



1 GGATCTGCATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATGGCCCGAGCCATGGCCGCGGCTGGCCGCTGCTGCT
1▶ M A R A M A A A W P L L L
601 GGTGGCGTACTGGTGTCTCTGGCCACCCAGGAACCGGGACGTCGTGTCGAGCGCCACCCAGGTGCCCGCTTCTGGGCGACTCCGTGACG
13▶ V A L L V L S W P P P G T G D V V V Q A P T Q V P G F L G D S V T
701 CTGCCCTGCTACCTACAGGTGCCAACATGGAGGTGACGCATGTGTACAGCTGACTTGGGCGCGCATGGTGAATCTGGCAGCATGGCCGCTTCCACC
47▶ L P C Y L Q V P N M E V T H V S Q L T W A R H G E S G S M A V F H
801 AAACGCAGGGCCCGACTATTCGGAGTCCAACCGGCTGGAATTCGTGGCAGCCAGACTGGGCGCGGAGCTGCGGAATGCCTGCTGAGGATGTTCCGGTT
80▶ Q T Q G P S Y S E S K R L E F V A A R L G A E L R N A S L R M F G L
901 GCGCGTAGAGGATGAAGCAACTACCTGCCTGTTTCGTCACGTTCCCGCAGGGCAGGAGCGTGGATATCTGGCTCCGAGTGTGCAAGCCCGAG
113▶ R V E D E G N Y T C L F V T F P Q G S R S V D I W L R V L A K P Q
1001 AACACAGCTGAGGTTTCAAGGTCAGCTCACTGGAGAGCCAGTGCCCATGGCCCGTGCCTCCACAGGGGTCGCCCCAGCCCAAATCACCTGGC
147▶ N T A E V Q K V Q L T G E P V P M A R C V S T G G R P P A Q I T W
1101 ACTCAGACCTGGGCGGGATGCCAATACGAGCCAGGTGCCAGGTTCTGTCTGGCACAGTCACTGTACCAGCCTCTGGATATTGGTCCCTCAAGCCA
180▶ H S D L G G M P N T S Q V P G F L S G T V T V T S L W I L V P S S Q
1201 GGTGGACGGCAAGATGTGACCTGCAAGGTGGAGCAGAGAGCTTGAAGAAGCCTCAGCTGCTGACTGTGAACCTCACCGTACTACCCCGAGAGGTA
213▶ V D G K N V T C K V E H E S F E K P Q L L T V N L T V Y Y P P E V
1301 TCCATCTCTGGCTATGATAACAACCTGGTACCTTGCCAGAATGAGCCACCCTGACCTGCGATGCTCGCAGCAACCCAGAGCCACAGGCTATAATTGGA
247▶ S I S G Y D N N W Y L G Q N E A T L T C D A R S N P E P T G Y N W
1401 GCACGACCATGGGTCCCCTGCCACCCTTGTGTGGCCAGGGCGCCAGCTCCTGATCCGTCCTGTGGACAAACCAATCAACACAACCTTAAATCTGCAA
280▶ S T T M G P L P P F A V A Q G A Q L L I R P V D K P I N T T L I C N
1501 CGTACCAATGCCCTAGGAGCTGCCAGGCAGAATGACCGTCCAGGTCAAAGAGGGACCTCCAGTGAAGCACTCAGGATGTCCCGTAAACGCCATCATC
313▶ V T N A L G A R Q A E L T V Q V K E G P P S E H S G M S R N A I I
1601 TTCTGTTCTGGGAATCCTGGTTTTCTGATCCTGCTGGGATCGGGATTTATTTCTATTGGTCAAATGTTCCCGTGAAGTCTTTGGCACTGTGATC
347▶ F L V L G I L V F L I L L G I G I Y F Y W S K C S R E V L W H C H
1701 TGTGTCCTCGAGTACAGAGCATGCCAGCGCCTCAGTAATGGGATGTCTCCTATTAGCTGTGAGCAGAGAGAACAGCTCTCCAGGATCCACAGAC
380▶ L C P S S T E H A S A S A N G H V S Y S A V S R E N S S S Q D P Q T

NheI (1818)

1801 AGAGGGCACAAAGGTGACAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAAATGCAATGAAAAATGCTTTATT
413▶ E G T R •
1901 GTGAAATTTGTGATGCTATTGCTTTATTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAAATTCATTTCATTTTATGTTTCAGGTTTCAGGG
2001 GGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTTCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCT
2101 CTAATTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTTATGAGGTT
2201 TAAGATATAGTGTATTTTCCCAAGTGTGAAGTACTGCTTTCATTTCTTATGTTTTAAATGCAGTACCTCCACATTCCTTTTATGAAAAATTACG
2301 AAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTG
2401 GACTTAGGAACAAAGGAACCTTTAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGGGGGATGAGTTCCTCAAT
141▶ • N R T Y K L P I L E E I
2501 GGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAAGCAGTCAAGGATGAGTCAAGATGAGTCTCTGCACATGCCACAGGGGCTGACCAC
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
2601 CTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTGTCTCACAGCAGACCAATGGCAATGG
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
2701 CTTGAGCAGACAGTACCTGCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCT
61▶ E A C V T V R G I Y A E I H V A S I E G T K T R I A A G V H H K
2801 GTTGTCCATAGCATGGTGTGATCTTCTCAGTGGCCACCTCCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGCTTTCATGGTGGCCCTCTATAG
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
2901 TGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCT
3001 CTGCTTATATAGACTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTGTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCT
3101 AAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCCAAACCGCATCAT

3201 CATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCA
3301 TTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAG
3401 TCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTATGT
3501 AACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTGGCGTTTTTCCATAGGCTCCGCC
3601 CCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTCCCTGGAAGCTCCCTCGT
3701 GCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTC
3801 AGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCA
3901 ACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTG
4001 GCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA
4101 CAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGT
4201 CTGACGCTCAGTGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAATAAAATATCTTTATTTTCA
4301 TTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCA
4401 GTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA