



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

HindIII (245)
Psp1406I (203) 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACCGCGCCGCCCTACCTGAGGCC
PvuII (239)
Bsu36I (291)
301 GCCATCCACGCCGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCCTCTGAATGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

AgeI (552) **BspHI (560)**
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCCATCATGAGTTCTCGCTGTTAATGTAAAAAGCTGAAGGT
1 M S S S P V N V K K L K V

PshAI (640)
601 GTCGGAGCTGAAAGAGGAGCTCAAGAAGCGACGCCCTTCTGACAAAGGCTCAAGGCCGAGCTCATGGAGCGACTCCAGGCTGCGCTGGACGACGAGGAG
13 S E L K E E L K K R R L S D K G L K A E L M E R L Q A A L D D E E

XhoI (781)
NcoI (716) 701 GCCGGGGCCCGCCCGCATGGAGCCCGGAACGGCAGCCTAGACCTGGCGGGGATTCCGCTGGGCGCTCGGGAGCAGGCCCTCGAGCAGGAGGCCGCGG
47 A G G R P A M E P G N G S L D L G G D S A G R S G A G L E Q E A A

StuI (777)
BsrBI (851)
801 CCGCGCGCATGAAGAGGAGGAAGAAGAGGAAGAGGAGGAGGAAGAACTCCGCTCTGGACGGCGACCAGATGGAGCTAGGAGAGGAGAACGGGGCCG
80 A G G D E E E E E E E E E E E E G I S A L D G D Q M E L G E E N G A A

Bsp120I (915)
901 GGGGGCGGCCACTCGGGCCCGATGGAGGAGGAGGCCCTCGGAAGACGAGAACGGCGACGATCAGGGTTTCCAGGAAGGGGAAGATGAGCTCGGG
113 G A A D S G P M E E E E A A S E D E N G D D Q G F Q E G E D E L G

SgrAI (1055)
1001 GACGAAGAGGAAGGCGCGGGCGACGAGAACGGGCACGGGGAGCAGCAGCCTCAACCGCGGCGACGACGAGCAACAGCCCCAACAGCAGCGCGGGGGCCG
147 D E E E G A G D E N G H G E Q Q P Q P P A T Q Q Q Q P Q Q Q R G A
1101 CCAAGGAGGCCCGGGGAAGAGCAGCGGCCACCTCGCTGTTGCGGGTACGGTGGCGCCCGCCGGGCGAGGCAGGGCCAGCAGCAGGCGGGAGGGGA
180 A K E A A G K S S G P T S L F A V T V A P P G A R Q G Q Q Q A G G D
1201 CGGCAAAACAGAACAGAAAGGCGGAGATAAAAAGAGGGGTGTTAAAAGACCACGAGAAGATCATGGCCGTGGATATTTTGTGATACATTGAAGAACAAAG
213 G K T E Q K G G D K K R G V K R P R E D H G R G Y F E I E E N K
1301 TATAGCAGCCAAATCTCCTCAGCCACCTGTTGAAGAAGAAGATGAACACTTCGATGACACAGTGGTTGTCTTGATACTTATAITGTGATCTACAT
247 Y S R A K S P Q P P V E E E D E H F D D T V V C L D T Y N C D L H
1401 TTTAAATATCAAGAGATCGTCTCAGTGTCTTCCCTTACAATGGAGAGTTTTGCTTTTCTTTGGGCTGGAGGAAGAGCATCTATGGTGTGTCAAAAGG
280 F K I S R D R L S A S S L T M E S F A F L W A G G R A S Y G V S K G
1501 CAAAGTGTGTTTTGAGATGAAGTTACAGAGAAGATCCAGTAAGGCATTTATATACAAAAGATATTGACATACATGAAGTTCGTATTGGCTGGTCACTA
313 K V C F E M K V T E K I P V R H L Y T K D I D I H E V R I G W S L
1601 ACTACAAGTGAATGTTACTTGGTGAAGAAGAAATTTCTTATGGGTATTCTCTAAAAGGAATAAAAACATGCAACTGTGAGACTGAAGATTATGGAGAAA
347 T T S G M L L G E E E F S Y G Y S L K G I K T C N C E T E D Y G E

BspLU11I (1723) 1701 AGTTTGATGAAAATGATGTGATTACATGTTTGGCTAACTTTGAAAGTATGAAGTAGAACTCTCGTATGCTAAGAATGGACAAGATCTTGGCGTTGCCTT
380 K F D E N D V I T C F A N F E S D E V E L S Y A K N G Q D L G V A F
1801 CAAAATCAGTAAGGAAGTTCTTGTGGACGGCCACTGTTCCCGCATGTTCTCTGCCACAACCTGTCAGTTGAATTTAATTTTGGTCAGAAGGAAAAGCCA
413 K I S K E V L A G R P L F P H V L C H N C A V E F N F G Q K E K P

BglIII (1782)
Bst1107I (1920) 1901 TATTTTCCAATACCTGAAGAGTATACTTTCATCCAGAACCTCCCTTAGAGGATCGAGTTAGAGGACCAAGGGCCGTAAGAGAAGAAAGATTGTGAAG
447 Y F P I P E E Y T F I Q N V P L E D R V R G P K G P E E K K D C E
2001 TTGTGATGATGATTGGCTTGCCAGGAGCTGAAAAACTACCTGGGTTACTAAACATGCAGCAGAAAATCCAGGAAAATATAACATTCTTGGCACAAAATAC
480 V V M M I G L P G A G K T T W V T K H A A E N P G K Y N I L G T N T

EcoO109I (1971)
2101 TATTATGGATAAGATGATGGTGGCAGGTTTTAAGAAGCAAATGGCAGATACTGAAAACTGAACACACTGTTGCAGAGAGCCCCCAGTGTCTTGGGAAA
513 I M D K M M V A G F K K Q M A D T G K L N T L L Q R A P Q C L G K
2201 TTTATTGAGATTGCTGCCGAAAAGAAGCGAAATTTTATTCTGGATCAGACAAATGTGTCTGCTGCTGCCAGAGGAGAAAAATGTGCCTGTTTGCAGGCT
547 F I E I A A R K K R N F I L D Q T N V S A A A Q R R K M C L F A G
2301 TCCAGCGAAAAGCTGTTGTAGTTTGCCAAAAGATGAAGACTATAAGCAAAGAACACAGAAGAAAGCAGAAGTAGAGGGGAAAGACCTACCAGAACATGC
580 F Q R K A V V V C P K D E D Y K Q R T Q K K A E V E G K D L P E H A
2401 GGTCTCAAAATGAAAGGAACTTTACCTCCAGAGGTAGCTGAGTGCTTTGATGAAATAACCTATGTTGAACTTCAAGAAGGAAAGCCAAAAAATC
613 V L K M K G N F T L P E V A E C F D E I T Y V E L Q K E E A Q K L
2501 TTGAGCAATATAAGGAAGAAAGCAAAGGCTCTCCACCAGAAAAGAAACAGAACACTGGCTCAAAGAAAAGCAATAAAAAAAGAGTGGCAAGAAC
647 L E Q Y K E E S K K A L P P E K K Q N T G S K K S N K N K S G K N

XcmI (2619)

MscI (2616)

2601 AGTTTAAACAGAGGTGGTGGCCATAGAGGACGTGGAGGATTCAATATGCGTGGTGGAAATTCAGAGGAGGAGCCCCTGGGAATCGTGGCGGATATAATAG
680▶ Q F N R G G G H R G R G G F N M R G G N F R G G A P G N R G G Y N R
2701 GAGGGCAACATGCCACAGAGAGGTGGTGGCGGTGGAGGAAGTGGTGAATCGGCTATCCATACCCTCGTGGCCCTGTTTTCTGGCCGTGGTAGTTAC
713▶ R G N M P Q R G G G G G S G G I G Y P Y P R A P V F P G R G S Y
2801 TCAAACAGAGGAACTACAACAGAGGTGAATGCCCAACAGAGGAACTACAACCAGAACTTCAGAGGACGAGGAAACAATCGTGGCTACAAAAATCAAT
747▶ S N R G N Y N R G G M P N R G N Y N Q N F R G R G N N R G Y K N Q

NcoI (2947)

2901 CTCAGGGCTACAACCAGTGGCAGCAGGGTCAATTCTGGGGTCAAGCCATGGAGTCAGCATTATCACCAAGGATATTATTGAATACCCAAATAAACGA
780▶ S Q G Y N Q W Q Q G Q F W G Q K P W S Q H Y H Q G Y Y •

MscI (3011)

NheI (3005)

3001 ACTGAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTAATTTGTGA

HpaI (3143) MfeI (3154)

3101 TGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACATTGCATTCTTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAG

EcoRI (3239)

3201 GTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAATACAGCATAGCAAACCTTAACCTCAAATCAAGCCTCTACTTGAATCCT

3301 TTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGT

SspI (3478) SmaI (3492)

3401 ATTTTCCAAGGTTTGAAGTCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTATAGTAAAATATTCAGAAAATAATTTAAAT

EcoO109I (3553)

3501 ACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACA

3601 AAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACC
141◀ • N R T Y K L P I L E E I T T K V

BstXI (3782)

3701 AGCTTGCCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGT

123◀ 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 S R D
3801 CCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAAATGGCAATGGCTTCAGCACAGAC
90◀ 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 C V

StuI (3917)

3901 AGTGACCCTGCCAATGTAGGCCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTTCCTCATAG
57◀ 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 E Y

XmnI (4059)

4001 AGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTAT
23◀ 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 (4125)

4101 ACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGA

SpeI (4280)

4201 CCTCCCACGTACAGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCC

4301 ATTGACGTCAATGGGGTGGAGACTTGAAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGTAATAGCG

SnaBI (4408)

4401 ATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAAGTCAATGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAG

NdeI (4513)

4501 GGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGAGTTTACCGTAAATACTCCACCATTGACGTCAATGAAAGTCCCTATTGGCGT

PacI (4699)

PstI (4692)

SdaI (4691)

4601 TACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCCTGCAGGT

BspLU11I (4709)

4701 TAAATTAAGAACATGTGAGCAAAAGGCCAGAAACCGTAAAAAGGCCGCTTGGTGGCGTTTTTCATAGGCTCCGCCCCCTGACGAGCA

4801 TCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCTGTT

4901 CCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGGTGAGG

ApaI (5023)

5001 TCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACA
5101 CGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGC
5201 TACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG
5301 GTAGCGGTGGTTTTTTTGGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTG

PacI (5439) SwaI (5448) NotI (5458)

5401 GAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTG
5501 TTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGG
5601 TGCCAGAACATTTCTCTATCGAA