



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTTCGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

BspHI (560)

AgeI (552)

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGTCATCAGACGCGCGGCTCTGGCTCTGGCTCTGCGCGCT
601 GGCCTGCGCCTGCAGCCTGCCCTCCGCAAATTGTGGCTGTAATGTTCTCTGAAGATCAGGATGGCTCTGGGGATGACTCTGACAACTCTCTGGC
13▶ A L R L Q P A L P Q I V A V N V P P E D Q D G S G D D S D N F S G
701 TCTGGCACAGGTCTTTGCCAGATACTTGTACGGCAGACACCTTCCACTTGAAGGACGTGTGGCTGTTGACAGCCAGCCACAGCTCCAGAGCCCA
47▶ S G T G A L P D T L S R Q T P S T W K D V W L L T A T P T A P E P
801 CCAGCAGCAACCGAGACTGCTTTTACCTCTGCTGCCAGCCGAGAGAAGCCCGAGGAGGAGAGCCTGTGCTCCATGTAGAAGCAGAGCCTGGCTT
80▶ T S S N T E T A F T S V L P A G E K P E E G E P V L H V E A E P G F
901 CACTGCTCGGACAAGGAAAAGGAGTCAACCAGGCCAGGAGACCGTGCAGTCCCATCACCAACGGGCTCAACAGTCAGAGTCACCACAGCC
113▶ T A R D K E K E V T T R P R E T V Q L P I T Q R A S T V R V T T A
1001 CAGGCAGTGTACATCTCATCCGACGGGGCATGCAACCTGGCCTCCATGAGACCTCGGCTCCACAGCACCTGGTCAACTGACCATCAGCCTCCAC
147▶ Q A A V T S H P H G G M Q P G L H E T S A P T A P G Q P D H Q P P
1101 GTGTGGAGGGTGGCGCACTTCTGTCATCAAAGAGTTGTGAGGATGGAAGTCCCAATCAGCTTCCCGCAGGAGAGGGCTCTGGAGAACAAGACTTCAC
180▶ R V E G G G T S V I K E V V E D G T A N Q L P A G E G S G E Q D F T
1201 CTTTGAACATCTGGGAGAACACAGCTGTGGCTGCCGTAGAGCCCGCTGCGGAATCAGCCCCGGTGGACGAAGGAGCCACAGGTGCTTCTCAGAGC
213▶ F E T S G E N T A V A A V E P G L R N Q P P V D E G A T G A S Q S
1301 CTTTTGGACAGGAAGGAGTGTGGAGGTGTATTGCCGAGGCCTAGTGGGCTCATCTTTGCTGTGCTGGTGGCTTTCATGCTGTACCGGATGA
247▶ L L D R K E V L G G V I A G G L V G L I F A V C L V A F M L Y R M

NheI (1498)

1401 AGAAGAAGGACGAAGGCAGTACTCCTTGAGGAGCCAAACAAGCCAATGGCGGTGCTACCAGAAACCCACCAAGCAGGAGGAGTTCTACGCTGAGC
280▶ K K K D E G S Y S L E E P K Q A N G G A Y Q K P T K Q E E F Y A •
1501 TAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGACAAACCAACTAGAATGCAGTGAAGGAAATGCTTTATTTGTGAAATTTGTGATGCTATT
1601 GCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTT
1701 AAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCATAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGA
1801 GGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCC
1901 CAAGTTTGAAGTACTCTTCTTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCAT
2001 TGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGAACAAAGGAAC
2101 CTTTAATAGAAATTTGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGC
141◀ • N R T Y K L P I L E E I T T K V L K G
2201 CATTATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGTCATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTC
121◀ 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 R I S R D V E
2301 ATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTTCAGCACAGACAGTACC
88◀ 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 E A C V T V
2401 CTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTGTTGCTCCTATAGAGCATGG
54◀ 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 N D E Y L M T
2501 TGATCTTCTCAGTGGCGACTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGGCTTTCATGGTGGCCCTCTATAGTGAAGTCTATTATACTATGC
21◀ I K E T A V E V L E L D Q Q S I N F T K M
2601 CGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACTCCCA
2701 CCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGTATTACTAGTCAAACAAACTCCCATTGACG
2801 TCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGTAATAGCGATGACTA
2901 ATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGT
3001 ACTTGGCATATGATACACTTGATGACTGCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATG
3101 GGAACATACGTCAATTATTGACGTCAATGGGGGGGCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTA

3201 GAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAA
←
3301 AATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCC
3401 TGCCGTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCG
3501 CTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTA
3601 TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTA
3701 GAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGG
3801 TGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAA
3901 AACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTT
4001 TTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGA
4101 ACATTTCTCTATCGAA