



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTCCGCCCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

SphI (566)
BspEI (558)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGCTCCGGACAGCATGCCGAAAGTAGTGTCTCGGTGAGTGTCTG
111 M P K V V S R S V V C
601 CTCTGACACTCGGGACCGGGAGGAATATGACGACGGCGAGAAGCCCTCCATGTTTACTACTGTTTGTGCGGCCAGATGGTCTAGTGTGGACTGCCAG
111 S D T R D R E E Y D D G E K P L H V Y Y C L C G Q M V L V L D C Q
701 TTAGAGAAATTGCCCATGAGGCCCGGGACCGGTCCCGTGTGATTGATGCTGCCAAACATGCCATAAGTTTTGTAACACAGAAGATGAGGAGACTATGT
45 L E K L P M R P R D R S R V I D A A K H A H K F C N T E D E E T M
801 ATCTGCGGAGACTGAAGGCATTGAACGACAGTACAGGAAGAAATGTCAAAGTGTGGACTGCCGCTCTTCTACCAATCCCAGCCAAAGAATGCTCCTGT
78 Y L R R P E G I E R Q Y R K K C A K C G L P L F Y Q S Q P K N A P V
901 TACCTTCATTGTGGATGGAGCAGTAGTCAAGTTTGGCCAGGGCTTTGGGAAAACGAACATATACTCAGAAAACAAGAGCCTCCTAAGAAGGTGATGATG
111 T F I V D G A V V K F G Q G F G K T N I Y T Q K Q E P P K K V M M
1001 ACCAAACGGACCAAAGACATGGGCAAGTTCAGTTCTGTACAGTGTCTACCATTGATGAAGAGGAAGAGGAGATTGAGGCTAGGGAAGTTGCTGACTCAT
145 T K R T K D M G K F S S V T V S T I D E E E E E I E A R E V A D S
1101 ATGCACAGAATGCCAAAGTATTGAAAAACAGCTGGAGCGCAAAGGCATGAGCAAGAGGCGACTGCAAGAGCTGGCTGAATTGGAAGCCAAGAAAGCGAA
178 Y A Q N A K V I E K Q L E R K G M S K R R L Q E L A E L E A K K A K

NheI (1245)

1201 AATGAAGGGACCTTGATTGACAACAGTTCAAATAACAGGCTGTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGACAAACCACAATA
211 M K G T L I D N Q F K •
1301 GAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAATTG
1401 CATTCAATTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAG
1501 CAAAACTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTT
1601 TGCAGCCTCACCTCTTTTCATGGAGTTAAGATATAGTGTATTTCCCAAGGTTTGAAGTACTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCC
1701 ACATTCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCT
1801 TCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGT
1901 GTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTTCATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCT
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
2001 CTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGT
103 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 N
2101 TGCTCACAGCAGACCAATGGCAATGGCTTCCAGCAGACAGTACCCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGT
70 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 I E G T K T
2201 CCTGATGGCCGCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTG
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
2301 AAGTCTTCATGGTGGCCCTCTATAGTGTGATGCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCTCTCCAG
3 F T K M
2401 CTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATT
2501 TTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAAACCGCTATCCACGCC
2601 ATTGATGTAAGTCCAAACCGCATCATCATGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAAGTTCATGTAAGTGGCA
2701 TAATGCCAGGCGGGCATTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTACCGTAAAT
2801 ACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATATTGACGTCAATGGCGGGGGTCTGTTGGCGGTCAGCC
2901 AGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGT
3001 TGCTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAAGGTTGGCGAAACCGACAGGACTATAAAGATACCA
3101 GCGTTTTCCCGTGAAGCTCCCTCGTGCCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTT

3201 TCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTGAGCCGACCGCTGCGCCT
3301 TATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGG
3401 CGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAA
3501 AGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAG
3601 AAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGC
3701 GGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAA
3801 ACAAACTAGCAAAATAGGCTGTCCCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA