



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

SphI (560)
AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCAGCATGCGATCGGGCACTCACTGGAGAGTTCTGGGCCTCTG
1 M Q S G T H W R V L G L C
601 CCTCTTATCAGTTGGTGTGGGGCAAGATGGTAATGAAGAAATGGTGGTATTACACAGACACCATATAAAGTCTCCATCTCTGGAACCACAGTAATA
13 L L S V G V W G Q D G N E E M G G I T Q T P Y K V S I S G T T V I
701 TTGACATGCCCTCAGTATCCTGGATCTGAAATACTATGGCAACACAATGATAAAAAACATAGGCGGTGATGAGGATGATAAAAAACATAGGCAGTGTAGG
47 L T C P Q Y P G S E I L W Q H N D K N I G G D E D D K N I G S D E
801 ATCACCTGTCACCTGAAGGAATTTTCAGAATTGGAGCAAAGTGGTATTATGTCTGCTACCCAGAGGAAGCAAACAGAAAGATGCGAACTTTATCTCTA
80 D H L S L K E F S E L E Q S G Y Y V C Y P R G S K P E D A N F Y L Y
901 CCTGAGGGCAAGAGTGTGTGAGAACTGCATGGAGATGGATGTGATGTCGGTGGCCACAATTGTCATAGTGGACATCTGCATCACTGGGGGCTTGTGCTG
113 L R A R V C E N C M E M D V M S V A T I V I V D I C I T G G L L L
1001 CTGTTTACTACTGGAGCAAGATAGAAAGGCCAAGGCCAAGCCTGTGACACGAGGAGCGGGTCTGGCGGCAAGGGGACAAAACAAGGAGAGGC
147 L V Y Y W S K N R K A K A K P V T R G A G A G G R Q R G Q N K E R
1101 CACCACCTGTTCCCAACCCAGACTATGAGCCATCCGAAAGGCCAGCGGGACCTGTATTCTGGCCTGAATCAGAGACGCATCTGACCTCTGGAGAAACA
180 P P P V P N P D Y E P I R K G Q R D L Y S G L N Q R R I •

NheI (1207)

1201 CTGCCTCGTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAACAAAATGCTTTATTTGTGAAATTTGT
1301 GATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTAACAACAACAAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTTGGG
1401 AGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATC
1501 CTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGGCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAGATATAGT
1601 GTATTTTCCCAAGGTTTGAAGTACTCTTCATTCTTTATGTTTAAATGCACCTGACCTCCACATTCCTCTTTTAGTAAAATATTCAGAAATAATTTAA
1701 ATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAA
1801 CAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGGAGGGGATGAGTTCCTCAATGGTGGTTTTGA
141 • N R T Y K L P I L E E I T T K V
1901 CCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCT
124 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 R
2001 GTCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTGTCTACAGCAGACCAATGGCAATGGCTTACAGACAG
91 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
2101 ACAGTGACCTGCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTGCTCCTCAT
57 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 Y
2201 AGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGCTTTCATGATGGCCCTCTATAGTGAGTCGTATT
24 L M T I K E T A V E V L E L D Q Q S I N F T K M
2301 ATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATA
2401 GACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGCGGAGTTGTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTC
2501 CCATTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGACTGCCAAAACCGCATCATCATGGTAATAG
2601 CGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCCATAAAGTCAATGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCAATGACGTCAAT
2701 AGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGC
2801 GTTACTATGGAAACATACGTCATTATTGACGTCAATGGGCGGGTGTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCTGCAG
2901 GTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAG
3001 CATCACAATAATCGACGCTCAAGTCAAGGTTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGGCTCTCCTG
3101 TTCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGA

3201 GGTTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTGAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGA
3301 CACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGGTGGCCTAACTACG
3401 GCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGC
3501 TGGTAGCGGTGGTTTTTTTGGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAG
3601 TGGAACGAAAACACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTG
3701 TGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCA
3801 GGTGCCAGAACATTTCTCTATCGAA