



1 GGACCTGCAGGGCTGAATAACCTCTGAAAGGGAAGCTTGGTTAGGTACCTTCTGAGCGGAAAGAACAGCTGTGGAATGTGTGTCAGTTAGGGTGTG
101 GAAAGTCCCAGGCTCCCAGCAGGCAGAAATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGAAAGTCCCAGGCTCCCAGCAGGCAG
201 AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCACTAGTCTCCGGTGGCCGTGAGTGGGAGAGCGCACATCGCCACAGTCCCCGA
301 GAAGTTGGGGGAGGGGTGCGCAATTGAACGGGTGCCTAGAGAAGGTGGCGGGGTAACCTGGGAAAGTGTGCTGTACTGGCTCCGCTTTTTCCC
401 GAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGGCCGAGAACACAGCTGAAGCTTCGAGGGGCTC
501 GCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCGCCATCCACGCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAACTGC
601 GTCCGCGCTTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCCTCACGCTTTGC

NcoI (799)
AgeI (791)

701 CTGACCCTGCTTGTCAACTCTACGTCTTTGTTTCTGTTTCTGTTCTGCGCAGTTACAGATCCAAGCTGTGACCGGGCCTACCTGAGATCACCGGTAC

XhoI (852)

801 CATGGAGATCAAGGTGCTGTTTGCCTCATCTGTATTGCTGTTGCTGAGGCACTCGAGATGTCCGATAATGGGCCACAGAACCAGAGAAATGCACCCCGG
1 M E I K V L F A L I C I A V A E A L E M S D N G P Q N Q R N A P R
901 ATCACCTTTGGAGGACCTTCAAGTCAACGGGAGTAAACCAGAATGGTGGAGGATCTGGAGCAGTCTAAACAACGAGACCTCAAGGGCTCCAAACA
15 I T F G G P S D S T G S N Q N G E R S G A R S K Q R R P Q G L P N
1001 ATACCGCTAGTTGGTTTACTGCTCTCACACAGCATGGGAAGGAGGATCTGAAATTTCCACGAGGGCAAGGCGTCCCATCAACACCAACTCCAGTCCCGA
48 N T A S W F T A L T Q H G K E D L K F P R G Q G V P I N T N S S P D
1101 CGACCAGATCGGATACTACAGACGAGCAACCCGGCGAATCAGAGGCGGCGATGGCAAATGAAAGACCTGAGTCCCAGGTGGTACTTCTATTATCTCGGA
81 D Q I G Y Y R R A T T R R I R G G D G K M K D L S P R W Y F Y Y L G
1201 ACAGGCCCTGAAGCCGGCTGCCCTATGGTCCCAACAAGACGGCATCATCTGGGTAGCTACTGAGGGAGCCTTGAACACTCCAAGGACACACATGGTA
115 T G P E A G L P Y G A N K D G I I W V A T E G A L N T P K D H I G
1301 CCCGCAATCCCGCAAACAATGCTGCCATTGTGCTCCAGTTGCCACAGGGCACAACCTTCCCAAGGGTTTCTATGCCGAAGGCTCTCGTGGAGGATCTCA
148 T R N P A N N A A I V L Q L P Q G T T L P K G F Y A E G S R G G S Q
1401 AGCCAGCAGCCGTTCTCTCTCGGAGTAGGAATAGCTCTCGAACTCCACACCAGGACAGCTCTCGCGGAACAAGCCCTGCTCGCATGGCCGTAACGGT
181 A S S R S S S R S R N S S R N S T P G S S R G T S P A R M A G N G
1501 GGTGATGCCGCTCTGGCTCTGCTTCTCTGGACAGGCTGAATCAGCTGGAATCCAAGATGTCTGGTAAGGGACAACAGCAACAGGGGAGACCGTTACCA
215 G D A A L A L L D R L N Q L E S K M S G K G Q Q Q G G Q T V T
1601 AGAAGTCCGCTGCAGAGCCAGCAAGAAACCTCGGCAGAAGAGGACCCCACTAAAGCCTACAACGTGACACAGGCCCTTCGGCAGGAGGGCCAGAACA
248 K K S A A E A S K K P R Q K R T A T K A Y N V T Q A F G R R G P E Q
1701 GACACAGGGCAACTTTGGTGATCAGGAGTTGATACGCCAGGGGACAGACTACAAGCATTGGCCACAGATTGCACAGTTCGCACCTCCGCTAGTGCCTTC
281 T Q G N F G D Q E L I R Q G T D Y K H W P Q I A Q F A P S A S A F
1801 TTTGGGATGTCACGGATAGGCATGGAGGTGACGCTAGCGGACATGGCTGACCTATACTGGAGTATCAAACCTGGATGACAAAGACCCCACTTCAAGG
315 F G M S R G I G M E V T P S G T W L T Y T G A I K L D D K D P N F K
1901 ACCAGTGATTCTGTGAATAAGCATTGATGCTCAAAAACGTTTCTCCTCAACTGAGCCCAAGGACAAGAAGAAGGCCGATGAAACCCAGGC
348 D Q V I L L N K H I D A Y K T F P P T E P K K D K K K A D E T Q A
2001 ACTTCTCAGAGGCAGAAGAAACAGCAGACCGTCACTCTTCTGCCGAGCGGATCTGGACGACTTCTCAAAACAGTTGCAACAGTCCATGAGCTCAGCC
381 L P Q R Q K K Q Q T V T L L P A A D L D D F S K Q L Q Q S M S S A

BamHI (2115)

NheI (2146)

2101 GATTCCACCAAGCTGGATCCGGCCATCATCATCACCATCACTAAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAC
415 D S T Q A G S G H H H H H H •
2201 AGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAAT
2301 GCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATA
2401 GCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGT
2501 TTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTCTTCATTTCTTTATGTTTAAATGCACTGACCTCC
2601 CACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCC
2701 TTCATAATATCCCCAGTTTAGTGTGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGG
2801 TGTACTTGAGGGGATGAGTTTCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTC
137 Y 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
2901 TCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCG
104 R 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
3001 TTGCTCAGCAGACCCAATGGCTTTCAGCAGCAGTGTGACCTGCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTGG
70 N 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 E G F T K T
3101 TCCTGATGGCCGCCCCGACATGGTGTGTTGCTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGT
37 R 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

3201 GAAGGTCTTCATGGTGGCCCTCCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACAGCGTGGATGGCGTCTCCA
4 F T K M
3301 GCTTATCTGACGGTTCACATAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACAT
3401 TTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAAATCCCCGTGAGTCAAACCGCTATCCACGCC
3501 CATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCATAAGGTCATGTAAGTGGGC
3601 ATAATGCCAGGCGGGCCATTTACCGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCAAGTGGGCAGTTTACCGTAAA
3701 TACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGC
3801 CAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCG
3901 TTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACC
4001 AGGCGTTTTCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCT
4101 TTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTAGCCCGACCGCTGCGCC
4201 TTATCCGGTAACTATCGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAA
4301 GCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAA
4401 AAGAGTTGGTAGCTCTTATCCGCAAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAA
4501 GAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAG
4601 CGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTTGTGTAATCGTAACTAACATACGCTCTCCATCAAACAAACGAAACAA
4701 AACAACTAGCAAAATAGGCTGTCCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA