



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

SphI (560)

AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCAGCATGCTTCGAGGATGGGGTGGCCCCAGTGTGGGTGTGTG
1 M L R G W G G P S V G V C
601 TGTGCGCACAGCACTGGGGTGTGTGCCTCTGCCTCACAGGAGCTGTGGAAGTCCAGGTCTCTGAAGACCCCGTGGTGGCCCTGGTGACACGGATGCC
13 V R T A L G V L C L C L T G A V E V Q V S E D P V V A L V D T D A
701 ACCCTACGCTGCTCTTTTCCCGAGAGCTGGCTTCACTGCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGACACCAAACAGCTGGTGACAGCTTCA
47 T L R C S F S P E P G F S L A Q L N L I W Q L T D T K Q L V H S F
801 CGGAGGGCCGGACCAAGGCAGTGCCTACTCCAACCGCACAGCGCTCTCCCTGACCTGTTGGTGAAGGCAATGCCTCTTGAAGCTGCAGCGCTCCG
80 T E G R D Q G S A Y S N R T A L F P D L L V Q G N A S L R L Q R V R
901 AGTAACCGACGAGGGCAGCTACACCTGCTTTGTGAGCATCCAGGACTTTGACAGCGTCTGTTAGCCTGCAGGTGGCCGCCCCCTACTCGAAGCCCAGC
113 V T D E G S Y T C F V S I Q D F D S A A V S L Q V A A P Y S K P S
1001 ATGACCCTGGAGCCCAACAGGACCTACGTCCAGGGAACATGGTGACCATCACGTGCTCTAGTACCAGGGCTATCCGGAGGCGGAGGTGTTCTGGAAGG
147 M T L E P N K D L R P G N M V T I T C S S Y Q G Y P E A E V F W K
1101 ATGGACAGGGAGTGCCCTTACTGGCAATGTGACCACATCCCAGATGGCCAACGAGCGGGGCTTGTTCGATGTTTACAGCGTCTGAGGGTGGTGTGGG
180 D G Q G V P L T G N V T T S Q M A N E R G L F D V H S V L R V V L G
1201 TGCTAACGGCACCTACAGCTGCCTGGTACGCAACCCGGTGTGAGCAAGATGCTCAGGCTCAGTACCATCACAGGCGAGCCCTGACATTCCCCCT
213 A N G T Y S C L V R N P V L Q Q D A H G S V T I T G Q P L T F P P
1301 GAGGCTCTGTGGTAACCGTGGGCTCTCTGTCTGTCTTGTGGTACTACTGGTGGCCCTGGCTTTCGTGTGCTGGAGAAAGATCAAGCAGAGCTGCGAGG
247 E A L W V T V G L S V C L V V L L V A L A F V C W R K I K Q S C E
1401 AGGAGAATGCAGGTGCCGAGGACAGGATGGAGATGGAGAAGATCCAAGACAGCTCTACGGCTCTGAAACCCTCTGAAAACAAAGAAGATGACGGACA
280 E E N A G A E D Q D G D G E G S K T A L R P L K P S E N K E D D G Q

NheI (1538)

1501 AGAAATTGCTTATTGGGAGCTGCTGCCTTCCAGGTCCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGCAAACCACAACCTAGAATGCA
313 E I A •
1601 GTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAT
1701 TTTATGTTTCAGGTTAGGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTTCTAAAATACAGCATAGCAAACCT
1801 TTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCC
1901 TCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTTGAAGTACTCTTCTTTTATGTTTAAATGCACTGACCTCCACATTCC
2001 CTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAAT
2101 ATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTG
141 • N R T Y K
2201 AGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACA
134 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 R C M
2301 TGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGATAGGGGTCCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCAC
101 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 S V
2401 AGCAGACCCAATGGCAATGGCTTCCAGCACAGACAGTACCCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTGATG
68 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 R I
2501 GCCGCCCCGACATGGTGTGTTGTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGGTCT
34 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 F T K
2601 TCATGGTGGCCCTCTATAGTGAGTCTGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCT
1 M
2701 GACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAA
2801 GTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATG
2901 TACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCCATAGGTCATGACTGGGCATAATGCC
3001 AGGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTGACTGCCAAGTGGGCGAGTTTACCGTAAATACTCCAC
3101 CCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTGAGCCAGGCGGG

3201 CCATTTACGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGC
3301 GTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTT
3401 CCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATA
3501 GCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCCTTATCCGG
3601 TAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCT
3701 ACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTG
3801 GTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCCGAGAAAAAAGGATCTCAAGAAGATCC
3901 TTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCA
4001 ATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACT
4101 AGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA