



1 CTCGAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACG
101 AAACAAAACAAACTAGCAAAATAGGCTGTCCCAAGTCAAGTGCGAGTGCAGAACATTTCTCTATCGAAGGACCTGCAGGCGTTACATAACTTACGGTA
201 AATGGCCCCGCTGGTGACCGCCCAACGACCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTC
301 AATGGGTGGAGTATTTACGGTAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCC
401 CGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGATGATGCGGTTTTGG
501 CAGTACATCAATGGGCGTGGATAGCGGTTTACTCACGGGATTTCCAAGTCTCCACCCATTGACGTCAATGGGAGTTTGTGTTTACTAGTCAGTGGCC
601 AGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGATCCGGTGCCTAGAGAAGTGGCGCGGGTAACTGGGAAAGTGAT
701 GTCGTGTAAGTGGCTCCGCTTTTCCGAGGTTGGGGGAGAACCCTATATAAGTGCAGTAGTTGCCGTGAACGTTTCCCTATCAGTGATAGAGATCTCCC
801 TATCAGTGATAGAGATCTTTCGAACGGGTTTCCGCCAGAACACAGCTGAAGCTTACCAGTCCACCATGGTGAAGCAAGGGCGAGGAGCTGTTACCAGGG
12 M V S K G E E L F T G
901 GTGGTGCCATCCTGGTTCGAGCTGGACGGCGACGTAACGGCCACAAGTTACGCGTGTCCGGCAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCC
12 V V P I L V E L D G D V N G H K F S V S G E G E G D A T Y G K L T
1001 TGAAGTTCATCTGCACCACCGCAAGCTGCCGTGCCCTGGCCACCCTCGTACCACCTGACCTACGGCGTGCAGTGTTCAGCCGCTACCCCGACCA
45 L K F I C T T G K L P V P W P T L V T T L T Y G V Q C F S R Y P D H
1101 CATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCGGAAGGTACGTCCAGGAGCGACCATCTTCTTCAAGGACGACGGCAACTACAAGCCCGCGCC
78 M K Q H D F F K S A M P E G Y V Q E R T I F F K D D G N Y K T R A
1201 GAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGACACAAGCTGGAGTACA
112 E V K F E G D T L V N R I E L K G I D F K E D G N I L G H K L E Y
1301 ACTACAACAGCCACAACGTCTATATCATGCGCCACAAGCAGAAGAAGCGCATCAAGGTGAACCTTCAAGATCCGCCACAACATCGAGGACGGCAGCTGCA
145 N Y N S H N V Y I M A D K Q K N G I K V N F K I R H N I E D G S V Q
1401 GCTCGCCGACCACTACCAGCAGAACCACCCCATCGGCGACGGCCCGTGTGCTGCCGACAACCCTACCTGAGCACCCAGTCCGCCCTGAGCAAAGAC
178 L A D H Y Q Q N T P I G D G P V L L P D N H Y L S T Q S A L S K D
1501 CCCAACGAGAAGCGCATCACATGGTCTGCTGGAGTTCGTGACCGCCCGCGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAGCTAGCTGGCCA
212 P N E K R D H M V L L E F V T A A G I T L G M D E L Y K
1601 GACATGATAAGATACATTGATGAGTTTGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTG
1701 TAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAATTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTA
1801 AAACCTCTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATA
1901 AGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTCCCAAGGTTTGA
2001 ACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTATGAAAAATTCAGAAATAATTTAAATACATCATTGCAATGAAA
2101 ATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGGAACCTTTAATAGA
2201 AATTGGACAGCAAGAAAGCGAGCTTCTAGCGAATTCTCGACTCATTCCCTTTCGCTCGGACGAGTGTGGGGCGTGGTTTTCCACTATCGGCGAGTACTT
342 E K A R P R T S P R R N G S D A L V E
2301 CTACACAGCCATCGGTCCAGACGGCCGCTTCTCGGGCGATTTGTGTACGCCCAGAGTCCCGGCTCCGGATCGGACGATTGCGTCGCATCGACCTG
322 V C G D T W V A A S R R A I Q T R G V T G A G S R V I A D C R G Q
2401 CGCCCAAGCTGCATCATGAAATTCGCTCAACCAAGCTGTGATAGAGTTGGTCAAGACCAATCGGGAGCATATACGCCCGAGCCGCGGATCCTGCA
289 A W A A D D F N G D V L S Q Y L Q D L G I R L M Y A R L R P S G A
2501 AGCTCCGGATGCCTCCGCTCGAAGTAGCGCTGCTGCTCCATACAAGCCAACCAGGCTCCAGAAGAAGATGTTGGGACCTCGTATTGGGAATCCC
255 L E P H R R E F Y R T Q Q E M C A L W P R W F F I N A V E Y Q S D G
2601 CGAACATCGCCTCGCTCCAGTCAATGACCGCTTATGCGGCCATTGTCGCTCAGGACATTGTTGGAGCCGAAATCCGCGTGCACGAGGTGCCGACTTC
222 F M A E S W D I V A T I R G N D T L V N N S G F D A H V L H R V E
2701 GGGGAGTCCCTCGGCCAAAGCATCAGCTCATCGAGACCTGCGGACGGACGCACTGAGCGGTGTGCTCCATCACAGTTTCCAGTGATACACATGGGGA
189 P C D E A W L M L E D L A Q A V S A S V T D D M V T Q W H Y V H P
2801 TCAGCAATCGCGCATATGAAATCACGCCATGTAGTGTATTGACCGATTCTTCCGGTCCGAATGGGCCGAACCGCTCGTGTGGTAAGATCGGCCGAG
155 D A I A C I F D R W T T Y Q G I G Q P G F P G F G S T Q S L D A A A
2901 CGATCGCATCCATGAGCTCCGCGACGGGTTGCAGAACAGCGGGCAGTTCGGTTTTAGGAGGCTTGAACGTGACACCCTGTGCACGGGGGAGATGCA
122 I A D M L E A V P Q L V A P L E T E P L D Q L T V G Q A R R S I C
3001 ATAGTCAAGGCTCTCGCTGAATTCCTCAATGCAAGCACTTCCGGAATCGGGAGCGGGCCGATGCAAGTGGCGATAAACAATACGATCTTTGTAGAAA
89 Y T L S E S F E G I D L V E P I P L A A S A F H R Y V Y R D K Y F
3101 CCATCGGCGAGCTATTTACCCGAGGACATCCACGCCCTCCTACATGAAAGTGAAGCAGGAGATTTCCGCTCCGAGAGCTGCATCAGGTCGG
55 G D A C S N V R L V Y G R G G V D F S F A R S E E G E S L Q M L D S
3201 AGACGCTGCGAATTTTCGATCAGAACTTCGCGACAGAGCTCGGGTGAAGTTCAGGCTTTTTCATGATGGCCCTCTATAGTGAGTCTATTATACTA
22 V S D F K E I L F K A V S T A T L E P K K M

3301 TGCCGATATACTATGCCGATGATTAATTGTCAATCCGGTTGCTTTGAATTAGCGGTGGTTTTCAACACCTAAAAAGGGTTAAAAGATACCTTTGAA
3401 CCGCTAAGAAGCCCGAGAATTAGCTCCGCTCAAAACTCAAGGGGACAAATTCAAAAATGACTTCCAGCGCCAGGCTGGCCTGACTAGTCTCCACCCACC
3501 AAATGTGAACAACTCCAACGCCATTACATCCCCTCCCCCGCCGCGACTAGCCGTGCTCAAAGCCCGAGGTGACTATTGCGGCCGATAGGACCACGGG
3601 GTCACAGGAAGCAGCAGCCGGTGAAGGACCAGGCCCTTCTCTTTGTGTGGTGACTCACCCGCCGCTCCACCGGGCTGCCGCTCTCCATTTTGAAG
3701 TCCTTGCAACAGGGCCCGGGAGCGGCCATCTTCCACGCACGCAACTGGTGCCGGACGGGATGGCCTCACCTAGTTAGGGAGGCAGGGCAACGCGGGC
3801 CGCCAAGCCAGATCGTGCCGGTGTGGGGCCACATGGCCTCGGCACGCTAACCCAGCCTGGTTGCTTCGGGAAAAACCCAGGCTCGCCCATCCAG
3901 GTGGCGTCGGACATGTGCTCCGAAGCGGGCGGGCCAGCCGCCACTCTGTCCCTCATTCTCCCAACCATGACCTCTCCGGGCTCCGGGCGAGCA
4001 AGCCCCGACCCTCCCTTTGTTAGCCCTATTGCTGAACGGCAATCGAAGGCAGCAGGGCAACAACAACAAAAAAGACCAGAGTGCAGCCGGGAG
4101 TAGCACGCGGGCGGGCGGGACACCACGCTAGGCCTCAAGCCGGACACGAGGCGAGGCTACGGGGTTGCCGCTAGGCCTCGCACTCTGCCTCCCGGCC
4201 GCCCCGAACTCGAAGCGGAATGCTCGCAGCTAATCCCCGCCGACGACAGCGGGGCCCGCCGCTCGGAGCAGGACCTCCAGCTCGGCGGCCCGGGAA
4301 GCCACACCCGCCCTCACCTGCGTTCTGACGGCAAGCTTCGGCGAAGAAGCTTGGCCTGCGGGTGCCTTGTGGTCTTTATAGCCGCTCGGCGTCAG
4401 GCCCGCCCGGCAATCAGCGCCGCCGCCGAGCCCGCTCTTCCGGTGGGCGGGGACCCCGCCCTGCTGTGGGGAGGGGCGGCCGCTGGAGGCC
4501 CTCGCGCTCTGGCGAACTAGT**CGACGGACTGGGCTACGGCCGCCCCGAGAGGCGCAGCCAGAGGCCGCGCTAGGAAGGGGCGGGCGCCGAGAAC**
4601 **ACGATCCCTCCCCACCCCTCGGACGTGACTCGGACCACATCCCGGGTCTGCTAGGGCCCTCCCTTCTGCTCTTTCCCCAGCCTGGCGGCTCTGG**
4701 **GGCGCCGTGACTCAGCCAGAATGTTGGCAATGGGGAGGGCGGAACGGGAAGTGGAGGACGCGGATGGAAAAGTCGGAAACGAAGGAAGCTGAGTTTCG**
4801 **CCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCC**
4901 **TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATAACAGGCGTTTTCCCCTGGAAGCTCCCTCGTGCGC**
5001 **TCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCATAGCTCACGCTGTAGGTATCTCAGTT**
5101 **CGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCC**
5201 **GGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCT**
5301 **AACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAA**
5401 **CCACCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGA**
5501 **CGCTCAGTGAACGAAAACCTCACGTTAAGGATTTTGGTCATGGCTAGT**