



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BspLU111 (560)

AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCAACATGTTGGGGACTGTTGTTGATGCTCGCGGTCCGCTT
1 M L G T L V W M L A V G F
601 CCTGCTGGCACTGGCGCCGGCCGCGCGGGCGCGTGTGAGGACCGGGAGGCGCCCGCGCGGGGACTGCGCGGACCGCCGGAGGAGCTCCTG
13 L L A L A P G R A A G A L R T G R R P A R P R D C A D R P E E L L
701 GAGCAGCTGTACGGCGGCTGGCGGCCGGCTGCTCAGCGCTTCCACCACACGCTGCAGCTCGGGCCGCGGAGCAGGCGCAATGCCAGCTGCCCGG
47 E Q L Y G R L A A G V L S A F H H T L Q L G P R E Q A R N A S C P
801 CCGGGGCGAGGGCCGCGGACCGCGCTTCCGGCCACCACCACTGCGCAGCGTGTCCGCTGGGCGTACAGGATTTCTACGACCTGCTCGCTTTCC
80 A G G R A A D R R F R P P T N L R S V S P W A Y R I S Y D P A R F P
901 GAGGTACCTGCCGAAGCTACTGCCTGTGCCGAGGCTGCCTGACCGGCTCTACGGGAGGAGGACTTCCGCTTTCGAGCACACCGCTTCTCTCCA
113 R Y L P E A Y C L C R G C L T G L Y G E E D F R F R S T P V F S P
1001 GCCGTGGTGTGCGGCGCACAGCGGCTGCGCGGGCGCGCTGTGTACGCGGAACACTACATCACCATCCCGGTGGGCTGCACCTGCGTCCCGAGC
147 A V V L R R T A A C A G G R S V Y A E H Y I T I P V G C T C V P E

NheI (1194)

1101 CGGACAAGTCCGCGGACAGTGCGAATCCAGCATGGACAAGTGTGCTGGGGCCCGGACAGGCTGCGGGGCGTGTGCCGGGACTGCCGGGACTGCCGCTAGC
180 P D K S A D S A N S S M D K L L L G P A D R P A G R •
1201 TGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCAACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTT

1301 TATTTGTAACCATTATAAGTGAATAAACAAGTTAACAACAACAATTGCATTATTTATGTTTCAGGTTTCCAGGGGAGGTGTGGGAGGTTTTTAAAG
1401 CAAGTAAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTGAACTCTTTCTGAGGGA
1501 TGAATAAGGCATAGGCATCAGGGGCTTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTCATGGAGTTTAAAGATATAGTATTTTCCCAAG
1601 GTTTGAACTAGCTCTTCATTTCTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTAGAAATAATTTAAATACATCATTGCA
1701 ATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTT

1801 AATAGAAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATT
141 • N R T Y K L P I L E E I T T K V L K G N
1901 CATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTGCAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCA
120 M E I L V F C D P A Y D S I L E R C M G C P S V R I S R D V E D
2001 GAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGC
86 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 R G
2101 CAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTGGTCTGATGGCCGCCGACATGGTCTTGTGCTCATAGAGCATGGTGTAT
53 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 I
2201 CTTCTCAGTGGGACCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGCTGGCCCTCTATAGTGAAGTCTATTACTATGCGCAT
20 K E T A V E V L E L D Q Q S I N F T K M
2301 ATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGT

2401 ACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAA
2501 TGGGGTGGAGACTTGGAAATCCCGTGAGTCAAACCGCTATCCAGCCATTGATGTACTGCCAAAACCGCATCATATGGAATAGCGATGACTAATAC
2601 GTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTGATGACTGGGCATAATGCCAGGCGGGCATTACCCTGATTGACGTCAATAGGGGGGCTACTT
2701 GGCATATGATACTTGTACTGCTCAAGTGGGAGTACCCTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAA

BspLU111 (2898)

2801 CATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGTCCAGCCAGGCGGCCATTTACCCTAAGTTATGTAACGCCCTGCAGGTTAATTAAGAAC
2901 ATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATC
3001 GACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTTCCCCTGGAAGCTCCCTGTCGCTCTCTGTTCCGACCCTGCC
3101 GCTTACCAGTACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCC

3201 AAGCTGGGCTGTGTGCACGAACCCCGTTAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGC
3301 CACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAG
3401 AACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGT
3501 TTTTTTGTGGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAA
3601 CACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAATCAGCGGCCCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTG
3701 TGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACAT
3801 TTCTCTATCGAA