



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGTGTCGTGTAAGTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCGCCCTACCTGAGGCC
HindIII (245)
Bsu36I (291)
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
NaeI (441) 401 GGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

KasI (535) 501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCCTACCTGAGATCACCGGTACCATTGGCTCTCTTCACTCCGTGGAAGTTGTCTCTCAGAA
AgeI (552) 1▶ M A L F T P W K L S S Q K
NcoI (560)
DraIII (573)
601 GCTGGGCTTTTTCTGGTGCATTTTGGCTTCAATTTGGGTATGATGCTTTCGCACTTTACCATCCAGCAGCGAACTCAGCTGAAAGCAGCTCCATGCTG
13▶ L G F F L V T F G F I W G M M L L H F T I Q Q R T Q P E S S S M L
701 CGCGAGCAGATCTGGACCTCAGCAAAAGGTACATCAAGGCACCTGGCAGAAAGAAAACGGAATGTGGTGGATGGGCCATACGCTGGAGTCATGACAGCTT
47▶ R E Q I L D L S K R Y I K A L A E E N R N V V D G P Y A G V M T A
801 ATGATCTGAAGAAAACCTTGTGTGTTATTAGATAACATTTTGCAGCGCATTGGCAAGTTGGAGTCAAGGTGGACAATCTTGTGTCAATGGCACC GG
80▶ Y D L K K T L A V L L D N I L Q R I G K L E S K V D N L V V N G T G

EcoRV (967)
901 AACAACTCAACCACTCACTACAGCTGTTCCAGCTTGGTTGCACTTGAGAAAATTAATGTGGCAGATATCATTAACCGGAGCTCAAGAAAAATGTGTA
113▶ T N S T N S T T A V P S L V A L E K I N V A D I I N G A Q E K C V
1001 TTGCTCCTATGGACGGCTACCCCTACTGTGAGGAAAAGATCAAGTGGATGAAAGACATGTGGCGTTCAGATCCCTGTACGCAGACTATGGAGTGGATG
147▶ L P P M D G Y P H C E G K I K W M K D M W R S D P C Y A D Y G V D
1301 GATCCACCTGCTCTTTTTTATTTACCTCAGTGAAGTTGAAAATTTGGTGCCTCATTACCTTGGAGAGCAAAAAATCCCTACGAAGAAGCTGATCATAA
180▶ G S T C S F F I Y L S E V E N W C P H L L P W R A K N P Y E E A D H N

EcoRI (1261)
1201 TTCATTGGCGAAATTCGTACAGATTTAATATTCTCTACAGTATGATGAAAAAGCATGAAGAATTCGGTGGATGAGACTACGGATCCGGCGAATGGCT
213▶ S L A E I R T D F N I L Y S M M K K H E E F R W M R L R I R R M A
1301 GACGCATGGATCCAAGCAATCAAGTCCCTGGCAGAAAAGCAGAACTTGAAGAAGAAAGCGGAAGAAAGTCCCTGTTTACCTGGGACTCCTGACCAAGG
247▶ D A W I Q A I K S L A E K Q N L E K R K R K K V L V H L G L L T K

XmnI (1450)
1401 AATCTGGATTTAAGATTGCAGAGACAGCTTTCAGTGGTGGCCCTCTGGTGAATTAGTTCAATGGAGTGATTTAATTACATCTCTGTACTTACTGGCCA
280▶ E S G F K I A E T A F S G G P L G E L V Q W S D L I T S L Y L L G H
1501 TGACATTAGATTTCAGCTTCACTGGCTGAGCTCAAGGAAATCATGAAGAAGTTGTAGGAAACCGATCTGGCTGCCAACTGTAGGAGACAGAAATGTT
313▶ D I R I S A S L A E L K E I M K K V V G N R S G C P T V G D R I V

SphI (1676)
1601 GAGCTCATTTACATTGATATTGTAGGACTTGTCAATTCAGAAAACCTTTGGACCATCTGGGTTTATTACCAGTGCATGCTCCGAGTCCCTGATTCAT
347▶ E L I Y I D I V G L A Q F K K T L G P S W V H Y Q C M L R V L D S
1701 TTGGTACTGAACCGAATTTAATCATGCAAAATATGCCAATCGAAAGGCCACAGACCCCTTGGGAAAATGGAATCTGAACCTCAGCAGTTTTATAC
380▶ F G T E P E F N H A N Y A Q S K G H K T P W G K W N L N P Q Q F Y T

EcoRV (1867)
1801 CATGTTCCCTCATACCCAGACAACAGCTTTCTGGGTTTGTGGTTGAGCAGCACCTGAACTCCAGTGCATCCACCACATTAATGAAATCAAAGGCAG
413▶ M F P H T P D N S F L G F V V E Q H L N S S D I H H I N E I K R Q

BglIII (1949) 1901 AACAGTCCCTTGTGTATGGCAAAGTGGATAGCTTCTGGAAGAATAAGAAGATCTACTGGACATTATTCACACATACATGGAAGTGCATGCAACTGTTT
447▶ N Q S L V Y G K V D S F W K N K K I Y L D I I H T Y M E V H A T V
SphI (1986)
2001 ATGGCTCCAGCACAAAGAATATCCAGTTACGTGAAAACCATGGTATCCTCAGTGGACGGGACCTGCAGTTCCTTCTCGAGAAAACCAAGTTGTTTGT
480▶ Y G S S T K N I P S Y V K N H G I L S G R D L Q F L L R E T K L F V

Bsp120I (2121) 2101 TGGACTTGGGTTCCCTTACGAGGGCCAGCTCCCTGGAAGCTATCGCAAATGGATGTGCTTTTCTGAATCCCAAGTTCAACCCACCCAAAAGCAGCAAA
513▶ G L G F P Y E G P A P L E A I A N G C A F L N P K F N P P K S S K

BstXI (2281) 2201 AACACAGACTTTTTTATTGGCAAGCAACTCTGAGAGAGCTGACATCCAGCATCCTTACGCTGAAGTTTTCATCGGGCGGCCACATGTTGGACTGTTG
547▶ N T D F F I G K P T L R E L T S Q H P Y A E V F I G R P H V W T V

NdeI (2372) 2301 ACCTCAACAATCAGGAGGAAGTAGAGGATGAGTGAAGCAATTTTAAATCAGAAGATTGAGCCATACATGCCATATGAATTTACGTGCGAGGGGATGCT
580▶ D L N N Q E E V E D A V K A I L N Q K I E P Y M P Y E F T C E G M L

NcoI (2439) 2401 ACAGAGAATCAATGCTTTTATTGAAAAACAGGACTTCTGCCATGGGCAAGTGTGTTGCCACCCCTCAGCGCCCTACAGGTCAAGCTTGTGAGCCCGG
613▶ Q R I N A F I E K Q D F C H G Q V M W P P L S A L Q V K L A E P G
2501 CAGTCTGCAAGCAGGTGTGCCAGGAGGCCAGCTCATCTGCGAGCCTTCTTTCTCCAGCACCTCAACAAGGACAAGGACATGCTGAAGTACAAGGTGA
647▶ Q S C K Q V C Q E S Q L I C E P S F F Q H L N K D K D M L K Y K V

XcmI (2621) 2601 CCTGCCAAAGCTCAGAGCTGGCCAAGGACATCTGGTGCCTCCTTTGACCCTAAGAATAAGCACTGTGTGTTTCAAGGTGACCTCCTGCTTTCAGCTG
680▶ T C Q S S E L A K D I L V P S F D P K N K H C V F Q G D L L L F S C
DraIII (2662)
HindIII (2482)
SrfI (2493)
XmaI (2494)

KasI (2704)
2701 TGCAGGCGCCACCCAGGCCAGGAGGGTCTGCCCTGCCGGACTTCATCAAGGGCCAGGTGGCTCTCTGCAAAGACTGCCTATAGCAGCTACCTGCT
713▶ A G A H P R H Q R V C P C R D F I K G Q V A L C K D C L •

NheI (2812)
2801 CAGCCCTGCACCCTAGCTAGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAA

HpaI (2950) MfeI (2961)
2901 TTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACCAAGTTAAACAACAACATTGCATTCTTTTATGTTTCAGGTTCAAGGGGAGGT

EcoRI (3046)
3001 GTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTT
3101 GAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGAT

Swal (3299)
3201 ATAGTGTATTTTCCCAAGTTTGAAGTACTGCTCTTCATTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTATGTAATAATTCAGAAATAA
3301 TTTAAATACATCATTGCAATGAAATAAATGTTTTTATTAGCGAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTA
3401 GGGAAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGT
141◀ • N R T Y K L P I L E E I T T

BstXI (3589)
3501 TTTGACCAGCTTGCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCTGATG
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
3601 GATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCAGCAGACCCCAATGGCAATGGCTTCAG
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 (3724)
3701 CACAGACAGTGACCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTGTG
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

BbsI (3870)
3801 CTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGATGGCCCTCTATAGTGAGTC
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
XmnI (3866)

3901 GTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTT

SpeI (4087)
4001 ATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAAC
4100 AAACCTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGG

SnaBI (4215)
4200 TAATAGCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCCATAGGTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCAATTGAC

NdeI (4320)
4300 GTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTAAGTGGCAGTTTACCGTAAATACTCCACCATTGACGTCAATGAAAGTCCCT

SdaI (4498)
4400 ATTGGCGTACTATGGAAACATACGTCATTATTGACGTCAATGGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGC

PstI (4499) PacI (4506)
4500 CTG 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 C A A A A G G C C A G G A A C C G T A A A A A G G C C G G T T G C T G G C G T T T T T C C A T A G G C T C C G C C C C
4598 CTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCTGGAAGCTCCCTCGTGCG
4698 CTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCCGGAAGCGTGGCGTTTCTCATAGCTCACGCTGTAGGTATCTCAGT

ApaLI (4830)
4798 TCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTGAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACC
4898 CGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCC
4998 TAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAA
5098 ACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTG

EagI (5266)
5198 ACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTCATTA
PacI (5246) Swal (5255) NotI (5265)
5298 CATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTG
5398 CAAGTGCAGTGCCAGAACATTTCTCTATCGAA