



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

BspHI (560)

Agel (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGTCATCATGAAGATGCACCTGCAAAGGCTCTGGTGGTCTGGC
1 M K M H L Q R A L V V L A
601 CCTGCTGAACCTTGCACGGTCAAGCTCTCTGTCCACTGCACCACCTGGACTTCGGCCACATCAAGAAGAAGAGGGTGAAGCCATTAGGGGACAG
13 L L N F A T V S L S L S T C T T L D F G H I K K K R V E A I R G Q
701 ATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCAACGGTATGACCCACGTCGCCCTATCAGGTCCTGGCCCTTACAACAGCACCCGGGAGCTGC
47 I L S K L R L T S P P E P T V M T H V P Y Q V L A L Y N S T R E L
801 TGGAGGATGCATGGGAGAGGGAGGAAGGCTGCACCCAGGAAACACCGAGTCGGAATACTATGCCAAAGAAATCCATAAATTCGACATGATCCAGGG
80 L E E M H G G E R E E G C T Q E N T E S E Y Y A K E I H K F D M I Q G
901 GCTGGCGGAGCACAACGAACTGGCTGTCTGCCCTAAAGGAATTACCTCCAAGTTTTCCGCTTCAATGTCTCAGTGGAGAAAAATAGAACCAACCTA
113 L A E H N E L A V C P K G I T S K V F R F N V S S V E K N R T N L
1001 TTCCGAGCAGAATCCGGGCTTTCGGGTGCCAACCCAGCTAAGCGAATGAGCAGAGGATCGAGCTCTCCAGATCTTCGGCCAGATGAGCACA
147 F R A E F R V L R V P N P S S K R N E Q R I E L F Q I L R P D E H
1101 TTGCCAAACAGCGCTATATCGGTGGCAAGAATCTGCCACACGGGCACTGCCGAGTGGCTGTCTTTGATGTCACTGACACTGTGCGTGAGTGGCTGTT
180 I A K Q R Y I G G K N L P T R G T A E W L S F D V T D T V R E W L L
1201 GAGAAGAGATCCAACCTAGGCTAGAAATCAGCATTCACTGTCCATGTACACCTTTCCAGCCAAATGAGATATCTGAAAAACATTACGAGGTGATG
213 R R E S N L G L E I S I H C P C H T F Q P N G D I L E N I H E V M
1301 GAAATCAAATCAAAGCGTGGACAATGAGGATGACCATGGCCGTGGAGATCTGGGCGCCTCAAGAAGCAGAAGGATCACCACAACCTCATCTAATCC
247 E I K F K G V D N E D D H G R G D L G R L K K Q K D H H N P H L I

BspHI (1400)

1401 TCATGATGATCCCCACACCGGCTCGACAACCCGGGCCAGGGGGTTCAGAGGAAGAAGCGGGCTTTGGACACCAATTACTGCTTCCGCAACTTGGAGGA
280 L M M I P P H R L D N P G Q G G Q R K K R A L D T N Y C F R N L E E
1501 GAACTGTGTGTCGCCCCCTACATTGACTTCCGACAGGATCTGGCTGGAAGTGGTCCATGAACCTAAGGGCTACTATGCCAACTTCTGCTCAGGC
313 N C C V R P L Y I D F R Q D L G W K W V H E P K G Y Y A N F C S G
1601 CCTTCCCATACCTCCGAGTGCAGACACAACCCACAGCACGGTGTGGGACTGTACAACACTCTGAACCTGAAGCATCTGCCTCGCCTTGTGCGTGC
347 P C P Y L R S A D T T H S T V L G L Y N T L N P E A S A S P C C V
1701 CCCAGACCTGGAGCCCCTGACCATCTGTACTATGTTGGGAGGACCCCAAGTGGAGCAGCTCTCCAACATGGTGGTGAAGTCTTTGAAATGTAGCTG
380 P Q D L E P L T I L Y Y V G R T P K V E Q L S N M V V K S C K C S •

NheI (1809)

1801 AGACCCACCGTAGCTGGCCAGACATGATAAGATACATTGATGAGTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTT
413
1901 GTGATGCTATTGCTTTATTTGTAACCATTATAAGTGAATAAACAAGTTAACAACAACAATTGCATTCATTTATGTTTCAGGTTACAGGGGAGGTGTG
2001 GGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTGAA
2101 TCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTTCATGGAGTTTAAAGATA
2201 GTGTATTTTCCAAGGTTTGAAGTAGCTCTTCATTTCTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTT
2301 AAATACATCATTGCAATGAAAAATAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGG
2401 AACAAAGGAACCTTAAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTT
141 • N R T Y K L P I L E E I T T K
2501 GACCAGCTTGCCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGAT
125 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 C P S V V R I S
2601 CTGTCCACCTCATCAGATAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTTCGCCGCTTGCACAGCAGACCCAAATGGCAATGGCTCAGCAC
91 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 S G I A I A E A C
2701 AGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTGATGGCCGCCCGACATGGTCTTGTGCTCTC
58 V T V 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
2801 ATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACAGCTCCAGATCTGCTGAGAGATGTTGAAGTCTTTCATGGTGGCCCTCTATAGTGAAGTCGTA
25 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
2901 TTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACGCTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATA
3001 TAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAAC
3101 TCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAAT

3201 AGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGGGGCCATTTACCGTCATTGACGTCA
3301 ATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTG
3401 GCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGGGGCCATTTACCGTAAGTTATGTAACGCCTGC
3501 AGGTTAAITAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACG
3601 AGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCC
3701 TGTTCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTG
3801 TAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAA
3901 GACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA
4001 CGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
4101 GCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTC
4201 AGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAATCAGCGGCCCAATAAAATATCTTTATTTTCATTACATCTG
4301 TGTGTTGGTTTTTTTGTGTAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCAGTGAAGTG
4401 CAGGTGCCAGAACATTTCTCTATCGAA