



1 GGACCTGCAGGGCTGAATAACCTCTGAAAGAGGAAGCTGGTTAGGTACCTCTGAGCGGAAAGAACAGCTGTGGAATGTGTGTCAGTTAGGGTGTG
101 GAAAGTCCCAGGCTCCCAGCAGGCAGAAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGAAAGTCCCAGGCTCCCAGCAGGCAG
201 AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCACTAGTCTCCGGTGGCCGTGAGTGGGAGAGCGCACATCGCCACAGTCCCCGA
301 GAAGTTGGGGGAGGGGTGCGCAATTGAACGGGTGCCTAGAGAAGGTGGCGGGGTAACCTGGGAAAGTGTGCTGTACTGGTCCGCTTTTTCCC
401 GAGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGGCCGAGAACACAGCTGAAGCTTCGAGGGGCTC
501 GCATCTCTCTTACGCGCCCGCCCTACCTGAGGCGCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAACTGC
601 GTCCGCGCTTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCCTCACGCTTTGC

NcoI (799)

Agel (791)

701 CTGACCCTGCTTGCTCAACTCTACGTCTTTGTTTCTGTTTCTGTTCTGCGCAGTTACAGATCCAAGCTGTGACCGGGCCTACCTGAGATCACCGGTAC

XhoI (852)

801 CATGGAGATCAAGGTGCTGTTTGCCTCATCTGTATTGCTGTTGCTGAGGCACTCGAGGTTAATCTACCACCCGAAGTCAACTCCCACCCGCATATACA
1 M E I K V L F A L I C I A V A E A L E V N L T T R T Q L P P A Y T
901 AATTCTTACCAGAGGAGTGTACTATCCTGACAAAGTGTTCGGTCAAGTGTCTCCACTTACTCAGGACCTTTTCTGCCTTTCTTTTCTAACGTTA
15 N S F T R G V Y Y P D K V F R S S V L H S T Q D L F L P F F S N V
1001 CATGGTTTCATGCAATCCATGTGTCTGGGACAAACGGCACAAACGCTTGCACAACTGTATTGCCATTCAATGATGGGGTGTACTTTGCCTCCACAGA
48 T W F H A I H V S G T N G T K R F D N P V L P F N D G V Y F A S T E
1101 GAAATCCAACATCATTGAGGATGGATTTTTCGGGACTACTCTGGACTCAAAGACACAGAGCCTGTGATCGTTAACAACGCCACAAACGTTGTATCAA
81 K S N I I R G W I F G T T L D S K T Q S L L I V N N A T N V V I K
1201 GTGTGCGAATTCAGTTTTGCAATGATCCCTTCTGGGAGTGTACTATCACAAGAATAACAAGTCTGGATGGAGAGCGAATTTCCGGTCTACAGCAGCG
115 V C E F Q F C N D P F L G V Y Y H K N N K S W M E S E F R V Y S S
1301 CAAACAACTGCACCTTTCGAGTACGTGAGTCAACCTTTCTGATGGACTGGAAGGAAACAGGAAACTTCAAGAACCTGAGAGAGTTTGTCTTTAAGAA
148 A N N C T F E Y V S Q P F L M D L E G K Q G N F K N L R E F V F K N
1401 CATCGACGGCTATTTAAGATCTATAGTAAGCATAACGCTATCAACCTGGTAAAGGATCTTCCCAGGGCTTTTTCAGCCCTGGAACCTTTGGTTGACTTG
181 I D G Y F K I Y S K H T P I N L V R D L P Q G F S A L E P L V D L
1501 CCTATTGGTATCAATATCACCAGATTTACAGACCTTTCGGCATTGCAICGGTCTTACTTACTCCAGGTGATTCCTCCTCCGGGTGGACTGCCGGCGCCG
215 P I G I N I T R F Q T L L A L H R S Y L T P G D S S P G D S W T A G A
1601 CTGCTACTATGTCGGCTATCTGCAACCAAGAAGTCTCTGCTCAAGTACAACGAAAACGGCACTATTACGGATGCTGTTGATTGTGCCCTGAGACCCCT
248 A A Y Y V G Y L Q P R T F L L K Y N E N G T I T D A V D C A L D P L
1701 GTCTGAGACTAAATGCACCTCAAGAGCTTTACCGTTGAGAAGGGGATTTACCAAACAGTAATTTCCGGGTCCAACCCAGAAAGCATTGTGCGGTTT
281 S E T K C T L K S F T V E K G I Y Q T S N F R V Q P T E S I V R F
1801 CCAAATATCACCATCTGTGTCCCTTTGGCGAAGTGTCAATGCTACAAGGTTTGTCTGTGTACGCATGGAATAGGAAACGCATCTCCAATTGTGTGCG
315 P N I T N L C A P F G C E V F N A T R F A S V Y A W N R K R I S N C V
1901 CTGATTACTCGTGTGTCAATTCGGCTCTTTCTCAACCTTCAAGTGTATTGGCGTTTACCTACCAAACCTAACGACCTGTGCTTCACTAATGTGTA
348 A D Y S V L Y N S A S F S T F K C Y G V S P T K L N D L C F T N V Y
2001 TGCCGACTCTTTTGTGATACGAGGCGATGAAGTGAACAGATTGCACAGGGCAGACCGGCAAAATTGCCGACTACAACCTACAAGCTTCCAGATGACTTT
381 A D S F V I R G D E V R Q I A P G Q T G K I A D Y N Y K L P D D F
2101 ACCGGATGTGTTATTGCATGAACTCAAACAATCTGGATTCCAAGGTGGGTGGCAACTATAACTACCTGTATAGACTGTTCCAGGAAATCCAACCTGAAAC
415 T G C V I A W N S N N L D S K V G G N Y N Y L Y R L F R K S N L K
2201 CATTGAGCGAGATATAAGCACAGAAATCTACCAGGCTGGAAGTACGCCCTGCAACGGCGTGAAGGGTTCAACTGCTACTTCCCATTGCAGAGTTACGG
448 P F E R D I S T E I Y Q A G S T P C N G V E G F N C Y F P L Q S Y G
2301 ATCCAGCCTACAAACGGGGTGGGTTACCAACCTATCGTGTCTGATGCTGAGTTTTGAGCTCCTCATGCCCCAGCCACAGTCTGTGGCCCAAGAAA
481 F Q P T N G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K K
2401 AGCACCAATCTGGTGAAGAACAATGCGTGAACCTTAACTTAAACGGACTCACAGGAACCGCGTATTGACGGAGAGTAACAAGAAGTTCTGCCATTCC
515 S T N L V K N K C V N F N F N G L T G T G V L T E S N K K F L P F
2501 AGCAGTTCGGTCCGATATTGCCGACACTACCGACGCTGTCCGAGATCCCAGACATTGGAGATTCTTGATATCACACCTGTAGTTTCCGGCGAGTGAG
548 Q Q F G R D I A D T T D A V R D P Q T L E I L D I T P C S F G G V S
2601 CGTGATTACGCCCGGAACCAATACCAGCAATCAGGTTGCCGCTCTGTATCAGGACGTGAATTCACCCAGGTTACCTGTCCGATCCACGCTGACCAACT
581 V I T P G T N T S N Q V A V L Y Q D V N C T E V P V A I H A D Q L
2701 ACACCCATGGCGAGTATATCCACCGGCTCCAACGCTTTTCCAGACAGTGTGGATGTCTGATCGGTGCAGAAACGTTAATAAGTACGAGTGTG
615 T P T W R V Y S T G S N V F Q T R A G C L I G A E H V N N S Y E C

BamHI (2868)

NheI (2899)

2801 ATATCCCATCGGTGCTGGAATATGCGCCTTATCAAACCTCAAACCAACTCTCTAGCGGGCAGTGGATCCGGCCATCATCATCACCATCACTAAAG
648 D I P I G A G I C A S Y Q T Q T N S P R R A R G S G H H H H H H •
2901 CTAGTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCAACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTAT
3001 TGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAATTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTTT
3101 TAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTAACTCCAATCAAGCCTCTACTGAATCCTTTTCTG

3201 AGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAAGATATAGTGATTTTC
3301 CCAAGGTTTGAAGTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCA
3401 TTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAA
3501 CCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTG
141 • N R T Y K L P I L E E I T T K V L K
3601 CCATTCATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCT
121 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 S R D V E
3701 CATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGAC
88 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 C V T V
3801 CCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCCTCCAGCTTTGGTCTGTATGGCCGCCCGACATGGTCTTGTGTCTCATAGAGCATG
55 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 E Y L M
3901 GTGATCTTCTCAGTGGCGACCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCTATTATACTATG
21 T I K E T A V E V L E L D Q Q S I N F T K M
4001 CCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCC
4101 ACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGATTTACTAGTCAAAAACAACTCCCATTGAC
4201 GTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACT
4301 AATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGTGTACTGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCG
4401 TACTTGGCATATGATACACTTGTACTGCAAGTGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTAT
4501 GGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCTGTTGGGCGTCCAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACCGCTGCAGGTTAATTA
4601 AGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAA
4701 AAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGTCCCTCGTGCCTCTCTGTTCCGACC
4801 CTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT
4901 GCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGCCGACCGTGGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACAGACTT
5001 ATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACT
5101 AGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCG
5201 GTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGA
5301 AAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTT
5401 TTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAACGAAACAAACAAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAG
5501 AACATTTCTATCGAA