



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGTGCTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

BspLU11I (560)

Agel (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACTCTGAGATCACCGGTCAACATGTACAGGATGCAACTCCTGCTTGCATTGCACTAAG  
130 M Y R M Q L L S C I A L S  
601 TCTTGCACTTGTACGAATTCGGTGTATCTCTCAGAGTGAAGACTGGGAATGAAAGAACTACAGAGGGACGATGCCAAAACAAAAATGGCATCACCC  
130 L A L V T N 19 V Y L S E C K T G N G K N Y R G T M S K T K N G I T  
701 TGCAAAAATGGAGTTCACCTTCTCCACAGACCTAGATTCTCACCTGCTACACACCCCTCAGAGGGACTGGAGGAGAACTACTGCAGGAATCCAGACA  
270 C Q K W S S T S P H R P R F S P A T H P S E G L E E N Y C R N P D  
801 ACGATCCGAGGGGCCCTGGTGTATACTACTGATCCAGAAAAGAGATGACTACTGCGACATCTTGTAGTGTGAAGAGGAATGTATGCATTGCAGTGG  
600 N D P Q G P W C Y T T D P E K R Y D Y C D I L E C E E E C M H C S G  
901 AGAAAATGATGACGGCAAAATTTCCAAGACCATGTCTGGACTGGAATGCCAGGCTGGGACTCTCAGAGCCCACAGCTCATGGATACATTCTTCCAAA  
930 E N Y D G K I S K T M S G L E C Q A W D S Q S P H A H G Y I P S K  
1001 TTTCCAAACAAGAACCTGAAGAAGAATTACTGTCGTAACCCCGATAGGAGCTGCGGCCTTGGTGTTCACCACCGACCCCAACAGCGCTGGGAACCTT  
1270 F P N K N L K K N Y C R N P D R E L R P W C F T T D P N K R W E L  
1101 GTGACATCCCCGCTGCACAACCTCCACCATCTTCTGGTCCACCTACCAGTGTCTGAAGGGAACAGGTGAAAATATCGCGGGAATGTGGCTGTAC  
1600 C D I P R C T T P P P S S G P T Y Q C L K G T G E N Y R G N V A V T  
1201 CGTGTCCGGGCACACCTGTGAGACTGGAGTGCACAGACCCCTCACACACATAACAGGACACCGGAAAATTTCCCTGCAAAAATTTGGATGAAAATAC  
1930 V S G H T C Q H W S A Q T P H T H N R T P E N F P C K N L D E N Y  
1301 TGCCGCAATCCTGACGAAAAAGGGCCCATGGTGCCATACAACCAACAGCCAAGTGCAGTGGGAGTACTGTAAGATACCGTCTGTGACTCTCCCCAG  
2270 C R N P D G K R A P W C H T T N S Q V R W E Y C K I P S C D S S P  
1401 TATCCACGGAACAATTGGCTCCACAGCACCACCTGAGCTAACCCCTGTGGTCCAGGACTGCTACCATGGTGTGACAGAGTACCAGGCACATCCTC  
2600 V S T E Q L A P T A P P E L T P V V Q D C Y H G D G Q S Y R G T S S  
1501 CACCACCACAGGAAAGAAGTGTGAGTCTTGGTCTATGACACCACACCGGCACAGAAAGCCAGAAAATACCCAAATGCTGGCCTGACAAATG  
2930 T T T T G K K C Q S W S S M T P H R H Q K T P E N Y P N A G L T M  
1601 AACTACTCAGGAATCCAGATGCCGATAAAGGCCCTGGTGTTTTACCACAGACCCAGCGTCAAGTGGGAGTACTGCAACCTGAAAAAATGCTCAGGAA  
3270 N Y C R N P D A D K G P W C F T T D P S V R W E Y C N L K K C S G

NheI (1720)

1701 CAGAAGCGAGTGTGTATAGGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTAT  
360 T E A S V V •  
1801 TTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAATTGCATTCAATTTATGTTTCAGGTTACAG  
1901 GGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGC  
2001 CTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAG  
2101 TTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTC  
2201 AGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGT  
2301 TGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCA  
141 • N R T Y K L P I L E E  
2401 ATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAAGCAGTCAAGGAGCATAGTCAAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCA  
128 I T T K V L K G 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  
2501 CCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAAATGGCAAT  
95 R I S R D V E 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  
2601 GGCTTCAGCACAGACAGTGCACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCACTTGGTCTGATGGCCGCCCGACATGGTGC  
62 A E A C V T V R G I Y A E I H V A S I E G T K T R I A A G V H H  
2701 TTGTTGCTCATAGAGCATGGTGTCTCTCAGTGGCCAGCTCCACAGCTCCAGACTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTCAT  
28 K N D E Y L M T I K E T A V E V L E L D Q Q S I N F T K M  
2801 AGTGAGTCGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAG  
2901 CTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAG  
3001 TCAAAAACAACTCCCATTTGACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATC  
3101 ATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAAGTTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGT

3201 CATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAA  
3301 AGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTAT

---

3401 GTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTGGCGTTTTTCCATAGGCTCCG  
3501 CCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTC  
3601 GTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATC  
3701 TCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCAGCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC  
3801 CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGG  
3901 TGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCA  
4001 AACAAACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGG  
4101 GTCTGACGCTCAGTGAACGAAAACACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGCCGCAATAAAATATCTTTATTTT  
4201 CATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCC  
4301 CAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA