



150

1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGGCTCTCCACGCTTTGCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

Mlul (558)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGACGCGTACCATGGGGTGGCTGGGGACTAGGCGTCTGTT
1▶ M G L A W G L G V L F
601 CCTGATGCATGTGTGGCACCAACCGCATTCCAGAGTCTGGCGGAGACAACAGCGTGTGACATCTTTGAACTCACCGGGCCGCCCGCAAGGGTCT
11▶ L M H V C G T N R I P E S G G D N S V F D I F E L T G A A R K G S
701 GGGCGCGACTGGTGAAGGGCCCGACCTTCCAGCCAGCTTCCGCATCGAGGATGCCAACCTGATCCCCCTGTGCTGATGACAAGTCCAAGACC
45▶ G R R L V K G P D P S S P A F R I E D A N L I P P V P D D K F Q D
801 TGGTGGATGCTGTGCGGGCAGAAAAGGTTTCTCCTTCTGGCATCCCTGAGGACAGATGAAGAAGACCCGGGGCAGCGTGTGGCCCTGGAGCGGAAAGA
78▶ L V D A V R A E K G F L L L A S L R Q M K K T R G T L L A L E R K D
901 CCACCTGGCCAGGTCTCAGCGTGGTGTCCATGGCAAGGCGGGCACCTGGACCTCAGCTGACCGTCCAAGGAAAGCAGCAGTGGTGTCTGTGGAA
111▶ H S G Q V F S V V S N G K A G T L D L S L T V Q G K Q H V V S V E
1001 GAAGCTCTCTGCAACCGCCAGTGAAGAGCATCACCTGTTTGTGAGGAAGACAGGGCCAGCTGTACATCGACTGTAAAAGATGGAGAATGCTG
145▶ E A L L A T G Q W K S I T L F V Q E D R A Q L Y I D C E K M E N A
1101 AGTTGGAGTCCCATCAAAGCTTCCAGCAGACCTGGCAGCATCGCCAGACTCCGCATCGCAAAGGGGGCGTCAATGACAAATTTCCAGGGGT
178▶ E L D V P I Q S V F T R D L A S I A R L R I A K G G V N D N F Q G V
1201 GCTGCAGAATGTGAGGTTTGTCTTTGAAACACACCAGAAGACATCCTCAGGAACAAAGGCTGCTCCAGCTCTACCAGTGTCTCCTCACCTTGACAAC
211▶ L Q N V R F V F G T T P E D I L R N K G C S S S T S V L L T L D N
1301 AACGTGGTGAATGTTCCAGCCCTGCCACTCCGACTAATACATTGGCCACAAGACAAGGACTTGAAGCCATCTCGGCATCTCCTGTGATGAGCTGT
245▶ N V V N G S S P A I R T N Y I G H K T K D L Q A I C G I S C D E L
1401 CCAGCATGGTCTGGAACCTCAGGGCCTGCGCACCATTTGTGACCACGCTGCAGGACAGCATCCGCAAAGTACTGAAGAGAACAAGAGTTGGCCAATGA
278▶ S S M V L E L R G L R T I V T T L Q D S I R K V T E E N K E L A N E
1501 GCTGAGGCGCCCTCCCCTATGCTATCACAACGGAGTTCAGTACAGAAATAACGAGGAATGGACTGTTGATAGTGCAGTGTACTGTGCAAGTCA
311▶ L R R P P L C Y H N G V Q Y R N N E E W T V D S C T E C H C Q N S
1601 GTTACCATCTGCAAAAAGGTGCTGCTGCCCATCATGCCTGTCCAATGCCACAGTTCCTGATGGAGAATGCTGTCTCGCTGTTGGCCAGCGACTCTG
345▶ V T I C K K V F S C P I M P C S N A T V P D G E C C P R C W P S D S
1701 CGACGATGGTGGTCCATGGTCCGAGTGGACCTCCTGTTCTACGAGCTGTGGCAATGGAATTCAGCAGCGCGCCCTCCTGCGATAGCTCAACAA
378▶ A D D G W S P W S E W T S C S T S C G N G I Q Q R G R S C D S L N N
1801 CCGATGTGAGGGTCTCCTCGTCCAGACCGGACCTGCCACATTGAGGAGTGTGACAAGAGATTTAAACAGGATGGTGGCTGGAGCCACTGGTCCCCGTGG
411▶ R C E G S S V Q T R T C H I Q E C D K R F K Q D G G W S H W S P W
1901 TCATCTTGTCTGTGACATGTGGTGTGATGATCACAAGGATCCGGCTCTGCAACTCTCCAGCCCCAGATGAACGGGAAACCTGTGAAGGCGAAG
445▶ S S C S V T C G D G V I T R I R L C N S P S P Q M N G K P C E G E
2001 CGCGGGACCAAAGCTGCAAGAAAGACCGCTCCCATCAATGGAGGTGGGGTCTTGGTACCATGGGACATCTGTTCTCACCTGTGGAGGAGG
478▶ A R E T K A C K A C K D A C P I N G G W G P W S P W D I C S V T C G G G
2101 GGTACAGAAACGTAGTCTCTGCAACAACCCACACCCAGTTTGGAGGCAAGGACTGCGTTGGTGTGTAACAGAAAACAGATCTGCAACAAGCAG
511▶ V Q K R S R L C N N P T P Q F G G K D C V G D V T E N Q I C N K Q
2201 GACTGTCCAATTGATGGATGCCTGTCCAATCCCTGCTTTGCCGGCGTGAAGTGTACTAGTACCTGATGGCAGTGGAAATGTGGTGTCTGTCCCCCTG
545▶ D C P I D G C L S N P C F A G V K C T S Y P D G S W K C G A C P P
2301 GTTACAGTGGAAATGGATCCAGTGCACAGATGTTGATGAGTCAAAGAAGTGCCTGATGCCTGCTTCAACCAATGGAGACACCGGTGTGAGAACAC
578▶ G Y S G N G I Q C T D V D E C K E V P D A C F N H N G E H R C E N T
2401 GGACCCCGGTACAACCTGCCCTGCCCTGCCCTCCCGCTCACAGCCCTTCGGCCAGGGTGTGAAACCTGGTGTGCGTGGCCAATGCGACTTA
611▶ D P G Y N C L P C P P R F T G S Q P F G Q G V E H A T A N K Q V C
2501 AAGCCCCGTAACCCCTGCACGGATGGGACCCACGACTGCAACAAGAAGCACAAGTGAACCTACCTGGGCCACTATAGCGACCCCATGTACCGTGTGAGT
645▶ K P R N P C T D G T H D C N K N A K C N Y L G H Y S D P M Y R C E
2601 GCAAGCCTGGCTACGCTGGCAATGGCATCATCTGCGGGGAGGACAGACCTGGATGGCTGGCCAATGAGAACCTGGTGTGCGTGGCCAATGCGACTTA
678▶ C K P G Y A G N G I I C G E D T D L D G W P N E N L V C V A N A T Y
2701 CCACTGCAAAAAGGATAATTGCCCAACCTTCCAACCTCAGGGCAGGAAGACTATGACAAGGATGGAATGGTGTGCTGTGATGATGACGATGACAAT
711▶ H C K K D N C P N L P N S G Q E D Y D K D G I G D A C D D D D N
2801 GATAAATCCAGATGACAGGACAACCTGTCCATTCCATTCAACCCAGCTCAGTATGACTATGACAGAGATGTGGGAGACCGCTGTGACAACCTGT
745▶ D K I P D D R D N C P P R F T G S Q P F G Q G V E H A T A N K Q V C
2901 CCTACAACCACAACCCAGATCAGGCAGACAGACAACAATGGGGAAGGAGACGCTGTGCTGCAGACATTGATGGAGACGGTATCCTCAATGAACGGGA
778▶ P Y N H N P D Q A D T D N N G E G D A C A A D I D G D G I L N E R D
3001 CAACTGCCAGTACGTCTACAATGTGGACAGAGAGACTGATATGGATGGGGTGGAGATCAGTGTGACAATGGCCCTTGGAAACACAATCCGGATCAG
811▶ N C Q Y V Y N V D Q R D T D M D G V G D Q C D N C P L E H N P D Q
3101 CTGGACTCTGACTCAGACCGCATTGGAGATACTGTGACAACAATCAGGATATTGATGAAGATGGCCACCAGAACAATCTGGACAACCTGTCCCTATGTGC
845▶ L D S D S D R I G D T C D N N Q D I D E D G H Q N N L D N C P Y V
3201 CCAATGCCAACAGGCTGACCATGACAAGATGGCAAGGGAGATGCCTGTGACCACGATGATGACAACGATGGCATTCTGATGACAAGGACAACCTGCAG
878▶ P N A N Q A D H D K D G K G D A C D H D D D N D G I P D D K D N C R

3301 ACTCGTGCCCAATCCCGACCAGAAGGACTCTGACGGCGATGGTCGAGGTGATGCCTGCAAAGATGATTTTGACCATGACAGTGTGCCAGACATCGATGAC
911▶ L V P N P D Q K D S D G D G R G D A C K D D F D H D S V P D I D D
3401 ATCTGTCTGAGAATGTTGACATCAGTGAGACCGATTTCCGCCGATTCCAGATGATTCCTCTGGACCCCAAAGGGACATCCCAAAATGACCCCTAACTGGG
945▶ I C P E N V D I S E T D F R R F Q M I P L D P K G T S Q N D P N W
3501 TTGTACGCCATCAGGGTAAAGAAGCTCCAGACTGTCAACTGTATCCTGGACTCGCTGTAGGTTATGATGAGTTTAAATGCTGTGGACTTCAGTGGCAG
978▶ V V R H Q G K E L V Q T V N C D P G L A V G Y D E F N A V D F S G T
3601 CTTCTTATCAACACCGAAAGGGACGATGACTATGCTGGATTTGTCTTTGGCTACCAGTCCAGCAGCCGCTTTTATGTTGTGATGTGGAAGCAAGTACC
1011▶ F F I N T E R D D D Y A G F V F G Y Q S S S R F Y V V M W K Q V T
3701 CAGTCCTACTGGGACCAACCCACGAGGGCTCAGGGATACTCGGGCTTTCTGTGAAAGTTGTAACCTCCACCACAGGGCCTGGCAGCACCTGCGGA
1045▶ Q S Y W D T N P T R A Q G Y S G L S V K V V N S T T G P G E H L R
3801 ACGCCCTGTGGCACACAGGAAACACCCCTGGCCAGGTGCGCACCTGTGGCATGACCTCGTCACATAGGCTGAAAGATTTACCGCCTACAGATGGCG
1078▶ N A L W H T G N T P G Q V R T L W H D P R H I G W K D F T A Y R W R
3901 TCTCAGCCACAGGCCAAAGACGGGTTTCATTAGAGTGGTGTATGAAGGGAAGAAAATCATGGCTGACTCAGGACCCATCTATGATAAAAACCTATGCT
1111▶ L S H R P K T G F I R V V M Y E G K K I M A D S G P I Y D K T Y A
4001 GGTGGTAGACTAGGGTTGTTTGTCTTCTCAAGAAATGGTGTCTTCTGACCTGAAATACGAATGTAGAGATCCCTAATCATCAAATTTGTTGATTGA
1145▶ G G R L G L F V F S Q E M V F F S D L K Y E C R D P •

NheI (4141)

4101 AAGACTGATCATAAACCAATGCTGGTATTGCACCTTCTGGAGCTAGCTAGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAAT
4201 GCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAAACAACAATTGCATT
4301 CATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTTCTAAAATACAGCATAGCAAA
4401 ACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGA
4501 GCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGTGTGAACTAGCTCTTCATTTCTTTATGTTTAAATGCACTGACCTCCACAT
4601 TCCTTTTTAGTAAAATATTAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCAT
4701 AATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTAC
141▶ • N R T Y
4801 TTGAGGGGATGAGTTTCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAAGCAGTCCAGGAGCATAGTCAGAGATGAGCTCTCTGC
135▶ 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 A Y D S I L E R C
4901 ACATGCCACAGGGGCTGACCACCTGATGGATCTGCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCT
102▶ M G C P S V V 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
5001 CACAGCAGACCAATGGCAATGGCTTCAGCAGACAGTACCCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTGGTCTG
69▶ 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 I I E G T K T R
5101 ATGGCCGCCCCGACATGGTGCTTGTGCTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGG
35▶ 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 L D Q Q S I N F T
5201 TCTTCATGGTGGCCCTCTATAGTGAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTA
2▶ K M
5301 TCTGACGGTTCACATAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGG
5401 AAAGTCCCCTGATTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCTGAGTCAAACCGCTATCCACGCCATTG
5501 ATGTAAGTCCAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCAAGTAGGAAAGTCCATAAGGTGATGACTGGGCATAAT
5601 GCCAGGCGGGCATTACCGTCAATGACGTCAATAGGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGAGTTTACCGTAAATACTC
5701 CACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTATTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTGAGCCAGGC
5801 GGGCATTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGT
5901 GCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGGC
6001 TTTCCCTTGAAAGTCCCTCGTGCCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCGCTTTCTCCCTTCCGGGAAAGCGTGGCGCTTCTC
6101 ATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCTTATC
6201 CGGTAACATATCGTCTTGAAGTCCAAACCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTATGTAGCGGT
6301 GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGTTACCTTCGAAAAAGAG
6401 TTGGTAGCTCTTGTATCCGGCAACAAACCACCGCTGGTAGCGGTGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGA
6501 TCCTTTGATCTTTTCTACGGGCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTGGTATGGCTAGTTAATTAACATTTAATCAGCGGCC
6601 GCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGTTTTTTGTGTAATCGTAACTAACATACGCTCTCCATCAAACAAACGAAACAAACAA
6701 ACTAGCAAAATAGGCTGTCCCGAGTCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA