



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTCCGCCCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BspLU111 (560)

AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCTACTCTGAGATCACCGGTCAACATGTCGCCGCTCTATGCTGCTTGGCTGCTCTCC
601 CAGCTCCTTCGACTCGGAGCGGCTCAGGAGACAGAAGACCCGGCCTGCTGCAGCCCCATAGTCCCCGGAACGAGTGAAGGCCCTGGCATCAGAGTGC
13▶ S L L R L G A A Q E T E D P A C C S P I V P R N E W K A L A S E C
701 GCCCAGCACCTGAGCCTGCCCTTACGCTATGTGGTGGTATCGCACACGGCGGGCAGCAGCTGCAACACCCCCGCTCGTGCCAGCAGCAGGCCCGAATG
47▶ A Q H L S L P L R Y V V V S H T A G S S C N T P A S C Q Q Q A R N
801 TGCAGCACTACCACATGAAGACACTGGGCTGGTGCACGTGGGTACAACCTCCTGATTGGAGAAGACGGGCTCGTATACGAGGGCCGTGGCTGGAACCT
80▶ V Q H Y H M K T L G W C D V G Y N F L I G E D G L V Y E G R G W N F
901 CACGGGTGCCACTCAGGTCCTTATGGAACCCATGTCCATTGGCATCAGCTTCATGGCAACTACATGGATCGGGTCCCCACACCCAGGCCATCCGG
113▶ T G A H S G H L W N P M S I G I S F M G N Y M D R V P T P Q A I R
1001 GCAGCCAGGGTCTACTGGCTCGGTGTGGCTCAGGGAGCCCTGAGTCCAACCTATGTCTCAAAGACACCGGGATGTGCAGCGTACACTCTCTCCAG
147▶ A A Q G L L A C G V A Q G A L R S N Y V L K G H R D V Q R T L S P

NheI (1167)

1101 GCAACCAGCTCTACCACCTCATCCAGAATTGCCACACTACCGCTCCCTGAGGCCCTGCTGATCCGCTAGCTGGCCAGACATGATAAGATACATTGAT
180▶ G N Q L Y H L I Q N W P H Y R S P •
1201 GAGTTTGGACAAACCACTAGAATGCAGTAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATA

1301 AACAAAGTTAAACAACAACAAATTGCATTCATTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTAT
1401 GGAATTCTAAAATACAGCATAGCAAACTTTAACTCCAAATCAAGCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTG
1501 TTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAACAGCTCTTCAATTTCTTTAT
1601 GTTTTAAATGCACCTGACCTCCACATTCCTTTTATGATAAATATTAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCA
1701 GAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGA

1801 GCTTCTAGCTTTAGTTCCTGGTACTTGGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCAATGAGCACAAAGCAGTCAGG
141▶ • N R T Y K L P I L G E I T T K V L K G N M E I L V F C D P
1901 AGCATAGTCAGAGATGAGCTCTGACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATG
111▶ 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 A V I
2001 GTGTCAAAGTCTTCTGCCGTTGCTCAGCAGACCAATGGCAATGGCTTCCAGCACAGACAGTACCCTGCAATGTAGGCTCAATGTGGACAGCAG
77▶ 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 H V A S
2101 AGATGATCTCCAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTC
44▶ 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 E V L E
2201 CAGATCCTGCTGAGAGATGTTGAAGTCTTCAATGGTGGCCCTCTATAGTGGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAA
11▶ L D Q Q S I N F T K M
2301 ACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTGCGTCAAT

2401 GGGCGGAGTTGTTACGACATTTGGAAAGTCCGTTGATTTACTAGTCAAAACAACCTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCGGTG
2501 AGTCAAACCGCTATCCACGCCATTGATGTAAGTCCGTTGATTTACTGCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCAGT
2601 CCATAAGGTCATGTAAGTGGCATAATGCCAGGCGGGCATTACCGTCAATGACGTCAATAGGGGGGCTACTTGGCATATGATACACTTATGTAAGTCC
2701 AAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGAAACATACGTCAATATTGACGTCAATGGGC

BspLU111 (2871)

2801 GGGGTCGTTGGGCGGTGAGCCAGGCGGCCATTTACCCTAAGTTATGTAACGCTGACAGTTAATAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCC
2901 AGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAAGTCAAGGTTGGCAAAAC
3001 CGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCTTTCT
3101 CCCTTCGGAAGCGTGGCGTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCAGCAACCCCC

3201 GTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA
3301 TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCT
3401 GAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACG
3501 CGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGGCTA
3601 GTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGGAATCGTAACTAACATACGCTCTC
3701 CATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA