



NotI (-1)
1 GCGGCCCAATAAATATCTTTATTTTCATTACATCTGTGTGTTGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACA
PvuI (172)
SgfI (171)
101 AAACAACTAGCAAAATAGGCTGTCCCCAGTCAAGTGCAGGTGCCAGAACATTTCTCTATCGAAGGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCA
AgeI (254)
201 GAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGCGCAATTGAACCGGTGCTAGAGAAGGTGGCGGGGTAACGGGAAAGTATG
Psp1406I (368)
301 TCGTGTACTGGCTCCGCCCTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAG
HindIII (410) Bsu36I (456)
401 AACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCCTACCTGAGGCCCGCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTT
501 CCCGCTGTGGTGCCTCTGAACTGCGTCCCGCTAGGTAAGTTAAAGCTCAGGTGCGAGCCGGGCTTTGTCCGGCGCTCCCTTGGAGCTACCTA
NaeI (606)
601 GACTCAGCCGGCTCTCCAGCCTTTGCTGACCTGCTTGTCAACTCTACGCTTTTGTGTTTCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACC
AgeI (717)
701 GCGCCCTACCTGAGATCACCGTAGGAGGCCATCATGACTAGCATCTTCCATTTTGCATTATCTTCATGTTAATACTTCAGATCAGAATACAATTATC
1 M T S I F H F A I I F M L I L Q I R I Q L S
801 TGAAGAAAGTGAATTTTGTAGTGTAGGTCAAAAAACGGTCTCATCCAGTTCCTAAAGACCTATCCAGAAAAACAACAATCTTAAATATATCGAAAAAT
22 E E S E F L V D R S K N G L I H V P K D L S Q K T T I L N I S Q N
EcoRV (987)
901 TATATATCTGAGCTTTGGACTTCTGACATCTTACTGTCAAAACTGAGGATTTTATAATTTCTCATAATAACAATCCAGTATCTTATGATCAGTGT
56 Y I S E L W T S D I L S L S K L R I L I I S H N T I Q Y L D I S V
1001 TCAAATCAACAGGAATTTGAACTTTGCTCCCAACAAGTTGGTGAAGATTTCTTCCACCTACTGTGAACCTCAAGCTTGGACCTTGGACCTGTC
89 F K F N Q E L E Y L D L S H N K L V K I S L C H P T V N L K H L D L S
1101 ATTAATGCATTTGATGCCCTGCTATATGCAAGAGTTTGGCAATATGTCTCAACTAAAATTTCTGGGTTGAGCACCACACACTAGAAAAATCTAGT
122 F N A F D A L P I C K E F G N M S Q L K F L G L S T T H L E K S S
EcoO109I (1279)
Bsu36I (1275)
1201 GTGTGCCAATTGCTCATTGAATATCAGCAAGGCTTGTGCTGGTCTTAGGAGAGACTTATGGGAAAAAGAAGACCCTGAGGGCCTCAAGACTTTAACA
156 V L P I A H L N I S K V L L V L G E T Y G E K E D P E G L Q D F N
1301 CTGAGAGTCTGCACATTTGTTCCCAACAACAAGAATCCATTTTATTTGGATGTGTCAGTCAAGACTGTAGCAAATCTGGAATCTAATATCAA
189 T E S L H I V F P T N K E F H F I L D V S V K T V A N L E L S N I K
1401 ATGTGTGCTAGAGAATAACAAATGTTCTTACTTCTTAAGTATTCTGGCAAACTTCAAAACAATCCAAAGTTATCAAATCTTACCTTAAACAACATTGAA
222 C V L E D N K C S Y F L S I A K L Q T N P K L S N L T L N N I E
XmnI (1508) BamHI (1522)
1501 ACAACTTGAATCTTTTCATTAGCTCCAGCTGGTTTGGCATACAACTGTATGGTATTTCTCAATTTCAAACGTGAAGCTACAGGGTCACTGGACT
256 T T W N S F I R I L Q L V W H T T V W Y F S I S N V K L Q G Q L D
StuI (1633)
1601 TCAGAGATTTGATTATTCTGGCACTTCTTGAAGCCTTGTCTATACCAAGTGTGAGGATGTGTTTCCGTTTTCCGCAAAGTTATATCTATGAAT
289 F R D F D Y S G T S L K A L L S I H Q V V S D V F G F P Q S Y I Y E I
BstBI (1703)
1701 CTTTTCGAATGAACATCAAATAATTTACAGTGTCTGGTACCGCATGGTCCACATGCTTTGCCATCCAAAATAGCCGTTCTGCATTTGGATTTT
322 F S N M N I K N F T V S G T R M V H M L C P S K I S P F L H L D F
1801 TCCAATAATCTTAAACAGACACGGTTTTTGAATAATGTGGGCACCTTACTGAGTTGGAGACACTTATTTTACAAATGAATCAATTAAGAAGACTTTCAA
356 S N N L G T D T V F E N C G H L T E L E T L I L Q M N Q L K E L S
1901 AAATAGCTGAAATGACTACACAGATGAAGTCTCTGCAACAATTTGGATATTAGCCAGAACTTCTGAAGCTATGATGAAAAAGAAAGGAGACTGTTCTTGGAC
389 K I A E M T T Q M K S L Q Q L D I S Q N S V S Y D E K K G D C S W T
2001 TAAAAGTTTATTAAGTTTAAATATGTCTTCAAATATACTTACTGACACTATTTTTCAGATGTTTCACTCCAGGATCAAGGTACTTGTATCTCACAGCAAT
422 K S L L S L N M S S N I L T D T I F R C L P P R I K V L D L H S N
HindIII (2137)
2101 AAAATAAAGAGCATTCCTAAACAAGTCGTAAGACTGGAAGCTTTGCAAGAACTCAATGTTGCTTTCAATTTCTTAACTGACCTTCTGGATGTGGCAGCT
456 K I K S I P K Q V V K L E A L Q E L N V A F N S L T D L P G C G S
2201 TTAGCAGCCTTTCTGTATTGATCATTGATCACAATTCAGTTTCCACCCATCGGCTGATTTCTCCAGAGCTGCCAGAAGATGAGGTCAATAAAAGCAGG
489 F S S L S V L I I D H N S V S H P S A D F F Q S C Q K M R S I K A G
2301 GGACAATCCATTCATGCTACCTGTGAGCTAGGAGAATTTGTCAAAATATAGACCAAGTCAAGTGAAGTGTAGAGGGCTGGCCTGATTCTTATAAG
522 D N P F Q C T C E L G E F V K N I D Q V S S E V L E G W P D S Y K
BspLUIII (2446) BstXI (2497)
2401 TGTGACTACCCGAAAGTTATAGAGGAACCTACTAAAGGACTTTTCCATGTCTGAATLCTCTGCAACATAACTCTGTGATCGTACCATCGTTGCCA
556 C D Y P E S Y R G T L L K D F H M S E L S C N I T L L I V T I V A
XcmI (2500) PstI (2534)
2501 CCATGCTGGTGTGGCTGTGACTGTGACCTCCCTCTGCACTACTTGGATCTGCCTTGGTATCTCAGGATGGTGTGCCAGTGGACCCAGCCGGCGCAG
589 T M L V L A V T V T S L C S Y L D L P W Y L R M V C Q W T Q T R R R
2601 GGCCAGGAACATACCCTTAGAAGAAGTCCAAAGAAATCTCCAGTTTTCATGATTTATTTTATATAGTGGGCAGGATCTTTCTGGGTGAAGAATGAATTA
622 A R N I P L E E L Q R N L Q F H A F I S Y S G H D S F W V K N E L
2701 TTGCCAAACCTAGAGAAAGAGGATGAGATTTGCCTTCAAGAGAGAACTTTGTTCTGGCAAGAGCATTGTGAAAAATATCATCACCTGCAATGAGA
656 L P N L E K E G M Q I C L H E R N F V P G K S I V E N I I T C I E
2801 AGAGTTACAAGTCCATCTTTGTTTGTCTCCAACTTTGTCCAGAGTGAATGGTGCATTATGAACCTACTTTGCCCATCACAACTCTTTTTCATGAAGG
689 K S Y K S I F V L S P N F V Q S E W C H Y E L Y F A H H N L F H E G
XmnI (2928) ScaI (2941) Ball (2983)
2901 ATCTAATAGCTTAATCCTGATCTTGTGGAACCCATTCCGCACTACTCCATTCTAGCAGTTATCACAAGCTCAAAGTCTCATGGCCAGGAGGACTTAT
722 S N S L I L I L L E P I P Q Y S I P S S Y H K L K S L M A R R T Y
SspI (3063) AseI (3060)
3001 TTGGAATGGCCCAAGGAAAAGAGCAAACGTGGCCTTTTTGGCTAACTAAGGGCAGCCATTAATATTAAGCTGACAGAGCAAGCAAGAAATAGTCTA
756 L E W P K E K S K R G L F W A N L R A A I N I K L T E Q A K K •
NheI (3107)
3101 GGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGACAACCCACAACCTAGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCT
HpaI (3239)
3201 ATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAATTCATTCTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTT
3301 TTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTTCAAAATACAGCATAGCAAACTTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTC

3401 TGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGATT

3501 TCCCAAGTTTGAAC TAGCTCTTCATTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTAGAAAATATTTAAATACAT

3601 CATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAG

3701 AACCTTTAATAGAAATTTGGACAGCAAGAAAGCGAGCTTCTAGCGAATTCTCGACTCATTCTTTGCCCTCGGACGAGTGCTGGGGCTCGGTTCCACTA

3801 TCGGGGAGTACTTCTACAGCCATCGGTCAGACGGCCGCGCTTCTGCGGGCGATTGTGTACGCCGACAGTCCCGGCTCCGGATCGGACGATTGCGT

3901 CGCATCGACCCTGCGCCCAAGCTGCATCATCGAAATGCGCTCAACCAAGCTCTGATAGAGTTGGTCAAGACCAATCGGGAGCATATACGCCGGAGCCG

4001 CGGGATCTGCAAGCTCCGGATGCCTCCGCTCGAAGTAGCGCTGCTGCTCCATACAAGCCAACACGGCCTCCAGAAGAAGATGTTGGCGACTCG

4101 TATTGGGAATCCCCAAGCTCGCTCGCTCCAGTCAATGACCGCTGTATGCGGCCATTGCTCCGTCAGGACATTGTTGGAGCCGAAATCCGCGTGCACGA

4201 GGTGCCGACTTTCGGGGCAGTCCCTCGGCCAAAGCATGAGTGCATCGAGAGCCTGCGCGACGGACGCACTGACGGTGTCTGTCATCACAGTTTCCAGTG

4301 ATACACATGGGGATCAGCAATCGCGCATATGAAATCACGCCATGTAGTGTATTGACCGATTCTTTCGGTCCGAAATGGGCCAACCCTCGTCTGGCTA

4401 AGATCGGCCGACGCGATCGCATCCATGAGCTCCGCGACGGGTTGCGAAGACAGCGGGCAGTTCCGGTTTCAGGCAGGTCTTCAACGTCACACCCTGTGAC

4501 GCGGGGAGATGCAATAGGTCAGGCTCTCGCTGAAATCCCAATGTCAGCACTCCGGAATCGGGAGCGCGGCGATGCAAAAGTCCGATAAACATAACG

4601 ATCTTTGAGAAACCATCGGCGCAGCTATTTACCCGACGAGCATATCCACGCCCTCTACATCGAAGCTGAAAGCACGAGATTCTTCGCCCTCCGAGAG

4701 TGCATCAGTTCGGAGACGCTGTGCAACTTTTCGATCAGAACTTCGCGACAGAGCTGCGGGTGAAGTTTCAGGCTTTTTCATGATGCCCTCTATAGTGAG

4801 TCGTATTACTACTGCGGATATACTATGCGGATGATTAATTGTCAAACACGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCCTAAACGAGCTCTGC

4901 TTATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAA

5001 CAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATG

5101 GTAATAGCGATGACTAATACGTAGTACTGCCAAGTAGGAAAGTCCATAAGGTGATGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCAATTGA

5201 CGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCTCAAGTGGGCGGTTTACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCC

5301 TATTGGCGTTACTATGGGAACATACGCTCATTATTGACGTCAATGGGCGGGGTCGTTGGCGGTCAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACG

5401 CCTGCAGGTTAATTAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTGTGGCGTTTTTCCATAGGCTCCGCCCCC

5501 TGACGAGCATCAGAAAATCGACGCTCAAGTCAAGGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCC

5601 TCTCCTGTTCCGACCTGCGGCTTACCGGATACCTGTCCGCTTTCTCCCTTGGGAGCGTGGCGCTTCTCATAGCTACGCTGTAGGTATCTCAGTT

5701 CGGTGTAGTTCGTTCCCAAGCTGGGCTGTGTGACGAACCCCGTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTAGTCCAAACC

5801 GGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCCT

5901 AACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTGCTGTAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTAGTCCGGCAACAAA

6001 CCACCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACGGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGA

6101 CGCTCAGTGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCA