



PvuI (7)
SgfI (6) 1 GGATCTCGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)
Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACCGCGCCGCCCTACCTGAGGCC
PvuII (239)
301 GCCATCCACGCGGTTGAGTGCAGTCTGCCGCTCCCGCCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCGCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

SphI (566)
MluI (558) 501 TCTGTTCTGCGCGGTTACAGATCCAAGCTGTGACCGCGGCTACCTGAGATCACCGGACGCGTCAGCATGCTCCCTGAAGCCAGTTCCTGTGGTCTCT
1▶ M L P E A S S L W L L
601 TCGGCTACTCCGGGATGTTCAACTGGCCAGTTTTACCGGCCATCCTTGAAGAACTCAACGTCCTCGGCCAGAATTTTGAAGTCTGAG
11▶ R L L R D V Q L A Q F Y R P I L E E L N V T R P E H F D F V R P E

AgeI (764)
Acc65I (761) 701 GACCTGGACAACATAGGCATGGGAGGCTGCCAGCGCAGACTTAATGAAGCCCTGAAGAGGTACCGGTCAGGGTCAAGTCCAAGAATTGGGTGTACA
45▶ D L D N I G M G R P A Q R R L N E A L K R Y R S G V K S G K N W V Y
801 AGATACTGGGGCTTTGCTCCCGAAGCAGAAAGGATCCCTCCACGCTCAGACAGTCTGTCTTCCATGAGCCAGAGGGGGAGCTCAAGTGTCTGAT
78▶ K I L G G F A P E Q K E I P P R S D S P L C F H E P E G G L K C L I

BsrGI (794)
901 TCCAGAGGTGCTGTGCGCAGAGGGGAGCTGCTGGGTTGAGGCTGCTTTGGTGTGTACATCGAGGGTATGGACTACTCCAGTGGGAGAGTATCCCA
111▶ P E G A V R R G E L L G S G C F G V V H R G L W T L P S G Q S I P

SandI (1023) 1001 GTGGCGTCAAGTCCCTCCGTGTGGTCCCGAGGGTCCGATGGGCACAGAAGTGGGCGACTTCTCGGGGAAGTATCTGTCTATGATGAAGCTAGAGCACC
145▶ V A V K S L R V G P E G P M G T E L G D F L R E V S V M M K L E H

BspHI (1079)
BbrPI (1102) 1101 CACACGTGCTGCGCTGCATGGCCTCGTACTGGCCAGCCTCTACAGATGGTGTGAACTTGCGCCACTGGGCTCCCTGCATACGCGCCTGACTGCCCC
178▶ P H V L R L H G L V L G Q P L Q M V M E L A P L G S L H T R L T A P

XcmI (1262)
NcoI (1262) 1201 AGCACCGACACCCCACTGCCTGTGGCCCTGCTGTGCCTTTTCTGCGTCAGTTGGCAGGACCATGGCGTACTTGGGGTCTGTGGGTTAGTGACCCGG
211▶ A P T P P L P V A L L C L F L R Q L A G A M A Y L G S C G L V H R

ApaLI (1291)
XmaI (1385) 1301 GATCTCGCCACCCGCAACTGCTGCTGGCCTCACCTCGAATGATCAAGGTGGCGGATTTTGGACTGGTTCGGCCACTGGGAGGCGCCGGGACGCTACG
245▶ D L A T R N L L L A S P R M I K V A D F G L V R P L G G A R G R Y
1401 TCATGGGCGGGCCCGTCTATCCCTATGCTGGTGTGCCAGAGGCTACGCGAGGAGCCTTCTGCTGCTGCTGATGTGGATGTTTGGCGT
278▶ V M G G P R P I P Y A W C A P E S L R Q G A F S S A S D V W M F G V
1501 GACTCTGGGAAATGTTTCTGGAGGTGAGGAACCTGGGCGAGTGTCCACCATATCTCATCTACAGCGACTGGAGAAGGACCGAGCCAGGCTGCCA
311▶ T L W E M F S G G E E P W A G V P P Y L I L Q R L E K D R A R L P
1601 AAGCTCCCTGTGTTCCAGGGCCCTACTCCCTGCGCTTGGCGTGTGGGCCCCACCCCTGCAGACCGGCTAGCTTTTCCAACCTGGAAGGGCTGC
345▶ K P P L C S R A L Y S L A L R C W A P H P A D R P S F S N L E G L

BstEII (1774)
1701 TTCAAGAGCTTGGCTTTCTGAGGGACGCTGTGTGAGGGAGGTACAGAGCCAGGTGCTTTGCGGATGGAACCCGGTACCCCATCACTATCATTGAGGG
378▶ L Q E A W L S E G R C V R E V T E P G A L R M E P G D P I T I I E G

SfiI (1825)
1801 CAGCCTGGACTGCAACTGAAAGGCCAGAATGGCCGACCTCAAAGTGGGCAACTTCCGGCTCCGAGTGACTAGCAGACTTGGGGGTTTCA
411▶ S L D T A T W K G Q N G R T L K V G N F P A S A V T L A D L G G S

NcoI (1935) 1901 CCAGTCAACCATCCAGCCACAGAGGTTCCCTGCCATGGAGAGAAATGCAGAGGGGGCACAGATGGGGACAGAGAAGGCCAACGTTACAAGACCTAC
445▶ P V T H P A H R G S P A H G E K C R G G T D G D R E K A T L Q D L

BstAPI (2034) 2001 CCCAGCAGCGGACACAGAACGAAGATGCCCCGTCAGAGGATGCGAGGCATTTCCAAGAGTCTGGAATCAGTCTGTCCCTGGGCCCTAGACCCACAGG
478▶ P P A R S H R T K M P L Q R M R G I S K S L E S V L S L G P R P T G
2101 AGTGGTTCAAGTCTCCAGAACTCGGCGTACAAGAGCTACCGGCAAGACTCCAGACTCCAGACTCCAGCCACCCAGAGACTGCTCCAGCCACCT
511▶ G G S S P P E L R R T R A M P Q R L P D L P P R P P D L P P R P P
2201 ATTATCTGCAACTCTTCTCAGCCACCCAGCCCAAAAGCCCGCCAAAAGAGAATCTCACATAATCACCAGCTGGAGCACCTGGAGCCAGCAAAG
545▶ I I C N S S Q P T Q P H K A R P K R E S S H N H R T G A P G A S K
2301 CCACTGTCCCTCTGGAGGTCCCTTGTGACACCCGAGTGGCAGAGGAAGTTGTAGAGGTAGAGTGTGATGGTGTACCTACCAGGAGTGCCA
578▶ A T V P S G G P L S D P E W Q R K V V E V E L S V H G V T Y Q E C Q

BspEI (2441) 2401 AGTGGACTCAGAACCCTGGGGAGATGTGGCTTCTGCTATCCGGAACCTCAAGGTAGACCAGCTCTCCACCTTAGTAACCGTCCCGAGCAGACTGT
611▶ V A L R T T G G D V A S A I R N L K V D Q L F H L S N R S R A D C

SapI (2463) 2501 CGTCGCATCTGGAGCATCACCAGTGGGACCTGTGACGAGCCAGTCGCTACATCTAGCCGGTCTGAACTCTGTATCTCAGGGATGGACAGCTAGCT
645▶ R R I L E H H Q W D L S A A S R Y I L A R S •

AgeI (2480)
PshAI (2494)
MscI (2599)
NheI (2593)

2601 GGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTT

2701 ATTTGTAACCATTATAAGCTGCAATAAACCAAGTTAACAAACAACCAATTGCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGC

2801 AAGTAAAACCTCTACAAATGTGGTATGGAATTCATAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGAT

2901 GAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTTCCAAGG

3001 TTTGAAGTCTCTTCATTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTATGAAAAATTTCAGAAATAATTTAAATACATCATTGCAA

3101 TGAAAAATAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTA

3201 ATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTC
 141 • N R T Y K L P I L E E I T T K V L K G N

3301 ATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAG
 119 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 S R D V E D S

3401 AGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCAGACAGCTGACCCTGCC
 86 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 C V A T V R G

3501 AATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAQTCTGGTCTGATGGCCGCCAGACATGGTGCCTTGTTCCTCATAGAGCATGGTGATC
 53 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 E Y L M T I

3601 TTCTCAGTGGCGACTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATA
 19 K E T A V E V L E L D Q Q S I N F T K M

3701 TACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCCACCGTA

3801 CACGCTACCGCCATTGCGTCAATGGGGCGGAGTTGTTACGACATTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCCATTGACGTCAAT

3901 GGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACG

4001 TAGATGTAAGTCCAAAGTAGGAAAGTCCATAAGGTCATGTAAGTGGCATAATGCCAGGCGGGCATTACCCTGATTGACGTCAATAGGGGGCGTACTTG

4101 GCATATGATACACTTGATGTAAGTCCCAAGTGGGCGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAAC

4201 ATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTACGCCAGGCGGGCCATTTACCCTAAGTTATGTAACGCTGCAGGTTAATTAAGAACA

4301 TGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCG

4401 ACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCCG

4501 CTTACCGGATACCTGTCCGCTTCTCCCTTCCGGGAGCGTGGCGTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTCCGCTCCA

4601 AGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCC

4701 ACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGA

4801 ACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACCACCGCTGGTAGCGGTGTT

4901 TTTTTGTTCGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACTC

5001 ACGTTAAGGGATTTTGGTATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTTGT

5101 GTGAATCGTAACTAACATACGCTCTCCATCAAACAAACGAAACAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATT

5201 TCTCTATCGAA