



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTTCGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

BspHI (560)

AgeI (552)

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGTCATCATGAGCTTTCCATGTAATTTGTAGCCAGCTTCCTTCT
1 M S F P C K F V A S F L L
601 GATTTTCAATGTTTCTTCCAAGGTGACGTCCTCCAAGAGATTACGAATGCCTTGAAACCTGGGGTGCCTTGGGTGACGACATCACTTGGACATTCCT
13 I F N V S S K G A V S K E I T N A L E T W G A L G Q D I N L D I P
701 AGTTTCAATGAGTGATGATATTGACGATATAAAATGGGAAAAACCTCAGACAAGAAAAAGATTGCACAATTGAGAAAAGAGAAAGAGACTTTCAAGG
47 S F Q M S D D I D D I K W E K T S D K K K I A Q F R K E K E T F K
801 AAAAGATACATATAAGCTATTTAAAAATGAACTCTGAAAATTAAGCATCTGAAGACCGATGATCAGGATATCTACAAGGTATCAATATATGATACAAA
80 E K D T Y K L F K N G T L K I K H L K T D D Q D I Y K V S I Y D T K
901 AGGAAAAATGTGTTGGAAAAATTTGATTTGAAGATTCAAGAGAGGGTCTCAAAACCAAGATCTCCTGGACTTGATCAACACAACCTGACCTGT
113 G K N V L E K I F D L K I Q E R V S K P K I S W T C I N T T L T C
1001 GAGGTAATGAATGAACTGACCCGAATTAACCTGTATCAAGATGGGAAACATCTAAAACCTTCTCAGAGGGTCATCACACAAGTGGACCACAGCC
147 E V M N G T D P E L N L Y Q D G K H L K L S Q R V I T H K W T T S
1101 TGAGTGCAAAATCAAGTGCACAGCAGGGAACAAAGTCAAGCAAGGAATCCAGTGTGAGCCTGTGAGTGTCCAGAGAAAGGTCTGGACATCTATCTCAT
180 L S A K F K C T A G N K V S K E S S V E P V S C P E K G L D I Y L I
1201 CATTGGCATATGTGGAGGAGGCAGCCTCTGATGGTCTTTGTGGCACTGCTCGTTTTCTATATCACAAAAGGAAAAACAGAGGAGTCGGAGAAATGAT
213 I G I C G G G S L L M V F V A L L V F Y I T K R K K Q R S R R N D
1301 GAGGAGCTGGAGACAAGAGCCACAGTAGTACTGAAGAAAAGGGCCGGAAGCCCAAAATTCAGCTTCAACCCCTCAGAATCCAGCAACTCC
247 E E L E T R A H R V A T E E R G R K P H Q I P A S T P Q N P A T S
1401 AACATCTCTCCACCACCTGGTTCATCGTTCCAGGCACCTAGTTCATGTCCTCCCGCCTCCTGGACACCGTGTTCAGCACCAGCCTCAGAAGAGGCCTCC
280 Q H P P P P P G H R S Q A P S H R P P P P G H R V Q H Q P Q K R P P
1501 TGCTCCGTCGGGCACACAAGTTCACCAGCAGAAAGGCCCGCCCTCCCGAGCCTCGAGTTACGCCAAAACCTCCCATGGGGCAGCAGAAAACCTATTG
313 A P S G T Q V H Q Q K G P P L P R P R V Q P K P P H G A A E N S L

NheI (1678)

1601 TCCCCTCTCTAATTAATAAAGATAGAACTGTCTTTTTCAATAAAAAGCACTGTGGATTTCTGCCCTCTGATGTGGCTAGCTGGCAGACATGATAA
347 S P S S N •
1701 GATACATTGATGAGTTTGGACAAACCACAACACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTAT
1801 AAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGTTTTTAAAGCAAGTAAAACCTCTAC
1901 AAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGC
2001 ATCAGGGGCTGTTGCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTGCTT
2101 CATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTT
2201 TTTATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTAAATAGAAATTTGGACAG
2301 CAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACA
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
2401 AAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGA
114 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 V
2501 CAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCCAGCAGACAGTACCCTGCCAATGTAGGCCTCAAT
81 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
2601 GTGGACAGCAGAGATGATCTCCCGAGTCTGGTCTGATGGCCGCCCGACATGGTGTCTGTTGTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACC
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
2701 TCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGGCTTTCATGGTGGCCCTCTATAGTGAAGTCTATTATACTATGCCGATATACTATGCCGATGAT
14 E V L E L D Q Q S I N F T K M
2801 TAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACTCCACCGTACACGCCTACCGCCA
2901 TTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGG
3001 AAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGACTGCCAAAACCGCATCATCATGTAATAGCGATGACTAATACGTAGATGACTGCCAA
3101 GTAGGAAAGTCCCATAAGGTCATGACTGGGCATAATGCCAGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGGCTACTTGGCATATGATACTT

3201 GATGTA CTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGA
3301 CGTCAATGGGCGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAAGAACATGTGAGCAAAAGGCC
3401 AGCAAAAGGCCAGGAACCGTAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAG
3501 GTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAAGTCCCTCGTGGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTG
3601 TCCGCCTTCTCCCTTCGGGAAGCGTGGCGTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGC
3701 ACGAACCCCGTTTACGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCAC
3801 TGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATC
3901 TGGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGC
4001 AGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTT
4101 GGTGATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAA
4201 CATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAAGTCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA