



NotI (-1)  
1 GCGGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACATAACGCTCTCCATCAAAACAAAACGAAACA  
PvuI (172)  
SgfI (171)  
101 AAACAACTAGCAAAATAGGCTGTCCAGTGCAAGTGCAGGTGCCAGAACATTTCTATCGAAGGATCTGCGATCGCTCCGGTCCCGTCACTGGGCA  
AgeI (254)  
201 GAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGTCCGCAATTGAACCGGTGCTAGAGAAGGTGGCGGGGTAACTGGAAAGTGATG  
Psp1406I (368)  
301 TCGTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCGGTGACGGTCTTTTTCGCAACGGGTTTCCGCCAG  
HindIII (410)  
401 AACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCGCCATCCACGCGGTTGAGTCCGCTTCTGCCGCT  
501 CCCGCTGTGGTCCCTGAACTGCGTCCGCGTCTAGGTAAGTTAAAGTTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGAGGCTACCTA  
NaeI (606) NcoI (695)  
601 GACTCAGCCGGCTCTCCACGCTTGGCTGACCCTGCTTCAACTCTACGCTTTTGTTCGTTTTCTGTTCTGCGCCGTTACAGATCCAGGCCACCATG  
1 M  
AvrII (720) XmnI (755)  
701 GGAGACCACCTGGACCTTCTCTAGGAGTGGTCTCATGGCCGGTCTGTGTTTGAATTCCTTCTCTCTTGTATGGCCGAATAGCCTTTATCGTT  
2 G D H L D L L L G V V L M A G P V F G I P S C S F D G R I A F Y R  
801 TCTGCAACCTCACCCAGTCCCCAGGTCCTCAACACCAGTGCAGAGGCTCTGCTGAGCTTCAACTATACAGGACAGTCACTGCTTATCCTTCCCCTT  
35 F C N L T Q V P Q V L N T T E R L L L S F N Y I R T V T A S S F P F  
Bst1107I (936) StuI (962)  
901 TCTGGAACAGCTGCAGCTGTGGAGCTCGGGAGCCAGTATACCCCTTGACTATTGACAAGGAGGCCTTCAGAAAACCTGCCAACCTTAGAATCTGGAC  
68 L E Q L Q L L E L G S Q Y T P L T I D K E A F R N L P N L R I L D  
1001 CTGGGAAGTAGTAAGATATACTTCTTGCATCCAGATGCTTTTCAGGACTGTTCCATCTGTTTGAACCTAGACTGTATTTCTGTGGTCTCTCTGATGCTG  
102 L G S S K I Y F L H P D A F Q G L F H L F E L R L Y F C G L S D A  
1101 TATTGAAGATGTTATTTAGAAATTTAAAGCTTTAACTCGTTGGATCTATCCAAAACAGATTCTGAGCCTTACCTTCACTTCTTATTTGGGAA  
135 V L K A D G G Y F R N L K A L T R L D L S K N Q I R S L Y L H P S F G K  
SspI (1239) XhoI (1264)  
1201 GTTGAATTCCTTAAAGTCCATAGATTTTTCTCCAACCAATATCTTGTATGTGAACATGAGCTCGAGCCCTACAAGGGAAAACGCTCTCTTTTTT  
168 L N S L K S I D F S S N Q I F L V C E H E L E P L Q G K T L S F F  
XmnI (1356)  
1301 AGCCTCGCAGTAATAGCTTGTATAGCAGAGTCTCAGTGGACTGGGAAAATGTATGAACCATTCAGAAAACATGGTCTGGAGACTAGATGTTCTG  
202 S L A A N S L Y S R V S V D W G K C M N P F R N M V L E I L D V S  
StuI (1457)  
1401 GAAATGGCTGGCAGTGGACATCACAGGAACTTGTAGCAATGCCATCAGCAAAAGCCAGGCTTCTCTTGTATCTTCCACCACATCATGGTGCCGG  
235 G N G W T V D I T G N F S N A I S K S Q A F S L I L A H H I M G A G  
Ball (1550)  
1501 GTTGGCTTCCATAACATCAAGATCCTGACCAGAACACATTTGCTGGCCTGGCCAGAAGTTCAGTGCAGACCTGGATCTTTCACATGGGTTTGTCTT  
268 F G F H N I K D P D Q N T F A G L A R S S V R H L D L S H G F V F  
1601 TCCCTGAACTCACGAGTCTTGTAGACACTCAAGGATTTGAAGTTTGAACCTTGCCTACAACAAGATAAATAAGATTGCAGATGAAGCATTTTACGGAC  
302 S L N S R V F E T L K D L K V L N L A Y N K I N K I A D E A F Y G  
AsuII (1755)  
1701 TTGACAACCTCCAAGTCTCAATTTGTATATAACCTTCTGGGGAACTTTACAGTTCGAATTTCTATGGACTACCTAAGGTAGCTACATTGATTTGCA  
335 L D N L Q V L N L S Y N L L G E L Y S S N F Y G L P K V A Y I D L Q  
1801 AAAGAATCACATTGCAATAATTCAAGCAAACTCAAAATTCCTGGAAAAATACAGACCTTGGATCTCCGAGACAATGCTTACAACCATTTT  
368 K N H I A I I Q D Q T F K F L E K L Q T L D L R D N A L T T I H F  
EcoRV (1915) SpeI (1938)  
1901 ATCCAAGCATACCCGATATCTTCTGAGTGGCAATAAACTAGTACTTGGCAAAAGTCAACCTTACAGCGAACCTCATCCACTTATCAGAAAACAGGC  
402 I P S I P D I F L S G N K L V T L P K I N L T A N L I H L S E N R  
XbaI (2007) Acc65I (2034)  
2001 TAGAAAATCTAGATATCTCTACTTCTCTACGGTACCTCATCTCCAGATTCTCATTTTAAATCAAATCGCTTCTCTCTCTGTAGTGGAGATCAAC  
435 L E N L D I L Y F L L R V P H L Q I L I L N Q N R F S S C S G D Q T  
2101 CCCTTCAGAGAATCCAGCTTAGAACAGCTTTTCTTGGAGAAAATATGTTGCAACTTGCCTGGGAACTGAGCTCTGTTGGGATGTTTTGAGGGACTT  
468 P S E N P S L E Q L F L G E N M L Q L A W E T E L C W D V F E G L  
2201 TCTCATCTCAAGTTCGTATTTGAATCATAAATCTTAATTCCTTCCACAGGAGTATTTAGCCATCTGACTGCATTAAGGGGACTAAGCCCTCACT  
502 S H L Q V L Y L N H N Y L N S L P P G V F S H L T A L R G L S L N  
2301 CCAACAGGCTGACAGTCTTCTCACAATGATTTACCTGCAATTTAGAGATCTGGACATATCCAGGAACAGCTCCTAGCTCCTAATCCTGATGATT  
535 S N R L T V L S H N D L P A N L E I L D I S R N Q L L A P D V F  
2401 TGTATCACTTAGTGTCTTGGATATAACTATAACAAGTTCATTTGTGAATGTGAACCTAGCACTTTTATCAATTGGCTTAATCACACCAATGCTACTATA  
568 V S L S V L D I T H N K F I C E C E L S T F I N W L N H T N V T I  
2501 GCTGGCCCTCTGCAGADATATTTGTGTACCTGACTCGCTCTCTGGGGTTCCCTTCTCTCTTCCAGGAAGTGTGATGAAGAGGAAGTCT  
602 A G P P A D I Y P D S L S G V S L F S L S T E G C D E E V  
BspHI (2657)  
2601 TAAAGTCCCTAAAGTCTCCCTTTTTCATTGTATGCACTGCTACTGCTGTTCCCTCATGACCATCTCACAGTCACAAAGTCCGGGGCTTCTGTTT  
635 L K S L K F S L F I V C T V T L T L F L M T I L T V T K F R G F C F  
BsrGI (2765)  
2701 TATCTGTTATAAGACAGCCAGAGACTGGTGTCAAGGACCATCCCCAGGGCACAGAACCTGATATGTACAAATATGATGCCTATTTGTCTTCCAGCAGC  
668 I C Y K T A Q R L V F K D H P Q G T E P D M Y K Y D A Y L C F S S  
2801 AAAGACTTCACATGGGTGCGAATGCTTTGCTCAACACCTGGACACTCAATACAGTGACCAAAACAGATTCAACCTGTGCTTGAAGAAAGAGACTTTG  
702 K D F T W V Q N A L L K H L D T Q Y S D Q N R F N L C F E E R D F  
2901 TCCAGGAGAAAACCGCATTTGCCAATATCCAGGATGCCATCTGGAACAGTAGAAAAGTCTGTTTGTCTTGTGAGCAGACACTTCTTAGAGATGGCTGGT  
735 V P G E N R I A N I Q D A I W N S R K I V C L V S R H F L R D G W C  
3001 CTTTGAAGCCTTCACTTATGCCCAGGGCAGGTGCTTATCTGACCTTAACAGTGTCTCATCATGGTGGTGGTGGGCTTGTCCAGTACCAAGTGTGATG  
768 L E A F S Y A Q G R C L S D L N S A L I M V V V G S L S Q Y Q L M  
BsrGI (3123) BstXI (3162)  
3101 AAACATCAATCCATCAGAGGCTTTGTACAGAAACAGCAGTATTTGAGGTGGCCTGAGGATCTCCAGGATGTTGGCTGGTTTCTCATAAACTCTCTCAAC  
802 K H Q S I R G F V Q K Q Y L R W P E D L Q D V G W F L H K L S Q  
3201 AGATACTAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG  
835 Q I L K K E K E K K K D N N I P L Q T V A T I S •  
Ball (3312)  
NheI (3306)  
3301 CAAGCCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCAACTAGAATGCAGTAAAAAAATGCTTTATTTGTGAAATTTGTG  
HpaI (3444)  
3401 ATGCTATTGCTTTATTTGTAACCATTATAAGTGAATAAACAAGTTAAACAACAATTGCATTCATTTATGTTTCAGGTTCCAGGGGAGGTGTGGGA

3501 GGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACCTTAACTCCAATCAAGCCTCTACTTGAATCC  
3601 TTTTCTGAGGGATGAATAAGGCATAGGCATCAAGGGCTGTTGCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTG  
3701 TATTTTCCAAGGTTTGAAGTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAAATATTAGAAAATTAATTTAAA  
3801 TACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAAC  
3901 AAAGGAACCTTAAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCGAATTCGACTCATTCTTTGCCCTCGGACGAGTGCTGGGCGTGGTTTC  
342 • E K A R P R T S P R R N G  
4001 CACTATCGGGAGTACTTCTACACAGCCATCGTCCAGACGGCCGGCTTCTGCGGGCGATTGTGTACGCCGACAGTCCCAGCTCCGGATCGGACGAT  
328 S D A L V E V C G D T W V A A S R R A I Q T R G V T G A G S R V I  
4101 TCGCTCGCATCGACCTCGCCCAAGTGCATCGAAATGCGCTCAACCAAGCTGTAGAGTTGGTCAAGACCAATCGGAGCATATACGCCGG  
295 A D C R G Q A W A A D D F N G D V L S Q Y L Q D L G I R L M Y A R  
4201 AGCCGCGCGGATCCTGCAAGTCCGGATCGCTCCGCTCGAAGTAGCGCTGCTGCTCCATACAAGCAACCAGGCCCTCCAGAAGAAGTGTGGCGA  
261 L R P S G A L E P H R R E F Y R T Q Q E M C A L W P R W F F I N A V  
4301 CCTCGTATTGGGAATCCCGAACATCGCTCGCTCCAGTCAATGACCGCTGTTATGCGCCATTGTCGTCAGGACATTGTTGGAGCCGAAATCCGCGTG  
228 E Y Q S D G F M A E S W D I V A T I R G N D T L V N N S G F D A H  
4401 CAGGAGTGCAGGATTCGGGGCAGTCTCGCCAAAGCATCAGCTCATCGAGAGCTCGCGGACGGACGCACTGACGGTGTGCTCCATCACAGTTTC  
195 V L H R V E P C D E A W L M L E D L A Q A V S A S V T D D M V T Q  
4501 CAGTGATACACATGGGGATCAGCAATCGGCATATGAAATCACGCCATGTAGTGTATTGACCGATTCTTGGTCCGAATGGGCCGAACCCGCTCGTCT  
161 W H Y V H P D A I A C I F D R W T T Y Q G I G Q P G F P G F G S T Q  
4601 GGCTAAGATCGGGCAGCGATCGCATCCATGAGCTCCGCGACGGTTCGAGAACAGCGGGCAGTTCCGTTTCAGGCAGGCTTTCACCGTGACACCCTG  
128 S L D A A A I A D M L E A V P Q L V A P L E T E P L D Q L T V G Q  
4701 TGCACGGCGGAGATGCAATAGTCAAGTCTCGCTGAAATCCCAATGCAAGCACTCCGGAATCGGGAGCGCGCGATGCAAAGTCCGATAAAC  
95 A R R S I C Y T L S E S F E G I D L V E P I P L A A S A F H R Y V  
4801 TAAGATCTTTGTAGAAACCATCGGGCAGCTATTACCCGAGGACATATCCAGCCCTCTACATCGAAGCTGAAAGCAGGAGTCTTCCGCTCCG  
61 Y R D K Y F G D A C S N V R L V Y G R G G V D F S F A R S E E G E S  
4901 AGAGTGCATCAGTTCGGAGACGCTGTCGAACTTTTCGATCAGAACTTCGCGACAGACGTCGCGGTGAGTTCAGGCTTTTTCATGATGGCCCTCTATA  
28 L Q M L D S V S D F K E I L F K A V S T A T L E P K K M  
5001 GTGAGTCGTATTATACTATGCCGATATACTATGCGGATGATTAATTGTCAAACACGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGC  
5101 TCTGCTTATATAGACCTCCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGATTACTAGT  
5201 CAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCA  
5301 TCATGGTAATAGCGATGACTAATACGTAGTGTACTGCCAAGTAGGAAAGTCCATAAGGTGATGACTGGGCATAATGCCAGGCGGGCCATTTACCGTC  
5401 ATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAA  
5501 GTCCCTATTGGCGTTACTATGGAAACATACGTCATTATTGACGTCAATGGCGGGGGTCTTGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATG  
5601 TAACGCTGCAGGTTAATTAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGAGCCGCGTGTGGCGTTTTTCCATAGGCTCCGC  
5701 CCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAAGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCG  
5801 TGGCTCTCCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCTTTTCCCTTCCGGAAGCGTGGCGTTTCTCATAGCTACGCTGTAGGTATCT  
5901 CAGTTCGGTGTAGTCTGCTCCAAGTGGGCTGTGTGACGAACCCCGTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAAGTCC  
6001 AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGT  
6101 GGCTAACTACGGCTACACTAGAAGAACGATTTTGGTATCTGCGCTGCTGTAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTTGTATCCGGCAA  
6201 ACAAAACCCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGG  
6301 TCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTATGGCTAGTTAATTAACATTTAAATCA