



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

NcoI (560) **AgeI (552)** NcoI (574)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATGGGACGCTGCCATGGCTTCTTGCCTTCTTATTCT
601 GGGTCTCCAGGCTTGGGATACTCCACCATCGTCTCCCGCAAGGAGTGGGGGCAAGACCGCTCGCTGCAGGGCCCTGCTGACCCTGCCTGTGGCCTAC
13▶ G L Q A W D T P T I V S R K E W G A R P L A C R A L L T L P V A Y
701 ATCATCACAGACCAGCTCCAGGGATGCACTGCCAGCAGCAGAGCGTTTGCAGCCAGATGCTCGGGGGTTCAGTCCATTCCGTCTACCCATAGGCT
47▶ I I T D Q L P G M Q C Q Q Q S V C S Q M L R G L Q S H S V Y T I G
801 GGTGCGAGTGGCGTACAACCTTCTGGTGGGATGATGGCAGGTTGATGAAGTGTGGCTGGAACATCCAAGGCTTGCACACCCAGGGCTACAACAA
80▶ W C D V A Y N F L V G D D G R V Y E G V G W N I Q G L H T Q G Y N N
901 CATTCCCTGGGCATCGCTTCTTGGCAATAAGATAGGCAGCAGTCCAGCCCTGCCTTATCAGCTGCAGAGGGTCTGATCTCCTATGCCATCCAG
113▶ I S L G I A F F G N K I G S S P S P A A L S A A E G L I S Y A I Q
1001 AAGGTACCTGTGCCCAGGTATATTAGCCACTTCTTCTGAAAGAAGACCTGCCTGGACCTCAACATCCAGTATGCCAGGAAGTTTGGCCCA
147▶ K G H L S P R Y I Q P L L L K E E T C L D P Q H P V M P R K V C P
1101 ACATCATCAAACGATCTGCTTGGGAAGCCAGAGAGACACACTGCCCTAAAATGAACCTCCAGCCAAATATGTCATCATCCACACCGCTGGCACAAG
180▶ N I I K R S A W E A R E T H C P K M N L P A K Y V I I I H T A G T S
1201 CTGCACTGTATCCACAGACTGCCAGACTGTCGTCCGAAACATACAGTCTTTTACATGGACACACGGAACCTTTGTGACATTGGATATCACTTCTGGTG
213▶ C T V S T D C Q T V V R N I Q S F H M D T R N F C D I G Y H F L V
1301 GGCCAGGATGGTGGCGTGTATGAAGGGTGGATGGCAGTCCAAGGCTCTCACACTTATGGATTCAACGATATTGCCCTAGGAATTGCCCTCATCGGCT
247▶ G Q D G G V Y E G V G W H I Q G S H T Y G F N D I A L G I A F I G
1401 ACTTTGTAGAAAAGCCTCAAATGCTGCAGCGTGGAGCGGCCAGGACCTGATCCAGTGTGCCGTGGTGGAGGGTACCTGACTCAAACCTACCTGCT
280▶ Y F V E K P P N A A A L E A A Q D L I Q C A V V E G Y L T P N Y L L
1501 GATGGGCCACAGTGCAGTGGTCAACATCCTGTCCCCTGGGCGAGCTTTGTATAACATCATCAGCACCTGGCCTCATTCAAGCACTGAAGGAGGCCCCAC
313▶ M G H S D V V N I L S P G Q A L Y N I I S T W P H F K H •

NheI (1618)

1601 TCCCTTTGAGACTGCCCTGTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTAAAAAATGCTTTATTT
1701 GTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAATTTATGTTTCAGGTTTCAGGG
1801 GGAGGTGTGGGAGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACCTTTAACCTCAAATCAAGCCT
1901 CTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTT
2001 TAAGATATAGTGATTTTCCCAAGTTTGAAGTACTTCTCATTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTATGAAAATATTCAG
2101 AAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTG
2201 GACTTAGGGAACAAGAACCTTTAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAAT
2301 GGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAGCAGTCAAGGATGAGTCTGACATGCCACAGGGGCTGACCACC
128▶ 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 G C P S V V
2401 CTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCCAATGGCAATGG
94▶ 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 A S G I A I A
2501 CTTGACACAGACAGTACCCTGCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTGGTCTGATGGCCGCCCCGACATGGTGCTT
61▶ 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 A A G V H H K
2601 GTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGTCTTCTGTTGGCCCTCTATAG
28▶ 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 M
2701 TGAGTCGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCT
2801 CTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTGATTTACTAGTCT
2901 AAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGACTGCCAAAACCGCATCAT
3001 CATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGGGGCCATTTACCGTCA
3101 TTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTACCGTAAATACTCCACCCATTGACGTCAATGAAAG

3201 TCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGT
3301 AACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCC
3401 CCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGT
3501 GCGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTCTCATAGCTCAGCTGTAGGTATCTC
3601 AGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCCGACCGTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCA
3701 ACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTG
3801 GCCTAACTACGGTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA
3901 CAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGT
4001 CTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAATAAAATATCTTTATTTCA
4101 TTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCA
4201 GTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA