



PvuI (7)
SgfI (6)

1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTIGCC

HindIII (245)

201 GTGAACGTTCTTTTTCGCAACGGGTTTGC GCCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGCTCAACTCTACGCTTTTGTTCGTTT
501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGTAGGAGGGCCATCATGATGTCTGCCTCGCGCTGGCTGGGACT
1 M M S A S R L A G T

NcoI (610)

601 CTGATCCCAAGCCATGGCCTTCTCTCCTGCGTGAGACCCGAAAGCTGGGAGCCCTGCGTGAGGTGGTTCCTAATATTACTTATCAATGCATGGAGCTGA
11 L I P A M A F L S C V R P E S W E P C V E V V P N I T Y Q C M E L
701 ATTTCTACAAAATCCCGACAACCTCCCTTCTCAACCAAGAACCTGGACCTGAGCTTTAATCCCTGAGGCATTTAGGCAGCTATAGCTTCTTCAGTTT
44 N F Y K I P D N L P F S T K N L D L S F N P L R H L G S Y S F F S F
801 CCCAGAAGTGCAGGTGCTGGATTATCCAGGTGAAAATCCAGACAATTGAAGATGGGCATATCAGAGCCTAAGCCACCTCTCTACCTTAATATTGACA
77 P E L Q V L D L S R C E I Q T I E D G A Y Q S L S H L S T L I L T
901 GAAACCCCATCCAGAGTTAGCCCTGGGAGCCTTTCTGGACTATCAAGTTACAGAAGCTGGTGGCTGTGGAGACAATCTAGCATCTCTAGAGAAT
111 G N P I Q S L A L G A F S G L S S L Q K L V A V E T N L A S L E N
1001 TCCCATTGGACATCTCAAACTTTGAAAGAATTAATGTGGCTCACAATCTTATCCAATCTTTCAAATTACTGAGTATTTTTCTAATCTGACCAATCT
144 F P I G H L K T L K E L N V A H N L I Q S F K L P E Y F S N L T N L
1101 AGAGTACTTGGACCTTCCAGCAACAAGATTCAAAGTATTTATGCACAGACTTGGGGTCTACATCAAATGCCCTACTCAATCTCTCTTAGACCTG
177 E Y L D L S S N K I Q S I Y C T D L R V I H Q M P L L N L S L D L
1201 TCCTGAACCTATGAACCTTATCCAACAGGTGCATTTAAAGAAATTAGGCTTCATAAGCTGACTTAAAGAAATAATTTGATAGTTAAATGTAATGA
211 S L N P M N F I Q P G A F K E I R L H K L T L R N N F D S L N V M
1301 AAACCTGTATTCAAGTCTGGCTGGTTAGAACTCCATCGTTTGGTCTGGGAGAATTTAGAAATGAAGGAACTGGAAAAGTTGACAAAATCTGCCT
244 K T C I Q G L A G L E V H R L V L G E F R N E G N L E K F D K S A L
1401 AGAGGGCTGTGCAATTTGACCATTGAAGAATCCGATTAGCATACTAGACTACTACCTCGATGATATTATTGACTTATTTAATTTGACAAAATGTT
277 E G L C N L T I E E F R L A Y L D Y Y L D D I I D L F N C L T N V
1501 TCTTCATTTCCCTGGTGTGACTATTGAAAGGGTAAAAGACTTTTCTTATAATTTCCGATGGCAACATTTAGAATTAGTTAACTGTAATTTGGAC
311 S S F S L V S V T I E R V K D F S Y N F G W Q H L E L V N C K F G
1601 AGTTTCCACATTGAACTCAAATCTCTCAAAAGGCTTACTTTCACCTCCAACAAAGTGGGAATGCTTTTTCAGAAGTTGATCTACCAAGCCTTGAGTT
344 Q F P T L K L K L K R L T T T S N K G G G N A F S E V D L P S L E F
1701 TCTAGATCTCAGTAGAAATGGCTTGAGTTCAAAGTTGCTGTTCTCAAAGTATTTGGGACAACCAGCCTAAAGTATTTAGACTGAGCTTCAATGGT
377 L D L S R N G L S F K G C C S Q S D F G T T S L K Y L D L S F N G
1801 GTTATTACCATGAGTTCAAATCTTGGGCTTAGAACAACCTAGAACATCTGGATTTCCAGCATTCAAATTTGAAACAATGAGTGAGTTTTAGTATTCC
411 V I T M S S N F L G L E Q L E H L D F Q H S N L K Q M S E F S V F
1901 TATCACTCAGAAACCTCATTACCTGACATTTCTCACTACACACAGAGTTGCTTTCAATGGCATCTTCAATGGCTGTCCAGTCTCGAAGCTTGAA
444 L S L R N L I Y L D I S H T H T R V A F N G I F N G L S S L E V L K

EcoRV (2037)

2001 AATGGCTGGCAATTTTCCAGGAAACTTCTTCCAGATATCTTCCAGAGCTGAGAAAATGACCTTCTGACCTTCTCAGTGTCAACTGGAGCAG
477 M A G N S F Q E N F L P D I F T E L R N L T F L D L S Q C Q L E Q
2101 TTGTCTCAACAGCATTTAACTCACTCTCCAGTCTTCCAGTACTAAATATGAGCCACAACAATCTTTTTATTGGATACGTTTCTTATAAGTGTCTGA
511 L S P T A F N S L S S L Q V L N M S H N N F F S L D T F P Y K C L
2201 ACTCCCTCCAGGTTCTTGATTACAGTCTCAATCACATAATGACTTCCAAAAACAGGAACTACAGCATTTCCTCAAGTAGTCTAGCTTTCTTAAATCTTAC
544 N S L Q V L D Y S L N H I M T S K K Q E L Q H F P S S L A F L N L T
2301 TCAGAATGACTTTGCTTGTACTTGTGAACACCAGAGTTTCTGCAATGGATCAAGGACCAGAGGCAGCTCTTGGTGGAAAGTTGAACGAATGGAATGTGGC
577 Q N D F A C T C E H Q S F L Q W I K D Q R Q L L V E V E R M E C A
2401 ACACCTTCAGATAAGCAGGGCATGCCTGTGCTGAGTTGAATATCACCTGTGAGATGAATAAGACCATCATTGGTGTGCTCGGTCTCAGTGTCTGTAG
611 T P S D K Q G M P V L S L N I T C Q M N K T I I G V S V L S V L V
2501 TATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTTTCCCTGATGCTTCTGCTGGCTGCATAAAGTATGGTAGAGGTGAAACATCTATGATGCCTT
644 V S V V A V L V Y K F Y F H L M L L A G C I K Y G R G E N I Y D A F
2601 TGTATCTACTCAAGCCAGGATGAGGACTGGTAAGGAATGAGCTAGTAAAGAATTTAGAAGAAGGGTGCCTCCATTTACAGTCTGCCTTCACTACAGA
677 V I Y S S Q D E D W V R N E L V K N L E E G V P P F Q L C L H Y R
2701 GACTTTATCCCGGTGTGCCATTGCTGCCAACATCCATGAAGGTTTCCATAAAAAGCCGAAAGGTGATTGTTGGTGTCCAGCACTTCAATCCAGA
711 D F I P G V A I A A N I I H E G F H K S R K V I V V V S Q H F I Q
2801 GCCGCTGGTGTATCTTTGAATATGAGATTGCTCAGACCTGGCAGTTTCTGAGCAGTGTGCTGGTATCATCTTATTGCTCCTGCAGAAGGTGGAGAAGAC
744 S R W C I F E Y E I A Q T W Q F L S S R A G I I F I V L Q K V E K T
2901 CCTGCTCAGGCAGAGGTGGAGCTGTACCGCTTCTCAGCAGGAACACTTACCTGGGGTGGGAGGACAGTGTCTGGGGCGGCACATCTTCTGGAGACGA
777 L L R Q Q V E L Y R L L S R N T Y L G W E D S V L G R H I F W R R
3001 CTCAGAAAAGCCCTGCTGGATGGTAAATCATGGAATCCAGAAGGAACAGTGGGTACAGGATGCAATTTGGCAGGAAGCAACATCTATCTGTACAAGGta
811 L R K A L L D G K S W N P E G T V G T G C N W Q E A T S I L Y K G

3101 agtcactgactgtctatgcctgggaaaggggtgggcaggagatggggcagtgaggagaaagtggcactatgaacccACTAGTTTGACAATTAATCATAAGC
844▶ K S L T V Y A W E R V G R R W G S A G K V A L • T H • F D N • S • A
3201 ATAGTATAATACTACTATAGcaatgtactaaccttcttctcttctcctcctgacagGAGGAGCCATCATGGCCGCTATGGAGATCGAGTGCCGC
877▶ • Y N T T H Y S N C T N L L L F P L L T G G A I M A A M E I E C R
3301 ATCACCGCACCTGAACGGCGTGGAGTTCGAGCTGGTGGCGGGGAGAGGGCACCCCGAGCAGGGCCGCATGACCAACAAGATGAAGAGCACCAAG
911▶ I T G T L N G V E F E L V G G G E G T P E Q G R M T N K M K S T K
3401 GCGCCCTGACCTTCAGCCCTACCTGCTGAGCCACGTGATGGCTACGGCTTCTACCCTTCCGACCTACCCAGCGGCTACGAGAACCCTTCTGCA
944▶ G A L T F S P Y L L S H V M G Y G F Y H F G T Y P S G Y E N P F L H
3501 CGCATCAACAACGGCGGTACACCAACACCCGCATGAGAAGTACGAGGACGGCGGTGCTGCACGTGAGCTTACGCTACCGCTACGAGGCGGCCGC
977▶ A I N N G G Y T N T R I E K Y E D G G V L H V S F S Y R Y E A G R
3601 GTGATCGGCGACTTCAAGGTGATGGGCACCGCTTCCCGAGGACAGCGTATCTTACCACCAAGATCATCCGACGCAACGCCACCGTGGAGCACCTGC
1011▶ V I G D F K V M G T G F P E D S V I F T D K I I R S N A T V E H L
3701 ACCATATGGCGATAACGATCGAGTTCGAGCTTACCCTGAGCTTACCCTGCGCAGCGGCTACTACAGCTCCGTGGTGACAGCCACATGCACTT
1044▶ H P M G D N D L D G S F T R T F S L R D G G Y Y S S V V D S H M H F
3801 CAAGAGCGCCATCCACCCAGCATCCTGCGAAGCGGGGCCCATGTTGCGCTTCCGCGCGTGGAGGAGGATCACAGCAACCCGAGCTGGGCATCGTG
1077▶ K S A I H P S I L Q N G G P M F A F R R V E E D H S N T E L G I V

NheI (3943)

3901 GAGTACCAGCACGCTTCAAGACCCCGGATGCAGATGCCTAAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGA
1111▶ E Y Q H A F K T P D A D A •
4001 ATGCAGTGAATAAATGCTTTATTGTGAAATTTGTGATGCTATTGCTTTATTGTAAACCATTATAAGCTGCAATAAACAAGTTAACAAACAATTGCA
4101 TTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTTCTAAAATACAGCATAGCA
4201 AAACCTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTG
4301 CAGCCTCACCTTCTTTCATGGAGTTTAAGATATAGTGTATTTTCCCAAGTGTGAACTAGCTCTTATTCTTTATGTTTTAAATGCACTGACCTCCCAC
4401 ATTCCCTTTTATGTAATAATTCAGAAATAATTTAAATACATCATTGCAATGAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTC
4501 ATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGGAACCTTAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGT
141▶ • N R T Y
4601 ACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCT
136▶ K L P I L E E I T T K V L K G N M E I L V F C D P A Y D S I L E R
4701 GCACATGCCACAGGGGCTGACCACCTGATGGATCTGCCACCTCATCAGATAGGGGTGCCTGACAGCCACAATGGTCAAAGTCTTCTGCCCGTTG
103▶ C M G C P S V V R I S R D V E D S Y P H R V A V I T D F D K Q G N
4801 CTCACAGCAGACCAATGGCAATGGCTTCCAGCAGACAGTACCCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCC
69▶ S V A S G I A I A E A C V T V R G I Y A E I H V A S I I E G T K T R
4901 TGATGGCCGCCCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAA
36▶ I A A G V H H K N D E Y L M T I K E T A V E V L E L D Q Q S I N F
5001 GGTCTTATGATGGCCCTCTATAGTGAAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCT
3▶ T K M
5101 TATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTT
5201 GGAAGTCCCCTGATTACTAGTCAAAACAACTCCATTGACGTCAATGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCCAT
5301 TGATGACTGCCAAAACGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGACTGGGCATA
5401 ATGCCAGGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCTCAAGTGGGCGAGTTTACCGTAAATAC
5501 TCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTGAGCCAG

BspLU111 (5647)

5601 GCGGGCCATTTACCGTAAGTTATGTAACGCCTGACGGTAAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTG
5701 CTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGG
5801 CGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTC
5901 TCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTGCTCCAAGTGGGTGTGTGCACGAACCCCGTTACGCCGACCGCTGCGCTTA
6001 TCCGTAACATATCGTCTTGTGATCCAACCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCG
6101 GTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAG
6201 AGTTGGTAGTCTTGTCCGCAAAACAAACCCGCTGGTAGCGGTGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAA
6301 GATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAACTACGTTAAGGGATTTTGGTATGGCTAGTTAATTAACATTTAAATCAGCGG

NotI (6396)

6401 CCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAC
6501 AACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA