



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTTCGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

NcoI (560)
AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTACCCATGGCTAGAGGACCTGGCCTCGCGCCGCCACCCTGCG
601 GCTGCCGCTGCTGCTTCTGGTGTTCGGCAGTGACCGGCCACACGGCTGCGCAGGACAACCTGCACGTGCCACCAACAAGATGACCGTGTGCAGCCCC
13▶ L P L L L L V L A A V T G H T A A Q D N C T C P T N K M T V C S P
701 GACGGACCCGAGGCCGCTGTCAGTGCCGCGCACTGGGCTCGGCATGGCGTTGACTGCTCCACGCTGACCTCCAAGTGTCTGCTCAAGGCACGCA
47▶ D G P G G R C Q C R A L G S G M A V D C S T L T S K C L L L K A R
801 TGAGCGCCCCAAGAACGCCGACGCTGGTGAAGCAGTGAGCAGCGCTCGTGGACAACGATGGCCTCTACGACCCCGACTGCGACCCCGAGGGCCG
80▶ M S A P K N A R T L V R P S E H A L V D N D G L Y D P D C D P E G R
901 CTTCAAGGCAGCCAGTGCAACCAGACGAGTGTGTGGTGGTGAAGTGGGCTGGCGCCGACGACAAGGGCGACCTGAGCCTACGCTCGCAT
113▶ F K A R Q C N Q T S V C W C V N S V G V R R T D K G D L S L R C D
1001 GAGCTGGTGGCACCACCATCTCATTGACCTGCGCCACCGCCCGCTGGCGCTTCAACCACTCAGACCTGGACGCGGAGCTGAGGCGGCTCT
147▶ E L V R T H H I L I D L R H R P T A G A F N H S D L D A E L R R L
1101 TCCGCGAGCGTATCGGTGCACCCCAAGTTCGTGGCGCTGTGACTACGAGCAGCCACCATCCAGATCGAGCTGGCGAGAACACGTCTCAGAAGGC
180▶ F R E R Y R L H P K F V A A V H Y E Q P T I Q I E L R Q N T S Q K A
1201 TGCCGGTACGTGGATATCGGCGATGCCGCTACTACTTCGAGAGGGACATCAAGGGCGAGTCTCTATTCCAGGGCCGCGGCGCCTGGACTTGC
213▶ A G D V D I G D A A Y Y F E R D I K G E S L F Q G R G G L D L R V
1301 CGCGAGAACCACTGCAGGTGGAGCGCACGCTCATCTATTACCTGGACGAGATVCCCCGAAGTTCCTCATGAAGCGCCTACCGCCGCTCATCGCG
247▶ R G E P L Q V E R T L I Y Y L D E I P P K F S M K R L T A G L I A
1401 TCATCGTGGTGGTGTGGTGGCCCTCGTGCCTGGCATGCCGCTCGTGGTATACCAACCGGAGAAGTGGGGAAGTACAAGAAGTGGAGATCAAGGA
280▶ V I V V V V V A L V A G M A V L V I T N R R K S G K Y K K V E I K E

NheI (1541)
Acc65I (1533)

1501 ACTGGGGGAGTTGAGAAAGAACCGAGCTTGTAGGTACCCGGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAGTAACTAGAA
313▶ L G E L R K E P S L •
1601 GCAGTGAAAAAATGCTTTATTTGTAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATT
1701 CATTTTATGTTTCAGGTTTCAAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAA
1801 ACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGA
1901 GCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTCTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCACAT
2001 TCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCAT
2101 AATATCCCCAGTTTAGTGTGGACTTAGGGAACAAAGAACCTTTAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTAC
141▶ • N R T Y
2201 TTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTTCATCTCAATGAGCACAAGCAGTCAAGGAGCATAGTCAAGATGAGCTCTCTGC
135▶ 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 R C
2301 ACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCT
102▶ 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 S
2401 CACAGCAGCCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTG
69▶ 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 R
2501 ATGGCCGCCCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGG
35▶ 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 F T
2601 TCTTCATGGTGGCCCTCTATAGTGAAGTCTTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTA
2▶ K M
2701 TCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGG
2801 AAAGTCCCCTGTTATTTACTAGTCAAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCCTATCCACGCCATTG
2901 ATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCAAGTAGGAAAGTCCCATAAAGTTCATGACTGGGCATAAT
3001 GCCAGGCGGGCATTACCGTCAATGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCAGTTTACCGTAAATACTC

3101 CACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGC
3201 GGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCT
3301 GGCCTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCG
3401 TTTCCCCCTGGAAGTCCCTCGTGCGCTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTC
3501 ATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCCTGCGCCTTATC
3601 CGGTAACATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGT
3701 GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAG
3801 TTGGTAGCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGTTTTTTTGTTCGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGA
3901 TCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCC
4001 GCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAA
4101 ACTAGCAAAATAGGCTGTCCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA