



PvuI (7)
SgfI (6) 1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAAGTGGCGCGGGTAAACTGGGAAAGTGTCTGTACTGGCTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTCACGCGCCCGCCCTACCTGAGGCC
HindIII (245)
Bsu36I (291)
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BstEII (555)
AgeI (552) 501 TCTGTTTGTCCGGCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTACCATTGGCGGCGACTACAGCTAACCCAGAAATGACATCAGA
NcoI (560)
1▶ M A A T T A N P E M T S D

SandI (612) 601 TGTACCATCGTGGTCCACCATTGCTTCTGAAACCCTGGACCTGGGATCAAGGTGGAGGAGCTGTTGTACAGAGGGCTATTAAGCGACGGTCAGGG
13▶ V P S L G P T I A S G N P G P G I Q G G G A V V Q R A I K R R S G
BsrGI (669)
701 CTGGATTTTGTATGATGAAGTAGAAGTGAACACTAAATTTTTGAGATGCGATGATGACCAGATGTGTAATGACAAAGAGCGGTTTGCCAGGGAAAATCATA
47▶ L D F D D E V E V N T K F L R C D D D Q M C N D K E R F A R E N H
BsrBI (775)
801 GTGAAATAGAACGGCGGCACGGAACAAGTACAGCTTACATCACAGAACTGTCAGACATGGTACCTACATGTAGTGCCTGGCTCGAAAACAGACAA
80▶ S E I E R R R R N K M T A Y I T E L S D M V P T C S A L A R K P D K
901 GCTAACCATCTTACGCATGGCCGTTTCTCACATGAAGTCTTGAGGGAACTGGCAACACATCTACTGATGGCTCCTACAAGCCATCTTCCCTCACTGAT
113▶ L T I L R M A V S H M K S L R G T G N T S T D G S Y K P S F L T D
1001 CAGGAAGTGAACATTTGATCTTGAGGCGAGCAGATGGCTTCTGTTATTGTCTCCTGTGAGACTGGACGGTGGTGTATGTCTCTGACTCAGTGACTC
147▶ Q E L K H L I L E A A D G F L F I V S G E T G R V V Y V S D S V T
ApaLI (1153)
1101 CCGTTTTGAACGACCCAGTCTGAATGGTTCGGGAGCACACTGTATGATCAGGTGCACCCAGATGATGTGGATAAATTCGAGAGCAGCTCTCTACATC
180▶ P V L N Q P Q S E W F G S T L Y D Q V H P D D V D K L R E Q L S T S
1201 AGAAAATGCCCTAACAGGGCGGGTCTGGATCTGAAGACTGAAACAGTGA AAAAGGAAAGGCCAGCAGTCTCCATGAGGATGTGCATGGGCTCACGAAGG
213▶ E N A L T G R V L D L K T G T V K K E G Q Q S S M R M C M G S R R
1301 TCGTTCATCTGCCGATGAGGTGGTACTAGCTCCGTGGACCCTGTTCCATGAATAGACTGAGCTTTTTGAGGAACAGATGCAGGAATGGGCTTGCT
247▶ S F I C R M R C G T S S V D P V S M N R L S F L R N R C R N G L G

DraIII (1420) 1401 CTGTGAAGGAAGGAGAACCTCACTTTGTGGTAGTCCACTGCACAGGCTACATCAAGGCTGGCCACCAGCAGGTGTCTCCCTCCAGATGATGACCCAGA
280▶ S V K E G E P H F V V V H C T G Y I K A W P P A G V S L P D D D P E
StuI (1454)
1501 GGCTGGCCAGGGGAGCAAATTCGCCTAGTGGCCATTGGCAGGCTGCAGGTAAGTCTCCCACTGTACAGACATGAGTAACATTTGTGAGCCAAACA
313▶ A G Q G S K F C L V A I G R L Q V T S S P N C T D M S N I C Q P T

BstAPI (1514) 1601 GAGTTCATCTCCGACACAACATTGAAGGGATATTCACCTTTGTAGACCATCGTTGTGTGGTACTGTTGGTACCAGCCACAGGAGCTCTTAGGGAAGA
347▶ E F I S R H N I E G I F T F V D H R C V A T V G Y Q P Q E L L G K
SspI (1699)
1701 ATATTGTAGAATTTTGTATCCTGAAGACCAACAATTCTAAGAGACAGCTTTCAGCAGGTGGTGAATTA AAAAGGTCAGGTGCTGTCCGTCATGTTCCG
380▶ N I V E F C H P E D Q Q L L R D S F Q Q V V K L K G Q V L S V M F R
1801 ATCCGATCTAAGACCCGAAATGGCTGTGGATGAGAACGAGCTCCTTTACCTTCCAAAACCTTATTCAGATGAAATTTAGTATATTATCTGACCAACA
413▶ F R S K T R E W L W M R T S S F T F Q N P Y S D E I E Y I I C T N
1901 ACCAATGTGAAGAACTTAGCCAGGAACACGGCCTACACTGTCCAACACCATCCCAAGGTCACAGCTAGGTCGACAGCCAAATTTATCCCTAGAGATGG
447▶ T N V K N S S Q E P R P T L S N T I P R S Q L G P T A N L S L E M
2001 GTACAGGGCAGCTGCCATCCAGGCGAGCAGCAGCAGCACACAGAAGTGGATATGGTACCAGGAAGAGATGGGCTGGCCAGCTATAATCATTCCAGGT
480▶ G T G Q L P S R Q Q Q Q H T E L D M V P G R D G L A S Y N H S Q V
2101 TTCTGTCCAGCCTGTGGCAAGTGCAGGATCAGAACACAGCAAGCCCTTGAGAAGTCAAGGCTCTTTGACAGGACAGATCCAAGGTTTCCAGAA
513▶ S V Q P V A S A G S E H S K P L E K S E G L F A Q D R D P R F F A
2201 ATCTATCCAGCATCACTGCAGATCAGAGTAAAGGCATCTCCTCCAGCACTGCTCCCTGCCACCAACAGCTGTTCTCCAGGGCAGCTACCTCCCTA
547▶ I Y P S I T A D Q S K G I S S S T V P A T Q Q L F S Q G S S F P P

NgoMIV (2307) 2301 ACCCCCGCCGGCAGAGAATTTAGGAATAGTGGTCTTACCCCTCCTGTAAACATTGTCCAGCCATCATCTTCTGCAGGGCAGATACTGGCCAGATTTCC
580▶ N P R P A E N F R N S G L T P P V T I V Q P S S S A G Q I L A Q I S
EagI (2305)
2401 ACGTCACTCCAACCTGCCAGGGATCAGCGCCAGCTGGACCTTAGCTCCCGCCAGGCTTTGCCGCCAGCAGGTGCCACCCAGGCTACAGCCAAG
613▶ R H S N P A Q G S A P T W T S S S R P G F A A Q Q V P T Q A T A K
2501 ACTCGTCTTCCAAATTTGGTGTGAACAATTTAGACTTCTCCTCCTCAGTGCTATGTCTCTCCGGGTGCTCCACTGCCTCATCTGGTACTGCTG
647▶ T R S S Q F G V N N F Q T S S S F S A M S L P G A P T A S S G T A

2601 CCTACCCTGCTCTCCCAACCGTGGCTCCAACCTTCTCTGAGACTGGACAGACCACAGGACAGTTCAGGCCCGGACAGCAGAGGGCGTGGGTGTCTG
680▶ A Y P A L P N R G S N F P P E T G Q T T G Q F Q A R T A E G V G V W
Bsu36I (2788)

2701 GCCACAGTGGCAGGGCCAGCAGCCCATCATCGGTCTAGTTCAGTGGACAGCATGTTTACGACAGACACAAGCACAAGCACCTAGCCAGCCTGAGGTCTTT
713▶ P Q W Q G Q Q P H H R S S S S E Q H V Q Q T Q A Q A P S Q P E V F
XmnI (2848)

2801 CAAGAAATGCTGTCCATGCTGGGAGACCAAAGCAACCTACAACAATGAAGAATTTCTGATCTAACTATGTTTCCCCCTTTTCCGAATAGAACTATT
747▶ Q E M L S M L G D Q S N T Y N N E E F P D L T M F P P F S E •
NheI (2919)

2901 GGGGTGAGGATAAGGGTGGGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAAATGCAGTGAAAAAATGCTTTATT

3001 TGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAACAAACAATTGCATTCATTTTATGTTTCAGGTTTCAGG
HpaI (3057) MfeI (3068)

3101 GGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCATAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCC
EcoRI (3153)

3201 TCTACTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCCTCACCTCTTTTCATGGAGT

3301 TTAAGATATAGTGTATTTTCCCAAGTTTGAAGTCTCTTTCATTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTATAGTAAATATTCA
SapI (3335) SspI (3392)

3401 GAAATAATTTAAATACATCATTGCAATGAAATAAATGTTTTTTATTAGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTT
SwaI (3406)

3501 GGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAA
141▶ • N R T Y K L P I L E E I
BstXI (3696)

3601 TGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCCAGGACATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCAC
128▶ 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 P S V V
3701 CCTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATG
95▶ 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 G I A I
StuI (3831)

3801 GCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTCT
61▶ 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 V H H K
XmnI (3973)

3901 TGTTGCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATA
28▶ 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
AseI (4039)

4001 GTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGC

4101 TCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGTATTACTAGT
SpeI (4194)

4201 CAAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGTATCCACGCCATTGATGTACTGCCAAAACCGCATCA
SnaBI (4322)

4301 TCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCACTGACTGGGCATAATGCCAGGCGGGCCATTTACCGTC

4401 ATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCTCAAGTGGGCAGTTTACCCTAAATACTCCACCATTGACGTCAATGGAAA
NdeI (4427)

4501 GTCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGGCGGGTCTGTTGGCGGTGACGACGGCGGGCCATTTACCCTAAGTTATG

4601 TAACGCTGCAGGTTAAATAAGAACATGTGAGCAAAGGCCAGAAAGCCGTAAGGCGCGTTGCTGGCGTTTTTCCATAGGCTCCGC
SdaI (4605) PacI (4613) BspLU11I (4623)

4701 CCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATAACAGGCGTTTCCCCTGGAAGCTCCCTCG

4801 TGGCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCT

4901 CAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCC
ApaLI (4937)

5001 AACCCGGTAAGACACGACTTATCGCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGT

5101 GGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAA

5201 ACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGG

EagI (5373)

PacI (5353) SmaI (5362) **NotI (5372)**

5301 TCTGACGCTCAGTGAACGAAAACACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAATAAAATATCTTTATTTTC

5401 ATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCC

5501 AGTGAAGTGCAGGTGCCAGAACATTTCTATCGAA