGACATTTTGAAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGC SpeI (4030)

4100 TATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAGGTCAT SnaBI (4158)

4200 TGTACTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGT NdeI (4263)

4300 TTACCGTAAATACTCCACCATTGACGTCAATGAAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGCGGGGTCGTTG

4400 GCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCC T G C A G 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 A G G C C A G G A A C C G T SdaI (4441) PacI (4449) BspLU11I (4459)

4498 AAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC

4598 TATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGG

4698 AAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTAGCCC

4798 GACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG

4898 CGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT

4998 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAA

5098 AAAGGATCTCAAGAAGATCCTTTGATCTTTTACGGGGTCTGACGCTCAGTGAACGAAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTA PacI (5189)

5198 CATTTAAATC AGCGGCCGAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAC SwaI (5198)

5298 AAACGAAACAAACAACTAGCAAATAGGCTGTCCCGAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA