



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGTGTCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC
HindIII (245)
Bsu36I (291)
301 GCCATCCACGCGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTTGTTCGTTT
NgoMIV (441)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGAACCGCTCACCATGGGTTTCTGCCGAGCGCCTGCACCCGCT
KasI (535) **MluI (558)**
1 M G F C R S A L H P L

601 GTCTCTCCTGGTGCAGGCCATCATGCTGGCCATGACCCTGGCCCTGGGTACCTTGCTGCCTTCTACCTGTGAGCTCCAGCCCCACGGCCTGGTGAAC
Acc65I (646)
11 S L L V Q A I M L A M T L A L G T L P A F L P C E L Q P H G L V N

701 TGCAACTGGCTGTTCTGAAGTCTGTGCCCCACTTCTCCATGGCAGCACCCCGTGGCAATGTACCAGCCTTCTTGTCTCCAACCGCATCCACCACC
DraIII (746)
45 C N W L F L K S V P H F S M A A P R G N V T S L S L S S N R I H H
801 TCCATGATTCTGACTTTGCCACCTGCCAGCCTGCGGCATCTCAACCTCAAGTGAAGTGCCTCCGCGGTTGGCTCAGCCCCATGCACCTTCCCTGCCA
78 L H D S D F A H L P S L R H L N L K W N C P P V G L S P M H F P C H

901 CATGACCATCGAGCCAGCACCTTCTTGGCTGTGCCACCCTGGAAGAGCTAAACCTGAGCTACAACAACATCATGACTGTGCCTGCGCTGCCAAATCC
BspHI (971)
111 M T I E P S T F L A V P T L E E L N L S Y N N I M T V P A L P K S

1001 CTCATATCCTGTCCCTCAGCCATACCAACATCCTGATGCTAGACTCTGCCAGCCTCGCCGGCCTGCATGCCCTGCGCTTCTATTATGAGCAGGCAACT
NgoMIV (1057)
145 L I S L S L S H T N I L M L D S A S L A G L H A L R F L F M D G N

1101 GTTATTACAAGAACCCTGCAGGCAGGCACTGGAGGTGGCCCCGGTGCCTCCTTGGCTGGCAACCTCACCCACCTGTCACTCAAGTACAACAACCT
SdaI (1115) **XmaI (1140)**
178 C Y Y K N P C R Q A L E V A P G A L L G L G N L T H L S L K Y N N L

1201 CACTGTGGTGGCCCGAACCTGCCTTCCAGCCTGGAGTATCTGCTGTTGTCTACAACCGCATCGTCAAACCTGGCGCCTGAGGACCTGGCCAATCTGACC
KasI (1272)
211 T V V P R N L P S S L E Y L L L S Y N R I V K L A P E D L A N L T
1301 GCCCTGCGTGTGCTGATGTGGCGGAAATTGCCGCGCTGCGACCACGCTCCCAACCCCTGCATGGAGTGCCTCGTCACTTCCCCAGCTACATCCCG
245 A L R V L D V G G N C R R C D H A P N P C M E C P R H F P Q L H P
1401 ATACCTTACAGCCACTGAGCCGCTTGAAGGCCTGGTGTGAAGGACAGTTCTCTCTCTGGCTGAATGCCAGTTGGTCCGTTGGGCTGGGAAACCTCCG
278 D T F S H L S R L E G L V L K D S S L S W L N A S W F R G L G N L R

Ppu10I (1533)
NsiI (1533) 1501 AGTGTGGACCTGAGTGAGAATTCCTCTACAAATGCATCACTAAAACCAAGGCCTTCCAGGGCCTAACACAGCTGCGCAAGCTTAACCTGTCCTTCAAT
311 V L D L S E N F L Y K C I T K T K A F Q G L T Q L R K L N L S F N
1601 TACCAAAAGAGGGTGTCTTTGCCACCTGTCTTGGCCCTTCTTGGGAGCCTGGTGCCTGAAGGAGCTGGACATGCACGGCATCTTCTTCCGCT
345 Y Q K R V S F A H L S L A P S F G S L V A L K E L D M H G I F F R
1701 CACTCGATGAGACCACGCTCCGGCCACTGGCCCGCTGCCATGCTCCAGACTCTGCTGCTGCAGATGAATTCATCAACCAGGCCAGCTCGGCATCTT
378 S L D E T T L R P L A R L P M L Q T L R L Q M N F I N Q A Q L G I F
1801 CAGGGCCTTCCCTGGCCTGCGCTACGTGGACCTGTCGGACAACCGCATCAGCGGAGCTTCCGAGCTGACAGCCACCATGGGGAGGCAGATGGAGGGGAG
411 R A F P G L R Y V D L S D N R I S G A S E L T A T M G E A D G G E
1901 AAGTCTGCTGACGCTGGGACCTTGTCCGGCCAGTGGACCTCCAGCTCGAAGACTTCAAGCCACTGCAGCCACCTCAACTTCCACTTGG
445 K V W L Q P G D L A P A P V D T P S S E D F R P N C S T L N F T L

2001 ATCTGTCACGGAACAACCTGGTGAACGCTGACGCGGAGATGTTTGGCCAGCTCTCGACCTGCAGTGCCTGCGCCTGAGCCACAACCTGCATCTCGCAGGC
BstEII (2019)
478 D L S R N N L V T V Q P E M F A Q L S H L Q C L R L S H N C I S Q A

2101 AGTCAATGGCTCCAGTTCCTGCGCCTGACCGGCTGACGGTGTGAGCTGTCCACAATAAGCTGGACCTTACCACGAGCACTATTACGGAGCTA
AgeI (2128)
511 V N G S Q F L P L T G L Q V L D L S H N K L D L Y H E H S F T E L

2201 CCACGACTGGAGGCCTGGACCTCAGCTACAACAGCCAGCCCTTTGGCATGCAAGGGCTGGGCCACAACCTCAGCTTCTGTTGGCTCACCTGCGCACCCCTGC
SphI (2246)
545 P R L E A L D L S Y N S Q P F G M Q G V G H N F S F V A H L R T L
2301 GCCACCTCAGCCTGGCCACAACAACATCCACAGCAAGTGTCCAGCAGCTCTGAGTACGCTGCTGCGGGCCCTGGACTTCCAGCGCAATGCAGTGGG
578 R H L S L A H N N I H S Q V S Q Q L C S T S L R A L D F S G N A L G

2401 CCATATGTGGCCGAGGGAGACCTCTATCTGCACCTTCTTCCAAGGCCTGAGCGGTTGATCTGGCTGGACTTGTCCAGAACCCTGCACACCCCTCCTG
NdeI (2401)
611 H M W A E G D L Y L H F F Q G L S G L I W L D L S Q N R L H T L L
2501 CCCCAAACCTGCACAACCTCCCAAGAGCCTACAGGTGCTGCTGCTCCGTGACAATACCTGGCCTTCTTAAAGTGGTGGAGCCTCACTTCTGCCCA
645 P Q T L R N L P K S L Q V L R L R D N Y L A F F K W W S L H F L P

2601 AACTGGAAGTCTCGACCTGGCAGGAAACAGCTGAAGGCCCTGACCAATGGCAGCCTGCCTGCTGGCACCCGGCTCCGGAGGCTGGATGTCAGCTGCAA
BspEI (2675)
678 K L E V L D L A G N Q L K A L T N G S L P A G T R L R R L D V S C N

2701 CAGCATCAGTTCGTGGCCCCGGCTTCTTTTCCAAGGCCAAGGAGCTGCGAGAGCTCAACCTTAGCGCCAACGCCCTCAAGACAGTGGACCACTCCTGG
711▶ S I S F V A P G F F S K A K E L R E L N L S A N A L K T V D H S W
2801 TTTGGGCCCTGGCGAGTGCCTGCAAACTACTAGATGTAAGCGCAACCTCTGCACTGCGCCTGTGGGGCGGCTTTATGGACTTCTGTGGAGGTGC
745▶ F G P L A S A L Q I L D V S A N P L H C A C G A A F M D F L L E V
2901 AGGCTGCCGTGCCGGTCTGCCAGCGGGTGAAGTGTGGCAGTCCGGGCCAGCTCCAGGGCCTCAGCATCTTTGACAGGACCTGCGCCTCTGCCTGGA
778▶ Q A A V P G L P S R V K C G S P G Q L Q G L S I F A Q D L R L C L D
3001 TGAGGCCCTCTCTGGGACTGTTTCGCCCTCTCCTGCTGCTGGCTGTGGCTTGGCCATGCTGCATCACCTCTGTGGCTGGGACCTCTGG
811▶ E A L S W D C F A L S L L A V A L G L G V P M L H H L C G W D L W
3101 TACTGCTTCCACCTGTGCTGGCTTCCCTGGCGGGCGCAAAGTGGCGAGATGAGATGCCCTGCCCTACGATGCCCTTCGTGGTCTTCGACA
845▶ Y C F H L C L A W L P W R G R Q S G R D E D A L P Y D A F V V F D

BsrGI (3227)

3201 AAACGCAGAGCGCAGTGGCAGACTGGGTGTACAACGAGCTTCGGGGGAGCTGGAGGAGTGCCGTGGCGCTGGGCACTCCGCTGTGCCTGGAGGAACG
878▶ K T Q S A V A D W V Y N E L R G Q L E E C R G R W A L R L C L E E R

Tth111I (3388)

RsrII (3386)

3301 CGACTGGCTGCCTGGCAAAACCTCTTTGAGAACCTGTGGCCTCGGTCTATGGCAGCCGCAAGACGCTGTTTGTGCTGGCCACACGGACCGGGTCACT
911▶ D W L P G K T L F E N L W A S V Y G S R K T L F V L A H T D R V S

BssHII (3408)

3401 GGTCTCTTGCAGCCAGCTTCTGCTGGCCAGCAGCGCTGTGGAGGACCGCAAGGACGCTGTTGCTGGTATCCTGAGCCTGACGGCCGCCGCT
945▶ G L L R A S F L L A Q Q R L L E D R K D V V V L V I L S P D G R R
3501 CCCGCTACGTGCGGCTGCCAGCGCCTCTGCCGCCAGAGTGTCTCCTCTGGCCCCACAGCCAGTGGTCAGCGCAGCTTCTGGGCCAGCTGGGCAT
978▶ S R Y V R L R Q R L C R Q S V L L W P H Q P S G Q R S F W A Q L G M

SandI (3648)

NheI (3688)

3601 GGCCTGACCAGGGACAACCACCACTTCTATAACCGAACTTCTGCCAGGACCCAGCCGAATAGCCGTGAGCCGGAATCTGCACGCTAGCTGGCCA
1011▶ A L T R D N H H F Y N R N F C Q G P T A E •

3701 GACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTG

HpaI (3826) MfeI (3837)

3801 TAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTA

EcoRI (3922)

3901 AAACCTCTACAAATGTGGTATGGAATTTCTAAAATACAGCATAGCAAACCTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATA
4001 AGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTCCCAAGTGTGA

SspI (4161)

SwaI (4175)

4101 ACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAA

4201 ATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGAACCTTTAATAGA

4301 AATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCATCTC

141▶

• N R T Y K L P I L E E I T T K V L K G N M E

4401 AATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAG

118▶ 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 R D V E D S Y

4501 GGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCAGACAGTGCACCTGCCAATGT

84▶ 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 C T A T V T V R G I Y

4601 AGGCCTCAATGTGGACAGCAGATGATCTCCAGTCTTGGTCTGTGCGCCCGGACATGGTGTGTTGCTCATAGATGATGATCTTCTC

51▶ 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 Y L M T I K E

BspHI (4750)

XmnI (4742)

4701 AGTGGCGACTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTA

18▶ T A V E V L E L D Q Q S I N F T K M

AseI (4808)

4801 TGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGC

SpeI (4963)

4901 CTACCGCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAACCTCCATTGACGTCAATGGGG

SnaBI (5091)

5000 TGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCCAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGA

NdeI (5196)

5100 TGTACTGCCAAGTAGGAAAGTCCATAAGGTACTGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCAATAGGGGGCTACTTGGCAT

5200 ATGATACACTTGATGTAAGTGGCAGTTTACCCTAAACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATAC

SdaI (5374)

PacI (5382)

BspLU11I (5392)

5300 GTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACCGCTG C A G G T T A T T A A G A A C A T G

5398 TGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAATAATCGAC

5498 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCGTGAAGCTCCCTCGTGCCTCTCTGTTCCGACCTGCCGCT

5598 TACCGGATACCTGTCGCTTTTCTCCCTCGGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAG

ApaI (5706)

5698 CTGGGCTGTGTGCACGAACCCCGTTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC
5798 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGCCTAACTACGGCTACACTAGAAGAAC
5898 AGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTT
5998 TTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCAC

PacI (6122) SwaI (6131) NotI (6141)

6098 GTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGTTTTTGTGT
6198 GAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTC
6298 TCTATCGAA