



100

1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTTCGCCGCTCCCGCTGTGGTGCCTCTGAAGCTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

BspHI (560)

AgeI (552)

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCATGAGTACTCCACAGAAGCCAAGATGCAGCCTCTTAG
1 M S D S T E A K M Q P L S
601 CTCCATGGACGATGATGAGTTGATGGTCAGCGGCAGCAGGTATTCTATTAAGTCCAGACTACGACCAAATCTGGAATCAAGTGTGGCAGGATGC
13 S M D D D E L M V S G S R Y S I K S S R L R P N S G I K C L A G C
701 TCGGGACACAGCCAAGTCCCCTTGGTCTCGAGCTGCTCTCCTTCTGTTCTGGCTGGGCTCCTGCTGATCATTCTTTCCAAGTCTCCAAAACCCAA
47 S G H S Q V P L V L Q L L S F L F L A G L L L I I L F Q V S K T P
801 ATACCGAGAGGCAGAAGGAACAAGAGAAGATCTCCAGGAAGTACCAGCTGACAGATGAGCTTACGTCAGGATCCCCTCTCCCAAGGGAAGAATGA
80 N T E R Q K E Q E K I L Q E L T Q L T D E L T S R I P I S Q G K N E
901 GTCCATGCAGGCGAAGATCACTGAGCAACTGATGCAGCTGAAAAGTGAAGTCTTCCAGGATCCCCTTCCAGGGCAGAATGATCCATACAAGAG
113 S M Q A K I T E Q L M Q L K T E L L S R I P I F Q G Q N E S I Q E
1001 AAGATCTCTGAGCAACTGATGCAGCTGAAGGCTGAAGTCTTTCCAAGATCTCCAGCTCCCGGTAAGGATGATTCTAAGCAGGAGAAGATCTACCAAC
147 K I S E Q L M Q L K A E L L S K I S S F P V K D D S K Q E K I Y Q
1101 AGCTGGTACAGATGAAGACTGAAGTCTTCCGCTGTGTCGACTCTGCCCTGGGACTGGACATTCCTCTAGGAAATGTTACTTCTTCTCAAGTCCCA
180 Q L V Q M K T E L F R L C R L C P W D W T F L L G N C Y F F S K S Q
1201 GCGGAAGTGAAGTACGCGCTCAGACTTCAAAGAAGTGAAGGCTCAACTAGTCAATCAATAGTGAAGAGCAGACCTTCTGAGCAGACTTCT
213 R N W N D A V T A C K E V K A Q L V I I N S D E E Q T F L Q Q T S
1301 AAGGCTAAAGGACCAACTGGATGGCCTGTGACAGCTGAAGAAGGAGGCCACGTGGCTCTGGGTAGATGGTCTACTCTGTCTCAGATCCAGAAAT
247 K A K G P T W M G L S D L K K E A T W L W V D G S T L S S R F Q K
1401 ATTGGAATAGAGGGGAGCCTAACACATCGGTGAGGAAGACTGTGCAATTTGCTGGGATGGTGAAGTACTCTAAATGTGAAGTCAAAAAGTCTG
280 Y W N R G E P N N I G E E D C V E F A G D G W N D S K C E L K K F W

NheI (1535)

1501 GATCTGCAAGAAGTCTGCAACCCATGCACTGAAGGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGACAAACCACAACCTAGAATGCAGTG
313 I C K K S A T P C T E G •
1601 AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTAACAACAACCTGCAATCATT
1701 ATGTTTCAGGTTCAAGGGGAGGTGTGGGAGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTA
1801 ACCTCAAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTGCAGCCTCA
1901 CTTCTTTTATGAGTTAAAGATATAGTGTATTTTCCCAAGTTTGAAGTACTCTTCAATTTCTTTATGTTTTAAATGCAGTACCTCCACATTCCTT
2001 TTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATC
2101 CCCAGTTTAGTAGTTGGACTTAGGGAACAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGG
141 • N R T Y K L
2201 GGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGACATGC
133 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 C M G
2301 CACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGC
100 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 S V A
2401 AGACCCCAATGGCAATGGCTTCAAGCAGACAGTACCTGCAATGTAGGCCTCAATGTGGACAGAGATGATCTCCAGTCTTGGTCTGATGGCC
67 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 I A
2501 GCCCGACATGGTGTGTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCAGCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTTCTCA
33 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 T K M
2601 TGGTGGCCCTCTATAGTGAAGTCTTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCTCTCCAGCTTATCTGAC
0
2701 GGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACAGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTC
2801 CCGTTGATTTACTAGTCAAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCTGAGTCAAACCGCTATCCACGCCATTGATGTAC
2901 TGCCAAAACCGCATCATCATGGAATAGCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGG
3001 CGGGCCATTTACCGTATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTACCGTAAATACTCCACCCA
3101 TTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTTGACGTCAATGGGGGGGGTCTGGGGCGGTGAGCCAGCGGGCCA

3201 TTTACGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTT
3301 TTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCC
3401 CCTGGAAGCTCCCTCGTGCCTCTCTGTTCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCATAGCT
3501 CACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCCTTATCCGGTAA
3601 CTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACA
3701 GAGTTCTTGAAGTGGTGGCCTAACTACGGTACACTAGAAGAAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTA
3801 GCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTCGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTT
3901 GATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATA
4001 AAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTTGTGTAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGC
4101 AAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA