



100

1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTCAACTCTACGCTTTGTTTCGTTT

BstEII (555)
AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCACCATGGCTTATGTTACTTGGCTTTTGGCTATTCTTCATGT
1 M A Y V T W L L A I L H V
601 GCACAAAGCACTGTGTGAAGAGACATTGTGGGACACAACAGTTCGGCTTTCTGAGACTATGACTCTGGAATGTGTATATCCATTGACGCATAACTTAAAC
13 H K A L C E E T L W D T T V R L S E T M T L E C V Y P L T H N L T
701 CAGGTGGAGTGGACCAAGAACACTGGCACAAGACAGTGGACATAGCAGTTTACAACCTAACCATAATATGCATATAGAATCTAACTACCTCCATAGAG
47 Q V E W T K N T G T K T V S I A V Y N P N H N M H I E S N Y L H R
801 TACATTCCTAAACTCAACAGTGGGTTCCGCAACATGAGCCTTTCTTTTACAATGCCTCAGAAGCAGACATTGGCATCTACTCTGCTTGTTCATGC
80 V H F L N S T V G F R N M S L S F Y N A S E A D I G I Y S C L F H A
901 TTTCCCAATGGACCTTGGGAAAAGAAGATAAAAGTAGTCTGGTCAGATAGTTTTGAGATAGCAGCACCCTCGGATAGCTACCTGTCTGCAGAACCTGGA
113 F P N G P W E K K I K V V W S D S F E I A A P S D S Y L S A E P G
1001 CAAGATGCACACTCACTTCCAGCTTCCAAGGACTTGGCCAGTGAACAAGTGCATATGGGAAAAAGTCCAGCCCATCAGGTAGACATCTTAGCTTCTCT
147 Q D V T L T C Q L P R T W P V Q Q V I W E K V Q P H Q V D I L A S
1101 GTAACCTATCTCAAGAGACAAGATACACTTCAAAGTACCTAAGACAAAACAGGAGCAACTGTAGCCAGGGGAGCATGAAGAGCATCCTCATATTCCAAA
180 C N L S Q E T R Y T S K Y L R Q T R S N C S Q G S M K S I L I I P N
1201 TGCCATGGCCGCTGACTCAGGACTTTACAGATGTCGCTCAGAGGCCATTACAGGAAAAACAAGTCCCTTTGTATAAGGCTGATCATAACTGATGGTGA
213 A M A A D S G L Y R C R S E A I T G K N K S F V I R L I I T D G G
1301 ACCAATAAACATTTTATCCTTCCCATCGTTGGAGGGTAGTTTCACTGTTACTTGTATCCTAATTATCATATTTTATTTATATAACAGGAAGAGAC
247 T N K H F I L P I V G G L V S L L L V I L I I I I F I L Y N R K R
1401 GGAGACAGGTGAGAATCCACTTAAAGAGCCAGGATAAACAGAGTAAGGTAGCCACCAACTGCAGAAGTCTACTTCTCCCATCCAGTCTACAGATGA
280 R R Q V R I P L K E P R D K Q S K V A T N C R S P T S P I Q S T D D

NheI (1579)

1501 TGAAAAAGAGGACATTTATGTTAACTATCCAACCTTCTCTCGAAGACAAAACCAAGACTCTAAGCTGCTCTTTTGGCCGCTAGCTGGCCAGACATGATA
313 E K E D I Y V N Y P T F S R R P K P R L •
1601 AGATACATTGATGAGTTTGGACAAACCACAACACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTA
1701 TAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAATTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTA
1801 CAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGG
1901 CATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTTGAAGTACTCT
2001 TCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTT
2101 TTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGGAACCTTTAATAGAAATTGGACA
2201 GCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTCCATTATCTCAATGAGCAC
141 • N R T Y K L P I L E E I T T K V L K G N M E I L V
2301 AAAGCAGTCAGGAGCATAGTCAGAGATGAGTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCCTG
115 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 P H R
2401 ACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAA
81 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 Y A E I
2501 TGTGGACAGCAGAGATGATCTCCAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTGTCCTCATAGAGCATGGTATCTTCTCAGTGGCGAC
48 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 T A V
2601 CTCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTTCATGGTGGCCCTCTATAGTGAGTCTATTACTATGCGGATATACTATGCCGATGA
15 E V L E L D Q Q S I N F T K M
2701 TTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACTCCACCGTACACGCCTACCGCCC
2801 ATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTG
2901 GAAATCCCGTGAGTCAAACCGCTATCCAGCCATTGATGTACTGCCAAAACCGCATCATATGTAATAGCGATGACTAATACGTAGATGTACTGCCA
3001 AGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCATTACCCTGATTGACGTCAATAGGGGGGCTACTTGGCATATGATACACT
3101 TGATGACTGCCAAGTGGGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATTATG

3201 ACGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAAITTAAGAACATGTGAGCAAAGGC
3301 CAGCAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGA
3401 GGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCGACCCTGCCGCTTACCGGATACCT
3501 GTCCGCCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTG
3601 CACGAACCCCGTTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCA
3701 CTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTAT
3801 CTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTGTTTGAAG
3901 CAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACACTCACGTTAAGGGATTT
4001 TGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTA
4101 ACATACGCTCTCCATCAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA