



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

SphI (560)
AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCAGCATGCACAGGAGGAGAAGCAGGAGCTGTCGGGAAGATCA
131 M H R R R S R S C R E D Q
601 GAAGCCAGTCATGGATGACCAGCGGACCTTATCTCCAACAATGAGCAACTGCCATGCTGGGCGGCGCCCTGGGGCCCGGAGAGCAAGTGCAGCCGC
131 K P V M D D Q R D L I S N N E Q L P M L G R R P G A P E S K C S R
701 GGAGCCCTGTACACAGGCTTTCCATCCTGGTACTCTGCTCCTCGCTGGCCAGGCCACCACCGCTACTTCTGTACCAGCAGCAGGGCCGCTGGACA
47 G A L Y T G F S I L V T L L L A G Q A T T A Y F L Y Q Q Q G R L D
801 AACTGACAGTCACCTCCAGAACCTGCAGCTGGAGAACCTGCGCATGAAGCTTCCCAAGCCTCCCAAGCCTGTGAGCAAGATGCGCATGGCCACCCCGCT
80 K L T V T S Q N L Q L E N L R M K L P K P P K P V S K M R M A T P L
901 GCTGATGCAGGCGCTGCCATGGGAGCCTGCCAGGGGCCATGCAGAATGCCACCAAGTATGGCAACATGACAGAGGACCATGTGATGCACCTGCTC
113 L M Q A L P M G A L P Q G P M Q N A T K Y G N M T E D H V M H L L
1001 CAGAATGCTGACCCCTGAAGGTGTACCCGCCACTGAAGGGGAGCTTCCGGAGAACCTGAGACACCTTAAGAACACCATGGAGACCATAGACTGGAAGG
147 Q N A D P L K V Y P P L K G S F P E N L R H L K N T M E T I D W K
1101 TCTTTGAGAGCTGGATGCACCATTGGCTCCTGTTTGAATGAGCAGGCACTCCTTGGAGCAAAAGCCACTGACGCTCCACCGAAAGAGTCACTGAACT
180 V F E S W M H H W L L F E M S R H S L E Q K P T D A P P K E S L E L

NheI (1289)

1201 GGAGGACCCGCTTCTGGGCTGGGTGTGACCAAGCAGGATCTGGGCCAGTCCCATGTGAGAGCAGCAGAGGCGGTCTTCAACATCCTGCTAGCTGGCC
213 E D P S S G L G V T K Q D L G P V P M •
1301 AGACATGATAAGATACATTGATGAGTTTGGACAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTT
1401 GTAACCATTATAAGTGAATAAACAAGTTAAACAACAACCAATTGCATTCAATTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGT
1501 AAAACCTCTACAAATGTGGTATGGAATTCTAAATAACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTGAACTCTTTTCTGAGGGATGAAT
1601 AAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTTCATGGAGTTTAAAGATATAGTATTTTCCAAGGTTTG
1701 AACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAA
1801 AATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAG
1901 AAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTCGGTGACTTGGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCT
141 • N R T Y K L P I L E E I T T K V L K G N M E
2001 CAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTA
118 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 D V E D S Y
2101 GGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCCAGCAGACAGTACCCTGCCAATG
85 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 V T V R G I
2201 TAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTGGTCTGATGGCCGCCCGACATGGTCTTGTGCTCATAGAGCATGGTATCTTCT
51 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 L M T I K E
2301 CAGTGGCAGCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCAATGGTGGCCCTCTATAGTGAGTCTATTATACTATGCCGATATACT
18 T A V E V L E L D Q Q S I N F T K M
2401 ATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACAG
2501 CCTACCGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGG
2601 TGGAGACTTGAAATCCCGTGAGTCAAACCGTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGA
2701 TGTAAGTCCAAAGTAGGAAAGTCCATAAGGTATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGCTACTTGGCAT
2801 ATGATACACTTGTACTGCAAGTGGGAGTTTACCCTAAATACTCCACCATGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATAC
2901 GTCATTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATAAGAACATGTG
3001 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGC
3101 TCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATAACAGCGTTTTCCCCTGGAAGCTCCCTGTCGCTCTCTGTTCCGACCCTGCCGCTTA

3201 CCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTGCTCCAAGCT
3301 GGGCTGTGTGCACGAACCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG
3401 GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAG
3501 TATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT
3601 TGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGT
3701 TAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTGTGTGA
3801 ATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTC
3901 TATCGAA