



PvuI (7)  
SgfI (6)  
1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGCGCAATTGAACGGTGCCTA  
101 GAGAAGGTGGCGGGGTAAGTGGAAAGTGATGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC

HindIII (245)  
Psp1406I (203) PvuII (239) Bsu36I (291)  
201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGGTGCGATCTCTCTTACAGCGCCCGCCGCTACCTGAGGCC  
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCCTGTGGTGCCTCTGAAGTCGCTCCGCGCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

KasI (535) AgeI (552)  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTTCATCATGAACCACCAACAGCAGCAGCAGCAGCAGCAAAGC  
1 M N H Q Q Q Q Q Q Q Q K A

PvuII (610)  
601 GGGCAGCAGCAGCTGAGCGAGCCCGAGGACATGGAGATGGAAGCTGGAGATACAGATGACCCACCAAGAATTACTCAGAACCCTGTCATCAATGGGAAT  
13 G E Q Q L S E P E D M E M E A G D T D D P P R I T Q N P V I N G N

Eco47III (791)  
701 GTAACCCTGAGTGATGGGCACAGCAACGCAGAAGAGGACATGGAGGATGACACCAGTTGGCGCTCCGAGGCAACCTTTTCACTGTCGAGCGCTTCA  
47 V T L S D G H S N A E E D M E D D T S W R S E A T F Q F T V E R F  
801 GCAGACTGAGTGAGTCGGTCTTAGCCCTCCGTTTGTGCGAAATCTGCCATGGAAGATTATGGTATGCCACGCTTTTATCCAGACAGACCACCA  
80 S R L S E S V L S P P C F V R N L P W K I M V M P R F Y P D R P H Q

DraIII (945)  
901 AAAAAAGCTAGGATTCTTCTCCAGTGCAATGTGAATCTGACTCCACGCTGCTGCTGCCACGCACAAGCAGTGTGAAGATAATAAATTACAGGGAT  
113 K S V G F F L Q C N A E S D S T S W S C H A Q A V L K I I N Y R D  
1001 GATGACAAATCATTACAGCCGGCAATCAGCCACTTGTTCTTCCACAAGAAAAGACTGGGGATTCTTAATTTTCATGGCTGGAGTGAAGTGACTGATC  
147 D D K S F S R R I S H L F F H K E N D W G F S N F M A W S E V T D

Tth111I (1123)  
XmnI (1104) ClaI (1114)  
1101 CTGAGAAAGGATTATCGATGATGACAAAGTACCTTTTGAAGTTTTGTACAGGCAGATGCTCCCATGGAGTTGCGTGGGACTCAAAGAAGCATACTGG  
180 P E K G F I D D D K V T F E V F V Q A D A P H G V A W D S K K H T G  
1201 CTATGTCGGTTAAAGAATCAGGAGCAACGTTACATGAATAGCTTGTGTCAGACCTGTTTTCACAAATCAGTTACGAAAGGCTGTGTACATGATG  
213 Y V G L K N Q G A T C Y M N S L L Q T L F F T N Q L R K A V Y M M  
1301 CCAACAGAGGGCGATGATTCATATAAAGCGTCCCTTTAGCATTACAAAGAGTGTCTATGAATTACAGCACAGCGATAAGCCCGTAGGAACAAAAAGC  
247 P T E G D D S S K S V P L A L Q R V F Y E L Q H S D K P V G T K K  
1401 TCACAAAGTCATTTGGGTGGGAACTTTAGATAGCTTCAACACGATGTTCAAGAGCTTTGTCGAGTGTGCTTGTATAATGTGAAAAAATGAAGTAA  
280 L T K S F G W E T L D S F M Q H D V A G E L C R V L L D N V E N K M K  
1501 AGGCACCTGTGTAGAGGCACATTCCAAGTTATCCGAGGCAAAATGGTGTCTATATTAGTGCAAAAGCAGTAGACTACCGTTCTGATAGGAGAGAA  
313 G T C V E G T I P K L F R G K M V S Y I Q C K D V D Y R S D R R E

EcoRV (1609)  
1601 GATTATTATGATATCCAGCTAAGTATAAAAAGGAAAGAAAACATATTTGAATCATTTGTGGATTATGTGGCAGTAGAACAACACTAGACGGTGACAATAAT  
347 D Y Y D I Q L S I K G K K N I F E S F V D Y V A V E Q L D G D N K

SdaI (1719)  
BstAPI (1713)  
1701 ACGATGCTGGGAGCATGGCCTGCAGGAGGAGAGAAAGGTTGAAATTCCTAACATTGCCCCAGTGTACACCTGCAACTTATGAGGTTTCATGTATGA  
380 Y D A G E H G L Q E A E K G V K F L T L P P V L H L Q L M R F M Y D

Bsu36I (1888)  
1801 CCCTCAGACGGACAAAAACATCAAGATCAATGACAGGTTTGTGTTCCAGAACAGTTACCACTTGATGAATTTTTGCAAAAAACAGACCCTAAGGACCCT  
413 P Q T D Q N I K I N D R F E F P E Q L P L D E F L Q K T D P K D P  
1901 GCAAAATTATTTCCATCGAGTCTGGTTCCAGTGGAGATAATCATGGTGGACATTCAGTGGTTTACCTAAACCCAAAGGGGATGGCAAAATGGGTGA  
447 A N Y I L H A V L V H S L D N H G G H Y V V Y L N P K G D G K W C  
2001 AGTTCGATGATGACGTGGTATCCAGGTGTAATAAAGAAGCCATTGAGCACAAATTATGGGGTTCATGATGATGATGATGATGATGATGATGATGATGATGAT  
480 K F D D D V V S R C T K E E A I E H N Y G G H D D D L S V R H C T N  
2101 TGCCTATATGTTAGTGTACATCAGGGAATCAAAGCTAAGTGAAGTGTACAAGCTGTACCCGACCATGATATTCTCAGCAGTGGTGGAAACGATTGCAA  
513 A Y M L V Y I R E S K L S E V L Q A V T D H D I P Q Q L V E R L Q

BsrBI (2233) BstXI (2283) MscI (2292)  
2201 GAAGAGAAAAGGATCGAGGCTCAGAAGCGGAAGGAGCGGCAGGAAAGCCATCTCTACATGCAAGTGCAGATAGTTGCAGAGGACCAAGTTTGTGGCCACC  
547 E E K R I E A Q K R K E R Q E A H L Y M Q V Q I V A E D Q F C G H  
2301 AAGGAAATGACATGTACGACGAGGAGAAAGTGAAGTACTGTGTCAAGGTTCTGAAGAACTCTCCCTGGCCGAGTTTGTTCAGAGCCTCTCCAGAC  
580 Q G N D M Y D E E K V R Y T V F K V L K N S S L A E F V Q S L S Q T  
2401 CATGGGTTTCCACAAGATCAGATTTCGATTGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT  
613 M G F P Q D Q I R L W P M Q A R S N G T K R P A M L D N E A D G N

SspI (2541)  
2501 AAAACAATGATTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT  
647 K T M I E L S D N E N P W T I F L E T V D P E L A A S G A T L P K  
2601 TTGATAAAGATCATGATGTAATGTTATTTTTGAAGATGTATGATCCTAAACCCGAGGCTTGAATTACTGTGGACATATCTACACACCATATCTGTAA  
680 F D K D H D V M L F L K M Y D P K T R S L N Y C G H I Y T P I S C K  
2701 AATACGTGACTTGTCCCGTATGTGTGACAGAGCAGGATTTATCCAGGATACTAGCCTTATCTCTATGAGGAAGTTAAGCCGAATTAACAGAGAGA  
713 I R D L L P V M C D R A G F I Q D T S L I L Y E E V K P N L T E R  
2801 ATTCAGGACTATGATGTCTCTTGATAAAGCCCTTGTGAGCTGATGGATGGTGACATCATAGTGTTCAGAAGGATGACCCAGAAAACGATAACAGTG  
747 I Q D Y D V S L D K A L D E L M D G D I I V F Q K D D P E N D N S

BglII (2928) SalI (2944)  
2901 AGCTGCCCACTGCTAAAGAGATTTTCAGAGATCTTACCACCGGGTGCAGCTCATCTTCTGTGACAAAACGATCCCCAATGACCCCTGGATTTGTGGTCAC  
780▶ E L P T A K E Y F R D L Y H R V D V I F C D K T I P N D P G F V V T  
3001 ATTATCAAATAGGATGAACATTTTTTCAGGTTGCAAAAGACAGTTGCACAGAGGCTCAACACAGACCCCATGCTGCTGACGTTTTTCAAGTCTCAAGGTTAT  
813▶ L S N R M N Y F Q V A K T V A Q R L N T D P M L L Q F F K S Q G Y

BglII (3150)  
3101 AGGGACGGCCCCGGTAATCCTCTTAGGCATAATTACGAAGGTAATTAAGAGATCTTCTACAATTTCAAGCCAAGACAACCTAAGAAACTTTACTATC  
847▶ R D G P G N P L R H N Y E G T L R D L L Q F F K P R Q P K K L Y Y  
3201 AGCAAATTAAGATGAAAATCACAGACTTTGAGAACAGGAGAAGTTTTAAGTGTATATGGCTAAACAGCCAATTTAGGGAAGAGGAAATAACACTATACCC  
880▶ Q Q L K M K I T D F E N R R S F K C I W L N S Q F R E E E I T L Y P  
3301 AGACAAGCATGGCTGTGTCCGGACCTGTTAGAAGAATGTAAGGAGCTGGAGCTCGGTGACAAAGCATCAGGGAGACTTAGGCTGCTAGAAATTTGTG  
913▶ D K H G C V R D L L E E C K K A V E L G D K A S G R L R L L E I V  
3401 AGCTACAAAATTGGTGTGCATCAAGAAGATGAACATTTAGAAATCTTATCTGCAACGAGTCCGACTTTTAGAATAGAGGAAATACCTTTGGACC  
947▶ S Y K I I G V H Q E D E L L E C L S P A T S R T F R I E E I P L D  
3501 AGGTAGACATAGATAAGGAAAACGAGATGCTAATCACAGTGGCACATTTCCACAAAGAGGTGTTGGGACGTTTGAATTTCCATTTTTGCTAAGGATACA  
980▶ Q V D I D K E N E M L I T V A H F H K E V F G T F G I P F L L R I H  
3601 CCAGGGTGAACATTTCCGAGAGGTGATGAAGCGGATACAGAGCCTGCTGGACATCCAGGAGAAGGAGTTTGAAGGTTTAAATTTGCAATTTGTCATGATG  
1013▶ Q G E H F R E V M K R I Q S L L D I Q E K E F E K F A I V M M  
3701 GGCCGGCACAGTACATCAATGAAGATGAGTATGAAGTAAATTTGAAGGACTTTGAGCCACAACCTGGTAACATGTCTCATCCTCGCCTTGGCTAGGGC  
1047▶ G R H Q Y I N E D E Y E V N L K D F E P Q P G N M S H P R P W L G

MscI (3892)  
3801 TTGACCACTTCAACAAAGCCCCAAAGAGGAGTGCCTACACTTACCTAGAAAAGGCAATTAATAATCCACAACCTGACTTCCCTCCCGTGC**TAGCTGGCCAGA**  
1080▶ L D H F N K A P K R S R Y T Y L E K A I K I H N •  
3901 CATGATAAGATACATTGATGAGTTTGGACAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTA

HpaI (4024)  
4001 ACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAA  
4101 ACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAAAACCTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAG  
4201 GCATAGGCATCAGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTTCATGGAGTTAAGATATAGTATTTTCCAAGGTTTGAAC

SapI (4302) SspI (4359) SwaI (4373)  
4301 TAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAT  
4401 AAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAA  
4501 TTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGGGGGATGAGTTCCCTCAATGGTGGTTTTGACCAGCTTGCATTTCATCTCAA  
141▶ • N R T Y K L P I L E E I T T K V L K G N M E I

BstXI (4663)  
4601 TGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGG  
117▶ 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 Y P

StuI (4798)  
4701 GTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGAACCTGCCAATGTAG  
84▶ 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 T V R G I Y  
4801 GCCTCAATGTGGACAGCAGATGATCTCCCAAGTCTTGGTCTGATGGCCGCCCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAG  
50▶ 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 K E T

BbsI (4944) XmnI (4940)  
4901 TGGCGACCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCCTATAGTGAGTCTATTATACTATGCCGATATACTATG  
17▶ A V E V L E L D Q Q S I N F T K M

AseI (5006)  
5001 CCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCT

SpeI (5161)  
5101 ACCGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCCATGACGTCAATGGGGTGG

SnaBI (5289)  
5201 AGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGTAGTAATAACGTAGATGT

NdeI (5394)  
5301 ACTGCCAAGTAGGAAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATG  
5401 ATACACTTGATGTACTGCCAAGTGGCAGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAAGTCCCTATTGGCGTTACTATGGGAACATACGTC

SdaI (5572) PacI (5580)  
5501 ATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAA**TAGAACATGTGAGC**  
5601 AAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCA  
5701 AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCTGGAAGTCCCTCGTGCCTCTCTGTTCCGACCTGCCGTTACCG  
5801 GATACCTGTCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCAAGCTGGG

**ApaI (5904)**

5901 CTGTGTGCACGAACCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCA  
6001 GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGCCTAACTACGGCTACACTAGAAGAACAGTAT  
6101 TTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGTTTTTTGT  
6201 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAACTCACGTTAA

**EagI (6340)**

PacI (6320) SwaI (6329) NotI (6339)

6301 GGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTAATC  
6401 GTAAC TAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTAT  
6501 CGAA