



3301 ATCAAACAATATGGTGATTGCCTTGGTGATATTGCTGCTAGAGACCTCATTGTGCACAAAAGTTTAAACGGCCTTACTGTTTTGCCACCTTTGCTCACAG
834▶ I K Q Y G D C L G D I A A R D L I C A Q K F N G L T V L P P L L T
3401 ATGAAATGATTGCTCAATACACTTCTGCACTGTTAGCGGGTACAATCACTTCTGGTTGGACCTTTGGTGCAGGTGCTGCATTACAAATACCATTTGCTAT
867▶ D E M I A Q Y T S A L L A G T I T S G W T F G A G A A L Q I P F A M
3501 GCAAAATGGCTTATAGGTTAATGGTATTGGAGTTACACAGAATGTTCTCTATGAGAACCAAAATGATTGCCAACCAATTTAATAGTGCTATTGGCAAA
900▶ Q M A Y R F N G I G V T Q N V L Y E N Q K L I A N Q F N S A I G K
3601 ATCAAGACTCACTTTCTCCACAGCAAGTGCACCTGGAAAACCTCAAGATGTGGTCAACCAAAATGCACAAGCTTTAAACACGCTTGTAAACAACCTTA
934▶ I Q D S L S S T A S A L G K L Q D V V N Q N A Q A L N T L V K Q L
3701 GCTCCAATTTTGGTGCAATTTCAAGTGTAAAATGATATCCTTTACGTCTTGACAAAGTTGAGGCTGAAGTCAAATGATAGGTTGATCAGGCGAG
967▶ S S N F G A I S S V L N D I L S R L D K V E A E V Q I D R L I T G R
3801 ACTTCAAAGTTTGCAGACATATGTGACTCAACAATTAATTAGAGCTGCAGAAATCAGAGCTTCTGCTAATCTTGTCTACTAAAATGTCAGAGTGTGTA
1000▶ L Q S L Q T Y V T Q Q L I R A A E I R A S A N L A A T K M S E C V
3901 CTTGGACAATCAAAAAGAGTTGATTTTTGTGGAAAGGGCTATCATCTTATGTCCTTCCCTCAGTCAGCACCTCATGGTGTAGTCTTCTTGCATGTGACTT
1034▶ L G Q Q S K R V D F C G K G Y H L M S F P Q S A P H G V V F L H V T
4001 ATGTCCTGCACAAGAAAAGAACTTCACAACTGCTCCTGCCATTTGTCATGATGAAAAGCACACTTCTCGTGAAGGTGCTTTGTTTCAAATGGCAC
1067▶ Y V P A Q E K N F T T A P A I C H D G K A H F P R E G V F V S N G T
4101 AACTGGTTTGTAAACACAAAGGAATTTTTATGAACCACAAATCATTACTACAGACAACACATTTGTGTCTGGTAACTGTGATGTTGTAATAGGAATTGTC
1100▶ H W F V T Q R N F Y E P Q I I T T D N T F V S G N C D V V I G I V
4201 AACACACAGTTTATGATCCTTTGCAACTGAATTAGACTCATTCAAGGAGGAGTTAGATAAATTTTTAAGAATCATACATCACCAGATGTTGATTTAG
1134▶ N N T V Y D P L Q P E L D S F K E E L D K Y F K N H T S P D V D L
4301 GTGACTCTCTGGCATTAACTGTTGTAACACTTCAAAAAGAAATGACCGCTCAATGAGGTTGCCAAGAATTTAAATGAATCTCTCATCGATCT
1167▶ G D I S G I N A S V V N I Q K E I D R L N E V A K N L N E S L I D L
4401 CCAAGAACCTGAAAAGTATGAGCAGTATATAAATGGCCATGGTACATTTGGCTAGGTTTTATAGCTGGCTTGATTGCCATAGTAATGGTGACAATTATG
1200▶ Q E L G K Y E Q Y I K W P W Y I W L G F I A G L I A I V M V T I M

NheI (4567)

4501 CTTTGTGTATGACCAGTTGCTGTAGTTGTCTCAAGGGCTGTTGTTCTTGTGGATCCTGCTGCTAAAGCTAGCTGGCCAGACATGATAAGATACATTGAT
1234▶ L C C M T S C C S C L K G C C S C G S C C
4601 GAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATA
4701 AACAAAGTTAACACAACAATTCATTATTTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAATGTGGTAT
4801 GGAATCTAAAATACAGCATAGCAAACTTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTG
4901 TTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAACACTGCTTTCATTTCTTTAT
5001 GTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTAGAAAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCA
5101 GAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGA
5201 GCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGG
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 F C D P
5301 AGCATAGTCAGAGATGAGCTCTGTCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATG
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
5401 GTGTCAAAGTCTTCTGCCGTTGCTCAGCAGACCCAAATGGCAATGGCTTCAGCAGACAGCTGACCTGCCAATTAGGCTCAATGTGGACAGCAG
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
5501 AGATGATCTCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTGTTGTCCTCATAGAGCATGGTGTCTTCTCAGTGGGACCTCCACCAGCTC
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
5601 CAGATCCTGCTGAGAGATGTTGAAGTCTTATGATGGCCCTCTATAGTGTGCTATTATACTATGCGGATATACTATGCCGATGATTAATTGTCAA
11▶ L D Q Q S I N F T K M
5701 ACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTACCGCCATTTGCGTCAAT
5801 GGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAACCTCCATTGACGTCAATGGGTGGAGACTTGGAAATCCCCGTG
5901 AGTCAAACCGCTATCCACGCCATTGATGTAAGTCCGATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCGATGACTGCAAGTAGGAAAGTC
6001 CCATAAGTTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCC
6101 AAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTATTATTGACGTCAATGGG
6201 GGGGTCGTTGGGCGTCCAGCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCTGACGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCC
6301 AGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAATAATCAGCGCTCAAGTCAAGGTGGCGAAACC
6401 CGACAGGACTATAAAGATACCAGGCGTTTCCCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCT
6501 CCCTTCGGGAAGCGTGGCGTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGGTGTAGGTGTTCCGCTCAAGCTGGGCTGTGTGACGAACCCCC
6601 GTTCAGCCGACCGCTGCGCTTATCCGTAATATCGTCTTGTAGTCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA

6701 TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCT
6801 GAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTGCAAGCAGCAGATTACG
6901 CGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGAACGAAAACACGTTAAGGGATTTGGTCATGGCTA
7001 GTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTAATCGTAACTAACATACGCTCTC
7101 CATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA