



1 GGATCTGCATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGCTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGGCTCTCCACGCTTTGCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

Agel (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCTGAGATGAAAAGCCCTGCTTTGCAACCCCTCAGCATGGCAGG
1 ▶ M K S P A L Q P L S M A G
601 CCTGCAGTCATGACCCCTGCTTCTCAACCAATGGGTCTTTCTTTGGACTGCCATGGCAACAAGAAGCAATTCATGATAACATTTATACGCCAAGAAAA
13▶ L Q L M T P A S S P M G P F F G L P W Q Q E A I H D N I Y T P R K
701 TATCAGGTTGAAGTCTGAAGCAGCTCTGGATCATAATACCATCGTCTGTTAAACTGGCTCAGGGAAGACATTTATTGCAGTACTACTACTAAAG
47▶ Y Q V E L L E A A L D H N T I V C L N T G S G K T F I A V L L T K
801 AGCTGTCCTATCAGATCAGGGGAGACTTCAGCAGAAATGAAAAAGGACGGTGTCTTGGTCAACTCTGCAAACAGGTTGCTCAACAAGTGTGAGCTGT
80▶ E L S Y Q I R G D F S R N G K R T V F L V N S A N Q V A Q Q V S A V
901 CAGAAGTCTCAGATCTCAAGTTGGGAATCTCAACCTAGAAGTAAATGCATCTGGACAAAAGAGAGATGGAACCAAGAGTTACTAAGCACCAG
113▶ R T H S D L K V G E Y S N L E V N A S W T K E R W N Q E F T K H Q
1001 GTTCTCATTATGACTTGTATGTCGCCTGAATGTTTTGAAAAATGGTACTTATCACTGTGACAGATTAACCTTTTGGTGTGTTGATGAGTGTGATCTTTG
147▶ V L I M T C Y V A L N V L K N G Y L S L S D I N L L V F D E C H L
1101 CAATCCTAGACCACCCCTATCGAGAAATTAAGCTCTGTGAAAATGTCCATCATGCTCCTCGATTTTGGGACTAACTGCTCCATTTTAAATGGGAA
180▶ A I L D H P Y R E I M K L C E N C P S C P R I L G L T A S I L N G K
1201 ATGTGATCCAGAGGAATTGGAAGAAAAGATTGAGAACTAGAGAAAATTTTAAAGAGTAATGTGAAACTGCAACTGACCTGGTGGTCTTAGACAGGTAT
213▶ C D P E E L E E K I Q K L E K I L K S N A E T A T D L V V L D R Y
1301 ACTTCTCAGCCATGTGAGATTGGTGGATTGGACCATTTACTGACAGAAGTGGGCTTTATGAAAGACTGCTGATGGAATTAGAAGAAGCACTTAATT
247▶ T S Q P C E I V V D C G P F T D R S G L Y E R L L M E L E E A L N
1401 TTATCAATGATTGTAATATATCTGTACATTCAAAAGAAAAGAGATTCTACTTTAATTTGAAAACAGATACTATCAGACTGTCGTGCCGATTGGTAGTCT
280▶ F I N D C N I S V H S K E R D S T L I S K Q I L S D C R A V L V V L
1501 GGGACCCTGGTGTGCAGATAAAGTAGCTGGAATGATGGTAAGAGAAGTACAGAAATACATCAACATGAGCAAGAGGAGCTGCACAGGAAATTTTATTG
313▶ G P W C A D K V A G M M V R E L Q K Y I K H E Q E E L H R K F L L
1601 TTTACAGACACTTTCTAAGGAAAATACATGCACTATGTGAAGAGCACTTCTCACCTGCCTCACTTGCCTGAAATTTGAACTCCTAAAGTAATCAAC
347▶ F T D T F L R K I H A L C E E H F S P A S L D L K F V T P K V I K
1701 TGCTCGAAATCTTACGCAATATAAACCATATGAGCGCAGCAGTTTGAAGCGTTGAGTGGTATAATAATAGAAATCAGGATAAATATGTCATGGAG
380▶ L L E I L R K Y K P Y E R Q Q F E S V E W Y N N R N Q D N Y V S W S
1801 TGATTCTGAGGATGATGATGAGGATGAAGAAATGAAGAAAAAGAGAAGCCAGAGACAAATTTTCTTCTCTTTTACCAACATTTTGTGCGGAATTATT
413▶ D S E D D D E D E E I E E K E K P E T N F P S P F T N I L C G I I
1901 TTTGTGAAAAGAAAGATACACAGCAGTTGTCTTAAACAGATTGATAAAGGAAGCTGGCAACAAGATCCAGAGCTGGCTTATATCAGTAGCAATTTTCATAA
447▶ F V E R R Y T A V V L N R L I K E A G K Q D P E L A Y I S S N F I
2001 CTGGACATGGCATTGGGAAGAATCAGCTCGCAACAACAGATGGAAGCAGAATTCAGAAAACAGGAAGGACTTAGGAAATTTTCGAGCAGATGAGAC
480▶ T G H G I G K N Q P R N K Q M E A E F R K Q E E V L R K F R A H E T
2101 CAACCTGCTTATTGCAACAGTATTGTAGAAGGGGTTGATATAACAAAATGCAACTTGGTGGTTGCTTTGATTTGCCACAGAAATATCGATCCTAT
513▶ N L L I A T S I V E E G V D I P K C N L V V R F D L P T E Y R S Y
2201 GTTCAATCTAAGGAAGAGCAAGGGCACCCATCTCAATTATATAATGTTAGCGGATACAGACAAAATAAAAAGTTTTGAAGAAGACCTTAAACCTACA
547▶ V Q S K G R A R A P I S N Y I M L A D T D K I K S F E E D L K T Y
2301 AAGCTATTGAAAAGATCTTGAGAAAAGTGTTCGAAGTCCGTTGATACTGGTGGAGACTGACATTGATCCTGTGATGATGATGACGTTTTCCACC
580▶ K A I E K I L R N K C S K S V D T G E T D I D P V M D D D D V F P P
2401 ATATGTGTGAGGCTGACGATGGTGGTCCACGAGTCAACATCAACACGGCCATTGGACACATCAATAGATACTGTGCTAGATTACCAAGTGTGATCGGTTT
613▶ Y V L R P D D G G P R V T I N T A I G H I N R Y C A R L P S D P F
2501 ACTCATCTAGCTCCTAAATGCAACCCGAGAGTTGCCTGATGGTACATTTTATTCAACTCTTTATCTGCCAATTAACCTCACCTCTTCGAGCCTCATTG
647▶ T H L A P K C R T R E L P D G T F Y S T L Y L P I N S P L R A S I
2601 TTGGTCCACCAATGAGCTGTGTACGATTGGCTGAAAGAGTTGAGCTCTCATTGCTGTGAGAAAAGTGCACAAAATGGCGAACTGGATGACCATTTGAT
680▶ V G P P M S C V R L A E R V V A L I C C E K L H K I G E L D D H L M
2701 GCCAGTTGGGAAAGAGACTGTTAAATATGAAGAGGAGCTTGTATGATGATGAAGAAGAGACCAGTGTCCAGGAAGACCAGGTTCCACGAAACGAAGG
713▶ P V G K E T V K Y E E E L D L H D E E E T S V P G R P G S T K R R
2801 CAGTGTACCCAAAAGCAATTCAGAGTGTGAGGAGTAGTTATCCAGACCTGATCAGCCCTGTACCTGTATGTGATGGAATGGTTTTAACTACAC
747▶ Q C Y P K A I P E C L R D S Y P R P D Q P C Y L Y V I G M V L T T
2901 CTTTACTGATGAACTCAACTTTAGAAGCGGAAGCTCTATCTCCTGAAGATACCACAAGATGCTTTGGAATACTGACGGCAAAACCCATACCTCAGAT
780▶ P L P D E L N F R R R K L Y P P E D T T R C F G I L T A K P I P Q I
3001 TCCACATTTCTGTGATACACGCTCTGGAGAGGTTACCATATCCATTGAGTTGAAGAAGTCTGGTTTCATGTTGTCTCTACAAATGCTTGAAGTTGATT
813▶ P H F P V Y T R S G E V T I S I E L K K S G F M L S L Q M L E L I
3101 ACAAGACTTACCAGTATATATTCTCACATATTCTCGGCTTGAACAACTGCACTAGAATTTAAACCTACAGACGCTGATTCAGCATACTGTGTTCTAC
847▶ T R L H Q Y I F S H I L R L E K P A L E F K P T D A D S A Y C V L
3201 CTCTAATGTTGTAATGACTCCAGCACTTTGGATATTGACTTTAAATTCATGGAAGATATTGAGAAGTCTGAAGCTCGCATAGGCATTCCAGTACAAA
880▶ P L N V V N D S S T L D I D F K F M E D I G A K S E A R I G I P S T K

3301 GTATACAAAAGAAACACCCCTTTGTTTTAAATTAGAAGATTACCAAGATGCCGTTATCATTCCAAGATATCGCAATTTTGATCAGCCTCATCGATTTAT
913▶ Y T K E T P F V F K L E D Y Q D A V I I P R Y R N F D Q P H R F Y
3401 GTAGCTGATGTGTACTGATCTTACCCCACTCAGTAAATTCCTTCCCTGAGTATGAAACTTTTGAGAATATTATAAAAACAAAGTACAACCTTGACC
947▶ V A D V Y T D L T P L S K F P S P E Y E T F A E Y Y K T K Y N L D
3501 TAACCAATCTCAACCAGCCACTGCTGGATGTGGACCACACATCTTCAAGACTTAATCTTTTGACACCTCGACATTTGAATCAGAAGGGGAAAGCGCTTCC
980▶ L T N L N Q P L L D V D H T S S R L N L L T P R H L N Q K G K A L P
3601 TTTAAGCAGTGTGAGAAGAGGAAAGCCAAATGGGAAAGTCTGCAGAATAAACAGATACTGGTTCCAGAACTCTGTGTATACATCCAATTCAGCATCA
1013▶ L S S A E K R K A K W E S L Q N K Q I L V P E L C A I H P I P A S
3701 CTGTGGAGAAAAGCTGTTTGTCTCCCCAGCATACTTTATCGCTTCTACTGCCTTTTGACTGCAGAGGAGCTAAGAGCCAGACTGCCAGCGATGCTGGCG
1047▶ L W R K A V C L P S I L Y R L H C L L T A E E L R A Q T A S D A G
3801 TGGGAGTCAAGTCACTTCTGCGGATTTTAGATACCCTAAGTACTGAGTCCGGTGGAAAAAATCTATTGACAGCAAATCTTTCATCTCAATTTCTAACTC
1080▶ V G V R S L P A D F R Y P N L D F G W K K S I D S K S F I S I S N S
3901 CTCTTCAGTGAATAATTACTGTAAGCACAGCACAATGTCCCTGAAATGCTGCACATCAAGGTGCTAATAGAACCTCTCTAGAAAATCAT
1113▶ S S A E N D N Y C K H S T I V P E N A A H Q G A N R T S S L E N H
4001 GACCAAATGTCTGGAAGTGCAGAAGTGTCTCAGCGAGTCCCTGGTAAAGTCCAGTGTGAAGTTCAGCAGATCTTACAGCAATTAATGGTCTTCTT
1147▶ D Q M S V N C R T L L S E S P G K L H V E V S A D L T A I N G L S
4101 ACAATCAAATCTCGCAATGGCAGTTATGATTAGCTAACAGAGACTTTTGCAAGGAAATCAGCTAAATTAACAAGCAGGAAATACCCGTGCAACC
1180▶ Y N Q N L A N G S Y D L A N R D F C Q G N Q L N Y Y K Q E I P V Q P
4201 AACTACCTCATATTCATTGAGAAATTTATACAGTTACGAGAACCAGCCCCAGCCAGCGATGAATGTACTCTCTGAGTAATAAATACCTGATGAAAT
1213▶ T T S Y S I Q N L Y S Y E N Q P Q P S D E C T L L S N K Y L D G N
4301 GCTAACAAATCACTCAGATGGAAGTCTGTGATGCGCGTAAATGCCTGGTACGACAGACTATTCAAGTGTCAAGGCGAGGATGGATTGAGCGA
1247▶ A N K S T S D G S P V M A V M P G T T D T I Q V L K G R M D S E Q
4401 GCCCTTCTATTGGGACTCCTCAAGGACTCTTGGCCCAATCTGGACTTATTCTTCAGGCTTTGACTCTGTCAAACGCTAGTGATGGATTTAACCTGGA
1280▶ S P S I G Y S S R T L G P N P G L I L Q A L T L S N A S D G F N L E
4501 GCGGCTTGAATGCTTGGCGACTCCTTTTTAAAGCATGCCATACCACATATCTATTTGCACTTACCCTGATGCGCATGAGGGCCGCTTTTCATATATG
1313▶ R L E M L G D S F L K H A I T T Y L F C T Y P D A H E G R L S Y M
4601 AGAAGCAAAAAGTCAAGTCAACTGTAATCTGTATCGCTTGGAAAAAGAGGGACTACCCAGCCGATGGTGGTGTCAATATTTGATCCCCCTGTGAAT
1347▶ R S K L V S N C N L Y R L T G K K K A G T G P S R M V V S I F D P P V N
4701 GGCTTCCCTGGTTATGTAGTAAATCAAGACAAAAGCAACAGATAAATGGGAAAAAGTGAATGACAAAAGACTGCATGCTGGCGAATGGCAAAT
1380▶ W L P P G Y V V N Q D K S N T D K W E K D E M T K D C M L A N G K L
4801 GGATGAGGATTACGAGGAGGAGGATGAGGAGGAGGAGCCTGATGTGGAGGGCTCCGAAGGAAGAGGCTGACTATGAAGATGATTTCTGGAGTATGAT
1413▶ D E D Y E E E D E E E E S L M W R A P K E E A D Y E D D F L E Y D
4901 CAGGAACATATCAGATTTATAGATAATATGTTAATGGGTCAGGAGCTTTGTAAAGAAAATCTCTTTTCTCTTTTCAACCACTGATTCGCATATG
1447▶ Q E H I R F I D N M L M G S G A F V K K I S L S P F S T T D S A Y
5001 AATGGAAAATGCCAAAAATCTCCTTAGTGTATGCAATTTTTCATCAGATTTTGAGGATTTTGACTACAGCTCTGGGATGCAATGTGCTATCTGGA
1480▶ E W K M P K K S S L G S M P F S S D F E D F D Y S S W D A M C Y L D
5101 TCCTAGCAAAGCTGTTGAAGAAGATGACTTTGTGGTGGGTTCTGGAATCCATCAGAAGAAAATGTTGGTGTGACACGGGAAAGCAGTCCATTTCTTAC
1513▶ P S K A V E E D D F V V G F W N P S E E N C G V D T G K Q S I S Y
5201 GACTTGACACTGAGCAGTGTATTGCTGACAAAAGCATAGCGGACTGTGTGGAAGCCCTGCTGGGCTGCTATTTAACCAGCTGTGGGAGAGGGCTGCTC
1547▶ D L H T E Q C I A D K S I A D C V E A L L G C Y L T S C G E R A A
5301 AGCTTTTCTCTGTTCACTGGGCTGAAGTGTCTCCGGTAATTTAAAGGACTGATCGGAAAAAGCCCTGTGCCCTACTCGGGAGAATTTCAACAGCCA
1580▶ Q L F L C S L G L G V L P V I K R T D R E K A L C P T R E N F N S Q
5401 ACAAAGAACCCTTCAGTGAGTGTGCTGCTGTGCCAGTTCACGCTCTTGTATTGAAAGACTCGGAATATGGTTGTTGAAGATTTCCACAT
1613▶ Q K N L S V S C A A A S V A S S R S S V L K D S E Y G C L K I P P
5501 AGATGTATGTTTGATCATCCAGATGCAGATAAAACACTGAATCACCTTATATCGGGTGTGAAAAATTTGAAAAGAAAATCAACTACAGATTCAGAATA
1647▶ R C M F D H P D A D K T L N H L I S G F E N F E K K I N Y R F K N
5601 AGGCTTACCTTCCAGGCTTTACACATGCCTCCTACCACTACAATACTATCACTGATTGTTACCAGCGCTTAGAATTCCTGGGAGATGCGATTTTGA
1680▶ K A Y L L Q A F T H A S Y H Y N T I T D C Y Q R L E F L G D A I L D
5701 CTACCTCATAACCAAGCACCTTTATGAAGACCCGCGCAGCACTCCCCGGGGTCTGACAGACCTGCGGTCTGCCCTGGTCAACAACACCATCTTTGCA
1713▶ Y L I T K H L Y E D E P R Q H S P G V L T D L R S A L V N N T I F A
5801 TCGTGGCTGTAAGTACGACTACCACAAGTACTTCAAAGTGTCTCTGAGCTTCCATGCTCATTGATGACTTTGTGCAAGTTTGTGAGTGTGAGAAGA
1747▶ S L A V K Y D Y H K Y F K A V S P E L F H V I D D F V Q F Q L E K
5901 ATGAAATGCAAGGAATGGATTCTGAGCTTAGGAGATCTGAGGAGGATGAAGAGAAAGAAGAGGATATTGAAGTTCAAAGGCCATGGGGATATTTTTGA
1780▶ N E M Q G M D S E L R R S E E D E E K E E D I E V P K A M G D I F E
6001 GTCGCTTGTGGTCCATTTACATGGATAGTGGGATGTCACTGGAGACAGTCTGGCAGGTGTACTATCCCATGATGCGGCCACTAATAGAAAAGTTTTCT
1813▶ S L A G A I Y M D S G M S L E T V W Q V Y Y P M M R P L I E K F S
6101 GCAAATGTACCCGTTCCCTGTGCGAGAATTGCTTGAATGGAACCAAGAACTGCAAAATTTAGCCCGGCTGAGAGAATTCACGACGGGAAGGTGAGAG
1847▶ A N V P R S P V R E L L E M E P E T A K F S P A E R T Y D G K K V R
6201 TCACGTGGAAGTAGTAGGAGGGGAAATTTAAAGGTGTTGGTCAAGTTTACAGGATTTGCCAAATCTGCAGCAGCAAGAAGAGCCCTCCGAAGCCTCAA
1880▶ V T V E V V G K G K F K G V G R S Y R I A K S A A A R R A L R S L K

NheI (6351)

6301 AGCTAATCAACCTCAGGTTCCCAATAGCTGAAACCGCTTTTTAAAATCAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCA
1913▶ A N Q P Q V P N S •
6401 CAACTAGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACA
6501 CAATTGCATTCATTTTATGTTTCAGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTTACAAATGTGGTATGGAATTCAAAATACA
6601 GCATAGCAAACTTTAACCTCCAATCAAGCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGCCAATGTGCATTA

6701 GCTGTTTGAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGATTTTCCAAGTTTGAAGTCTTTCATTTCTTTATGTTTTAAATGCACTGA
6801 CCTCCACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAA
6901 GGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTT
7001 CCTGGTGTACTTGAGGGGATGAGTTCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATG
139 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 A Y D S I
7101 AGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCCTTCT
105 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 T D F D K Q
7201 GCCCGTTGCTCACAGCAGACCAATGGCAATGGCTT CAGCACAGACAGTGACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGT
72 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 I I E G T
7301 CTTGGTCTGATGGCCGCCCCGACATGGTGTGTTGTCTCATAGACATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAG
39 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 L D Q Q S
7401 ATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGT
5 I N F T K M
7501 CTCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTAC
7601 GACATTTTGAAAAGTCCCGTTGATTTACTAGTCAAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAAATCCCGTGAGTCAAACCGCTATCC
7701 ACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAAGTCCCATAAAGTCATGTAC
7801 TGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACC
7901 GTAAATACTCCACCCATTGACGTCAATGAAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCTGTGGGCGG
8001 TCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG
8101 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAG
8201 ATACCAGGCGTTTCCCCCTGGAAGTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTG
8301 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTGAGCCGACCGCT
8401 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTA
8501 TGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGTGAAGCCAGTTACCTTC
8601 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATTACGGCAGAAAAAAGGAT
8701 CTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAA
8801 ATCAGCGCGCAATAAAAATATCTTTATTTTATTACATCTGTGTGGTTTTTTTGTGTAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGA
8901 AACAAAACAACTAGCAAAATAGGCTGTCCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA