



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
HindIII (245)
301 GCCATCCACGCCGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT
NgoMIV (441)
NaeI (441)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTACCATTGAGAGTGGTGGGCGACTTCGAGTACTGCAAGCGGGA
KasI (535) **BstEII (555)** **NcoI (560)** **AgeI (552)** **Scal (584)**
1 M E V V G D F E Y C K R D
AseI (682)
601 CCTCGTGGGACACGGGCTTCCGCTGTGGTCTTCCGGGGCGGCACCGCCAGAAAAGTATTGGGAGTGGCTATTAAGTATTAATAAAAAGAACTTG
13 L V G H G A F A V V F R G R H R Q K T D W E V A I K S I N K K N L

701 TCAAAATCACAATTCTGCTTGGAAAGAAATAAAAATCTTAAAGGAGCTTCAGCATGAAAACATCGTAGCGCTCTATGATGTTCAAGAAATGCCAACT
Eco47III (768)
AfeI (768)
47 S K S Q I L L G K E I K I L K E L Q H E N I V A L Y D V Q E L P N
801 CTGCTTTCTGGTATGGAGTATTGCAATGGTGGAGACCTGGCAGATTATTTGCAAGCTAAAGGAACTCTGAGTGAAGATACTATCAGAGTGTTCCTCA
80 S V F L V M E Y C N G G D L A D Y L Q A K G T L S E D T I R V F L H
901 TCAGATTGCGGCAGCCATGCAATCCTGCACAGCAAAGGATAATCCACAGGATCTCAAACACAGAATATCCTGTTGTCTTATGCCAATCGAAGGAAG
113 Q I A A A M R I L H S K G I I H R D L K P Q N I L L S Y A N R R K

1001 TCGAATGTCAGTGGTATTGCTATTAATAAGTGTGATTTGGTTTGCACCGTACATAGTAACACAATGGCAGCGACACTGTGGATCCCCAATGT
Acc65I (1049) **BamHI (1087)**
147 S N V S G I R I K I A D F G F A R Y L H S N T M A A T L C G S P M
1101 ACATGGCTCCCGAGGTTATTATGCTCAACATTATGATGCTAAGGCAGATTTATGGAGCATAGGAACAGTGTATCAATGCCTAGTTGGAAAACCACC
180 Y M A P E V I M S Q H Y D A K A D L W S I G T V I Y Q C L V G K P P
1201 TTTTCAGGCTAATAGTCTCAGGACCTAAGGATGTTTTATGAAAAAAGCAGGAGCTTAATGCCTAGTATTTCCAGAGAAAACATCACCTTACTTGGCTAAT
213 F Q A N S P Q D L R M F Y E K N R S L M P S I P R E T S P Y L A N
1301 CTCCTTTGGGTTGCTTCAGAGAAATCAAAGGATAGAAATGGACTTTGAAGCATTTCAGCCATCCTTTCCCTTGGAGCAAGTTCAGTTAAAAAATCTT
247 L L L G L L Q R N Q K D R M D F E A F F S H P F L E Q V P V K K S
1401 GCCAGTCCAGTGCCTGTGATTCTGGCCCTGTCCCTGGAAGCTCCTGCAGCAGCTACCATTGTGCGTTTGGCTTCCACCATCCCTTCCAGATAT
280 C P V P V P V Y S G P V P G S S C S S S P S C R F A S P P S L P D M
1501 GCAGCATATTCAGGAAGAAAATCTCCACCCGCTGGTGGTCTCCCAACTATCTACAGGTGTCCAAAGACTCTGCAGTAATAGTAGCAAGAATCTCT
313 Q H I Q E E N L S S P P L G P P N Y L Q V S K D S A S N S S K N S
1601 TCTTGTGACACGGATGACTTTGTTTGGTTCACACAACATCTCGTCAGACCCTCATATGACATGCCAATGGGGACTACGGCCAGACGTGCTTCAAATG
347 S C D T D D F V L V P H N I S S D H S Y D M P M G T T A R R A S N
1701 AATTCTTTATGTTGGAGGGCAGTGTCAACCTACTGTGCACCTCACAGCGAAACAGCCCAATTCCAGTTCCTACTCAAGTAAGGAATTATCAGCGCAT
380 E F F M C G G Q C Q P T V S P H S E T A P I P V P T Q V R N Y Q R I

1801 AGAACAGAATCTTATATCCACTGCCAGCTCTGGCACAAACCCACATGGTTCTCCAAGATCTGCAGTAGTACGAAGGTCTAATACCAGCCCCATGGGCTTC
BglII (1855) **NcoI (1889)**
413 E Q N L I S T A S S G T N P H G S P R S A V V R R S N T S P M G F

1901 CTCCGGTTGGTCTGCTCCCTGTACCAGGAGACAGTGCAGACAGGAGGACGAAGACTCTCTACTGGCTTCCAGGCTTACTACCATCCCCTT
StuI (1978)
447 L R V G S C S P V P G D T V Q T G G R R L S T G S S R P Y S P S P

2001 TGGTTGGTACCATTCTGAAACAGTTTAGTCAAGTGTGCTGTGGACATCCTCAGGGCCATGAAGCCAGGAGTAGGCACTCCTCAGGTTCTCCAGTACCACA
Acc65I (2005)
480 L V G T I P E Q F S Q C C C G H P Q G H E A R S R H S S G S P V P Q

2101 GACCCAGGCACCACAGTCACTCTTACTGGGTGCTAGACTGCAGAGTGCACCCACCCTCACCGATATCTATCAGAACAAGCAGAAGCTCAGAAAGCAGCAC
EcoRV (2161)
BsaBI (2161)
513 T Q A P Q S L L L G A R L Q S A P T L T D I Y Q N K Q K L R K Q H
2201 TCTGACCTGTGTGCTCCATGCTGGAGCTGGGTATAGTACTCACCTCAGCCTAGTCCGGCTGGCAGCCTTGGACCTCTCCACCAAGCACAGG
547 S D P V C P S H A G A G Y S Y S P Q P S R P G S L G T S P T K H T
2301 GGTCTCTCCACGGAATCTGACTGGTTCTTTAAACTCCTTTACCAACAATCATTGGCTCTCCTACTAAGACTACAGCTCCTTTCAAATCCCTAAAAC
580 G S S P R N S D W F F K T P L P T I I G S P T K T T A P F K I P K T
2401 ACAAGCATCTTAACTGTAGCTTGGTTACTCGTCATGGGCTGTGAAAGCCAGTCCAAAGATGGGAATGACCCCTGTGAGTGTCCCACTGCCTC
613 Q A S S N L L A L V T R H G P A E S Q S K D G N D P R E C S H C L

2501 TCAGTACAAGGAAGCGAGAGGCATCGATCTGAGCAGCAGCAGAGCAAGGCAGTGTGGCAGATCTGTGAGTACTGGGAGTTATCAGAACAACAAGTAA
ClaI (2522) **BglII (2560)** **Scal (2569)**
647 S V Q G S E R H R S E Q Q S K A V F G R S V S T G K L S E Q Q V

2601 AGGCACCTTTAGTGGACACCAGGGCAGCACGGATAGTTTAAACACAGAACGACCAATGGATGTAGTCTCTGCAGGAGCCTGTGGTGTATGCTGGCATT
PmeI (2636)
680 K A P L G G H Q G S T D S L N T E R P M D V A P A G A C G V M L A L

2701 GCCAGCAGGAACAGCAGCAAGCGCCAGAGCTGTCTCTTACCCTGGGGTCTCTCCACACAGTGCCACAGCCCCACTTGTACTCATATGGTCTTCGA
713 P A G T A A S A R A V L F T V G S P P H S A T A P T C T H M V L R
2801 ACAAGAACCACCTCAGTGGGGTCCAGCAGCTCAGGAGTTCTTGTGTCTGCAAGTGGCCGAGTATGTGTGGCTCCCTCGACAGGGTTGGCT
747 T R T T S V G S S S S G G S L C S A S G R V C V G S P P G P G L G
2901 CTTCCCCACAGGAGCAGAGGGAGCTCCAGCCTAAGATACGTGCCTTATGGTGTCCACCACCCAGCCTAGAGGGTCTCATCACCTTTGAAGCCCTGA
780 S S P P G A E G A P S L R Y V P Y G A S P P S L E G L I T F E A P E
3001 ACTACCAGAGGAGACTGATGGAGCGAGACACAGACACCTTACGCCATCTGAACATGATGTTAATGTTTACTGAGTGTGTGGACCTGACGGCA
813 L P E E T L M E R E H T D T L R H L N M M L M F T E C V L D L T A
3101 GTGAGGGTGGGAACCTGAGCTGTGCACATCTGCTGTCTTGTACAGATTAGGAGAGTGTAGTGTGGACCAGATCAGCCAGCTAAGCAAAGATT
847 V R G G N P E L C T S A V S L Y Q I Q E S V V V D Q I S Q L S K D

BsrGI (3224)

3201 GGGGGCGGGTGGAGCAGCTGGTGTGTACATGAAGGCAGCACAGCTGCTGGCGGCTTCCCTGCATCTCGCAAAGCTCAGGTCAAGTCTGGGAAGCTGAG
880 W G R V E Q L V L Y M K A A Q L L A A S L H L A K A Q V K S G K L S

NcoI (3305)

3301 CCCATCCATGGCTGTGAAACAAGTTGTTAAAAATCTGAATGAAAGATACAAAATCTGCATCACCATGTGCAAGAACTTACAGAAAAGCTGAATCGCTTC
913 P S M A V K Q V V K N L N E R Y K F C I T M C K K L T E K L N R F
3401 TTCTCCGATAAACAGAGATTTATTGATGAAATCAACAGTGTGACTGCAGAGAACTCATATAATTGTGCTGTGAAATGGTCAATCTGCAGCCCTGG
947 F S D K Q R F I D E I N S V T A E K L I Y N C A V E M V Q S A A L
3501 ATGAGATGTTTCAGCAGACTGAAGACATCGTTTATCGTACCACAAGGAGCCCTTCTTTTGAAGGCTTAAGTAAGATCCTGCAGGACCTACAGATGT
980 D E M F Q Q T E D I V Y R Y H K A A L L L E G L S K I L Q D P T D V
3601 TGAATGTGCATAAGTATAATGTAGTATTGAAAGAAGATTGTGAGCACTCTGCTGTAGCACTGCAACTGTGTGAGTACAGGCTTGTCTGTGGACTGG
1013 E N V H K Y K C S I E R R L S A L C C S T A T V •

MscI (3718)

NheI (3712)

3701 CATGGAACAGGAGCTAGCTGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAA

HpaI (3850) MfeI (3861)

3801 TTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAGTTAAACAACAACATTGCATTCATTTTATGTTTCAGGTTACAGGGGAGGT
3901 GTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTT
4001 GAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGAT

SspI (4185)

SwaI (4199)

4101 ATAGTGTATTTTCCCAAGTTTGAAGTACTGCTTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTATGTAATAATTCAGAAATAA
4201 TTTAAATACATCATTGCAATGAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTA
4301 GGGAAACAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGCTGTTACTTGAGGGGGATGAGTTCCTCAATGGTGGT
4401 TTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATG
126 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 P S V V R I
4501 GATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAG
92 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 G I A I A E A

StuI (4624)

4601 CACAGACAGTGACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTCTTGTGTG
59 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

BspHI (4774)

4701 CTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGATGGCCCTCTATAGTGAGTC
26 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

AseI (4832)

4801 GTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTATCTGACGGTTCCTAAACGAGCTCTGCTT

SpeI (4987)

4901 ATATAGACCTCCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCCTGTTGATTTACTAGTCAAAAC
5000 AAACCTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGG

SnaBI (5115)

5100 TAATAGCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCCATAAGGTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGAC
5200 GTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCT
5300 ATTGGCGTTACTATGGAAACATACGTCATTATTGACGTCATGGGGGGGTCGTTGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGC

PaeI (5406)

BspLU11I (5416)

5400 CTGCAGGTTAAATAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCC
5498 CTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACAGGCGTTTTCCCTGGAAGCTCCCTCGTGCG
5598 CTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTCTCCCTTGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGATAGGTATCTCAGT
5698 TCGGTGATAGTCTGCTCCAAGCTGGGCTGTGTGCACGAACCCCGCTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGTAGTCCAACC

5798 CGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCC

5898 TAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAA

5998 ACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTG

EagI (6166)

PacI (6146) SmaI (6155) **NotI (6165)**

6098 ACGCTCAGTGAACGAAAACACGTTAAGGGATTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTCATTA

6198 CATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTG

6298 CAAGTGCAGGTGCCAGAACATTTCTCTATCGAA