



125

1 GGATCTGCATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTCCGCTTCCGCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGGCTACCTAGACTCAGCGGGCTCTCCACGCTTTGCCTGACCTGCTTGTCAACTCTACGCTTTGTTTCGTTT

Mlul (558)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGCACGCGTACCATGGCCAAGGGTCTATATTTCCAAGTCCCT
1▶ M A K G F Y I S K S L
601 GGCATCTGGGATCCTCCTGGGCGTGGCAGCCGTGTGCACAAATCATCGACTGTAGTGGTGTACTCCAGGAGAAGAACAAGAACGCCAACAGCTCC
11▶ G I L G I L L G V A A V C T I I A L S V V Y S Q E K N K N A N S S
701 CCCGTGGCTCCACCACCCGTCGCGCTCAGCCACCACCAACCCGCTCGGCCACCCTTGGACCAAAGTAAAGCGTGAATCGTTACCGCTCCCA
45▶ P V A S T T P S A S A T T N P A S A T T L D Q S K A W N R Y R L P
801 ACACGCTGAAACCGATTCTACCAGGTGACGCTGAGACCGTACCTCACCCCAATGACAGGGGCTGTACGTTTTAAGGGCTCCAGCACCGTCCGTTT
78▶ N T L K P D S Y Q V T L R P Y L T P N D R G L Y V F K G S S T V R F
901 CACCTGCAAGGAGGACTGACGTCATCATCCACAGCAAGAAGCTCAACTACACCTCAGCCAGGGGCACAGGGTGGTCTGCGTGGTGGGAGGC
111▶ T C K E A T D V I I I H S K K L N Y T L S Q G H R V V L R G V G G
1001 TCCAGCCCCCGACATTGACAAGACTGAGCTGGTGGAGCCACCGAGTACCTGGTGGTGCACCTCAAGGGTCCCTGGTGAAGGACAGCCAGTATGAGA
145▶ S Q P P D I D K T E L V E P T E Y L V V H L K G S L V K D S Q Y E
1101 TGGACAGCGAGTTCGAGGGGAGTTGGCAGATGACCTGGCGGCTTCTACCGCAGCGAGTACATGGAGGGCAATGTCAGAAAGTGGTGGCCACTACACA
178▶ M D S E F E G E L A D D L A G F Y R S E Y M E G N V R K V V A T T Q
1201 GATGCAGGCTGCAGATGCCCGAAGTCTTCCCATGCTTCGATGAGCCGGCCATGAAGGCCGAGTTCAACATCACGCTTATCCACCCCAAGGACCTGACA
211▶ M Q A A D A R K S F P C F D E P A M K A E F N I T L I H P K D L T
1301 GCCCTGTCCAACATGCTTCCAAAGGTCAGCAGCCCACTTCCAGAAGACCCCACTGGAATGCTACTGAGTTCACACCCACGCCAAGATGTCACGT
245▶ A L S N M L P K G P S T P L P E D P N W N V T E F H T T P K M S T
1401 ACTTGCTGGCCTTATTGTCAGTGTGACTGAGTTCGACTACGTGGAGAAGCAGGCATCCAATGGTGTCTTGATCCGGATCTGGGCCCGCCAGTGCCATTGCGGC
278▶ Y L L A F I V S E F D Y V E K Q A S N G V L I R I W A R P S A I A A
1501 GGGCCACGGCGATTATGCCTGAACGTGACGGGCCCATCTTAACCTCTTGTGGTATTATGACACACCTACCCTCCAAAATCAGACCAGATT
311▶ G H G D Y A L N V T G P I L N F F A G H Y D T P Y P L P K S D Q I
1601 GGCTGCCAGACTTCAACGCGGCGCCATGGAGAAGTGGGACTGGTACCTACCGGGAGAAGTCCCTGCTGTTGACCCCCCTGCTCCTCCAGCAGCA
345▶ G L P D F N A G A M E N W G L V T Y R E N S L L F D P L S S S S S
1701 ACAAGGAGCGGGTGGTACTGTGATTGCTCATGAGTGGCCACCATGTTGTCGGAACTGGTGACCATAGAGTGGTGAATGACCTGGTGGTGAACGA
378▶ N K E R V V T V I A H E L A H Q W F G N L V T I E W W N D L W L N E
1801 GGGCTTCGCTCCTACGTGGAGTACCTGGTGTGACTATGCGGAGCCACCTGGAAGTGAAGACCTCATGGTGTGAATGATGTACCGCGTATG
411▶ G F A S Y V E Y L G A D Y A E P T W N L K D L M V L N D V Y R V M
1901 GCAGTGGATGACTGGCCTCCTCCACCCGCTGTCACACCCGCTCGGAGATCAACACGCGCCGAGATCAGTGTGCTGTTGACGCCATCTCCTACA
445▶ A V D A L A S S H P L S T P A S E I N T P A Q I S E L F D A I S Y
2001 GCAAGGGCGCTCAGTCTCAGGATGCTCTCCAGTCTCCTGTCCGAGGAGTTCAGGAGGCTGAGGAGGCTGGGCTCCTACCTCCACACCTTTGCCTACCAGAA
478▶ S K G A S V L R M L S S F L S E D V F K Q G L A S Y L H T F A Y Q N
2101 CACCATCTACCTGAACCTGTGGGACCCTGCAGGAGGCTGTGAACAACCGTCCACTCAACTCCCCACCACCGTGGGGGACATCATGAACCGCTGGACC
511▶ T I Y L N L W D H L Q E A V N N R S I Q L P T T V R D I M N R W T
2201 CTGAGATGGGCTTCCGGTTCATCAGGTGGATACCAGCAGGGGACCTTTCCAGGAGCACTTCTCCTTGACCCGATTCCAATGTTACCGCCCT
545▶ L Q M G F P V I T V D T S T G T L S Q E H F L L D P D S N V T R P
2301 CAGAATCAACTACGTGTGATTGTGCCATCATCATCAGAGATGGCAGACAGCAGGACTACTGGCTGATAGATGTAAGAGCCAGAACGATCT
578▶ S E F N Y V W I V P I T S I R D G R Q Q Q D Y W L I D V R A Q N D L
2401 CTTCAGCATCAGGCAATGAGTGGTCTGCTGAACCTCAATGTGACGGGCTATTACCGGTGAACACGAGGAGAACTGGAGGAAGATTGAGACT
611▶ F S T S G N E W V L L N L N V T G Y Y R V N Y D E E N W R K I Q T
2501 CAGCTGCAGAGAGACCTCGGCCATCCTGTGATCAATCGGGCAGATCATTAATGACGCCTTCAACTGGCCAGTGGCCATAAGTCCCTGTCACTC
645▶ Q L Q R D H S A I P V I N R A Q I I N D A F N L A S A H K V P V T
2601 TGGCGTGAACAACCCCTTCTCCTGATTGAAGAGAGACAGTACATGCCTGGGAGGCGCCCTGAGCAGCCTGAGCTACTCAAGCTCATGTTTGACCG
678▶ L A L N N T L F L I E E R Q Y M P W E A A L S S L S Y F K L M F D R
2701 CTCGAGGTCTATGGCCCATGAAGAATACCTGAAGAAGCAGGTACACCCCTTCTTATTCACTTCAAGAAATAACCAACAAGTGGAGGGAGATCCCA
711▶ S E V Y G P M K N Y L K K Q V T P L F I H F R N N T N N W R E I P
2801 GAAAACCTGATGGACAGTACAGCGAGTTAATGCCATCAGCACCCTGCTCCAACGGAGTTCAGAGTGTGAGGAGATGGTCTCTGGCCTTTTCAAGC
745▶ E N L M D Q Y S E V N A I S T A C S N G V P E C E E M V S G L F K
2901 AGTGGATGGAGAACCCCAATAAACCAGTCCACCCCACTGCGGTCCACCGTCTACTGCAACGCTATCGCCAGGGCGGGGAGGAGTGGGACTT
778▶ Q W M E N P N N N P I H P N L R S T V Y C N A I A Q G G E E E W D F
3001 CGCCTGGGAGCAGTTCGAAATGCCACTGGTCAATGAGGCTGACAAGCTCCGGGACCGCTGGCCTGCAGCAAAGAGTGTGGATCCTGAACAGGTAC
811▶ A W E Q F R N A T L V N E A D K L R A A L A C S K E L W I L N R Y
3101 CTGAGCTACACCTGAACCCGACTTAATCCGAAGCAGGACGCCACCTTACCATCATCAGCATTACCAACAACGTCATTGGGCAAGGTCTGGTCTGGG
845▶ L S Y T L N P D L I R K Q D A T S T I I S I T N N V I G Q G L V W
3201 ACTTTGTCCAGAGCAACTGGAAGAAGCTTTTAAACGATTATGGTGGTGGCTGTTCTCCTTCCAACTCATCCAGGAGTACACGACGATTCTCCAC
878▶ D F V Q S N W K K L F N D Y G G S F S F S N L I Q A V T R R F S T

3301 CGAGTATGAGCTGCAGCAGCTGGAGCAGTTCAAGAAGGACAACGAGGAAACAGGCTTCGGCTCAGGCACCCGGGCCCTGGAGCAAGCCCTGGAGAAGACG
911▶ E Y E L Q Q L E Q F K K D N E E T G F G S G T R A L E Q A L E K T
NheI (3491)

3401 AAAGCCAACATCAAGTGGGTGAAGGAGAACAAGGAGGTGGTGTCCAGTGGTTCACAGAAAACAGCAAATAGTCCCCAGCCCTGAAGTCAGCTAGCTGG
945▶ K A N I K W V K E N K E V V L Q W F T E N S K •

3501 CCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTAT

3601 TTGTAACCATTATAAGCTGCAATAAACAAAGTTAAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAA

3701 GTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACCTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGA

3801 ATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTT

3901 TGAAGTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTATGAAAATATTCAGAAAATAATTTAAATACATCATTGCAATG

4001 AAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAAT

4101 AGAAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCAT
141◀ • N R T Y K L P I L E E I T T K V L K G N M

4201 CTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGACATGCCACAGGGGTGACCACCTGATGGATCTGCCACCTCATCAGAG
119◀ E I L V F C D P 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

4301 TAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCTGCCAA
85◀ 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 I

4401 TGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTGGTCTGATGGCCGCCGACATGGTGTGTTGTCTCATAGAGCATGGTATCTT
52◀ 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 K

4501 CTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATA
19◀ E T A V E V L E L D Q Q S I N F T K M

4601 CTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCCTAAACGAGCTCTGCTTATATAGACCTCCACCCTACA

4701 CGCTACCGCCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGG

4801 GGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTA

4901 GATGTAAGTCCAAGTAGGAAAGTCCCATAGGTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGC

5001 ATATGATACACTTGATGTAAGTGGCAGTTTACCCTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACAT

5101 ACGTCATTATTGACGTCAATGGGCGGGGCTGTTGGCGGTGACGCCAGGCGGGCCATTTACCCTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATG
CCTGCAGGTTAATTAAGAACATG

5201 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAGAAAATCGAC

5301 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCTGCCGCT

5401 TACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAG

5501 CTGGGCTGTGTGCACGAACCCCGTTAGCCGACCGCTGCGCTTATCCGTAACATCGTCTTGAAGTCCAACCCGTAAGACACGACTTATGCCAC

5601 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAAC

5701 AGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGAAGTCCAACCCGTAAGACACGACTTATGCCAC

5801 TTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTACGGGTCTGACGCTCAGTGGAAACGAAAACTCAC

5901 GTTAAGGGATTTTGGTATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGTTTTTGTGT

6001 GAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTCCCCAGTGAAGTGCAGGTGCCAGAACATTTCT

6101 TCTATCGAA