



PvuI (7) SgfI (6) MfeI (82) EcoNI (96)  
1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245) Bsu36I (291)  
Psp1406I (203) PvuII (239) EcoNI (287)  
201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACCGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCGGTTGAGTGCAGTCTGCCGCTCCCGCCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMI (441) NaeI (441)  
401 GGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGTCTTTGTTTCGTTT

NcoI (560) BstEII (555) AgeI (552)  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATGGCTGTCAGCGACGCGCTGCTCCCATCTTTCTCCAC  
601 GTTCGCGTCTGGCCCGCGGGAAGGAGAAGACTGCGTCAAGCAGGTGCCCGAATAACCGCTGGCGGGAGGAGCTCTCCACATGAAGCGACTTCCC  
13▶ F A S G P A G R E K T L R Q A G A P N N R W R E E L S H M K R L P  
701 CCAGTGTCTCCCGCGCCCTATGACCTGGCGGCGGCGACCGTGGCCACAGACTGGAGAGCGGCGGAGCGCGTGGCGTTGCGGGGTAGCAACCTGG  
47▶ P V L P G R P Y D L A A A T V A T D L E S G G A G A A C G G S N L  
801 CGCCCTACCTCGAGAGAGACCAGGAGTTCAACGATCTCCTGGACCTGGACTTTATTCTCTCCAATTGCTGACCCATCCTCCGGAGTCAGTGCCCGC  
80▶ A P L P R R E T E E F N D L L D L D F I L S N S L T H P P E S V A A  
901 CACCGTGTCTGTCAGCGTCAGCTCTCTTCTGTCGCGCTGAGCAGCGCCCTGCCAGCGCCCTCCACCTGCAGCTTACCTATCCGATCCGG  
113▶ T V S S S A S A S S S S S P S S S G P A S A P S T C S F T Y P I R

XmaI (1011) SmaI (1011) StuI (1041) Eco147I (1041)  
1001 GCCGGAAACGACCCGGGCGTGGCGCCGGCGGCGACGGCGGAGGCTCTCTATGGCAGGGAGTCCGCTCCCCCTCCGACGGCTCCCTTCAACCTGGCGG  
147▶ A G N D P G V A P G G T G G G L L Y G R E S A P P P T A P F N L A

BsrGI (1166)  
1101 ACATCAACGACGTGAGCCCTCGGGCGGCTTCTGTCGCGGAGCTCTGCGGCCAGAAATGGACCCGGTACATTCGCCCGCAGCAGCCGCGAGCCGCCAGG  
180▶ D I N D V S P S G G F V A E L L R P E L D P V Y I P P Q Q P Q P P G  
1201 TGGCGGGCTGATGGCAAGTTCTGTCGTAAGGCGTCTGAGCGCCCTGGCAGCGAGTACGGCAGCCGTCGGTTCATCAGCGTCAGCAAAGGCGCCCT  
213▶ G G L M G K F V L K A S L S A P G S E Y G S P S V I S V S K G S P

DraIII (1309) BbrPI (1350) XcmI (1392) ApaLI (1386)  
1301 GACGGCAGCCACCCGGTGGTGGCGCCCTAACCGCGGGCCGCCGCGCAGCTGCCCAAGATCAAGCAGGAGGCGGTCTCTTCGTGACCCACTTGG  
247▶ D G S H P V V V A P Y N G G P P R T C P K I K Q E A V S S C T H L

NgoMI (1431) NaeI (1431)  
1401 GCGCTGGACCCCTCTCAGCAATGGCCACCGCGGCTGCACACGACTTCCCTGGGGCGGAGCTCCCGCAGGACTACCCGACCTGGGTCTTGA  
280▶ G A G P P L S N G H R P A A H D F P L G R Q L P S R T T P T L G L E

Bsp120I (1567) EcoO109I (1566) XmaI (1563) SmaI (1563)  
1501 GGAAGTGTGAGCAGCAGGACTGTCACCTGCCCTGCCGCTTCTCCCGGCTCCATCCCCACCCGGGGCCCAATTACCCATCCTTCTGCCGATCAG  
313▶ E V L S S R D C H P A L P L P P G F H P H P G P N Y P S F L P D Q

SphI (1655) PvuI (1690)  
1601 ATGCAGCCGCAAGTCCCGCGCTCCATTACCAAGACTCATGCCACCCGGTCTCTGCATGCCAGAGGAGCCCAAGCCAAAGAGGGGAAGCAGTCTGTCG  
347▶ M Q P Q V P P L H Y Q E L M P P G S C M P E E P K P K R G R R S W  
1701 CCCGAAAAGGACCCACCACTTGTGATTACCGGGCTGCGGCAAACTACACAAAGAGTTCCATCTCAAGGACACCTGCGAACCCACACAGG  
380▶ P R K R T A T H T C D Y A G C G K T Y T K S S H L K A H L R T H T G  
1801 TGAGAAACCTTACCACTGTGACTGGGAGCGTGTGGATGGAAATTCGCCCGCTCAGATGAAGTACCAGGCACTACCTGAAACACACGGGCGCCCGC  
413▶ E K P Y H C D W D G C G W K F A R S D E L T R H Y R K H T G H R P

BspEI (1975)  
1901 TTCCAGTGC AAAAATGCGACCGAGCATTTTCCAGGTCGGACCCTCGCCTTACACATGAAGAGGCATTTT GAGTCCGGAGGGGGTCTCTCTGGCA  
447▶ F Q C Q K C D R A F S R S D H L A L H M K R H F E S G G G G S P G

BamHI (2032) Tth111I (2000) SacII (2028) NheI (2041)  
2001 GACGTCGTCGCCGAGACGAGCGCGCGCGGATCTGAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAAT  
480▶ R R R R R R R R R R R R G S •

HpaI (2179) MfeI (2190)  
2101 GCAGTGA AAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACATTGCATT

EcoRI (2275)  
2201 CATTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAA

2301 ACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGCTGTTGCCAATGTGCATTAGCTGTTTGCA

2401 GCCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTTCCAAGTTTGAAGTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACAT

2501 TCCCTTTTATAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCAT

2601 AATATCCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTAC

2701 TTGAGGGGGATGAGTTCCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGC  
135 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 I L E R C

2801 ACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCT  
102 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 K Q G N S

2901 CACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGT  
69 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 T K T R

3001 ATGCCC GCCCGACATGGTGTCTTGTCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGG  
35 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 S I N F T

3101 TCTTCATGATGGCCCTCCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTT  
2 K M

3200 ATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTG

3300 GAAAGTCCC GTTGATTTACTAGTCAAAAACAACTCCATTGACGTCAATGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCAT

3399 TGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATA

3499 ATGCCAGGCGGGCCATTTACCGTATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGATTTACCGTAAATAC

3599 TCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGAAACATACGTCAATATTGACGTCAATGGGCGGGGCTGTTGGCGGTCAGCCAG

3699 GCGGGCCATTTACCGTAAGTTATGTAACGCCTG 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 A A A G G C C A G G A A C C G T A A A A G G C C G C T

3797 TGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCA

3897 GCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTT

3997 TCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGTTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGCCGACCGCTGCGCCT

4097 TATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGG

4197 CGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAA

4297 AGAGTTGGTAGCTCTTGTCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAG

4397 AAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTTCATGGCTAGTTAATTAACATTTAAATC AGC

4497 GGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAAACAAAACGAAACAAA

4597 ACAAACCTAGCAAATAGGCTGTCCCGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA